

MOL Uploader Manual

Overview

This tool is designed to help you easily publish your biodiversity data to Map of Life. Here you can map your data, compare and integrate it with other datasets, and set sharing and download permissions.

We currently support two types of observational data for upload:

1. **Occurrence Data:** incidental species records with known coordinates and date.
2. **Inventory data:** records from a survey or sampling event that targeted a group of species over an explicit area.

The steps required to upload a dataset are:

1. **Basic Information**
 - **Select the taxonomic group** (e.g., birds, mammals, butterflies, etc.) from the list of groups currently supported by MOL. If your data include multiple taxonomic groups, please prepare and upload separate datasets for each taxonomic group.
 - **Select the dataset type.** Occurrence or Inventory.
 - **Select the spatial data format** (inventory datasets only). Choose the file format in which the spatial data are stored.
 - Shapefile: polygon(s) of study area(s) in a zipped file with .dbf, .prj, .shp, and .shx files.
 - CSV: reference coordinate(s) of study area(s) (to be buffered) stored as latitude and longitude. If this option is selected, you will be prompted to provide information about the shape and size of your study area.
2. **Upload Files**
3. The file upload requirements differ depending on dataset type:
 - Occurrence: a CSV file
 - Inventory with spatial data in a shapefile: a CSV file and zipped geometry files (.dbf, .prj, .shp, and .shx)
 - Inventory with spatial data in a CSV: a single CSV file or two CSV files
4. **Match Column Names**
5. Check that the column names in the upload files are correctly matched to the corresponding Map of Life columns. Map of Life's data standards closely conform to Darwin Core and Humboldt Core, depending on the dataset type.
6. **Review Geometry**
7. Explore the map preview to verify that the spatial information you provided in the uploaded files is correct. If the study areas are not plotting correctly, please go back to the upload files step and upload corrected spatial data. If you continue to have problems, please contact us at info@mol.org.
8. **Metadata**

9. Fill in the required and optional metadata fields. After uploading your data files, you will be prompted to fill out dataset-level information (e.g., Permissions, License, Dataset Title). There is a short list of required fields, and you can choose to provide more detailed metadata (e.g. Citation, Sampling Description). Providing additional metadata makes the dataset more visible and searchable in Map of Life. Additionally, should you make your dataset public, providing detailed metadata will assist other users in understanding the dataset.
10. **Review Dataset**
11. View your dataset in the dataset overview page. Check the table summaries, map visualizations, and metadata. At this time, we do not support editing the data. If you see major errors that need to be fixed, please delete the dataset and re-upload it.
12. **Harmonize Names**
13. The final step is to taxonomically harmonize your species list. From the dataset overview, select Harmonize to start the taxonomic name matching process. Choose from a list of recognized taxonomies and resolve unmatched names.
14. **Share your dataset!**
15. Now you can visualize your dataset in detailed species maps and invite other users to view it through the permissions feature.

