AntWeb Documentation

Specimen Data Documentation

Summary

Specimen data can be submitted to AntWeb in two ways: (1) emailed as a spreadsheet or (2) uploaded directly as a text file. If you manage a growing (active) dataset, it will be easier to learn to upload your data directly to Antweb. If you have completed a dataset for a specific project (e.g. revision) that you want to include in Antweb, you can email the data to Antweb for uploading.

In order to upload your data directly to AntWeb, you must be sure to name the headers on your spreadsheet exactly as indicated in these instructions, and include the specified linking fields (in addition to Specimen codes you must include Locality, collection and species codes). If you email your data, we will ensure your data is complete and suitable before uploading it to AntWeb. Below we provide instructions on how to organize data and submit it to AntWeb via direct uploading and email submission.

Specimen codes (unique object identifier) are required in all types of data submission. The last section discusses different approaches and case scenarios for using specimen identifiers.

Outline:

Part I: Guidelines for ant data collection
Part II Spreadsheet preparation
Part III Specimen data submission
        email spreadsheet to CAS
        Direct text file upload
Part IV Creating and using specimen codes
Part V: Special instructions for Biota users
Part VI: FAQ's

Part I: Guidelines for ant data collection

A certain amount of essential information must be associated with every specimen. I outline below how to collect this information in the field.

Overview data to be collected.

The following locality and field data should be recorded for each collection:

(1) Country
(2) First administrative district
(3) Locality name
(4) Latitude and longitude
(5) Elevation
(6) Locality code (unique code for locality: optional but useful)
(7) Date or range of dates
(8) Habitat

(9) Collecting method or microhabitat
(10) Collector(s)
(11) Collection code (Field number or Collector number: essential code)

This data can be organized into two levels: Locality and Collection data.

Locality Data
Locality data include [example]:

(1) Country [Madagascar]
(2) First administrative district: State or Province [Province d'Antsiranana]
(3) Locality name: (a) name of area if available, and (b) distance and direction from
    nearest village (be sure to use a village that can be located on readily available
    published maps of the region [Réserve Spéciale de Manongarivo, 12.8 km 228° SW
    Antanambao]
(4) Latitude and longitude: obtained through GPS receiver
(5) Elevation: obtained through altimeter, GPS, or Google Earth maps [780 m]
(6) Locality code [Manongarivo 780]

a) A locality is defined as a site having a unique elevation, latitude and longitude. All
collections from a site with the same elevation, latitude and longitude are from the same
locality.

b) Locality codes are unique code to each locality. They provide a shorthand method of
referring to a site in field notes as well as facilitate the inclusion of specimen records in
the database. The system I often use includes a descriptive name and elevation. For
example, the locality code "Manongarivo 780" refers to the following locality described
in the label below. But any system that generates unique locality codes can be used. The
code is arbitrary but plays an important role in databasing specimens. The code can also
be generated at the time of data entry.

MADG'R: Prov. Antsiranana
R.S. Manongarivo 12.8 km
228° SW Antanambao 780 m
13°58.6'S 48°25.4'E 11.x.1998
B.L.Fisher#1859
ex rotten log
rainforest

Collection Data
Collection data include [example]:

(1) Collection code [BLF1859]
(2) Date or range of dates [11.x.1998]
(3) Habitat [rainforest]
(4) Collecting method or microhabitat [ex rotten log]

(5) Collector [B.L.Fisher]

a) A collection event is defined as the collection of a specimen or specimens having a unique date (or range of dates), habitat, collecting method/microhabitat, and collector at a locality.

b) Collection field numbers are unique numbers given to each collection and are linked to the collector by the collector's initials. The preferred method is to use sequential numbers. It is very important to never repeat a number. In the above example, the collection code BLF1859 refers to collection 1859 by B. L. Fisher.

c) Collectors should keep a record of collection data and use a collection code system (field number) to organize this data. For example, in the above label, 1859 is the collection number that refers to the collection at locality "Manongarivo 780" by B. L. Fisher on 11.x.1998, from a rotten log, in rainforest.

d) Collection BLF1859 could be the only collection or one of many collections at locality "Manongarivo 780."

e) Collection BLF1859 could consist of one or many specimens that all share the same collection and locality information.

