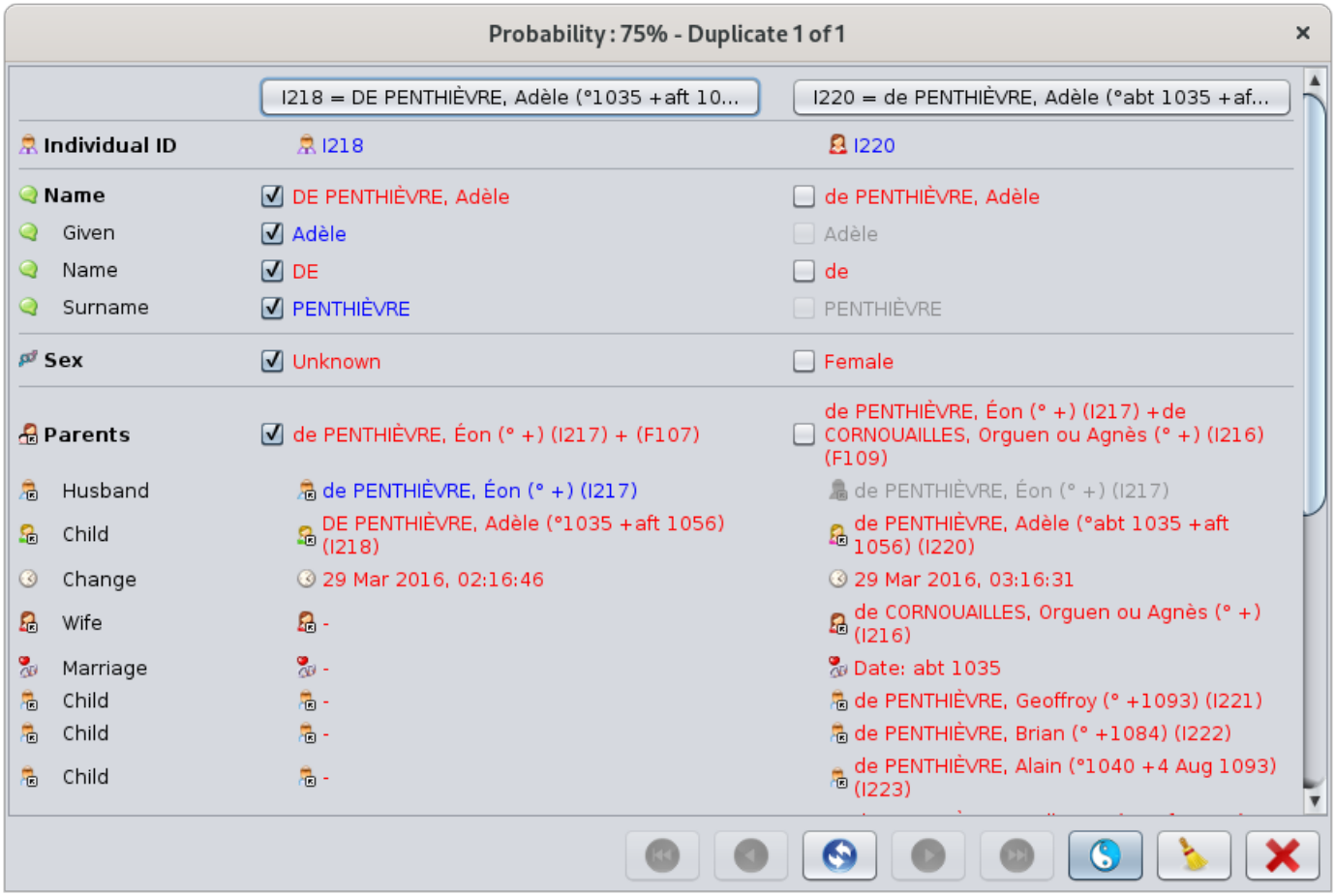
Merge duplicates

This tool detects duplicate entities and allows you to merge them.



Description

> This tool gives the list of entities likely to be duplicates, from the most certain pair of duplicates to the least certain pair of duplicates, by category of entity. For each pair of similar entities, Ancestris gives you a similarity percentage.

This tool does not give 100% certain duplicates. Even a human being can sometimes have difficulty certifying that two individuals are the same or certainly not.

Of course, one could stop at saying that two individuals with exactly the same surname, first name and date of birth are duplicates. But one of these pieces of information may be missing for one of the individuals, or it may be imprecise.

