iDigBio External Advisory Board Summit 2017

State of iDigBio

Primary issues for EAB

    Sustainability

    Data use

    Metrics

1

Sustainability of iDigBio

Issue: What is to be sustained and at what level?
How will sustainability be achieved?
What are the pressing issues now?

Data Use

Issue: What aggregated data are being used?
How are they being used?
How might use be increased?

Metrics

Issue: What's happening?

iDigBio

1. **iDigBio's cooperative agreement with NSF and budget extend to July 2021**

2. **Budget decreases annually beginning in Year 8**

3

iDigBio

1. ADBC = 10 year program; $100 million

2. TCNs will be funded through year 10,
   with those in the final years likely
   to go at least through 2025

3. iDigBio likely to receive funding by NSF
   to work with these TCNs and others in the
   collections community through 2025
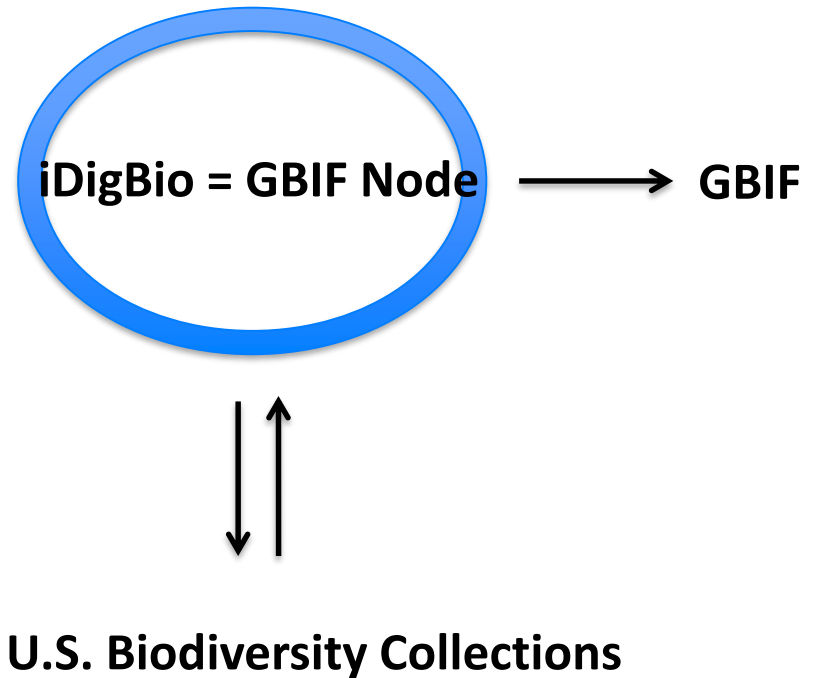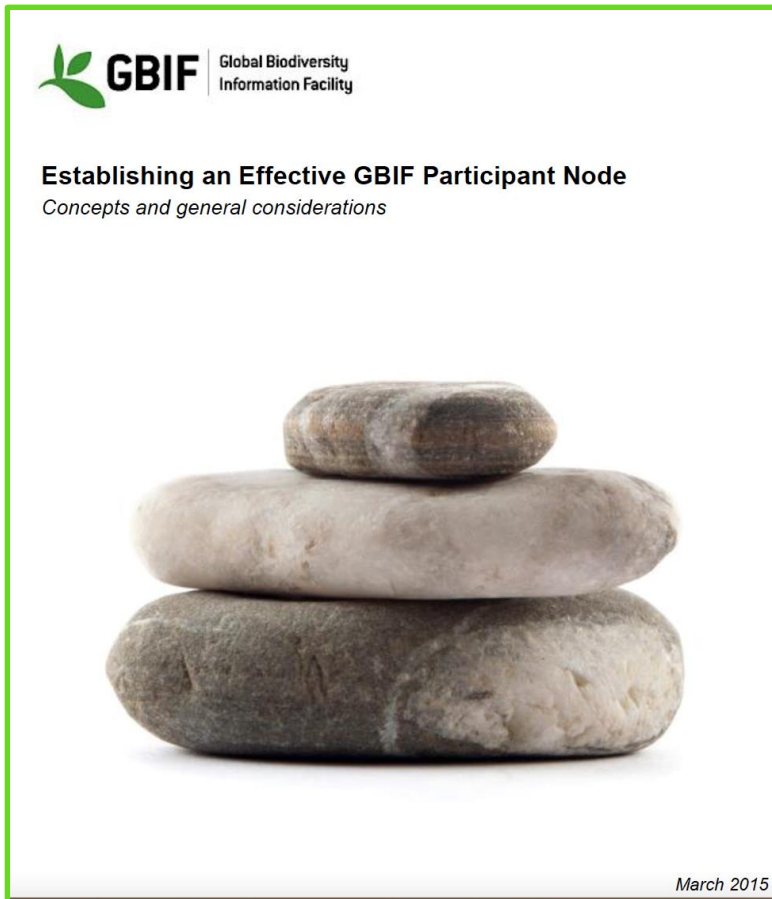
4

