Overview of iDigBio Phase 3 Project

NSF Award DBI-2027654 (2021-2026)

Gil Nelson, Director

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.

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)

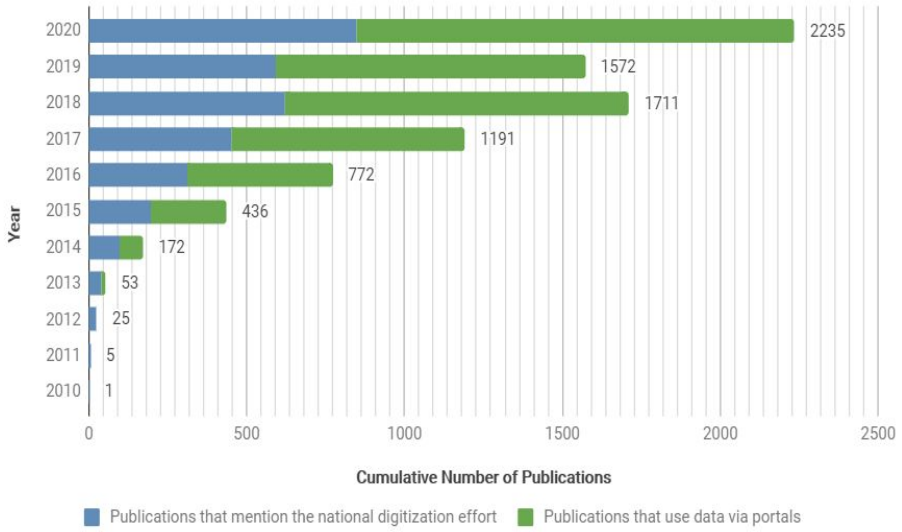
Contents: 224-page proposal, including 46-page *Project Execution Plan*
NSF Program: Sustaining Infrastructure for Biological Research (Sustaining)
Award Period: 9/1/2021–8/31/2026
Final Budget: $19,995,068 (AWD10160) with $3,346,436 to FSU (SUB00002725) and $2,478,801 to ASU (SUB00002726)

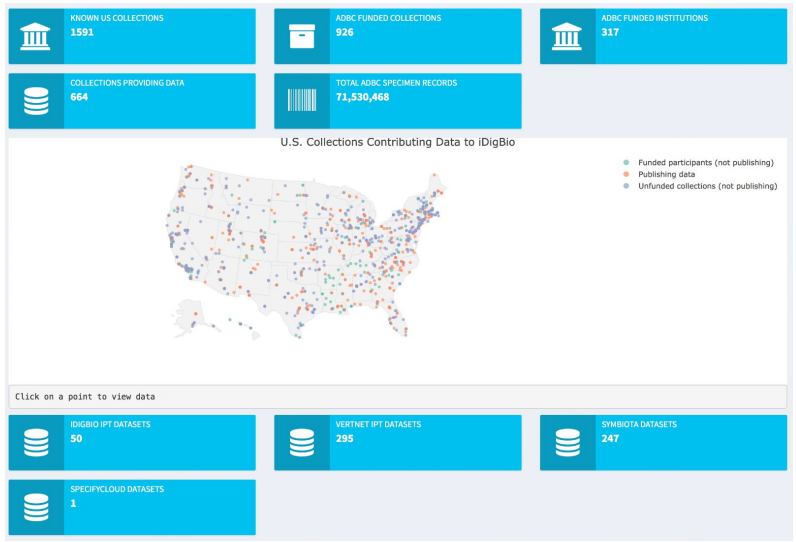
Phase 3 builds on strong results of prior support

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

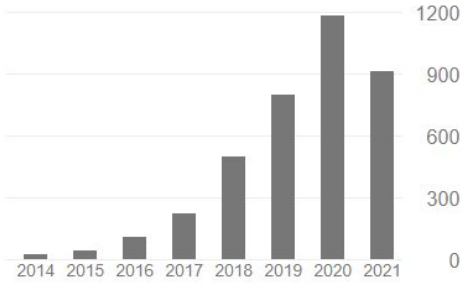
iDigBio
Integrated Digitized Biocollections

Status of the National Biodiversity Collections Digitzation Effort

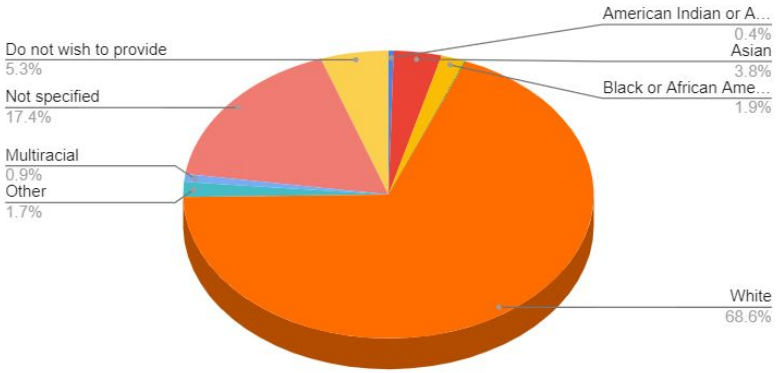




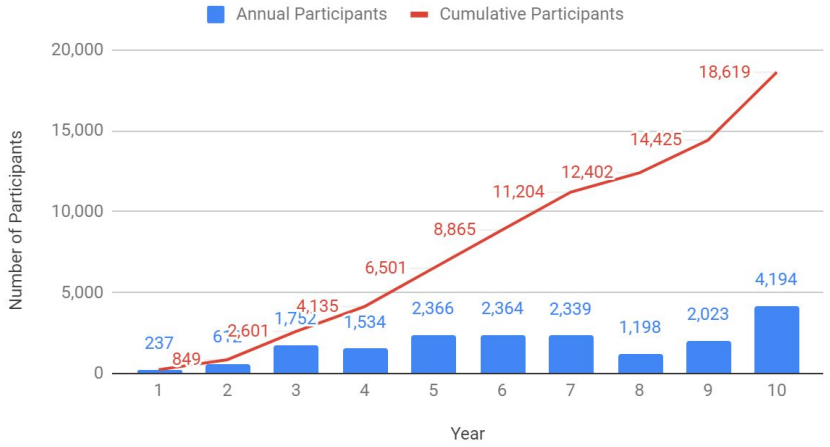
	All	Since 2016
Citations	4158	3769
h-index	34	31
i10-index	94	88



