

Instructions for Documenting Study Databases (v7)

A Database consists of one or more data tables, GIS layer(s), remote sensing image(s) and metadata (documentation). Structured and comprehensive metadata are necessary to support database archival and retrieval, the functional re-use of the data both by owners and secondary users, and data quality control. The Forest Science Data Bank (FSDB) requires complete documentation for long-term maintenance and distribution of study databases. Metadata elements requested in this guide are generally consistent with national standards for metadata, specifically the Ecological Metadata Language (EML) <http://knb.ecoinformatics.org/software/eml/> and the Federal Geographic Data Committee (FGDC) <https://www.fgdc.gov/metadata>

Primary steps for preparing and submitting study metadata and data

1. Contact one of the Andrews Forest Information Managers to establish a database code for your study data.
 - a. Suzanne Remillard (suzanne.remillard@oregonstate.edu)
 - b. Don Henshaw (don.henshaw@oregonstate.edu)
2. Provide Study Description metadata (Database Level Information). This metadata documents general information about a study (title, abstract, purpose, dates, methods, site characteristics, etc.)
3. Provide Database Table or GIS information
 - a. Provide Database Table metadata (Entity and Attribute Level Information). This metadata defines the individual data tables (entities) and describes the specific variable fields within each table (attributes).

Or,

 - b. Provide Geographic Information System (GIS) metadata (Spatial Reference Information). This metadata documents information related specifically to ESRI GIS data layers.)

Data Access Policy: It is the intention of the FSDB to display database metadata on a publicly available web site. Sensitive information can be withheld from public view on request. Public online access to the actual data will only be provided with permission from the data provider or study principal investigator. Please refer to the Data Access Policy. (<http://andrewsforest.oregonstate.edu/data/policy>)

Step 1. Contact one of the Andrews Forest Information Managers:

It is important to talk to the data manager before you begin developing the metadata and data files. A unique database code will be assigned to your study. The database manager is also available to review or answer questions regarding the preparation of database level or entity and attribute level information. The data manager can also review or recommend strategies for organizing your data.

Step 2. Provide Study Description metadata (Database Level Information):

The LTER Administrative Interface provides web forms that are established consistent with Database Level Information contained in this guide. It is recommended that required information be assembled in a Word or text file and then use “copy and paste” to place into the web forms (this alleviates timing out issues with the application). This assembled file may also be useful for the information manager in reviewing your metadata. Please answer information requests as completely as possible. Where an element does not apply, responses such as “Unknown” or “Not Applicable” are acceptable.

LTER Administrative interface: <http://andlter.forestry.oregonstate.edu/webadmin>

Web form instructions can be found at:

<http://andrewsforest.oregonstate.edu/data/metadata>

Step 3a. Provide Database Table metadata (Entity and Attribute Level Information):

An Excel Workbook (a sample template) is established to describe your data tables (entities) and your data fields (attributes). The workbook includes a separate worksheet to describe the list of entity names for your study database. For each entity another worksheet is needed to define and characterize table attributes. Some attributes require an enumerated domain, that is, are coded to a specific set of valid entries, and these valid entries must be listed (enum). Additionally, specific site locations must also be listed (place).

Link to Excel Workbook:

<http://andrewsforest.oregonstate.edu/data/metadata#data>

Data can be provided in this Excel Workbook as well. Data can also be accepted in other structured formats, e.g., Access, FoxPro, etc. Data in the database tables/worksheets may need to be normalized to take advantage of advanced database features. The goal of data normalization is to reduce or even eliminate data redundancy. The advantage of having a highly normalized data is that information is stored in one and only one place, reducing the possibility of inconsistent data. For more information about normalized data, search on the web or go to http://en.wikipedia.org/wiki/Database_normalization.

Or,

Step 3b. Provide GIS metadata (Spatial Reference Information):

ESRI ArcCatalog is used to document ArcGIS data sets. Much of this information can be electronically obtained directly from the layers or images themselves.

Please consult with the FSDB GIS specialist when submitting this type of information. Steps 1 and 2 are still necessary to allow spatial data integration into the online database catalog.

Link to Spatial Data Documentation Guidelines PDF:

<http://andrewsforest.oregonstate.edu/data/metadata#data>

Detailed Descriptions for Study Metadata Using the Andrews LTER Administrative Interface (Step 2)

The Andrews LTER Administrative Interface includes a series of web forms where study metadata can be described. The Administrative Interface is password protected and enables users to 'Update Study' metadata for studies to which they are associated. The interface will show associated studies in a list with the ability to 'Update' each study.