What you are expecting from Ancestris is to alert you and say :"***It is not certain, but given the similarities in the information between these two individuals, they might be duplicates. And this is the level of confidence that they are***". Then it's up to you to decide.

That is the purpose of this tool.

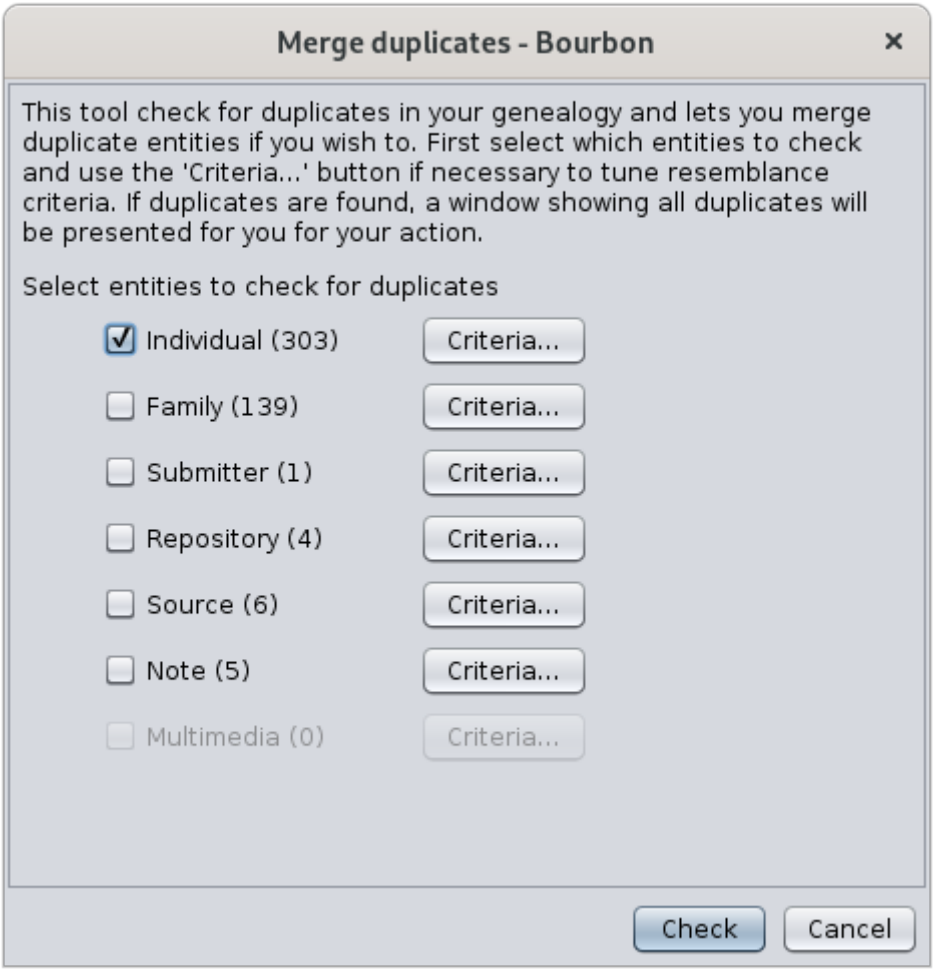
Usage

The duplicate merge tool works in two steps.

First you specify the **detection criteria**, then you choose how to **merge duplicates**.

Detection criteria

When the tool is launched, the criteria selection window is displayed.

