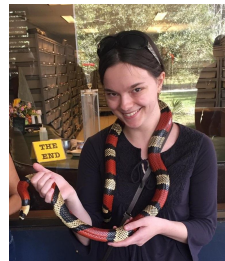




How To Get Your Data To iDigBio



Cat Chapman
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

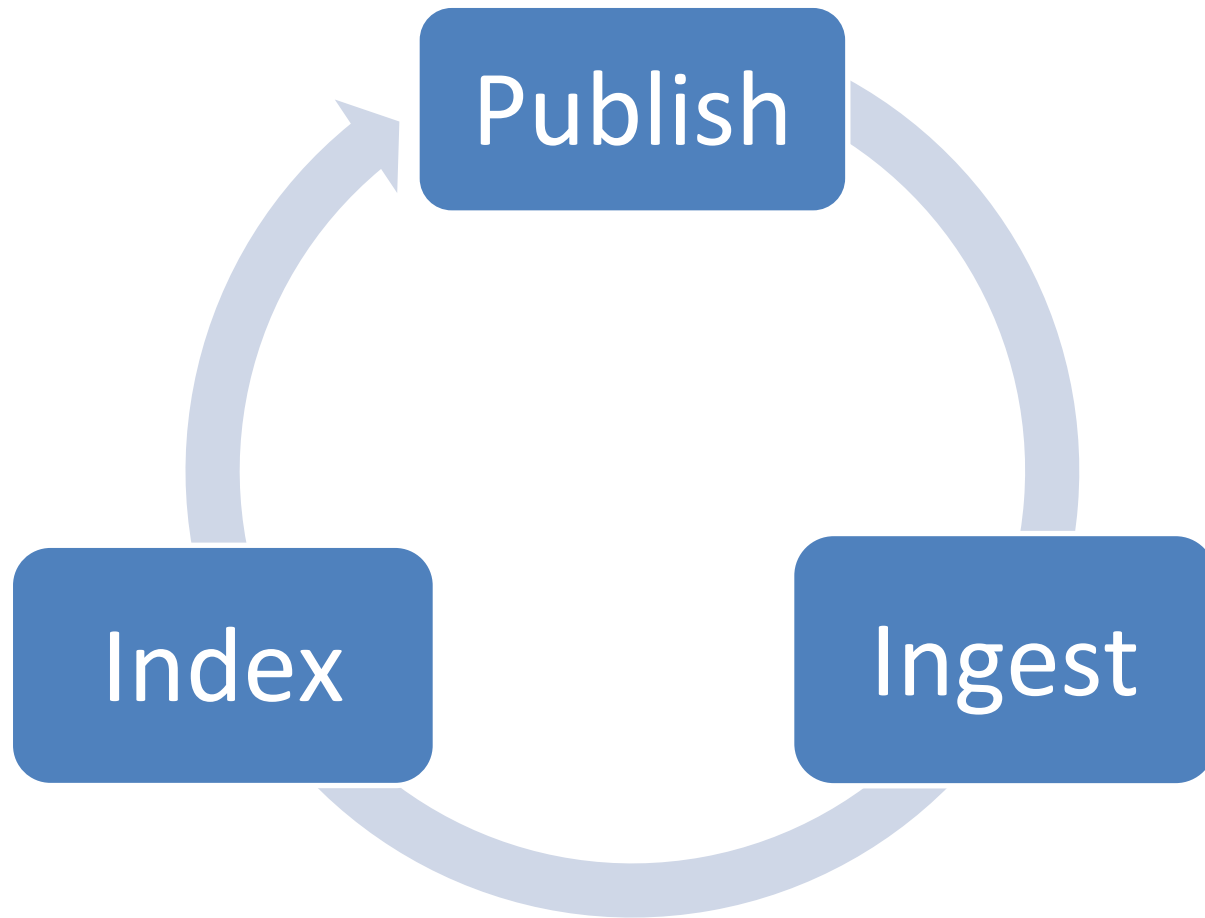
Everything you wanted to know about preparing data for ingestion:

https://www.idigbio.org/wiki/index.php/Data_Ingestion_Guidance

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



Ingestion Process

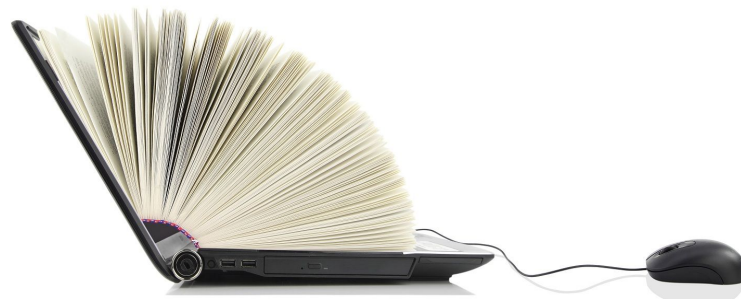




What do we mean by **publishing** data?

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

Consortium of Midwest Herbaria

Home >> Species Search >> Darwin Core Archive Publishing

Darwin Core Archive Publishing

The following downloads are occurrence data packages from collections that have chosen to publish their complete dataset as a Darwin Core Archive (DwC-A) File. A DwC-A File is a single compressed ZIP file that contains one to several data files along with a meta.xml document that describes the content. The archives below contain three comma separated (CSV) files containing occurrences, identifications (determinations), and image metadata. Fields within the occurrences.csv file are defined by the Darwin Core exchange standard. The identification and image files follow the DwC extensions for those data types.

Data Usage Policy:
Use of these datasets requires agreement with the terms and conditions in our Data Usage Policy. Locality details for rare, threatened, or sensitive records have been redacted from these data files. One must contact the collections directly to obtain access to sensitive locality data.

RSS Feeds: <http://midwestherbaria.org/portal/web/services/dwc/rss.xml>

Consortium of Midwest Herbaria DwC-Archive Files

| Code | Collection Name | DwC-Archive | Metadata | Pub Date |
|--------|-------------------------------------------------------|---------------|----------|------------|
| ALSC | Allen College | DwC-A (383) | EML | 2017-09-05 |
| CALVIN | Calvin College | DwC-A (83M) | EML | 2017-09-05 |
| CMC | Central Michigan University | DwC-A (200) | EML | 2017-08-31 |
| EMUC | Eastern Michigan University Herbarium | DwC-A (53M) | EML | 2017-09-05 |
| GVSC | Grand Valley State University | DwC-A (83M) | EML | 2017-09-05 |
| HSUO | Hillsdale College Herbarium | DwC-A (58M) | EML | 2017-09-05 |
| HDHM | Hopkins College | DwC-A (83M) | EML | 2017-09-05 |
| HUNT | Huntington University Herbarium | DwC-A (83M) | EML | 2017-09-05 |
| ILSB | Illinois Natural History Survey | DwC-A (911M) | EML | 2017-08-31 |
| IND | Indiana University Herbarium (Dorr Herbarium) | DwC-A (11M) | EML | 2017-09-05 |
| MIN | L. J. Sell Museum of Natural History Herbarium | DwC-A (14.1M) | EML | 2017-08-31 |
| MSU | Michigan University, Wilbur Smithman Tunnel Herbarium | DwC-A (33M) | EML | 2017-09-05 |
| MOR | Morton Arboretum | DwC-A (14.4M) | EML | 2017-09-05 |

GBIF INTEGRATED PUBLISHING TOOLKIT (IPT)
Free and open access to biodiversity data

