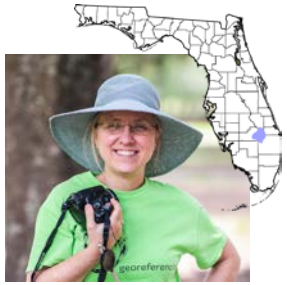
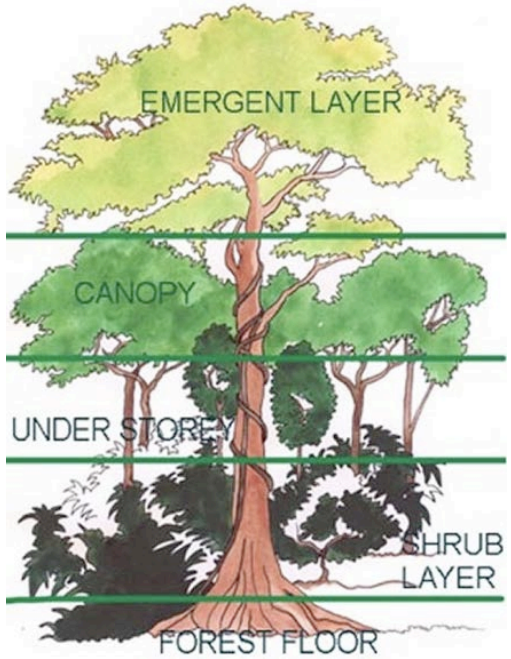
Dashboard Confessionals: Looking at measuring the success of a national scale digitization program



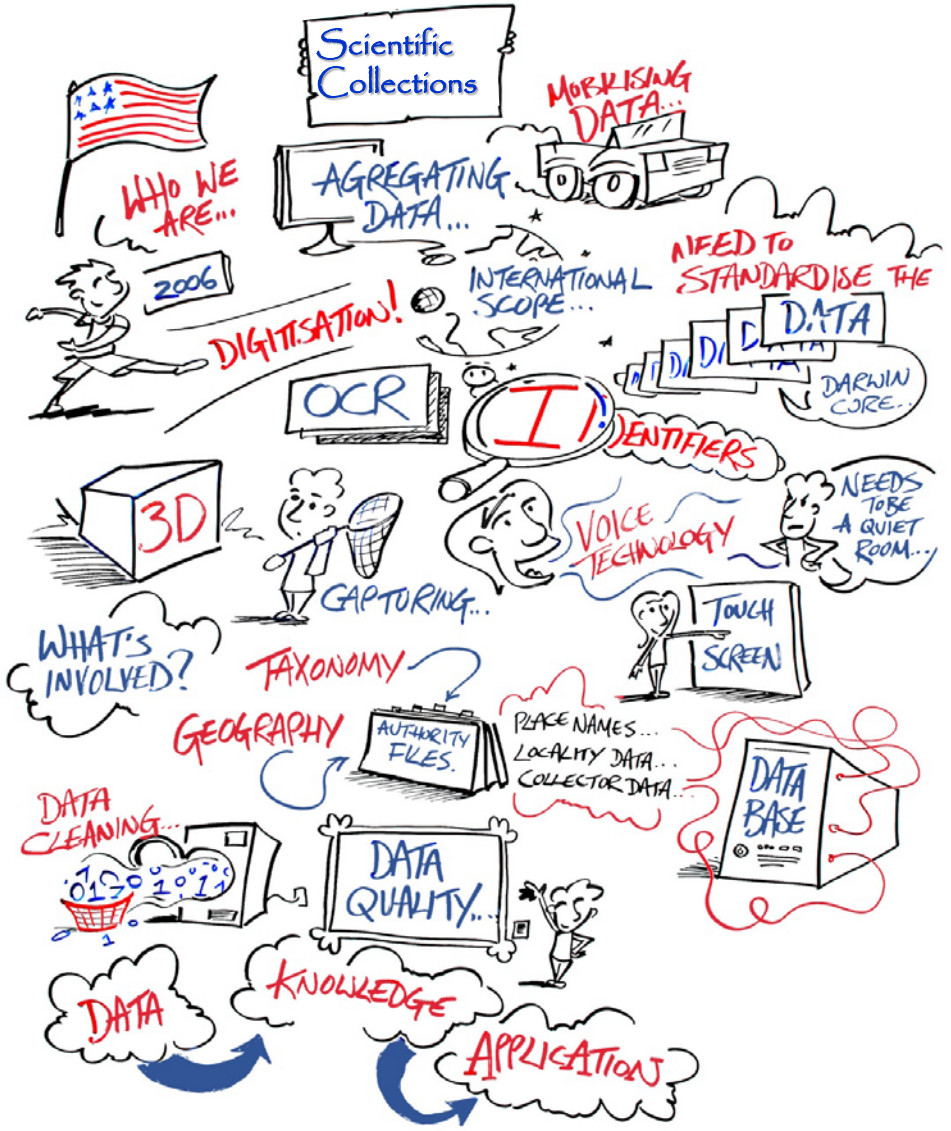
Deborah Paul
iDigBio, Florida State University
ChicagoSPNHC2019 29 May 2019
@idbdeb @iDigBio #metricsmatter



