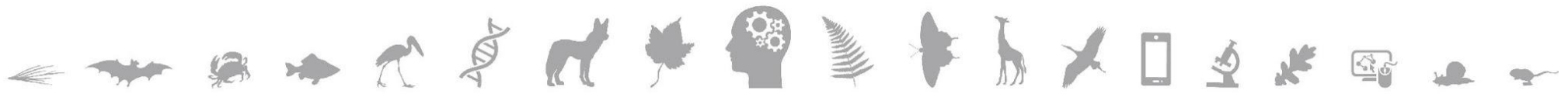




Overview of iDigBio Phase 3 Project

NSF Award DBI-2027654 (2021-2026)

Gil Nelson, Director



Phase 1 & Phase 2 ⇨ Advancing Digitization

In an effort to make collections universally accessible to taxonomists, ecologists, researchers, and the general public, in 2011, the U.S. National Science Foundation launched a \$100 million, 10-year, Advancing Digitization of Biodiversity Collections (ADBC) program and named the University of Florida and Florida State University jointly as the coordinating center and national resource for digitization.

The scope of our work was limited to public, non-federal, U.S. collections, though NSF has encouraged us to develop domestic and international collaborations, including especially with GBIF, ALA, DiSSCo and the National Museum of Natural History - Smithsonian.

The goal of ADBC was to digitize and make available via the Internet records for all biological and paleontological collection objects in North America over the 10-year life of the project.



Phase 3 ⇒ **Sustaining** the Momentum

iDigBio Phase 3: Sustaining the digitization, mobilization, accessibility, and use of biodiversity specimen data in U.S. museum and academic collections

Proposal submitted on 3/23/2020 by UF with subawards to FSU and ASU
Award made on 5/5/2021 (NSF DBI-2027654)

Contents: 224-page proposal, including 46-page *Project Execution Plan*

NSF Program: Sustaining Infrastructure for Biological Research (Sustaining)

Award Period: 9/1/2021–8/31/2026

Final Budget: \$19,995,068 (AWD10160) with \$3,346,436 to FSU (SUB00002725)
and \$2,478,801 to ASU (SUB00002726)

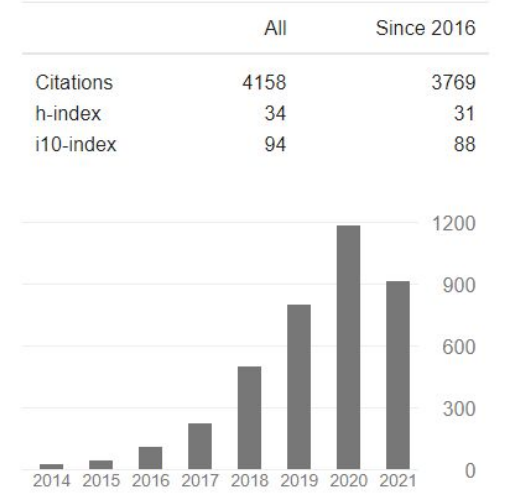
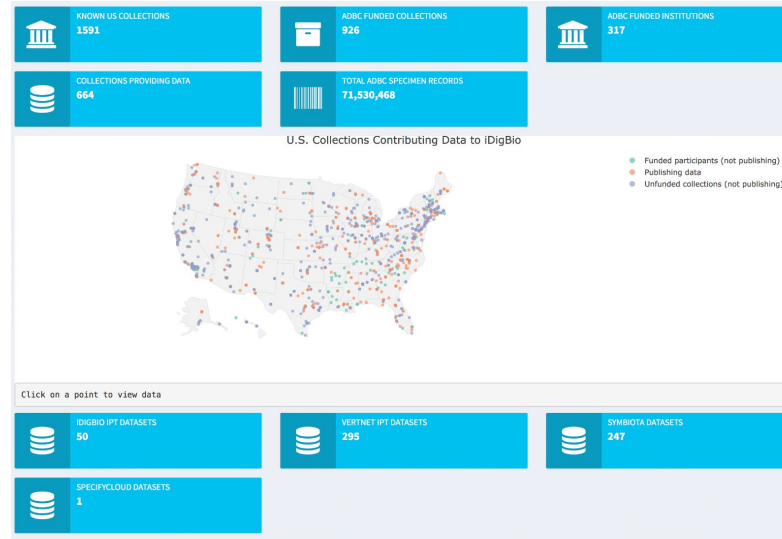
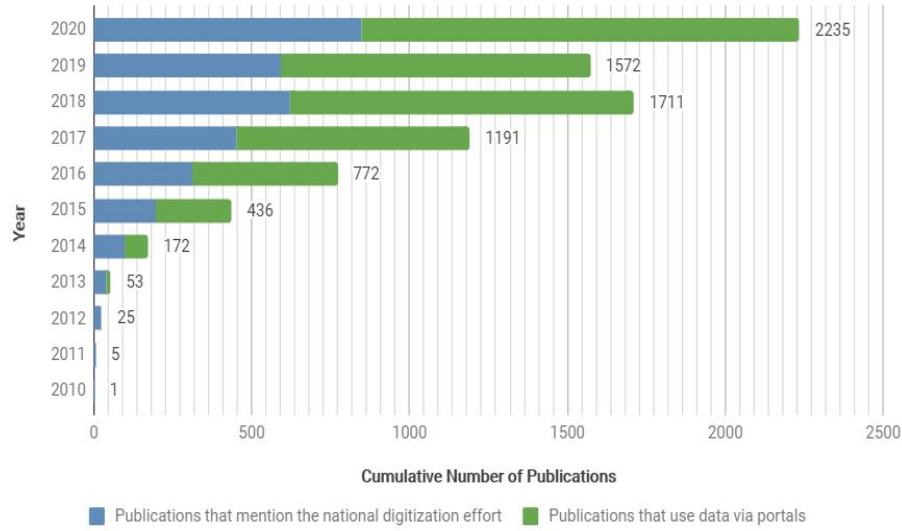


Phase 3 builds on strong results of prior support

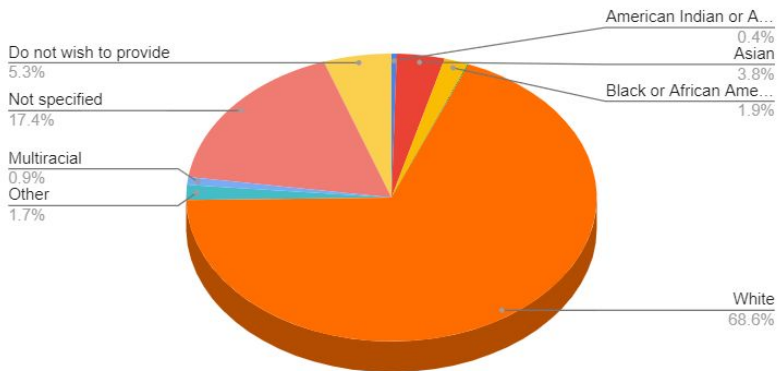
1. Produced more than 72 peer-reviewed publications
2. Trained more than 33 graduate students and 10 postdocs
3. Engaged over 18,600 participants in 482 workshops, webinars, symposia, and other events
4. Produced annual Digital Data in Biodiversity Research Conference
5. Produced annual Biodiversity Summit
6. Supported 33 working and interest groups
7. Aggregated, served, and enabled searching and visualization of 131 million specimen records with 43 million associated media from nearly 1,700 datasets
8. Supported 33 Thematic Collections Networks (TCNs) and 54 Partners to Existing Networks (PENs) involving more than 900 collections in more than 300 institutions across the country



