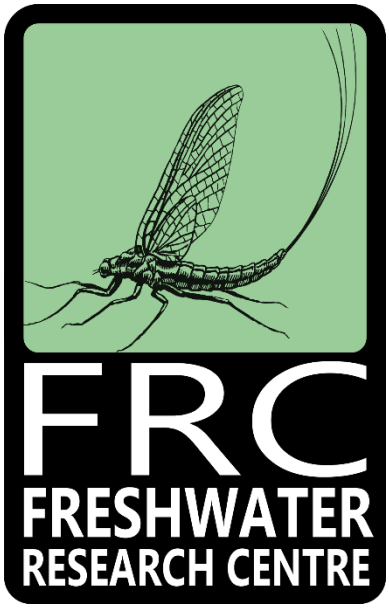
Administration guidelines for the Rwanda Biodiversity Information System (RBIS)

Helen Dallas



October 2020

Prepared for the Center of Excellence in Biodiversity and Natural Resource Management, University of Rwanda

Table of Contents

1 Introduction

These document serves as a guideline for the administration of biodiversity data in the Rwanda Biodiversity Information System (RBIS). The sections outlined in this document are intended to serve as a guide for RBIS administrators and biodiversity data administrators to mobilise and ingest biodiversity data into RBIS. The steps outlined are sequential and each provides details on the process and key considerations. It is the intention to include each of these steps in the online administration in RBIS.

Only registered users with **super user status**, typically the site administrators, are able to view the administration sections in RBIS and undertake the following steps related to the mobilisation and ingestion of biodiversity data into RBIS.

The sections covered include:

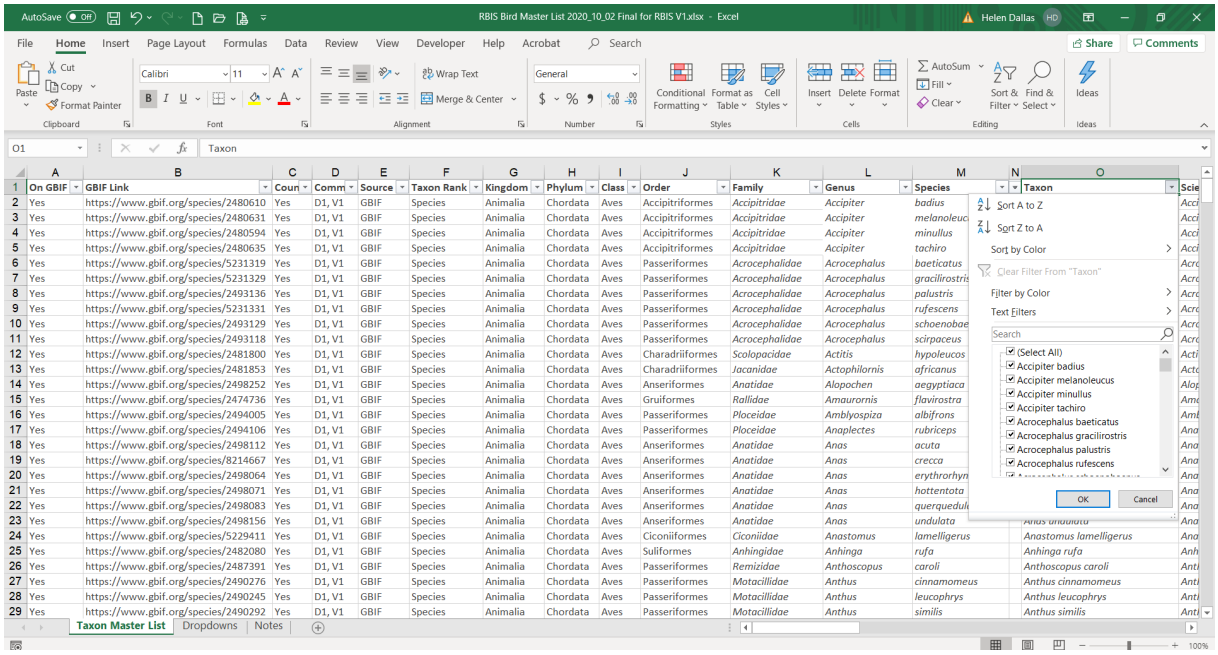
- Preparing and checking a Master List of Taxa before uploading to RBIS
- Preparing and checking an Occurrence Data File before uploading to RBIS
- Uploading a new Taxon Group (Module) and adding a Master List of Taxa for the Taxon Group
- Uploading Occurrence Data
- Harvesting GBIF Data
- Managing taxa in Taxon Management

2 Preparing and checking a Master List of Taxa before uploading to RBIS

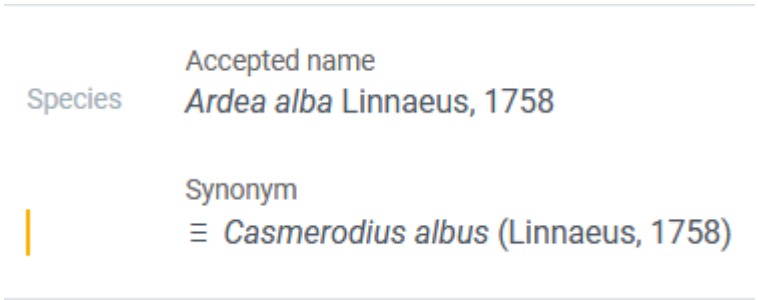
A taxonomic Master List is a list of all species and /or taxa within a particular group such as birds, fish, invertebrates, wetland plants, algae, etc. The purpose and details of the Master list have been outlined in the Data Management Guidelines (Dallas 2020a), and this section in intended to highlight issues specific checks to improve accuracy of the Master List. The format of the Master List is important to ensure consistency for ingestion of data into the information system. The columns included in the Master Lists are detailed in the Data Management Guidelines (Dallas 2020a).

To ensure the Master list is accurate, several steps should be taken before uploading taxonomic data to RBIS. After consolidating the master list, you should check the following:

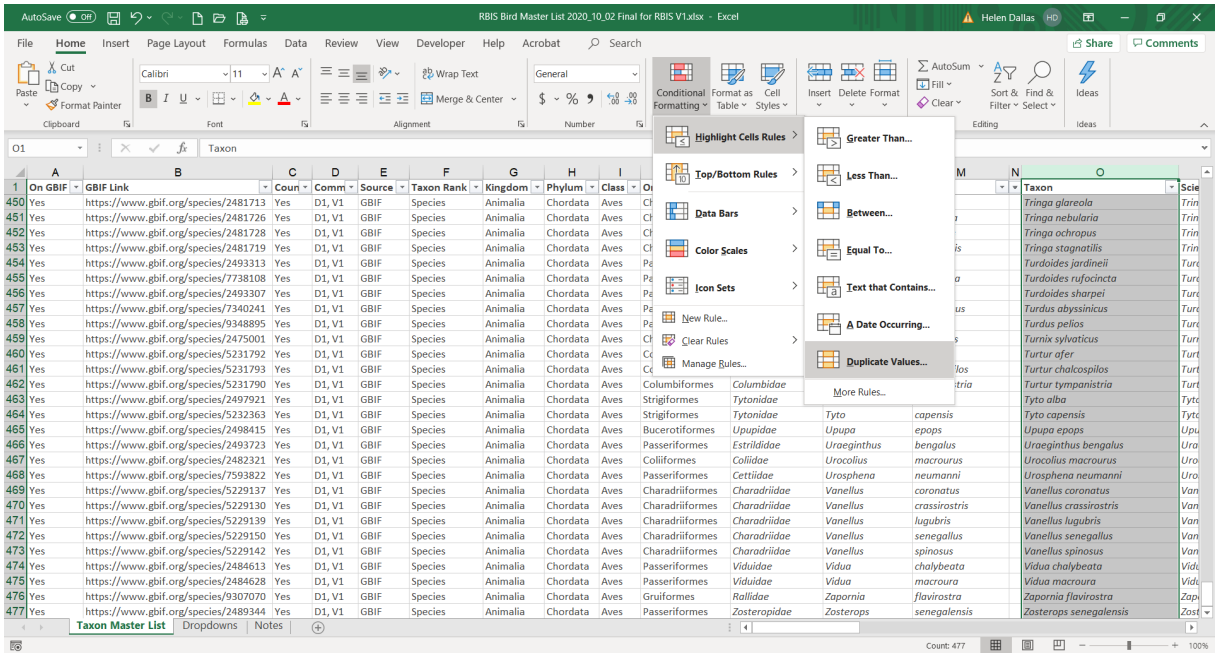
Apply filters for checking the data by highlighting the header row, clicking **Data, Filter.** All columns should be checked for consistencies and typos. Systematically work from column A to W. In particular, check consistency of the Taxon Rank and taxonomic hierarchy (Kingdom, Phylum, Class, Order, Family, Genus, Species, SubSpecies, Taxon).



It is important to **check the GBIF taxonomy for accepted names and synonyms**. For example, in the avian master list, *Ardea alba* - is the accepted name, whereas *Casmerodius albus* is the synonym. Preferably only accepted names should be included in the Master List of Taxa.

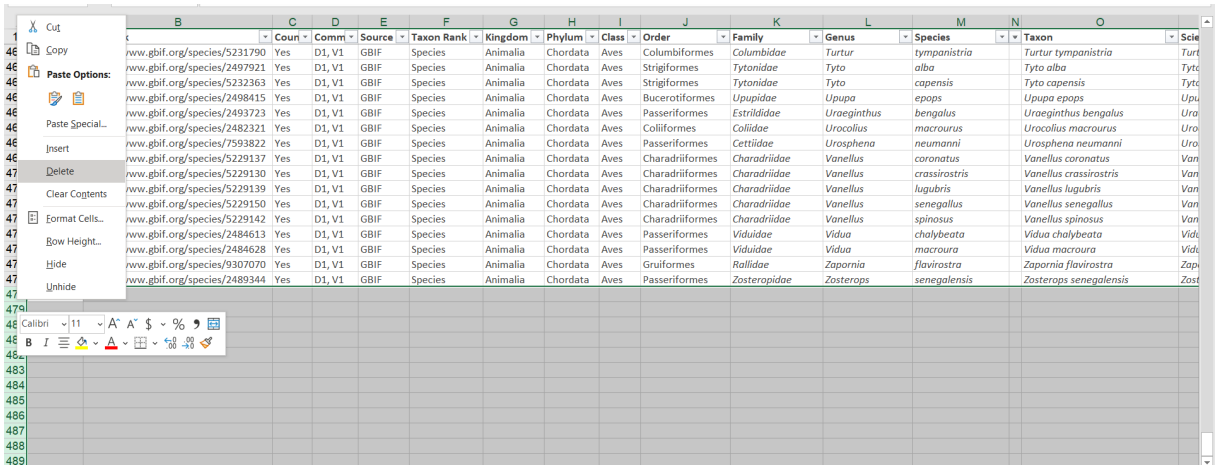


Taxa should be **checked for duplicates** by highlighting the Taxon column, and from the Home Menu, selecting **Conditional Formatting, Highlight Cells Rules, Duplicate Values.**



Note: All taxa can be updated after ingestion into RBIS through the Taxon Management section (see section 7 in this document).