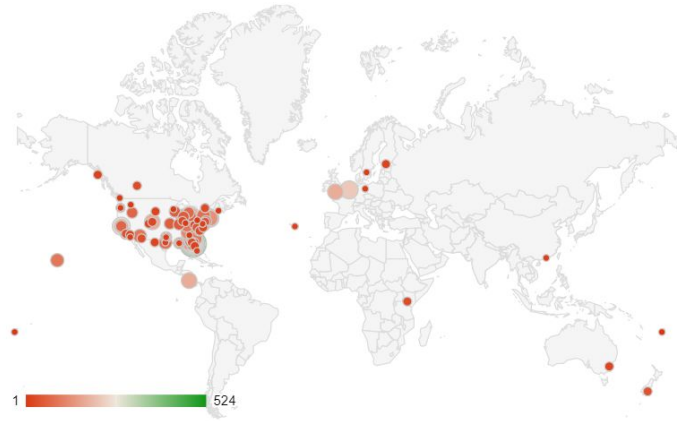
Racial Identity of iDigBio Event Participants (2018-2020)



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)**

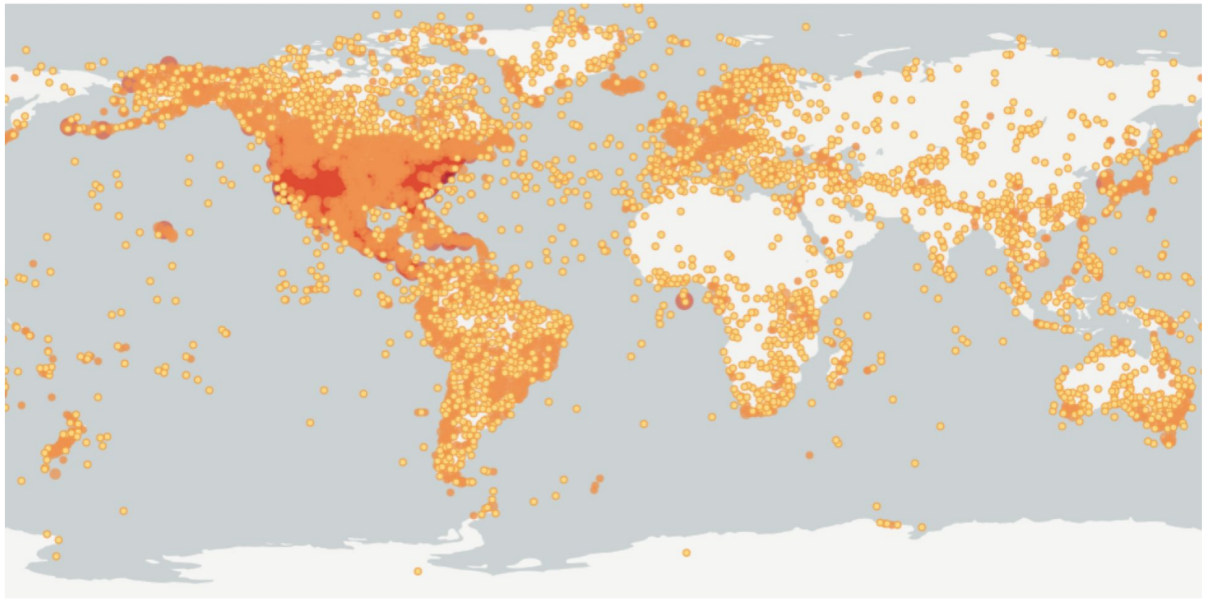
Introducing the Symbiota Support Hub ("SSH")

- 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.

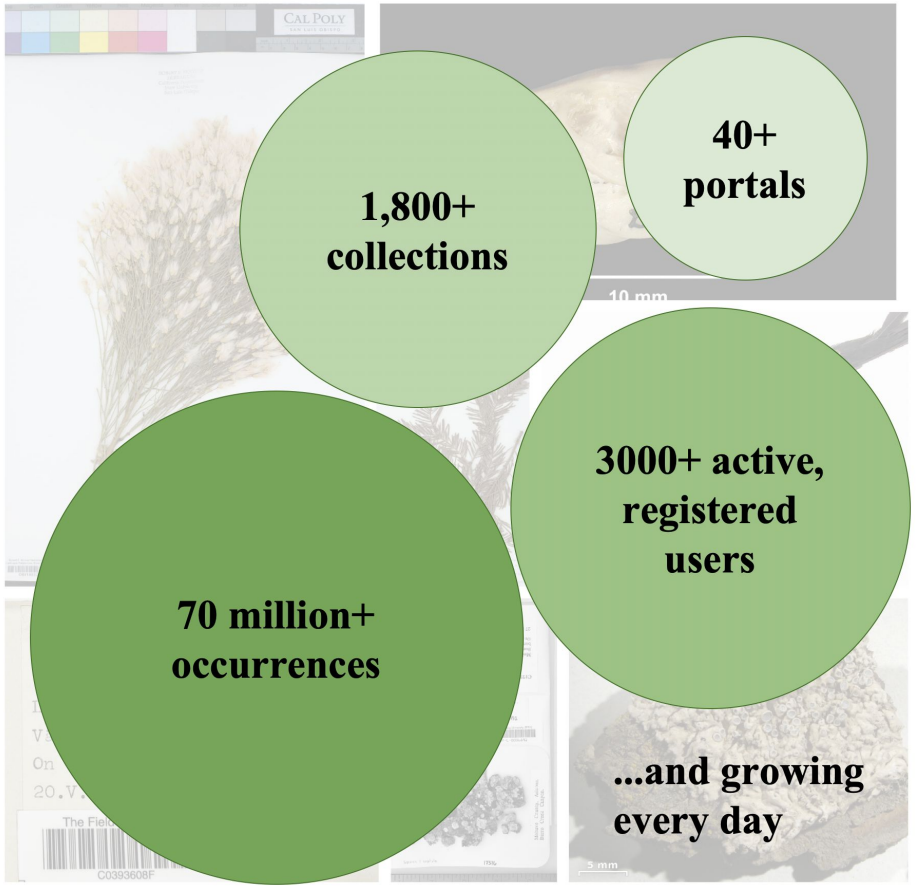
Present reach

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



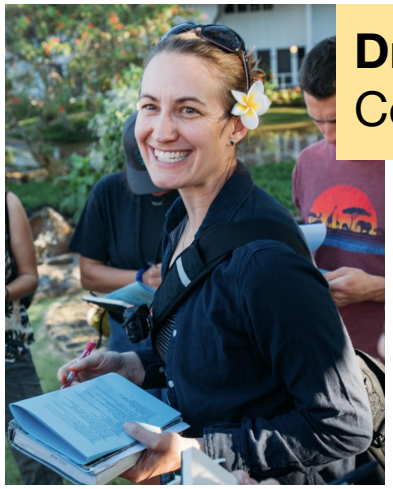
Symbiota portals contribute over 12 million occurrence records to GBIF.

