**User manual for data encoding in Darwin excel import template**

# Summary

The excel template has been developed to enter data in a temporary file, before importing them to DaRWIN database of Africa Museum.

Template contains 7 sheets with all main fields that could be used in a specimen collecting expedition.

Data are entered in the sheet with the use of 2 different forms (a minimal and a more complete) called by buttons on first sheet.

2 other buttons on first sheet are there to create files needed to import in DaRWIN and to check taxonomy with a web service.

# General presentation

DaRWIN database has been developed ten years ago at the Institute of Natural Sciences of Brussels.

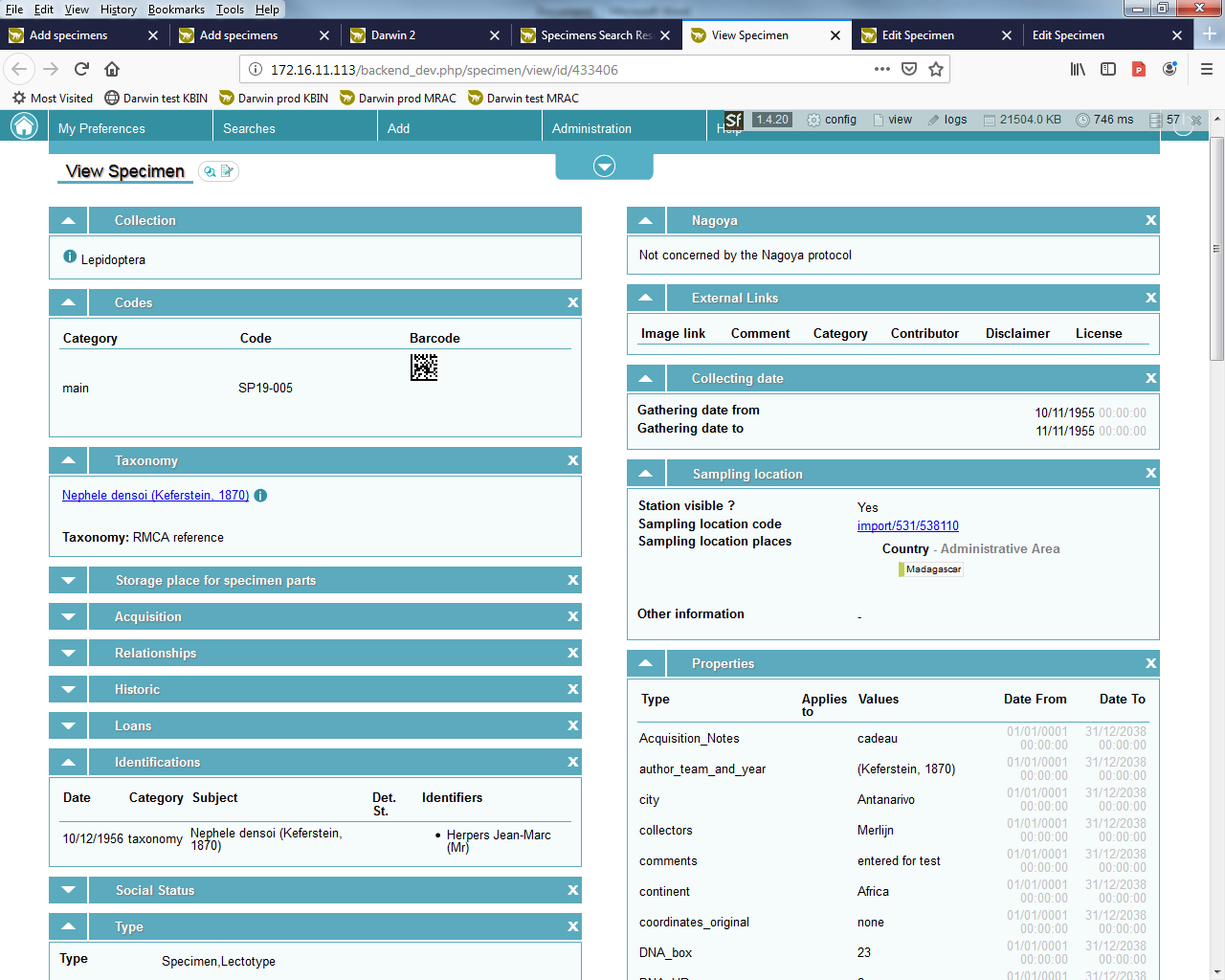
Aim of this database was to manage all the collections of the institute : biology, paleontology, geology,…

A quite complex interface has therefore been developed on basis of a PostgreSQL database. That interface is an intranet site developed in Symfony 1.4 framework(PHP)

5 years ago, Africa museum was interested by that tool and began to use it, but only for biological collections.