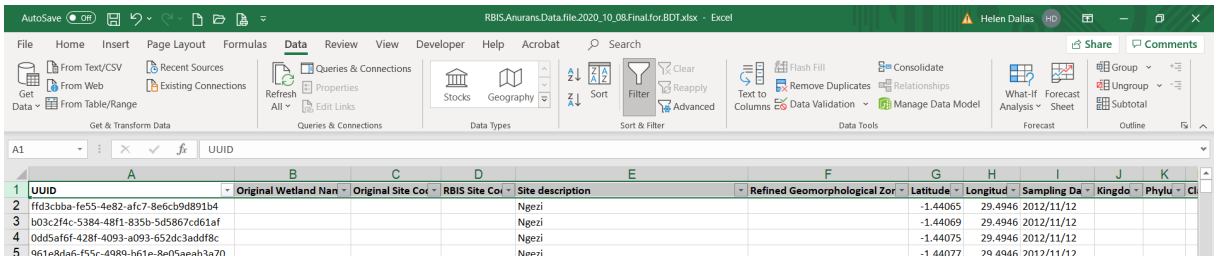
Delete blank rows and columns. Lastly, ensure that there are no extra blank rows or columns, by deleting them.



3 Preparing and checking an Occurrence Data File before uploading to RBIS

To ensure that data are accurate, several steps should be taken before uploading occurrence data to RBIS. After consolidating the occurrence data in the data file, you should check the following:

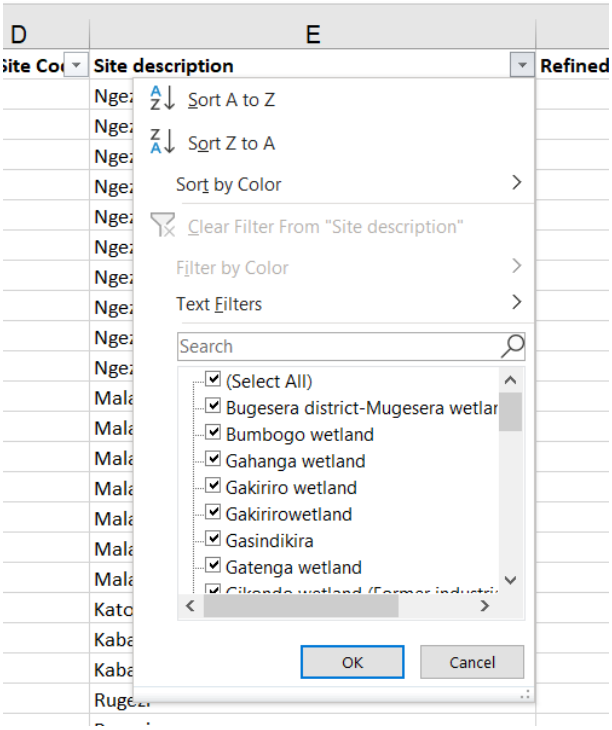
Apply filters for checking the data by highlighting the header row, clicking **Data, Filter**

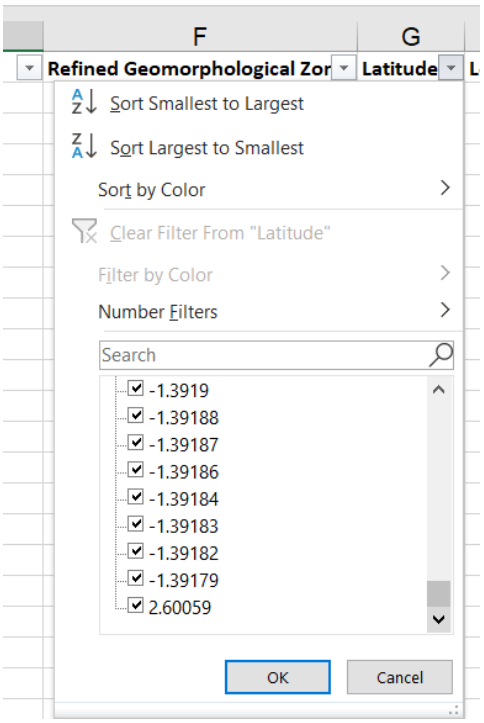


UUID. This is a unique if for each occurrence record. It needs to be copied and pasted so that the formula sued to generate it is saved as a number. See below for guidance.

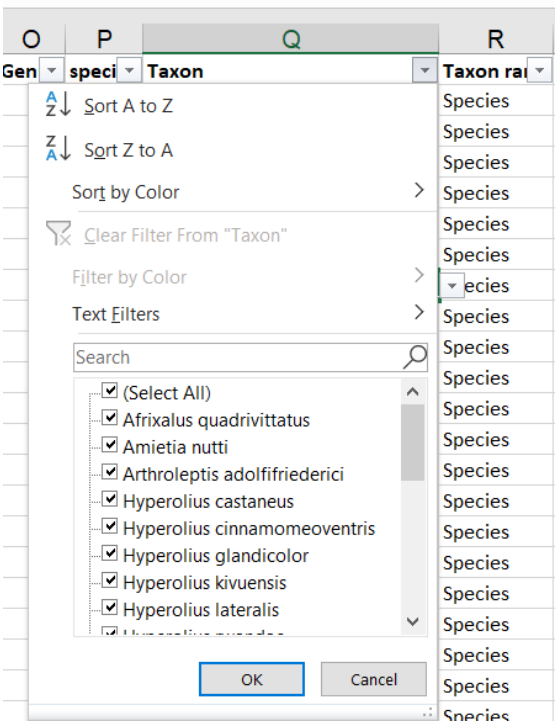
https://docs.rbis.kartoza.com/batch-importing-taxon-occurrence-data-bims/dealing-unique-identifiers/

Systematically check each column using the dropdown arrows, and look for inconsistencies. Some common issues include, #num in UUID column instead of the UUID, incorrect spelling in the **Site description** column (e.g. Gakiriro wetland, Gakirirowetland), latitude with missing "-" (e.g. 2.60059 as latitude is incorrect – should be -2.60059), longitude.

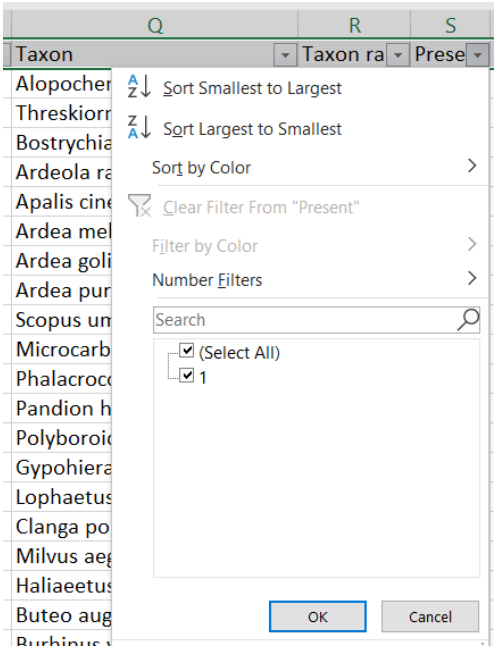


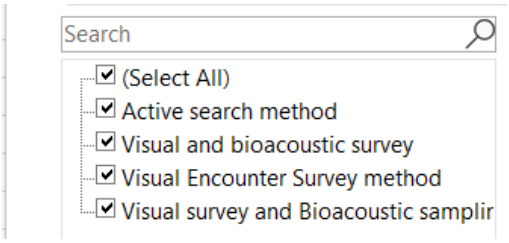


Check that all taxa are correct and are present in the Master List. If the dropdown of master taxa list was used then this should not be an issue. Check that the Taxon rank is correct.

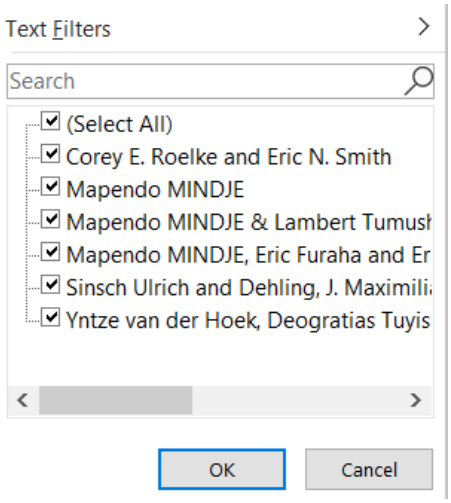


Check presence is all "1", check **sampling method** is correct.





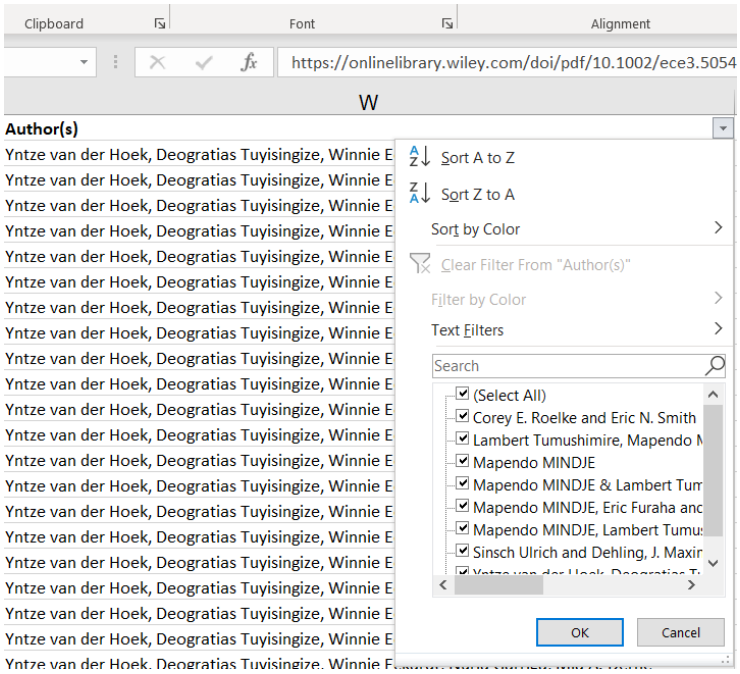
Check Collector/Owner and Collector/Owner Institute. Ideally CAPITALS should not be used, First name Surname if known. Do not use middle initial and punctuation.