Data served in GBIF. org. Accessed 9/15/2021



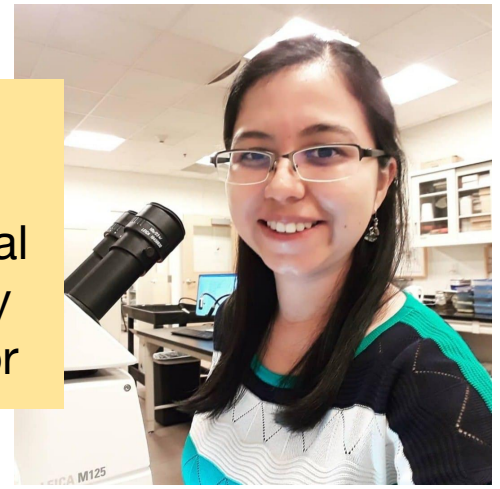
SSH team

+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**

Biodiversity Knowledge Integration Center
Arizona State University

CAL POLY
SAN LUIS OBISPO

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		▲ ▲	▲ ▲		

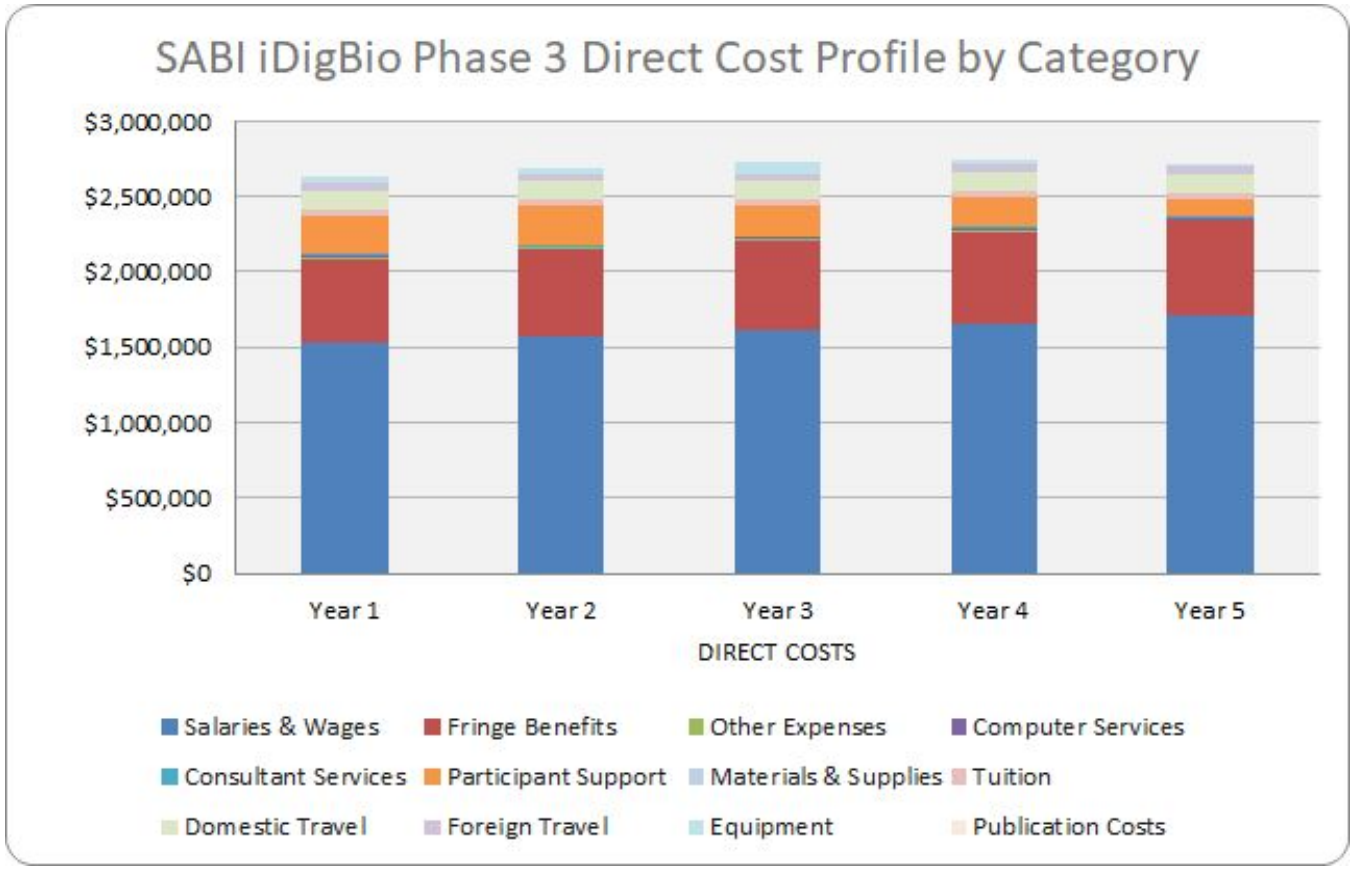
Domain 5: Education, Outreach, Diversity, and Inclusion (EODI)

