Transcribing MSU's specimens through Symbiota's Consortium of Midwest Herbaria Portal

Matt Chansler – 1/19/2021

Some background:

I. First of all, <u>thank you</u> for your help with this big job!

II. About MSU's fern collection (Pteridophytes).
   a. We have about 15,000 fern specimens.
   b. Our goal is eventually to make the entire collection online and accessible to botanists all over the world.

III. We are on a large grant that is providing resources to digitize ferns at many institutions
   a. This is a stepwise process
      1. Image specimens (mostly done!)
      2. **Transcribe labels (what we need help with!)**
      3. Georeference using label data (for later)
   b. Transcription makes it possible for botanists to download large sets of data that can be used to answer important questions about wild plants. These include:
      1. How many kinds of plants are there?
      2. Where do plants grow?
      3. How do plants' physical traits change over time?
      4. How do plants' ranges change over time?
      5. Where are rare or invasive plants found?
      6. Where do we need to collect more?
      7. Which institutions hold which specimens (for physical loans)?

To access our specimens:

I. Our ferns are viewable at http://midwestherbaria.org.
II. To access them:
   a. Create a username and password and log in
   b. Navigate to the "Crowdsource" tab at the top.
   c. Look down to the bottom where it says, "Your Stats By Collection" and click on "open records" from Michigan State University (MSC).
   d. This shows a new table with specimens that are available to transcribe. You can browse them by clicking on their "Symbiota ID"

To transcribe specimens:

I. Start by adding the MSU accession number into the "**Other Cat. #s**" box. These are usually stamped in the upper left portion of the specimen.

II. Change the "**Status Auto-Set**" dropdown box at the bottom to "Stage 2".

III. Zoom in on the label in the lower right corner. You can hold shift+left mouse button to zoom in and out. It's also possible to change the size of the window so you can see more of the image!

IV. Enter these boxes first:

> *Is the specimen illegible?*
>
> *If it seems like it will take a long time to interpret, only fill **MSU accession number** and **Collector**, and then change **Status Auto-Set** to "Unprocessed/NLP" before moving on. We're going to deal with illegible ones at a later time!*

   a. **Collector**. Different collectors write their names in different ways; just copy it how they wrote it. If it says "Doe, John" don't change it to "John Doe". Sometimes, there might be multiple people here and if a single one isn't obviously the main collector, just add all their names separated by semicolons. If no collector is named, fill with "unknown".

   b. **Number**. This will usually be written nearby the name of the collector, but not always. If there is no collector number given, write "s.n."

   c. Once Collector and Number are filled, click the "Dupes?" button. This will check to see if a *different* herbarium has a duplicate specimen they've already entered. If they do, you might be able to fill in blank boxes at the press of a button! **Make sure to verify the boxes that get filled in actually match what's on the specimen's label.**

V. If duplicate matching doesn't work, enter the rest of the label

   a. **Date**. This must be entered in the format of "2020-03-05". If only a month is given "2020-03-00"; if only a year "2020-00-00". If there is no date at all, leave blank.

   b. **Associated Collectors**. These were other people tagging along with the Collector. Add with a semicolon separating multiple people.

   c. **Verbatim Date**. The date as written by the collector. If no date is given, fill with "s.d."

   d. **Exsiccati Title** and **Number**. Don't worry about these fields; Exsiccati specimens that are part of a set and we don't track them for plants, only fungi.

   e. **Scientific name**. This should already be filled in. If you find a specimen whose name does not match anything on the sheet, (especially paying attention to additional labels glued or penciled-in above the main one), let Matt C. know.

   f. **Country, State, County**, and **Municipality** (access Municipality by clicking "Long-form" up above "Dupes?"). You may sometimes need to google what's written on the label to determine which field to put it in. As you start typing into the field Symbiota will automatically pull the name up if you're on the right track for all except Municipality. If you are unsure about what category a place belongs in, you may simply include it in Locality.

g. **Locality**. This is the specific location the plant was collected, often akin to directions. Examples are "5mi S of logging road" or "Sanford Natural Area, adjacent to Hagadorn Rd". Be sure to split this apart from any Habitat or Substrate data.

h. **Latitude, Longitude, Uncertainty**, and **Verbatim Coordinates**. These are all fields for GPS coordinates. Press the "F" box, where you can enter GPS coordinates that are given as degrees, minutes, seconds and then click "Insert Lat/Long Values". This will automatically convert to decimal degrees. Township, Range, and Section maybe be entered similarly, but I don't know how to convert those to GPS coordinates yet. We'll deal with it later!