Development goes further in the interface, while new collections are introduced regularly in DaRWIN.

Data encoding can be done directly in Darwin, specimen by specimen or via an excel containing a lot of specimens.



Example of a specimen in Darwin (small part of all the fields…)

Main problem to encode data on the field is dependency of DaRWIN on an internet connection because the database is on a remote server. A simple tool had to be developed to be able to encode locally data, on a simple laptop without web connection.

It’s the reason for the development of an excel allowing to enter data on the field.

# Template presentation.

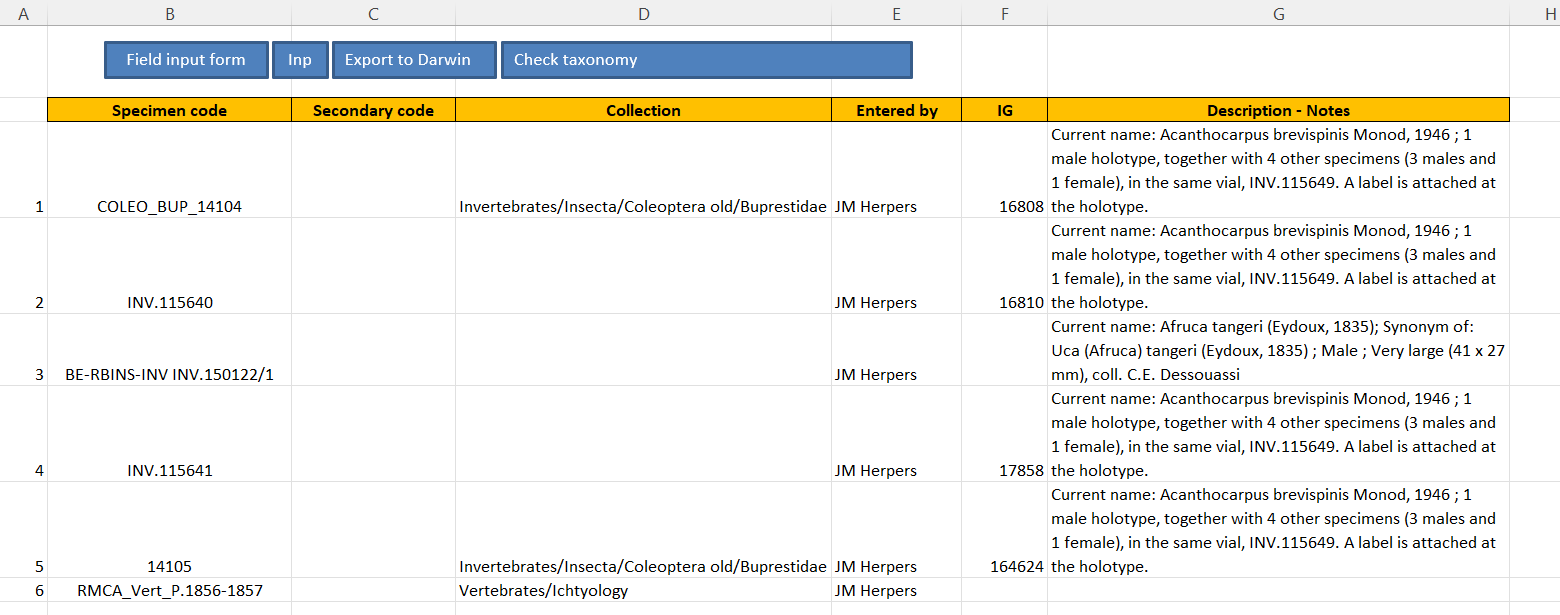
Template has been developed in Excel 2010.

It contains 7 different sheets : Code, Location, DNA, Ecology, Taxonomy, Counts\_Storage and Acquisition.