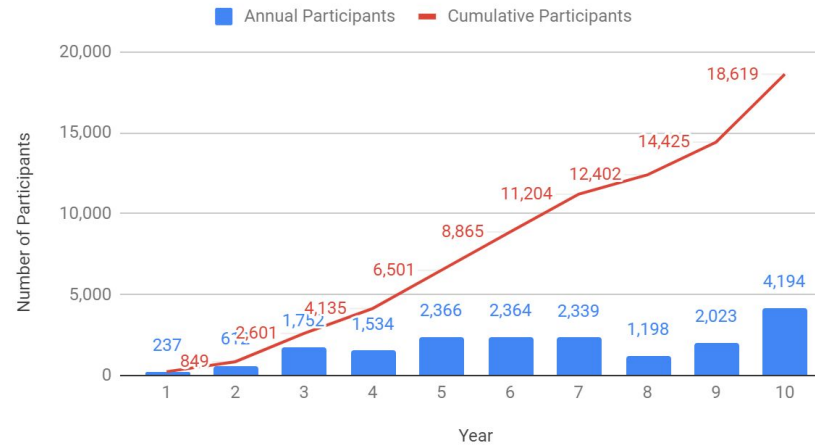
Status of the National Biodiversity Collections Digitization Effort



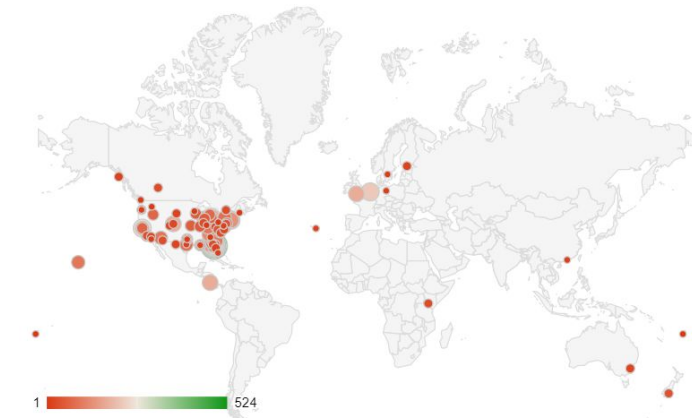
Racial Identity of iDigBio Event Participants (2018-2020)



Participants in iDigBio Events



iDigBio Onsite Events by Location

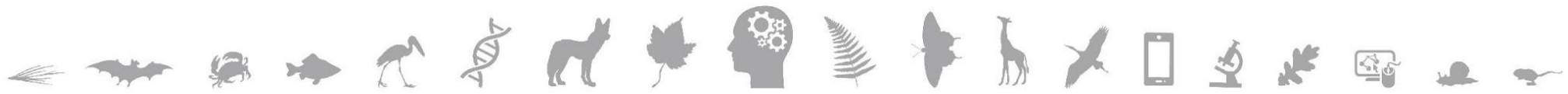




Major goal remains the same!

Catalyze excellence in **digitization, mobilization, and usage** of data about the roughly 1 billion biodiversity specimens curated in the 1,600 U.S. biodiversity collections for **research and education.***

* We affirm that “catalyzing excellence” fundamentally must include ***broadening participation*** in the activities!



Organizing strategy is updated

1. Administration, Sustainability, and Community Coordination Domain (UF)
2. Digitization, Workforce Development, and Citizen Science Domain (FSU)
3. Cyberinfrastructure Domain (UF)
4. Promoting and Facilitating Research using iDigBio Data Domain (UF)
5. Education, Outreach, Diversity, and Inclusion Domain (UF)
6. **Symbiota Support Hub (ASU)**





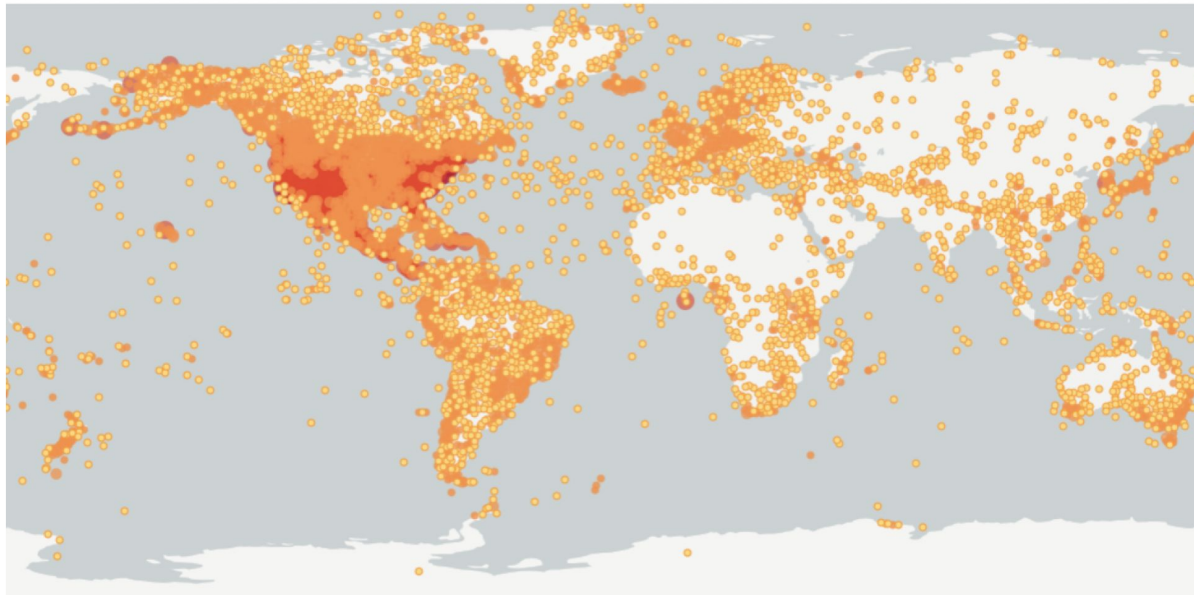
Introducing the Symbiota Support Hub ("SSH")

- Symbiota is an **open source software** for managing and mobilizing biodiversity data. See <https://doi.org/10.3897/BDJ.2.e1114>
- Key traits:
 - **Integrated content management system** - data are managed collaboratively via a browser-based user interface.
 - **Theme-based data aggregation** - data sharing within and among self-managed communities.
 - **Coalesced user communities** - each portal has a specific taxonomic, geographic, and social scope.
 - **FAIR data principles** - DwC data in Symbiota portals are findable, accessible, interoperable, and reusable.