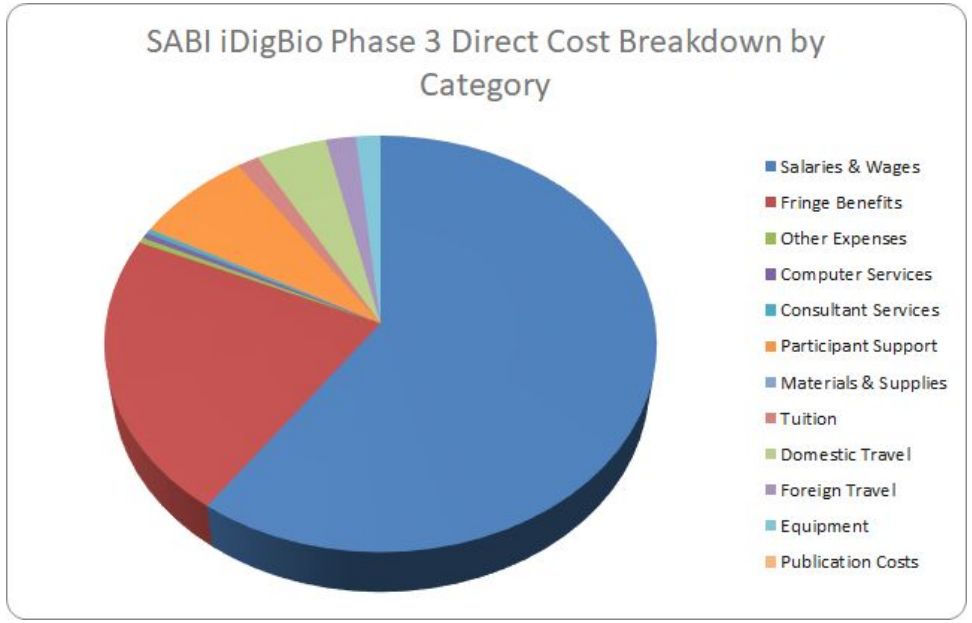
(Alnycea Blackwell, Caitlin Chapman, Adania Flemming, Molly Phillips)

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

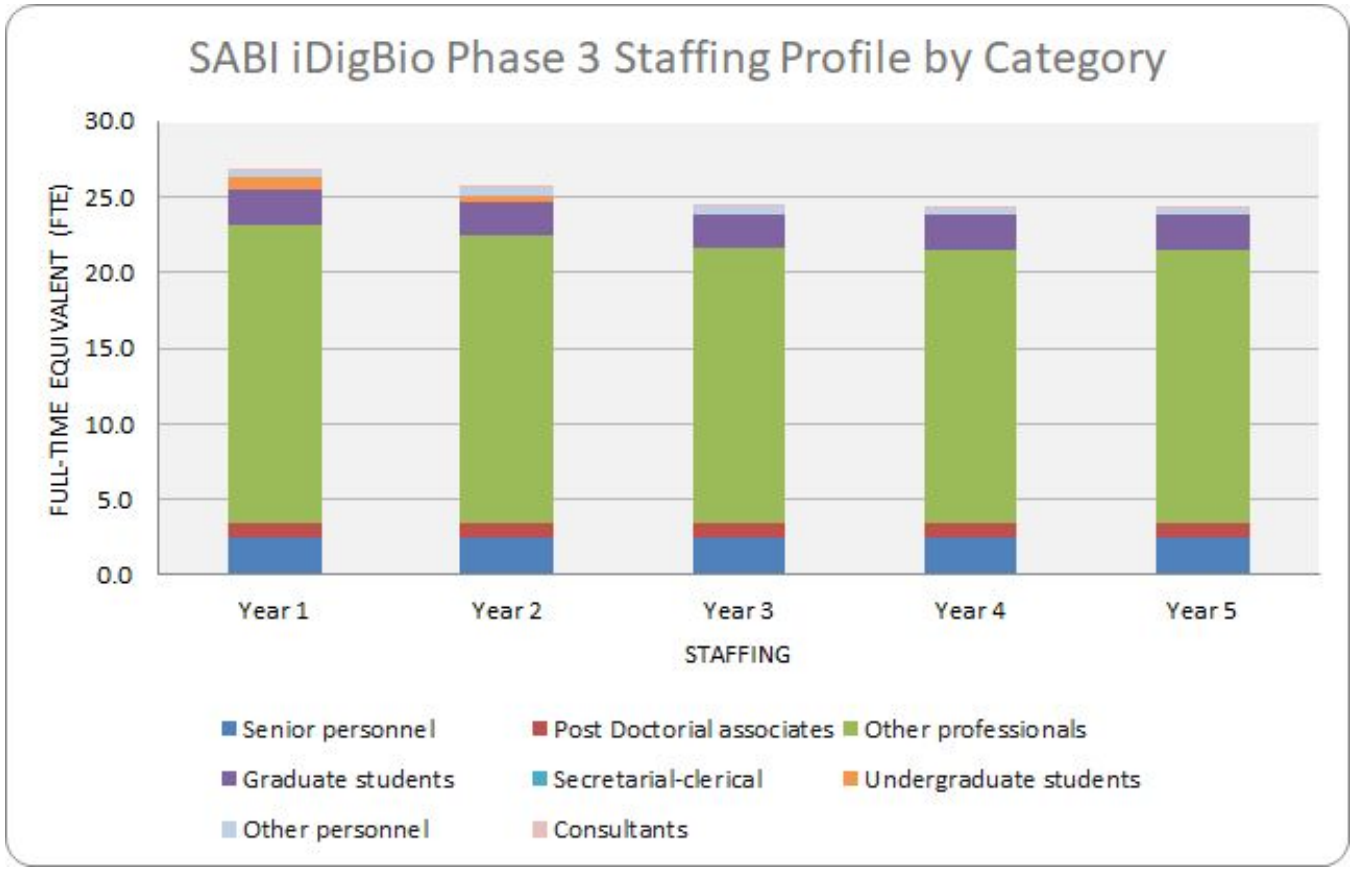
5.0	Education, Outreach, Diversity, and Inclusion	Y1	Y2	Y3	Y4	Y5
5.1	Increase awareness of digitized biodiversity collections across a diverse array of communities	▲	▲	▲	▲	▲
5.2	Empower new users through creation of multiple entry points to the data and resources	▲	▲	▲	▲	▲
5.3	Invite new people to the 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	▲	▲	▲	▲	▲

