Workflow Detail: Imaging (flat sheets and packets)

Module 1 Pre-digitization curation	Module 2 Imaging station setup/camera	Module 3 Imaging station setup/scanner	**Module 4 Image capture**

	Module 5 Image processing	Module 6 Data capture	Module 7 Data enrichment
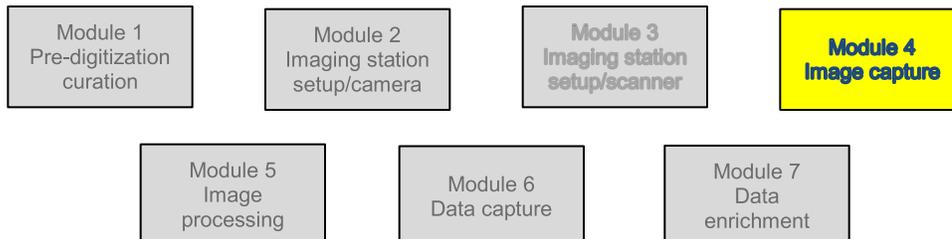
Module 4: Image Capture Task List

Task ID	Task Description	Explanations and Comments	Resource(s)
T1	Capture machine readable information (from 1D barcodes, DataMatrix, or QRcodes) associated with folders (see T2 in Module 1), if present.	The information gathered here is folder- or cabinet-level data to be associated with specimen-level records (see T5, this module).This step accommodates the workflow of the New England Herbaria TCN (NE-TCN), which captures cabinet- and folder-level data as part of pre-digitization curation, Module 1.	QRcode, DataMatrix, or barcode scanner.
T2	Remove specimen from folder.	Specimens have been previously moved to the imaging station (Module 1, T7). Ensure maintenance of original folder order and specimen order within folders via reverse stacking or some other institutionally specified consistent sequence as specimens are removed. Handle folders and specimens carefully. Do not turn folders or sheets upsidedown.	
T3	Stamp to indicate the specimen has been imaged. Some institutions stamp the folder, meaning its entire content, as imaged.	Desirable for indication of divergence of original image/record from physical specimen, which may have additions of future annotations, or insect damage. Completion of this task varies among institutions. If included, there are several methods by which it can be accomplished.	"IMAGED" stamp, "IMAGED + DATE" date-roller stamp, or printed slip to fit in imaging frame area.

| | | Strategy if re-imaging is planned (such as after future annotations):

• Ink-stamping the actual sheet with "IMAGED."
• Writing in pencil the imaged date "YYYY-MM-DD" immediately below this stamp.(The date can be erased and changed before re-imaging.)
• Alternately, some institutions write the date in permanent ink each time an image is recorded to ensure a record of imaging episodes for each specimen.

Strategy if re-imaging is not planned:
• Ink-stamping the actual sheet with "IMAGED YYYY-MM-DD" using a date-roller-type stamp.

Strategy if permanent indication of imaged date on actual sheet is not desired:
• Pre-printing a small slip of paper to fit onto the border of the imaging field for the scanner or camera (near color bar/grayscale) with "This image generated YYYY-MM-DD"

Additional option: include notation of imaging technique and resolution, e.g. "This image generated YYYY-MM-DD by a scanner at 600 PPI." | |
| **T4** | Apply specimen barcodes, if not already applied. | For some institutions, this task may have been completed previously, either as a step in pre-digitization curation (Module 1, T11), or as a separate barcode application workflow. The NE-TCN and others apply barcodes here.

It is preferable to apply barcodes near the label at the bottom of the sheet to facilitate future OCR or other data extraction technologies, | Barcodes. |

		though this practice is not followed by all institutions.	
T5	Scan specimen barcode into database.	Including task T4 in the workflow may trigger T5 to occur. Task T14 provides an alternate approach to capturing the barcode value into a database.	
T6	Create skeleton record via one of several entry techniques, e.g. keystroke entry, pick lists, voice recognition.	Skeletal record must contain: • Barcode value. Skeletal record might also contain: • Associated cabinet- and folder-level data obtained during Module 1, T1. • Critical data items transcribed from sheet (e.g., collector name, collector number, date collected, broad geographic description such as state or country) entered via keyboard or voice recognition sotware. Use controlled vocabularies and pick-lists when applicable. Another alternative, as that used by the Lichen/Bryophyte project, collects most recent ID, country, state, and sometimes collector during imaging using a project-specific application. Other institutions do not enter data at this juncture, though some suggest that capturing values for family and genus, followed immediately by creating a batch of records in sequential barcode order for subsequent data entry might improve efficiency.	Microphone for voice recognition, computer and keyboard, open database. Speech recognition software.
T7	Place and align specimen in imaging frame, light box, light tent, copy stand, or scanner.		Lightbox, copy stand, frame, light tent, and/or scanner.
T8	Place or ensure placement of scale and color standard and	GPI requires visible scale and color standard and prefers that they are on	Scale, color standard.

