

**Wet Collections Digitization Workshop**  
**March 4–6, 2013**  
**University of Kansas**  
**Lawrence**

This workshop is the second in a series of preparation-specific workshops focusing on organizing, launching, maintaining, and enhancing a biological collections digitization program. The primary goal is to prepare participants with the necessary skills and knowledge to launch or maintain a digitization program individually, through collaboration with an existing Thematic Collections Network (TCN) or Partner with an Existing Network (PEN), or through collaboration with iDigBio. The target audience includes wet collections managers, curators, directors, and digitization specialists.

Workshop content includes: 1) an overview of iDigBio, ADBC, and nationwide digitization activities, 2) how to organize and launch a digitization program, 3) a detailed overview of digitization tasks, 4) use and configuration of imaging systems and lighting, 5) an introduction to and overview of several database management systems and related tools, 6) demonstration of several imaging software packages, 7) a consideration of image file-naming strategies, 8) how to develop effective workflows and protocols, 9) strategies for evaluating hardware and software, 10) a detailed overview of procedures for image and data capture, including a detailed overview of image types, image manipulation, and image conversion, 11) strategies and considerations for georeferencing, including use of online collaborative tools, 12) assigning globally unique identifiers to records and objects, and 12) methods for moving digitized data to the internet via data aggregators, harvesters, and portals.

Workshop planners include: Gil Nelson (iDigBio), Andy Bentley (KU), Nelson Rios (Tulane), and Rob Robins (FLMNH), and Mark Sabaj-Perez (Academy of Natural Sciences)

Links to the workshop Wiki and workshop handouts and resources can be accessed through a link at: [https://www.idigbio.org/wiki/index.php/Digitization\\_Training\\_Workshops](https://www.idigbio.org/wiki/index.php/Digitization_Training_Workshops)

The workshop's collaborative workshop notes page, which can be edited by all workshop participants, can be found at the following link and via the Wiki page, above:

5 March: <http://tinyurl.com/WetNotes1>

6 March: <http://tinyurl.com/WetNotes2>

**Agenda**

**Day 1 – March 4, 2013**

Time	Activity	Responsible
Various	Vans pick up at airport	Gil Nelson, Andy Bentley,

		<b>Cathy Bester</b>
<b>6:00–7:00 p.m.</b>	<b>Reception &amp; Welcome</b> <b>Venue: Museum</b> <ul style="list-style-type: none"> <li>• Panorama</li> <li>• Participant introductions</li> <li>• Heavy Hors d'oeuvre</li> </ul>	<b>Leonard Krishtalka</b> <b>Andy Bentley</b> <b>Gil Nelson</b>
<b>7:00–9:00 p.m.</b>	<b>Dinner (on your own)</b>	

## Day 2 – March 5, 2013

<b>Time</b>	<b>Activity/Venue</b>	<b>Responsible</b>
<b>6:00–8:00 a.m.</b>	<b>Breakfast</b>	<b>Oread</b>
<b>8:15–8:30 a.m.</b>	<b>Commons Orientation</b> <ul style="list-style-type: none"> <li>• Workshop logistics</li> <li>• Introduction to iDigBio</li> <li>• overview of existing TCNs</li> </ul>	<b>Gil Nelson</b>
<b>8:30–9:45 a.m.</b>	<b>Getting Started:</b> <ul style="list-style-type: none"> <li>• Why digitize,</li> <li>• Startup costs and considerations,</li> <li>• Importance of institutional and administrative support,</li> <li>• IT support,</li> <li>• Anticipated and unanticipated benefits,</li> <li>• Anticipated and unanticipated obstacles,</li> <li>• Leveraging institutional resources,</li> <li>• Global decisions for launching and sustaining a digitization program,</li> <li>• Digitization task clusters</li> <li>• Image and data</li> <li>• Personnel recruitment,</li> <li>• Extent of additions to professional workload,</li> <li>• Leveraging the library and information sciences</li> </ul>	<b>Gil Nelson (30 min)</b> <b>Chris Johnson (15 min)</b> <b>Sally Bjork (15 min)</b> <b>Sandra Brantley (15 min)</b>
<b>9:45–10:00 a.m.</b>	<b>Coffee Break</b>	
<b>10:00 a.m.–12:30 p.m.</b>	<b>Venue: Commons Imaging</b> <ul style="list-style-type: none"> <li>• Scanning specimen-related 35mm slides,</li> <li>• InvertNet Imaging Station</li> <li>• Software tools: Digital Photo Pro, Camera Control Pro, Nikon Capture NX2, Nikon View NX2,</li> </ul>	<b>Mark Sabaj-Perez</b> <b>Kyle Luckenbill</b> <b>Chris Taylor</b> <b>Laura Monahan</b>

