This workflow was developed at an iDigBio workshop in January 2015. The most recent version is available at https://github.com/iDigBioWorkflows/FlatSheetsDigitizationWorkflows and https://www.idigbio.org/content/workflow-modules-and-task-lists.

Appendix S14. Module 14: Proactive Digitization

Proactive digitization involves moving digitization processes ahead in the collecting process, resulting in digitization activities being completed prior to accession, hence avoiding the creation of additional legacy data. The order of tasks may depend on the requirements and limitations of the database used. Example implementations from two institutions are included as appendices to this workflow.

Task ID	Task Description	Explanations and Comments	Resources
T1	Determine institutional requirements and policies.	Check with institution where specimens are to be deposited about data format expectations and policies, including any permitting requirements. Institution may provide a template data file with desired fields and formats.	Collections manager or database administrator. Institutional collection policies.
T2	Determine what data will be collected for each specimen.	In general, record copious data. Commonly recorded categories of information include: ● collector name(s) and number ● event date ● locality ● GPS coordinates ● elevation or depth ● habitat, site description ● associated species ● specimen description ● tentative identification Geospatial data should include, at a minimum: ● latitude/longitude in decimal degrees or UTM. ● extent of site and/or accuracy of coordinate (e.g., GPS error) and units used ● geospatial datum ● source of coordinate (e.g., GPS unit, Google Earth, measured on paper map)	Specimen Field Notebook guidelines: http://collectionseducation.org/field-notebook/. Darwin Core standard: http://www.tdwg.org/standards/450/. Berkeley MVZ, "Guide for Recording Localities in Field Notes": https://www.idigbio.org/sites/default/files/workshop-presentations/ttt2/GoodBadLocalities.doc.

		If data do not precisely conform to the Darwin Core standard, they should be easily mapped to it.	
T3	Create unique identifier specimen labels before going into the field.	For example, vials for DNA samples can be pre-labeled with identifiers or barcodes, and later associated with field notes while collecting. This is an optional step depending on collector or institutional policies and procedures.	DNA Barcoding Assistant: http://www.dnabarcodingassistant.org/.
T4	Record collection data in the field.	Could be in a field notebook, in an app on smart phone or tablet, or using a spreadsheet on a tablet or laptop. Each physical specimen should include an unambiguous reference to the corresponding field notes entry (for example, the collector's name and number written on the newspaper or a tag).	List of apps and software for field data collection: http://brunalab.org/apps/.
T5	Capture in-situ photographs of specimens.	If this is desired, one needs to consider: ● standard set of images to document occurrence, ● metadata (e.g., camera model and settings, format, date and time), ● copyright, usage rights, license, ● file naming scheme, ● method for linking photos to field notebook entries and database records.	See: Baskauf and Kirchoff (2008). http://www.cals.ncsu.edu/plantbiology/ncsc/vulpia/pdf/Baskauf_%26_Kirchoff_Digital_Plant_Images.pdf.
T6	Transcribe field notebook into electronic format.	This step can be skipped if data were directly entered digitally in T4.	Excel spreadsheet, Specify Workbench, temporary database (e.g., Access or FileMaker).
T7	Enter data into database.	This depends on form of data and requirements of database. Many institutions will not give the collector direct access to the database, in which case the collections manager or database staff can perform this step, usually as an electronic upload from a spreadsheet submitted by the collector, the key point being that the collector is	Database. Excel spreadsheet. Specify Workbench. Darwin Core standard.

		supplying electronic data for each deposited specimen to prevent the creation of new legacy data. Field notes can be cleaned and augmented for completeness at this step, e.g., to refine the locality description or coordinates.	
T8	Produce labels.	This can be accomplished using various software, e.g., Microsoft Word mail merge, Symbiota, Specify, or institutional database software. Depending on the specific workflow, this step may occur between steps T6 and T7 using a mail merge on an Excel spreadsheet.	See: Creating labels in Microsoft Word using mail merge from data in Excel spreadsheet: http://ww2.valdosta.edu/~rcarter/HERB/Two%20birds%20with%20one%20stone.pdf.
T9	Record permit information, if available.	In situations where a collecting permit has been obtained, permit data might also be captured at this stage. Follow institutional policies and procedures. A reference or link to permit information may also be stored with each specimen record in the database and/or printed on the specimen label.	

Literature Cited

Baskauf, S. J., and B. K. Kirchoff. 2008. Digital plant images as specimens: toward standards for photographing living plants. *Vulpia* 7: 16–30. http://www.cals.ncsu.edu/plantbiology/ncsc/vulpia/pdf/Baskauf_%26_Kirchoff_Digital_Plant_Images.pdf.