Sustainability of iDigBio

A proposal

Sustainability of iDigBio

A proposal

1. Become a U.S. node for GBIF

2. Transition research, E&O to FLMNH

1. Become a U.S. node for GBIF

GBIF | Global Biodiversity Information Facility

Establishing an Effective GBIF Participant Node
Concepts and general considerations

March 2015

iDigBio = GBIF Node → GBIF

U.S. Biodiversity Collections

1. Services provided by a GBIF node

GBIF | Global Biodiversity Information Facility

Establishing an Effective GBIF Participant Node
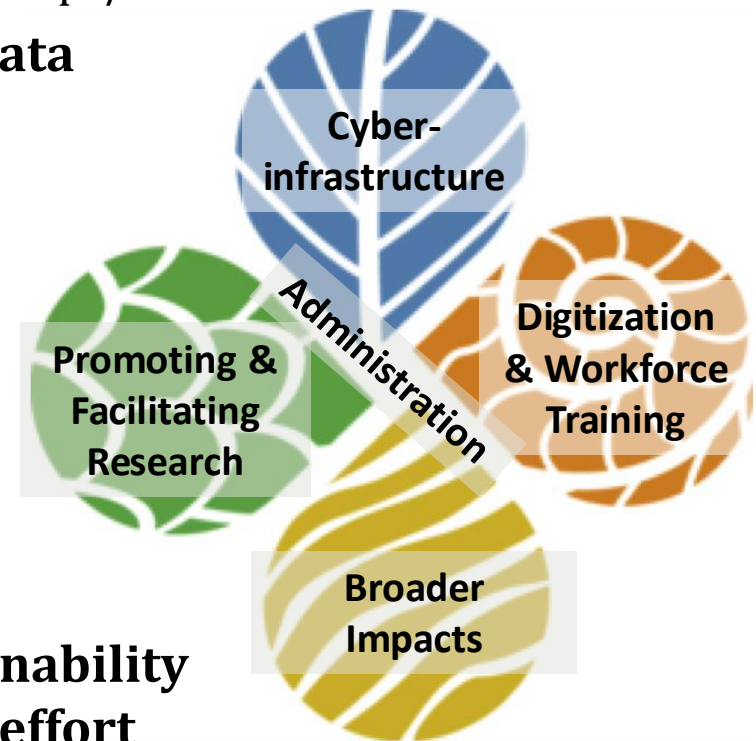*Concepts and general considerations*

March 2015

1. **Support biodiversity data mobilization**
2. **Support biodiversity data management and curation**
3. **Support biodiversity data analysis and use**
4. **Coordinate the landscape of biodiversity-related initiatives including participating in the GBIF network**

What does iDigBio do?

- **Enable digitization of biodiversity collections data**
  - Develop efficient & effective standards & workflows
  - Workforce education & training via workshops/webinars
- **Provide portal access to biodiversity data in a cloud computing environment**
  - Enable data access & discoverability
  - Respond to cyberinfrastructure needs
- **Promote use of biodiversity data to address key environmental and economic challenges**
  - Researchers, educators, general public, policy-makers, etc.
- **Assist in planning for long-term sustainability of the national digitization network & effort**
  - Expand participation: partners, data sources, general public, etc.
  - Proliferate and broaden uses of biodiversity data

What does iDigBio do?