Step 1 - Basic Information

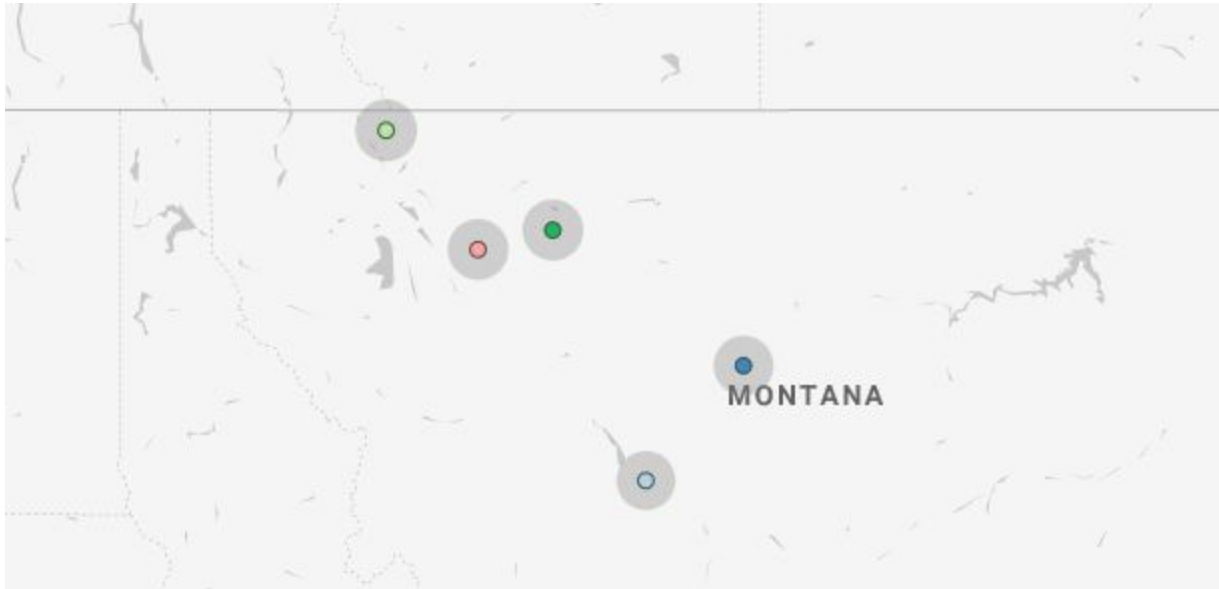
Picking a Taxonomic Group

The Map of Life platform supports a predefined list of taxonomic groups. Some of these groups are nested within others (e.g., Turtles are nested within Reptiles), while other groups are not yet complete in the database (e.g., Plants (selection)). As more taxonomies and range maps are added to the Map of Life database, these groups will become more refined. We require that uploaded datasets represent a single taxonomic group as this allows us to cross-reference the scientific names in a dataset against a published taxonomy for that group. This is done in the final taxonomic harmonization step before a dataset is finalized.

- Birds
- Mammals
- Amphibians
- Reptiles
- Turtles (non-marine)
- Butterflies
- Spingid Moths
- Bumblebees
- Dragonflies
- Fishes
- Plants
- Trees
- Cacti
- Conifers
- Palms

