



SciencesPo.

CEE

# EUROPOLIX

-----

## USER GUIDE



## Table of Contents

What is Europolix?.....	3
Overview of the website.....	4
Welcome page.....	4
Login page.....	4
Import page.....	5
Act ids validation page.....	7
Ministers' attendance validation page.....	8
Act data validation page.....	9
Database management page.....	10
Export page.....	11
History page.....	11
Fields to retrieve.....	13

## What is Europolix?

This section explains what is the project and what it is for.

Europolix is a project started by Sciences Po and Centre d'Études Européennes (CEE) in October 2012. Its aim is to build a database gathering European acts from *eurlex* (council) and *oeil* (parliament) - and *prelex* until the beginning of 2015 (commission).

This project can be seen as a mini website where information is automatically retrieved from the websites above and where experts validate it, act after act.

On top of the creation of a common database, a first goal is to cross validate data between the different European institutions and with an expert to highlight errors. In the end, we wish to have a clean database with no error.

A second goal is to do a statistical analysis of the cleaned acts. To this end, the website allows to extract all the acts of the database into a file readable in a spreadsheet tool.

## Overview of the website

This section shows how the website works with screenshots and explanations.  
The website is accessible at <http://may.cdsp.sciences-po.fr/europolix>.

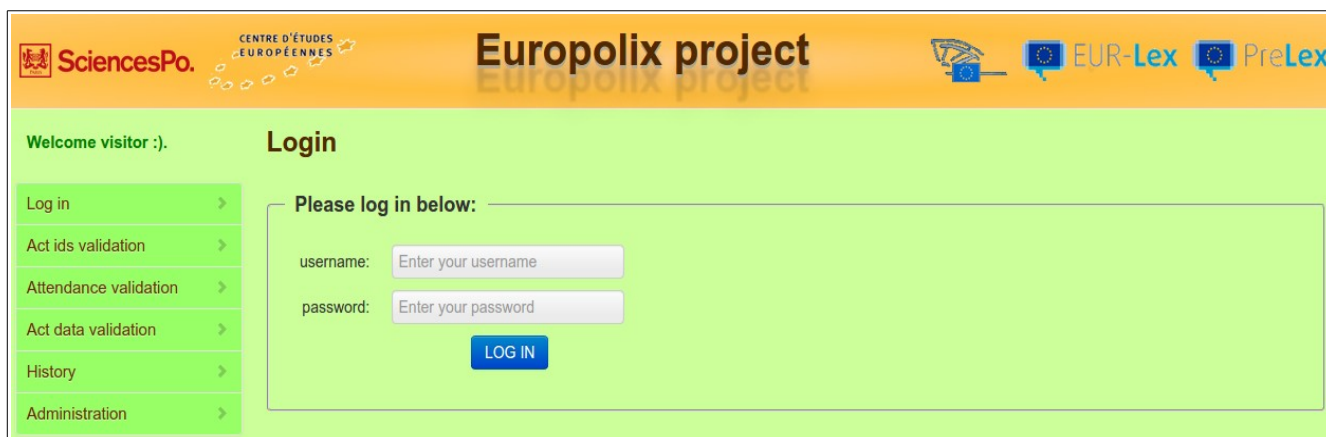
### Welcome page



The screenshot shows the homepage of the Europolix project. The header is orange and contains the SciencesPo logo, the text 'CENTRE D'ÉTUDES EUROPÉENNES', the title 'Europolix project', and logos for EUR-Lex and PreLex. The main content area has a light green background. On the left, there is a vertical menu with the following items: 'Log in', 'Act ids validation', 'Attendance validation', 'Act data validation', 'History', and 'Administration'. To the right of the menu, the text reads: 'Welcome visitor :). Welcome to europolix! This is the Europolix project main page. It aims to create a unique database to store European acts from the European institutions Eurlex, Oeil and Prelex. The acts are retrieved from these 3 websites, crossed and then validated by experts before being saved in the database. Please log in and choose an action from the left menu.'

This is the homepage of the website. The left menu shows the available actions on the website. **These actions must be performed in order to validate and then export act data.** The following details each of these actions.

### Login page



The screenshot shows the login page of the Europolix project. The header is identical to the welcome page. The main content area has a light green background. On the left, the same vertical menu is present. To the right of the menu, the text reads: 'Welcome visitor :). Login Please log in below:'. Below this text, there are two input fields: 'username: Enter your username' and 'password: Enter your password'. A blue button labeled 'LOG IN' is positioned below the password field.

This is the login page (click on “Log in” from the menu). You must log in to get access the other functionalities of the website. If you have no id, you must ask the administrator to use the website. Enter your username and password. Then click on “Log in”.

## Import page

This is the import page. You can import different csv files that are first saved on the server and then processed.