	<ul style="list-style-type: none"> <li>etc.,</li> <li>Lightroom, Photoshop, Aperture, Picasa Viewer, etc.</li> <li>Image storage formats (camera raw, jpg, tif, dng, etc.).</li> <li>Available capture devices (cameras, etc.) for various collections.</li> <li>Lighting.</li> <li>Image naming strategies.</li> <li>Image capture standards and common practices.</li> <li>Image processing dos and don'ts (color balance, cropping, saturation, contrast adjustment, etc.).</li> <li>Ethics of image enhancement.</li> <li>Stacking: Helicon Focus, AutoMontage, free software.</li> <li>Camera setup and use</li> <li>Archiving images (best formats),</li> <li>Simultaneous capture of label and specimen in a single image.</li> </ul> <p><b>Hardware</b></p> <ul style="list-style-type: none"> <li>Overview of what's available,</li> <li>Cameras &amp; lenses</li> <li>Photo Tanks ("squeeze boxes"), Copy stands, light boxes, diffusers, and related equipment,</li> <li>Lighting,</li> <li>Microscope and lenses,</li> <li>Computer equipment</li> </ul>	
<b>12:30–1:30 p.m.</b>	<b>Lunch</b>	
<b>1:30–3:15 p.m.</b>	<b>Museum Tours and Demonstrations</b>	<b>Andy Bentley</b>
<b>3:15–3:30 p.m.</b>	<b>Coffee Break</b>	
<b>3:30–4:30 p.m.</b>	<b>Tours and Demonstrations continued</b>	<b>Andy Bentley</b>
<b>4:30–5:30 p.m.</b>	<p><b>Workflows and Protocols</b></p> <ul style="list-style-type: none"> <li>Elements of workflow design</li> <li>Importance of written protocols</li> <li>How to develop, write, and manage workflows</li> <li>Strategies for getting feedback on improvements from technicians</li> <li>Sample and suggested workflows</li> <li>Image to data vs data from specimens</li> <li>Workflow and data entry efficiencies</li> </ul>	<p><b>Gil Nelson</b> <b>Sally Bjork</b></p>

6:00–8:00 p.m.	Dinner	On your own
----------------	--------	-------------

### Day 3 – March 6, 2013

Time	Activity/Venue	Responsible
6:00–8:00 a.m.	Breakfast	Oread
8:15–9:00 a.m.	<b>Digitization Workflows and Protocols: Practical Applications</b> UT & FoTX workflows Oregon State workflows Digitizing Field Books	Gil Nelson Andrew Short (15 min) Jessica Rains (15 min) Brian Sidlauskas (15 min)
9:00–10:00 a.m.	<b>Identifying and moving digitized objects to the web via aggregators and harvesters</b> <ul style="list-style-type: none"> <li>• GUIDs</li> <li>• Darwin Core</li> <li>• Collaborating with TCNs and PENs</li> <li>• Contributing to iDigBio</li> <li>• Data Aggregators</li> <li>• Guidelines for selecting a database</li> </ul>	Deb Paul
10:00–10:15 a.m.	<b>Data cleaning and validation</b> <ul style="list-style-type: none"> <li>• Data cleaning</li> <li>• Google Refine</li> </ul>	Deb Paul
10:15–10:30 a.m.	Coffee Break	
10:30–11:45 a.m.	<b>Database Overviews:</b> <ul style="list-style-type: none"> <li>• Guidelines for evaluating collections databases</li> <li>• Specify, including workbench for new and legacy data</li> <li>• Migrating Excel to Specify</li> </ul>	Andy Bentley
11:45–12:45 p.m.	<b>Database overviews: Symbiota</b> <ul style="list-style-type: none"> <li>• Migrating Excel and other flat data stores into Symbiota</li> <li>• OCR advances and limitations</li> <li>• Data entry from ilabel mages</li> </ul>	Ed Gilbert
12:45–1:45 p.m.	Lunch	
1:45–2:45 p.m.	<b>Georeferencing</b> <ul style="list-style-type: none"> <li>• Geolocate</li> <li>• Google Maps</li> </ul>	Nelson Rios
2:45–3:00 p.m.	Coffee Break	

<b>3:00–3:30 p.m.</b>	<b>Pre-digitization curation and strategies for simultaneous collection digitization and management</b>	<b>Rob Robins</b>
<b>3:30–4:30 p.m.</b>	<b>Discussion and Q&amp;A</b>	<b>Discipline specific discussion groups</b>
<b>4:30–5:00 p.m.</b>	<b>Wrap up</b>	<b>Gil Nelson</b>
	<b>Dinner (on your own)</b>	