The interface is a tab oriented interface. A series of tabs across the top of the page access various types of information related to necessary metadata for the study. The tabs in order across the top (and described below) are: General, Personnel, Theme Keys, Place keys, Taxon Keys, Access, Cross-Ref, Methods, Related, and Return... Each tab must be saved independently. You will be prompted to save, if you switch to another tab without saving changes.

Please use the Word Document Form available online to fill out your metadata and then use "copy and paste" into the web forms (this alleviates timing out issues with the application).

Study Description Metadata

1. General

The top of the General tab shows the FSDB database code with the study title and includes the data of last edit.

1.1. Identification

1.1.1. Study Title

The title should be descriptive and should describe the data collected, geographic context, research site, and time frame (what, where, and when).

1.1.2. Associated Project(s)

You can associate your study to one or more projects. HJ Andrews Experimental Forest is the default. You may also choose to include LTER, if appropriate. If your study is part of a larger project, then select that project (for example, WS01 Airshed, WS10 Hillslope Hydrology, DIRT, or Microbial Observatory).

1.2. Description

1.2.1. Abstract

The abstract will be useful for full-text searches, and it should be rich with descriptive text. The measured parameters should also be included. Extensive description should include what, when, and where information as well as whether the dataset is ongoing or completed, some taxonomic information, and some methods description (what, where, when, and why plus parameters). If there are too many parameters for a dataset, use categories of parameters instead of listing all parameters (ex. – use

nutrients instead of nitrate, phosphate, calcium, etc.) in combination with the parameters that seem most relevant for searches.

1.2.2. Purpose

A summary of the intentions with which the study was developed. Describe the “why” aspects of the data. For example, Purpose: Provide a baseline for characterizing distribution and variation in snow depth, moisture, and duration in the Andrews Forest for climatological and hydrological modelling and examine differences in snow accumulation in clearcut and adjacent closed forest stand microclimates.

1.2.3. Geographic Extent

Provide a short description of the geographic area of the study should include the extent of all sites and not individual sites. Examples include, “H. J. Andrews Experimental Forest, Western Cascades, Oregon”, “Western Cascades in Oregon and Washington”, or “Deschutes River Basin”.

1.2.4. Site Description

Provide a comprehensive description of the multiple sub-regions where site(s) have been samples including location (country, county or province, city, state), general topography, landmarks, rivers, and other relevant information. This could be excerpted from publications where study sites are described.

1.2.5. Supplemental Information

Other descriptive information about the study such as historical data sources or alternative time scale information. Geologic time scales, such as radiocarbon dating, or tree ring chronology dating back hundreds of years, should be denoted as “years before present”.

1.3. Time Period

The period of time (range of dates) during which the data were collected, published, observed, or dates of image creation. Form brings up a calendar box for selection or edit date directly. If don't know exact day, estimate by selecting the middle of the month. If data collection is not complete, make an estimate.

1.3.1. Beginning Date

Beginning date of data collection

1.3.2. Ending Date

Ending date of data collection

1.3.3. Currentness

The basis for which the dates of data collection is determined. Choose one of the following:

- (1) Ground condition - the range of dates during which the site was visited and data collected.
- (2) Publication - the range of dates during which the data is prepared for publication (this basis is used when generating data from a variety of sources, such as soil classifications, or if specific collection dates are unknown).
- (3) Observed - the range of dates during which the data is observed (this basis is used when there is no physical presence at the collection site, and the information may be collected by instrument or other tool; for example, the range of dates

during which satellite imagery, aerial photographs, or other images forming the base study data are created).

1.4. Collection and Update

1.4.1. Progress

The state of study data collection. Choose one of the following:

- (1) Complete – Data collection is completed and no new collection is planned
- (2) Ongoing – Data collection continues or further data collection is planned
- (3) Planned – Data collection is planned or underway, but no data is ready for archival

1.4.2. Measurement Frequency

The periodicity of data collection. Indicate the intervals of data collection (examples: continuous, 15 minutes, hourly, daily, weekly, monthly, annually, biennially, 5 years, irregular). Be more specific if necessary. This field is limited to 50 characters.