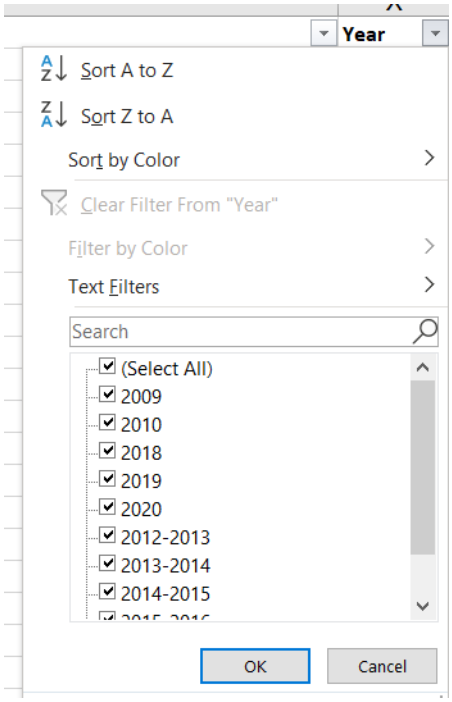
Check the metadata (Author(s), Year, Source, Title, Reference category, URL, DOI, Document Upload Link). For each study reference type, you need to populate the following columns:

- Peer-reviewed scientific article (Collector/Owner; Collector/Owner Institute; Author(s); Year; Source; Title; DOI or URL (if DOI is not available). For Peer-reviewed scientific article the Source is the Journal, For Peer-reviewed scientific article the Title is the title of the article.
- Published report (Collector/Owner; Collector/Owner Institute; Author(s); Year; Source; Title; URL or Document Upload Link). Note the Document Upload Link is obtained after the report is added to RBIS.
- Thesis (Collector/Owner; Collector/Owner Institute; Author(s); Year; Source; Title; URL or Document Upload Link)
- Database (Collector/Owner; Collector/Owner Institute; Author(s); Year; Source)
- Unpublished data (Collector/Owner; Collector/Owner Institute; Author(s); Year; Source)

Check format of Author(s). It needs to be: Surname + Initials, no punctuation. (e.g. Tumushimire L, Mindje M, Sinsch U & Dehling JM not Lambert Tumushimire, Mapendo MINDJE, Prof. Ulrich Sinsch & Julian Maxmillian Dehling). It is important to get the authors correct (e.g. Sinsch Ulrich and Dehling, J. Maximilian, Lümkemann Katrin, Rosar Katharina, Christiane Schwarz should be Sinsch U, Lümkemann K, Rosar K, Schwarz C & Dehling M as per the doi).

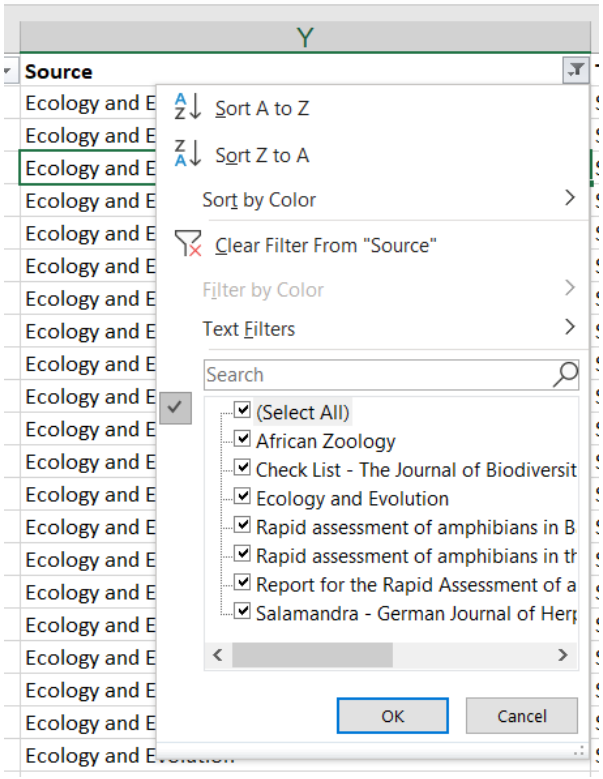


Check the Date: This is the publication date (so 2012-2013 should be 2019 as this is when the article was published - Ecology and Evolution. 2019. Same with all other data from this study).

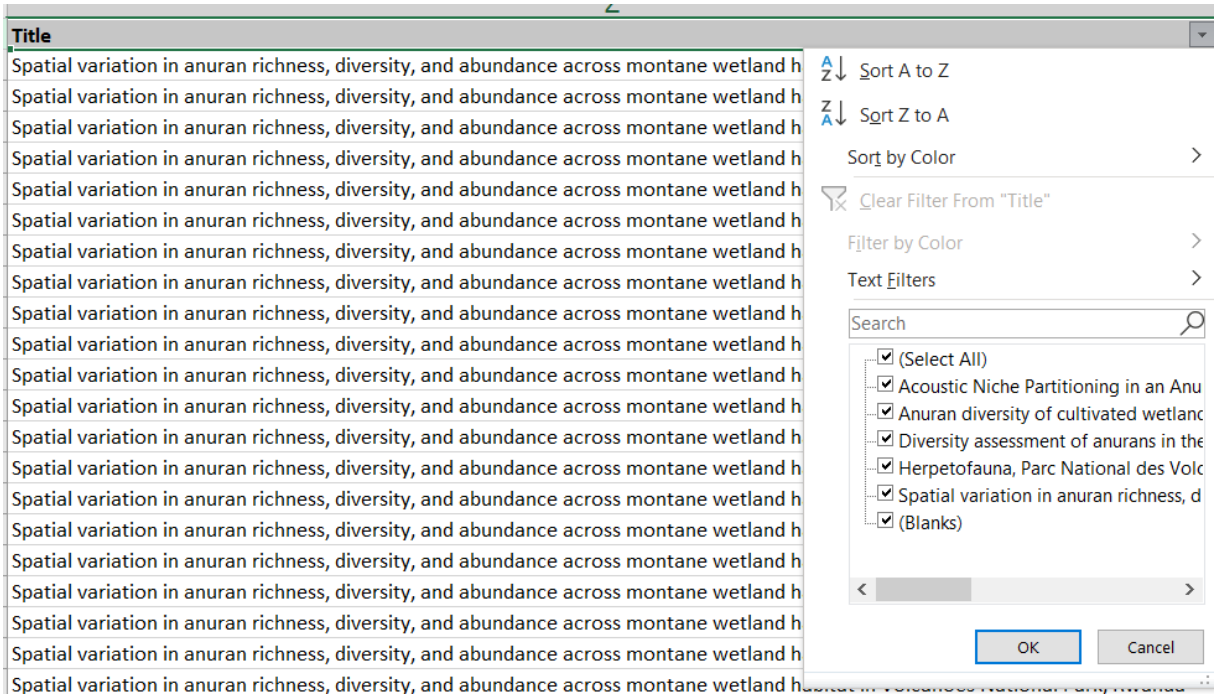


Check the Source. Please note when to include source or not, and what to include. (e.g. Mindje, M., Tumushimire, L., & Sinsch, U. (2020). Diversity assessment of anurans in the Mugesera wetland (eastern Rwanda): impact of habitat disturbance and partial recovery. Salamandra, 56, 27-38. Should be **Salamandra**)

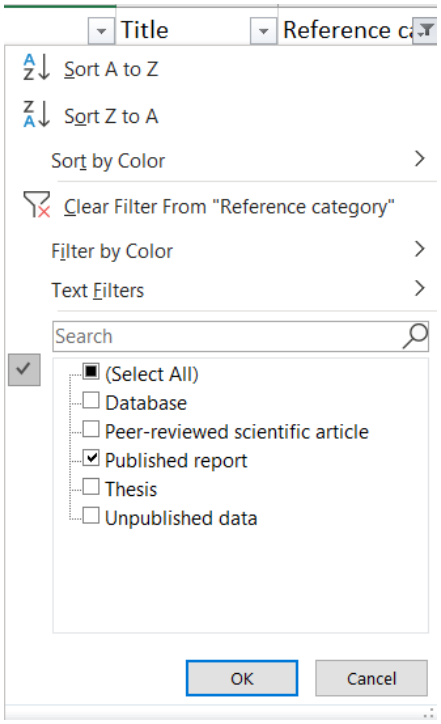
- For Peer-reviewed scientific articles - the Source is the Journal.
- For Published Reports and Theses - the Source is the publisher of the Report.
- For Unpublished Data - the source is the title of the study.



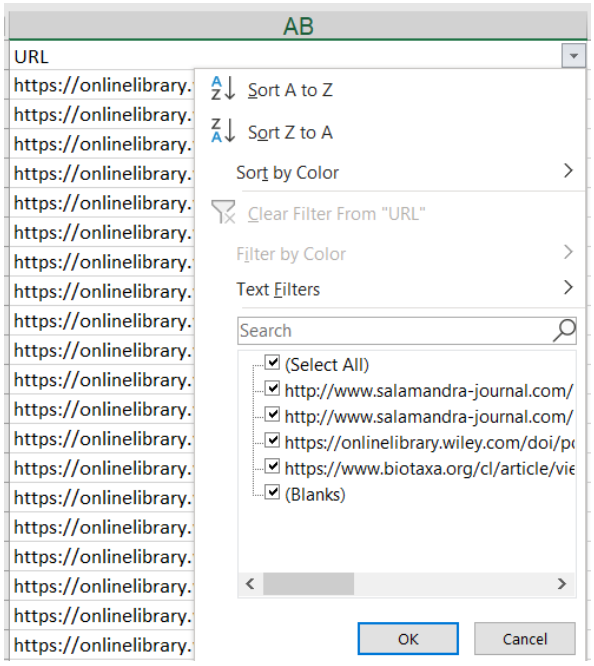
Check the Title. For Peer-reviewed scientific article the Title is the title of the article, for Published reports or theses, it is the title of the thesis. Unpublished data don't need a title.