iDigBio
Integrated Digitized Biocollections

Florida Museum of Natural History
University of Florida
Florida State University
352-273-1906

	make certain they are clean and visible.	the sheet. Some institutions opt to affix the scale and standard to framing outside the margin of the specimen but clearly visible to reduce manual manipulation steps and increase efficiency. Many institutions affix the scale and standard to the imaging frame or within the lightbox, also to reduce excessive and inefficient mechanical manipulation. An affixed (rather than unattached) scale and color standard is recommended.	
T9	• Focus (if not autofocused). • Adjust aperture as necessary. • Adjust plant parts obscuring label or barcode, or remove parts that are obscuring the specimen and place them in a fragment packet, if this is in accordance with institutional policy.	This may be a repetitive task if more than a single image is recorded per specimen. Focus may be manual or auto, depending upon camera selected and institutional preference. Autofocus ensures automatic adjustments to varying depths of field between bulky and flat specimens.	Fragment packet.
T10a	Release shutter to record entire specimen sheet.	It is important during this task not to physically touch and potentially shake the camera. A wireless or tethered remote shutter release (recommended) is often helpful here, or a mouse-operated shutter release that can be activated within the imaging software.	Wireless, tethered, or mouse-activated shutter release.
T10b	Scan complete specimen sheet.	This alternate imaging method is for those institutions using scanner technology, often in association with Global Plants Initiative (GPI). The complete protocol can be found in the JSTOR Plants Handbook. The steps in the scanning process are performed in place of T10–T15 and	

		include:	
		- preview, - check image, adjust if necessary, - use selection tool to drag an area around the herbarium sheet, - when the image is satisfactory, click AUTOFOCUS, - scan, - perform quality control: check for pixilation, blurriness, lines in the scan, green color in corners, and color separation along edges (See JSTOR Handbook for examples; check in-depth the first scan and selected scans at regular intervals thereafter, - set image format (TIFF), image compression (NONE), byte order (IBM PC) - save image, - resume at T16	
T11	Image fragment packet.	Completion of this task varies among institutions. If included, there are several methods by which it can be accomplished. One strategy includes; - opening the packet and spreading the enclosure, - ensuring that the expanded packet tabs do not obscure important plant material, - weighting the packet tabs to hold them down, - recording the image. Another strategy includes: - opening the packet, - removing the packet contents to a paper tray that is the same dimensions as the packet, - closing the packet, - placing the paper tray on top	

		of the packet (with weights, if necessary), • recording the image. Bryophyte and lichen packets may be free, or fastened in lots to herbarium sheets. Packets may be • opened during image capture to reveal the contents or labels on the inside of the packet tab, • kept closed with only an image of the label recorded, • or some combination of both. If images are recorded of open bryophyte/lichen packets, contamination of succeeding specimens should be avoided by cleaning the substrate between images, especially when the contents have been removed from the packet.	
T12	Image the specimen label for later OCR processing.	These optional label images might result in greater OCR accuracy than images of entire sheets. Assuming that the barcode has been placed near the label at the bottom of the specimen sheet, it will also be included within this image. Images intended for later OCR processing should contain about 20 megapixels, with a height in pixels of the lower case "x" not less than about 15 pixels.	
T13	Check image quality, including focus, exposure, and presence/visibility of barcode.	This task is one of several quality control checks.	
T14	Rename file.	As a practical matter, many institutions include the herbarium acronym as a prefix to the barcode value and rename the image file accordingly, as is the convention for	

		GPI, as XXX123456, where XXX is the herbarium acronym.	
		Some institutions utilize camera control functions that name files to barcode sequence specifications as the image is recorded, hence eliminating the need to rename files.	
		Other institutions utilize a file renaming application (BCR, Bardecodefiler) to rename the image file to match the barcode value. This can sometimes be an iterative step through execution of a batch operation to process numerous files.	
		Alternatives to file renaming include storing the camera-generated filename in a database record and associating that record with the specimen's barcode. This assumes that camera-produced names remain unique, even if new or multiple cameras are put into service.	
T15	Scan barcode into database record.	Some institutions create a database record here by scanning the barcode into the database, a step than can also precede image capture.	
		If task T4 has been completed, task T15 should be skipped.	
T16	Check for damage to specimen that might have occurred during imaging.	If damage occurred, route to the conservation sub-module. Upon re-insertion, continue with T17.	
T17	Return specimen to the collection/folder.	Ensure maintenance of original folder order and specimen order inside folder via reverse stacking or other strategy as noted in task T2.	
		Do not turn folders or sheets upside down.	
		Do not shake folders to even out sheets.	
		Some institutions keep specimens unfiled until images and data have been subjected to final quality control	

		procedures effected during the image processing module. This strategy is dependent upon space available, protocol, pace of operation, and quantity of images processed.	