- **Enable digitization of biodiversity collections data**
  - ✓ **1. Support biodiversity data mobilization**

- **Provide portal access to biodiversity data**
  - ✓ **2. Support biodiversity data management and curation**

- **Facilitate use of biodiversity data**
  - ✓ **3. Support biodiversity data analysis and use**
  - ✓ **4. Coordinate biodiversity-related initiatives**

Sustainability of iDigBio

A proposal
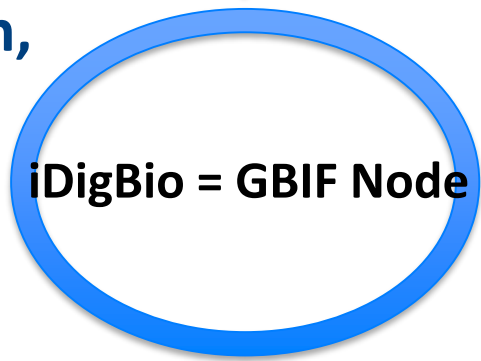
2. Transition research, E&O to FLMNH
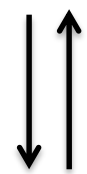
11

Research   Education   Outreach

FLMNH (or Biodiversity Institute or eMuseum)

2. Transition research, E&O to FLMNH

iDigBio = GBIF Node → GBIF

U.S. Biodiversity Collections
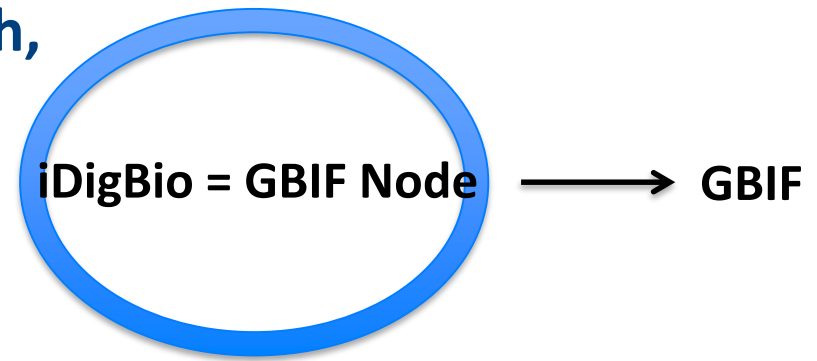
12

Research   **Education**   **Outreach**

Museum and external funding sources

FLMNH (or Biodiversity Institute or eMuseum)

2. Transition research, E&O to FLMNH

iDigBio = GBIF Node → **GBIF**

U.S. Biodiversity Collections

Research   **Education**   **Outreach**

FLMNH (or Biodiversity Institute or eMuseum)

Minimal staffing requirements (iDigBio)

iDigBio = GBIF Node

1. Project manager (Library of archived resources)

3. Database/Portal manager

2. Mobilization manager

4. Training and workforce development

U.S. Biodiversity Collections

14

Overall issues for iDigBio & ADBC

1. Find and engage remaining collections
(How many collections and specimens are there really?)

2. Funding model for ADBC

3. Improving quality of data in iDigBio, especially taxonomy

4. More response from TCNs on research & E&O
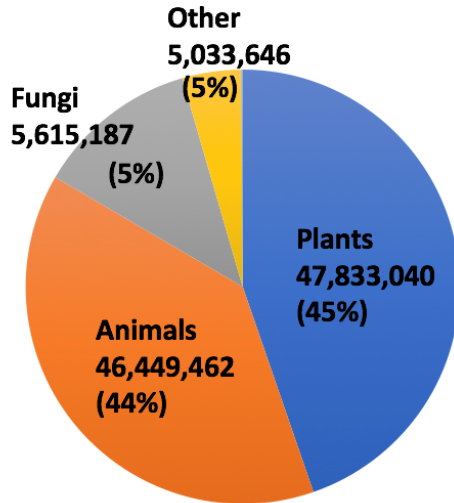
5. International collaboration

15

Overall issues for iDigBio & ADBC

1. Find and engage remaining collections
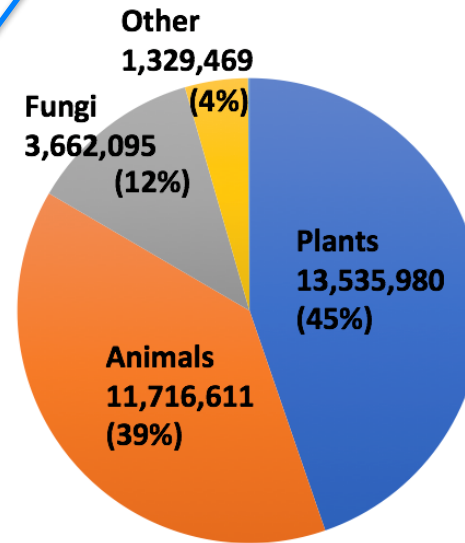   *(How many collections and specimens are there really?)*