1.4.3. Update Frequency

The frequency with which changes and additions are made to the study database after the initial database is established. Choose one of the following:

- (1) Continually
- (2) Daily
- (3) Weekly
- (4) Monthly
- (5) Annually
- (6) Unknown
- (7) As needed
- (8) Irregular
- (9) None planned

1.4.4. Completeness Report

Describe the general completeness of available data affecting the overall quality of the data set, tasks that are yet to be accomplished, and other insights describing problems or informing potential users on best use of data set.

1.4.5. Related Material

List any samples, specimens, maps, or other related materials and their location, or list citations for other reference material that are not in our bibliography. See "Cross Reference" to link this study to other publications or databases that are in FSDB.

1.5. Taxonomic

1.5.1. General Taxonomic Information

Describe the range of taxa addressed in the study.

1.5.2. Taxonomic System

Reference or field guide used to identify organisms in this study. Multiple references are allowed. If you use one not in the list, you may submit your reference to the Databases Managers.

2. Personnel

Contains information about the people involved in the conception, design, implementation, and documentation of a study and data tables. List each person involved in the study and their roles. Select a person, then select a role, then click on 'Add to Study'. The person will show in the list of 'Current Study Members'. A person may be assigned to multiple roles (i.e., Originator and PI). Persons may also be removed from the current study members list by selecting the person and clicking on 'Remove From Study'. If the person is not in the list, click 'Add Person'.

These are three required researcher roles:

Originator –The person who originated or initiated the study. This person may or may not be currently involved.

Principal Investigator (PI) – The person responsible for conducting the scientific study or reviewing the study data. This person may or may not be the originator of the study.

Data Set Contact Person – The contact person regarding access or use of the data.

The following roles should be assigned as appropriate:

Other Researchers – Major contributors to the study who are currently involved and should be credited.

Former Investigators – Major contributors to the study who are no longer involved but should be credited.

Metadata Contact – The person who prepares or maintains the study metadata and is the contact for questions about the data description (metadata.) Not needed if Data Set Contact is also Metadata Contact.

Methods Contact - The contact person regarding protocols, methods, models, tools, or instrumentation used for the study. Not needed if Data Set Contact is also Methods Contact.

Abstractor - The person who prepared the original data abstract (metadata)

3. Theme Key (Theme Keywords)

Provide theme keywords for searching based on the subject matter of the database. Please choose at least 3 theme keywords from the Andrews LTER Preferred Keyword List (PDF file available from this tab and is much easier to search). New keywords (not listed) can be suggested to the data manager. The study keywords show in the box on the right. Keywords can be added or removed from the study keyword list. Once a keyword is selected the 'Add' or 'Remove' button at the bottom of the box will become active.

4. Place Key (Place Keywords)

Provide place-based keywords for searching based on location of the study. Please indicate general study area name, as well as more specific site location names. New place keywords from your study will be added to the global list on left once your attribute documentation is submitted to the Data Managers (see 'Place' tab in Excel template for Entity and Attribute Information).

5. Taxon Key (Taxonomic Keywords)

Provide common-use words or phrases that describe the taxonomy covered by the study, e.g., anadromous fish, arthropods, birds, conifers, lichens, mammals, mosses, shrubs, trees. The list is very general, but please be as specific as possible.

6. Access (Database Access)

6.1. Access Constraint:

Access to the data will either be public (online) or restricted. Choose one of the following (copy and paste above):

- (1) Online
- (2) Restricted due to proprietary and/or publication issues
- (3) Restricted due to legal (i.e., location of sensitive resources) or copyright issues
- (4) Restricted due to QA/QC issues.

Our goal is to make most data available via the Internet in 2-3 years after collection, to conform to the LTER Network standards, CC-0, (<https://creativecommons.org/share-your-work/public-domain/cc0>).

6.2. Available Time Period

Specify the projected date for releasing online if access constraint is not 'online' above; be as specific as possible (month, year). These data will not be released without authorized permission.

6.3. Use Constraint

Describe restrictions and legal prerequisites for using the data after access is granted. These include any access constraints applied to assure the protection of privacy or intellectual property, and any special restrictions or limitations on using the data.

6.4. Database Credit

Acknowledge any persons, institutions, funding agencies, or grant numbers that contributed or supported collection or development of the database..

7. Cross-Ref (Cross References; Related Databases and Publications)

Existing Andrews online databases or publications can be cross-referenced. Select databases by the database_code or publications by the pub_number for any cross reference.

Assign whether this cross-reference is a direct reference or related secondarily.

