Project Management and Data Mobilization

David Jennings, iDigBio Project Manager
Cat Chapman, iDigBio Biodiversity Informatics Coordinator

November 3, 2021
iDigBio Orientation Webinar Series

October 6 - Welcome to iDigBio Phase 3

October 20 - Ongoing iDigBio collaborations and events

November 3 - Project management and data mobilization considerations for digitization networks

November 17 - Resources for digitizing biodiversity collections and managing the digitized data

December 1 - Increasing diversity and inclusion in biodiversity collections

December 15 - Accessing and using digitized data from biodiversity collections

https://www.idigbio.org/content/idigbio-orientation-2021
Today’s agenda

• This orientation session doubles as our regular quarterly TCN meeting

• We are discussing timely topics relevant to all TCNs, PENs, and other digitization projects:
  – Digitization project management
  – TCN responsibilities and reporting
  – Importance of your Data Management Plan
  – How to get your data to iDigBio

• Questions & Answers
Digitization Project Management

David Jennings
iDigBio Project Manager
djennings@flmnh.ufl.edu
Project management is important for success!

- We generally present something at the annual orientation
- We presented a webinar in our *Adapting to COVID-19* series that specifically focused on project management:
  - Highly recommended viewing!
  - Includes a broader overview of project management (presented by me)
  - Includes a series of **tips and tricks** on managing projects (presented by me, Diego Barroso, and Jen Zaspel)
Why do we keep bringing up project management?

• Common challenges with scientific and research projects:
  – Scientists want to do research, not administration
  – Project management and administration can be time consuming
  – Administration distracts scientists from their main interests
  – Scientists often lack training or experience in project management
  – Funding agencies desire accurate cost estimates and predictable outcomes

• There must be a better way:
  – Educate emerging scientists in the basics of project management and administration
  – Transfer knowledge and lessons from experienced scientists to those in emerging projects

Project management is not as abstract as you think…

“Plan-do-check-act (PDCA) represents an intersection between the scientific method and everyday operations. Given an objective, whatever process is defined to achieve that objective can be—and often unknowingly is—subject to PDCA.”

https://search.proquest.com/openview/1d159396d94f73d3d12b23ead675e43e/1?pq-origsite=gscholar&cbl=34671
Project Management is often intertwined with Leadership

- **Management** is about *coping with complexity* to bring order and predictability to a situation.
- **Leadership** is about *coping with change* to survive or compete effectively in a new environment.

<table>
<thead>
<tr>
<th>Management</th>
<th>Leadership</th>
</tr>
</thead>
<tbody>
<tr>
<td>Planning and budgeting</td>
<td>Setting direction</td>
</tr>
<tr>
<td>Organizing and Staffing</td>
<td>Aligning people</td>
</tr>
<tr>
<td>Control and problem-solving</td>
<td>Motivation</td>
</tr>
</tbody>
</table>

[https://hbr.org/2001/12/what-leaders-really-do](https://hbr.org/2001/12/what-leaders-really-do)
Project management starts with a **Workplan**

• A workplan serves as a map and guide for your team:
  – **What** needs to be done?
  – **Who** is doing the work?
  – How much will the work cost (**budget**)?
  – When will the work be done (**schedule**)?
  – How the work will be done (**workflows**)?
  – How will you manage **communications**?
  – How will you manage **risk**?
  – How will you manage **change**?
  – What **metrics** will you use to track and measure **progress**, **quality**, and **scope**?
  – What are your internal and external **dependencies**?

• Remember that a workplan is a living document → adaptable

• The planning process is vitally important for your team
A few tips and tricks from our collective experience...

- Manage your risks *early* and *often*
- Be flexible → you can act on *incomplete* data, but *insufficient* data does not compute
- Seek to build relationships → know who is who and what makes them unique
- Create a series of templates/forms for emails, reports, advertisements, etc.
- Use a centralized repository to facilitate access → documentation, training resources
- Create reports as you go (not all at the end)
- Prioritize and keep costs in mind
- Hold regular meetings
- Create a workplan (and a status dashboard)
- Keep your stakeholders in the loop → newsletter

TCN Responsibilities and Reporting

David Jennings
iDigBio Project Manager
djennings@flmnh.ufl.edu
TCN Responsibilities

1. Maintain your TCN wiki page
   - https://www.idigbio.org/wiki/index.php/TCNs
   - Have your Project Manager email us to get an account to allow editing

2. Submit requested info for the annual Summits

3. Provide feedback via community survey and other solicitations
TCN Responsibilities (cont.)

4. Prepare your annual report for NSF
   - Are you the lead institution or not? R.T.F.D.
   - You can find the above document linked on the TCN Resources page: https://www.idigbio.org/wiki/index.php/TCN_Resources
   - If in doubt, email Reed Beaman (rsbeaman@nsf.gov) or your fellow TCN PIs (TCNPIS-L@lists.ufl.edu)
TCN Responsibilities (cont.)

5. Participate in the quarterly meetings
   - First Wednesday of every February, May, August, and November @ 2:00 PM Eastern
   - Schedule including links to subscribe to the meeting events:
     - [https://www.idigbio.org/content/2021-internal-advisory-committee-meetings](https://www.idigbio.org/content/2021-internal-advisory-committee-meetings)
     - [https://www.idigbio.org/content/2022-internal-advisory-committee-meetings](https://www.idigbio.org/content/2022-internal-advisory-committee-meetings)
   - Minutes are posted on the wiki: [https://www.idigbio.org/wiki/index.php/Internal_Advisory_Committee](https://www.idigbio.org/wiki/index.php/Internal_Advisory_Committee)
TCN Responsibilities (cont.)