Home About

Hosted resources available through this IPT

| Logo | Name | Organisation | Type | Subtype | Records | Last modified | Last publication | Next publication |
|------|---------------------------------------------------------------------------------|--------------|------------|----------|---------|---------------|------------------|------------------|
| | AMNH Invertebrate Paleontology Collection | AMNH | Occurrence | Specimen | 8,982 | 2017-01-09 | 2017-01-09 | -- |
| | BMSM Bailey-Mathews National Shell Museum | BMSM | Occurrence | Specimen | 122,166 | 2017-07-11 | 2017-07-11 | -- |
| | BRPC Black Rock Forest Consortium Herbarium | BRPC | Occurrence | Specimen | 623 | 2016-08-29 | 2016-08-29 | -- |
| | CMC Cincinnati Museum Center Invertebrate Paleontology | CMC | Occurrence | Specimen | 60,165 | 2016-11-12 | 2016-11-10 | -- |
| | DAV UC Davis Center for Plant Diversity | DAV | Occurrence | Specimen | 129,178 | 2016-11-21 | 2016-11-21 | -- |
| | FSUCML Zoological Collections | FSUCML | Occurrence | Specimen | 1,015 | 2017-07-20 | 2016-03-08 | -- |
| LSU | LSU Shirley C. Tucker Herbarium at Louisiana State University - Algae | LSU LSUM | Occurrence | Specimen | 59 | 2017-05-31 | 2017-05-31 | -- |
| LSU | LSU Shirley C. Tucker Herbarium at Louisiana State University - Bryophytes | LSU LSUM | Occurrence | Specimen | 5,199 | 2017-05-31 | 2017-05-31 | -- |
| LSU | LSU Shirley C. Tucker Herbarium at Louisiana State University - Lichens | LSU LSUM | Occurrence | Specimen | 31,190 | 2017-05-31 | 2017-05-31 | -- |
| LSU | LSU Shirley C. Tucker Herbarium at Louisiana State University - Vascular Plants | LSU LSUM | Occurrence | Specimen | 118,109 | 2017-05-31 | 2017-05-31 | -- |
| LSU | LSUM Bernard Lowy Mycological Herbarium at Louisiana State University - Fungi | LSU LSUM | Occurrence | Specimen | 18,107 | 2017-05-31 | 2017-05-31 | -- |

