DATA MANAGEMENT PLAN

The Phases I-III of the project proposed here represent the first steps in enacting RWLC’s strategic plan, which calls for prioritization of collections safety and security (of both humans and specimens). A second, but paramount, priority is working toward full digital accessibility. To accomplish this we will use the computing power described in “Facilities” with regular back ups on several redundant external hard drives. During the project we will be connecting different types of data and introducing new forms of managing those data (i.e. barcodes). Initial data capture (Phase I) will involve scanning individual specimen cards from hard copy card files. Each individual scanned image will then be linked with a hand-keyed RWLC catalog number in the Specify database (on a Dell desktop PC). The Collections Intern inputting the scanned image will check the catalog number against the entered number once upon opening of the record, and once again before the closing of the record. During the Intern’s work period, the Collections Assistant will check in with the Intern to address any questions or concerns. At the end of the day the Assistant will open up 30% of the records entered to ensure match between the RWLC number the scanned images and the keyed number, and address any problems with workflow the following day, in consultation with PI Rundell. Given that the card file is well-organized we expect this work to go smoothly and quickly.

The next step (Phase II) in the data pipeline involves affixing printed barcodes to specimens, scanning the barcodes (printer and scanner purchased through this grant), and linking them with RWLC numbers in Specify. If sufficient new blank label space does not exist on the specimen, an additional archival foot tag will be carefully attached to the specimen so that the barcode may be affixed. The specimen’s cabinet number and drawer number (from the old Illick RWLC) will be recorded so that no specimens are misplaced during the process and specimens can be easily located during the taxonomic reorganization process in Phase III. The specimen will then be graded on a scale between 1 and 10 (1=conservation problem) so that our worst-case specimens can receive special treatment and so we can better assess priorities, both for the current project and for the coming decades. This quantitative assessment process is important in this new era for the Roosevelt Wild Life Collections, since we want to ensure we are making steady progress toward modern standards. Such assessment is also an invaluable tool for budgetary purposes. Assessment data (number) will be entered in a separate field. A list of common problems will also be included in a separate field so that we can identify these problems quickly through a search, rather than all of the information ending up in a “Notes” field. Hard-copy label data (i.e. foot tags) will also be carefully assessed at this stage. Since these sometimes delicate tags are about to move with the specimen into a bag and into the freezer for IPM, the Collections Manager will carefully supervise this process to ensure that specimens with flimsy or detaching tags get moved to a separate staging area for repair prior to freezing, so that their tags are not lost when they come out of the bags. PI Rundell will check on workflow efficiency and data quality on a regular basis, both through checking in with the Collections Manager, and directly.

In Phase III the RWLC specimens have received pest treatment and are safely in the new collections Center, still organized as they were in the old RWLC. New cabinets will be temporarily labeled with the new taxonomic reorganization (using Tree of Life and related publications), allowing plenty of room between drawers for expansion. Drawers coming from the freezer will be moved to a staging area for the final thaw. When they are ready, individual bags will be opened systematically, barcodes scanned, and foot tag data will be hand-keyed into the
Specify database. Any delicate tags at this stage will be fitted with a small narrow archival envelope to protect them. Ideally we will keep the original labels with the specimen, but if a tag does need to be separated for some reason, it will first be imaged, and then stored together in a single file drawer in the Center, within a labeled envelope in a plastic archival hanging folder. Skulls and other skeletal material may also require reboxing at this stage (archival-quality bone boxes) so that label information is keyed into the database at the same time as the reboxing occurs, and by the same person. This will help minimize risky handling of the specimens.

Data quality will be checked every couple of hours by the Collections Manager, (and each morning as the next round of digitization begins) while the Collections Assistant or Intern takes a break, lines and moves new drawers or works on freezer or organization tasks. Periodically PI Rundell will also assess the workflow and work with the Collections Manager to ensure that staging areas do not become overcrowded with specimens awaiting digitization, and to address those problems as they arise. Specimens that have finished being digitized then move to fresh cabinets lined with archival shelf liner. Interns who turn out to be especially mistake-prone in the digitization effort will be moved to other important collections tasks such as preparing new specimens and the reorganization effort. Data quality is the primary concern, and if Collections Interns are too inconsistent, the bulk of the digitization work will be handled by the Collections Assistant or Collections Manager.

PI Rundell, the Collections Manager, and the grad Collections Assistant will work together to ensure the best configuration of fields in the database and to modify our approach as necessary. We will check in with each other to ensure regular back ups. The new Roosevelt Wild Life Collections Research and Education Center will also be fully outfitted with wireless access (unlike Illick Hall) and computers will be fully networked, password-protected and maintained by our on-campus Computer Networking Service and IT professionals, who will be able to trouble shoot any database problems should they arise. We will plan for regular backups on an internal server. We will also still plan on maintaining multiple hard drive backups, one of which will be networked and maintained off-site in case of fire or other unforeseen disaster.

Data will be made available to the public online through our new RWLC website, currently in draft stage and in development (SD.16). Data will not be made public until it has been vetted by our collections staff. The working database will remain password-protected and only accessible by the collections staff on campus. The public database will be a copy of those data that cannot be edited. Any sensitive information about endangered species localities will be discussed by PI Rundell and the Collections Manager before it is released to the public. In a small number of cases, such data may be made more general for the public database in order to protect specific localities from poaching.

Our SUNY-ESF Communications Department is in charge of integrating data to be made available over any of our ESF websites and so they will interface with our RWLC team to make our first complete digital records from Phases I-III available to the public online. Our Wild Life on the Web high-resolution images will also be linked to our website and iDigBio (linked images possible in Specify 6.5). All TIFF images will be backed up on a SUNY-ESF server as well as external hard drives (LaCie d2Quadra 2TB). As our nestling RWLC database grows and becomes more complex (e.g. new collections data entries with GPS points allowing for precise georeferencing of localities), we expect to solicit funding for a database expert who can work closely with RWLC to address our needs and ensure that public accessibility (including speed of and quality of record access) and data security is maintained.