1 - Directly references the named catalog object, e.g., a publication that describes or is primarily based on a cross-referenced data set, OR a database collected with similar methodology or within the same study design.

2 - Related indirectly or secondarily to the named catalog object, e.g., a publication secondarily refers to or cites a cross-referenced data set, OR a database very similar in focus.

8. Methods (Methodology)

*Describe only those methods used to create the data. Do not include methodology for analyses you perform for papers or manuscripts unless you are including the analysis or model results within this database. Include instrumentation and appropriate citations where necessary. **All datasets should include, at least, Experimental design and either Lab and/or Field Methods.***

The following methods can be described:

- 1- *Experimental design: Describe the experimental design and provide appropriate citations (if any)*
- 2- *Field Methods: Describe the field methods and provide appropriate citations, such as published procedures or protocols. Provide a full description of the instruments used, including manufacturer, model, calibration dates and accuracy. Changes in instrumentation and dates of changes should be mentioned under field methods.*
- 3- *Lab Methods: Describe the lab methods and provide appropriate citations, such as published procedures or protocols. Provide a full description of the instruments used, including manufacturer, model, calibration dates and accuracy. Changes in instrumentation and dates of changes should be mentioned under lab methods.*
- 4- *Statistics: Describe statistical methods used to produce the data and provide appropriate citations (if any)*
- 5- *Models or Algorithms: Describe models or other algorithms used to produce the data and provide appropriate citations (if any)*
- 6- *Permanent plots: Describe the use or establishment of any permanent plots associated with this study, or the proximity to existing permanent plots, and provide appropriate citations (if any)*
- 7- *Processing: Describe any pre-/post-processing steps used to prepare the data and provide appropriate citations (if any)*
- 8- *Quality Assurance: Describe quality assurance procedures employed at the study or attribute level and provide appropriate citations (if any)*
- 9- *Data Entry Procedures: Describe any methods regarding the digital capture or key entry of data and provide appropriate citations (if any)*

9. Related (Related Files)

Supporting files can be provided (e.g., pdf, spreadsheet, Word, Text, etc.) or other web pages which further enhance the database description. These files or web pages will be linked from the primary study database web page (below the entities). The file cannot be uploaded at this time and must be physically sent to the Data Managers.

10. Return... (Return to List)

Returns back to the list of study databases with which you are associated.

Detailed Entity and Attribute Level Information (Step 3)

An Entity is a table/spreadsheet/file or layer. Studies/databases can have one or several Entities. Attributes are the fields or columns in each table/spreadsheet/file. You will want to generate a summary of your attributes for each table/spreadsheet/file. This summary will have the attribute names, type of field (number, text), width (or length), and a listing of any allowable codes. You should be prepared to define any codes you use. If the details of the fields and values are already captured in a data dictionary or other report, please provide a digital copy to the information managers. You may provide a citation if one is available.

You will want to meet with an Andrews Data Manager before you begin the process of documenting your entity and attribute level information.

The Data Manager will help you determine if your study data needs to be restructured to be more efficient and to work with the Andrews QA/QC programs. Data designed to be used with various analysis packages may need to be restructured so that the fields are unique. Also, a given field in a given record can only contain one value. This prohibits any sort of repeating group within a single record. It may be necessary to create additional tables so that groups do not repeat.

The Data Manager will get you started in filling out the Excel spreadsheet that is available on the webpage (<http://www.fsl.orst.edu/lter/data/metadata.cfm#data>). The following are descriptions to help you fill out the spreadsheet.

For each table/spreadsheet/file or layer (entity):

These are the fields in the spreadsheet tab called 'entity'. This tab will describe each table/spreadsheet/file that is a part of the database. The letter before the heading is the column that it is located in the spreadsheet.

(A) Database Code:

*Each data table will have a unique database code assigned by the database manager. **Leave blank until assigned.***

(B) Number:

Each data table will have a unique entity number. Assign consecutive numbers to data tables starting with entity '1'.

(C) Title:

The title of the database table/spreadsheet/file. The title should be descriptive and should describe the specific data collected, research site, and time frame (what, where, and when), but does not need to reiterate the dataset title.

(D) Beginning Date:

Earliest or only date for which the attribute values are current. In cases when a range of dates is provided, this is the earliest date for which the information is valid.