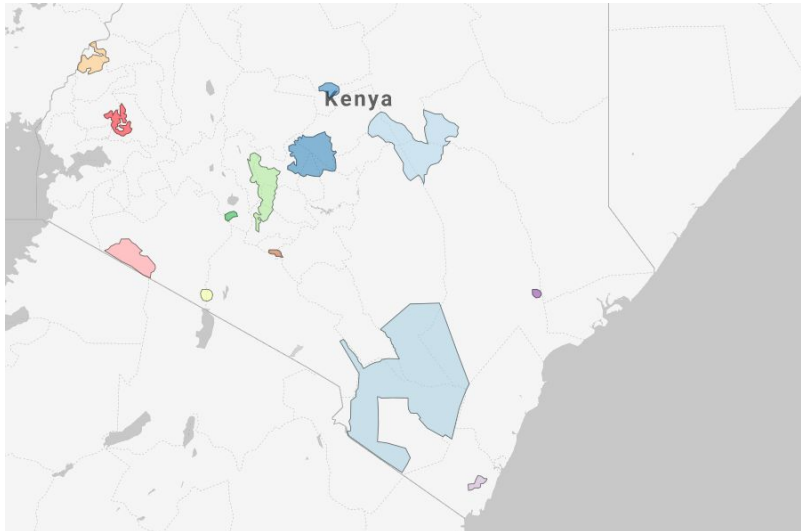
Picking a Dataset Type

Occurrences

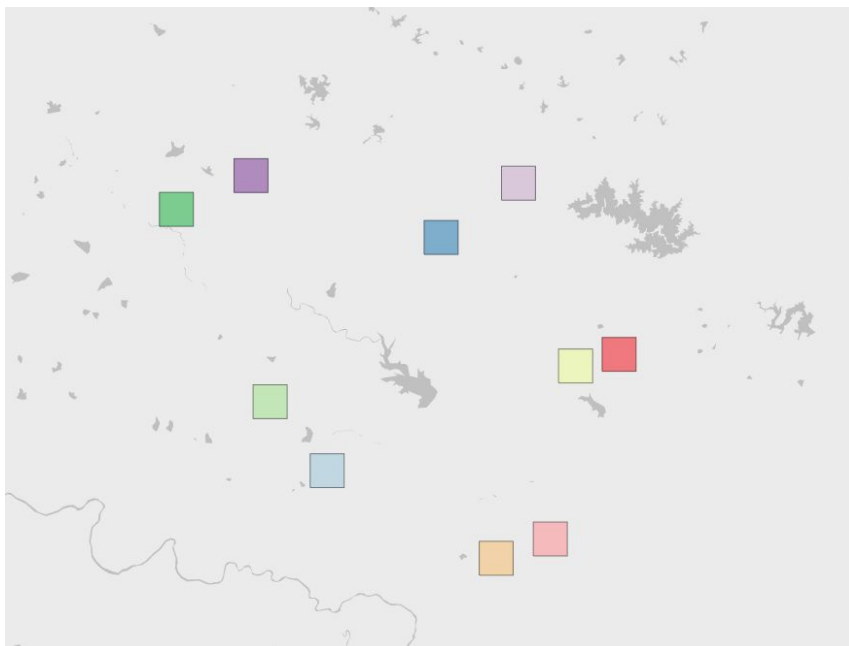


Species observations made at specific point locations (i.e., at a given latitude and longitude) with a date for each observation. An example of an occurrence dataset is a list of coordinates where grizzly bears were observed in Montana.

Inventories



Example of inventories at national parks in Kenya with the spatial data stored in shape files.



Example of inventories at uniform study areas with the spatial data stored as reference coordinates in a CSV file.

A list of species observed within an explicit area (defined by a polygon) and of a defined taxonomic scope (e.g., mammals, raptors, etc.). An example of an inventory dataset is a checklist of birds within a national park. Multiple file formats are supported:

- Shapefile: a zipped folder with .dbf, .prj, .shp, and .shx files
- CSV: reference coordinate(s) to be buffered to a circle or square to represent uniform study area(s) and stored as decimal latitude and decimal longitude. If this option is selected, you will be prompted to provide information about the shape (circle or square) and size (area, radius or edge length) of your study area.

Step 2 - Upload Files

Preparing Data for Upload

Data preparation steps vary depending on the dataset type. The following guidelines provide detailed information on how the data should be formatted for each type of dataset. We strongly recommend that you label the columns in your data tables according to the example provided. You will have a chance to check that the column names in the dataset are being matched to Map of Life fields, but the upload process will be easier if you format the files accordingly before upload. Additional fields can be included, preferably with the field names in lower camel case. Additional column names must not contain whitespace.

Occurrence Requirements

For an occurrence dataset, you only need a CSV file formatted to MOL specifications.

Table Format

Format column names and data to match the required fields below. You can have additional columns, but you must have at least the columns listed below. The file must be a table (spreadsheet) saved in CSV format.

1. `scientificName` - The scientific name of the species. This field must be a binomial (Genus species; e.g. 'Branta canadensis'). If you wish to record subspecies level information, it should be included as a separate column.
2. `decimalLatitude` - Latitude in decimal degrees.
3. `decimalLongitude` - Longitude in decimal degrees.
4. `eventDate` - Observation Date in ISO 8601 format (most commonly yyyy-mm-dd).

5. **basisOfRecord** - The specific nature of the data record. Possible values: "PreservedSpecimen", "FossilSpecimen", "LivingSpecimen", "HumanObservation", "MachineObservation".

Example

	A	B	C	D	E
1	scientificName	decimalLatitude	decimalLongitude	eventDate	basisOfRecord
2	Puma concolor	42.12	5.13	2014-05-13	PreservedSpecimen
3	Puma concolor	42.18	5.11	2014-05-16	FossilSpecimen
4	Puma concolor	42.19	5.13	2014-05-12	LivingSpecimen

Additional Columns

If additional information is available, add extra columns in the .CSV file to the right of the species names. We recommend including:

1. **individualCount**: number of individuals or abundance
2. **coordinateUncertaintyInMeters**: the horizontal distance (in meters) of uncertainty from the given decimalLatitude and decimalLongitude. Zero is not a valid value for this term.
3. **geodeticDatum**: geodetic datum or coordinate system used (e.g., NAD83, WGS84, ED50, GCJ-02)
4. **recordedBy**: name of the person that recorded the observation
5. **identifiedBy**: name of the person that identified the species

Troubleshooting

Date Format

To convert your dates to the correct format, you can use this function in Excel:

```
=TEXT(A1, "yyyy-mm-dd")
```

where A1 refers to the cell being converted.