6. Submit quarterly reports to iDigBio
   – *Due by the quarterly meeting*
   – Instructions for completing and submitting your report are here: [https://www.idigbio.org/content/TCN-progress-reports-idigbio](https://www.idigbio.org/content/TCN-progress-reports-idigbio)
   – Compilations are posted on the wiki: [https://www.idigbio.org/wiki/index.php/Internal_Advisory_Committee](https://www.idigbio.org/wiki/index.php/Internal_Advisory_Committee)
Importance of Your Data Management Plan

Cat Chapman
iDigBio Biodiversity Informatics Coordinator
cchapman@floridamuseum.ufl.edu
What nobody wants to think about… 🧵🙈🙉🙊

• “What happens after the $$$ runs out?”
  – What will happen to your data?
  – Where will your data live?
  – Where will your documents live?
    • Protocols? Workflows?
• iDigBio can provide training and guidance
  – We cannot provide funding
• Your proposal has a data management plan
  – Implementation is your responsibility
  – But we CAN help you
Some ideas

- Internet Archive (archive.org)
- Your institution
- If all else fails... GitHub
  - Third-party private entities probably aren’t the best, but better than nothing

Any other ideas? 👀
**TL;DR**

- Data Management Plans = very important!
- Your DMP = your responsibility to fulfill!
- We CAN help you out!
  - Other, established TCNs are also a valuable knowledge resource
- Think of the long-term
- Data loss makes the biodiversity informatics world sad
- Data persistence makes the biodiversity informatics world happy
How to Get Your Data to iDigBio

Cat Chapman
iDigBio Biodiversity Informatics Coordinator
cchapman@floridamuseum.ufl.edu
What’s In This For You?

Meet the iDigBio Staff
Overview of the ingestion process

Learn how to get your data published
iDigBio Data Mobilization Staff

Caitlin “Cat” Chapman
cchapman@floridamuseum.ufl.edu

Biodiversity Informatics Coordinator

Dan Stoner
dstoner@acis.ufl.edu

Data Integration Expert
data@idigbio.org
The go-to guide for data ingestion

Everything you wanted to know about preparing data for ingestion:


– Identifiers
– Darwin Core – occurrence data (specimen records)
– Audubon Core - media
Ingestion Process

Index → Publish → Ingest
What do we mean by publishing data?

making biodiversity data publicly accessible & discoverable, in a standardized form, via a URL.

* that is reproducible and automated
Data publishing – where to begin

• Email data@idigbio.org “I’m ready”
• Review your data and publishing options together
Ingestion Queue


Milestones:
• Negotiating
• Mobilizing
• Evaluating
• Ingesting

iDigBio Recordset
**Negotiating**
- Starts with an email, call, conversation
- Ends with submitting data for preparation and/or inspection

**Mobilizing**
- Data made deliverable (e.g. DwC-A)
- Submitted to iDigBio for evaluation

**Evaluating**
- Data is evaluated to see that it meets Darwin Core standards
- Data is put into the ingestion queue and formally submitted for ingestion

**Ingesting**
- Data is ingested into the iDigBio Portal
- Ingestions are run periodically so that data is kept up-to-date
- Once data is submitted it cannot be withdrawn

Return data to provider for fixing if necessary

If there are errors with ingestion, evaluate
DATASET INFO: info about the provider (metadata)

Document your dataset metadata with your provider information:

• responsible parties (name, address, email, role)
• institution name, institution code, collection code, logo
• URL to the collection at your institution
• descriptive paragraph about the institution, collection, and the dataset
DATASET INFO: rights

• Use Creative Commons standards:
  
  – CC0 for data (not copyrightable)
  
  – CC BY for media (at least)
IDENTIFIERS

Every specimen and media record needs an identifier. [Robust and persistent]

- Think of a unique identifier as a Social Security Number for your catalogued object!
- If it changes, it needs a “paper trail”!

Example UUID:

urn:uuid:2d5d3a8f-7a18-4825-a129-4a32b4ae58b8
Remember, when you’re ready:

data@idigbio.org
Thank You!

www.idigbio.org

facebook.com/iDigBio

twitter.com/iDigBio

vimeo.com/idigbio

idigbio.org/rss-feed.xml

webcal://www.idigbio.org/events-calendar/export.ics

iDigBio is funded by grants from the National Science Foundation (DBI-1115210 (2011-2018), DBI-1547229 (2016-2022), & DBI-2027654 (2021-2026)). 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. © 2011-2021 iDigBio
Join us for the next episodes in the Orientation series...

October 6 - Welcome to iDigBio Phase 3

October 20 - Ongoing iDigBio collaborations and events

November 3 - Project management and data mobilization considerations for digitization networks

November 17 - Resources for digitizing biodiversity collections and managing the digitized data

December 1 - Increasing diversity and inclusion in biodiversity collections

December 15 - Accessing and using digitized data from biodiversity collections

https://www.idigbio.org/content/idigbio-orientation-2021
Orientation Webinar Series Survey:

https://ufl.qualtrics.com/jfe/form/SV_8dj1FtMT6Sp45T0