- (E) Ending Date:
Latest date for which the information are current. Used in cases when a range of dates are provided.
- (F) File Name:
Name of data table file submitted to FSDB.
- (G) File Format:
Type of file submitted, i.e. spreadsheet, dbase file, ASCII file, ARC/INFO Export format
- (H) Overview:
A longer more descriptive title of what is in the table.

For each attribute (column, field) or item in each table:

These are the fields in the spreadsheet tab called "table" and describe each attribute (column/field) or item in each entity. The tab name (table_x) will match the entity that is being described, where x is the entity number from the entity tab. The letter before the heading is the column that it is located in the spreadsheet. **At a minimum, each attribute should include name of attribute, definition, domain type and unit of measurement, if a range domain.**

- (A) Name of the Attribute:
The name of the field (attribute). Should be somewhat descriptive. Best to limit to 10 characters and should not include spaces or special characters.
- (B) Definition of the Field:
The description of the attribute. Be clear and unambiguous. If appropriate, please list the authority for the attribute definition.
- (C) Primary Key (Y/N): (yes or no)
The primary key defines what fields the database are sorted on (i.e., date,site,plot). This key will be used to build the file when download via the web.
- (D) Units of Measurement:
The standard of measurement for an attribute value. This element only applies to "Range" Data Types (next attribute) . Date and time data consists of valid date or time combinations, i.e., YYYYMMDD HH:MM:SS or MM/DD/YYYY HH:MM:SS. If only a date, only include representation for the date.
- (E) Data Type:
Assign one of the following: Character, Integer, Numeric, Date.
Character: *C(character width), i.e. text field that holds up to 10 characters = C10*
Integer: *Whole number values, i.e. I5 = integers up to 99999*
Floating Decimal: *Fixed precision and scale numbers, Fp.s, i.e. F7.2 = 2546.86 or -146.43*
p (precision): Specifies the maximum total number of decimal digits that can be stored, both to the left and to the right of the decimal point and includes both the decimal point and negative signs. The precision must be a value from 1 through the maximum precision. The maximum precision is 38.
s (scale): Specifies the maximum number of decimal digits that can be stored to the right of the decimal point. Scale must be a value from 0 through p. The default scale is 0; therefore, 0 <= s <= p. Maximum storage sizes vary, based on the precision.
Datetime: *D8 for date only*

(F) Nullability (Y/N):

Indicates whether an attribute can be left blank (missing values).

(G) Domain Type:

Assign one of the 5 domains that best represent the attribute. Be sure to provide codesets or references if necessary.

Enum: *Or enumerated; a finite list of codes or values. Need to define list or provide source.*

Range: *Comprised of a sequence, series, or scale of (usually numeric) values between limits. Values of measured or calculated data with a minimum and maximum.*

Taxa: *Predefined codesets (i.e., Garrison et al., 1972 or USDA Plants Database). Need to provide list of codes and a reference to scientific name, common name, authority, family.*

Place: *Predefined codesets (i.e., LTER Reference Stands). Need to provide list of codes and a reference to geographic information.*

Freetext: *Data table values cannot be represented. Reasons include attributes whose values do not exist in a known, predefined set (for example, a "phonebook" of people's names), or attributes whose values cannot be depicted using the forms of representation (available character set, etc) used for the metadata. In these cases, the information content of the set of values should be provided.*

(H) Measurement Scale:

Measurement scale is dependent on the domain type.

nominal – *categories (represents named categories, a list of coded values, or plain text descriptions)*

ordinal – *ordered categories (values are also named and in a set order with reference to one another, but the distance between them is not indicated (e.g., low, medium, high))*

interval – *This is a numeric value, and uses equal-sized units on a scale between values. The starting point is arbitrary, so a value of zero is not meaningful. Example: temperature.*

ratio – *Measurements have a meaningful zero point, and the ratio comparisons between values are legitimate. For example, concentration is a ratio measurement because a solution at 10 micromolesPerLiter has twice as much substance as one at 5 micromolesPerLiter.*

datetime – *dates or date-time values from the Gregorian calendar.*

Domain type and Measurement scale relationship:

enumerated – *nominal or ordinal*

freetext – *nominal or ordinal*

taxonomic – *ordinal*

place – *ordinal*

range – *interval, ratio, or datetime*

(I) Number Type:

Description of number type for interval and ratio measurement scales only. Use the most restrictive description defined below.

whole – *only positive integers, including zero*

natural – *all whole numbers, except zero (starts at 1)*

integer – *all non-fractional positive or negative numbers*

real – *all numbers that contain fractional or decimal elements*

(J) Minimum Value:

The smallest value that the attribute can be assigned from a continuum of valid values. Do not include units, but must be in the same units defined under 'units'.