iDigBio is funded by a grant from the National Science Foundation's Advancing Digitization of Biodiversity Collections Program (Cooperative Agreement EF-1115210). 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. All images used with permission or are free from copyright.

4 Chapters – aggregator view

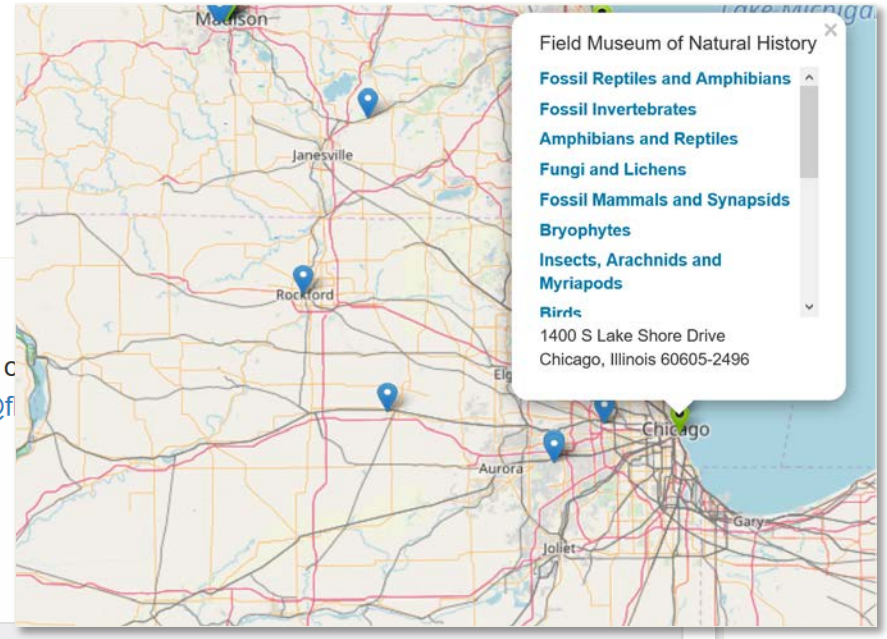
- Building a US collections list
- Aggregator metrics
- Everyone needs credit
- Carrots and benchmarks
  - your survey input
  - your participation
  - **strategic plans require benchmarks**



Chapter One

- Building a US Collections Resource
  - at iDigBio
  - for the Advancing the Digitization of Biological Collections (ADBC) vision