- Import acts to validate: Import a file containing information on acts (information found in monthly summaries and then gathered in a spreadsheet document). **This file is usually called RMC and stands for Relevé Mensuel du Conseil.** Two steps:
  - 1/ the content of the file is saved in the database.
  - 2/ for each act saved, we retrieve corresponding data (ids only) from the *eurlex* or *oeil* websites and update the database with this new information. A message is displayed if errors are encountered during the import.
- Import DosId: Import a file containing *prelex* unique id (called *dos\_id*) and the *NoCelex* variable to do the mapping ( *NoCelex* is contained either in the *VAL\_CHP\_SUP\_01* or *VAL\_CHP\_SUP\_03* column).
- Import CodeAgenda\*: This file is used to do the mapping between the *CodeAgenda* variables and the *CodeSecteur* variable.
- Import RespPropos\* and relative data: This file fills the data about each *RespPropos* variable (*eurlex*). Each *RespPropos* has a name, nationality, national party and party family.
- Import DGProposition\* and SiglesDG\*: This file associates an act with a *DG* variable composed of a *name* and a *sigle* (and a *DG* with a number for old *dg* names, if any).
- Import ConfigCons: This file is used to do the mapping between *ConfigCons* variables and *CodeSect* variables.
- Import AdoptPCAbs and AdoptPCContre: This file links an act to its *AdoptPCContre* and *AdoptPCAbs* variables.
- Import NationGvtPoliticalComposition: This file contains 3 columns: a start date, and end date and a government composition column. It represents the main political parties of each European nation with their respective dates of appearance and disappearance. The government composition column contains the *country* and its list of *parties* and *party families* for a given period of time (defined by the *start date* and *end date* variables). It is used to show the governments in power at the date of adoption of a given

act.

- Import opal file (NP variables): This file links an act to its *NP* variables through its *NoCelex* variable. There is a *CaseNumber*, an *ActType* and an *ActDate* variable.
- Import new Attendance of ministers: This file describes the ministers that attended council meetings. For each council meeting, there is the *country* of each minister, along with their *verbatim* and *status*. Each council meeting is linked to an act through the act ids (*releve ids*) and / or *NoCelex*.
- Import party families of Rapporteurs: link each *Rapporteur* variable with a *PartyFamily*. The file contains matchings between *Party* and *PartyFamily* variables.
- Update Attendance of ministers: Same thing than Import new Attendance of ministers but is used to update already existing ministers' attendances (because they were saved with wrong data for instance). In this case, the link with an act is made with the *NoCelex*.

The web form indicates the variables and their order for each file to import. At the end of the import process, a message is displayed if errors were encountered (in case of already existing record, incorrect data, etc.).

First of all, choose the import to perform in the drop down list. A textual help about the file format to upload appears on the right. Then click on “Choose file” and select the file to import (it must be a csv file). Finally click on “Import” to launch the import. It can take a while...

After the step Import acts to validate, you can validate the ids of the acts you just imported (see Act ids validation page).

## Act ids validation page

Welcome Romain :).

[Go to the bottom](#)

Log out >

CSV import >

**Act ids validation >**

Attendance validation >

Act data validation >

CSV export >

History >

Administration >

**Act to validate**

75 act(s) to validate!

ReleveAnnee=2014, ReleveMois=4, NoOrdre=20 ▼

**Search a validated act to modify**

ReleveAnnee ReleveMois NoOrdre

☒ *releve ids*

ProposOrigine ProposAnnee ProposChrono

☐ *propos ids*

**MODIF**

-	NoCelex	NoUniqueType	NoUniqueAnnee	NoUniqueChrono	ProposOrigine	ProposAnnee	ProposChrono	DosId	url exists?	url link
Index file	32014R0374	COD	2014	90	COM	2014	166		-	-
eurlex	32014R0374	OLP	2014	90	COM	2014	166	None	True	eurlex
oeil	32014R0374	COD	2014	90	COM	2014	166	-	True	oeil
<b>VALIDATION</b>	32014R0374	COD	2014	90	COM	2014	166			<b>Update</b>

Notes:

**SAVE THE ACT**

[Go to the top](#)

You can perform two actions on this page:

- validate ids of an act that has already been imported through a **RMC (releve mensuel du conseil) file** (see Import page): select an act in the drop down list to validate it.
- modify ids of an act that has already been validated: select *releve ids* to find an act with its *ReleveAnnee*, *ReleveMois* and *NoOrdre* variables or select *propos ids* to find an act with its *ProposOrigine*, *ProposAnnee* and *ProposChrono* variables. Fill the corresponding text boxes and click on the MODIF button to validate the act.