Putting the data on the specimens
The collection number relates the specimen to the appropriate locality and collection data and will facilitate labeling mounted or etoh specimens with the appropriate locality label. Thus after you make your collection, the collection number should stay with the specimens as they are processed and mounted.

An example of how to collect the data
The method described below involves taking brief notes in a field book and taking more detailed notes in a collectors notebook. In both cases, notetaking is facilitated by the use of locality codes. Locality codes are defined only once in the field and collectors book, and thus do not need to be rewritten for each collecting event. In addition, this approach to collection data is organized to facilitate data capture for uploading to AntWeb.

1) In the field: brief field notes
a) Record locality data and define locality code.

b) Record date, locality code and habitat to be sampled.

c) For subsequent collections at the noted locality code, record: collection number, what was collected (genus of ant, etc.), and method/microhabitat. Use abbreviations as often as possible when taking these field notes.

c) Temporarily mark the collection code on the vial using a fine black permanent marker or insert prefabricated numbers.

d) Below are examples of entries in the field notebook [explanation in brackets].

In the first example, all collections on 25.x.1998 were from the same locality.

25 Oct 1998 Manongarivo 1175 [MADAGASCAR: Province d'Antsiranana, Réserve
           Spéciale de Manongarivo, 14.5 km 220° SW Antanambao 1175 m, 14°15.9'S
           48°25.7'E] montane rainforest [date, locality code, habitat]
1960   *Odontomachus* [collection number, field ID]
           ex rotten log [microhabitat]
1961   *Hypoponera* [collection number, field ID]
           ex rotten log [microhabitat]
1962   *Camponotus* [collection number, field ID]
           on low vegetation [microhabitat]
1963   Malaise trap in tree fall clearing [collection number, method]
           trap collected on 30 Oct 1998

In the second example, collections on 15.x.1998 were from two different localities. The
date is simply repeated when more than one locality is visited per day or range of dates.

15 Oct 1998 Manongarivo 780 [MADAGASCAR: Province d'Antsiranana, Réserve
           Spéciale de Manongarivo, 12.8 km 228° SW Antanambao, 13°58.6'S 48°25.4'E
           780 m] rainforest [date, locality code, habitat]
1899   *Pyramica ludovici* [collection number, field ID]
           ex rotten log [microhabitat]
1900   *Hypoponera* [collection number, field ID]
           ex rotten log [microhabitat]

15 Oct 1998 Manongarivo 610 [MADAGASCAR: Province d'Antsiranana, Réserve
           Spéciale de Manongarivo 12.1 km 223° SW Antanambao, 13°57.4'S 48°25.1'E
           610 m] rainforest
1901   *Camponotus*, *Leptogenys*
           ground foragers

25 Oct 98 Manongarivo 1175 m montane rainforest
1960 Odontomachus ex rotten log [collection number, genus of ant, microhabitat]
1961 Hypoponera ex rotten log [collection number, genus of ant, microhabitat]
1962 Camponotus on low vegetation [collection number, genus of ant, microhabitat]
1963 Malaise trap in clearing [collection number, method, microhabitat]

2) Back at base camp:
Back in camp, two things should be done as soon as possible:

a) Write out the collection codes on label paper with indelible ink and place the
appropriate collection number label inside each vial or whirl pack bag. I prefer to preprint
collection codes before going into the field.

b) Review field data and transfer the collection data from the field notebook to the larger collectors notebook if necessary or type out in a database or word document. Be complete as possible when databasing, and avoid abbreviations used in the field notebook.

(C) Habitat Data
Habitat data refers to general vegetation classification. There is no standard classification available, but even a simple description such as forest, grassland, or desert will provide some information on the habitat structure of the species. This large-scale habitat description can often be noted as part of the locality description. However, if the locality includes multiple habitats, the habitats will need to be noted for each collecting event or set of collecting events.