i. **Elevation in Meters** and **Verbatim Elevation**. Enter Verbatim first with units, and then click the left arrow. Watch it fill in automatically!

j. **Depth in Meters** and **Verbatim Depth**. Only for aquatic species, generally not ferns except for maybe *Salvinia* and *Marsilea*.

k. **Habitat**. Characteristics of where the plant is growing. *Examples are* "In dry brush", "under dimly lit canopy", "Banks of the Rio Ramos", "Wet calcareous banks", "under trees, edge of irrigation ditch".

l. **Substrate**. Overlaps with habitat, leave blank.

m. **Description**. Field characteristics of the plant. *Examples are* height, flower color, number of flowers. Usually appearance.

n. **Notes**. Any other notes about the specimen, particularly abundance information (e.g. common, rare, frequent). This are remarks that don't fit into any other category.

o. **Cultivated.** If the label says the plant was cultivated, make sure to check that box.

When to open the long-form page with additional boxes:

There might be some additional information on the labels that don't fit in any boxes in the short form version and you need to open the long form:

1. **Annotations**:
    a. **Scenario 1.** Only original label present, and a different determiner is not specified. Leave "Identified by" and "Date identified" blank.
    b. **Scenario 2.** Only original label present, but someone besides the collector identified the specimen, often indicated with something like "Det. John Doe, 1976. Enter these data into "Identified By" and "Date Identified".
    c. **Scenario 3.** An additional annotation label is glued or penciled in above the main one that gives the most recent name of the plant specimen, and a person's name/initials are given. Enter these data into "Identified By" and "Date Identified".
    d. **Scenario 4.** An additional annotation label is glued or penciled in above the main one that gives the most recent name of the plant specimen, a person's name isn't given, but an "=" sign is present. Fill "Identified By" with "**Nomenclature update**".
    e. **Scenario 5.** An additional annotation label is glued or penciled in above the main one that gives the most recent name of the plant specimen, a person's name isn't given, nor is an "=" present. Fill "Identified By" with "**Unknown**".


Spore-producing structures under leaf of hart's tongue fern, an endangered species.

Some additional Matt thoughts and examples:

Google is your friend! The two biggest challenges are 1) Interpreting which information goes where and 2) Deciphering poor handwriting. Here are some examples of easy, medium, and hard labels:
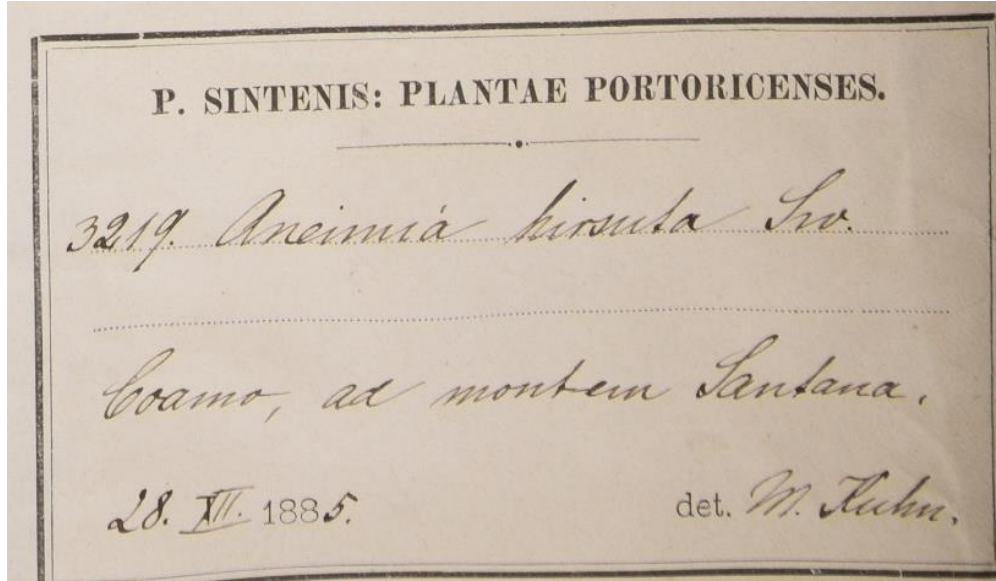
Easy: Typed, not faded, categories are all clear, all in English. ADDITIONALLY, upon entering collectors (D. Dunn; C. Dziekanowski; M. Pennell) and Collector number (23043), duplicate matching works.
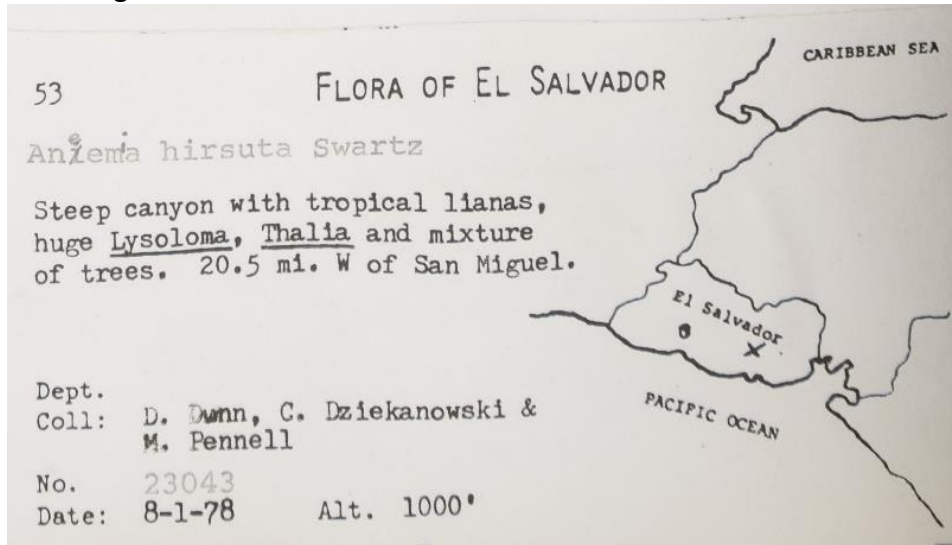
Medium: Typed, but slightly faded. Requires some googling of "Branganca Paulista" to verify what that means.
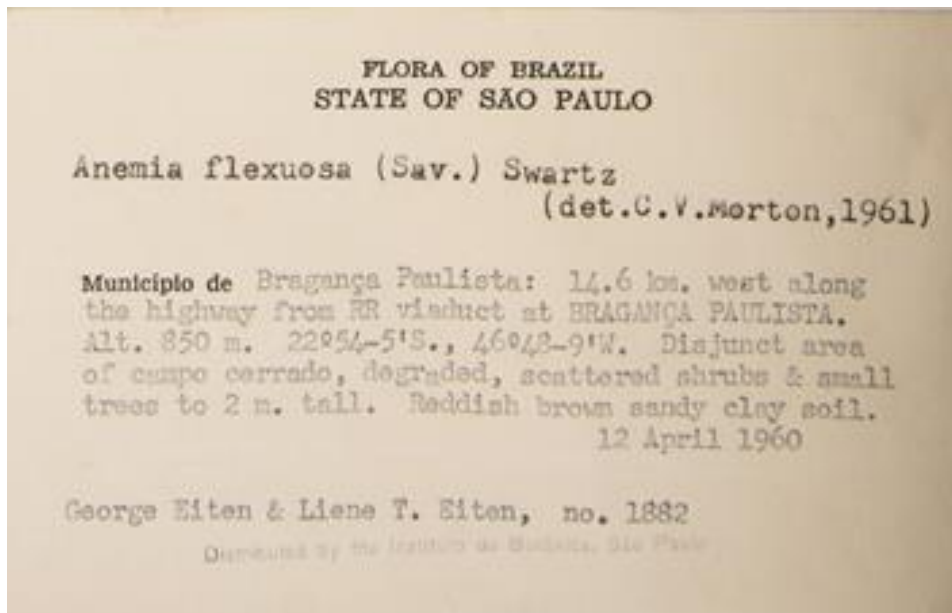
Hard: Written in cursive, Latin. Text says "Coamo, at Santana Mountain" in English, so it's Locality information. I had to Google "Boamo" and "Goamo" first because I couldn't tell what that first letter was, then I realized it said "Coamo". **Transcribe "P. Sintenis" into the Collector field, "M. Kuhn" into the Identified By field, "Puerto Rico" into the Country field, "Coamo" into the County field, and "ad montem Santana" into the Locality field.**

P. SINTENIS: PLANTAE PORTORICENSES.

3219 Aneimia hirsuta Sw.

Coamo, ad montem Santana.

28. XII. 1885.   det. M. Kuhn.

Ground cedar (not a cedar!), part of the so-called "fern allies" (not true ferns). They are spore-producers and are one of most ancient remaining lineages of vascular plants.