The format string is sensitive to region/language settings. For example, if Excel's region is set to Germany, you'll need to use this function instead:

```
=TEXT(C2; "jjjjMMtt")
```

If your date information is stored in separate year, month, and day columns, you will first need to concatenate the information before applying the date conversion function. To do this, use:

```
=TEXT(CONCATENATE(year, "-", month, "-", day), "yyyy-mm-dd")
```

where year refers to the cell containing the year, month to cell with the month, and day to the cell with the day. Fill down to change the coordinates for all records. Select the column, copy it, and then in the Edit toolbar use “Paste Special...” and select Values to remove the formula.

Coordinates in a Different Format

If your coordinates are not already in decimal degrees, you will need to convert them. There are online wizards where you can convert one coordinate at a time, but if you have a lot of coordinates, you will want to convert them all in one batch.

If you have degrees, minutes, and seconds in separate columns, you can write a simple formula in Excel. In the first cell of the column where you want the decimal degrees to go, use this formula:

```
=ROUND((degrees+minutes/60)/60+seconds,6)
```

Where degrees refers to the cell with degrees, minutes to the cell with minutes, and seconds to the cell with seconds. If you don't have seconds, leave that part out. The “Round” function restricts the results to 6 decimal places. Fill down to change the coordinates for all records. Select the column, copy it, and then in the Edit toolbar use Paste Special... and select Values to remove the formula.

If you have degrees, minutes, and seconds all in one column, it is easiest to use Visual Basic to convert the values to decimal degrees. You can follow this tutorial for Excel [here](#).

If you have UTM coordinates, you can use the US National Geodetic Survey tool [here](#).

Subspecies Included in Binomial

If the species names include trinomials, we ask that you remove the subspecies to a separate column. You can do this in Excel by first using “Text to Columns” with a space as the delimiter, and then to get the binomial back in one cell use “Concatenate”. To use “Text to Columns”, select the column you want to split and then in the Data menu select “Text to Columns”. Choose “Delimited” and on the next page select “Space” or whatever symbol you used to separate words. On the next screen, choose the destination cells where you want the results to go and then “Finish”. You should then have three columns with the genus, specific epithet, and subspecies. To get the binomial in one cell, use the “Concatenate” function. In the first cell in the column where you want the binomial, use the following:

```
=CONCATENATE(genus, “ “, specific epithet)
```

Where genus refers to the cell containing the genus, the quotes add a space between the words, and the specific epithet refers to the cell containing the specific epithet. Fill down to create the binomial for all records. Select the column, copy it, and then in the Edit toolbar use Paste Special... and select Values to remove the formula.

Incomplete Dates

If you have historical data with incomplete dates we ask you to use the following guidelines:

- If you have month and year, but no day, use the first of the month.
- If you have year, but no month or day, use the first of January.

- If you have no information, leave blank.

Inventory Requirements

We recognize that people may have their inventory data stored in multiple formats. To make our upload process more user-friendly, we are currently offering three options. In each option, we expect the species information to be stored in a .CSV file while the spatial data can be stored as:

- Polygons in a shapefile
- Coordinates in the species table
- Coordinates in a separate table from the species data

We provide instructions for each type of upload.

Shapefile

To upload an inventory dataset with the spatial data stored in a shapefile, you will need to prepare a .CSV and a shapefile.

Species Table Format

Format columns column names and data to match the required fields below. You can have additional columns, but you must have at least the columns listed below. The file must be a table (spreadsheet) saved in .CSV format.