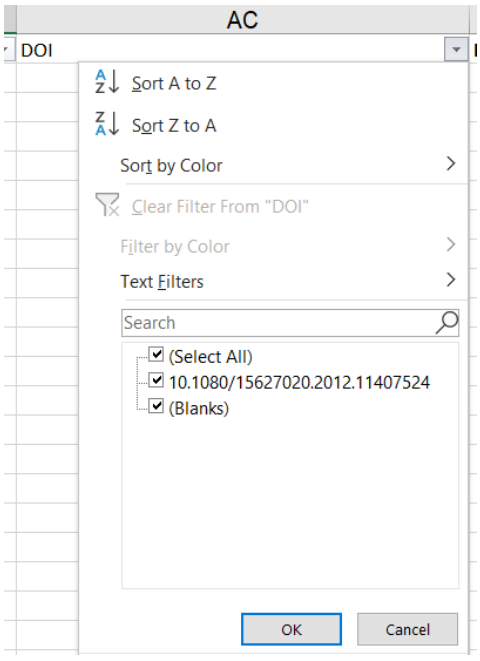


Check all Reference Categories are correct: options include:



Check URL and DOI. Use a DOI if it is available, URL – only needed for Peer-reviewed scientific article if there is no DOI. For the DOI you only need to include the number part, so 10.1080/15627020.2012.11407524, not https://doi.org/10.1080/15627020.2012.11407524.



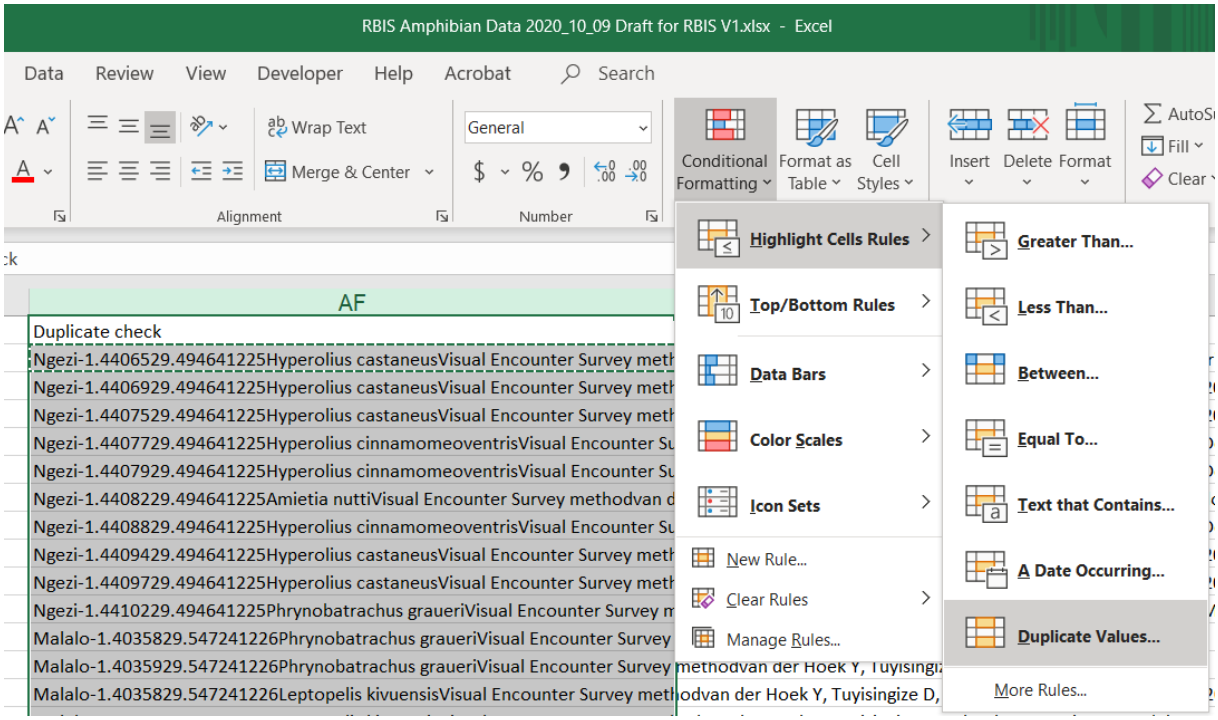


Check the document upload link is correct. Note the Document Upload Link is obtained after the report is added to RBIS. Reports are only uploaded when there is no DOI or URL to link the data to.

Checking for duplicate occurrence records. Use this formula for checking for duplicates. This is a combination of Site description, latitude, longitude, sampling date, Taxon, sampling method, author, year, source and title. Copy and paste the formula below into a new column at the end and name it "Duplicate check".

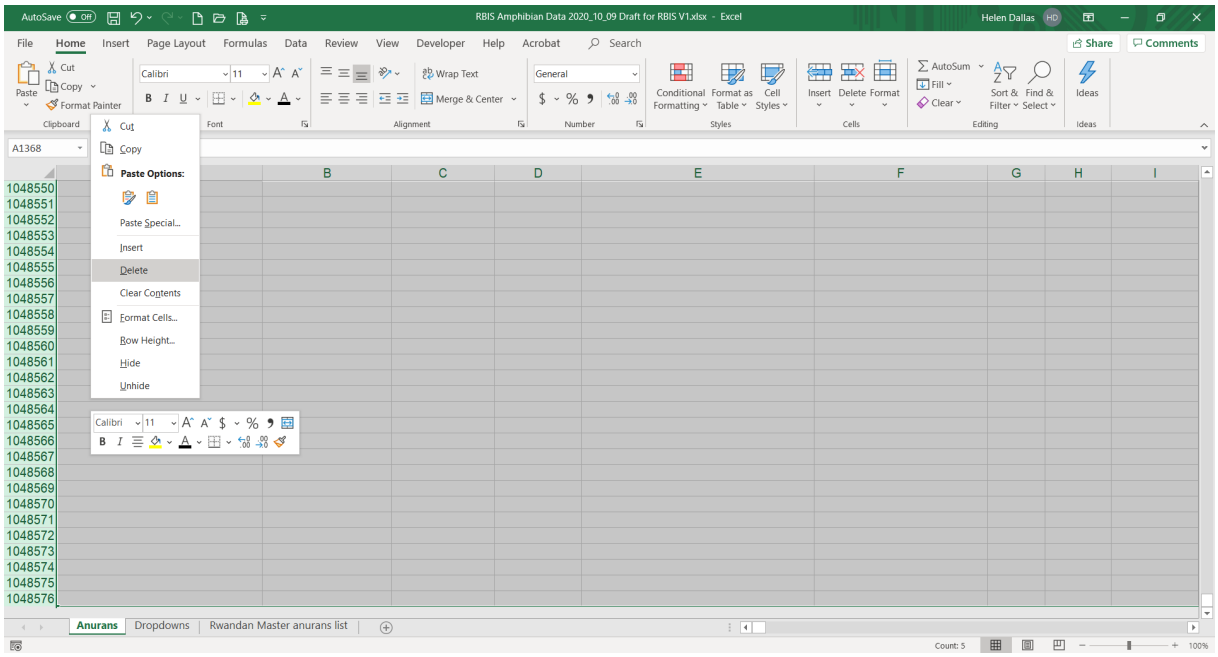
=CONCATENATE(E2,G2,H2,I2,Q2,T2,W2,X2,Y2,Z2)

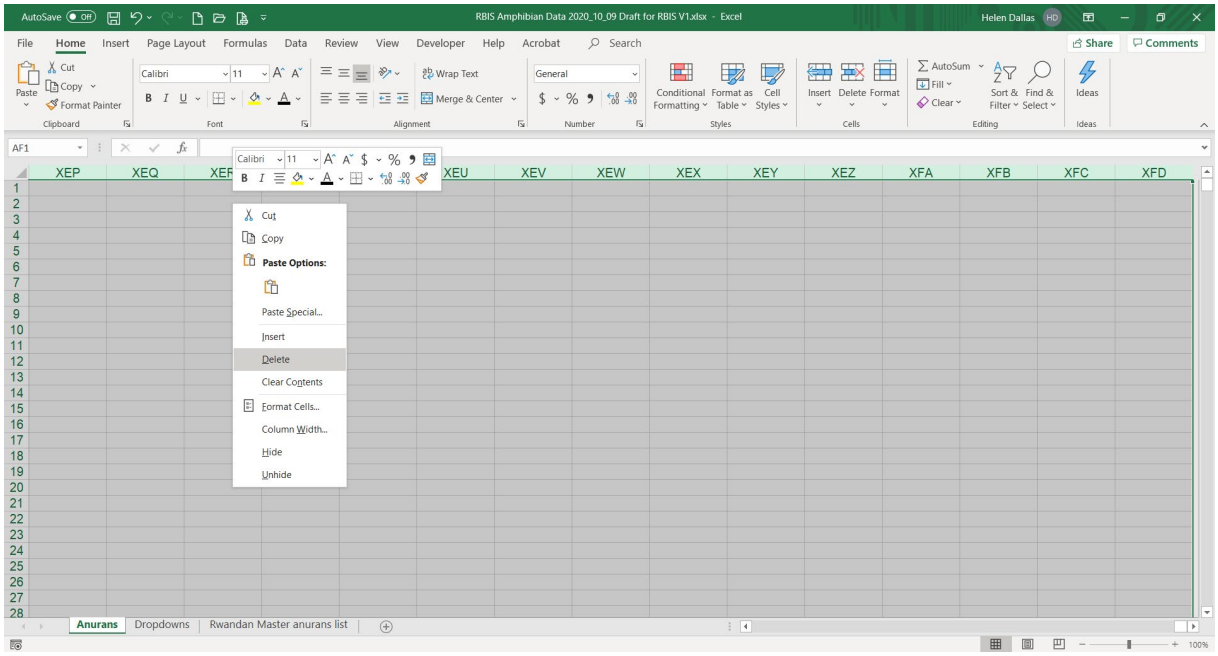
Then copy and paste the formula down to the end of the data rows. Then Highlight the column, and from the Home menu, select **Conditional Formatting, Highlight Cells Rules, Duplicate Values.**