Building a US Collections Resource



Collections

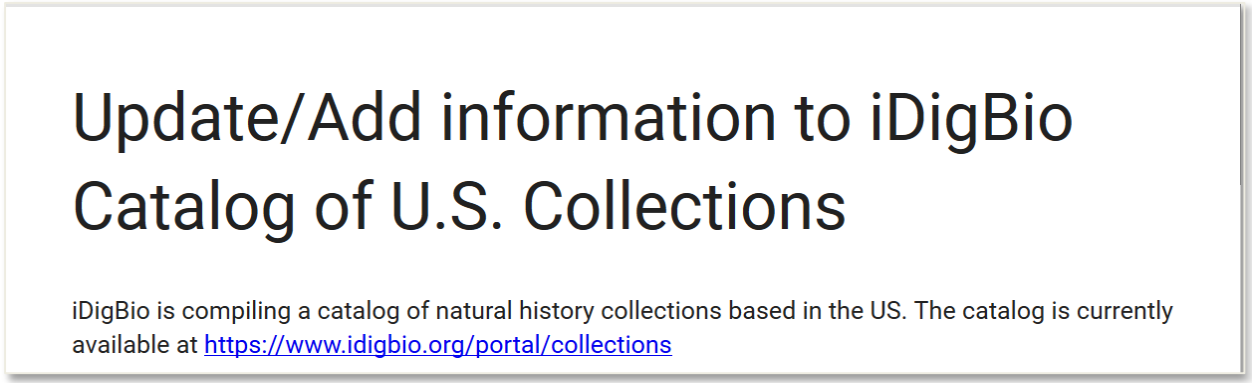
This is intended to be a comprehensive list of all natural history collections in the United States of America. If you know c
list, please complete this form. If you have any questions or encounter any difficulty, please email Kevin Love at klove@f

Show Map of Collections

field museum

Institution ▲	Collection	Contact	Contact Role	Update/Add Information	Show On Map	Recordsets
Field Museum of Natural History	Fossil Reptiles and Amphibians	William F. Simpson	Collections Manager, Fossil Vertebrates	Update/Add Information	Show On Map	
Field Museum of Natural History	Fossil Invertebrates	Paul Mayer	Collections Manager	Update/Add Information	Show On Map	Search Recordset(s)
Field Museum of Natural History	Amphibians and Reptiles	Alan Resetar	Collections Manager	Update/Add Information	Show On Map	Search Recordset(s)
Field Museum of Natural History	Fungi and Lichens	Robert Lücking	Collections Manager & Adjunct Curator	Update/Add Information	Show On Map	Search Recordset(s)

Building a US Collections Resource



Collections Metadata

institution	recordset Query	types
institution code	…	contacts
collection	who	provenance
collection code	what	geocoordinates
recordsets	where	…

Stakeholders

- You (as an institution)
- You (as an individual)
- Aggregators
- Funders
- Your peers
- Researchers

https://www.idigbio.org/portal/collections

Building a US Collections Resource
some reasons for this list – use case examples

- iDigBio uses for this list
  - **track** progress of contributors (automated if possible)
  - **connect** specimen recordsets to institutional collections
  - **plan** programs to **mobilize** collections not yet publishing data
  - as a **guide** when planning to reach historically black colleges and universities, for example

Building a US Collections Resource
some reasons for this list – use case examples

- UK Small and Regional Collections
  - **planning visit** to US collections
  - using list to develop visit itinerary
    - seeking models
      - sustainability, digitization, citizen science, outreach, …
    - expertise sharing
    - looking for collaborators
    - methods, workflows, protocols sharing

Building a US Collections Resource
challenges and some (as yet) unknowns

- Keeping US Collections data up-to-date
  - **– requires human upkeep**
  - – potential for more automation
  - – responsibility long term

field museum

Institution ▲	Collection	Contact	Contact Role	Update/Add Information	Show On Map	Recordsets
Field Museum of Natural History	Fossil Reptiles and Amphibians	William F. Simpson	Collections Manager, Fossil Vertebrates	Update/Add Information	Show On Map	
Field Museum of Natural History	Fossil Invertebrates	Paul Mayer	Collections Manager	Update/Add Information	Show On Map	Search Recordset(s)
Field Museum of Natural History	Amphibians and Reptiles	Alan Resetar	Collections Manager	Update/Add Information	Show On Map	Search Recordset(s)
Field Museum of Natural History	Fungi and Lichens	Robert Lücking	Collections Manager & Adjunct Curator	Update/Add Information	Show On Map	Search Recordset(s)

