iDigBio
Integrated Digitized Biocollections

Exploring the Who What Where: aka Alphabet Soup

Quick overview of some acronyms, projects, software, etc. affiliated with natural history collections data use and digitization initiatives

Orientation Session, ADBC Summit VI
Deborah Paul and Shelley James
Tuesday 1 November 2016

UF UNIVERSITY of FLORIDA
FLORIDA STATE UNIVERSITY 1851
FLORIDA MUSEUM OF NATURAL HISTORY
NSF

iDigBio is funded by a grant from the National Science Foundation's Advancing Digitization of Biodiversity Collections Program. 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.

iDigBio
Integrated Digitized Biocollections

An acronym rich environment (ARE)

A framework for this walk-through

- ADBC, TCNs and PENs
- Digitization related acronyms
- Aggregators and beyond
- And just a few more…

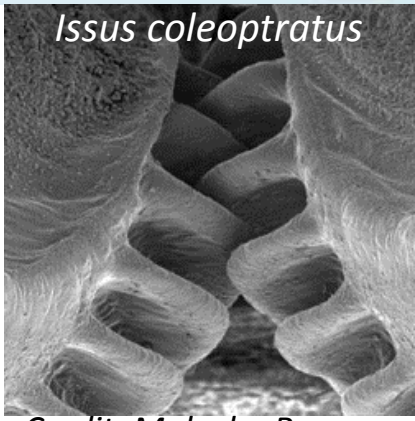
- Glossary in program
- Full glossary on the Wiki

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

From NIBA to the ADBC & BCoN

- Networked Integrated Biocollections Alliance (NIBA)
- Resulted in Advancing the Digitization of Biological Collections (ADBC)
- Biodiversity Collections Network (BCoN)
  - AIBS (American Institute for Biological Sciences), iDigBio & NSC Alliance (Natural Science Collections Alliance)

http://bcon.aibs.org/



Credit: Malcolm Burrows

Advancing the Digitization of Biological Collections (ADBC)

- iDigBio, that's Integrated Digitized Biocollections
  - Funded by National Science Foundation (NSF) to develop an integrated national infrastructure and resource for digitization of existing biodiversity collections in the U.S. iDigBio
- Thematic Collection Networks (TCNs)
  - A network of collaborative institutions
  - Digitize around a particular research theme.
  - Charged with defining and delineating subprojects for technology development or content generation; identifying deliverable goals, metrics for assessment, and specific needs for community support; providing technical support; and strengthening communications and outreach to other collections
- Partners to Existing Networks (PENs)
  - Must increase the number of specimens digitized in the ongoing project and fill gaps not addressed by the specimens in the existing TCN.

iDigBio – the Home Uniting Biocollections (HUB)

- Sustainability, digitization, facilitating research and educational use of collections data, cyberinfrastructure, …
- Advanced Computing and Information Science (ACIS)
  - Implementing the cyberinfrastructure for iDigBio
- Libraries of Life (a free app)
  - Collections coming to life in 3-D, classroom applications, collaboration with TCNs, ASU (Arizona State University)
- WeDigBio (Worldwide Engagement for Digitized Collections)
  - Planet-wide collaborative 4-day event to engage the public, including students, to transcribe collections data and meet scientists
- The FOSSIL Project
  - Engaging amateur fossil clubs and societies as downstream users of digitized data

Thematic Collections Networks (TCNs) and Partners to Existing Networks (PENs)

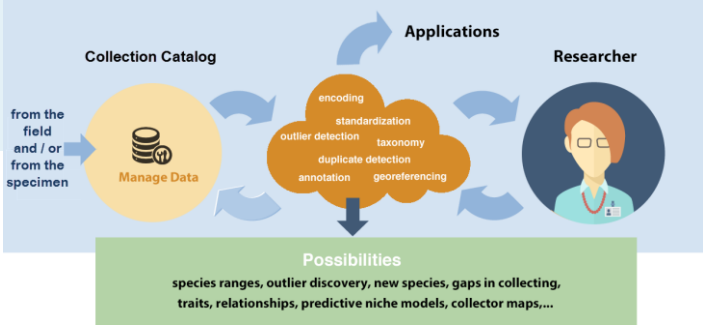
Cretaceous World	
Fossil Insect Collaborative	1
Fossil Marine Invertebrates (EPICC)	
Great Lakes Invasives	
InvertEBase	1
InvertNet	1
LepNet	
Lichens & Bryophytes	4
Macroalgal Consortium	
Macrofungi Consortium	2
Microfungi Consortium	
Mid-Atlantic Megalopolis (MAM)	
NEVP	2
PALEONICHES	2
SCAN	4
SERNEC	
Tri-Trophic	
Vouchered Animal Communication Signals	

TCN: network of institutions strategically digitizing information for a particular research theme, such as impacts of climate change or biota of a region.

