

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 S1. Module 1: Pre-digitization Curation

Task ID	Task Description	Explanations and Comments	Resources
T1	Select specimens to digitize.	For herbaria, this may include all specimens. If this is not the case, selection should follow the institution's pre-determined digitization policies or project management plan. For example, specimens from certain geographic regions, taxonomic groups, collectors, or times may be targeted. This can be defined by a proposal, research project, algorithm, or other. Fitness of use of data is framed by project scope.	Digitization policy manual or project management plan.
T2	Visually identify collections, cabinets, shelves, folders, and/or specimens to incorporate into the digitization program.	<p>Apply signage to determine what is in and out of the scope of the project or relevant to the particular stage or sequence in the project. This step is useful when systematically digitizing an entire collection and otherwise potentially helpful with herbarium inventory. Aids for indicating storage locations may be helpful, for example one-dimensional barcodes or QR codes indicating storage locations and/or digitizing destination (rooms, cabinets, shelves, folders, drawers, etc.).</p> <p>This task is especially helpful for collections that are digitizing in random order, only portions of the collection related to specific projects, or anticipate significant physical or temporal separation between the Pre-digitization Curation, Data Capture, and Imaging modules.</p> <p>This step might also include developing protocols to communicate</p>	<p>QR codes, DataMatrix codes, or 1D barcodes. Flagging accessories.</p> <p>See: Diazgranados and Funk (2013). http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3819127/.</p>

		<p>progress in the digitization process. For example, signage could be applied to note where collections, cabinets, folders, or sheets are in the digitization process. Marking completed sections may prevent filing of new specimens that have not been digitized. It is ideal to always complete a task for everything in a folder.</p>	
<p>T3</p>	<p>Establish barcode format and order barcodes.</p>	<p>What a barcode represents at an institution may vary. Some use barcodes as a means to assign a locally unique identifier to a specimen (as defined by the ICN) and give the barcode a broader level of uniqueness (but not truly globally unique) by including an institution and collection code with the locally unique value. Barcodes may also be used as <code>dwc:catalogNumber</code> in the institutional database and thus tie the electronic record to the physical object. Some use legacy accession numbers as barcode numbers. In such cases, each sheet (irrespective of whether or not it corresponds to a specimen) may have a locally unique barcode.</p> <p>It should be noted that some institutions choose not to use barcodes, preferring instead to use an unduplicated accession number as the catalog number. Such accession identifiers might be generated by self-incrementing stamps or otherwise. A few institutions combine barcodes and accession identifier by forcing the embedded barcode value to match the existing accession identifier.</p> <p>It is recommended to pad the barcode number with leading zeroes to a length that accommodates the number of sheets within the collection and include these zeroes in human readable text below barcode.</p>	<p>See: iDigBio Specimen Barcode and Labeling Guide: https://www.idigbio.org/wiki/index.php/Specimen_Barcode_and_Labeling_Guide.</p> <p>McNeill (2011). International Code of Nomenclature for algae, fungi, and plants (ICN). http://www.iapt-taxon.org/nomen/main.php?page=title. See especially the ICN Glossary of terms used in this code for the definition of “specimen”: http://www.iapt-taxon.org/nomen/main.php?page=glo.</p>

		<p>Use archival-quality barcodes.</p> <p>Some database systems generate barcodes. For example, Specify uses its iReports feature for this and can generate barcodes in more than 20 styles. See http://specifyx.specifysoftware.org/wp-content/static/Gallery%20of%20Printed%20Label%20and%20Report%20Formats.pdf.</p> <p>See Module 6, T5 for further detail regarding barcodes and their uses.</p>	
T4	Associate/insert machine-readable barcodes/documents with/into folders.	<p>Some institutions create machine-readable documents to gather data at the cabinet and/or folder level. Documents might contain such information as family, higher geography, and current identification (“filed-as name”). These data will be read and associated with individual collection records in Imaging or Data Capture modules.</p> <p>Tasks T1 or T4 might also include determining whether specimens are out on loan or otherwise removed from cabinets or folders. This might be done either by physical tracking (e.g., inserting markers within the cabinets) or by electronic database loan tracking. Since such specimens might be removed for an extended period, an efficient strategy for tracking their return and ensuring their reinsertion into the digitization workflow should be designed and documented. It should be noted that an increasing number of institutions have alleviated the need to track loan re-insertions by digitizing specimens prior to fulfilling loan requests and recording annotations as loans are returned, prior to re-insertion.</p>	<p>QR codes, DataMatrix codes, 1D barcodes, or OCR-readable documents for insertion into specimen folders.</p> <p>See: Diazgranados and Funk (2013). http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3819127/</p>
T5	Pull specimen	Further decisions and filtering might be	Digitization policy manual