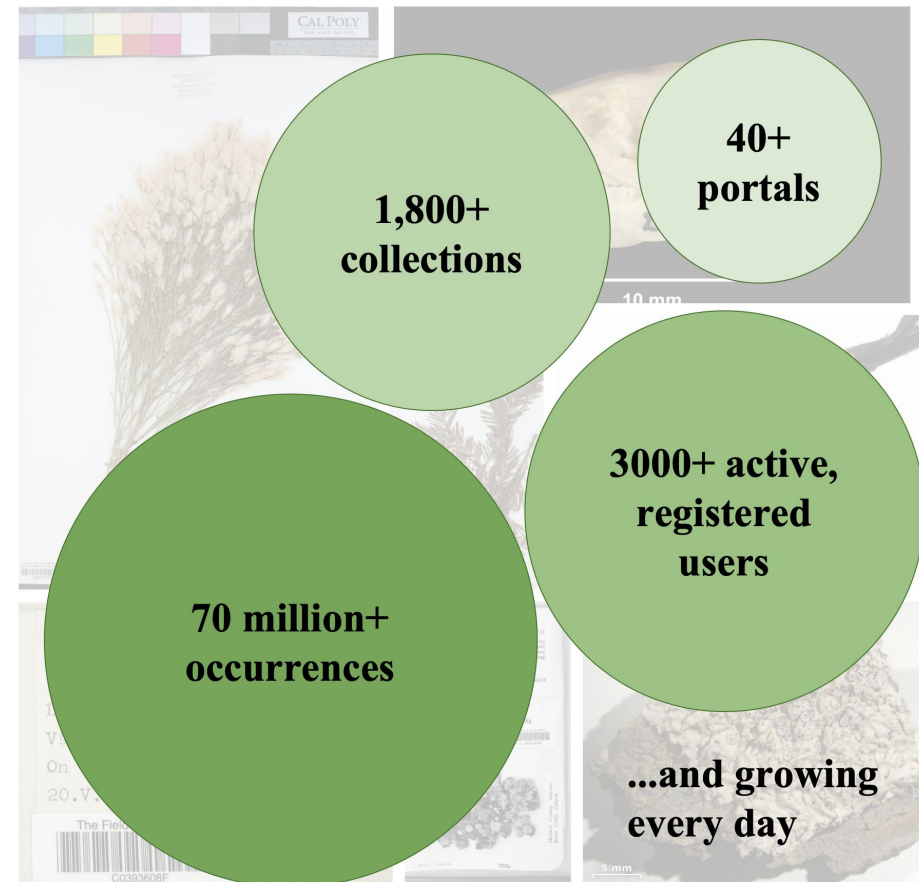
Present reach

- Symbiota portals support **22 Thematic Collections Networks**; including 700+ live-managed collections.



Symbiota portals contribute over 12 million occurrence records to GBIF.

Data served in [GBIF.org](https://www.gbif.org). Accessed 9/15/2021





SSH team

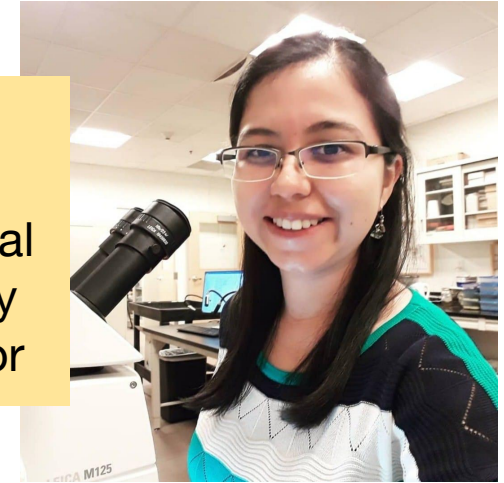
+2 open searches: IT Manager & Community Manager



Dr. Jenn Yost
Community Lead



Katie Pearson
Data Manager



Samanta Orellana
International
Community
Coordinator



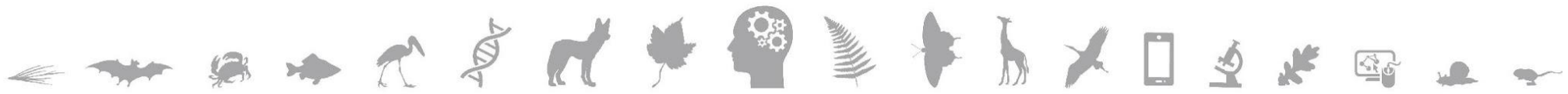
Dr. Nico Franz
Management
@ ASU

Ed Gilbert
IT Management
Lead



Dr. Laura Prado
Biodiversity
Informatician
(NEON Biorepo)





SSH - iDigBio Phase 3 and beyond

- **New IT infrastructure capabilities** - portal and **image hosting**.
- **Services for *all* Symbiota portal users.**
 - *Help Desk support* - access, management, publication, sharing.
 - Scalable *documentation* for all contributors and users.
 - *Community capacitation* - webinars, trainings, workshops.
 - Implementation of sustainable *business models*.
- Parallel, project-specific **software development** - *Extended Specimen Network*.
- **Email us at symbiota@asu.edu**



Domain 1: Administration, Sustainability, and Community Coordination

(Alnycea Blackwell, Caitlin Chapman, Libby Ellwood, Jill Goodwin, David Jennings)

1. Support 50 workshops (10/yr)
2. Produce 2 annual conferences (Digital Data & Summit)
3. Enable the Symbiota Support Hub
4. Maintain & expand global partnerships
5. Sustainability planning

(but also NSF reporting, project administration, and much more)

1.0 Administration, Sustainability, and Community Coordination	Y1	Y2	Y3	Y4	Y5
1.1 Project administration, management, and coordination	▲	▲	▲	▲	▲
1.2 Track overall progress towards goals and objectives	▲	▲	▲	▲	▲
1.3 Continual engagement and coordination with other aggregators and organizations	▲	▲	▲	▲	▲
1.4 Provide logistical support for workshops, webinars, symposia, and other events	▲	▲	▲	▲	▲
1.5 Conduct annual Biodiversity Summit	▲	▲	▲	▲	▲
1.6 Conduct annual Digital Data in Biodiversity Research Conference	▲	▲	▲	▲	▲
1.7 Lead Collections Data Infrastructures Working Group	▲	▲	▲	▲	▲
1.8 Communication, integration, systems optimization, and acceptance of protocols/standards	▲	▲	▲	▲	▲
1.9 Support the creation of Symbiota Support Hub at ASU	▲	▲			
1.10 Coordinate with the Symbiota Support Hub to bring new collections online and help manage the demands of existing collections	▲	▲	▲	▲	▲
1.11 Plan for long-term sustainability	▲	▲	▲	▲	▲