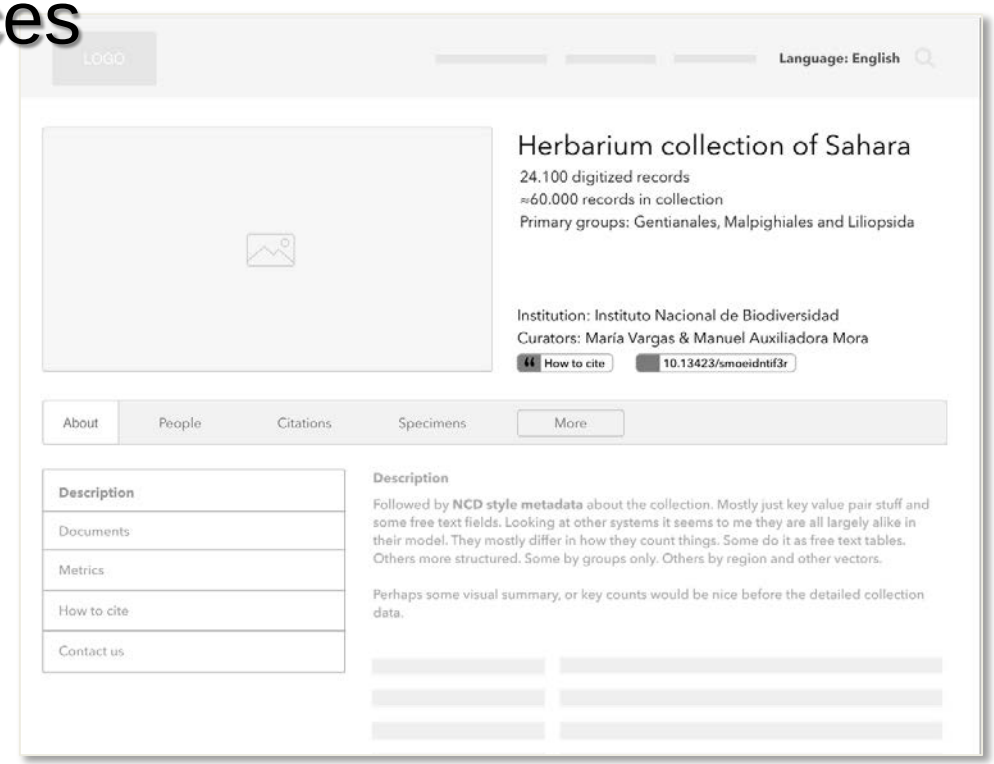
https://www.idigbio.org/portal/collections

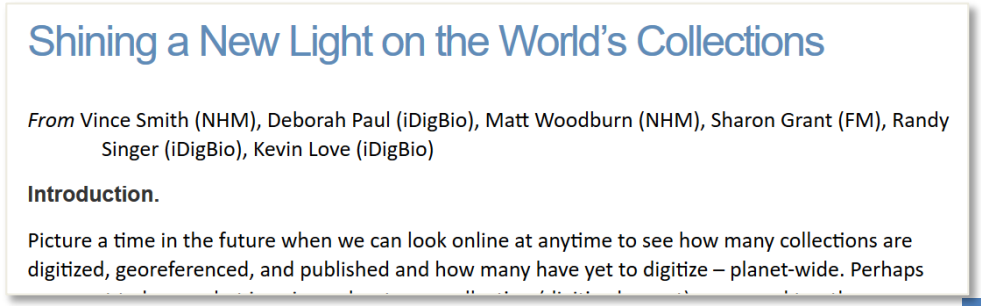
Building a US Collections Resource
potential carrots to entice upkeep