Particularity of this excel template is that it can be filled in by forms. On the first sheet, you will find 4 buttons. A complete description of these buttons will be given in the next section. But already now, notice than data must by preference be added in the excel by the forms. Following description of all the fields of the excel is only there to explain you the function of these fields. You can also check data on these sheets when you have finished to enter data with the forms.

## Code.

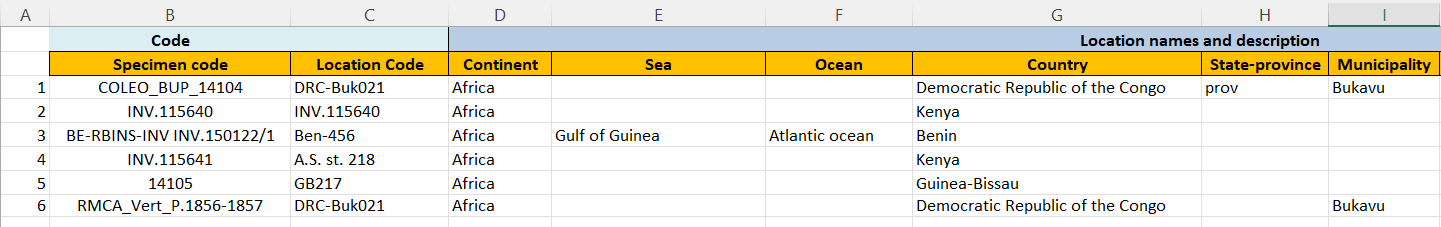
This sheet contains general information.

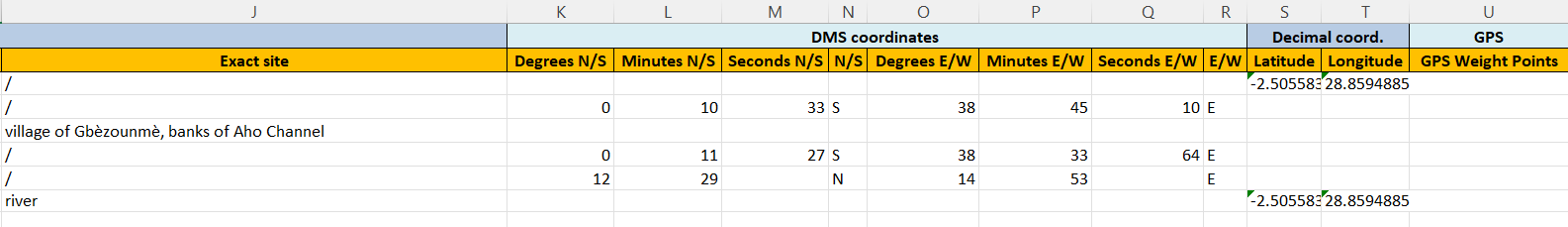


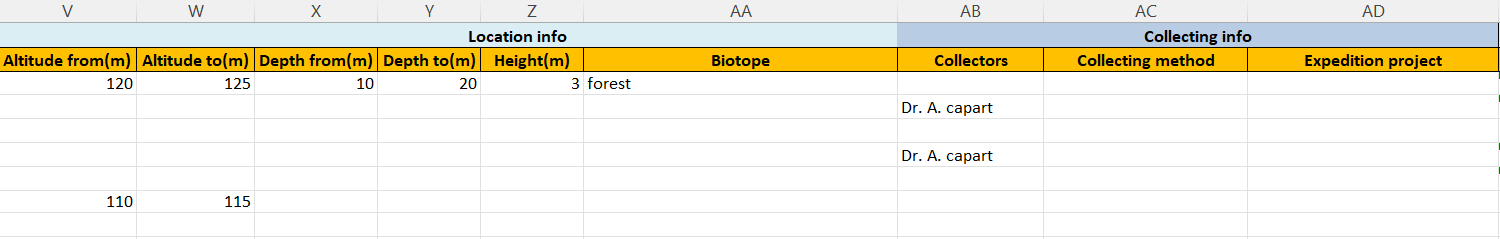
* 2 code fields. **Specimen code** can contain a field code or the final code that will be used in DaRWIN, if you already know it. **A secondary code** can be entered if necessary
* **Collection** : The DaRWIN collection
* **Entered by** : the person who encoded the specimen in this excel
* **IG** : The Inventory General code.
* A last field, **Description - Notes**, is there for the notes. You can enter here data that you couldn’t enter in another column. There is a Notes field on every sheet.