Specimen records in iDigBio

Other
5,033,646
(5%)

Fungi
5,615,187
(5%)

Plants
47,833,040
(45%)

Animals
46,449,462
(44%)

Total portal: 106 million

- Plants
- Animals
- Fungi
- Other
- Protista **143,086**

Other
1,329,469
(4%)

Fungi
3,662,095
(12%)

Plants
13,535,980
(45%)

Animals
11,716,611
(39%)

ADBC-funded: 30 million

- Plants
- Animals
- Fungi
- Other
- Protista **30,211**

Plant records in iDigBio

Algae
1,417,705
(3%)

Bryophytes
3,550,644

(7%)

Angiosperms
42,860,549
(90%)
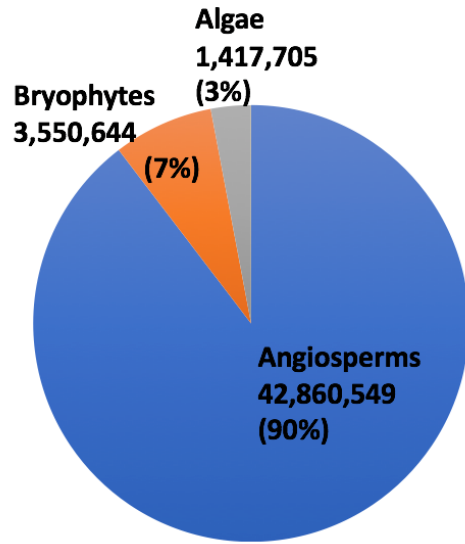
Total portal: 48 million

- Angiosperms
- Bryophytes
- Algae
- Ferns: **4,037**
- Gymnosperms: **105**

Algae
431,965
(3%)

Bryophytes
2,157,253
(16%)

Angiosperms
10,944,017
(81%)
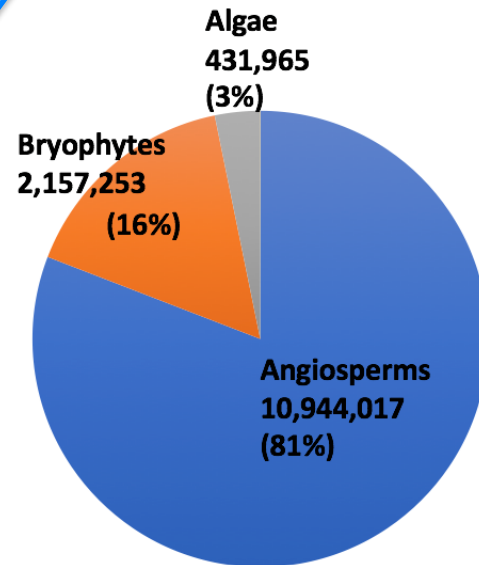
ADBC-funded: 14 million

- Angiosperms
- Bryophytes
- Algae
- Ferns: **2,644**
- Gymnosperms: **101**

Animal records in iDigBio

Marine Inverts
2,128,843
(5%)

Mollusca
5,218,825
(11%)

Chordata
20,975,443
(45%)

Arthropoda
17,179,436
(37%)
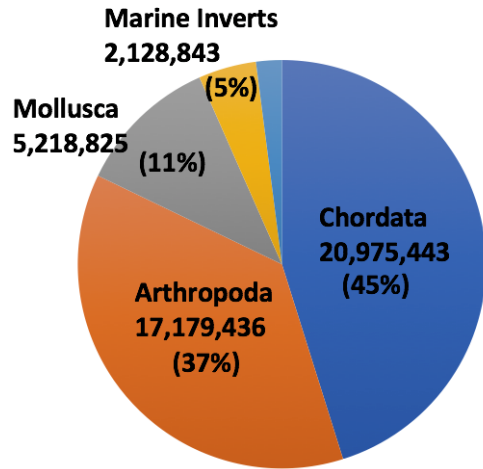
Total portal: 46 million

■ Chordata

■ Arthropoda

■ Mollusca

■ Marine Inverts

■ Non-marine
Inverts: **946,915**

Marine Inverts
693,590
(6%)

Mollusca
2,179,731
(19%)