- discovery
- awareness
- outreach for planning
- … even … funding
- perhaps badging / branding
  - strengthen value of community membership

Building a World of Collections' Resources
what's in the works

- A Global Collections Resource
  - Re-imagining GRBio
  - BIS TDWG Collection Descriptions Task Group
  - interoperable vision
  - automate where possible
  - how will this resource work?
- how will these resources work together?
  - Central European Taxonomic Facilities (CETAF)
  - DiSSCo Collection Digitization Dashboards
  - Index Herbariorum
  - US Collections List at iDigBio
- Read more about more such projects
  - Shining a New Light on the World's Collections




Shining a New Light on the World's Collections

From Vince Smith (NHM), Deborah Paul (iDigBio), Matt Woodburn (NHM), Sharon Grant (FM), Randy Singer (iDigBio), Kevin Love (iDigBio)

Introduction.

Picture a time in the future when we can look online at anytime to see how many collections are digitized, georeferenced, and published and how many have yet to digitize – planet-wide. Perhaps

Chapter Two

- iDigBio as a Aggregator – insights
  - challenges and observations
  - benefits
  - what's next for data feedback to aggregators?

iDigBio as an Aggregator

Providers get feedback

Flag	
idigbio_isocountrycode_added	61275
dwc_country_replaced	58013
dwc_datasetid_replaced	35508
dwc_parentnameusageid_added	35508
dwc_taxonid_added	35508
dwc_taxonomicstatus_added	35508
gbif_canonicalname_added	35508
gbif_genericname_added	35508
gbif_taxon_corrected	35508
dwc_taxonrank_added	35490
dwc_scientificnameauthorship_added	33559

Why add these data steps in the first place?

Mainly, it's about making biodiversity data searchable in the aggregate.

119,000,000+

iDigBio as an Aggregator
challenges and insights from providing data feedback

- Data providers get insights into their data
- Raw data in downloads provides check
- Aggregators learn about metrics needs
  - note *current ADBC-iDigBio survey out for your input*
- Feedback* may
  - be difficult for provider to implement
  - not be deemed relevant to local needs
  - be in error
  - be enlightening
- Why not just provide the raw data? (which aggregators do)
  - discoverability, need to be able to search aggregate
- Taxonomy is particularly challenging
  - Catalogue of Life+ may help

see SPNHC-TDWGNZ DQ Feedback Symposium

iDigBio as an Aggregator – benefits of aggregation
possible metrics and measures

Attribution
Citation
Data checks against standards
Data enhancement opportunity
Data used to improve data
Discoverability
Discover collection uniqueness
Discover holdings
Gap knowledge
Greater audience
Linked data
LOCKSS
Potential for increased data use
Searchability
Visibility

iDigBio as an Aggregator
the future of feedback

- harmonize data checks
  - the Global Biodiversity Information Facility
  - the Atlas of Living Australia
  - iDigBio, and others plan use same data tests built by the
  - Biodiversity Information Standards (TDWG) Data Quality Task Group

…. but …. a challenge

- come to the SPNHC SIG on Friday: Collection Management Solutions Special Interest Group Meeting to discuss (at least)
  - potential for **applying these Data Quality rules at the source**
  - figuring out how to share more linked / enhanced data