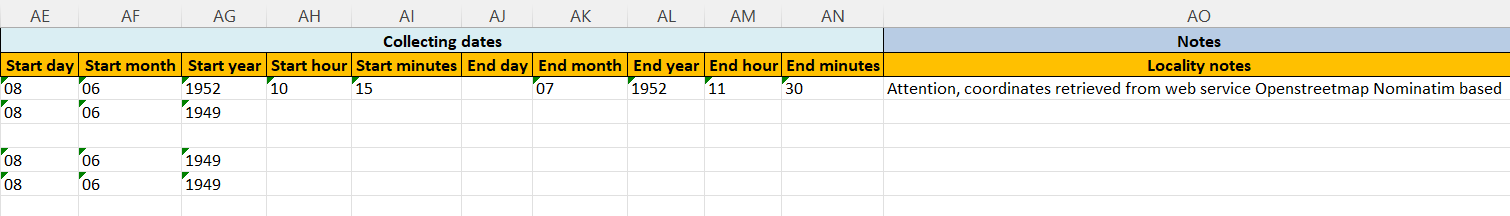
## Location.

This sheet contains all info’s about the sampling place and methods





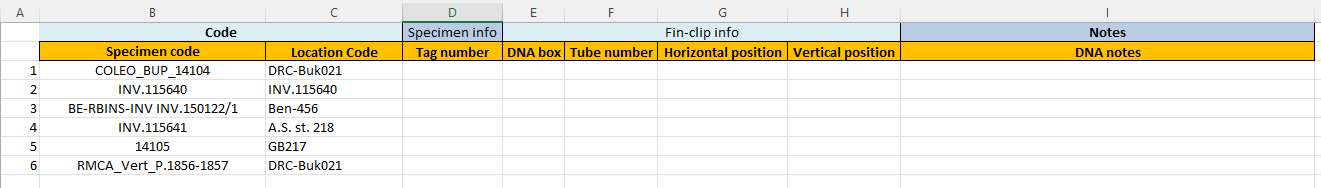




* **Code** : As on every following sheet, the 2 first columns are the specimen code and the location code. This will allow you to find more easily the specimen.
* **Location names and description**: Description of the sampling place, country, city,…
* **DMS coordinates** : Fields containing the coordinates in degree-minutes-seconds, with indication of N/S, E/W
* **Decimal coord**. : latitude and longitude in decimal degrees.
* **GPS** : You can enter here remarks about GPS position
* **Collecting info** : altitude, depth, height (if specimen in a tree for example), biotope, collectors, collecting method and expedition name.
* **Collecting dates** : From and To dates and hours of the collecting period.
* **Notes**

## DNA.

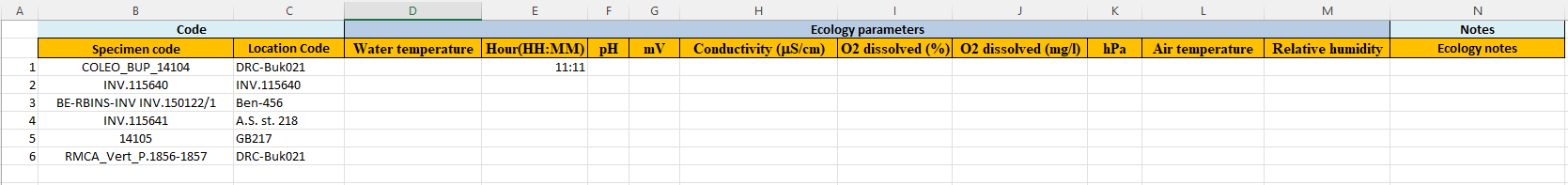
This sheet contains some info about a DNA number and container of DNA sample



* **Specimen info** : contains the tag number, to make the link between DNA sample and the specimen
* **Fin-clip info** : Info about the container : ID of box, tube, position
* **Notes**