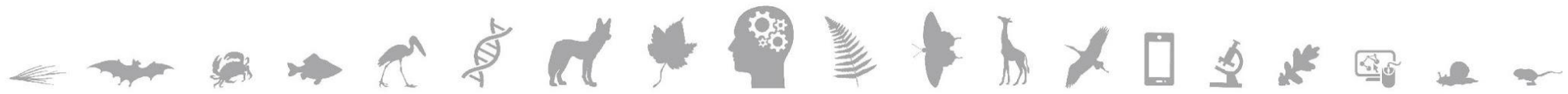
Domain 2: Digitization, Workforce Development, and Citizen Science

(Robert Bruhn, Michael Denslow, Erica Krimmel)

1. Knowledgebase development
2. Digitization gap-filling
3. Digitization Academy
4. Professional development strategic planning
5. WeDigBio growth
6. ML and other AI-enhanced digitization

2.0 Digitization, Workforce Development, and Citizen Science		Y1	Y2	Y3	Y4	Y5
2.1	Digitization training & optimization events	▲	▲	▲	▲	▲
2.2	Challenge-focused events to address community needs	▲	▲	▲	▲	▲
2.3	Digitization gap analysis within biodiversity community	▲	▲			▲
2.4	Evolution of digitization resources knowledge base	▲	▲	▲	▲	▲
2.5	Support collections access to the digitization resources	▲	▲	▲	▲	▲
2.6	Conduct semi-annual Strategic Planning courses	▲	▲	▲	▲	▲
2.7	Support semi-annual WeDigBio events	▲	▲	▲	▲	▲
2.8	Continue evolution of BIOSPEX with partner infrastructure	▲	▲	▲	▲	▲
2.9	Explore high-value AI incorporations into digitization, including an anomaly-description-in-text flagging and reporting workflow	▲	▲	▲	▲	▲
2.10	Communicate the value proposition of digitization, data management, and workforce development	▲	▲	▲	▲	▲





Domain 3: Cyberinfrastructure

(Ron Canepa, Caitlin Chapman, Michael Elliott, Erica Krimmel, Chris Wilson)

1. iCy architecture
2. iCy development and operational environment
3. Scaling-up services
4. Integrating new software, data, and hardware technologies
5. Leveraging partnerships
6. New data mobilization

3.0 Cyberinfrastructure		Y1	Y2	Y3	Y4	Y5
3.1	Data mobilization and ingestion	▲	▲	▲	▲	▲
3.2	Outreach to new collections	▲	▲	▲	▲	▲
3.3	Expansion, modernization, and convergence of services	▲	▲	▲	▲	▲
3.4	Increased automation and data integration beyond specimen records	▲	▲	▲	▲	▲
3.5	Integration and support of new software, data, and hardware technologies	▲	▲	▲	▲	▲
3.6	Convergence of cyberinfrastructure approaches and resource sharing by multiple aggregators	▲	▲	▲	▲	▲
3.7	Maintain infrastructure and services	▲	▲	▲	▲	▲
3.8	Plan for long-term sustainability		▲	▲	▲	▲





Domain 4: Promoting and Facilitating Research using iDigBio Data

(Tori Ford, Shelly Gaynor, Jill Goodwin, Mckenzie Mabry)

1. Trait data developments
2. Ontology developments
3. ML and other AI developments
4. Tracking data use

4.0 Promoting and Facilitating Research	Y1	Y2	Y3	Y4	Y5
4.1 Engage the research community in adoption of iDigBio services, infrastructure, tools, resources, and data	▲	▲	▲	▲	▲
4.2 Facilitate both traditional and novel uses of digitized specimen data	▲	▲	▲	▲	▲
4.3 Document use cases of research applications of specimen data, and track research use of data	▲	▲	▲	▲	▲
4.4 Artificial intelligence and machine learning	▲	▲	▲	▲	▲
4.5 Trait extraction and integration with specimen data	▲	▲	▲	▲	▲
4.6 Foster development of ontologies to enable new approaches in trait extraction and analysis		▲	▲	▲	▲



Domain 5: Education, Outreach, Diversity, and Inclusion (EODI) (Alyncea Blackwell, Caitlin Chapman, Adania Flemming, Molly Phillips)

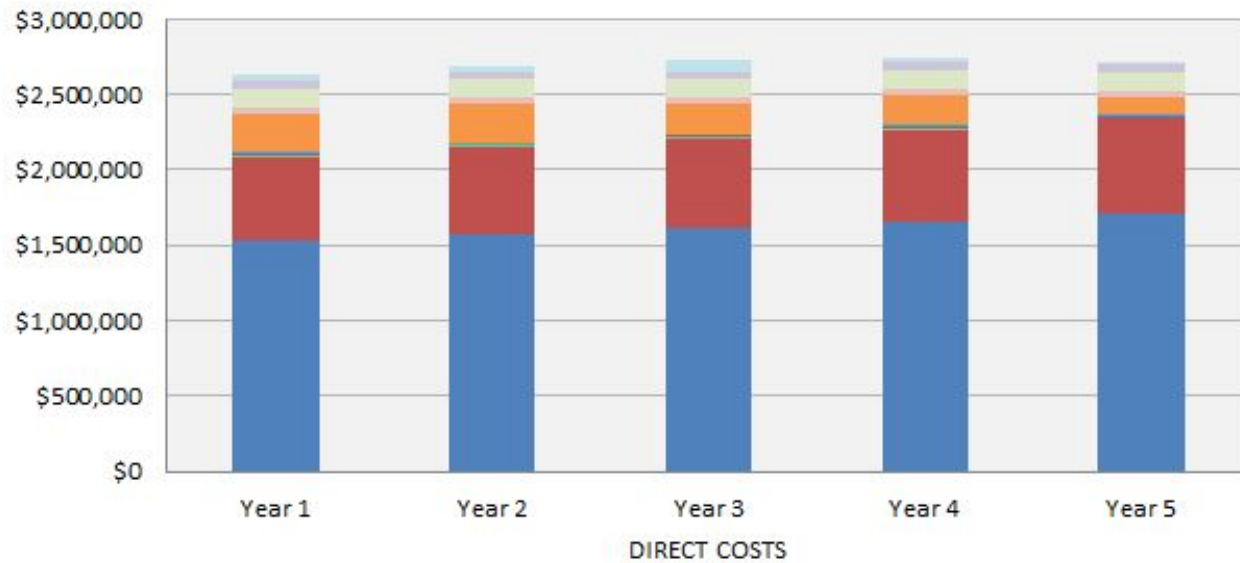
1. Creating, curating, and disseminating educational resources
2. Fostering community and scholarship around these resources
3. Recruiting and supporting at broadening participation events
4. Offering diversity and inclusion events at professional conferences