Act ids related to the selected act is then displayed regarding its source in row. The first source (row) corresponds to the RMC file (which can be imported – see previous paragraph), the second row corresponds to the ids retrieved on the *eurlex* page and the third one on the *oeil* page. The last row called “Validation” is used to validate and save the true data of the act. By default, it takes the values of the RMC file.

The *DosId* variable does not exist in the RMC file. The drop down list in the *DosId* column displays all the values that match the *NoCelex* of the selected act, from the imported file containing *prelex* unique ids (see Import page). The *DosId* displayed is the one used by default in the validation row. When information do not match (between sources), data is displayed in red, which means that there is an error and an expert is required to validate the act. If there is an error, the expert can change the value of the field. When everything is checked, the act can be saved.

You can also automatically update the ids of an act if you think they have been updated on *eurlex* or *oeil* recently or if one of these sources was not available at the import of the act or if you made a mistake when you first validated it. Click on the Update button at the bottom right corner of the table and wait for the update of the ids of the act.

Select the act to validate in the drop down list (or to modify by entering the corresponding *releve ids* in the text boxes) and check the corresponding information in the table below. Modify fields if

needed in the row “Validation”. You can check or / and add notes regarding the act. Finally, click on “Save” to validate and save the act.

After this step, you can validate the ministers' attendance of any act you have already imported (see below).

## Ministers' attendance validation page

Welcome Romain :).

**Ministers' attendance validation**

This page lets you validate / update ministers' attendance for each act. The data is then saved and displayed inside the act data validation form. Click on the "Update status" button to update all the empty statuses whenever the associated country+verbatim couples are found in the database and have a status.

[Go to the bottom](#)

**Act to validate**

0 act(s) to validate!

Select an act to validate

**Search a validated act to modify**

☒ releve ids

☐ propos ids

Country	Status	Verbatim	Delete
Select a country ▼	Select the status ▼	<input type="text"/>	<input type="checkbox"/>
Select a country ▼	Select the status ▼	<input type="text"/>	<input type="checkbox"/>
Select a country ▼	Select the status ▼	<input type="text"/>	<input type="checkbox"/>

This page displays a list of countries and minister's status and verbatim that attended the council meeting concerning an act to validate. This is possible thanks to the extraction of a pdf whose url is included in RMC files that were imported previously. **In order to validate the ministers' attendance of new acts, you must ask the webmaster to import them first.**

You can perform two actions on this page:

- validate the ministers' attendance of an act that has already been imported (see Import page): select an act in the drop down list and click on the OK button to validate it.
- modify the ministers' attendance of an act whose ministers' attendance have already been validated: select *releve ids* to find an act with its *ReleveAnnee*, *ReleveMois* and *NoOrdre* variables or select *propos ids* to find an act with its *ProposOrigine*, *ProposAnnee* and *ProposChrono* variables. Fill the corresponding text boxes and click on the MODIF button to validate the act.

Information related to the selected act is displayed in 3 columns: country, status and verbatim. For each minister attending the council meeting, we have retrieved his country, verbatim (position title) and status (code created to facilitate the statistical use of the verbatim – see codebook).

You can also update the status by clicking on the Update status button. This will search matches of couples (verbatim+country) and will try to associate the corresponding status to all the same couples (verbatim+country) with no status.



There is the possibility to delete one row by ticking the checkbox Delete in the last column (and clicking on the SAVE button) and adding extra rows using the 3 blank lines at the bottom.

Select the act to validate in the drop down list (or to modify by entering the corresponding *releve ids* or *propos ids* in the textboxes) and check the corresponding information in the table below. Modify fields if needed (choose another country or status in the corresponding drop down lists or change the verbatim). Finally, click on “Save” to validate and save the act.

After this step, you can retrieve information of the acts you just validated for the statistical analysis (see below).

## Act data validation page

**Act ids:**

ReleveAnnee	ReleveMois	ReleveMoisInitial	NoOrdre
2014	2	2	2

**Eurlex:**

NoCelex	url exists?	url link
32014R0254	True	<a href="#">eurlex</a>

<b>TitreEn</b>	Regulation (EU) No 254/2014 of the European Parliament and of the Council of 26 February 2014 on a multiannual consur		
<b>CodeSect01</b>	15.20.10.00	<b>CodeAgenda01=1525</b>	
<b>CodeSect02</b>	Select a CodeSect*	<b>CodeAgenda02=None</b>	
<b>CodeSect03</b>	Select a CodeSect*	<b>CodeAgenda03=None</b>	
<b>CodeSect04</b>	Select a CodeSect*	<b>CodeAgenda04=None</b>	
<b>RepEn1</b>	Environment, consumers and health protection; Consumers; General		
<b>RepEn2</b>			
<b>RepEn3</b>			
<b>RepEn4</b>			
<b>TypeActe</b>	REG		
<b>BaseJuridique</b>	12012E169; 12012E294		
<b>NombreMots</b>	8526		

You can perform two actions on this page:

- validate data of an act that has already been imported (see Import page): select an act in the drop down list to validate it.
- modify data of an act that has already been validated: select *releve ids* to find an act with its *ReleveAnnee*, *ReleveMois* and *NoOrdre* variables or select *propos ids* to find an act with its *ProposOrigine*, *ProposAnnee* and *ProposChrono* variables. Fill the corresponding text boxes and click on the MODIF button to validate the act.