(D)Microhabitat and methods
Microhabitat and/or methods should be recorded for each collecting event. This is in contrast to the general habitat description, which may need to be noted once as part of the locality description, and not for each subsequent collecting event.

Microhabitat and methods descriptions vary widely depending on habitat. As an example, the most common microhabitats and methods used in collecting ants are presented below. "ex" is Latin for "from": ex rotten log refers to insects collected from a rotten log.

The list below also provides a good summary of where and how to look for ants.

under stone   on tree trunk(s)
ground forager(s)   ex rot pocket above ground
ex rotten log   ex rotting tree stump
on low vegetation   ex dead branch above ground
ex dead twig   at light
ex live stem ____   pitfall trap
ex rotten stick on ground   Malaise trap
ground nest   ex termite mound
sifted litter (leaf mold, rotten wood)   beating low vegetation
under tree bark, live tree   yellow pan trap
carton nest on foliage   sweeping

AntWeb Documentation

Part II: Filling in the spreadsheet for AntWeb submission

The following field names (spreadsheet column headers) are allowed in the specimen upload text file. You can copy the headers from the example file called **Specimen headers and examples**.

Priority fields (columns in the spreadsheet) are indicated in red. Blue fields are necessary when uploading data directly to AntWeb. If your dataset is missing the blue fields, then your spreadsheet must be emailed to AntWeb so that these fields can be completed before uploading.

SpecimenCode
> Unique number: see section on specimen codes below.

CollectionCode
> Unique number. Collector's initials and collection number is one option; this number can also be generated for you by AntWeb.

Genus
> Current valid genus name. For undescribed genera, see permitted morphospecies section

SpeciesName
> Current valid species name. Include genus and subspecies separated by a space in a single cell. Every specimen must include a species name. Accepted names include undet and morphospecies codes: see the discussion in the **Antweb format for species names/codes** section of the Regional Authority Species list help documentation.

SpeciesCode
> Species code created by combining genus name and species name fields seperated by a period. Use a space to separate species from subspecies ex. Anoplolepis.steingroeveri gertrudae

LifeStageSex
> Castes can be abbreviated or spelled out: aQ, dQ, w, m or alate queen, dealate queen, workers, male
> For pins or vials with multiple ants, separate castes by a "," for example: 39w, 2dQ, 4aQ, 5m

Medium
> pin, etoh

CollectionOptional [optional field for collection data]
> Use this field for additional data that you would like to associate with collections.

LocatedAt
> acronym of museum, university, or collector's initials.

OwnedBy
> acronym of museum or collector's initials.

TypeStatus
> lectotype, paratype, holotype, syntype or paralectotype

CollectedBy
> Name of collector or collectors

DateCollected

    month/day/year format

DateCollEnd

    month/day/year format

Method

    How were the specimens collected, by what method or in what microhabitat.

Habitat

    Broad habitat classification.

SpecimenOptional [optional field for specimen data]

    Use this field for additional data you would like to associate with specimens.

LocalityCode

must be a unique code.

LocalityName

Name of locality, distance and direction from a locality, etc.

County/sub-territory

    2nd administrative distinct:

    http://en.wikipedia.org/wiki/Table_of_administrative_divisions_by_country

State/Province/Territory

    Current name for first administrative district:

    http://en.wikipedia.org/wiki/Table_of_administrative_divisions_by_country

Country

    current approved ISO name of country:

    http://www.iso.org/iso/english_country_names_and_code_elements

Elevation(m)

    record in meters, include space after elevation then the letter "m"

Latitude

    record as decimal degrees

Longitude

    record as decimal degrees

Lat/Long accuracy

    how latitude and longitude were obtained, eg. map, GPS, Google Earth

Biogeographic Region

    Afrotropical, Australian, Palearctic, Indomalaya, Malagasy, Nearctic, Neotropical, Oceania

Data Provider
    Institution or individual responsible for specimen data curation: email optional

Part III Specimen data submission

Ideally, specimens bear the specimen code that is recorded in the spread sheet. For example, at CAS, we add a CASENT number to all pinned and vial specimens and record this number in AntWeb. However, there will be cases where this number is not present on the specimens. In these cases, AntWeb can provide a number (such as a CASENT# or AntWeb#) for you to apply to the specimen. If you have already recorded the data, and can't feasibly add a specimen code label to each specimen, you can add an AntWeb# in the spreadsheet. These records are referred to as collection "observations."