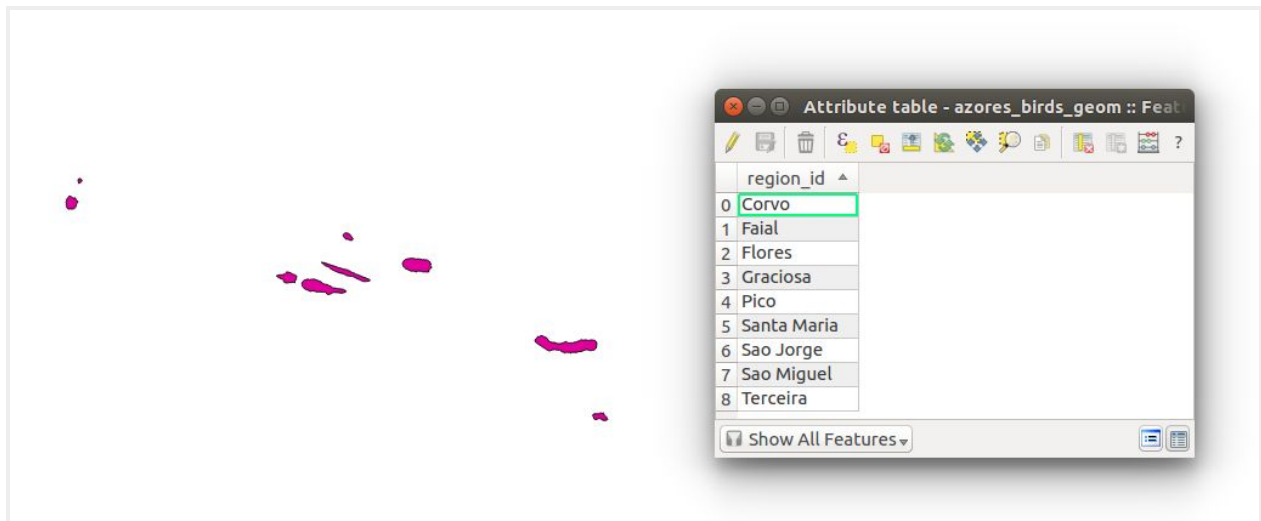
1. **scientificName**: the scientific name of the species. This field must be a binomial (Genus species; e.g. 'Branta canadensis'). If you wish to record subspecies level information, add it in a separate column.
2. **region_id**: this field will match the location of the species to the respective polygon(s) uploaded with this dataset. The **region_id** values in the .CSV and the attribute table of the shapefile should match.

	A	B
1	scientificName	region_id
2	Aix sponsa	Corvo
3	Anser albifrons	Pico
4	Anser anser	Corvo
5	Anser brachyrhynchus	Flores
6	Anser fabalis	Corvo
7	Aythya affinis	Flores
8	Aythya collaris	Corvo
9	Aythya ferina	Corvo

Shapefile

A shapefile is a geospatial file created by many GIS software products to describe geospatial shapes.

- We require a zipped folder of the shapefile containing the study area polygon(s). It must contain all of the necessary files (.dbf, .prj, .shp, and .shx).
- If there are multiple polygon features, each feature should have a unique `region_id` in the attribute table that matches the `region_id` for that observation in the CSV.
- If there is a single study area, the `region_id` can be set to '1' or 'locationA' in the shapefile and CSV.



An example of a shapefile and its attribute table in QGIS. Note how the `region_id` in the attribute table matches the `region_id` specified in the .CSV.

Coordinates in the Species Table

To upload an inventory dataset with the spatial data stored in the .CSV file with the species information, you will need to prepare a single .CSV file.

Format column names and data to match the required fields below. **You can have additional columns**, but you must have at least the columns listed below. The file must be a table (spreadsheet) saved in .CSV format.

1. `region_id`: a unique identifier for a survey location (i.e. plot name).
2. `scientificName`: the scientific name of the species. This field must be a binomial (Genus species; e.g. 'Branta canadensis'). If you wish to record subspecies level information, it should be included as a separate column.
3. `decimalLatitude`: latitude in decimal degrees.
4. `decimalLongitude`: longitude in decimal degrees.