5.0 Education, Outreach, Diversity, and Inclusion		Y1	Y2	Y3	Y4	Y5
5.1	Increase awareness of digitized biodiversity collections across a diverse array of communities	▲	▲	▲	▲	▲
5.2	Empower new users through creation of multiple entry points to the data and resources	▲	▲	▲	▲	▲
5.3	Invite new people to the collections community through specific broadening representation efforts	▲	▲	▲	▲	▲
5.4	Create and disseminate best practices, recommendations, and exemplar resources for using biodiversity data in the classroom	▲	▲	▲	▲	▲
5.5	Provide hands-on training on incorporating digitized biodiversity data into the classroom	▲	▲	▲	▲	▲
5.6	Conduct educator training in association with national conferences	▲	▲	▲	▲	▲
5.7	Partner on broadening participation events	▲	▲	▲	▲	▲
5.8	Reach broader audiences and disciplines with biodiversity data	▲	▲	▲	▲	▲



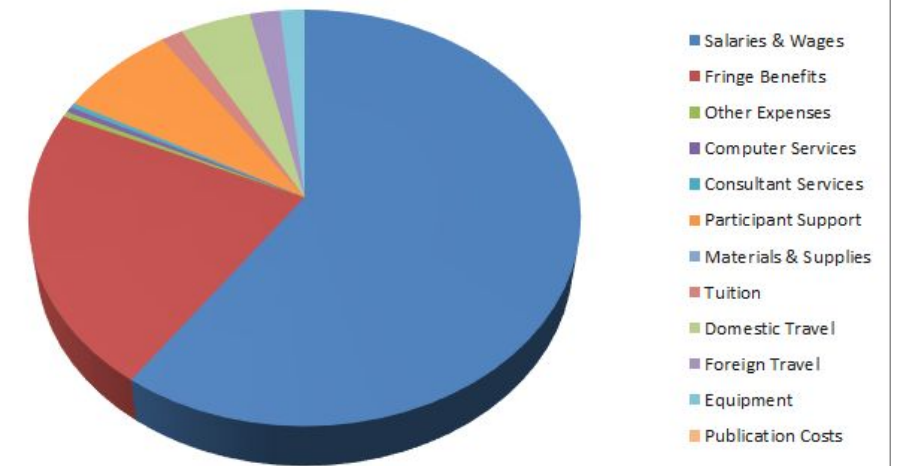
Budget Highlights

SABI iDigBio Phase 3 Direct Cost Profile by Category



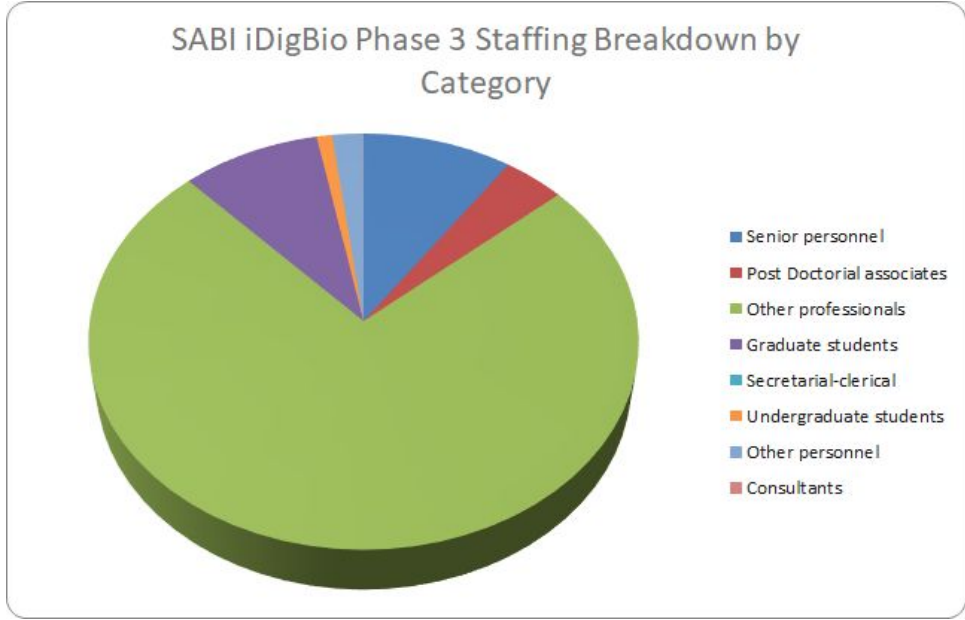
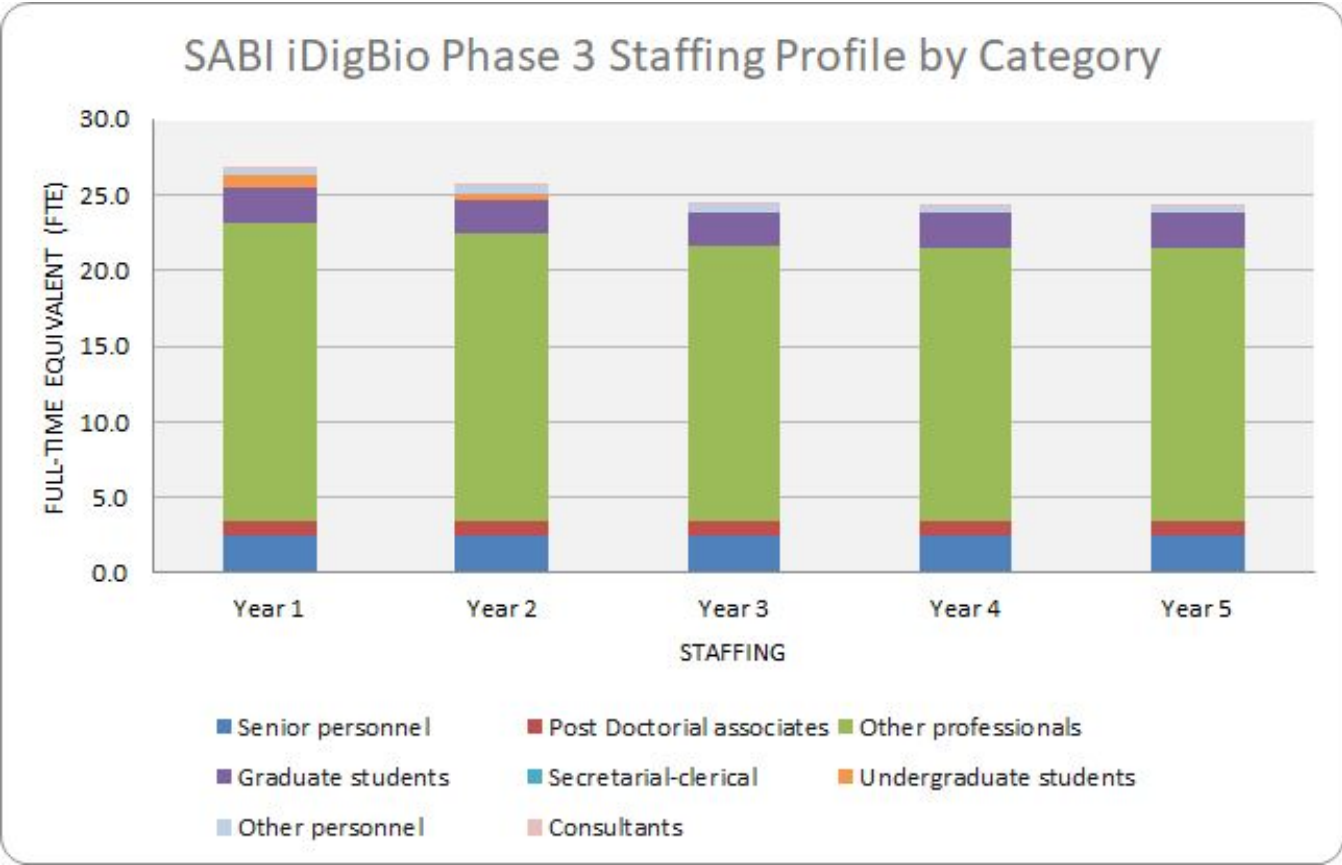
- Salaries & Wages
- Fringe Benefits
- Other Expenses
- Computer Services
- Consultant Services
- Participant Support
- Materials & Supplies
- Tuition
- Domestic Travel
- Foreign Travel
- Equipment
- Publication Costs