(K) Maximum Value:

The greatest value that the attribute can be assigned from a continuum of valid values. Do not include units, but must be in the same units defined under 'units'.

(L) Value Accuracy:

An estimate of the accuracy of the assignment of attribute values (+/- some value). Do not include units, but must be in the same units defined under 'units'.

(M) Measurement Precision (resolution):

The precision of the measurement; the smallest unit increment to which an attribute value is measured (e.g., to 1m). Do not include units, but must be in the same units defined under 'units'. For example, for an attribute with unit "meter", a precision of "0.1" would be interpreted as precise to the nearest 1/10th of a meter, and a precision of "1" would be interpreted as precise to the nearest 1 meter.

Example(s):

0.1

0.5

1

(N) Attribute Measurement Frequency:

The frequency with which attribute values are collected. Be as specific as possible. Use words or phrases like single collection, continuous updated at 15-minute intervals, hourly, daily, weekly, monthly, annually, seasonal, weekly throughout summers, irregular, unknown).

(O) Missing Value Code:

Characters used to denote missing values (e.g., -999). Null values to represent missing data are the best representation; if this is the case, leave blank.

For domain types of enumerated ("enum"):

These are the fields in the spreadsheet tab called "enum" and describe and codes used in the dataset (all tables/entities). If the code set is a place/location or taxa, then use the appropriate tab to describe (place or taxa tabs). The letter before the heading is the column that it is located in the spreadsheet.

(A) Attribute:

The name of the field (attribute). Should be somewhat descriptive. Best to limit to 10 characters and should not include spaces or special characters.

(B) Code:

The code that is used in the dataset.

(C) Definition/Source:

The description of the attribute. Be clear and unambiguous. If appropriate, please list the authority for the attribute definition.

For domain types of “Place”:

These are the fields in the spreadsheet tab called ‘Place’ and describe each place or location (site, sitecode, plot, watershed, etc.) in the dataset (all tables/entities). If the place or location already exists (an already established site in our database), then there is no need to enter that location here. The letter before the heading is the column that it is located in the spreadsheet.

- (A) Location Code:
The code used in the dataset.
- (B) Location Name:
The full name of the location. Be clear and unambiguous.
- (C) West Bounding Coordinate or Latitude (decimal degrees):
West bounding coordinate or latitude in decimal degrees.
- (D) East Bounding Coordinate (decimal degrees):
East bounding coordinate in decimal degrees.
- (E) North Bounding Coordinate or Longitude (decimal degrees):
North bounding coordinate or longitude in decimal degrees.
- (F) South Bounding Coordinate (decimal degrees):
South bounding coordinate in decimal degrees.
- (G) Coordinate Method:
The method of how the coordinates were obtained (map, GPS, etc.). Please state the datum, if known.
- (H) Elevation maximum (m):
The maximum elevation of the site.
- (I) Elevation minimum (m):
The minimum elevation of the site.
- (J) Elevation Method:
The method of how the elevation was obtained (map, GPS, etc.).
- (K) Slope (%):
The general slope of the site in percent.
- (L) Aspect (degrees):
The general aspect of the site in degrees.
- (M) Geology:
Describe the geology of the site.
- (N) Landform:
Describe the landform of the site.
- (O) Soils:
Describe the soils of the site.
- (P) Hydrology:
Describe the hydrology of the site.
- (Q) Vegetation:

Describe the vegetation of the site.

(R) Climate:

Describe the climate of the site.

(S) Site history:

Describe any known site history.

(T) Site Description:

Describe any other site characteristics not accounted for above.

For domain types of taxonomic (“Taxa”):

These are the fields in the spreadsheet tab called ‘Taxa’ and describe each taxonomic entity in the dataset (all tables/entities). The letter before the heading is the column that it is located in the spreadsheet.

(A) Species Code:

The code used in the dataset to represent the taxonomic entity.

(B) Scientific Name:

The accepted scientific name of the taxonomic entity.

(C) Common Name:

The common name of the taxonomic entity.

(D) Author:

Author of the taxonomic entity.

(E) Family:

List the family of the taxonomic entity.

(F) Order:

List the order of the taxonomic entity.