The Cretaceous World

The Mid-Atlantic Megalopolis

iDigBio
Integrated Digitized Biocollections



Digitization

Collection Management Software

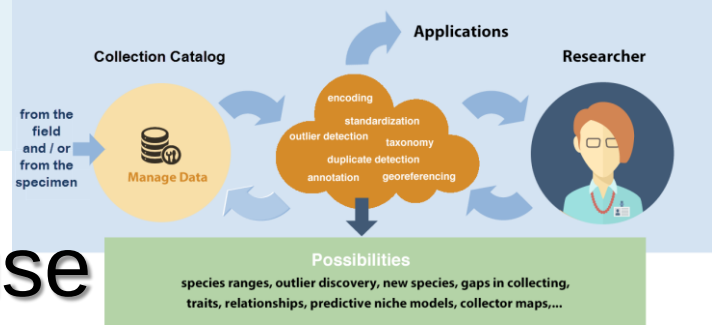
- Arctos
- Symbiota
- Specify
- Axiell EMu
- collectionspace

Community Building

- SPNHC
- GRBio
- SCNet
- ECN
- Working Groups: DROID, GWG etc.

Data Mobilization

- TDWG – Biodiversity Information Standards (tdwg)
- DwC - Darwin Core
- AC - Audubon Core
- IPT – Integrated Publishing Toolkit (from GBIF)
- Kurator and Akka
- GEOLocate
- DAMS – Digital Asset Management
- ABBYY
- Unique identifiers - GUID, UUID, ARK, IGSN



Data aggregation and Data use

Biodiversity Data Mobilization and Use

- **EoL** - Encyclopedia of Life
- **BHL** – Biodiversity Heritage Library
- **NEON** – National Ecological Observatory Network
- **CyVerse** (iPlant) - facilitating scientific research in the cloud

Aggregating Specimen Data

- **GBIF** – Global Biodiversity Information Facility
- **iDigBio**
- **VertNet** – Vertebrate data portal, Data mobilization and cleaning services
- **BISON** – Biodiversity Serving our Nation – GBIF North American Node
- **USVH** – United States Virtual Herbarium
- **ALA**, **DigiVol** – Atlas of Living Australia
- **Canadensys** - unlocking specimen information held by Canadian university-based biological collections
- **SiBBr** – Brazilian Biodiversity Information System or Sistema de Informação sobre a Biodiversidade Brasileira)
- **CONABIO** (Mexico)
- **CRIA** (Brazil) and **SpeciesLink**

Education, Outreach, Inreach

- AIM-UP!
  - Advancing Integration of Museums into Undergraduate Programs; exploring the use of collections data in undergraduate education
- Notes from Nature
  - Crowdsourcing collections transcription, creating communities
- iNaturalist
  - observation app, connecting scientists and the general public
- ePandda
  - Enhancing Paleontological and Neontological Data Discovery API, connecting scientific literature with specimens
- Data Carpentry & Software Carpentry (DC and SwC)
  - **Biodiversity informatics skills** for graduate students and researchers. Focus on tidy data, fit for reproducible research. DC – novices, SwC for self-taught programmers in the sciences.

Publishing

- Pensoft
  - ARPHA Writing Tool (AWT): Authoring, Reviewing, Publishing, Hosting and Archiving
- DataCite, DOI (Digital Object Identifier)
  - Find, share and reuse, cite data, connect and get credit
- Mendeley
  - Free reference manager; organize papers, read & annotate your PDFs, collaborate in private or open groups, and securely access your research from everywhere.
- ORCiD
  - Open Researcher and Contributor ID
- DataONE
  - Data Observation Network for Earth (DataONE) – data repository and data management best practices
- Data Dryad
  - Curated resource making the data underlying scientific publications discoverable, freely reusable, and citable' general-purpose home for a wide diversity of datatypes
- Figshare
  - online digital repository where researchers can preserve and share their research outputs, including figures, datasets, images, and videos. Open data

iDigBio
Integrated Digitized Biocollections

Got more?

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

idigbio.org/wiki

facebook.com/iDigBio

twitter.com/iDigBio

vimeo.com/iDigBio

idigbio.org/rss-feed.xml

idigbio.org/events-calendar/export.ics

iDigBio is funded by a grant from the National Science Foundation's Advancing Digitization of Biodiversity Collections Program. 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.

