



Digitizing from Source Materials Workshop

10-12 March 2014

Yale Peabody Museum

iDigBio and Yale Peabody Museum are pleased to sponsor this workshop focused on the digitization of original, specimen-related source materials in natural history collections, to include field notebooks, catalogs, ledgers, cards, and other archival materials that contain specimen-related data. Source documentation consists of the original records of scientific research and discovery. In disciplines related to biodiversity research, original source documentation describes specimen collecting activities and, in addition to the critical taxon, locality and event data, they often include descriptions of habitats, species abundance, co-occurrence, meteorological data and, sometimes cultural and political information.

The goals of the workshop are to:

- Review existing projects that incorporate original source documentation content into digitization workflows,
- Assess the design of workflows, including the evaluation of cost/benefit and the identification of required resources,
- Document and evaluate the outcomes (acquisition rates, digital surrogates, shared efficiencies) of using field book content in digitization workflows,
- Prepare an outline for a best practices document that will be completed by a writing team agreed upon by the workshop attendees and distributed through the iDigBio portal and various professional societies,
- Prepare a comprehensive list of use cases that describe the value of original source documentation to TCN collections digitization processes, biodiversity studies, and systematic and taxonomic research efforts, and
- Collaborate with the Society for the Preservation of Natural History Collections (SPNHC) on a workshop for the 2015 annual meeting at the Florida Museum of Natural History in Gainesville, Florida.

Invitees include:

- Individuals with existing knowledge of source materials digitization and content,
- Representatives from the several biological disciplines and collection types,
- Representatives from funded TCNs, and
- Individuals from the library sciences and other Council for Library & Information Resources (CLIR) Hidden Collections funded projects.

Pre-Meeting of workshop planning team, Sunday, 9 March 2014.

Day 1 Monday, 10 March 2014

Time	Activity/Venue	Responsible
7:45 a.m.	Transportation departs Omni for West Campus	Tim White Gil Nelson
8:30-8:45 a.m.	Welcome Twitter #	Tim White Rusty Russell
8:45-9:00 a.m.	Introduction to iDigBio	Gil Nelson
9:00-9:20 a.m.	Workshop goals, overview, logistics	Tim White Rusty Russell
9:20-9:35 a.m.	Defining source materials: Creating a reference point Synopsis of the survey and homework results. What are we digitizing? What are the criteria for primary source materials?	Rusty Russell Michelle Koo Margaret Landis, scribe
9:35-10:00 a.m.	Group Discussion Discuss survey and homework data, and what else is missing from list of source materials currently being digitized? (perhaps go around the table to the various collections and institutions represented)	Tim White, lead Dan Drew, co-lead Margaret Landis, scribe
10:00-10:30 a.m.	Coffee break	
10:30-11:00 a.m.	Whole group work session: Developing a shared definition for what is being digitized. By end of this session, create a consensus definition, to be recorded on a publicly accessible Google Doc.	Rusty Russell, lead Margaret Landis, scribe
11:00-11:10	Why digitize source materials 1. identify & reduce to text a shared vision 2. things that make our daily lives easier 3. protect material being digitized.	Rusty Russell, lead Margaret Landis, scribe
11:10-11:30 a.m.	Whole group work session: Why digitize source materials?	Rusty Russell, lead Michelle Koo, co-lead Margaret Landis, scribe
11:30 a.m.-Noon	Whole group activity Recording a manifesto or shared	Rusty Russell, lead Michelle Koo, co-lead Margaret Landis, scribe

	vision of why source materials digitization is important. Participants and designated scribe contribute to a publicly accessible outline recorded in a Google Doc.	
Noon-1:00 p.m.	Lunch (Catered)	
1:00-1:45 p.m.	Delineate 5-7 categories of commonalities across workflows for types of source materials Whole group activity staging for breakouts to follow on Tuesday morning. Product expected: 5-7 annotated categories of commonalities recorded in a Google Doc.	Gil Nelson, lead Deb Paul, co-lead Christina Fidler, scribe
1:45-2:45 p.m.	Breakout sessions organized around categories delineated in previous session	Gil Nelson, lead Deb Paul, co-lead Christina Fidler, scribe
2:45-3:30 p.m.	Reports from category discussions in previous session	Gil Nelson, lead Deb Paul, co-lead Christina Fidler, scribe
3:30 p.m.	Depart for main campus for collection tours	
5:30-8:00 p.m.	Reception Yale Peabody Museum	Tim White Gil Nelson

