

APE_x



Deliverable

Project Acronym: APE_x
Grant Agreement number: 297355
Project Title: Archives Portal Europe network of excellence

D5.3 Data Preparation Tool Manual v2.0

Revision: final

Authors:

- Kerstin Arnold (Bundesarchiv, Germany)
- Lucile Grand (Ministère de la Culture et de la Communication, France)
- Silke Jagodzinski (Bundesarchiv, Germany)
- Yoann Moranville (Ministère de la Culture et de la Communication, France)
- Stefan Papp (Bundesarchiv, Germany)
- Susanne Waidmann (Bundesarchiv, Germany)

Project co-funded by the European Commission within the ICT Policy Support Programme		
Dissemination Level		
P	Public	

Table of content

1. Introduction	4
2. Installation and launch of the tool	4
3. Overview of the main page	6
4 The menus	7
4.1 File menu	7
• Select file or directory (shortcut: Ctrl+O or Cmd+O):	7
• Close selected file(s) (shortcut: Ctrl+D or Cmd+D):	8
• Save selected APE file(s) (shortcut: Ctrl+S or Cmd+S):	8
• Save report for selected file(s) (shortcut: Ctrl+R)	8
• Quit:	9
4.2 Options menu	9
• Country code:	9
• Repository code:	9
• Digital object type / Rights:	9
• General options for EDM conversion	10
• List of date conversion rules:	10
• Use own style sheet:	10
• Use own schema:	10
• Choose default style sheet:	10
• Choose default schema:	10
• Checks when loading files:	11
• Default folder for saved files:	11
• Languages:	11
4.3 Actions menu	11
• Validate (shortcut: Ctrl+N or Cmd+N)	11
• Convert (shortcut: Ctrl+M or Cmd+M)	11
• apeEAC-CPF	11
• Create EAG 2012	12
4.4 Windows menu	13
4.5 About menu	14
5. The list	14
5.1 Upload files in the list	15
5.2 Remove files from the list	15
5.3 Functions of the list	15
5.4. Save file(s)	16
6. The tabs	17
6.1 Summary tab	17

6.2 Validation tab.....	17
6.3 Conversion tab	18
6.4 Europeana conversion tab.....	18
6.5 Edition tab	18
7. Additional functionalities of the tool	20
7.1. Holdings guide creation	20
7.2. EAG creation form	21
7.3. EAC-CPF creation form	22
7.4. EDM conversion	22
7.5. List the index terms of your files	23
8. Update of the tool.....	23
9. Troubleshooting	23
10. Annexes.....	24
Annex 1: Validation errors, meanings and possible fixes	24
Annex 2: Conversion of dates - normalisation rules	25
Annex 3: Remarks on <eadid>	26
Annex 4: Remarks on external links	27

1. Introduction

The Data Preparation Tool is meant to help you prepare your data locally, off-line, to be compliant with the minimum requirements needed to publish data in the portal : one apeEAG file describing your institution, and apeEAD files describing your archives.

The DPT offers the same range of functionalities as the Dashboard, with some differences though. Basically, it offers more possibilities than the Dashboard.

This tool is altogether a validation tool, a conversion tool, an edition and a creation tool. These different actions can be combined, processed in batch etc. Different options give the possibility to check one's data on different points, before or after conversion.

- Validation tool:

Against existing schemas; this allows you to check whether your data are compliant to the EAD schema for instance.

After a conversion; this is a mandatory step to obtain correct apeEAD, EAG 2012 or apeEAC-CPF files that can afterwards be uploaded and published in the Portal.

- Conversion tool:

You can convert from a lot of formats to a lot of others. You can even import your own style sheet to convert your data: from EAD 2002 to apeEAD, from apeEAD to EDM etc. You can also create your own date conversion table to get normalised dates (for instance from 20th century to 1901/2000).

- Edition tool:

Once the conversion is done, you can edit your file and check that everything went smoothly; on top of that, the tool provides you a report on the points that can be improved (data quality check) within the tool: normalise the dates, add titles and the type of digital objects.

- Creation tool:

The tool offers you the possibility to create from scratch different types of files: holdings guides based on your finding aids, EAG2012 files, EAC-CPF files and EDM files from your apeEAD files.

Please note that the tool is able to work with XML only. Other data cannot be processed by the tool and have to undergo additional actions before being ready to be used in the DPT

The tool is available via manual download on the project website (currently: www.apex-project.eu). The interface is translated so far in English, French, German, Dutch, Greek and Hungarian.

2. Installation and launch of the tool

The tool is distributed as a ZIP archive file, *APE_data_preparation_tool_X.Y.Z.zip*, X.Y.Z being the version number. The last version is downloadable at this address: <http://www.apex-project.eu/index.php/outcomes/tools-and-manuals>

Requirements in order to use the standalone tool:

- Have Java 6 installed on your computer, or a more recent Java version. You can verify which version is installed on your computer by launching the console (Windows) or terminal (UNIX) and typing *java -version*.

Please do the following once you have the ZIP archive:

- Uncompress the ZIP archive using *7-zip*, *winzip* or any program able to uncompress ZIP archives.
- You will have a folder containing 2 folders (*lib* and *xls*) and 6 files.
- Depending on your Operating System (OS):
 - Using Windows: Double-click on the *launch.bat* file,
 - Using UNIX: Using your terminal launch the *launch.sh* file (i.e. Type *shlaunch.sh* – make sure that the file has at least the execution rights, *chmod +x launch.sh*),
 - Using OS X: Double-click on the *APE_data_preparation_tool-jar-with-dependencies.jar* file¹.
- On the first launch of the tool, a dialogue box will ask first for your country code (2-letter country code, in capital characters, compliant to iso3166), then an identifier for your archives: please enter your registered ISIL identifier (if you have one, else you will still need to enter a valid iso15511² code). Note that you can change those anytime you want in the *options*. Be aware that the dialogue box might stay in the background so that you could not see it at first.
- The program should be now running (see figure 1), if that's not the case, please refer to the troubleshooting section of this manual.
- The folder now contains 5 folders (*apeoptions*, *lib*, *output*, *temp* and *xsl*) and 6 files.

¹ Double-clicking the *APE_data_preparation_tool-jar-with-dependencies.jar* file can also be used as an alternative to the *launch.bat* file with Windows.

² The ISO standard 15511 consists of up to 14 characters, usually starting with a two-letter country code (or a one-three- or four-letter organisational code), followed by a hyphen, followed by an alphanumeric code for the single institution. The latter can contain the digits 0 to 9, the characters A to Z and the characters -, / or \.