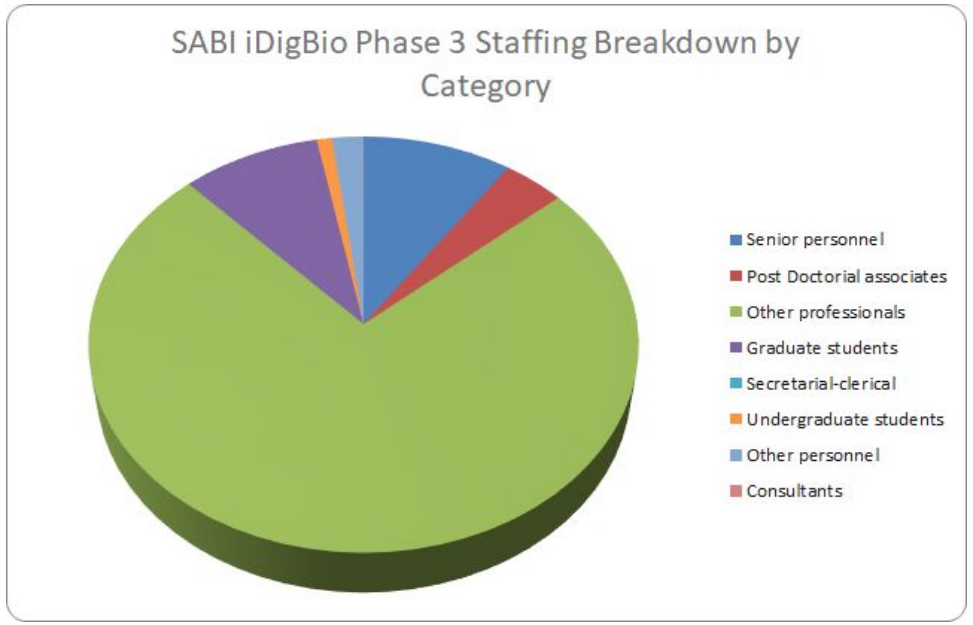
Budget Highlights



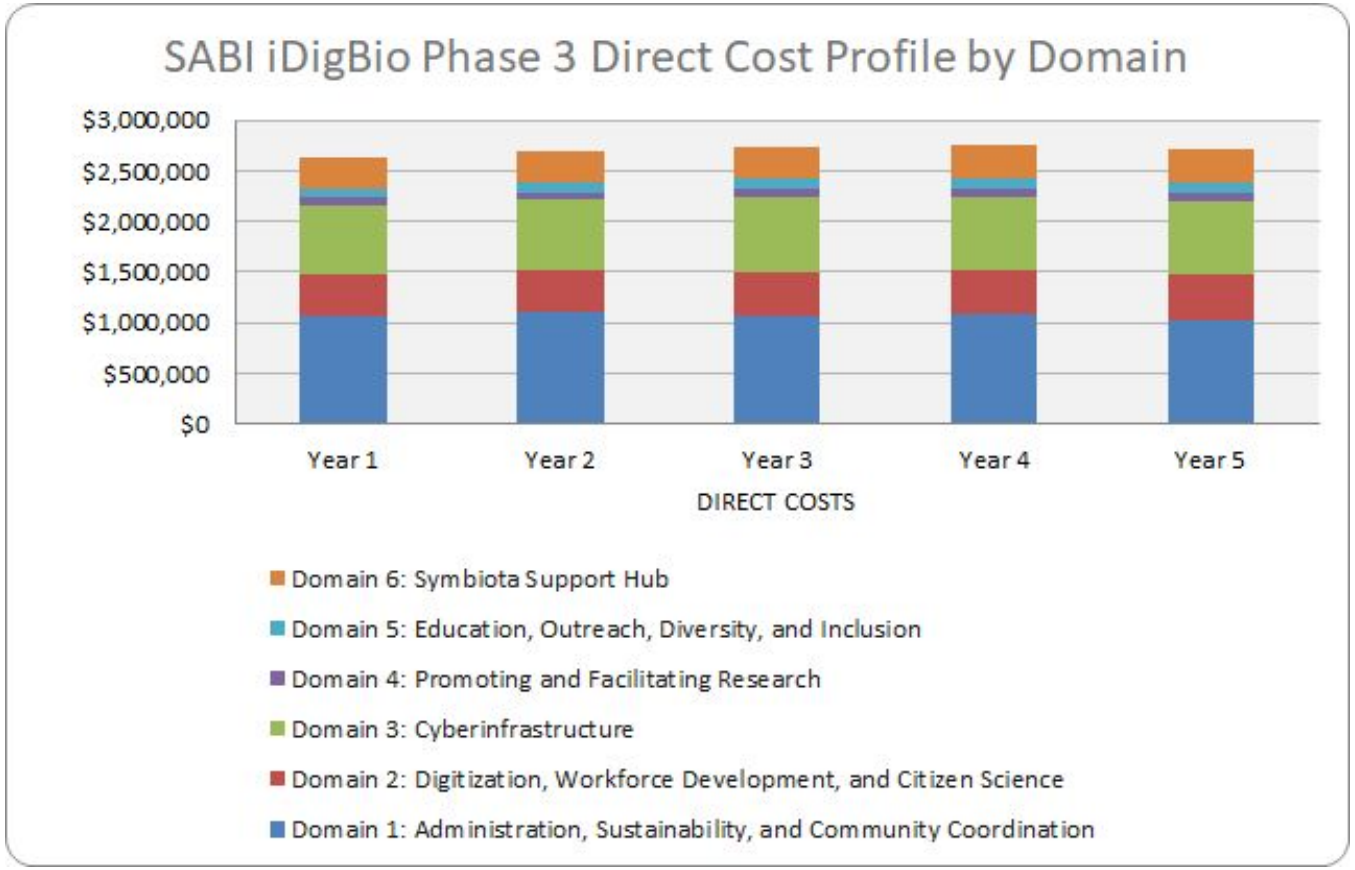


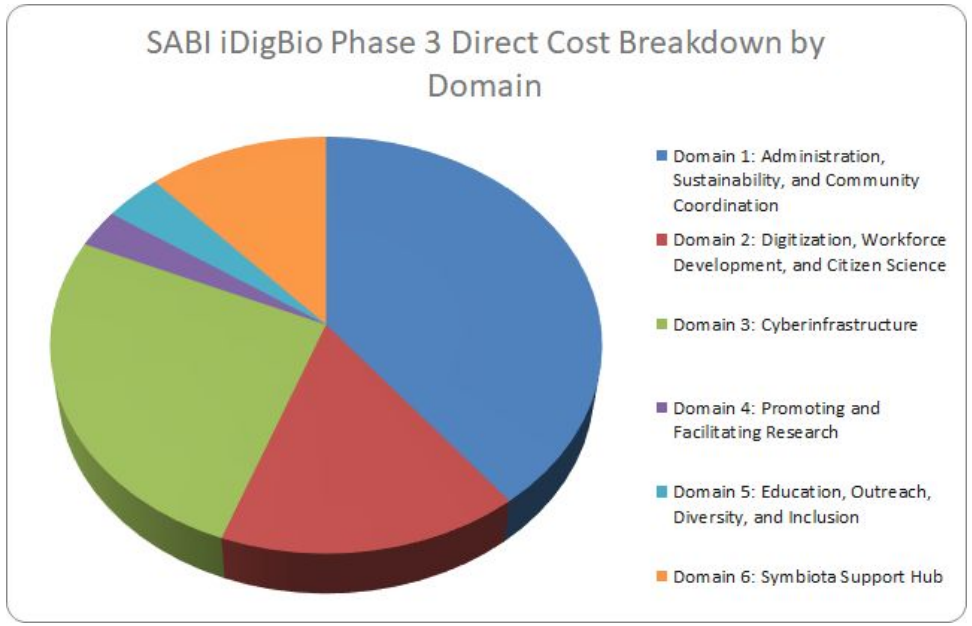
Staffing Highlights

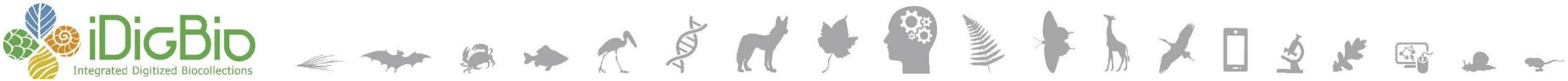


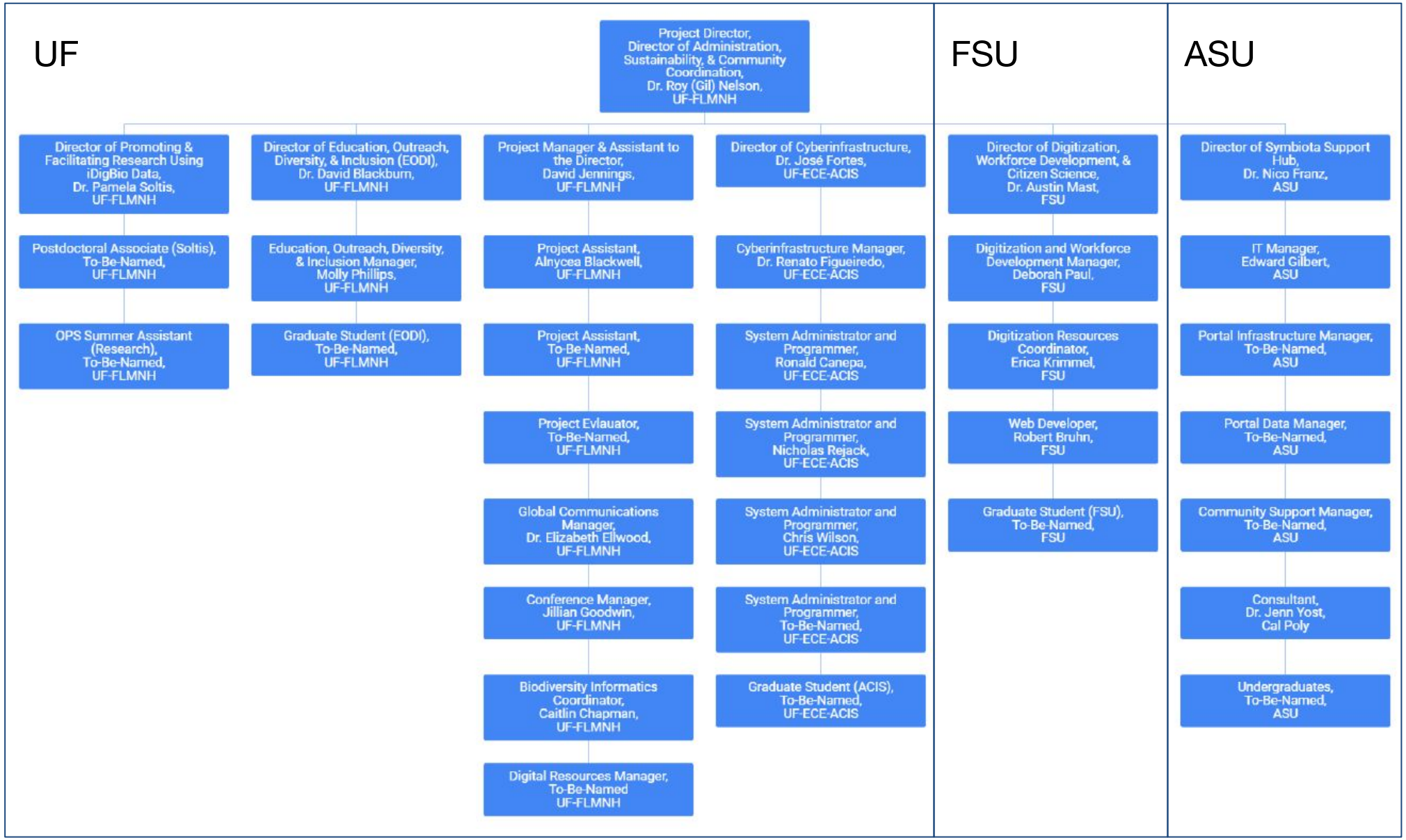


Domain Budget Highlights






iDigBio
Integrated Digitized Biocollections


UF
Project Director, Director of Administration, Sustainability, & Community Coordination, Dr. Roy (Gil) Nelson, UF-FLMNH
Director of Promoting & Facilitating Research Using iDigBio Data, Dr. Pamela Soltis, UF-FLMNH
Postdoctoral Associate (Soltis), To-Be-Named, UF-FLMNH
OPS Summer Assistant (Research), To-Be-Named, UF-FLMNH
Director of Education, Outreach, Diversity, & Inclusion (EODI), Dr. David Blackburn, UF-FLMNH
Education, Outreach, Diversity, & Inclusion Manager, Molly Phillips, UF-FLMNH
Graduate Student (EODI), To-Be-Named, UF-FLMNH
Project Manager & Assistant to the Director, David Jennings, UF-FLMNH
Project Assistant, Alnycea Blackwell, UF-FLMNH