## Ecology.

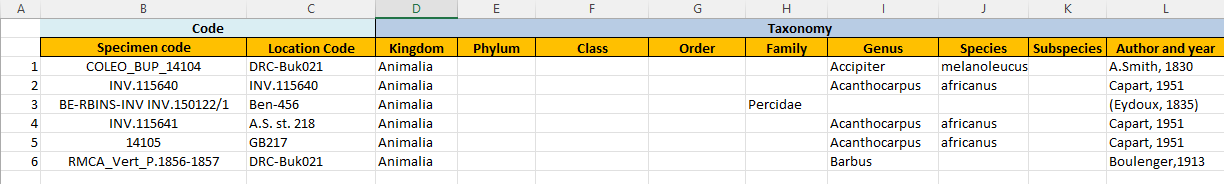
This sheet contains a lot of ecological parameters

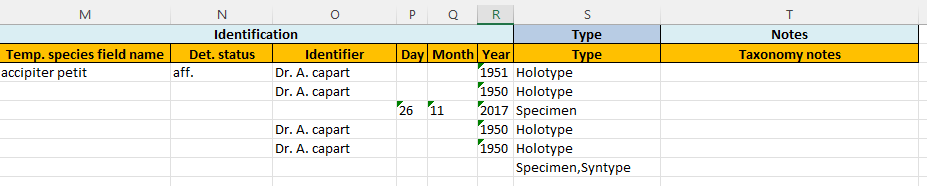


* **Ecology parameters** contains parameters used mainly for fresh or marine waters but also for other organisms.
* **Notes**

## Taxonomy.

This sheet contains taxonomy and identification info

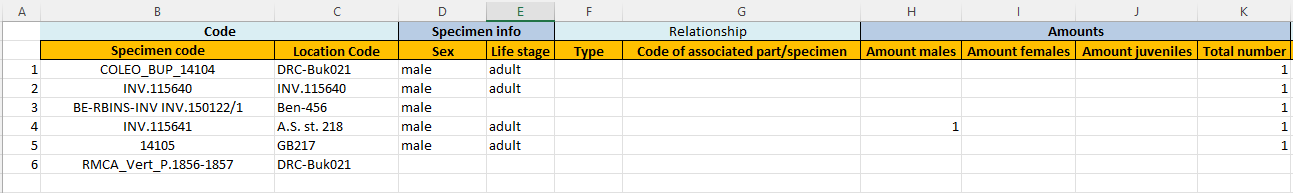


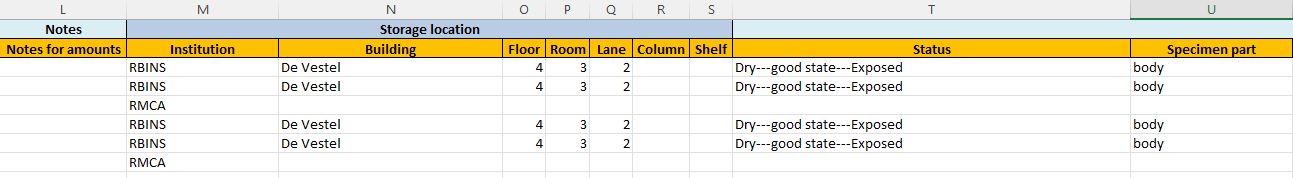


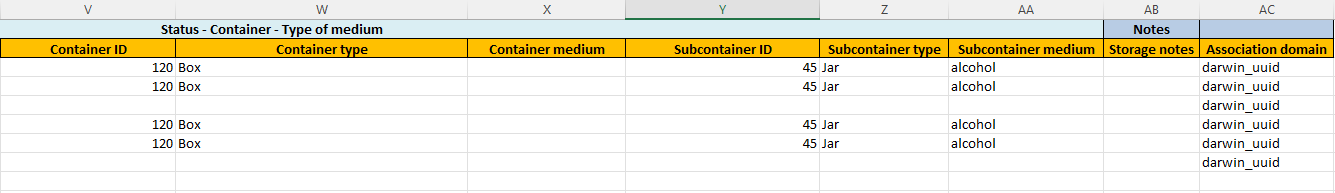
* **Taxonomy** : Taxonomic levels, from Kingdom to Subspecies. Last field contains Author and year.
* **Identification** : A first field, **Temp. species field name**, can contain an uncertain taxon or even a non taxon text as “blue tilapia with white spots”. It’s only a temporary field to give a name on the field. Other fields are determination status (aff., sp., cf., v., nomen nudum,...), name of identifier and date of identification.
* **Type** : specify whether specimen is a type or not
* **Notes**

## Counts\_Storage.

This sheet contains various info’s about the specimen, amounts and where it is stored.