3. Overview of the main page

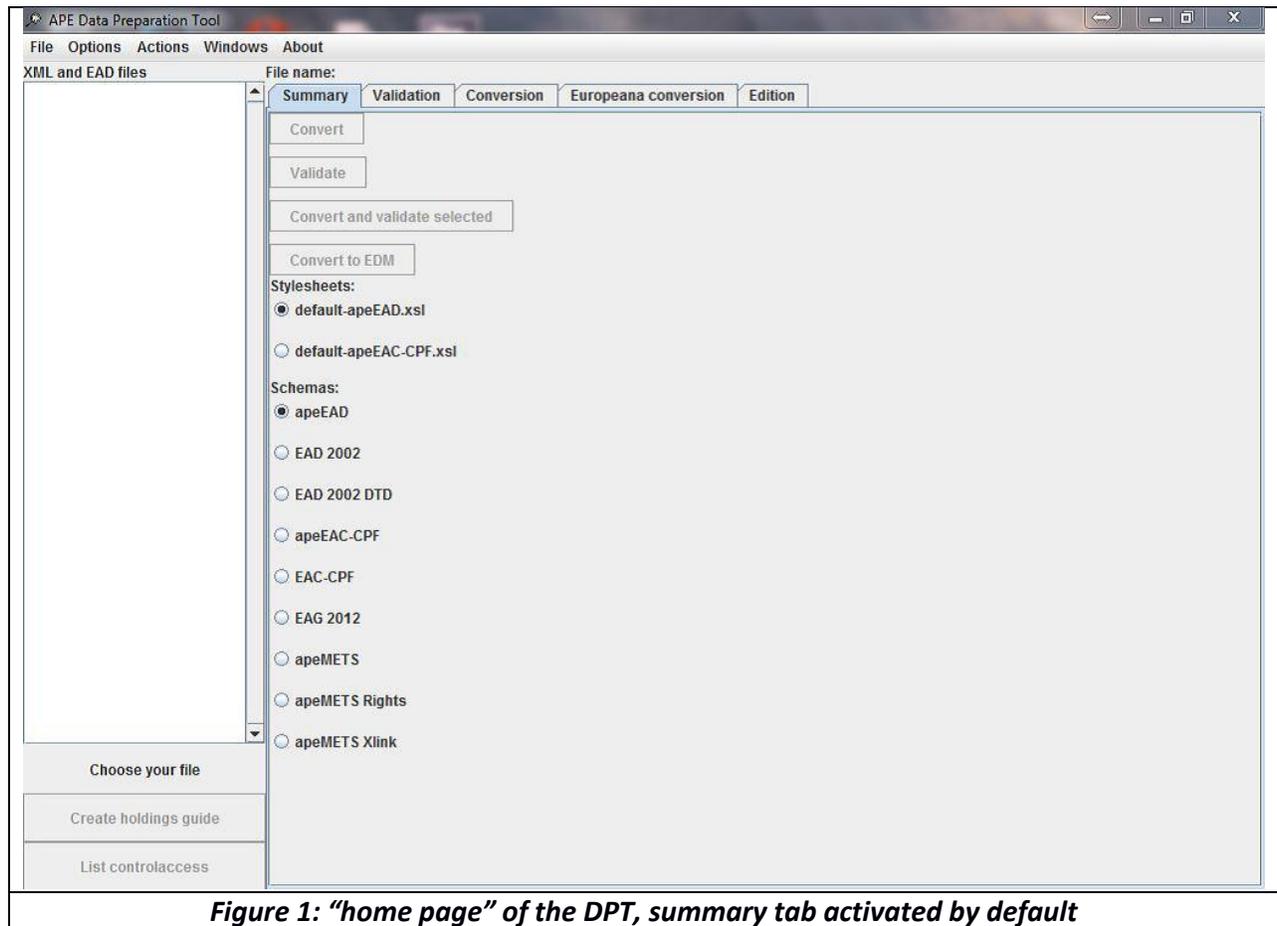


Figure 1: “home page” of the DPT, summary tab activated by default

In this main page, you see **5 menus at the top**: *File, Options, Actions, Windows* and *About*.

The rest of the screen below is divided in two parts:

- On the left side **one empty list** named “XML and EAD files” with 1 action button at the bottom: *Create holdings guide*. This button is placed just below the indication “Choose your file”.
- The right side contains **five tabs** at the top: *Summary, Validation, Conversion, Europeana conversion* and *Edition*.

Just above these five tabs, you can see the label “File name: “. It is meant to display the name of the file that is currently selected (only if one file selected, not several).

The summary tab has:

- 5 action buttons: *Convert, Validate, Convert and validate selected, Convert to EDM, Save message report*
- 2 radio button lists:
 - Stylesheets* for choosing the stylesheet needed: *default-apeEAD.xsl, default-apeEAC-*

CPF.xsl

Schemas for choosing the schema to be used: *apeEAD*, *EAD 2002*, *apeEAC-CPF*, *EAC-CPF*, *EAG 2012*, *apeMETS*, *apeMETS Rights*, *apeMETS Xlink*, plus the *EAD2002 DTD*.

The other tabs stay empty until a file is processed. They then display the according information.

4 The menus



There are five menus in the data preparation tool.

4.1 File menu

You have 4 items in the File menu.

- **Select file or directory (shortcut: Ctrl+O or Cmd+O):**

This is where you can open your files or directory to be able to process them in the tool. You have the possibility to select one or several files, or one directory. If you select a directory, it will import all the files located in this directory; however, subdirectories will not be included. Once you have done your selection, the files appear in the "XML and EAD" list on the left side of the window and are ready to be processed.

Please note that you can process hundreds or even thousands of files together.

Once you get files in the list, selecting one or several files will activate either the buttons on the right, either the one on the bottom (see figure 2).