Arthropoda
5,929,788
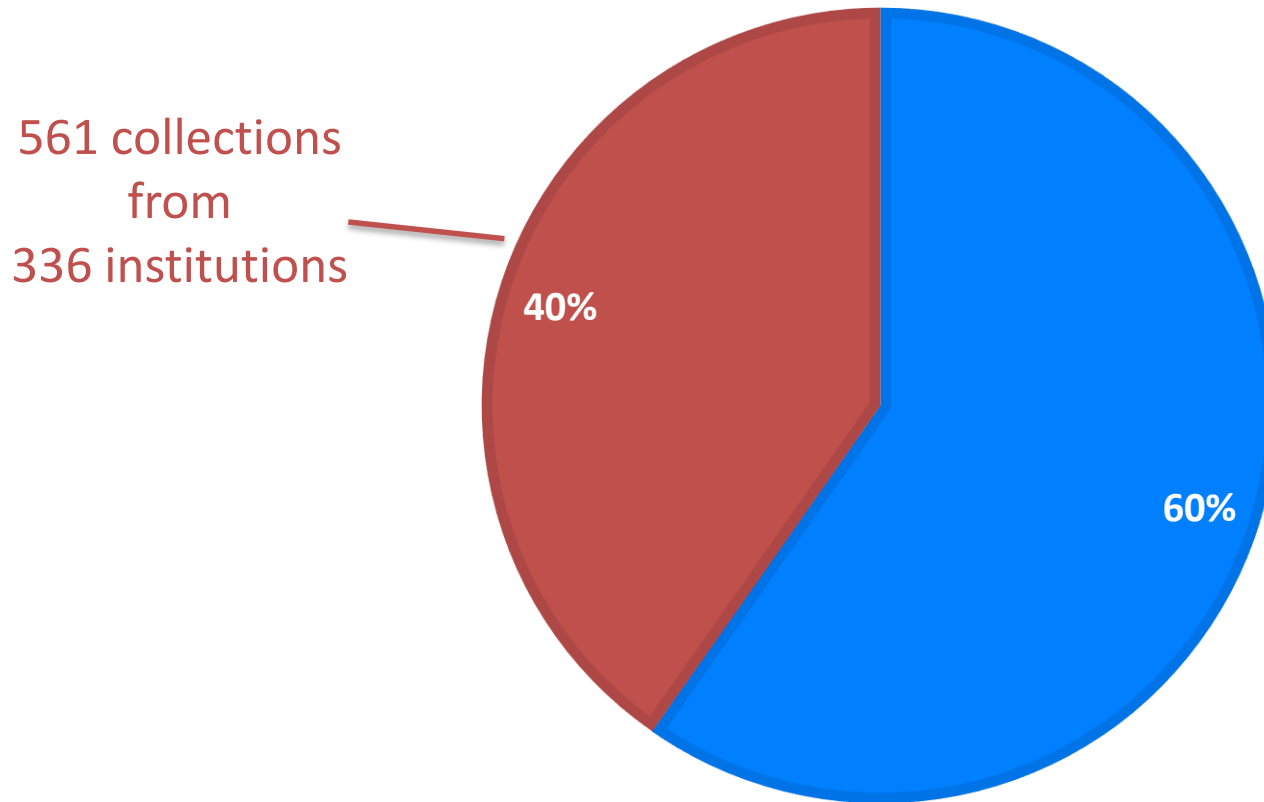**(51%)**

Chordata
2,806,614
(24%)
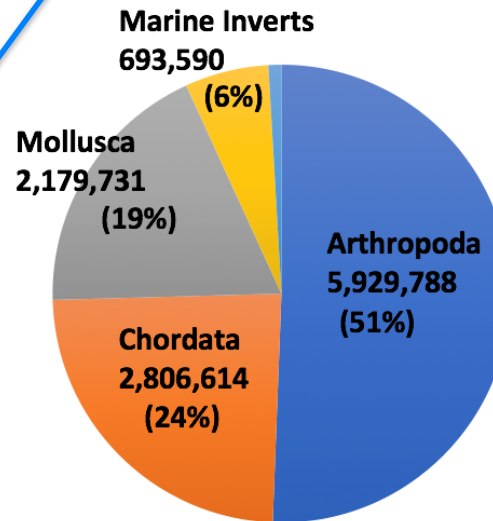
ADBC-funded: 12 million