Antweb currently does not accept the following:
Specimen records recorded from literature or field observations that are not backed up with a specimen.

email spreadsheet to CAS
If you have a completed dataset that you no longer plan to update on a regular basis, you can submit the data via email to "Michele Esposito" <MEsposito@calacademy.org> for uploading to AntWeb. We will check over the data to ensure it fits the AntWeb format. If in the future you do update the data associated with your specimens, you can resubmit the data. Resubmission will add new records or update old records but will not delete existing records. If you delete specimen records, you need to let us know which specimen codes to delete.

Before submitting the data to Michele, you data should be formatted with the appropriate AntWeb column headers. If your data lacks locality codes or collection codes, these will be assigned at the time of submission. If we assign specimens locality codes or collection

codes, we will send a new version of the spreadsheet back to you with these codes. You will need to use this new version if you plan to update these records in the future.

Direct text file upload

If you maintain an active specimen database in Excel, Biota, or any other database application that can exported a tab-delimited file, you could potentially upload directly to AntWeb. However, if you upload directly to AntWeb, you must adhere strictly to column header names and you must include locality, collection and specimen codes. Databases such as Biota also require these fields. Though any database can be used, special instructions are provided for Biota users.

The advantage of direct upload is that you maintain control of all of your data. If you reupload your data, specimens that have been deleted are automatically deleted (not so with emailing specimen data). We strongly encourage direct file upload if you plan to maintain a growing database of specimen records.

Curating specimen data
Once you submit specimen data i to AntWeb, you may need to go back and update specimen records, identifications, or adjust locality information based on feedback from users of AntWeb. You can receive these comments directly if you provide your email address as a data provider. Otherwise, most people will contact AntWeb with the update.

**Part IV Creating and using Specimen codes**

Specimens identifies (specimen codes) used by AntWeb are registered unique identifiers. If you do not already have at your institution unique identifies, we can provide you CASNET or ANTWEB identifiers for you use. We can allocate you a set of numbers or even print the numbers on archival paper number for you to use.

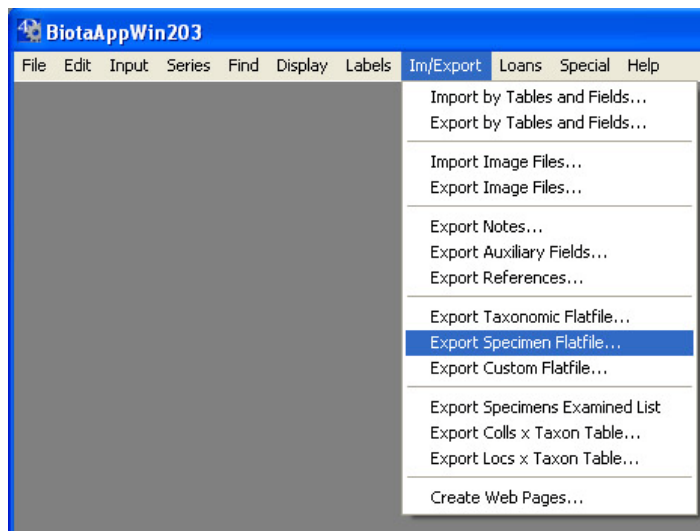**Part V Special instruction for Biota Users**

Uploading BIOTA specimen data to Antweb:

1. From the BIOTA **Find** menu option, select **All Specimens**.
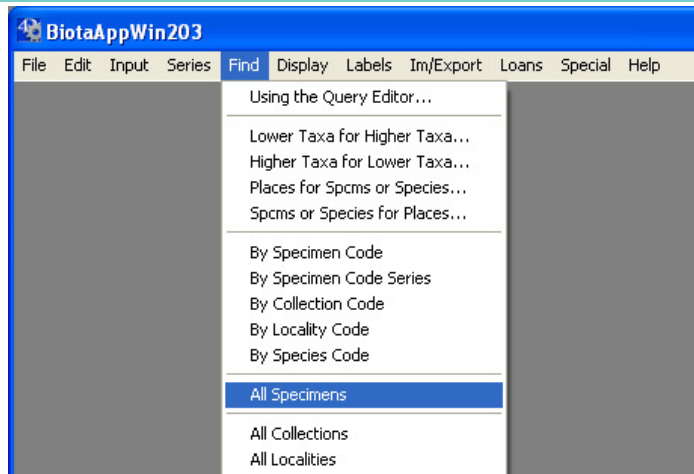
2. From the BIOTA **Import/Export** menu option, select **Export Specimen Flatfile**.

> i. Within the **Export Specimen Flatfile** window, click Select All to select all fields. IMPORTANT: un-select the **[Genus] Genus** box before exporting.

Export Specimen Flatfile

Check the box next to each field you want to export for the current Specimen Record Set

☐ [Genus] Genus	☑ [Order] Order	☑ [Phylum] Phylum Cust 1
☑ [Genus] Tribe	☑ [Order] Superorder	☑ [Phylum] Phylum Cust 2
☑ [Genus] Subfamily	☑ [Order] Subclass	☑ [Kingdom] Kingdom
☑ [Genus] Genus Cust 1	☑ [Order] Order Cust 1	☑ [Kingdom] Superkingdom
☑ [Genus] Genus Cust 2	☑ [Order] Order Cust 2	☑ [Kingdom] Kingdom Cust 1
☑ [Genus] Genus Cust 3	☑ [Order] Order Cust 3	☑ [Kingdom] Kingdom Cust 2
☑ [Family] Family	☑ [Class] Class	
☑ [Family] Superfamily	☑ [Class] Subphylum	
☑ [Family] Suborder	☑ [Class] Class Cust 1	
☑ [Family] Family Cust 1	☑ [Class] Class Cust 2	
☑ [Family] Family Cust 2	☑ [Phylum] Phylum	
☑ [Family] Family Cust 3	☑ [Phylum] Subkingdom	

Screen 3 of 3

Clear All   Select All   First Screen of Fields...   Second Screen of Fields...   Cancel   Export Fields

       b. Click the **Export Fields** button.
       c. Export as a text (txt) file and note the date in the file name eg., Biota dump032510.

3. Open the text file in Excel and replace the field names (column headers) with the corresponding AntWeb field names.

4. To check for problems within the data, such as additional returns, select all data in Excel and sort by Specimen code. Scroll down to the end of the data and check for mismatched field data.
       a. If no issues are found, close the spreadsheet and **DO NOT** save.Saving after sorting can cause errors with uploading to AntWeb.

5. Go to the AntWeb Admin Tool page and login.

AntWeb Admin Tool

Please Login to AntWeb

Username:	
Password:	•••••••

Submit

6. **Browse** to your specimen flat file on the Upload Files page.

a. **DO NOT** click the Update entire biot file
b. **Do** click Reload Pulldowns

7. Click **Submit** and wait for the "**Success**" screen.

AntWeb Documentation

Part VI: Specimen Data FAQ's

- **Is "100w" ok to put in the life stage/sex column? Is that the format you use for multiple individuals/collection event? How do you denote collections that include multiple semaphorontssuch as 39w, 2dQ, 4aQ, 5m? What are the acceptable abbreviations?**

  You can use abbreviations or full text works for caste such as aQ, dQ, w, m or alate queen, dealate queen, workers, male. You can also include the number of each caste such as for a record for a vial: 100w, 1dq, 3m.