SABI iDigBio Phase 3 Direct Cost Breakdown by Category





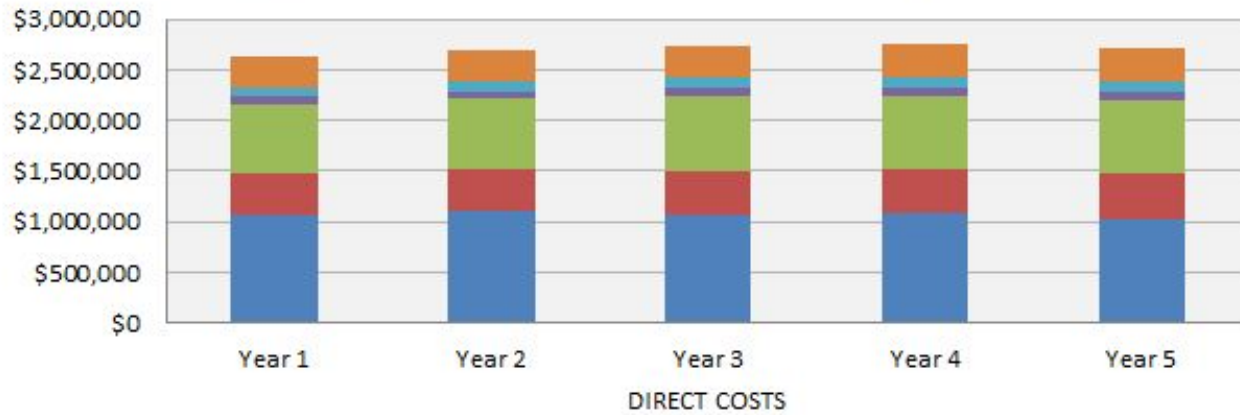
Staffing Highlights





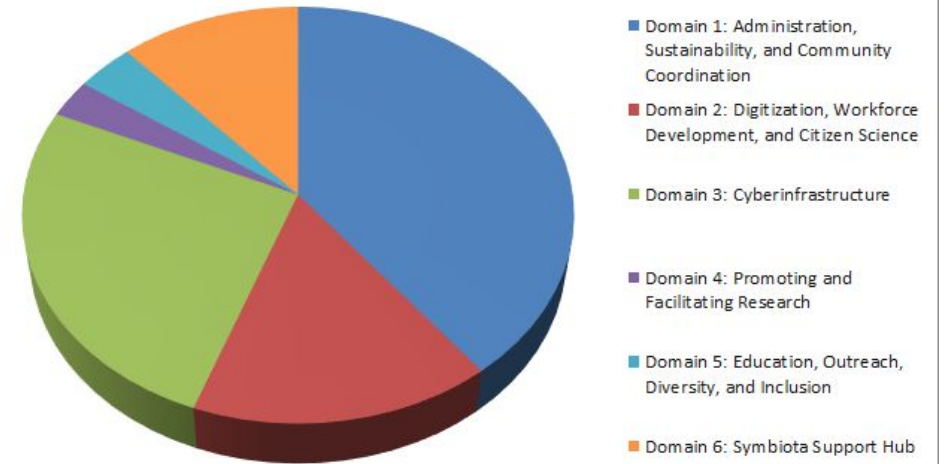
Domain Budget Highlights

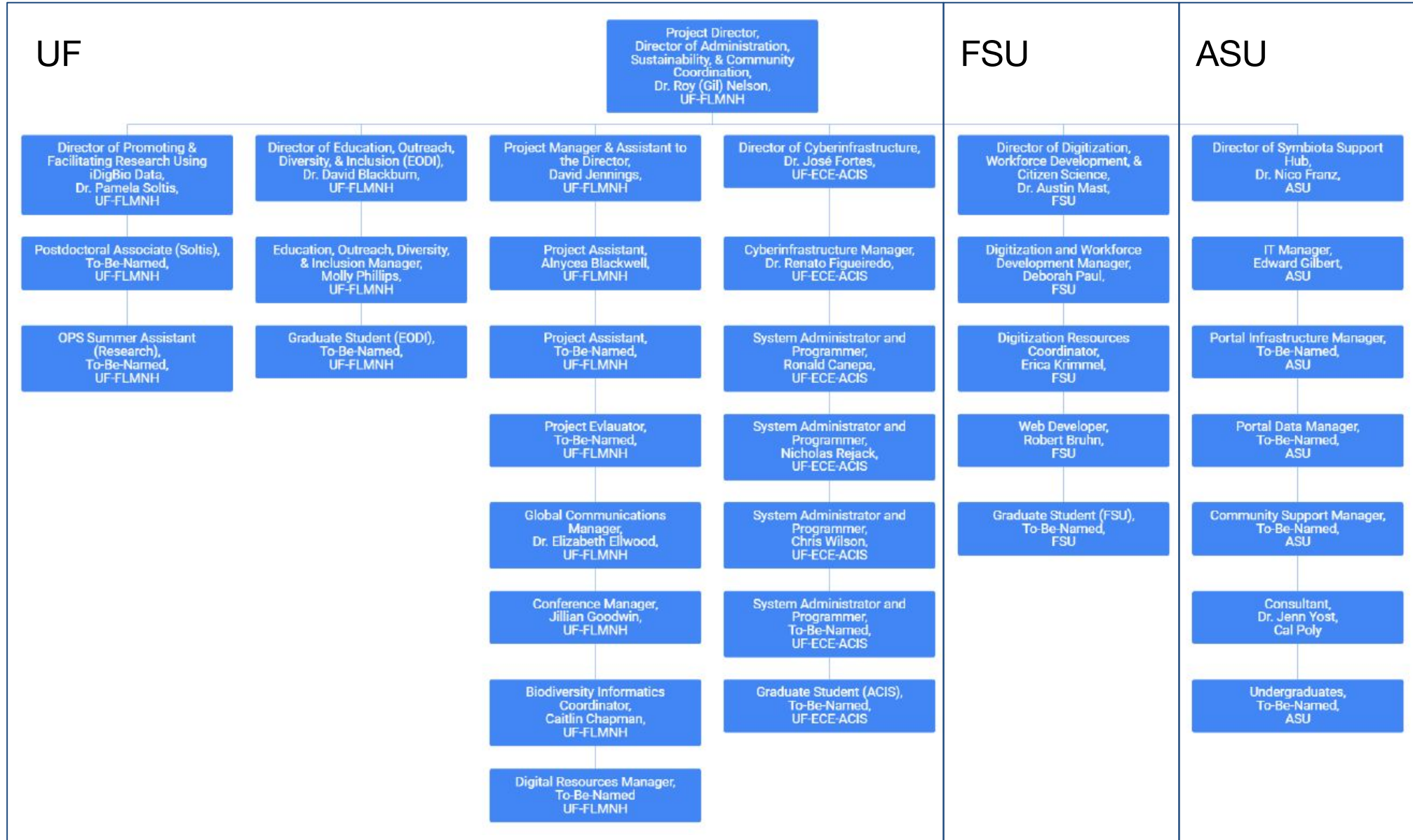
SABI iDigBio Phase 3 Direct Cost Profile by Domain