Chapter Three

- Everyone needs metrics
  - data providers/collections, aggregators, people

Everyone needs use metrics
data providers / collections

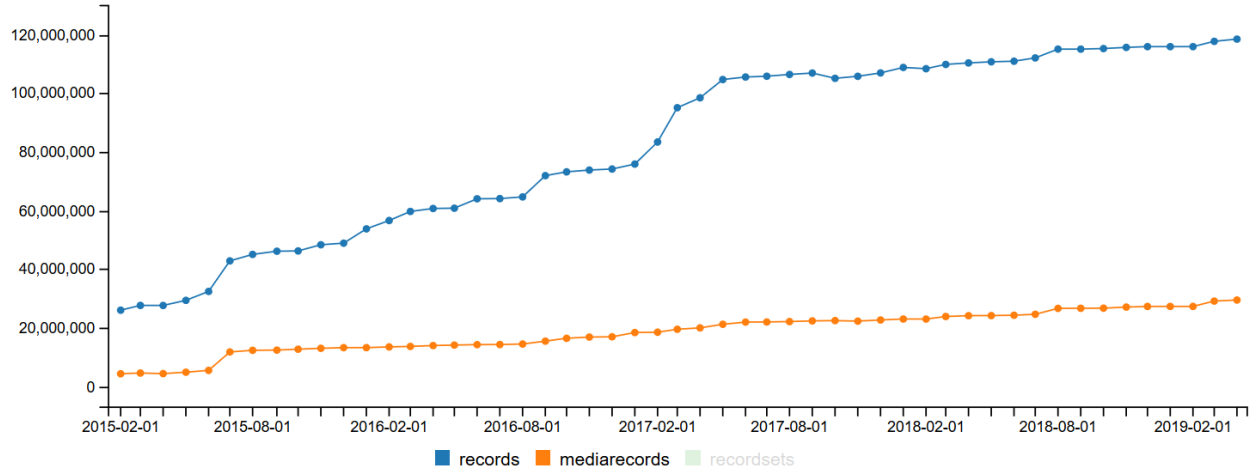
Data Corrected | Data Use | Raw

The table below represents monthly iDigBio portal use statistics for this recordset. ***Search*** indicates in how many instances a record from this recordset matched a search query. ***Download*** indicates in how many instances a record from this recordset was downloaded. ***Seen*** indicates in how many instances a record from this recordset appeared (visually) in the search results in a browser window. ***Records Viewed*** and ***Media Viewed*** indicate how many specimen and media records were opened and viewed in full detail. Note: Monthly statistics aggregation began on Jan 15th 2015; therefore, the month of (01 / 2015) represents approximately half a month of statistics reporting.

Month of	Search	Download	Seen	Records Viewed	Media Viewed
01 / 2019	6,265,608,794	207,222	7,400	17	5
02 / 2019	7,954,090,947	125,842	13,567	32	8
03 / 2019	8,452,334,889	43,452	144,035	27	7
04 / 2019	7,958,074,617	2,352	8,960	50	22
05 / 2019	5,425,003,833	34,890	211,117	16	12

Everyone needs use metrics *aggregators*

Data Ingestion



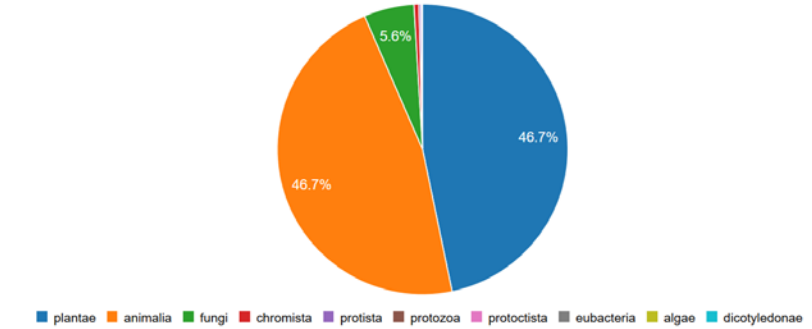
Cumulative:☑

Start: 01/16/2015 **End:** 04/30/2019

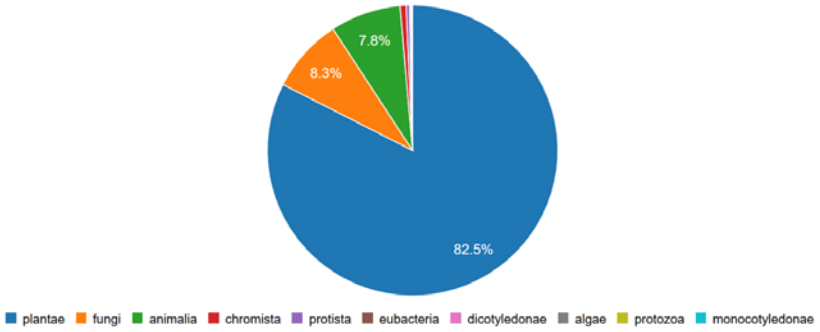
Total Searches:	**816596681**	Total Downloads:	**355417**
Total records searched: 49706232455303			
Total records seen:	**645595318**	Total specimen records viewed:	**112262961**
Total records downloaded:	**838917975**	Total media records viewed:	**14956006**