- **How do morphospecies identifiers fall out in the different columns? e.g., for Amblyopone mg01, is "current valid species" left blank, while the species code = Amblyopone.mg01?**

  Morphospecies names are recorded in the **SpeciesName** column. The species code field is only necessary if uploading directly to AntWeb. Species codes are composed of the genus and species name field with a "." separating the words. Spaces are permitted between species and subspecies names, thus the species code for Pachycondyla darwini madecassa would be Pachycondyla.darwini madecassa.

- **Does every specimen have to be identified to species/morpho-species? Although I am coming closer to understanding what is going on with Palauan Pheidole, many of my specimen records are still only identified to "Pheidole".**

  The Species name field cannot be left blank.
  For taxa that have not been named, even to species code, use "undet". Thus instead of just "Camponotus" or "Camponotus sp." use Camponotus undet.

- **Can I update the taxonomy later (e.g., what if I discover that Pheidole jcm01 is really Pheidole umbonata? What if I discover that there are several cryptic species within what I was calling "Pheidole sexspinosa"?) Will I be able to update the specimen records by uploading a new spreadsheet?**

  One advantage of a specimen-level database is that you can update specimen identifications. You can re-upload or email your data as many times as you want to change the identify of specimens. Specimens my go from undet, to morphospecies code, to named species, and each upload or submission will reflect these changes to your data. If you do email your specimens to AntWeb, note that if you delete specimen records, you must indicate in the email which specimen codes should be deleted.

- **What strategy should be used when multiple users are uploading data linked to the same collections/specimens? For example, if I upload new Stenamma**

determinations and then my colleague, as result of how he imports data to AntWeb, uploads his older determinations, will mine be overwritten?

We want to avoid having two separate data providers manage the same specimen data. However, this situation is bound to occur eventually. A taxonomist may visit collections from multiple institutions when working on a revision of a genus, and then submit the data to Antweb. Meanwhile, institutions may be submitting data from their own collections. Currently, only the most recent submission of a specimen record is shown on AntWeb.

One way to avoid this is to submit corrections and updates to instituitions that house the specimens.

- **Is there a way to be clear about the source and precision of GPS coordinaes? By source, I mean whether the data come from a gazetteer, a GPS unit, or Google Earth. Is this information included in Biota as an auxiliary field? For precision, is it possible to indicate the number of sig figs for every entry? A problem of the current system is that if you convert latitude/longitude information from degrees and minutes to decimal degrees, you often end up with coordinates that are too precise. AntWeb should make clear to users how precise the data are.**

  Currently, the field **Lat/Long accuracy** is the only field available to put information about the accuracy and precision of coordinate information. This field does not currently appear on AntWeb specimen pages, but is slated to appear soon.

- **How should locality codes be assigned? As locality data become more and more precise, it becomes very tedious and less useful to assign locality codes. Could AntWeb automatically assign locality codes for each collection event if the user does not include this information?**

  Locality codes are unique to each locality. They provide a shorthand means of refering to a locality in field notes, and facilitate the databasing of specimen records. A locality is defined as a site having a unique elevation, latitude and longitude. That is, all collections from a site with the same elevation, latitude and longitude are from the same locality The system I often use to generate a locality code in the field includes a descriptive name and elevation. For example, the locality code "Manongarivo 780" refers to a site at 780 m in the park Manongarivo in Madagascar. Any system can be used that generates unique locality codes. The code is arbitrary and but plays an important role in databasing specimens. The code can also be generated at the time of data entry. If you are generating a locality code at the time of data entry, and you can use a simple code such as PSWLOC1245.

- **What are the auxiliary fields commonly used in AntWeb? What extra data can be uploaded to AntWeb besides the standard fields currently presented?**

  In your specimen data, you can include one collection level optional field with header **CollectionOptional** and one specimen level optional field **SpecimenOptional.**

- **Is there a convention for stage/sex data? Many of you use a one specimen/pin system, but when there are multiple specimens/pin, as is common on older material, how should one indicate this? 1dq, 2w?**

In the **LifeStageSex** field, you can include both caste and number of individuals. However, this practice causes confusion when images of a single specimen are recorded yet the LifeStageSex field indicates more than one specimen. In the future we hope to include a field to indicate the caste of the imaged specimen.