iDigBio is the Coordinating Center for the NSF’s Advancing Digitization of Biodiversity Collections Program

Larry M. Page, Director
Goal of ADBC

- To remove the inaccessibility of specimen-based data in natural history collections through digitization: information online for researchers, educators, policymakers, etc.

$100 million over 10 years non-federal collections (1st funding in 2011)
What does iDigBio do?

- Engage the collections community and facilitate digitization of biodiversity collections data

- Provide portal access to biodiversity data in a cloud computing environment

- Facilitate use of biodiversity data to address key environmental and economic challenges

- Plan for long-term sustainability of the national digitization network & effort
Thematic Collections Networks (TCNs)

Groups of non-federal institutions

- Digitize specimen-based data around a research theme

Digitizing = databasing, georeferencing, imaging

- Provide data to iDigBio’s search portal
Thematic Collections Networks (TCNs)

- 18 TCNs funded
- 495 collections at 289 institutions
ADBC National Digitization Network

iDigBio Portal has 1,581 recordsets containing 99M records for ≈300M specimens with 21M associated media records.

612 participating collections in 313 institutions.
Progress

- Engage the collections community
- Facilitate digitization and mobilization of data
- Provide a search portal
- Promote research and outreach
Cumulative Data Ingested by Year

- **2015**
  - Recordsets: 1M
  - Specimen Records: 1G
  - Media Records: 100M

- **2016**
  - Recordsets: 10k
  - Specimen Records: 1G
  - Media Records: 100M

- **2017**
  - Recordsets: 1k
  - Specimen Records: 1G
  - Media Records: 100M

Legend:
- Blue: Recordsets
- Black: Specimen Records
- Green: Media Records
Top 10 Kingdoms and Families in Media Records

- Plantae, plantae
- Eubacteria
- Protozoa
- Prototista
- Bacteria
- Protista
- Chromista
- Fungi
- Animalia
- Plantae
Progress

- Engage the collections community
- Facilitate digitization and mobilization of data
- Provide a search portal
- Promote research and outreach
Engage the collections community
Facilitate digitization and mobilization of data

- 3,184 participants from 482 institutions at 85 workshops
- 2,586 participants at 78 webinars
Future

- **Continue to engage collections community**
  - Go international?
- **Improve quality of the data**
  - *(e.g. taxonomy; linking to other data)*
- **Reward data providers**
  - *(Credit for use in publications, etc.)*
- **Demonstrate value through research and outreach**
  - *(e.g., this conference)*
Small silos of data (institutional collections) that were difficult to access are connected to make data much more accessible.
Good news:
Small silos of data (e.g., institutional collections) that were difficult to access are connected to make data much more accessible.

Bad news:
Small silos replaced by larger silos that are difficult to cross-search to combine data in ways required for integrative research.
International Collaborations

*Exploring Synergies for Biodiversity Information Systems*

- Compare the structure and goals of the various BIS
- Enhance collaboration and interoperability,
- Globally connect information for research and outreach
Collaboration!