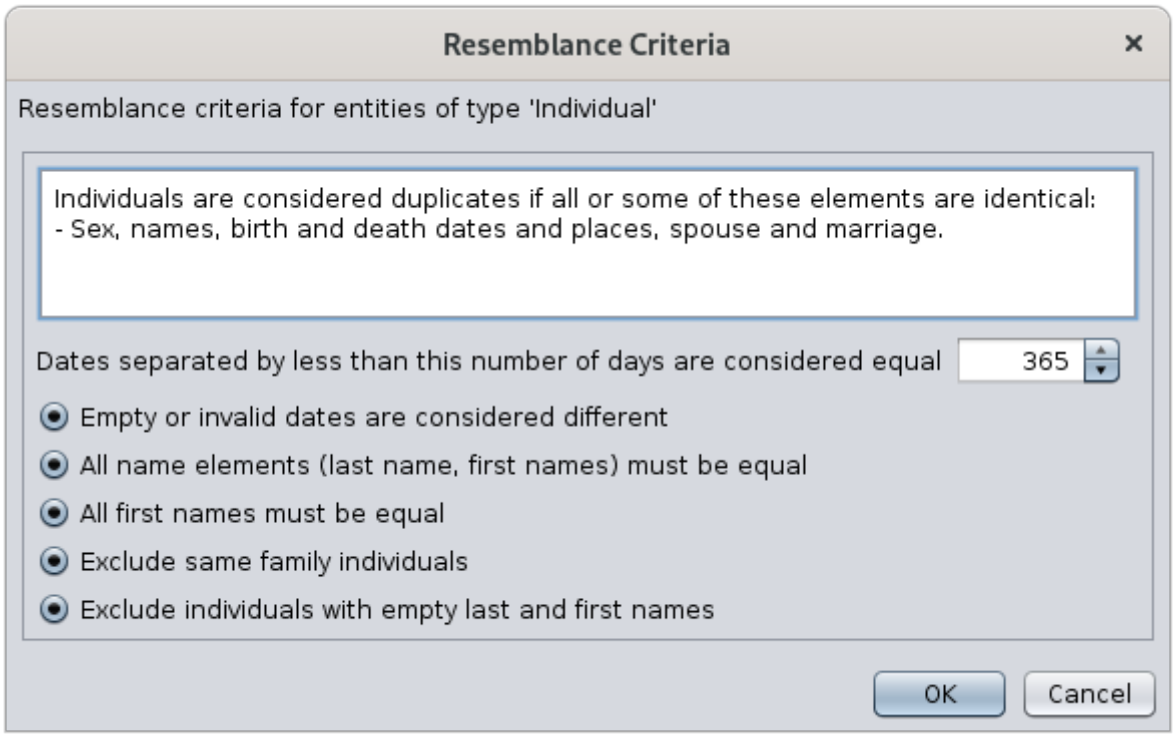


Check the entities for which you want to search for duplicates.

Only the entities that are present in the Gedcom file are available. In the example above, as there are no media entities, the corresponding Criteria button is unavailable.

Then check one by one the detection criteria for each category of entity.

The most sophisticated criteria are those of individuals. Here they are.



The criteria are as follows.

Identical dates

When are two dates considered identical? When their difference in number of days is close or zero.

If you indicate 365 days for example, i.e. 1 year, two dates will be equal if their difference is less than a year.

If you indicate 30 days, two dates will be equal if they differ by less than a month.

Empty or invalid dates

If a known date is compared to an unknown date, Ancestris will consider them different.

Name elements

Forces all elements of the name to be identical. Conversely, can be identical if only some elements of the name are identical.

First names

Forces all first names to be identical. Conversely, can be identical if only some first names are identical.

Exclusion of individuals from the same family

Individuals from the same sibling or parent-child relationship are not compared.