Acts data can be retrieved and then validated on this page. **You can select acts for which ids have been validated only.** Data related to one act is displayed. For each source (*eurlex* and *oeil*), a first table recalls the ids and the url while a second table displays the corresponding retrieved data. The *NationGvtPoliticalComposition* variable, at the end of the *eurlex* table, shows the composition of each national government when the act was adopted. The two other tables show the *ministers'*

*attendance* data (country and status only) and the *opal* data with one row per country.

All the retrieved fields are displayed in row. The first column corresponds to labels and the second one corresponds to retrieved data. If there is an error, the expert can change the value of any field. When everything is checked, the act is ready to be saved.

Select the act to validate in the drop down list (or to modify by entering the corresponding *releve ids* or *propos ids* in the textboxes) and check the corresponding information in the tables below. Modify fields if needed in the second column of the *eurlex* or *oeil* table. *NationGvtPoliticalComposition*, *ministers' attendance* and *opal* data cannot be modified, but *ministers' attendance* data can actually be modified in another form (see Ministers' attendance validation page). You can check or / and add notes regarding the act. Finally, click on “Save” to validate and save the act.

In fact there is one more action you can perform on this page...

## Database management page

On the Act data validation page, after selecting an act, you can add a new *Rapporteur*, *DG*, or *RespPropos* variable:

- in front of each *Rapporteur* variable, you can click on “Add a new Rapporteur\*” to add a new *Rapporteur* variable (*name*, *country* and *party*) .
- in front of each *DG* variable, you can click on “Add a new DGProposition\*” to add a new *DG* variable (*dg* and *dg\_sigle*) .
- in front of each *RespPropos* variable, you can click on “Add a new RespPropos\*” to add a new *RespPropos* variable (*name*, *country*, *party* and *party\_family*) .

After this step, the last step is to export the acts you already imported and validated in a csv file (see below).

## Export page

This is the export page. Once acts have been imported, ids validated and data retrieved, you can choose this tool to export them under the csv format. This action exports acts saved and validated only.

To download the validated file, click on the Export button to start the download.

You can also get the latest export of the database by clicking on the link at the bottom of the form. This file is exported on the server automatically, every Sunday. Choose this option if you have problems exporting the acts.

After this step, acts are ready for the statistical analysis :).

## History page

Date	Time	Action	Form	Act	User
Nov. 12, 2014	6:26 p.m.	modif	attendance	ReleveAnnee=2009, ReleveMois=3, NoOrdre=21	romain.lalande
Nov. 7, 2014	3:40 p.m.	add	data	ReleveAnnee=2014, ReleveMois=2, NoOrdre=1	selma.bendjaballah
Nov. 7, 2014	3:34 p.m.	add	ids	ReleveAnnee=2014, ReleveMois=4, NoOrdre=17	selma.bendjaballah
Nov. 7, 2014	3:32 p.m.	add	attendance	ReleveAnnee=2014, ReleveMois=4, NoOrdre=19	selma.bendjaballah
Nov. 7, 2014	3:31 p.m.	add	attendance	ReleveAnnee=2014, ReleveMois=4, NoOrdre=18	selma.bendjaballah
Nov. 7, 2014	3:30 p.m.	add	attendance	ReleveAnnee=2014, ReleveMois=4, NoOrdre=17	selma.bendjaballah
Nov. 7, 2014	3:29 p.m.	add	attendance	ReleveAnnee=2014, ReleveMois=4, NoOrdre=16	selma.bendjaballah
Nov. 7, 2014	3:28 p.m.	add	attendance	ReleveAnnee=2014, ReleveMois=4, NoOrdre=15	selma.bendjaballah

This page shows a history of the last 100 actions performed, useful for validators to help them

remember their most recent tasks.

There are 6 columns:

- the *date* of the performed action
- the *time* of the performed action
- the *action* itself: *add* when an act is added (selection of an act from the drop down list) or *modif* when an act is modified (after entering its *releve ids* or *propos ids* in the textboxes and clicking on the Modif button)
- the *form* used: *ids* for the Act ids validation form, *data* for the Act data validation form and *attendance* for the Ministers' attendance validation form
- the *act* column displays the *releve ids* of the act
- the *user* column displays the user who has performed the action

## Fields to retrieve

The explanation of each field used in the application is given in the **codebook**.