Project Assistant, To-Be-Named, UF-FLMNH
Project Evlauator, To-Be-Named, UF-FLMNH
Global Communications Manager, Dr. Elizabeth Ellwood, UF-FLMNH
Conference Manager, Jillian Goodwin, UF-FLMNH
Biodiversity Informatics Coordinator, Caitlin Chapman, UF-FLMNH
Digital Resources Manager, To-Be-Named UF-FLMNH
Director of Cyberinfrastructure, Dr. José Fortes, UF-ECE-ACIS
Cyberinfrastructure Manager, Dr. Renato Figueiredo, UF-ECE-ACIS
System Administrator and Programmer, Ronald Canepa, UF-ECE-ACIS
System Administrator and Programmer, Nicholas Rejack, UF-ECE-ACIS
System Administrator and Programmer, Chris Wilson, UF-ECE-ACIS
System Administrator and Programmer, To-Be-Named, UF-ECE-ACIS
Graduate Student (ACIS), To-Be-Named, UF-ECE-ACIS
FSU
Director of Digitization, Workforce Development, & Citizen Science, Dr. Austin Mast, FSU
Digitization and Workforce Development Manager, Deborah Paul, FSU
Digitization Resources Coordinator, Erica Krimmel, FSU
Web Developer, Robert Bruhn, FSU
Graduate Student (FSU), To-Be-Named, FSU
ASU