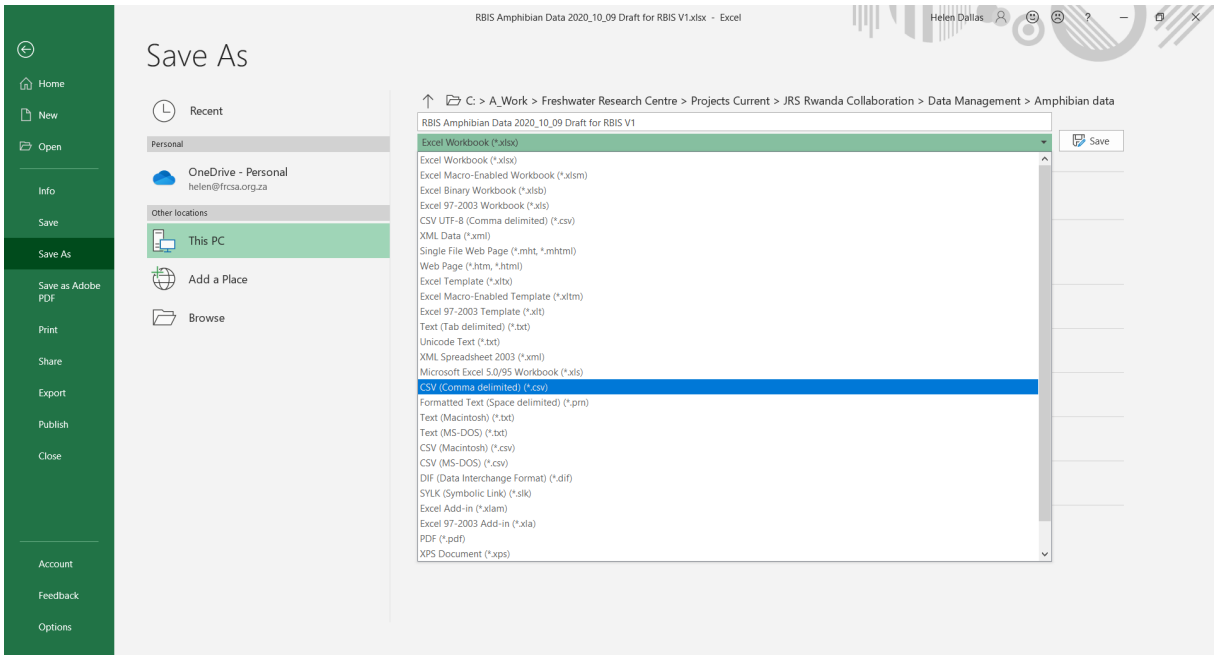
Any duplicates will be highlighted. Check and delete duplicate occurrence records. Then delete the Duplicate Check column.

Delete blank rows and columns. Lastly, ensure that there are no extra blank rows or columns, by deleting them.


RBIS Amphibian Data 2020_10_09 Draft for RBIS V1.xlsx - Excel
Cut
Copy
Paste Options:
Paste Special...
Insert
Delete
Clear Contents
Format Cells...
Row Height...
Hide
Unhide
Anurans
Dropdowns
Rwandan Master anurans list


RBIS Amphibian Data 2020_10_09 Draft for RBIS V1.xlsx - Excel
Cut
Copy
Paste Options:
Paste Special...
Insert
Delete
Clear Contents
Format Cells...
Column Width...
Hide
Unhide
Anurans
Dropdowns
Rwandan Master anurans list

Remove the data filter, save the file in excel, and save the file as csv file.



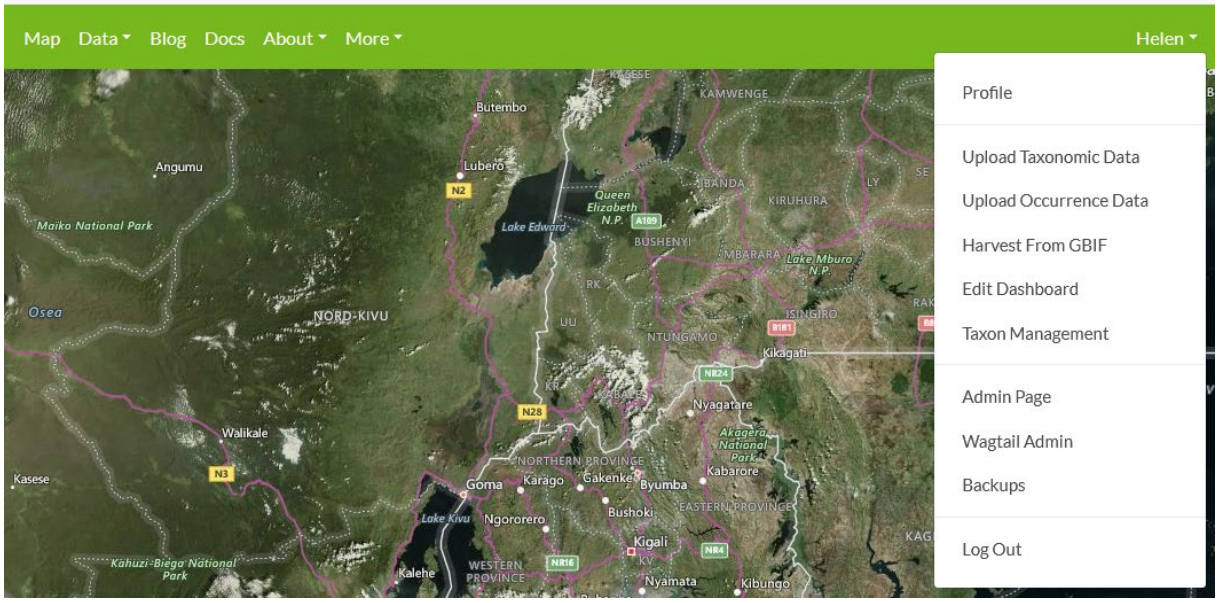
Data are always first uploaded to the RBIS Staging Site to make sure it is all working, and to check for errors.

4 Uploading a new Taxon Group (Module) and adding a Master List of Taxa for the Taxon Group

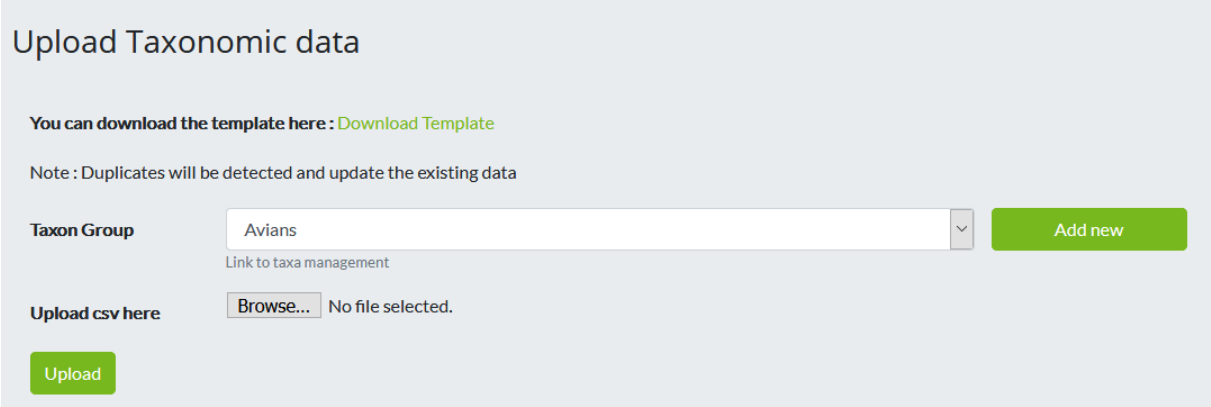
Only registered users with **super user status** are able to do this, typically the site administrators .

Steps:

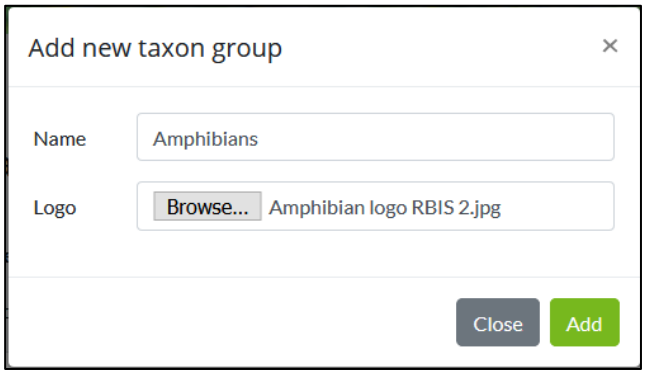
Click on your name and select **Upload Taxonomic Data**



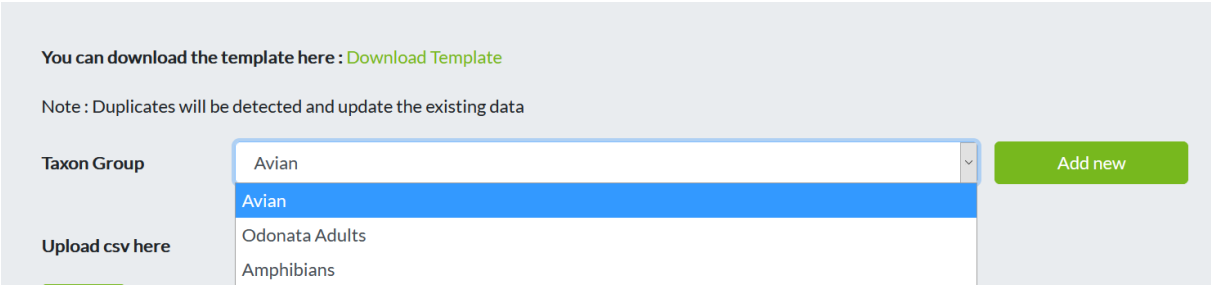
Click Add new – to add a new Taxon Group



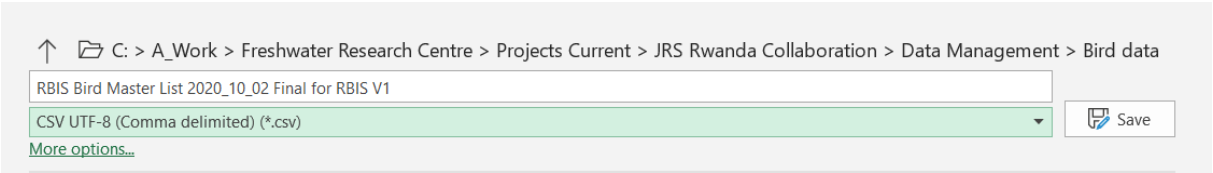
Add the new taxon group by adding the Taxon Group name and selecting the logo using the "Browse"