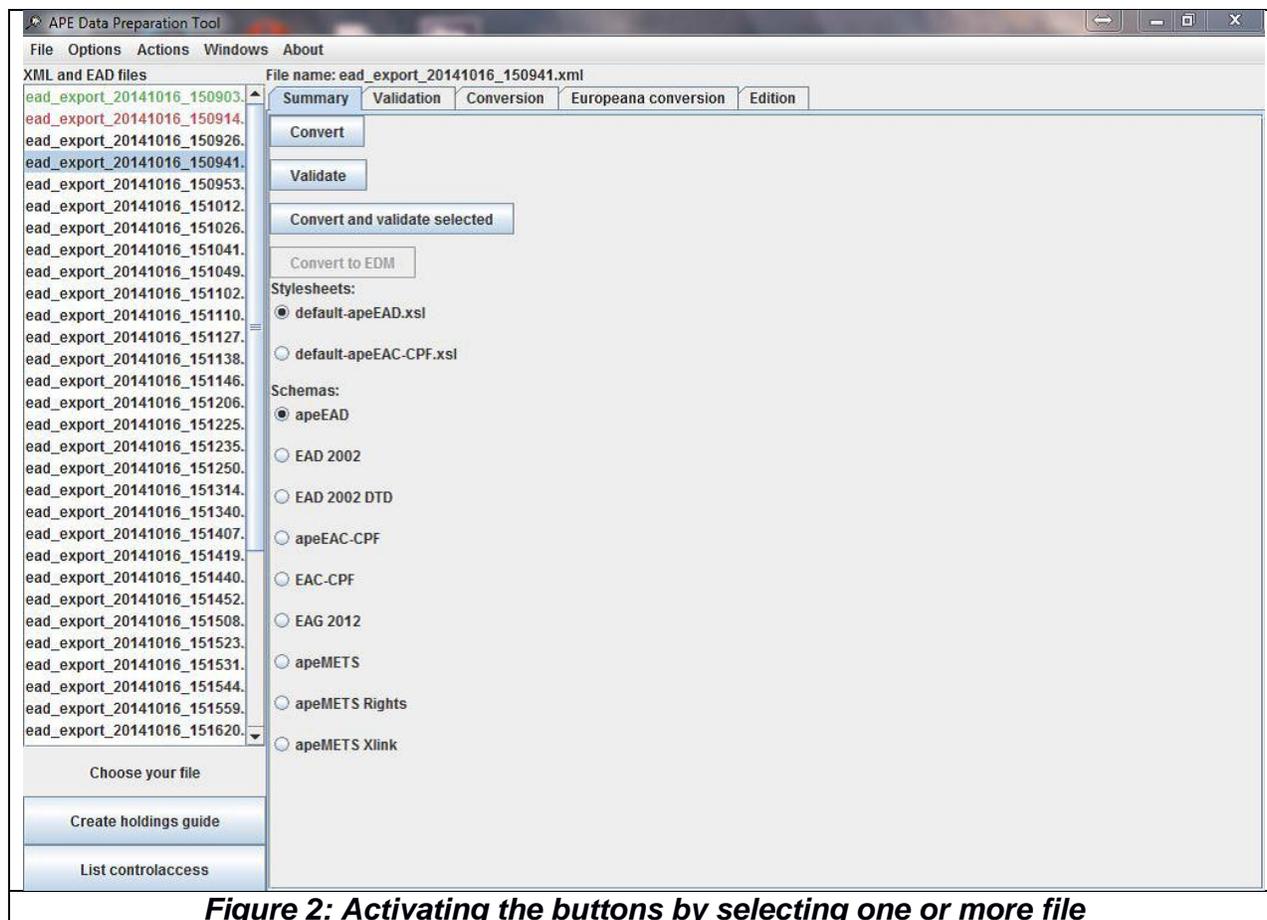


Figure 2: Activating the buttons by selecting one or more file

- **Close selected file(s) (shortcut: Ctrl+D or Cmd+D):**

Suppress the file(s) from the list.

- **Save selected APE file(s) (shortcut: Ctrl+S or Cmd+S):**

Once a file has been converted and/or validated, it has to be saved. The newly created files will be saved in the output directory by default but the options allow you to indicate another place. The name of the file changes when it is saved. If the file has been validated, its saved name will be *apeEAD_{original_name}.xml*, if not it becomes *NOT_apeEAD_{original_name}.xml*. Of course, this does not affect your original files.

- **Save report for selected file(s) (shortcut: Ctrl+R)**

All reports on the files processed in the tool (and selected in the list) are saved in one file, place by default in the output directory of the tool, even if the files did not pass the validation or conversion steps.

- **Quit:**

Closes the program.

4.2 Options menu

You have 12 items in the Options menu. They are meant to set up and keep your own parameters in the DPT, allowing you to apply specific rules during the processing of your files.

- **Country code:**

Here you can change the 2-letter country code, the one you had to enter at the first launch of the tool.

- **Repository code:**

Here you can change the identifier for your archives, the one you had to enter at the first launch of the tool. It's useful when you manage more than one institution.

- **Digital object type / Rights:**

This option allows you to add specific indication on the type and on the rights, mandatory for Europeana and strongly recommended for the Archives portal Europe.

Be aware of the fact that these indications will then be applied by default to all files processed in the DPT until you change or suppress them.

The digital object type indicates whether the digitised archives are Text, Image, Sound, Video or 3D. If you don't use this option, the default value will be Unspecified. This option is used within the Archives Portal Europe for refining the search results. Those values are reported inside your apeEAD XML files after you convert them. The chosen type will then be available inside your apeEAD XML file within each DAO element as attribute @xlink:role. In case, your local data already contain this information, you also have to specify that this shall be used if found.

Note: Converting the file and then choosing a type will *NOT* change the type inside the newly created file. A new conversion is needed.

The Rights information allows you to add default rights statement for the digital objects as well as for the finding aids themselves thanks to a dropdown list (containing the main licenses or indication of rights such as public domain mark or the different creative commons), that can be completed by a description and the name of the rights holder.

- **General options for EDM conversion**

You can specify by default the source of the identifiers as well as the base URL for the landing pages other than the digital objects.

- **List of date conversion rules:**

This option lists in a dialogue box the rules that will be applied during the conversion on the date elements of your files in order to transform text dates into normalized dates.

In the left column, you indicate the values read (to be found in the original files), in the right column, you indicate the values converted (to be found in the converted files). For instance you can indicate that 18th Century will be converted in 1701-1800, circa 1877 in 1877, or in 1876-1878 etc.

These conversion rules can be saved to be re-used later on. You can modify them as often as needed.

Please note that this is an extra possibility, to complete the normalization of dates already done during the conversion by the tool (See annex 2 for the full list of conversion rules). These added rules have priority on the default rules.

- **Use own style sheet:**

This option allows you to use your own style sheet (XSLT) in order to do your conversions instead of the default one. This option opens a dialogue box where you can select a style sheet and copy it in the xsl folder of the tool. The stylesheet then appears in the list of the summary tab and can be chosen to convert your files.

- **Use own schema:**

This option allows you to use your own schema (this must be a valid XSD file). This option opens a dialogue box where you can select a schema and copy it in the xsl directory of the tool. The schema then appears in the list of schemas of the summary tab.

- **Choose default style sheet:**

either default-apeEAD.xsl, or default-apeEAC-CPF.xsl;

By default the style sheet selected for newly imported files is the default-apeEAD style sheet, you can use this option if you wish to use another default style sheet on a regular base.

- **Choose default schema:**

lists the different possibilities offered by the tool to validate the files *EAD 2002*, *apeEAD*, *apeEAC-CPF*, *EAC-CPF*, *EAG 2012*, *apeMETS*, *apeMETS Rights*, *apeMETS Xlink*, *EAD2002 DTD*.

By default the system sets the data type to apeEAD files, you can change this by using this option.

- **Checks when loading files:**

This option allows the tool to check if the files that you are uploading are valid XML files. This is only an option because it might take time if you have to upload numerous files. Before displaying the file in the list, the tool parses the file and rejects the non XML files. For each file, a pop-up window warns you and allows you to copy the error message.

- **Default folder for saved files:**

The default folder where the tool saves the files as well as the report on the XML quality is the output folder of the tool (one of the six folders mentioned above in the chapter “Launch of the tool”).

This option allows you to change the default folder.

- **Languages:**

The tool is available in six languages: Dutch, English, French, German, Greek and Hungarian. The tool will open in the language of your operating system if it is Dutch, French, German, Greek or Hungarian and if not found will roll back to English.

4.3 Actions menu

You have 4 items in the Actions menu:

- **Validate (shortcut: Ctrl+N or Cmd+N)**
- **Convert (shortcut: Ctrl+M or Cmd+M)**

Those items do the same as the convert and validate buttons in the main summary tab (see figure 1).

- **apeEAC-CPF**

Has only one submenu that opens a form in a pop-up window (figure 3a and 3b).

You have to first choose one of the three possibilities: person, family, or corporate body. Then the corresponding form is displayed. The form is subdivided in four tabs that have to be completed and correspond to the ISAAR-CPF standard: Identity, Description, Relations, Control.

Start

Create a file for record creators (apeEAC-CPF)

Create instance for*: a person a family a corporate body

Please set the default language and script for this file (this can be individually changed for each item later)

Select default language for this file: --- Select default script for this file: ---

Exit Go

**Figure 3a: EAC-CPF creation
choose the type**

apeEAC-CPF

Identity Description Relations Control

Description of a person

Name

Full name or part of the name* Part of the name ---

Add further parts

Form of the name --- Select a language English

Add dates of use for the name Add single date Add timespan

Add further names

Authorised identifier

Authorised identifier of the person Authority agency

Add further authorised identifiers

Date(s) of existence*

Timespan from* to*

Date(s) is/are known unknown known unknown person is alive

Standardised dates

Add single date Add timespan

Exit Next tab Save

Figure 3b: EAC-CPF form

- Create EAG 2012

has two sub-menus: *Select an existing EAG (2012) file, Create a new EAG 2012 file.*

- Select an EAG file (2012) allows you to open the file that you want to process in the tool (completing or correcting existing elements for instance).