* **Specimen info** : 2 fields to enter the sex and life stage of the specimen
* **Relationship** : If the specimen is host or parasite of another species, you can use these fields to specify the species. Code can be used if it’s already in DaRWIN (it’s the DaRWIN code)
* **Amounts** : number of males, females, juveniles, total. If you have a male juvenile (so, 1 specimen), write 1 in male and 1 in juvenile but write 1 also in total (not 2).
* **Notes** : some notes about the amounts.
* **Storage location** : Place where the specimen is stored
* **Status - Container - Type of medium** : A specimen can be cut in different parts and these parts can be stored at different places, the body somewhere and the head at another place for example. Each part of the specimen must be entered separately, as it was a different specimen. So a fish with the whole body and the tail cut has to be entered on 2 lines, one for the body and one for the tail.

In **status**, you write how the specimen is prepared, in which state it is and whether it is at his storage place or not (it can be used in the museum, be exchanged with another institution,…).

**Specimen part** contains part of the body

**Container ID** is the identification of the container

**Container type** : the container can be a box, a jar, a tube,…

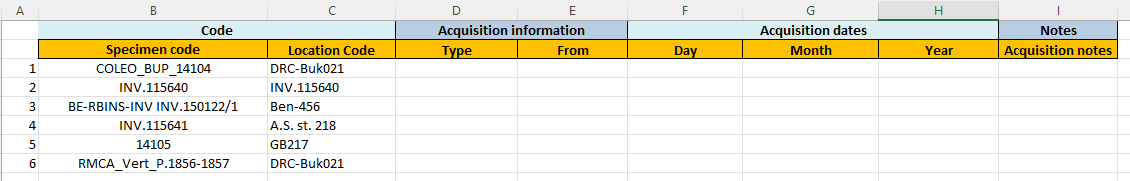
**Container medium** : How is the specimen preserved in the container : in alcohol, glycerin, medium of the microscope slide,…

The 3 following fields are used if there is a subcontainer (small tubes contained in a jar for example).

* **Notes** for the storage.
* **Association domain** : more technical field filled automatically if forms are used. Normally contains “Darwin\_uuid”.

## Acquisition.

This last sheet contains information about specimens that could be obtained by another way than sampling on the field



* **Acquisition information** : Which kind of acquisition (donation, specimen found on holidays, purchase,…) and who gave the specimens.
* **Acquisition dates** : When specimens are acquired.
* **Notes**

# Buttons

As already said in section 2, data should be entered in this excel via forms.

These forms can be called on the first sheet that contains 4 buttons:



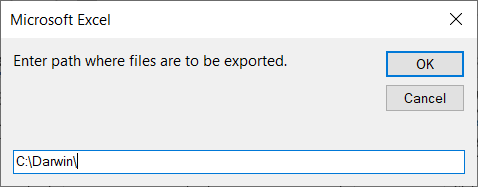
The first 2 buttons open 2 different forms:

- **Input form** is a complete form with all the fields from the different sheets, to be used mainly in the institution.

- **Field input form** is a more limited form that can be used on the field, to enter only a limited amount of data.

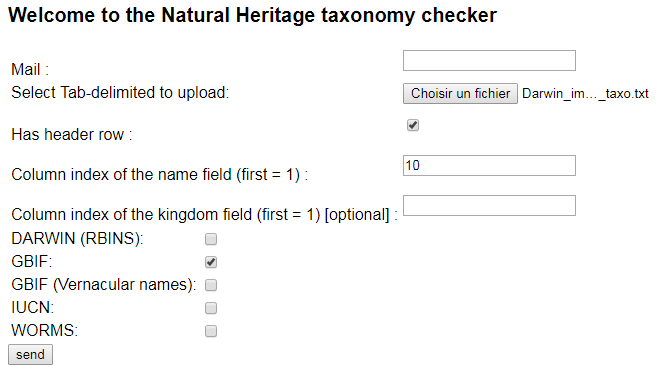
The 2 other buttons are :

- **Export to Darwin**. When you have finished entering the data, you only have to click on that button to export the files needed for a Darwin importation. A popup appears asking you where you want to save the files :



When you click on OK, 3 files are generated : *Excel name*\_taxo.txt, *Excel name*\_Location.txt, *Excel name*\_specimens.txt. These 3 CSV files are used in Darwin to import respectively taxonomy, sampling location and specimen data.