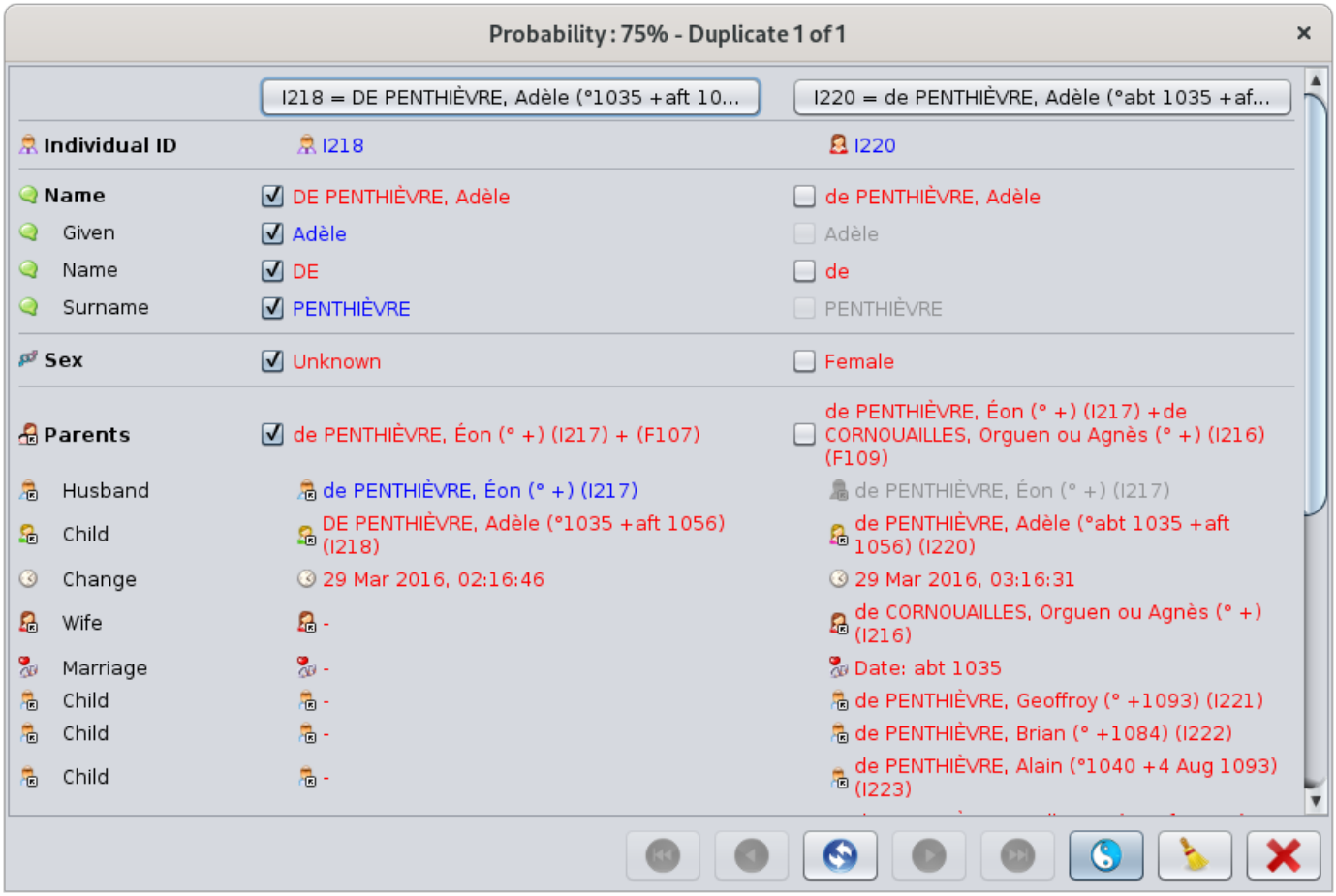
Exclusion of individuals without first or last name

Individuals without first or last names are not compared.

The criteria for other entities are either a sub-part of these criteria or are not modifiable.

Merge window

After starting the search for duplicates, the following window appears.



Window

The title of the window indicates the duplicate pair number displayed and the confidence that the two entities of this pair are in fact the same, and therefore to be merged.

The two entities of the supposed duplicate pair are in the two columns.

A button allows to select each of the entities in the editors for more details.

For each property of the entities, the window displays the values of the property for each of the two entities of the supposed duplicate.

In red are displayed the values that are different.

In blue are displayed the identical values for the left entity, in grey for the right entity.

> The purpose of the comparison is to merge the right entity into the left one.

For this purpose the check boxes select the information of each entity to keep after merging.

The buttons at the bottom navigate within the duplicate pairs, merge them or ignore them.

Toolbar



Go to first duplicate Button

Displays the first duplicate in the order of the confidence index, i.e. the most likely duplicate.

Go to previous duplicate Button

Displays the previous duplicate.

Swap Left and Right Entities Button

Swap the left and right entities in order to merge the two entities on the left one. This is useful if most of the information to be kept after the merge is on the right hand side.

Go to next duplicate Button

Displays the next duplicate.

Go to last duplicate Button

Displays the last duplicate in the confidence index, therefore the least likely duplicate.

Remove duplicate Button

Removes the potential duplicate from the displayed list.

If the duplicate search is restarted, it will reappear.

Close Button

Closes the window.

Merger

By clicking the Merge button, the left entity is removed from the Gedcom file and the information checked on the right is added to the left entity.

For information that can only exist once (e.g. birth), it is only possible to keep the information from one of the two entities.

As soon as the merge is done, the window displays the same duplicate with the result of the merge so that you can check that everything has been kept as you wanted.

You can then move on to the next duplicate.

Customization

The personalization elements are the criteria.

The criteria used are stored for the next time.

There is no other customization option.

Revision #1
Created 14 November 2021 14:36:34 by frederic
Updated 3 June 2023 14:56:43 by frederic