- Create a new EAG 2012 file opens a form in a pop-up window (figure 4). The form is subdivided in seven tabs that have to be completed following the logical order of the ISDIAH standard: *Institution, Identity, Contact, Access and Services, Description, Control, Relations*. You move from one tab to the other by clicking on the next/previous buttons.

The tab “Your institution” includes all mandatory elements, plus some optional elements
Once saved, the EAG file is added to the files in the list and can be validated afterwards

The screenshot shows a web-based form titled "Institution" with a tabbed interface. The active tab is "Your institution". The form is divided into several sections:

- Your institution:** Includes fields for "Person / institution responsible for the description", "Country code*" (FR), "Identifier of the institution" (FR-FRSIAF), "Name of the institution*", "Parallel name of the institution", and "Is this your registered ISIL code?" (yes). There is a button "Add further IDs" and a field "ID used in APE" (FR-FRSIAF).
- Visitors address:** Includes fields for "Street*", "City / Town (with Postalcode)*", "Country*", "Coordinates (latitude)", and "Coordinates (longitude)". There are buttons "Add further visitor address (translation)" and "Add postal address, if different".
- Contact and Access:** Includes fields for "Continent*" (Europe), "Telephone", "E-mail address", "Webpage", "Opening hours*", "Closing dates", "Access conditions (Accessible for the public)*" (yes), and "Disabled access*" (yes).
- Links:** Includes fields for "Link to your institutions holdings guide" and "Link title" (multiple instances).

At the bottom, there are buttons: "Add new repository", "Exit", "Next tab", "Save", and "Go to next repository".

Figure 4: EAG 2012 creation form

4.4 Windows menu

You have 5 items in the Windows menu:

- Summary (shortcut: Ctrl+1 or Cmd+1)
- Validation (shortcut: Ctrl+2 or Cmd+2)
- Conversion (shortcut: Ctrl+3 or Cmd+3)
- Europeana conversion (shortcut: Ctrl+4 or Cmd+4)
- Edition (shortcut: Ctrl+5 or Cmd+5)

In order to change to a selected tab, you can either click on it or select it in this menu, or use the shortcut.

4.5 About menu

You have 2 items in the About menu:

- **Visit project website:**

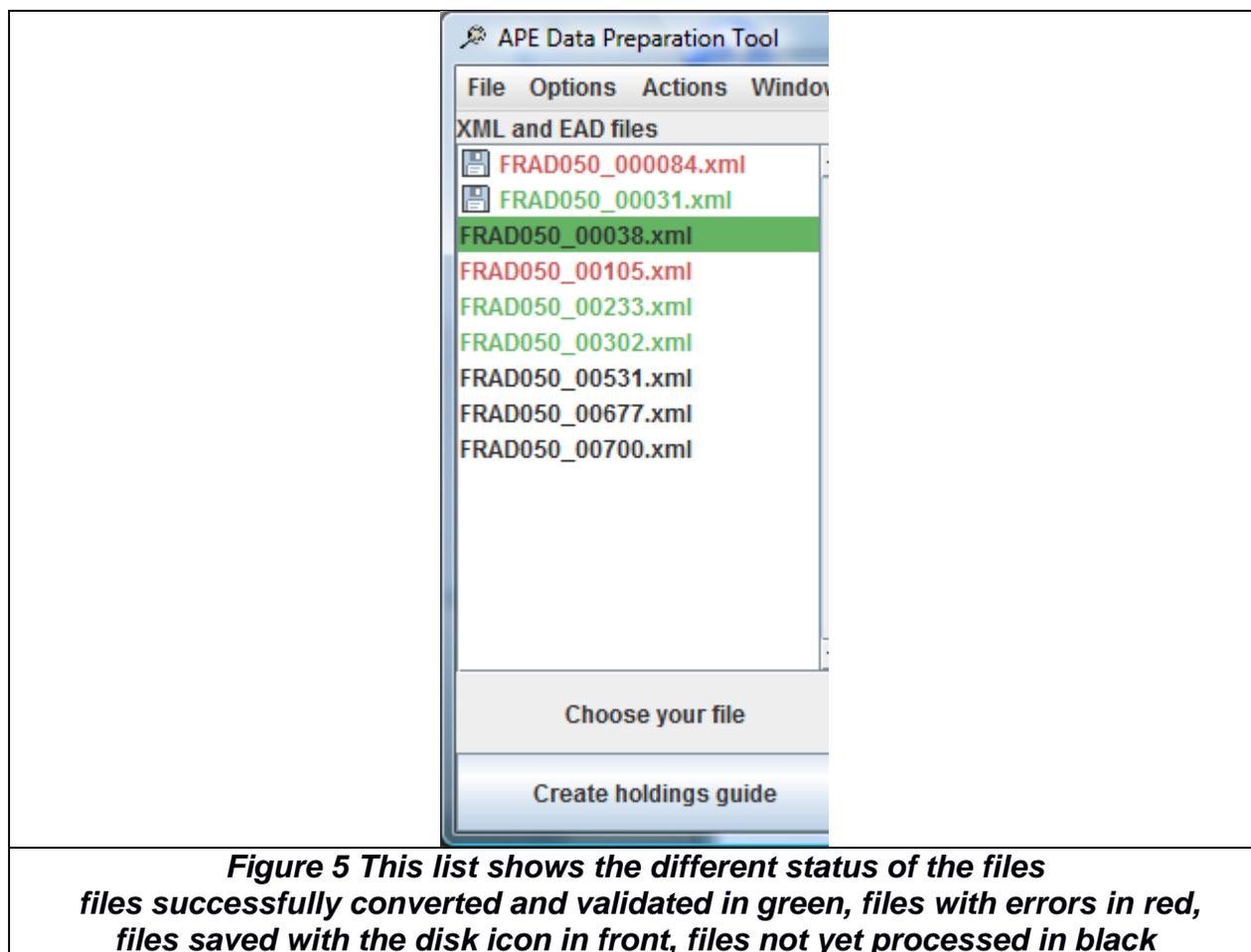
This option opens a page in your browser (you need an internet connection) and you will access the APEx project website (www.apex-project.eu).

- **APE DPT v2.0.x:**

This indicates the version of the DPT you are working with.

We will see in a following chapter (Chapter 8 *Update of the tool*) that an automatic check is done at the opening of the program to see if a new version is available.

5. The list



On the left side of the window there is an empty list that you populate either by opening files or directories or with the files created within the DPT: holdings guide, EAG, EAC-CPF or EDM files.

5.1 Upload files in the list

By clicking on “File” -> “Select file or directory”, you can upload one or several file(s) or a whole directory. The last path used when opening files is memorized by the tool.

If the option *Checks when loading files* is set to yes, only XML files are added to the list, and non-XML files are discarded.

It’s not possible to load two files with the same name. An according error message is displayed if it happens and you are asked to whether cancel the action or overwrite.