	A	B	D	E
1	region_id	scientificName	decimalLatitude	decimalLongitude
2	1	Litocranius walleri	19.4076664712	76.2147859221
3	1	Otomys	19.4076664712	76.2147859221
4	1	Genetta genetta	19.4076664712	76.2147859221
5	2	Dendrohyrax arboreus	19.4998515292	76.1341603944
6	2	Loxodonta africana	19.4998515292	76.1341603944
7	2	Oryx beisa	19.4998515292	76.1341603944
8	3	Proteles cristata	19.7563447497	76.0018517986
9	3	Mormopterus	19.7563447497	76.0018517986
10	3	Equus grevyi	19.7563447497	76.0018517986
11	4	Neoromicia rendalli	19.3165345162	76.529741966
12	4	Equus asinus	19.3165345162	76.529741966
13	4	Loxodonta africana	19.3165345162	76.529741966
14	5	Colomys goslingi	19.5629225386	76.6269039862

Coordinates and Species Data in Separate Tables

Species Table Format

Format column names and data to match the required fields below. **You can have additional columns**, but you must have at least the columns listed below. The file must be a table (spreadsheet) saved in .CSV format.

1. **scientificName**: the scientific name of the species. This field must be a binomial (Genus species; e.g. 'Branta canadensis'). If you wish to record subspecies level information, it should be included as a separate column.
2. **region_id**: this field will match the location of the species to the respective polygon(s) uploaded with this dataset. The **region_id** values in the species table and the coordinates table should match.

	A	B	C
1	region_id	scientificName	extraField
2	1	Litocranius walleri	example value
3	1	Otomys	example value
4	1	Genetta genetta	example value
5	2	Dendrohyrax arboreus	example value
6	2	Loxodonta africana	example value
7	2	Oryx beisa	example value
8	3	Proteles cristata	example value
9	3	Mormopterus	example value
10	3	Equus grevyi	example value

Coordinates Table Format

Format column names and data to match the required fields below. The file must be a table(spreadsheet) saved in .CSV format.

1. **decimalLatitude**: latitude in decimal degrees.
2. **decimalLongitude**: longitude in decimal degrees.
3. **region_id**: this field will match the location of the species to the respective study area(s).
The **region_id** values in the species table and the coordinates table should match.

	A	B	C
1	region_id	decimalLatitude	decimalLongitude
2	1	19.4076664712	76.2147859221
3	2	19.4998515292	76.1341603944
4	3	19.7563447497	76.0018517986
5	4	19.3165345162	76.529741966
6	5	19.5629225386	76.6269039862
7	6	19.2906980791	76.4533876467
8	7	19.7911802629	76.4847860654
9	8	19.8013185892	76.107114122
10	9	19.5474074649	76.5654409914

Additional Columns

If you have additional information you'd like to include, simply add extra columns in the .CSV file to the right of the species names.

	A	B	C
1	scientificName	region_id	notes
2	Aix sponsa	Corvo	Occasional Migrant
3	Anas acuta	Corvo	Regular Wintering
4	Anas americana	Corvo	Regular Wintering
5	Anas carolinensis	Corvo	Regular Wintering
6	Anas clypeata	Corvo	Regular Wintering
7	Anas crecca	Corvo	Regular Wintering
8	Anas discors	Corvo	Regular Wintering
9	Anas penelope	Corvo	Regular Wintering

Troubleshooting

Date Format

To convert your dates to the correct format, you can use this function in Excel:

```
=TEXT(A1, "yyyy-mm-dd")
```

where A1 refers to the cell being converted.

The format string is **sensitive to region/language settings**. For example, if Excel's region is set to Germany, you'll need to use this function instead:

```
=TEXT(C2;"jjjjMMtt")
```

If your date information is stored in separate year, month, and day columns, you will first need to concatenate the information before applying the date conversion function. To do this, use:

```
=TEXT(CONCATENATE(year, "-", month, "-", day), "yyyy-mm-dd")
```

where year refers to the cell containing the year, month to cell with the month, and day to the cell with the day. Fill down to change the coordinates for all records. Select the column, copy it, and then in the Edit toolbar use "Paste Special..." and select Values to remove the formula.

Coordinates in a Different Format

If your coordinates are not already in decimal degrees, you will need to convert them. There are online wizards where you can convert one coordinate at a time, but if you have a lot of coordinates, you will want to convert them all in one batch.

If you have degrees, minutes, and seconds in separate columns, you can write a simple formula in Excel. In the first cell of the column where you want the decimal degrees to go, use this formula:

```
=ROUND((degrees+minutes/60)/60+seconds,6)
```

Where degrees refers to the cell with degrees, minutes to the cell with minutes, and seconds to the cell with seconds. If you don't have seconds, leave that part out. The "Round" function restricts the results to 6 decimal places. Fill down to change the coordinates for all records. Select the column, copy it, and then in the Edit toolbar use Paste Special... and select Values to remove the formula.

