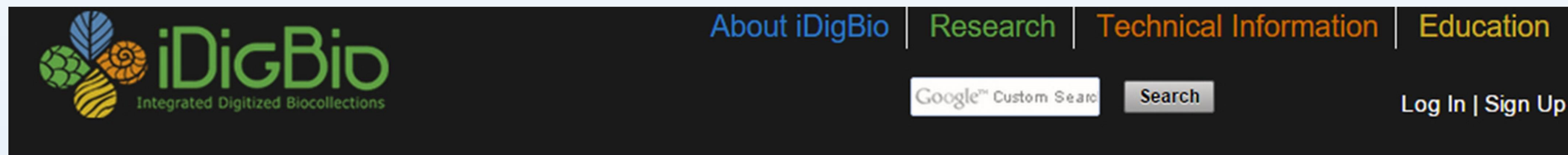


# DROID Workflow Working Groups



## Digitization Workflow Workshop Report

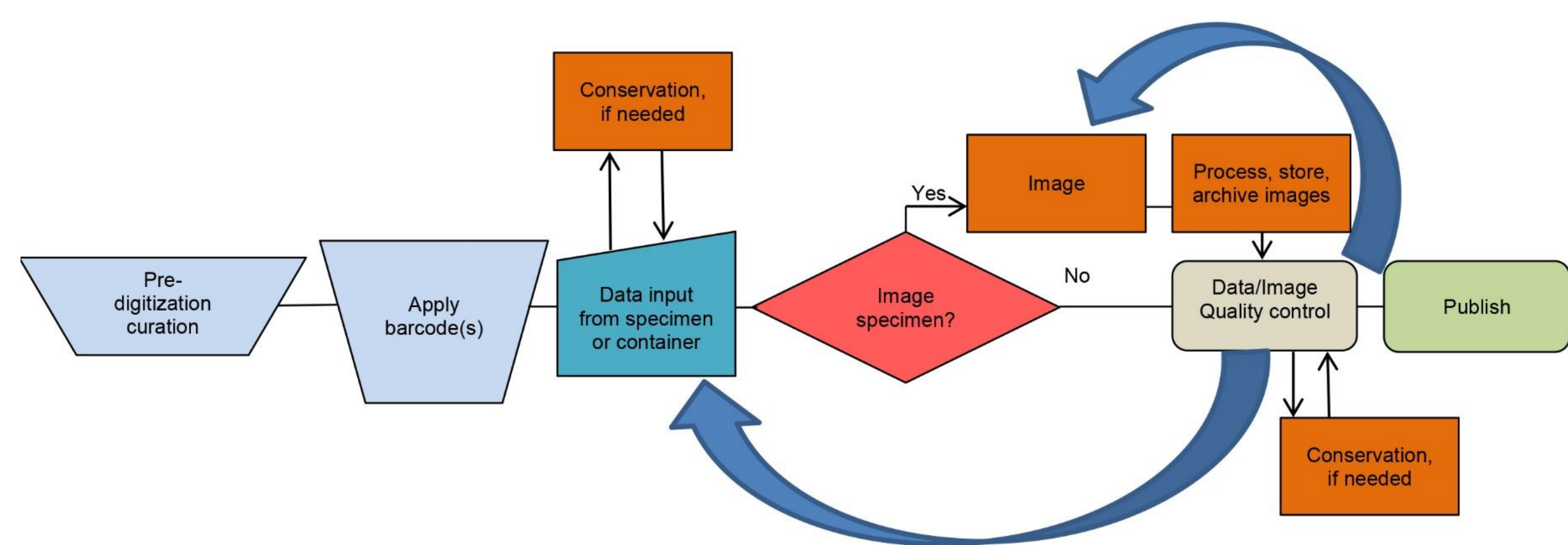


### Developing Robust Object-to-Image-to-Data (DROID) Workflow Workshop 30-31st May 2012, Florida Museum of Natural History, University of Florida (FLMNH)

During its first year of funding, iDigBio, in conjunction with the **Botanical Research Institute of Texas (BRIT)**, **Biodiversity Institute at the University of Kansas**, and **Yale Peabody Museum** co-sponsored a 2-day workshop in Gainesville, FL, focused on Developing Robust Object to Image to Data (DROID) workflows for digitizing biodiversity collection objects. The initial DROID workshop resulted in a series of workflow development working groups and workshops and several sets of preparation-specific workflows, including those for:

- Flat Sheets and Packets
- Pinned Things in Trays and Drawers
- Things in Spirits and Jars
- Three-dimensional Objects in Trays and Boxes

#### DROID Object to image to data workflow



Nelson et al. Applications in Plant Sciences 2015 3(9): 1500065-Data Supplement S7- Page 1