5.2 Remove files from the list

Select the file(s) by clicking on the name of one or more files and use the right click of the mouse, or after having selected it (them), chose *Close selected file(s)* in the File menu

5.3 Functions of the list

Selecting files to process them:

Select the file you want to process, and then convert and/or validate it by using the radio buttons in the summary tab on the right part of the window and clicking on the action buttons for processing the files. The radio buttons indicate which schema and style sheet you want to apply to your file(s). By default, the buttons are selected to convert or validate EAD files in apeEAD by using the default style sheet.

The action buttons on the right part are activated when at least one file is selected and depending on the status of your file (e.g. Convert to EDM will only be available for validated apeEAD files).

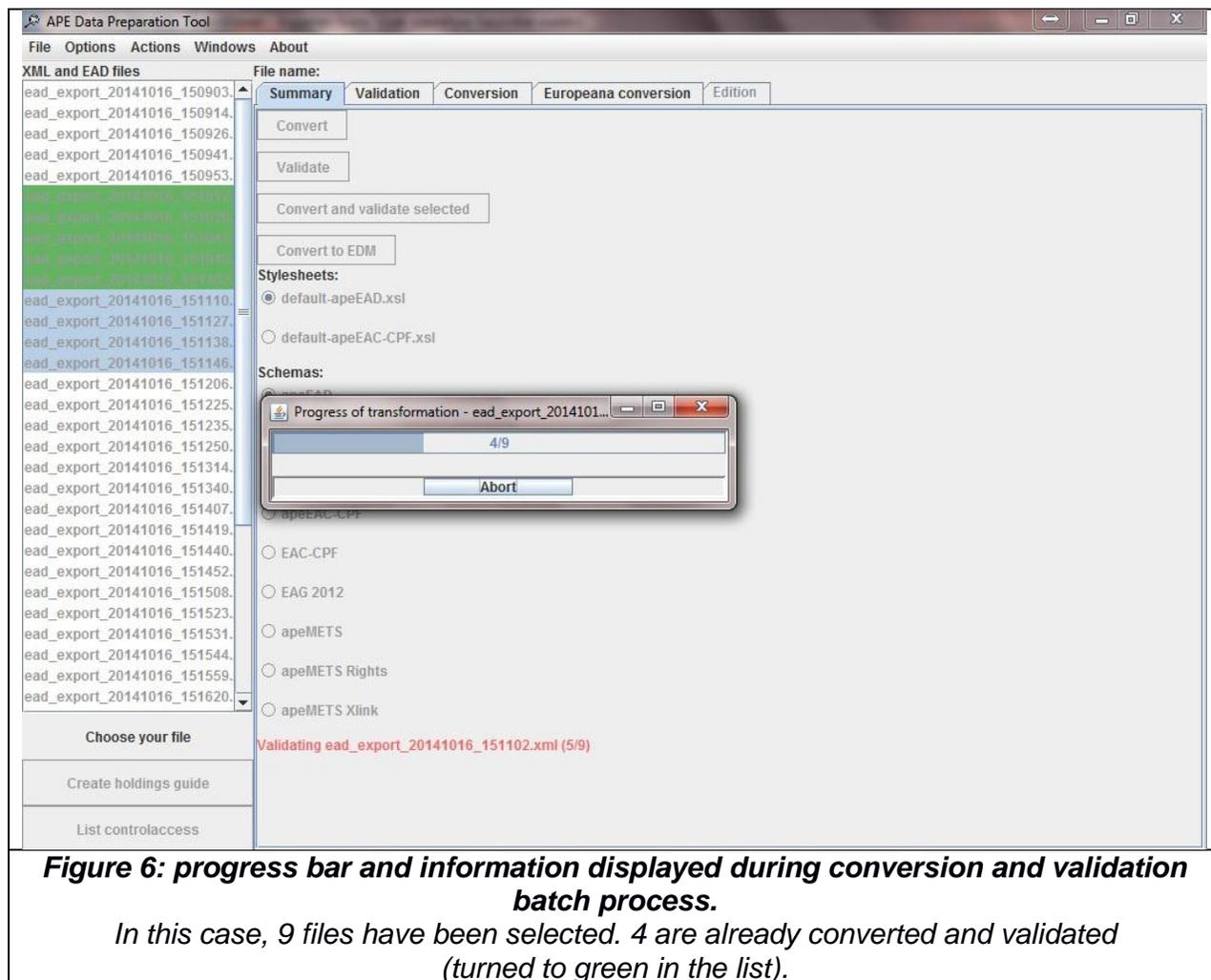
The action button *Convert to EDM* is activated when a file containing digitised object has been successfully converted to apeEAD and validated.

The action button at the bottom of the list, *Create holdings guide*, is activated when at least one file is selected in the list. For more details on this functionality, please see chapter 7 *Additional functionalities of the tool*.

Indicating the result of the action processed:

When a file has been successfully converted and validated, its colour changes in the list display to green. If an error occurred, its colour changes to red. Same for the colour of the tab corresponding to the process conducted.

On the main part of the window, you will see the progression and status of the operations thanks to a progress bar. You can stop the program when you want by pressing the *Abort* button that is made available during the process, but please note that after pressing the *Abort* button, the tool will finish its last operation before stopping.



5.4. Save file(s)

To save your converted or validated files, you have to choose “Save selected APE file(s)” in the File menu. It saves the file(s) currently selected in the list.

- If the file is valid, the file will be saved and renamed as follows in the output folder: ape{schema_used}_{original_filename}.{original_extension} and an according message will be shown.
- If the file is not valid the file will be saved and renamed as follows in the output folder: NOT_ape{schema_used}_{original_filename}.{original_extension} and an according message will be shown.

{schema_used} can be “EAD”, “EAG” or “EAC-CPF”. Exemple: apeEAD_FRAD015_2_4_AV.xml or apeEAG_FR-FRAD015.xml

The saved files are marked with a small disk symbol in front of the file name in the list. If the edition tab is used, the disk logically disappears.

6. The tabs



You have five tabs: *Summary*, *Validation*, *Conversion*, *Europeana conversion*, *Edition*. Depending on the process, some of them turn green or red to indicate that a process is completed and that you might check them.

6.1 Summary tab

The first and main tab is the **Summary**, and is opened by default when you enter the tool. See chapter 3. This is where you can apply operations to your XML files. When a file is selected in the list, you can, from this tab, convert the file to the APE format, pass it to the validator or convert it to Europeana format. You have the possibility to choose the style sheet to be used for the conversion –of course only if you installed your own style sheet (see chapter 4.2 *Options menu - Use own style sheet*). The same goes for the schemas.

The status of the current transformation or validation is displayed beneath those buttons and radio buttons.

Please note that the action buttons and radio buttons are available only if the action is possible for the selected file.

6.2 Validation tab

The second tab is **Validation** and here you will find the results of a validation for the selected file. It can either contain a sentence saying the validation succeeded or a list of all the errors found during the validation process, with the indication of the lines and columns where the errors have been found. You can find in the annex 1 a list of the possible errors, their meaning and the actions to be taken to avoid them.

Please note that if you validate a file before the conversion, the lines indicated in the error report will be the ones of your original file, but if you validate a file after the conversion, the lines indicated will be the ones of the new converted file, therefore will not exactly correspond to your original file.

Besides, in case of a successful validation, a report on the data quality or completeness of the file will be displayed, indicating the number of missing or wrong information for four elements:

```
----- Data quality -----  
Missing content for title information (unittitle): 0  
Missing normalised date information (unitdate@normal): 0  
Missing information on digital object type (dao@xlink:role): 0  
Wrong information on digital object link (dao@xlink:href):0
```

This information also can be downloaded via the button “download report of XML quality” in the tab: it is saved in the report file to be found in the output folder of the tool. To save this information for several files, you have to use the same command from the File menu.