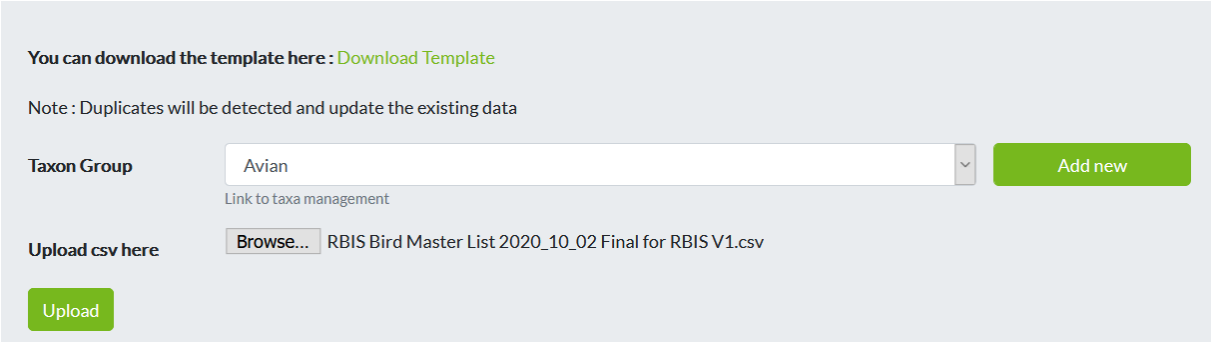


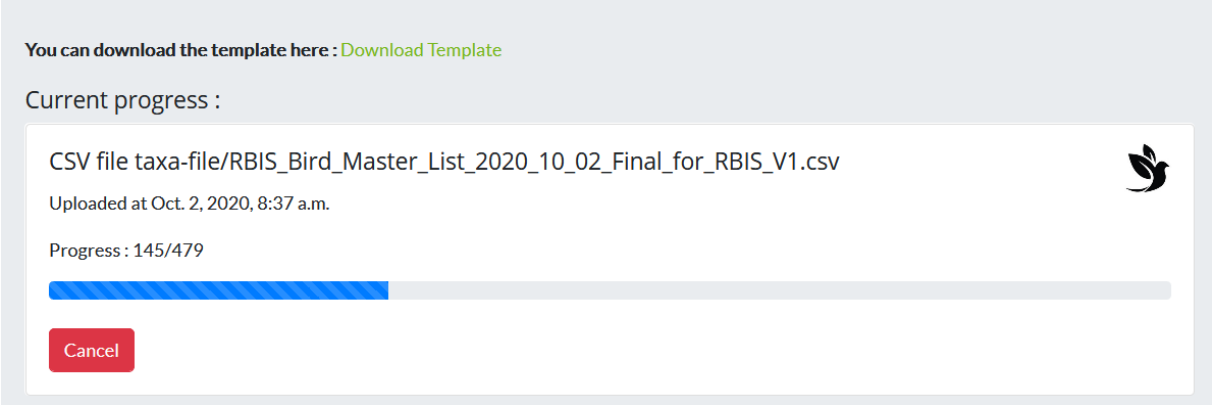
Select the new Taxon Group using the dropdown and browse to the file containing the Master List of Taxa for the Taxon Group, click upload



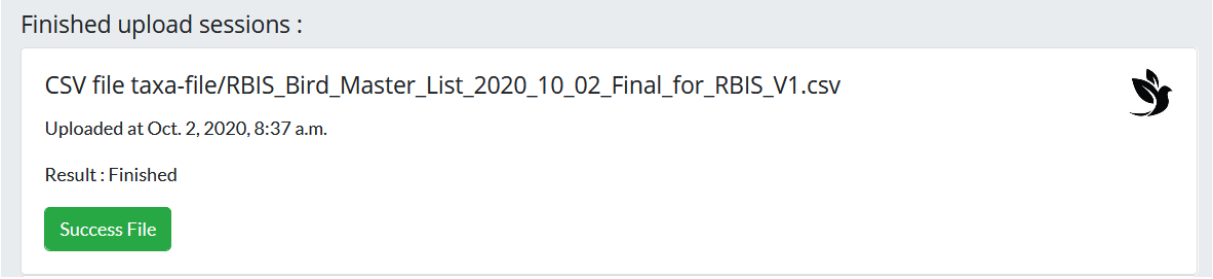
Note: The Excel file needs to be saved as a csv using the following option:







A **Success** file will indicate taxa uploaded to the system. An **Error** file will give details of taxa not uploaded. The last column in this csv file provides an indication of the reason the taxon was not uploaded. The administrator then needs to check the data and correct before uploading again.



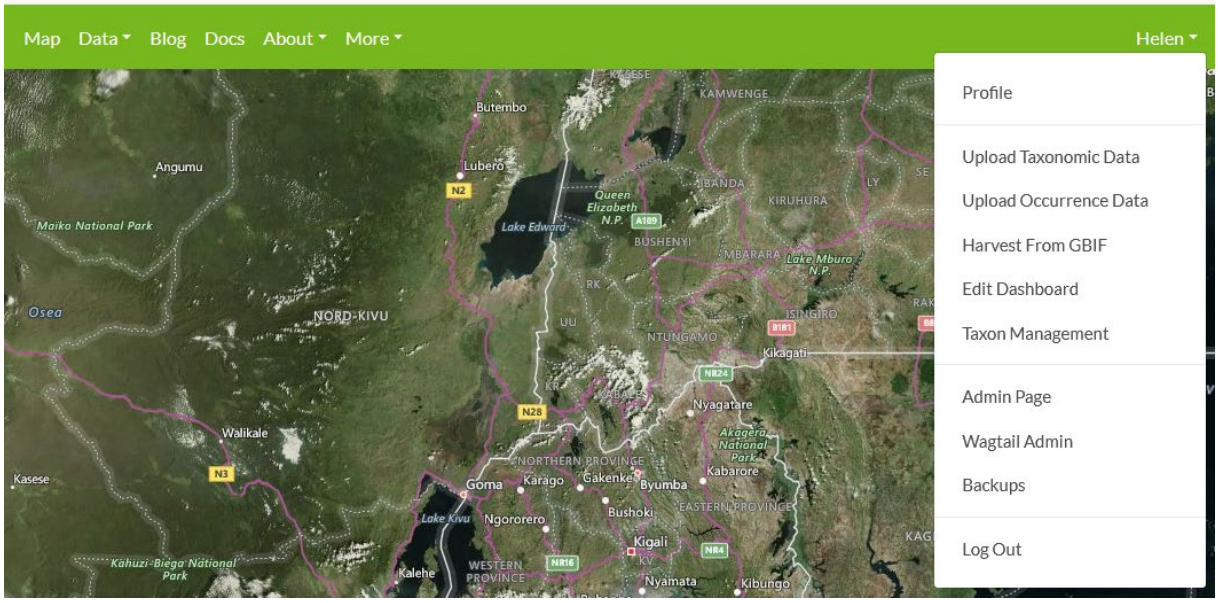
Once the Master List of Taxa has been added it can be viewed in the Taxon Management section.

New taxa may be added individually using the **Taxon Management** (see section 7 in this document), or if there are multiple new taxa to be added, then the steps from (4) can be repeated, to upload the additional new taxa for the Taxon Group.

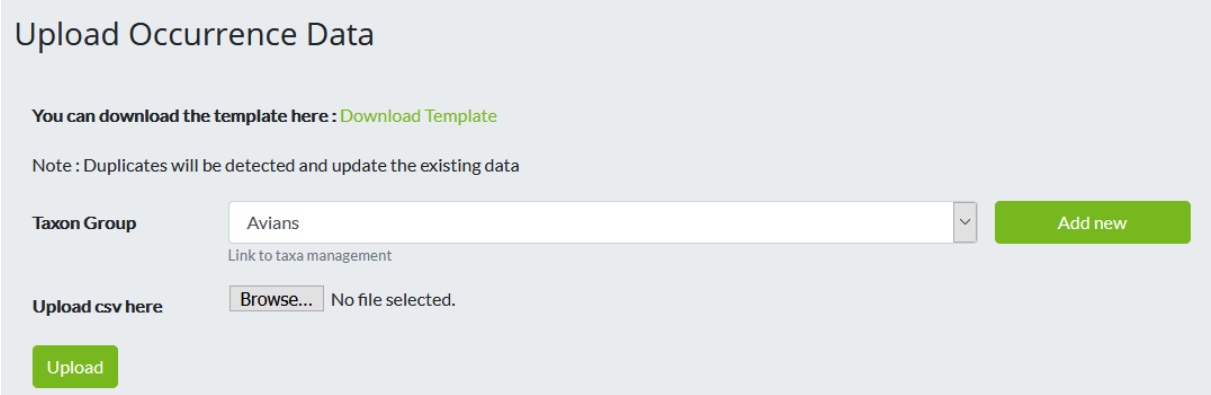
5 Uploading Occurrence Data

Steps:

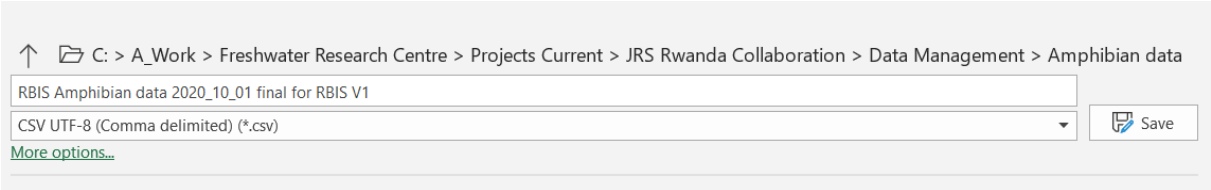
Click on your profile and select **Upload Occurrence Data**

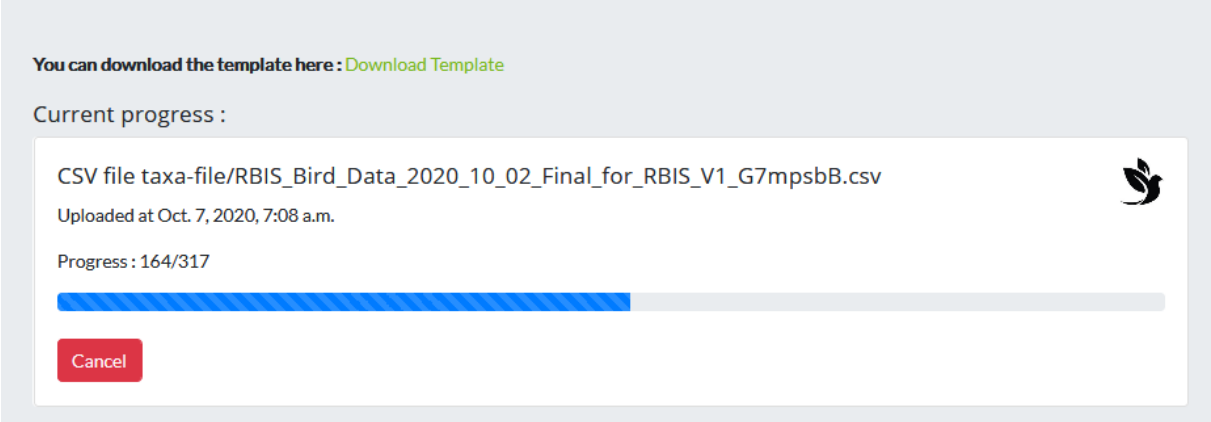


Select the Taxon Group using the dropdown and browse to the file containing the occurrence data for the taxon group, click upload



Note: The Excel file needs to be saved as a csv using the following option:





A **Success** file will indicate occurrence records uploaded to the system. An **Error** file will give details of occurrence records not uploaded. The last column in this csv file provides an indication of the reason the occurrence record was not uploaded. The administrator then needs to check the data and correct before uploading again.