Working and Interest Groups (wg)

DIGITIZATION

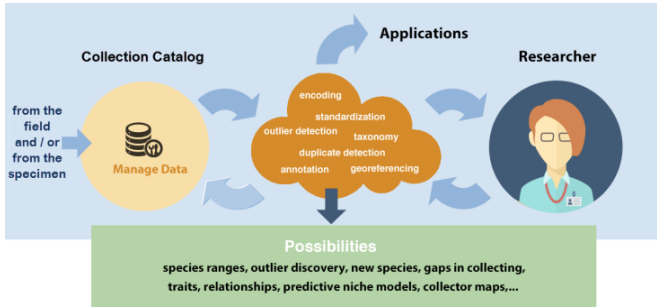
- Developing Robust Object to Image to Data (DROID1,2,3,4)
- International Whole-Drawer Digitization Interest Group (WDD)
- Fluid-preserved Arthropod and Microscopic Slide Imaging Interest Group
- Paleo Digitization Working Group (PaleoDigi)
- North American Network of Small Herbaria Working Group (NANSH)
- Biodiversity Informatics Management (BIM) Working Group
- Georeferencing Working Group (GWG)
- Symbiota Working Group (SWG)
- Augmenting OCR (aOCR)
- Outlier Detection and Documentation by Collectors Working Group (ODD Collectors)

E & O

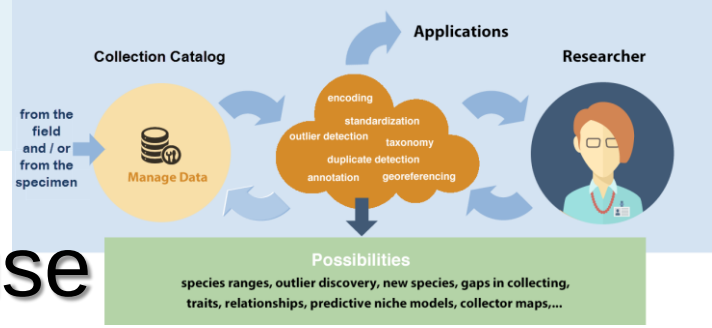
- Education & Outreach (E&O)
- Augmented Reality Public Education/Outreach Working Group (ARPEO)
- User Engagement for Public Participation in Digitization (CitSciEngage)
- Interoperability for Public Participation in Digitization (CitSciInterop)

VISION

- Cyberinfrastructure (CYWG)
- Data Management Interest Group (DMI)
- Project Management Interest Group (PMIG)
- Sustainability Working Group (SWG)



iDigBio
Integrated Digitized Biocollections



Data aggregation and Data use

Biodiversity Data Mobilization and Use

Aggregating Specimen Data

- EoL
- BHL
- NEON
- CyVerse (iPlant)

- GBIF
- iDigBio
- VertNet
- BISON
- USVH
- ALA, DigiVol
- Canadensys
- SiBBr
- CONABIO (Mexico)
- CRIA (Brazil) and SpeciesLink

Research tools and resources

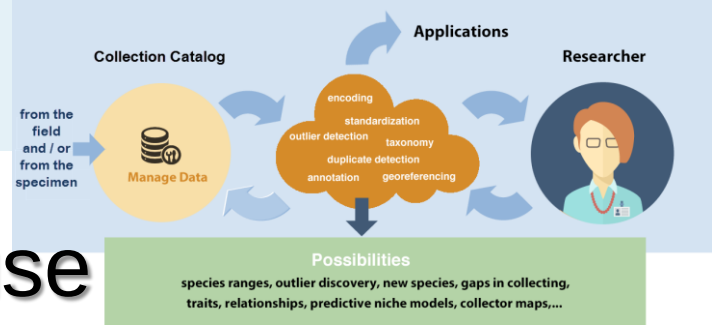
Implementation

- Lifemapper
- Phylolink (was Phylojive)
- Carto (aka CartoDB)
- QGIS
- API tools
  - riDigBio, rGBIF, rVertNet, ePandda

Architecture

- Python
- R, Rstudio, ROpenSci
- MySQL, SQL, SQLite
- GitHub

iDigBio
Integrated Digitized Biocollections



Data aggregation and Data use

Biodiversity Data Mobilization and Use

- EoL
- BHL
- NEON
- CyVerse (iPlant)

- GBIF
- iDigBio
- VertNet
- BISON
- USVH
- ALA, DigiVol
- Canadensys
- SiBBr
- CONABIO (Mexico)
- CRIA (Brazil) and SpeciesLink

Research tools and resources

Implementation

- Lifemapper
- Phylolink (was Phylojive)
- Carto (aka CartoDB)
- QGIS
- API tools
  - riDigBio, rGBIF, rVertNet, ePandda

Architecture

- Python
- R, Rstudio, ROpenSci
- MySQL, SQL, SQLite
- GitHub