Director of Symbiota Support Hub, Dr. Nico Franz, ASU
IT Manager, Edward Gilbert, ASU
Portal Infrastructure Manager, To-Be-Named, ASU
Portal Data Manager, To-Be-Named, ASU
Community Support Manager, To-Be-Named, ASU
Consultant, Dr. Jenn Yost, Cal Poly
Undergraduates, To-Be-Named, ASU

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

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.

Get Involved!

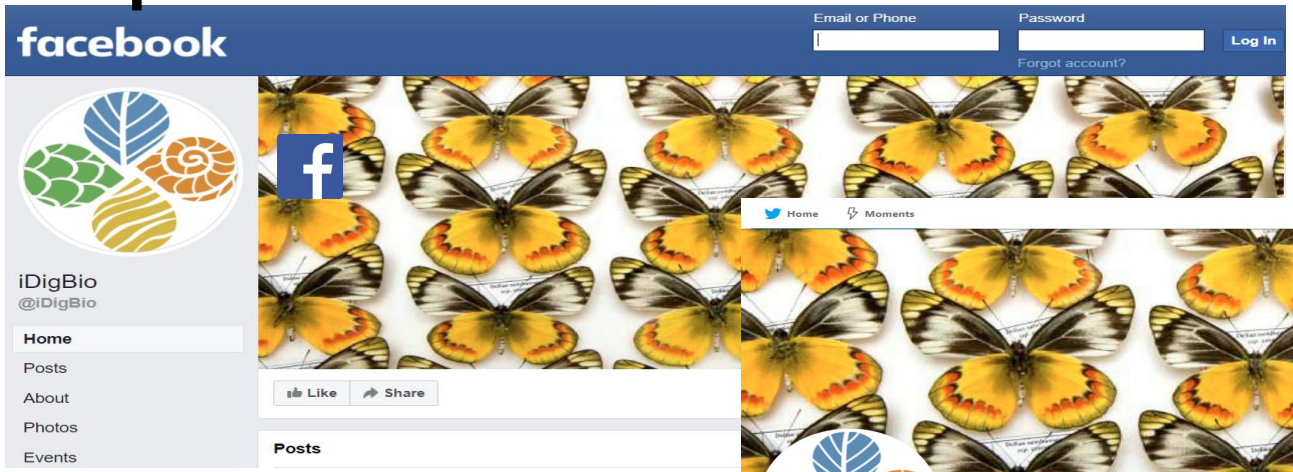
Alnycea Blackwell “Allie”
Project Assistant
ablackwell@floridamuseum.ufl.edu

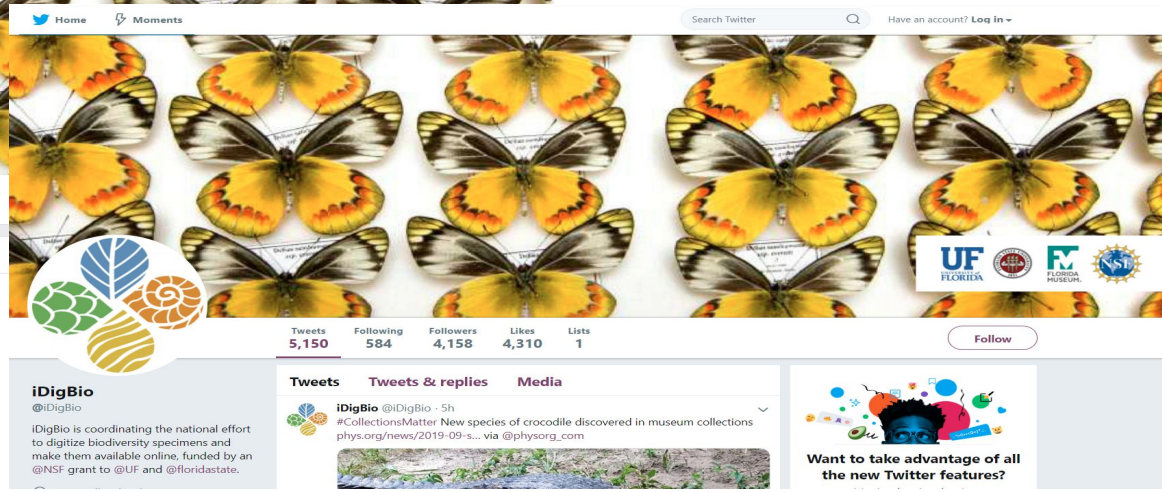
Step 1: Sign up for the iDigBio Newsletter

- 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





vimeo.com/idigbio

idigbio.org/rss-feed.xml

idigbio.org/events-calendar/export.ics

www.idigbio.org/wiki

Step 3: Get involved with a Community Working Group

- 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

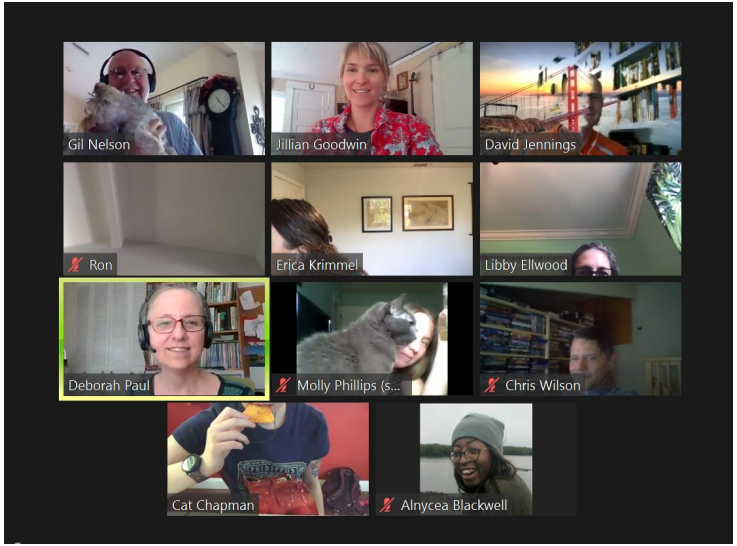
Once a month

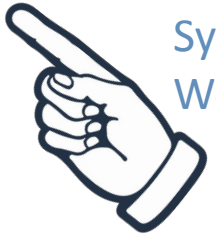
Every Third Wednesday!