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

Appendix S7 Module 7: Image Processing

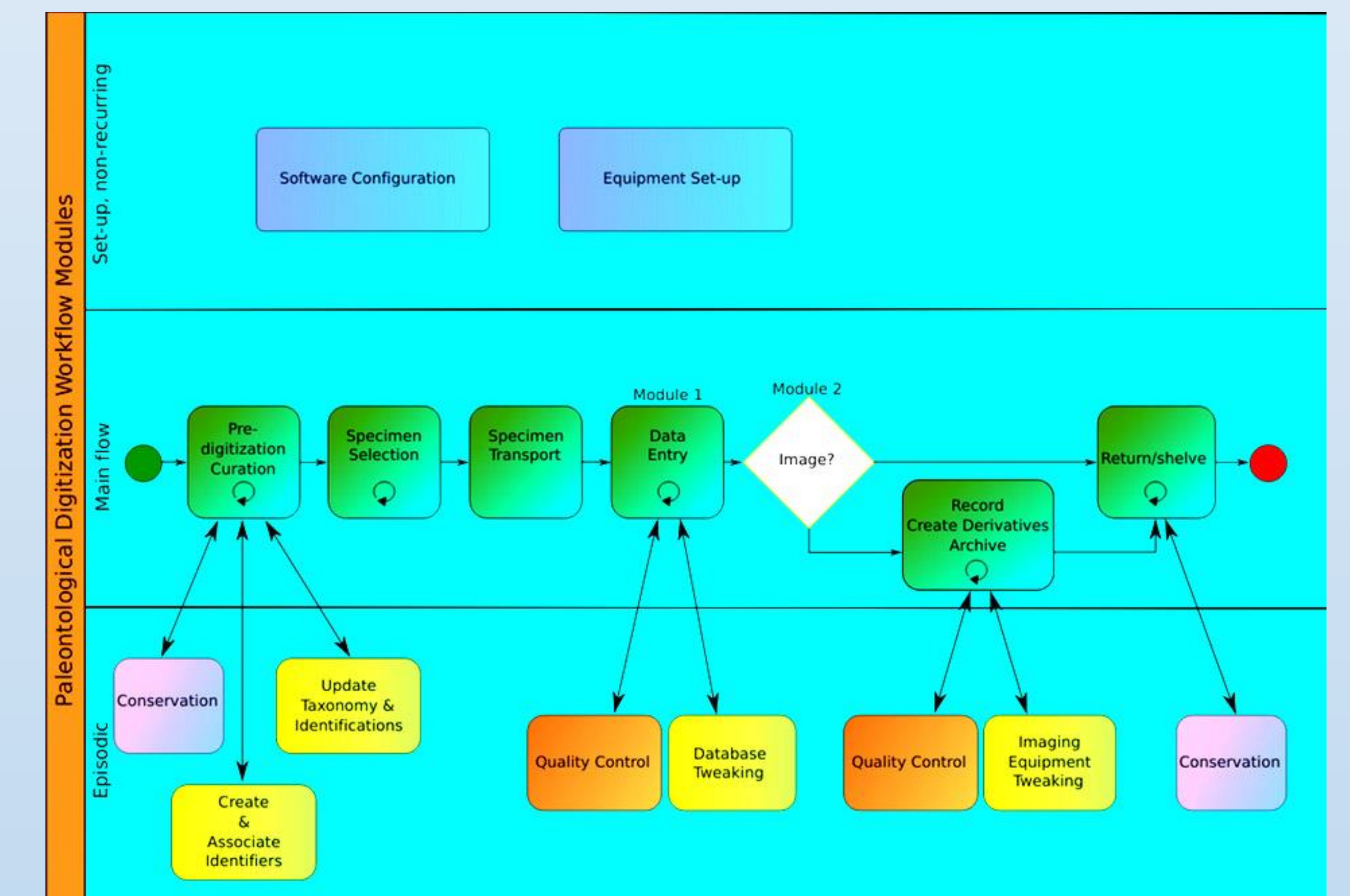
Task ID	Task Description	Explanations and Comments	Resource(s)
T1	Determine information flow and image archiving considerations.	Define sequence of folders/stages through which images and metadata will flow before imaging commences. Consider dividing images into batches for processing. Maintain logs of image processing steps and batches processed.  Consider generating a set of web-accessible derivative images (e.g., in JPEG format). A master image can be archived separately and does not necessarily need to have online or even near line accessibility (see Module 9: Image Archiving).  Even if images have been slated to be archived, a short-term backup plan is needed prior to their deposition in a digital preservation repository (see Data ONE: Best Practices for backing up your data).  When estimating storage requirements, consider archival master images, derivatives, and any replication that occurs due to backing up and digital preservation. Derivation of storage size needed should be made by recording an image with the intended camera, converting the image to PNG (or their adopted archival format) for archiving, creating a display-quality JPEG and any other derivatives, summing the sizes of the resulting files and multiplying by the anticipated number of images.	See: DataONE Best Practices for backing up data: <a href="https://www.dataone.org/tables/best-practices/back-up-your-data">https://www.dataone.org/tables/best-practices/back-up-your-data</a>
T2	Ensure file name quality control/assurance	Ensure image file names are correctly assigned based on barcode, accession number, etc. Document how the image filename and the physical specimen are	

All DROID workflows follow a modular format. Given the variety in institutional and collection infrastructure, ensuring flexibility and making workflow customization easy is essential. In the workflows presented here, adopters are encouraged to adapt as necessary by adjusting order of execution and tasks to be executed.

#### Business process modeling depiction Of paleo working group workflows

##### Paleo Digitization Working Group

Talia Karim (Chair)  
Roger Burkhalter  
Susan Butts  
Una Farrell  
Ann Molineux  
Gil Nelson  
Jessica Utrup



The first set of workflows to be developed following the DROID workshop focused on the digitization of specimens stored in flat sheets and packets, primarily in herbaria. In January 2015, approximately 30 herbarium professionals gathered at Valdosta State University to review and update the Flat Sheets and Packets workflows. Three days of onsite collaborative work led to the beginning of a paper that was refined over the following 4 months and accepted for publication by *Applications in Plant Sciences*, a journal of the Botanical Society of America. More recently, the **Paleo Digitization Working Group** completed a set of workflows for paleontological collections, depicted by the figure above and under revision for publication. The published versions of the paleo workflow documents will also be distributed through iDigBio and GitHub.



iDigBio is funded by a grant from the National Science Foundation's Advancing Digitization of Biodiversity Collections Program (Cooperative Agreement EF-1115210). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