You can improve the data quality of your files by editing them (Edition tab, see chapter 6.5), or by adding rules for date conversion and by indicating the digital object type (Options, see chapter 4.2).

6.3 Conversion tab

The third tab is **Conversion** and here is displayed information on the transformation operated on the files and the problems encountered:

- The tab lists the elements (and their attributes, if applicable) that are being excluded either because they are not recognised, either because they do not comply to the apeEAD profile which is more restrictive than the EAD2002.
- The tabs also indicate the problems due to the c element hierarchy, when a c level is not understood.

Please note that the conversion tab does not list all changes made to the file, but only the elements discarded during the conversion.

6.4 Europeana conversion tab

The fourth tab is **Europeana conversion** and here is displayed information on the EDM conversion. Files that are transformed into EDM format are automatically saved in the output folder (or an alternatively chosen default folder for saved files, see chapter 4.2 Options). Their names stay the same with an additional "-edm" before the file extension.

For more information regarding the conversion to EDM, please refer to chapter 7.4.

6.5 Edition tab

The last tab is **Edition**

When an EAD file is selected in the list, it is read and displayed in a tree like view. Some elements only can be edited, they are marked in green or blue (see figure 6):

- eadid @mainagencycode

- eadid @countrycode
- eadid / text
- language / language
- language / language / text
- langmaterial / language
- langmaterial / language / text
- unitdate@normal
- titleproper / text
- c @level

The eadid is mandatory to validate your file.

The elements and attributes related to the language are useful for the display and refinement of the search results.

Same goes for the date: a date is searchable only if normalised.

An empty titleproper will be displayed as “no title”.

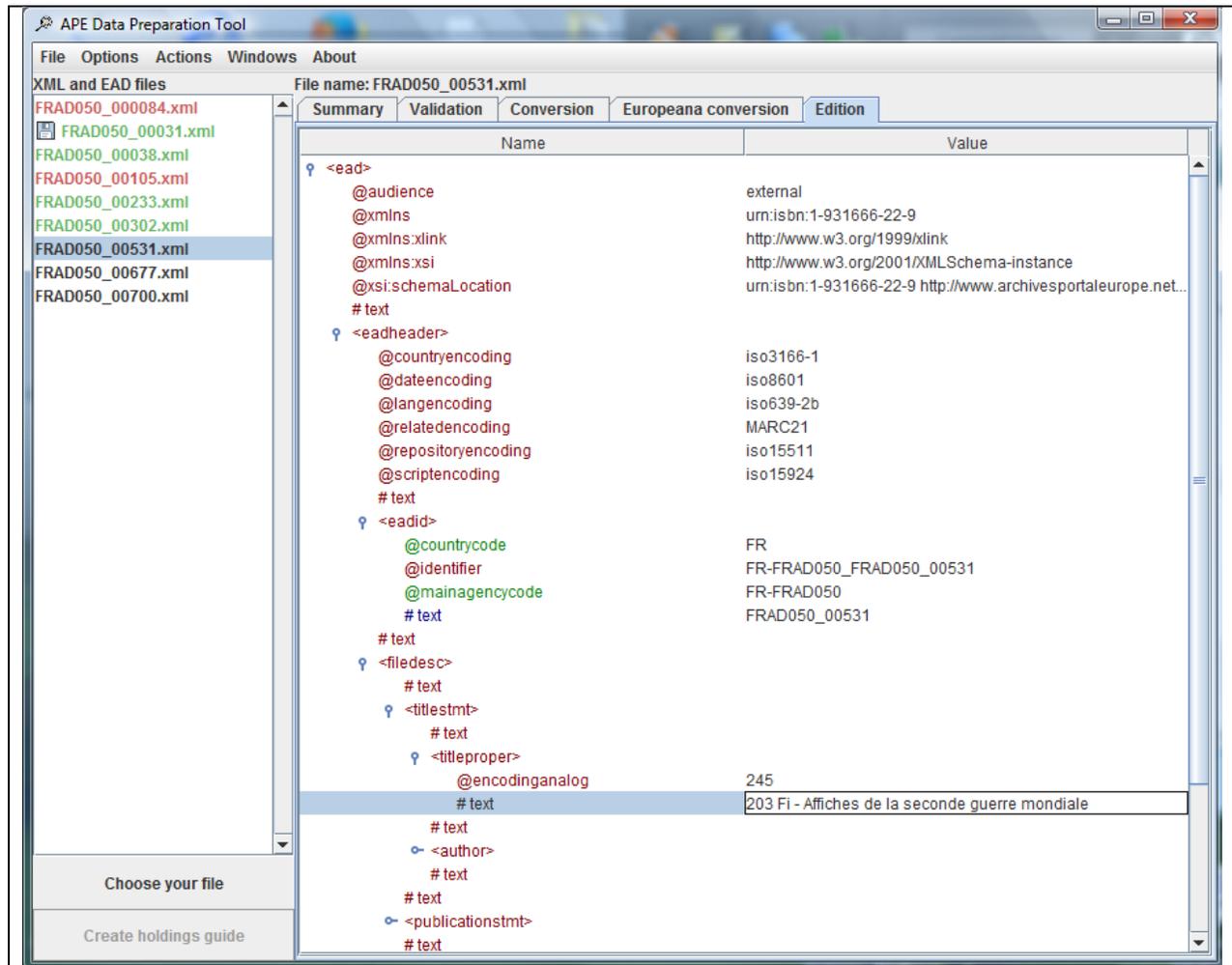


Figure 7: XML tree:
editable elements are indicated in blue, editable attributes are indicated in green

Please note that after editing your file, you should validate it again and save it.

7. Additional functionalities of the tool

7.1. Holdings guide creation

The button “Create Holdings Guide” in the summary tab becomes available once at least one valid apeEAD file has been selected. Clicking on the button opens a new window containing the file(s) selected to be integrated in the Holdings guide (It’s preferable to save your finding aids before beginning).