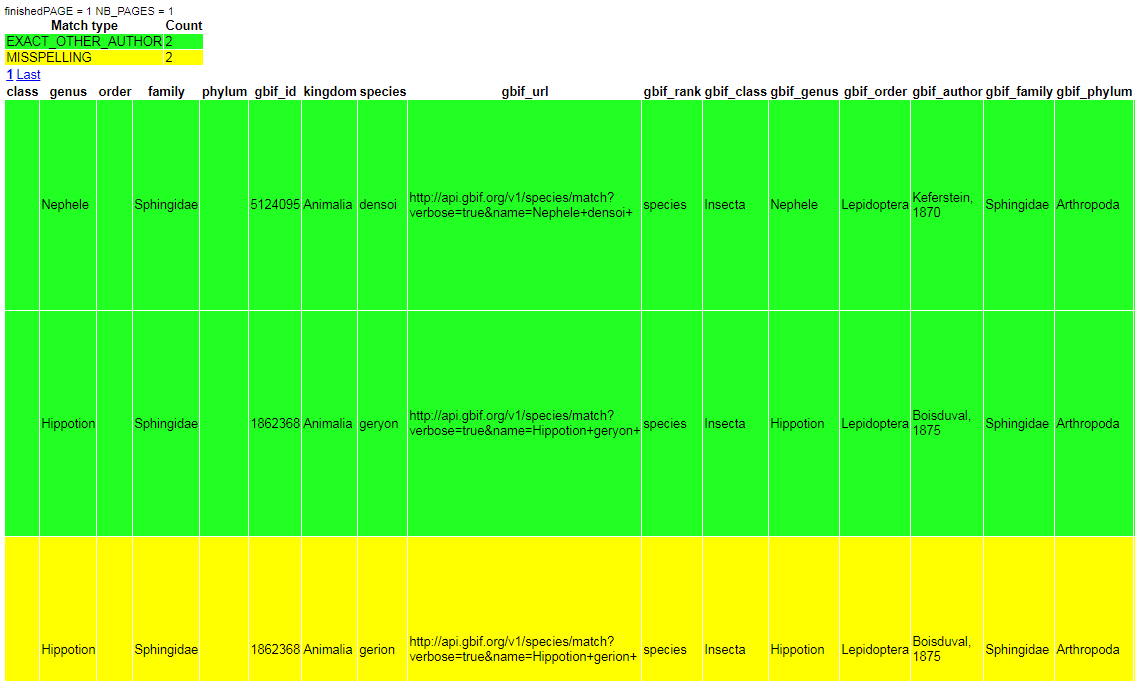
- **Check taxonomy**. This last button is used to check taxonomy with the help of a web service. Once you have exported files with the above procedure, you can click on this button to open a web page allowing you to check taxonomy against taxonomy of GBIF, WORMS, IUCN and also DaRWIN. Following page is opened:



On this page, you need first to choose the file concerning taxonomy that you have exported (*Excel name*\_taxo.txt). Then, write 10 in column index of the name field (full taxon name is in column 10 in the file)

Choose which taxonomy you want to check and click on send.

You will receive a page similar to this:



You receive a table with color code. Color code is explained at top of the page. Green means here for example that taxon was found back in GBIF. Yellow means that it was found but with a small problem (here name of species was misspelled). White is when taxon was not found at all.

# Forms

Forms are used to enter data in the sheets. They are called by buttons on first sheet.

## 5.1 Field input form

This form is for use on the field. So, amount of fields is limited.

You can find back the field description in section 3 They are gathered on 1 screen to facilitate encoding.