Common errors:

- Taxon not in Master list
- Taxon Rank incorrect

Occurrence records may be updated by uploaded corrected records in the csv file. As long as the UUID is the same then the old occurrence record will be updated.

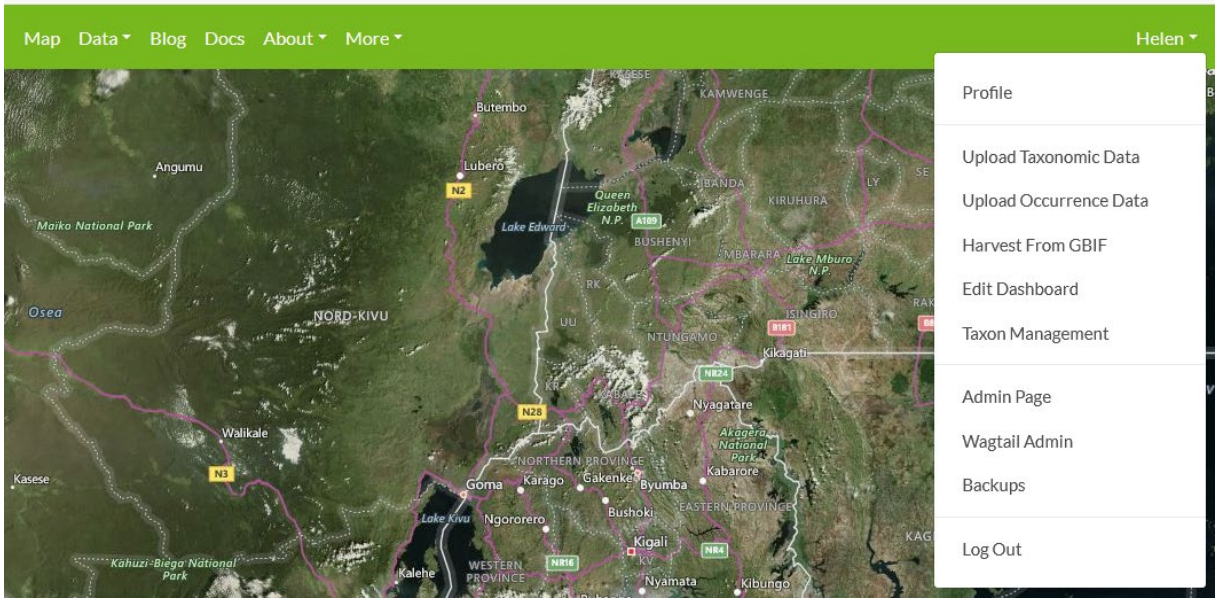
New occurrence records may be added by repeating the steps from (4), to upload the additional occurrence records for the Taxon Group.

Note that once the occurrence data has been uploaded, geocontext data (i.e. information about the site such as province, catchment etc.) are automatically populated for each site. This takes time and it is recommended that the next step (i.e. harvesting from GBIF), is done at least 24 hours after uploading occurrence data.

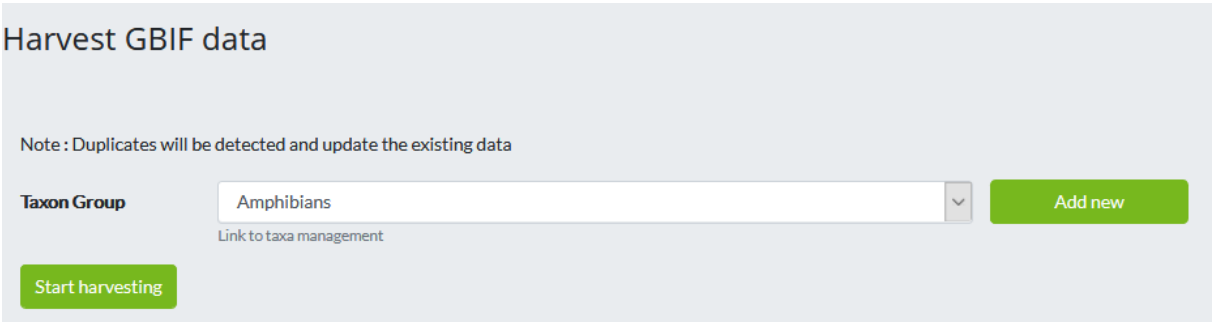
6 Harvesting GBIF Data

Steps:

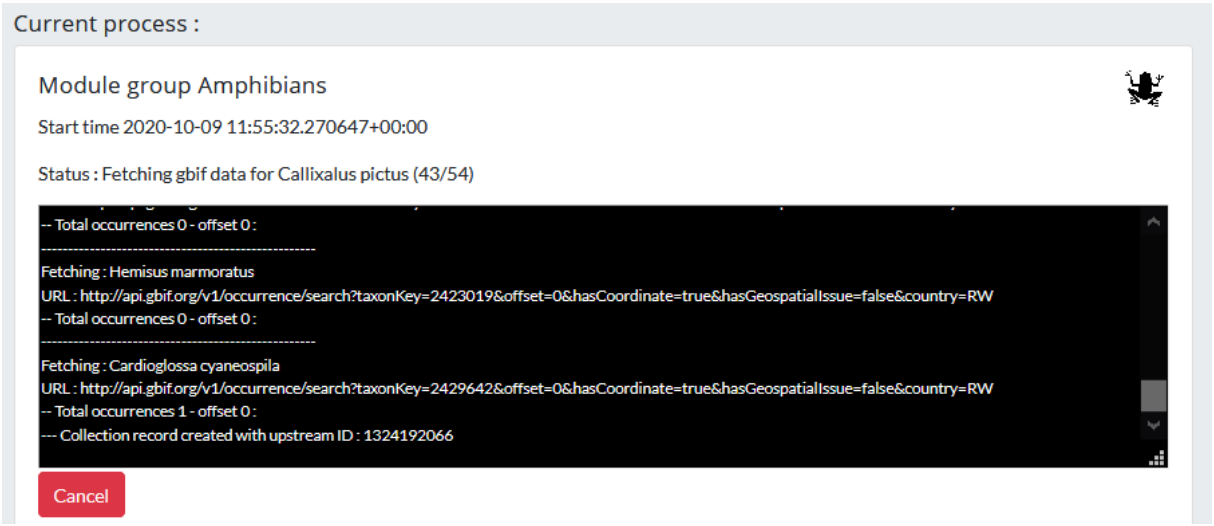
Click on your profile and select **Harvest from GBIF**



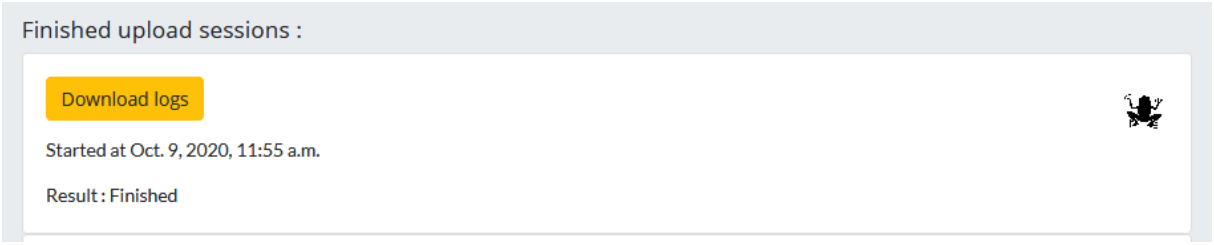
Select the Taxon Group using the dropdown and click Start harvesting.



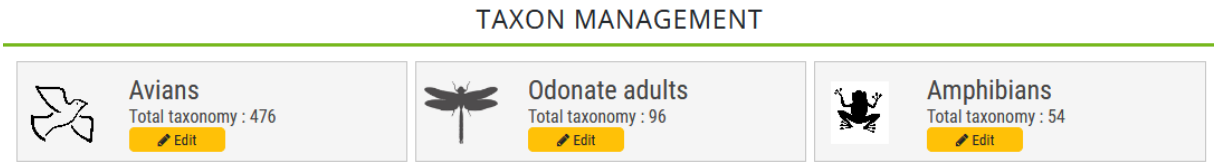
You can keep track of progress. The more taxa in the master list, the longer the time needed for harvesting data from GBIF. You can keep it running in the background and continue with other work as it harvests the data.