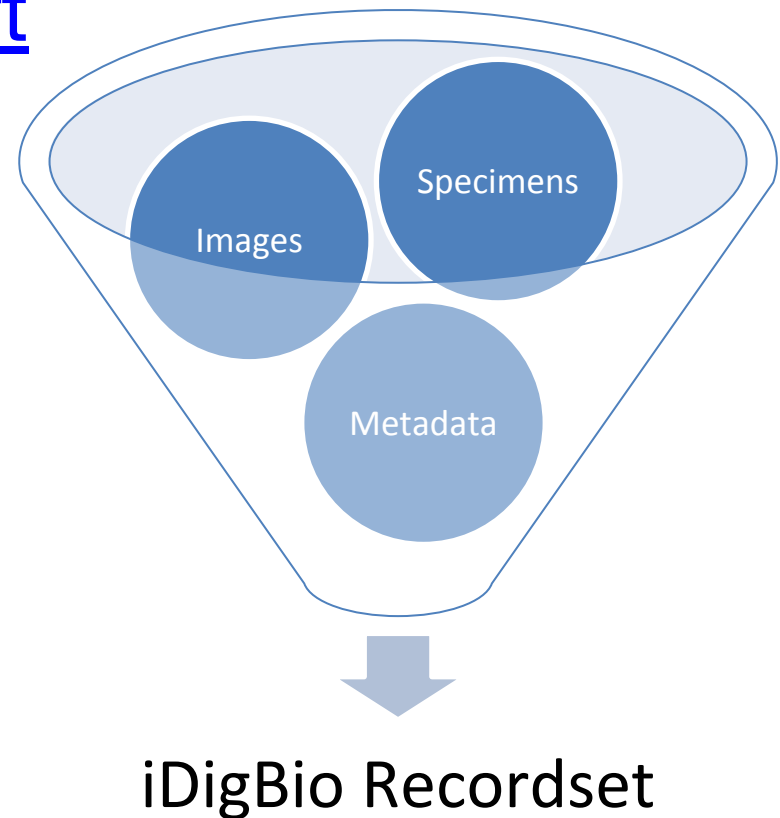


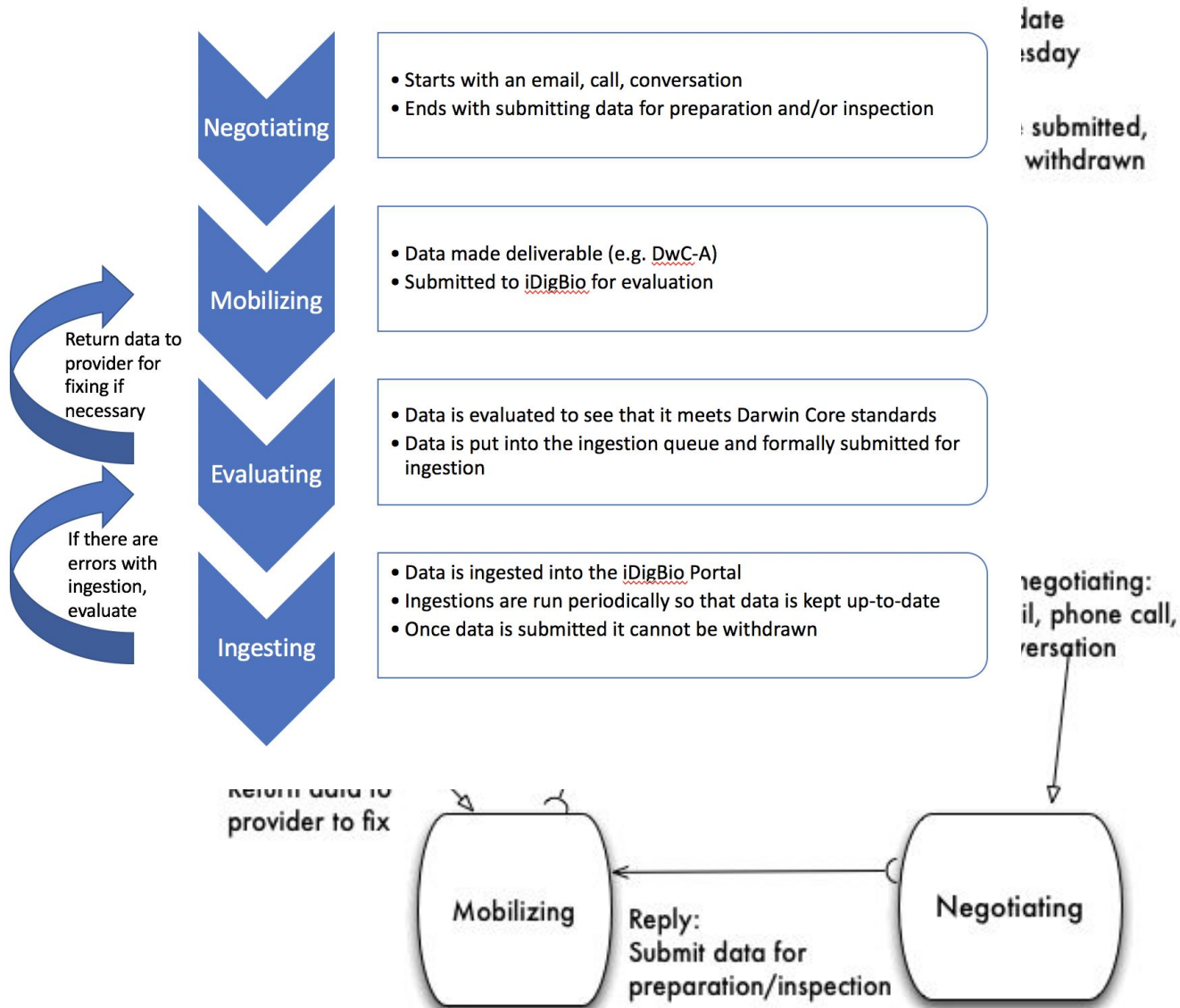
Ingestion Queue

- https://www.idigbio.org/wiki/index.php/Data_Ingestion_Report

Milestones:

- Negotiating
- Mobilizing
- Evaluating
- Ingesting







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]

We like UUIDs with a prefix:

urn:uuid:2d5d3a8f-7a18-4825-a129-4a32b4ae58b8



Remember, when you're ready:

data@idigbio.org