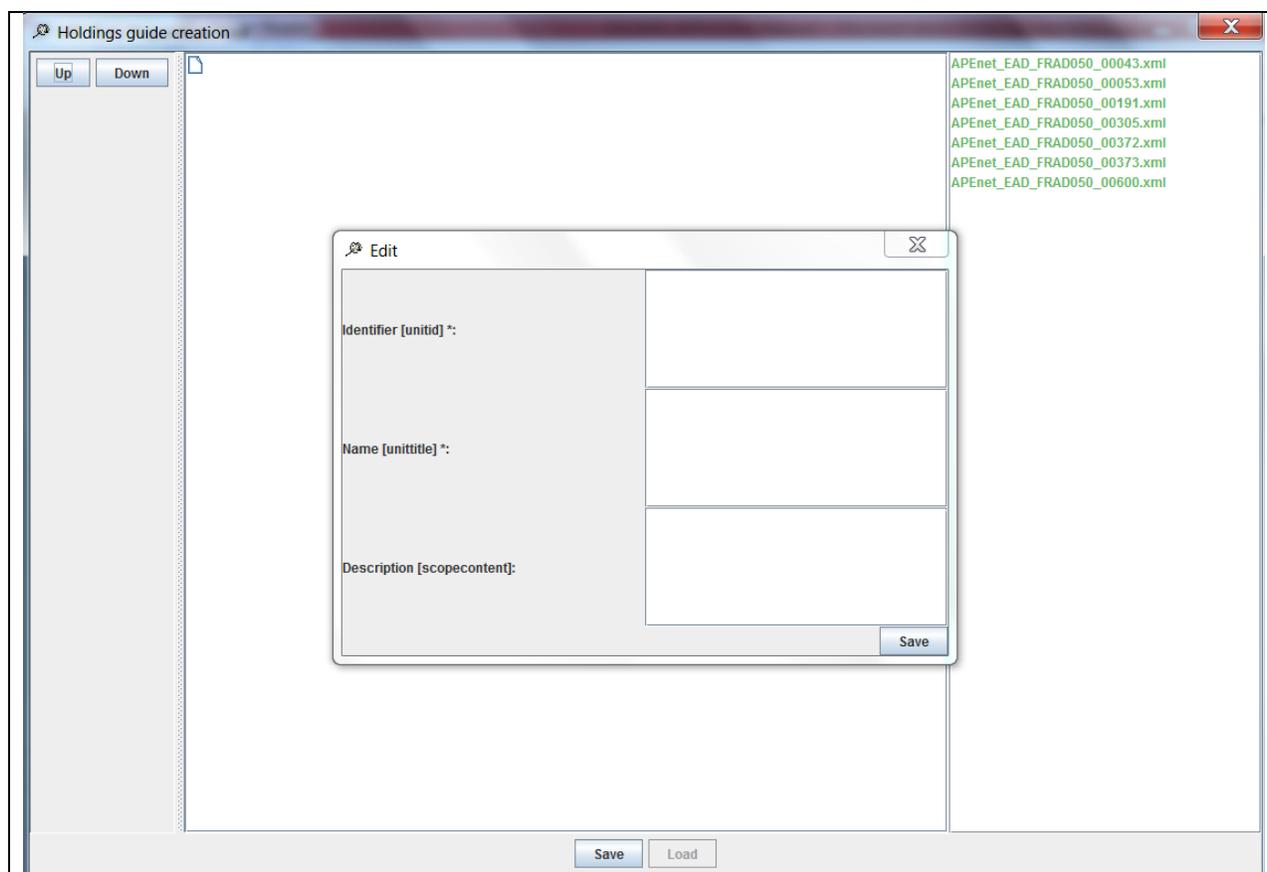


Figure 8: Holdings guide creation form.

*The screen is divided in three parts:
on the right: list of selected files
in the middle: the tree structure of the holdings guide
on the left: ordering buttons*

An extra dialogue box (Edit) allows to enter the elements of each description level

On the right side, you have the list of all finding aids to be linked within the holdings guide. You have first to provide in the *edit* window some information to describe your holdings guide: identifier (<unitid>), name (<unittitle>) and description (<scopecontent>). This has to be repeated for the lower levels if you do create lower levels.

- You can create as many levels and sub-levels as needed to organise your holdings guide. The structure will be displayed in the middle part of the window as a tree form.
- Once you have created your hierarchy, you simply drag and drop the files in the corresponding levels.
- If you need to rearrange the hierarchy or the order of your files: the up and down buttons on the left enable you to move files or levels inside the same level from one place to the other. To move an entire level to another level, you have to use the drag and drop facility.
- The right click on a level or a file enables you to add, edit or remove a level, or remove a file. In this last case, the file(s) will be available again in the right part of the screen for further use.
- Once all finding aids have been inserted in the holdings guide, click on save. The finding aids left in the right part won't be integrated.
- Back to the summary tab, validate your file that is now at the bottom of the list on the left.
- Save it before closing the tool.

The holdings guide obtained has the following characteristics: all information of each level (identifier, title and description) are transformed into c levels (unitid, unittitle and scopecontent) and all information of the file levels are retrieved from the files themselves (unitid, unittitle and scopecontent) but the <eadid> of each file becomes part of the otherfindaid/extref element used in the holdings guide for linking the finding aids.

Please note that during the save, a serialised file is also saved for loading the file at a later stage if wanted. The file is saved as a ".ser" file in the output directory. This .ser file can serve as a base to further modify or complete the holdings guide, if it has not been overwritten or moved from the directory. One can always save it and replace it in the output directory for further modification.

7.2. EAG creation form

The EAG 2012 is based on the ISDIAH standard. Please refer to it for more detailed information (<http://www.ica.org/10198/standards/isdiah-international-standard-for-describing-institutions-with-archival-holdings.html>) or see <http://www.apex-project.eu/index.php/outcomes/52-public/about-the-project/outcomes/35-tools-and-manuals>.

The first tab (*Your institution*) contains all mandatory elements required to create a valid EAG file. Numerous fields can be translated; therefore you have the possibility to select a language on the right of these fields.

You move from one tab to the other by clicking on the next/previous tab buttons and can save your file from any of the tabs by clicking on the save button without going through all tabs. The exit button closes the pop-up window without saving the file.

You have to validate and save your file in order to upload it in the Dashboard.

See figures 3a and 3b

7.3. EAC-CPF creation form

The apeEAC-CPF is based on the ISAAR (CPF) standard. Please refer to it for more detailed information (<http://www.ica.org/10203/standards/isaar-cpf-international-standard-archival-authority-record-for-corporate-bodies-persons-and-families-2nd-edition.html>) or see <http://www.apex-project.eu/index.php/outcomes/52-public/about-the-project/outcomes/35-tools-and-manuals>.

You can create EAC-CPF files in the DPT by filling a form. Please remember to validate your file then save it for further upload in the Dashboard of the Archives portal Europe.

See figure 4

7.4. EDM conversion

Based on valid apeEAD files, the tool can create EDM files for Europeana. Use the EDM button in the summary tab. A new window opens (see figure 8) where you have to select different options. Please note that **all fields are mandatory**, except *Specify additional rights information*.

Please choose a conversion mode <input checked="" type="radio"/> minimal <input type="radio"/> full	
All fields marked with an asterisk (*) must be filled	
Data provider:*	Archives départementales du Cantal <input checked="" type="checkbox"/> Take from file (<repository>) if existing
Type:*	<input type="radio"/> TEXT <input type="radio"/> IMAGE <input checked="" type="radio"/> VIDEO <input type="radio"/> SOUND <input type="radio"/> 3D <input checked="" type="checkbox"/> Take from file (<dao@xlink:role>) if existing
Inherit language:*	<input checked="" type="radio"/> Yes <input type="radio"/> No <input type="radio"/> Provided <input checked="" type="checkbox"/> Take from file (<language@langcode>) if existing Abkhazian Afar Afrikaans
Specify license:*	<input type="radio"/> Creative Commons <input type="radio"/> Creative Commons CC0 <input type="radio"/> Creative Commons Public Domain Mark <input checked="" type="radio"/> Europeana rights statements Restricted access
Specify additional rights information:	

Figure 9: Preferences for the conversion to EDM

Two conversion modes are available, minimal and full.

The **minimal conversion** will only include the information that is mandatory in EDM (the title of the object, the language of the object, the type of the digital object (TEXT, IMAGE, VIDEO, SOUND, 3D), the link to the digital object and the right information on the digital object) plus the information on the data provider and the provider (by default “Archives Portal Europe”). For all of these aspects, apart from the rights information, it is possible to have the DPT check for existing information in the EAD file. When the according checkboxes are checked, these existing information are taken instead of any value provided via the tool and the latter will only be used in cases, where there is no according information in the file.