Day 2 Tuesday, 11 March 2014

Time	Activity/Venue	Responsible
7:45 a.m.	Bus departs Omni	
8:30 a.m.	Organization of the day	Michelle Koo, lead
8:30-10:00 a.m.	Participant divide into 5+ groups to flesh out small group topics from Day 1: (e.g., what to digitize, why digitize, workflows, workflows, workflows!) Group 1: Define what we are digitizing Group 2: why digitize	Michelle Koo, lead Appoint 1 scribe per group

	components of workflow: Group 3: conservation Group 4: materials discovery Group 5: prioritization Group 6: metadata and standards Group 7: imaging/ data extraction Product expected: Minimally an annotated outline and rough draft of a document from Day 1 discussions by group.	
10:00-10:30 a.m.	Coffee break	
10:30-11:30 a.m.	Reports from working groups and further discussion	James Macklin, lead (volunteer), scribe
11:30-12:30 p.m.	Lightning round: The nexus of biodiversity, library and archival standards. 1. Overview 2. Audubon Core and IDs 3. Library standards, TaxPub 4. Archival standards 5. Dissemination standards for sharing/accessing data and their tools	James Macklin, lead Rusty Russell, co-lead Deb Paul, co-lead (volunteer), scribe 1. James Macklin 2. Greg Riccardi 3. Terry Catapano 4. Christina Fidler 5. Larry Gall
12:30-1:30 p.m.	Lunch (catered)	
1:30-2:30 p.m.	Discussion of standards addressed in previous session	James Macklin, lead (volunteer), scribe
2:30-2:45 p.m.	Coffee break	
2:45-3:45 p.m.	Strategies for justifying source material digitization Short talks, lightning rounds, followed by discussion, scribing, and writing 1. Paleontology 2. Botany 3. Zoology 4. Negative data	Tim White, lead Susan Butts, co-lead (volunteer), scribe 1. Pat Holroyd 2. Richard Rabeler 3. Andy Bentley 4. Sarah Schmits
3:45-5:00 p.m.	Workflow discussion...continued	Michelle Koo, Lead

Day 3 Wednesday, 12 March 2014

Time	Activity/Venue	Responsible
8:30-8:45 a.m.	Orientation to day	Carolyn Sheffield, lead Sarah Schmits, scribe
8:45-10:45 a.m.	Discussion about public participation and citizen science projects and their place in source materials digitization Products: inventory of tools available, future developments	Carolyn Sheffield, lead Libby Ellwood, co-lead Michelle Koo, co-lead Sarah Schmits, scribe Speakers 1. Rob Guralnick 2. Andrea Thomer 3. Ben Brumfield 4. Christina Fidler
10:45-11:15 a.m.	Coffee break	
11:15-12:30 p.m.	West Campus Tours (Workshop organizers plan the rest of the day)	Tim White, lead
12:30-1:30 p.m.	Lunch (Catered)	
1:30 a.m.-3:00 p.m. (three 20-minute talks followed by group discussion)	The “F word” - Funding: How to sell the justification Internal and external funding 1. Barbara Thiers 2. Rusty Russell 3. Patricia Pedersen What funding have you utilized for this? What are (untapped) potential sources of funding? Do you write collaborative funding proposals? How do we sustain the digital resource? Once legacy data are done, is money built into future funding initiatives to include born-digital documents that accompany specimens?	Deb Paul, lead Tim White, co-lead Paul Mayer, scribe 1. Barbara Thiers 2. Rusty Russell 3. Patricia Pedersen
3:00-5:00 p.m.	Organize to produce a cohesive document from results of the workshop	Rusty Russell