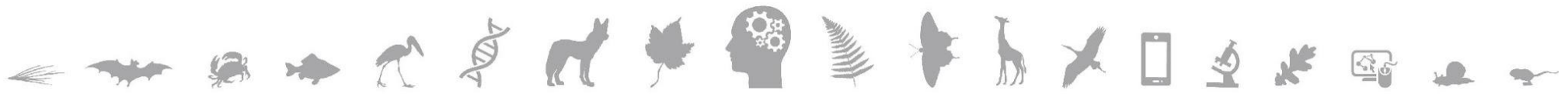


- Domain 6: Symbiota Support Hub
- Domain 5: Education, Outreach, Diversity, and Inclusion
- Domain 4: Promoting and Facilitating Research
- Domain 3: Cyberinfrastructure
- Domain 2: Digitization, Workforce Development, and Citizen Science
- Domain 1: Administration, Sustainability, and Community Coordination

SABI iDigBio Phase 3 Direct Cost Breakdown by Domain







Acknowledgements

- Thanks to NSF for their continued support of digitization of biodiversity collections through awards to TCNs, PENs, and iDigBio!
- Thanks to Austin Mast and David Jennings who prepared and presented previous version of this story for the Phase 2 and Phase 3 EAB in 2020 and 2021!
- Thanks to the continual hard work by the iDigBio team!
- Thanks to the new and existing TCNs and PENs helping iDigBio continue its journey!



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

Get Involved!



Alyncea Blackwell “Allie”
Project Assistant
ablackwell@floridamuseum.ufl.edu

Step 1: Sign up for the iDigBio Newsletter

- TCN and digitization news
- Upcoming workshops and webinars
- Event recaps
- Articles featuring innovative collections-based research
- Biodiversity Spotlights



<https://www.idigbio.org/newsletter-subscribe>

Step 2: Social media

facebook

Email or Phone

Password

Log In

Forgot account?



iDigBio
@iDigBio

Home
Posts
About
Photos
Events

Like Share

Posts

Home Moments

Search Twitter

Have an account? Log In



Tweets 5,150
Following 584
Followers 4,158
Likes 4,310
Lists 1

Follow

iDigBio

@iDigBio

iDigBio is coordinating the national effort to digitize biodiversity specimens and make them available online, funded by an @NSF grant to @UF and @floridastate.

Tweets Tweets & replies Media



@iDigBio · 5h

#CollectionsMatter New species of crocodile discovered in museum collections
phys.org/news/2019-09-s... via @physorg_com



Want to take advantage of all the new Twitter features?



vimeo.com/idigbio



idigbio.org/rss-feed.xml



idigbio.org/events-calendar/export.ics

Wiki

www.idigbio.org/wiki

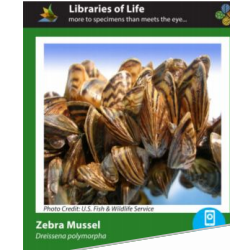
Step 3: Get involved with a Community Working Group



The screenshot shows the iDigBio website interface. At the top is the iDigBio logo with the tagline 'Integrated Digitized Biocollections'. Below the logo is a navigation bar with links for 'iDigBio Home', 'Wiki', 'Working Groups', 'Workshops', and 'Wiki Formatting Help'. On the left is a sidebar menu with categories like 'Wiki Home', 'Workshops', 'Working Groups', and 'iDigBio Working Groups'. A hand icon points to the 'Working Groups' link in the sidebar. The main content area displays the 'iDigBio Working Groups' page, which includes a list of working groups and their descriptions.

iDigBio Working Groups

- 1 Overview
- 2 Forming or Dissolving a Working/Interest Group
- 3 Active Working Groups
 - 3.1 Arctos Working Group
 - 3.2 Augmented Reality Public Education/Outreach Working Group (ARPEO)
 - 3.3 Augmenting OCR (aOCR)
 - 3.4 Biodiversity Collection Management Solutions Working Group
 - 3.5 Biodiversity Informatics Management (BIM) Working Group
 - 3.6 Data Management Interest Group (DMI)
 - 3.7 Developing Robust Object to Image to Data (DROID)
 - 3.7.1 DROID1: Flat Sheets and Packets
 - 3.7.2 DROID2: Pinned Specimens in Trays and Drawers
 - 3.7.3 DROID3: Things in Spirits
 - 3.7.4 DROID4: 3D objects in Trays
 - 3.8 Education & Outreach (E&O)
 - 3.9 Fluid-preserved Arthropod and Microscopic Slide Imaging Interest Group
 - 3.10 Georeferencing Working Group (GWG)
 - 3.11 Integrating Collections and Ecological Research (ICER)
 - 3.12 International Whole-Drawer Digitization Interest Group (WDD)
 - 3.13 Interoperability for Public Participation in Digitization (CitSciInterop)
 - 3.14 North American Network of Small Herbaria Working Group (NANSH)



Documentation
API development
Workflows
Standards
Best practices
Hackathons



Paleo Digitization Happy Hour



**This is an
informal,
biweekly
discussion.**

**Every other
Thursday!**

<https://www.idigbio.org/outreach-events-sidebar>

Collections Education Coffee Break Series



Collections Education
Coffee Break Series

Join us monthly for a quick session
introducing collections-based
education and outreach resources!