The **full conversion** allows you to additionally specify context information which shall be included.

Basically, this will refer to copying information from higher levels in the EAD hierarchy in case they are not available on the lower level, which is linked to the digital object. Typically, this refers to a concatenation of the titles of the higher levels to be provided with the single unit or the name of the records' creator.

7.5. List the index terms of your files

An option allows you to let the tool read the files and list some subelements of <controlaccess>: <subject>, <persname>, <corpname>, <famname>. This is interesting in two aspects.

- Prepare the "topics" within the Dashboard, namely the tag cloud displayed on the home page of the portal. The tags are based on the <controlaccess> content. Therefore, if an indexation term matches a topic, it will automatically be linked to it after publication in the Dashboard.
- Rapidly check the indexation terms of your file(s).

This option is available after having selected one or more files via a specific button *List controlaccess*, on the bottom of the list, under the *Create holdings guide* button.

8. Update of the tool

A semi-automatic update is available in the tool. It is only semi-automatic because once detected, it requires the user to download the new version and install it. Please note that you need to be connected to the Internet for this.

At the launch of the tool, the tool will check on the Internet if a new version is available. This check is done at a minimal interval of 2 weeks. If it does find a new version, then the tool proposes to go to the Internet page in order to download the new version.

Please note that downloading the tool will also get the last version of the default style sheet where additions might have been made.

9. Troubleshooting

If problems were to be found, please send us an email via <http://www.archivesportaleurope.net/contact> and select "Technical issues" in the proposed subjects.

If the tool contains bugs or you suspect it does, please do the following before sending an email:

- Go to the output directory of the tool and retrieve the file named *errors.log*,
- Copy and paste the content of that file into the email with some explanations on the error that occurred.

10. Annexes

Annex 1: Validation errors, meanings and possible fixes

Text found in the validation tab	Meaning	Possible way to fix
You are using the forbidden value 'XX' in the attribute 'YY' of the element 'ZZ'.	You are using a wrong attribute value. For example, you might be using a "c" element with attribute "level" with a value "wrongValue", which is not allowed.	The value of the attribute must be correct, so try to: - Fix this in the original file if possible; - Modify the conversion stylesheet used for your purposes; - Create your own stylesheet.
The element 'XX' is missing at least one of the following sub-elements: 'YY'.	Some elements must contain mandatory elements. This message means you are not using all the mandatory elements needed. So here, element 'XX' does not contain the mandatory element 'YY'.	- Fix in original file (this element must be already missing in there); - A stylesheet does not do its job correctly and deletes elements (fix stylesheet).
The element 'XX' has been found but should not appear here, only one (or more) of the following element(s) are accepted here: 'YY'.	An element appears where it should not. For example, you can have an XML containing: <c> <scopecontent>[...]</scopecontent> <did>[...]</did> </c> This is not allowed because <did> should appear before <scopecontent>.	- Fix in original file so the order of the elements is correct; - A stylesheet passes elements earlier in the chain and should wait for its turn (Fix stylesheet).
The element 'XX' has been found but should not appear here, no child element is allowed here.	An element which cannot contain sub-elements does contain one. For example, element "persname" contains a sub-element "unitdate".	- Fix in original file; - Add rule in stylesheet to ignore the sub element and only copy the data of that element; - Add rule in stylesheet to copy the element or its data to another wrapper element, where it would also be fitting and allowed;
The declaration of 'XX' was not found, your file might not contain the schema	The element 'XX' was not found at the root of the document, which means the file you are trying to	- You can try to first convert your file, then validate it;

declaration corresponding (xmlns).	validate is not using the schema.	- Your XML file does not contain the mandatory schema namespace; - You maybe try to validate an EAD file with another schema (EAG 2012 for example).
You are using the value 'XX' in an attribute that can only contain one of those values: 'YY'.	An attribute contains a wrong value. Only a value from a list can appear here and you are using a forbidden value not existing in that list.	- Fix in original file - Add rule in stylesheet to transform that special value into a not forbidden value.
Missing content for title information (unittitle): {0}	The element unittitle is empty. Note that this can be « normal », depending on your FA and the way it is structured and displayed in your system, but the result will be the display of « No title specified » in the hierarchy	- Fix in original file
Missing normalised date information (unitdate@normal): {0}	The element date is not normalised and contains only text ; the consequence is that the file is not searchable by date. Note that this can be « normal », depending on your FA and the way it is structured.	- Fix in original file - Add rules in the option "list of date conversion" - see chapter 4.2
Missing information on digital object type (dao@xlink:role): {0}	The type of dao will not be analysed (thumbnail or full image)	
Wrong information on digital object link (dao@xlink:href):{0}	The link to the original document will not work because it does not begin with http://, http:// or ftp://	- Fix in original file or - Add a rule when exporting your file to include the full url or uri

Annex 2: Conversion of dates - normalisation rules

These rules are applied automatically to the dates during the conversion of the file.

Value read in the original file	Value transformed in the apeEAD file
01.01.1985	1985-01-01

1.1.1985	1985-01-01
1985.01.01	1985-01-01
01.01.1985/02.01.1985	1985-01-01/1985-01-02
1.1.1985 - 2.1.1985	1985-01-01/1985-01-02
1985.01.01/1985.01.02	1985-01-01/1985-01-02
01011985	1985-01-01
01011985/02011985	1985-01-01/1985-01-02
19850101	1985-01-01
19850101/19850102	1985-01-01/1985-01-02
1985—1986	1985/1986
1985 – 1986	1985/1986
1985 - 1987, 1990	1985/1990
1985, 1987 - 1990	1985/1990
03.07.1985 bis (to, à, etc.) 06.07.1985	1985-07-03/1985-07-06
3.7.1985 bis (to, à, etc.) 6.7.1985	1985-07-03/1985-07-06
1985 bis (to, à, etc.) 1988	1985/1988
avant 1985	0000/1985
après 1985	1985/2999
around 1985	1985
15th century/siecle/Jhd/...	1400/1500
century timespan (15th to 17th century)	1400/1700
(1985) 1987 – 1990	1985/1990
1985 - 1987 (1990)	1985/1990

Annex 3: Remarks on <eadid>

The <eadid> element is mandatory to manage the files within the portal. Therefore, a check on

it is made during the conversion or the validation of each file.

If there is no eadid, the tool asks you to provide one manually *via* a pop-up window. The tool proposes by default the identifier found in the element archdes/did/unitid, if it exists, but any other <eadid> can be provided.

Please note that **white spaces** as well as these characters **cannot** be part of the <eadid>:

* : / \ [] + % @ \$ # ^ & () ! ~

Annex 4: Remarks on external links

The elements <dao>, <extref> and <extptr> should contain the attribute @xlink:href.

This attribute must be well formatted and contain the full url, beginning with http://, https:// or ftp:// in order to provide a valid link to the original data on your website.

You can place external links for different purposes:

- to link back to your FA: this can be done at any level of description
- to link back to your digitized documents and display a thumbnail in the portal. In this case, the dao element should be, in the case of an image (could be also SOUND, TEXT, VIDEO, 3D, UNSPECIFIED):
<dao xlink:href="<http://.....>" xlink:title="thumbnail"/>
<dao xlink:href="<http://.....>" xlink:role="IMAGE" xlink:title="{Any title you want to give to your image, optional but recommended for accessibility reason}"/>