	folders from cabinet.	applied here if policy or project management plan stipulations limit the number of specimens to be digitized per taxonomic, geographic, or other unit.	or project management plan.
T6	Iteratively curate collection in place, including nomenclatural standardization and annotations.	<p>This task might be accomplished before, after, or during the digitization process. This task may include</p> <ul style="list-style-type: none"> • determining if the specimen label is consistent with the folder “filed-as” name, • updating the nomenclature to a standard source (e.g., Flora of North America), • identifying and annotating the specimen, including routing the specimen to a taxonomist or other competent personnel, • deaccessioning specimens. <p>It should be noted that specimens routed to a taxonomist might be removed from the active digitization workflow for an extended period to allow identification/annotation activities to be completed, and may necessitate the design of an efficient re-insertion strategy for returning these specimens to the workflow.</p> <p>This curation step might potentially encompass all manner of curatorial activities including: adding missing accession stamps; rehousing specimens from damaged folders, boxes, etc.; re-filing incorrectly housed specimens; re-organization of herbarium to updated filing system.</p>	Workspace. Taxonomist.
T7	Remove specimen from collection and bring to digitization station area.	Usually involves several to numerous folders, depending upon storage space at imaging station. Can use carts, a cabinet on a dolly, or a dedicated cabinet adjacent to imaging station.	Cart, mobile cabinet, or dedicated cabinet.
T8	Document/flag location of	This occurs at the cabinet level.	Flagging accessories. Collection filing policy.

	removed specimens to ensure accurate return of the specimen to the appropriate storage location.		
T9	Separate specimens that need barcodes/imaging from those already barcoded or imaged.	Some institutions have previously barcoded and/or imaged specimens from prior projects.	
T10	Establish protocol for repairing and rerouting specimens in the digitization process.	For example, a protocol might establish the priority of imaging the specimen as is (arrange on sheet) and image prior to routing for repair. Otherwise, consider beforehand how repaired specimens will be re-inserted into the imaging process and refiled after imaging. This activity could also take place in T6 (Iterative Curation).	Repair tools and accessories. Conservation protocols/policy.
T11	Apply barcode with accession/catalog/barcode number to individual collection objects (if not already barcoded).	Depending on protocol, application of barcodes is sometimes accomplished earlier as an independent iterative process. However, it is recommended that barcodes be used to denote an imaged specimen.	Barcodes.
T12	Create skeletal database record.	If included, this task may be accomplished by scanning the barcode value into a new database record and optionally populating the record with a select set of skeletal data that might include any or all of the following: taxon name (usually the “filed as” or “filed under” name), collection date, collector name, collector number, and a high-level geographic description (e.g., country and/or state), the latter to aid in prioritizing specimens for data entry or further digitization activities.	Barcode scanner, open database.

Literature Cited

Diazgranados, M., and V. A. Funk. 2013. Utility of QR codes in biological collections. *PhytoKeys* 25: 21–34. doi:10.3897/phytokeys.25.5175.
<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3819127/>.

iDigBio. 2015. Specimen Barcode and Labeling Guide.
https://www.idigbio.org/wiki/index.php/Specimen_Barcode_and_Labeling_Guide. Accessed 1 May 2015.

McNeill, J., F. R. Barrie, W. R. Buck, V. Demoulin, W. Greuter, D. L. Hawksworth, P. S. Herendeen, S. Knapp, K. Marhold, J. Prado, W. F. Prud'homme van Reine, G. F. Smith, J. H. Wiersema, and N. J. Turland. 2011. International Code of Nomenclature for algae, fungi, and plants (Melbourne Code). International Association for Plant Taxonomy. <http://www.iapt-taxon.org/nomen/main.php?page=title>.