<https://bit.ly/3cP9xXd>

Every third Wednesday
@ 11:30 am ET



Once a month

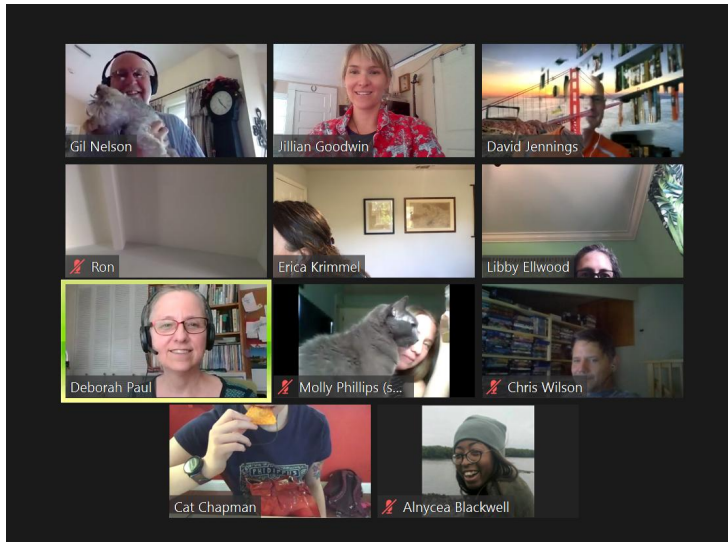
**Every Third
Wednesday!**

<https://www.idigbio.org/outreach-events-sidebar>

Step 4: Watch a webinar...or star in one!

<https://www.idigbio.org/tags/webinar>

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



Workshops
Symposia
Webinars

preferably with a microphone!

Step 5: Contribute to the iDigBio website

Bering Land Bridge and the MyCoPortal

Contributed by: Teresa Iturriaga, Rhianna Baldree, Alex Kuhn, Andrew Miller

- Submit an article for the **Research Spotlight**
- Write an article about **your project**
- Contribute your **workflows**
- **Update** your individual TCN wiki pages
- Write about your **iDigBio experience**
- Post an **event**
- Share education/outreach **resources**



Mycologists long to collect
areas remote to most men
where fungi today may thrive
keeping plants, trees, and cycles alive.

Bridges are to their liking
since one can go underneath
connecting with what lies beneath.
About fungi this is most striking.

In summer some may float
if the bridge is over a moat.
Fungi are versatile and persistent
to new niches they aren't resistant.

Step 6: Use the portal for research and data cleaning – feedback!

The screenshot shows the top navigation bar of the iDigBio portal with links for About iDigBio, Research, Technical Information, and Education. Below this is a green banner with a survey request: "Take our 30-second survey! The U.S. National Science Foundation and iDigBio are required to collect information on use of digitized collections-based specimen data. Please help us meet this requirement every time you use this search portal. Sustainability of the national digitization effort depends on evidence of data use! Maybe later." Below the banner is a green navigation bar with links for iDigBio Home, Portal Home, Search Records, Learning Center, Data, Research Collaboration, and Feedback.

data@idigbio.org

The screenshot shows the "Search Records" interface. On the left, there are search filters for "Scientific Name" (with a dropdown for "dwc:scientificName" and checkboxes for "Present" and "Missing"), "Date Collected" (with "Start:" and "End:" dropdowns and checkboxes for "Present" and "Missing"), and "Country" (with a dropdown for "dwc:country" and checkboxes for "Present" and "Missing"). In the center, there is a heatmap visualization showing "Record Density" with a color scale from 1 (blue) to 948 (red). The heatmap shows a high density of records in the eastern United States and parts of South America.

Family	Scientific Name	Date Collected	Country	Institution Code
"	"	1997-11-14	Brasil	IAC
Unplaced	"Acer" knowltoni	no data	United States	UF
Hamamelidaceae	"Acer" (Liquidambar) lesquereuxi	no data	United States	UCMP
Hamamelidaceae	"Acer" (Liquidambar) lesquereuxi	no data	United States	UCMP
Achatinellidae	"achatiniella" sp.	no data	no data	NHMUK
Achatinellidae	"achatiniella" sp.	no data	no data	NHMUK
Achatinellidae	"achatiniella" sp.	no data	no data	NHMUK
Unplaced	"Almont samara"	no data	United States	UF
Unplaced	"Almont samara"	no data	United States	UF

The screenshot shows the "Recordset" page with tabs for "Data Corrected", "Data Use", and "Raw". Below the tabs is a table with columns for "Flag", "Records With This Flag", and "(%) Percent With This Flag". The table lists various data quality flags and their corresponding record counts and percentages.

Flag	Records With This Flag	(%) Percent With This Flag
idigbio_isocountrycode_added	67961	98.832
dwc_continent_added	67932	98.79
geopoint_datum_missing	60241	87.605
dwc_datasetid_added	15170	22.061
dwc_kingdom_added	15170	22.061
dwc_parentnameusaged_added	15170	22.061
dwc_taxonid_added	15170	22.061
dwc_taxonomicstatus_added	15170	22.061
dwc_taxonrank_added	15170	22.061
gbif_canonicalname_added	15170	22.061
gbif_genericname_added	15170	22.061
gbif_taxon_corrected	15170	22.061
dwc_phylum_added	14947	21.737
dwc_scientificnameauthorship_added	14714	21.398
dwc_class_added	14460	21.028
dwc_multimedia_added	8706	12.661
taxon_match_failed	8593	12.496
dwc_order_replaced	8162	11.87
gbif_vernacularname_added	7878	11.457

Step 7: Collaborate!



[iDigBio Planned Network Maintenance](#) 09/25/2019 - 17:00 to 20:00

Research | Portal Home | **Research Collaboration** | Learning Center | Genetic Resources



- Researchers**
Browse our specimen portal
- Collections Staff**
Learn how your collection can benefit from our work
- Teachers & Students**
Learning resources & opportunities to engage

iDigBio Collaborations Enabling Research

Thu, 2014-07-24 16:15 -- ammatsum
To facilitate the study of biodiversity, a number of software products are being collaboratively developed with researchers and projects. These websites, tools, and workflows take advantage of the data being digitized at US and global institutions and made available by iDigBio through our [data services](#). Many other tools and services can be found through the [Biodiversity Catalogue](#). If you have a great idea, [contact us](#) or submit a proposal!



WordPress Leaflet Map Plugin Using iDigBio Data

iDigBio has collaborated with the [Atlas of Ordovician Life](#) project, part of the [PALEONICHES-TCN](#) to create a mapping plugin for [WordPress](#) that can generate maps of iDigBio specimen data on the fly. [Leaflet Map](#) enables map generation within WordPress webpages, and [Leaflet iDigBio geojson data plugin](#) developed by iDigBio enables an API query to iDigBio specimen occurrence coordinates. An example of the information about the plugin is available on <https://github.com/iDigBio/leaflet-iDigBio-geojson>.