If you have degrees, minutes, and seconds all in one column, it is easiest to use Visual Basic to convert the values to decimal degrees. You can follow this tutorial for Excel [here](#).

If you have UTM coordinates, you can use the US National Geodetic Survey tool [here](#).

Subspecies Included in Binomial

If the species names include trinomials, we ask that you remove the subspecies to a separate column. You can do this in Excel by first using "Text to Columns" with a space as the delimiter, and then to get the binomial back in one cell use "Concatenate". To use "Text to Columns", select the column you want to split and then in the Data menu select "Text to Columns". Choose "Delimited" and on the next page select "Space" or whatever symbol you used to separate words. On the next screen, choose the destination cells where you want the results to go and then "Finish". You should then have three columns with the genus, specific epithet, and subspecies. To get the binomial in one cell, use the "Concatenate" function. In the first cell in the column where you want the binomial, use the following:

```
=CONCATENATE(genus, " ", specific epithet)
```

Where genus refers to the cell containing the genus, the quotes add a space between the words, and the specific epithet refers to the cell containing the specific epithet. Fill down to create the binomial

for all records. Select the column, copy it, and then in the Edit toolbar use Paste Special... and select Values to remove the formula.

Incomplete Dates

If you have historical data with incomplete dates we ask you to use the following guidelines:

- If you have month and year, but no day, use the first of the month.
- If you have year, but no month or day, use the first of January.
- If you have no information, leave blank.

Step 3 - Match Columns

We adhere to multiple metadata standards depending on the type of data involved, including Darwin Core, Humboldt Core, and Dublin Core. Following these standards helps us simplify the process of uploading, publishing, and sharing biodiversity datasets across the Map of Life platform.

In the upload form we've provided a tool for quickly mapping your column names to these metadata terms. If you have columns that match these values, but are named something else, you can match them using our tool. Refer to the instructions for descriptions and file formatting specifications.

For inventory datasets, the following terms are required:

- scientificName
- region_id
- decimalLatitude (if no shapefile)
- decimalLongitude (if no shapefile)

For occurrence datasets, the following terms are required:

- scientificName
- eventDate
- basisOfRecord
- decimalLatitude
- decimalLongitude

Please be assured that if you upload data that does not match our existing metadata terms, they will still be uploaded to the database and stored with the rest of your data.

Step 4 - Review Geometry

Explore the map preview to verify that the spatial information you provided in the uploaded files is correct. If the study areas are not plotting correctly, please go back to the upload files step and upload corrected spatial data. If you continue to have problems, please contact us at info@mol.org.

Step 5 - Metadata

Permissions

There are three types of permissions that can be assigned to a dataset. You can select a combination or no options.

- **Allow public visualization.** Allow the dataset to be displayed on Map of Life and platforms supported by Map of Life.
- **Allow MOL modeling use.** Allow Map of Life and its associated partners (e.g., [Global Mountain Biodiversity Assessment](#)) to use the dataset in spatial analyses and modeling.
- **Allow public download.** Allow other registered Map of Life users to download the dataset.

To make a dataset 'Public', select 'Allow public visualization'. If you do not select 'Allow public visualization, the dataset is 'Private'.

License

You must specify a data license for when the data are made public. Three Creative Commons licensing options are currently offered:

- **CC0:** This license waives copyright of the dataset and, as much as is allowable by law, places the dataset in the public domain.
- **CC-BY:** This license means that users may copy, distribute, and make derivative works only if they give credit to the author or licensor (attribution).
- **CC-BY-NC:** This license indicates that users may copy, distribute, and make derivative works (see CC-BY above), while attribute the author and for non-commercial use only.

Release Policy

- **Set full release date.** If you are not ready to immediately allow public visualization, modeling, or download of your dataset, we ask that you provide a release date. On the date you select, all permissions (visualization, modeling, and download) will be allowed. The full release will occur at 00:00 GMT on the day you select.
- **Don't set release date.** If your work is still *in prep* and you cannot commit to a release date, you can select this option and we will send you an email reminder every 6 months.