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



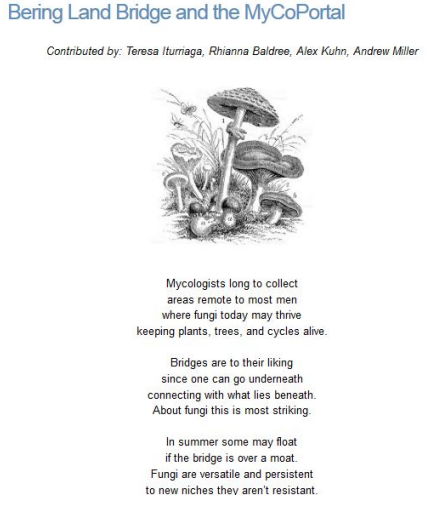


Workshops
Symposia
Webinars

preferably with a microphone!

Step 5: Contribute to the iDigBio website

- 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**



Bering Land Bridge and the MyCoPortal

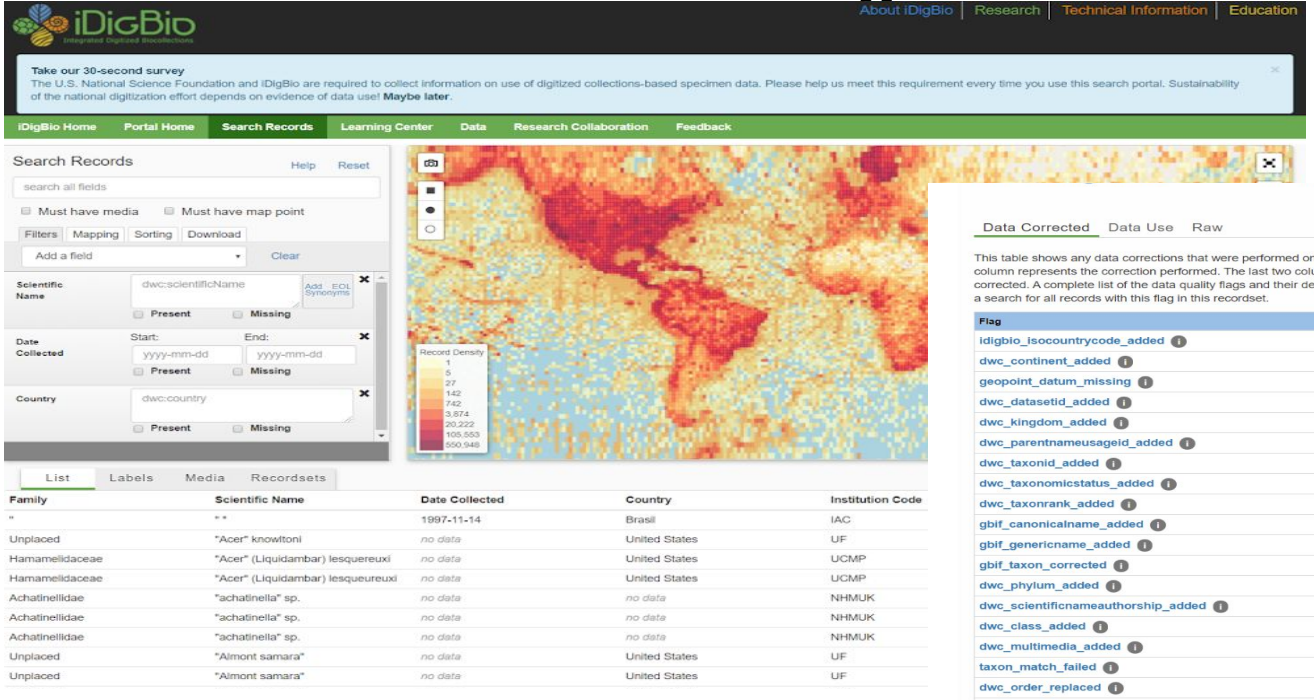
Contributed by: Teresa Iturraga, Rhianna Baldree, Alex Kuhn, Andrew Miller

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

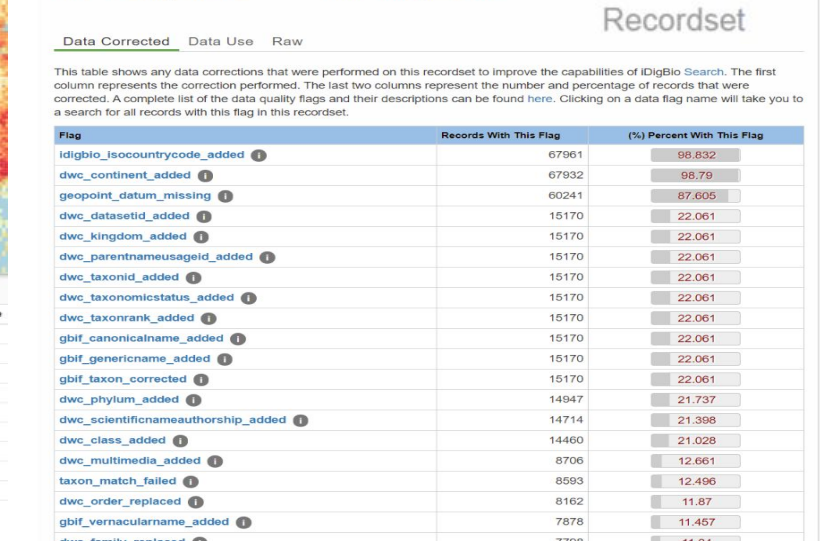
Bridges are to their liking
since one can go underneath
connecting with what lies beneath.
About fungi this is most striking.

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!



data@idigbio.org



Step 7: Collaborate!