■ Arthropoda

■ Chordata

■ Mollusca

■ Marine Inverts

■ Non-marine
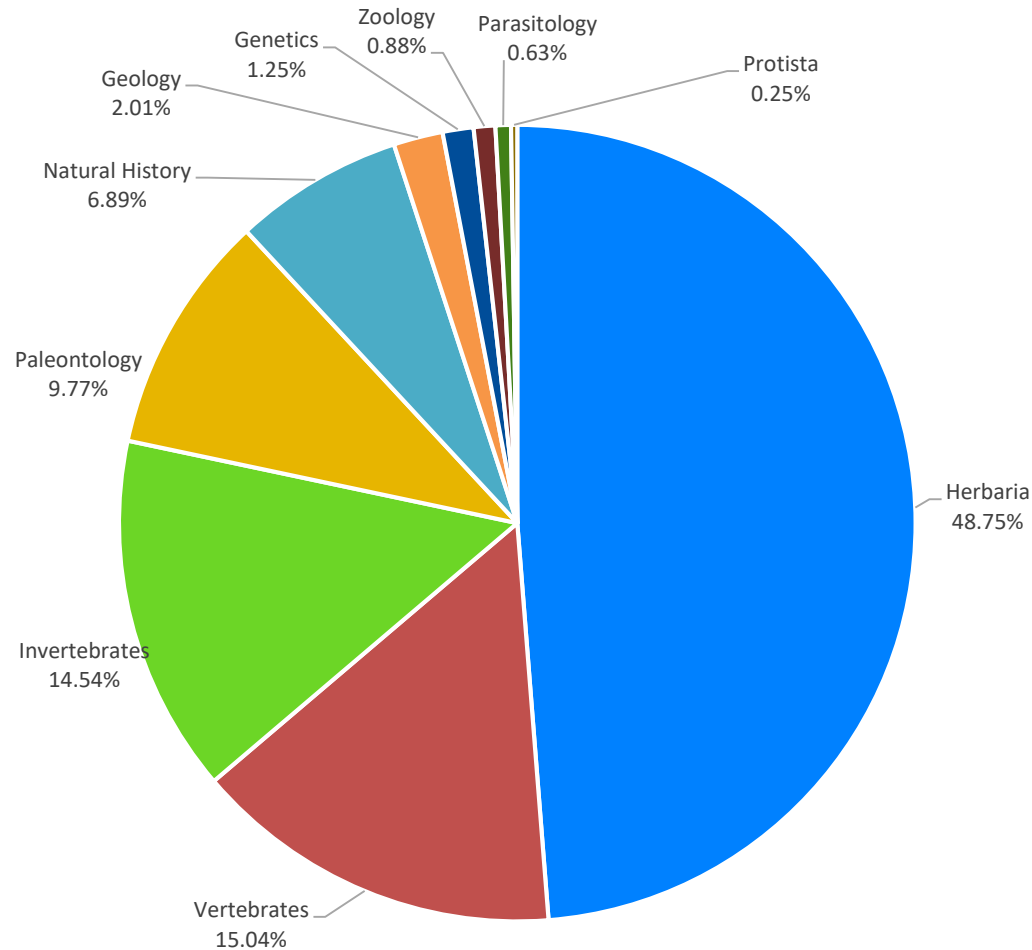Inverts **106,888**

Progress in National Digitization Effort

1,386 TOTAL KNOWN U.S. COLLECTIONS

■ Have NOT provided data in iDigBio ■ HAVE provided data in iDigBio

561 collections from 336 institutions

40%

60%

What We Know About What We Don't Have

Breakdown of Collections in iDigBio Catalog
that have NOT yet Contributed Data in iDigBio

- Genetics 1.25%
- Zoology 0.88%
- Parasitology 0.63%
- Geology 2.01%
- Protista 0.25%
- Natural History 6.89%
- Paleontology 9.77%
- Herbaria 48.75%
- Invertebrates 14.54%
- Vertebrates 15.04%

Future:

iDigBio is working with 561 collections in 336 institutions

Of how many?
Currently 1,386 collections in iDigBio's Catalog of U.S. Collections (561/1,386 collections in U.S. = 41%)

Specimen records: 106 million
(315 million specimens of 1 billion) *one-third?*

Remaining institutions are <u>mostly</u> small.
Many remaining collections are not!

Overall issues for iDigBio & ADBC

1. Find and engage remaining collections
(How many collections and specimens are there really?)

2. Funding model for ADBC

3. Improving quality of data in iDigBio, especially taxonomy (next "moonshot" for collections?)

4. More response from TCNs on research & E&O

5. International collaboration

5. **International collaboration**

1. Node for GBIF

2. Frankfurt meeting white paper to be submitted to:

"**National Science Foundation Seeks Community Input on Accelerating Research through International Network-to-Network Collaboration**"