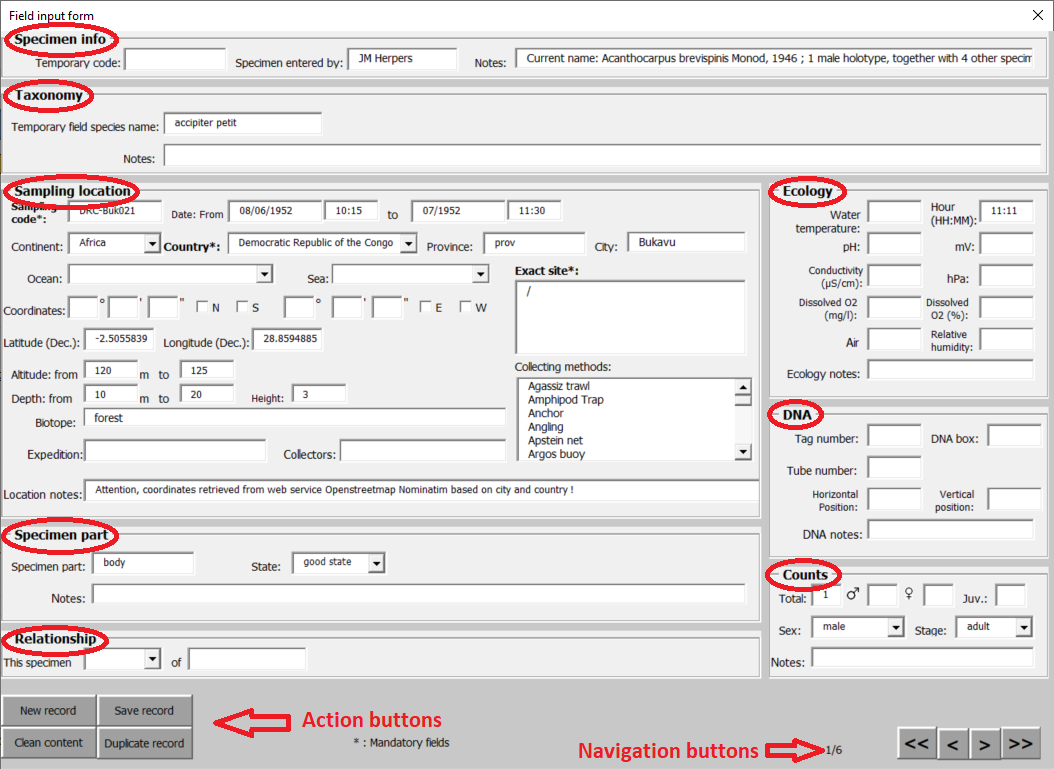
You can navigate from field to field with the mouse or with the tabulation.

Most of the fields are text fields. Some are combo boxes with predefined lists. You can receive errors on some fields if you entered incorrect data, as 10/35/2010 for a date or 95° for a latitude degree. A popup appears then to mention the error.

Only 3 fields are mandatory : sampling code, country, exact site.

If you have data that you can’t enter in a predefined field, write it in the note field available in every section (for example historical name of a country, that should be noted in Location notes)

As said in section 3, each part of the specimen must be entered separately. In this form, enter name of the part and state.



When you open the form, data of first line from each sheet are shown.

To navigate in the excel lines, click on the buttons at right bottom of the form (4 buttons : First line, previous line, next line, last line)

Action buttons:

* “Save record” : Each time you have finished editing the form, click on button “Save record” at left bottom of the form, otherwise changes won’t be sent to the excel sheets. Attention, this doesn’t save the entire document!
* “New record” : To create a new row in the sheets.
* “Duplicate record”: If you want copy all data from 1 row(of every sheet) click on “Duplicate record”. Record will be saved and form emptied and filled with a copy of the data that you can edit. Don’t forget to click on “Save record” at the end.
* “Clean content” only empties all the fields of the form, it doesn’t touch the data of the sheets.

## 5.2 Input form

This form contains all the fields mentioned on the sheets and should be used if you have already every data for DaRWIN. As there are many fields, there are 2 tabs in place of one , plus a tab “search”. This form can be used as the Field input form.A screenshot of a computer

Description automatically generated with medium confidence

Tab 1 “General info” of input form

A screenshot of a computer

Description automatically generated with medium confidence

Tab 2 “Secondary info” of input form

If there are many records, it could be useful to use the tab “Search”. Some fields have been selected, with the “free search” field to search for records.

Graphical user interface

Description automatically generated

Tab 3 “Search” of input form

If records are found, new buttons appear at the bottom, replacing the normal navigation buttons. Navigation is then done in the results of the search. Edit the records as usual. To exit the search results, go on tab “Search” and click on Reset.

