**Setting up a “live” collection on SCAN**

October 9, 2017

Neil Cobb

**Predigitization Steps**

# Register an institution code for your university/museum on The Global Registry of Biodiversity Repositories (GRBio) <http://grbio.org/> if it is not already registered. Ideally, all the collections at an institution should share the same institution code, but if not at least be aware of the institution code used be all the biological collections at your institute. It may be helpful to peruse the institution and collection codes already in use by museums on GRbio and SCAN. SCAN museums are listed after the names of collections on the index page <http://symbiota4.acis.ufl.edu/scan/portal/collections/index.php> . For example, Texas A&M University Insect Collection (TAMU-TAMUIC) has the institution-collection codes that are four and six letter acronyms. Unfortunately there is no regulation by GRBio or coordination to enforce standards.

At some point register your institution and collection with GBIF, but you can do this after your collection is set up on SCAN and you have records to serve. You do not need to register your institution/collection with iDigBio, we will make your data available to iDigBio and you will be registered with iDigBio the first time they harvest your data. They will use information from your SCAN registration.

Establishing a catalog code is important; you want to have a format that works as a Globally Unique IDentifier (GUID) so that your records can always be traced back to your collection. A triplet standard was attempted that included the institution code, collection code and an alphanumeric that adds to the GUID value of the entire catalog number. An example of a GUID type format is described here that we use for the insect collection. <http://www.lep-net.org/?page_id=43>

Precurate your collection to make transcribing efficient, this primarily means organizing by taxa.

Create catalog labels and apply them to specimens. You can add a 2D barcode label as well as an alphanumeric label or put both on one label. Typically you do not want to get ahead of yourself and label too many specimens before you enter them into the database. Typically people label and then enter a drawer at a time. You never want to be in the situation where you find a specimen with a pin label that has been moved several times and you find out that it and a drawer’s worth of specimens have been re-distributed throughout the collection and none of them have been entered into the database.

**Digitization**

You do not have to initially enter data into the SCAN database one record at a time, the primary reason to enter your data into a spreadsheet instead of Symbiota is if you have a lot of carry-over information from one record to another. Symbiota does allow for some carry-over information but not a lot. The downside of a spreadsheet it that it does not easily accommodate controlled vocabulary to prevent misspellings, which can be rampant for some fields (e.g., taxa names). Excel also has the capacity to change content (e.g., dates) when copying from Excel to any other database, including Symbiota. We are working to make Symbiota data entry so good that there is never a reason to use Excel spreadsheets.

There are several user guides on Symbiota.org and [www.lep-net.org](http://www.lep-net.org) that will help navigate someone through the data entry form and for most fields there is a pop up that defines the function of the dialog box.