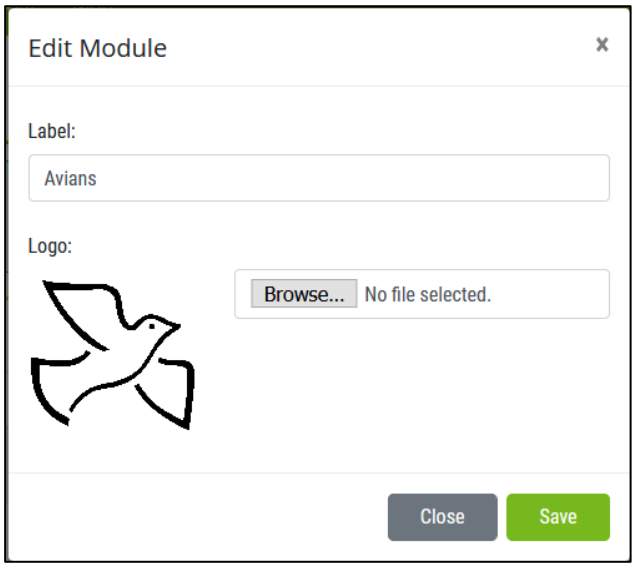
You can view the GBIF data harvested via the **Download Logs.**



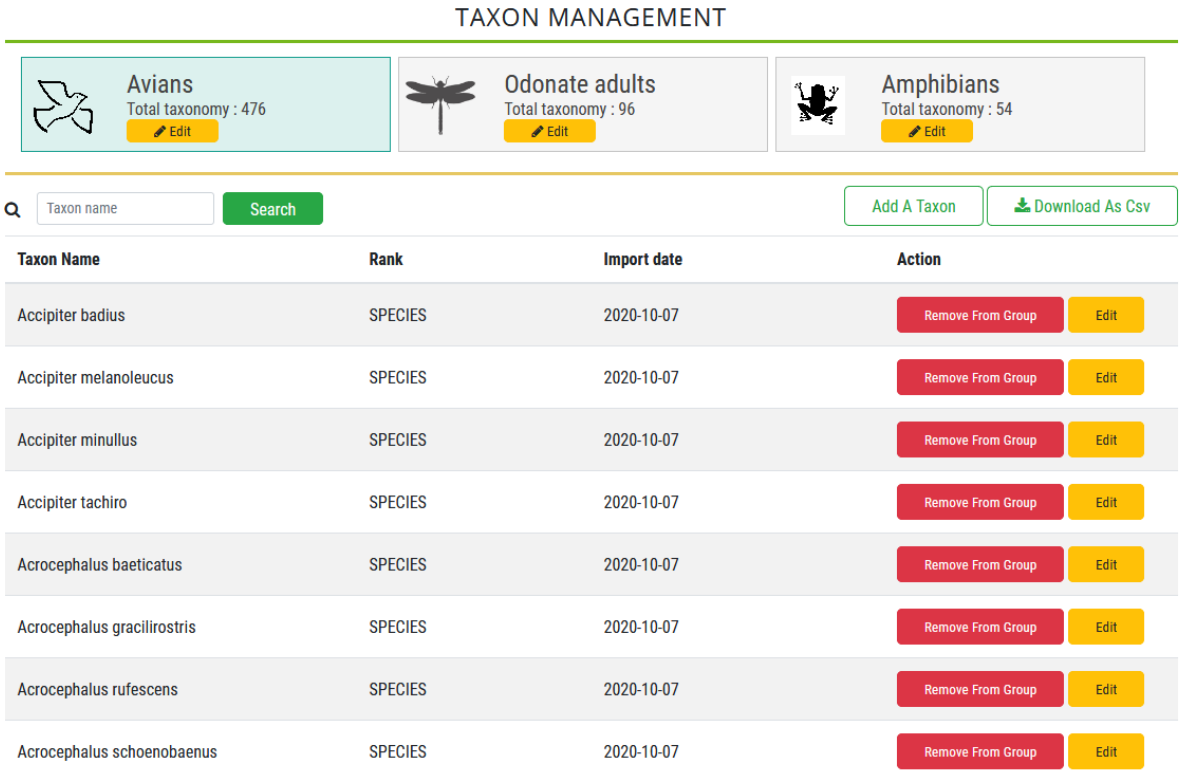
7 Managing taxa in Taxon Management



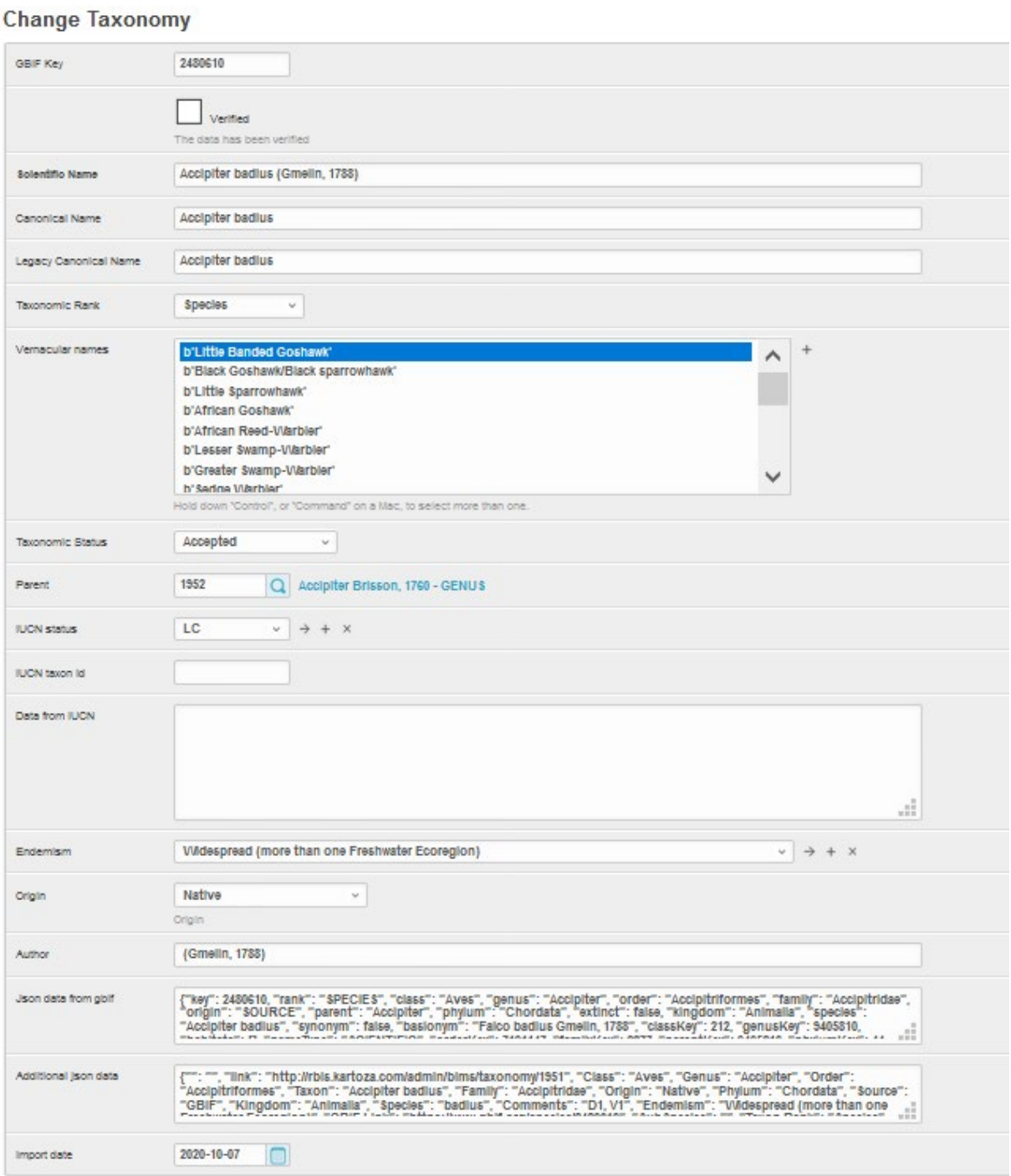
The label for the module and the logo may be updated, by clicking **Edit**:



Individual taxa may be edited using the Taxon Management section. The module is selected, which is then highlighted and all taxa in the module are served alphabetically. The rank and import date are shown for each taxon. The search button is used to search for specific taxa.



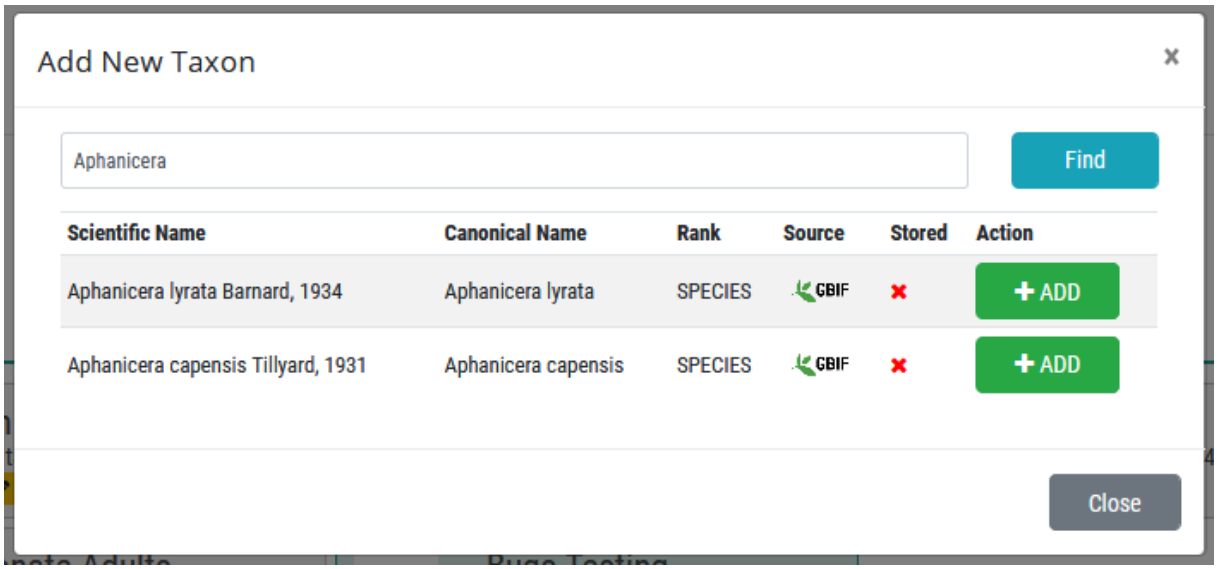
Details for each individual taxon are accessed using the Edit button. These include GBIF Key, Verification, Scientific Name, Canonical Name, Legacy Canonical Name, Taxonomic Rank, Vernacular names, Taxonomic Status, Parent (Taxon), IUCN status, Endemism, Origin and Author. Other details related to the data harvesting are shown as well as the import date.



You can also Download a Master List of Taxa for a Group using the **Download As CSV** button.

You can add a new taxon using the **Add A Taxon** button.

This functionality still needs to be finalised, but for now you type in the taxon name and it links to GBIF to provide options, which you then Add.



.

8 References

Dallas HF. 2020. Data management guide for developing the Rwanda Biodiversity Information System (RBIS) for housing and serving freshwater biodiversity data. Prepared for the Center of Excellence in Biodiversity and Natural Resource Management, University of Rwanda