Batch Uploading of Specimen Records

Herbaria that prefer to enter their data in an offline database (and there are many good reasons why this might be so, can upload multiple records at once, the only restriction being that the file size must not be larger than 2GB. One can also enter data into a spreadsheet but, for several reasons, it is better to use a database such as Access or Base, the database suite in OpenOffice.

Once the upload process has been established for your csv files, batch uploading is very simple UNLESS the format of the upload file is changed. If it is changed, it will be necessary to establish a new protocol.

Format

Files to be uploaded must be comma delimited (i.e., *.csv). Excel and probably other spreadsheets can be saved in this format; database export from database programs can be exported to this format. Be aware of one problem: a comma, ",", can be an integral part of many fields (locality, habitat, people's names). If saving from an Excel spreadsheet, text fields will be enclosed in quotation marks (" …"). If generating your csv file from some other program, check that the text fields are delimited in this manner.

Fields

The upload function allows for more variability in uploaded files than suggested by the data entry page. For example, the first collector can be split into CollectorsInitials and CollectorFamilyName, where initials can also include names that are completely spelled out (e.g., G. Ledyard)and family names can include suffixes (e.g., Stebbins, Jr.). Date can be entered as three columns (day, month, year in any order). Month can be a number; year should ALWAYS include 4 digits. For dates such as Sept. 4-5

Although there is a field called "municipality", it is best not to use it because it is not considered when using the Geolocate tool. Instead, concatenate it with other text locality information.

There is a field "Verbatim coordinates" for UTM, TRS, and degrees, minutes and seconds. It would be better to calculate decimal latitude and decimal longitude from the degrees, minutes, seconds information and place them in two fields decimallatitude and decimallongitude. Remember than longitudes in the western hemisphere are negative and latitudes are negative in the southern hemisphere. Do not include more decimal places than are reasonable. At 40°N, 1 sec of longitude is

about 24m. It is about 31m at the equator and 0m at the poles. For latitude, there is little variation, 1 sec corresponding to about 20m.

Remember that it must be clear whether elevations are in meters or feet. One way is to have separate fields for minimumelevation, maximumelevation, and elevation unit. Another is to convert elevations in feet to their corresponding values in meters; again, be reasonable as to how many significant places are used.

Reproductive state or phenology? Botanists are more apt to speak in terms of phenology but, whichever is used, the information will be stored in the same DarwinCore field.

Comments such as "abundant in this area", "Flowers usually white, varying to pink" belong in the OccurrenceRemarks field. Use the phenology/reproductivestatus field to state whether flowers are present on the specimen.

The file must have a primary key field that contains no duplicates. The problem with accession numbers is that, all too often, there are duplicates within a collection. The easiest thing to do is to add a field "primary key" and fill it with sequential numbers. The field need not be imported. It is just to ensure that none of the records in the file will overwrite each other.

Have file, will upload

It requires administrative privileges to upload a record file. If you have these, login, click on the name of your collection, go to the Administration Control Panel and click on Import/Update specimen records. (Fig. 1). If no one has used the batch upload for your herbarium yet, clicking will bring up a screen saying that "There are no upload profiles associated with this collection" and a second line saying "Click here to create an upload profile".

Clicking on "here" in the second line will enable you to create a profile. You will have to give it a name. Use something like "Trial" for your first attempt. You will then be directed to select a file, identify the primary key, and then upload the file. When you do this, you will be shown a table in which Symbiota for matching the field names in your file with those used by the system. At the bottom of the page there is a button to try automatic mapping but, whether you use that or not, check that each of your fields goes where you think it should – and that no fields are skipped. Once you are satisfied, click the "Save mapping" button at the bottom of the page.

Fields will be skipped if you have a "column" with no name. Symbiota will think it has come to the end of your fields. IF everything looks great, upload the records.

I strongly suggest trying a file with just a few records first. If all goes well, add the rest; if all does not go well, there will be few records to correct (or delete). Note the problems and figure out which fields or mappings you need to amend and try again.

If there is already a profile for your collection, select it. This will take you to the screen asking you to select the file for upload. Once you upload it, you will be shown the mapping screen with fields that Symbiota knows how to map (because the mapping is store under the profile name) in white and fields that are new in yellow. As when first creating a profile, you need to check that your fields are mapping to the fields they should go to. If they do not, start over and create a new profile. If they do, then complete the upload. You will be shown a table stating how many records appear to be brand new and how many are updates. If the numbers seem correct, go ahead and transfer the records to the main table. When you do this, the statistics for your collection will be updated. If they do not, it is a matter of starting over.