Taxonomic Coverage

Records

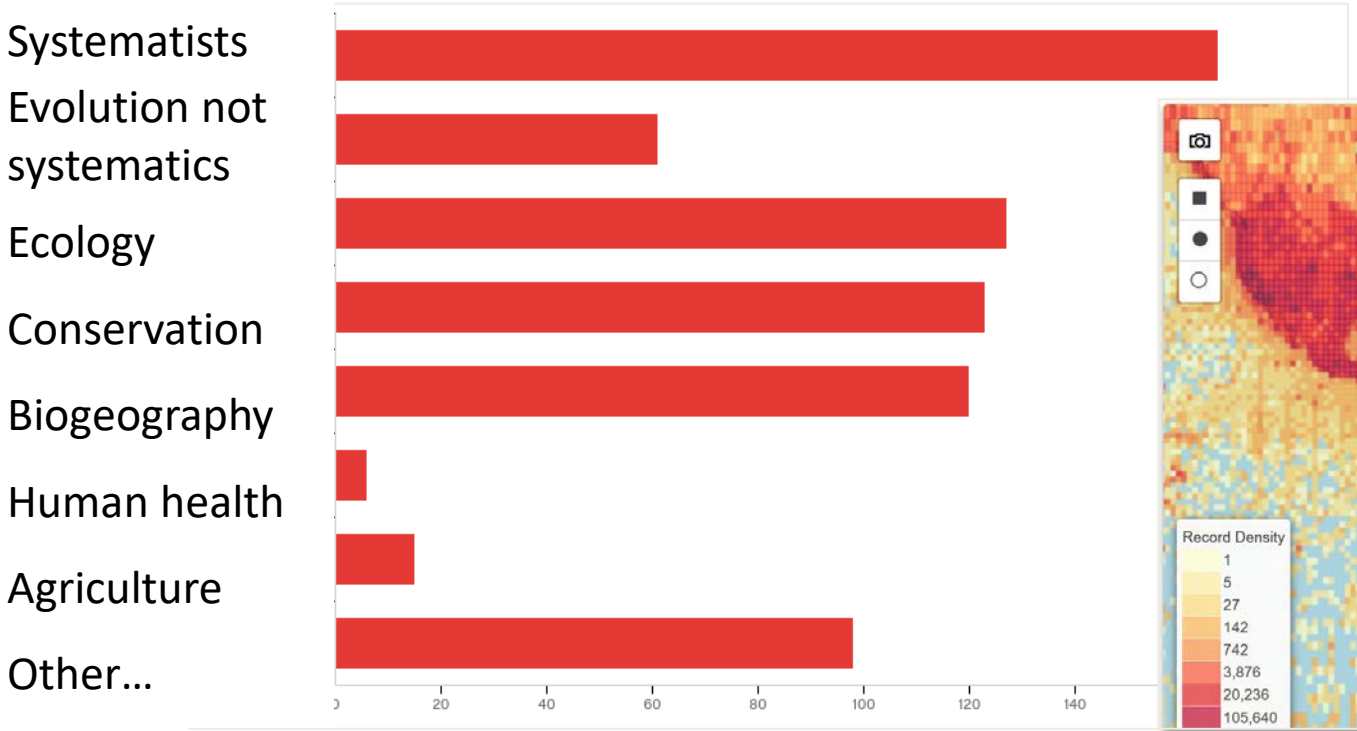


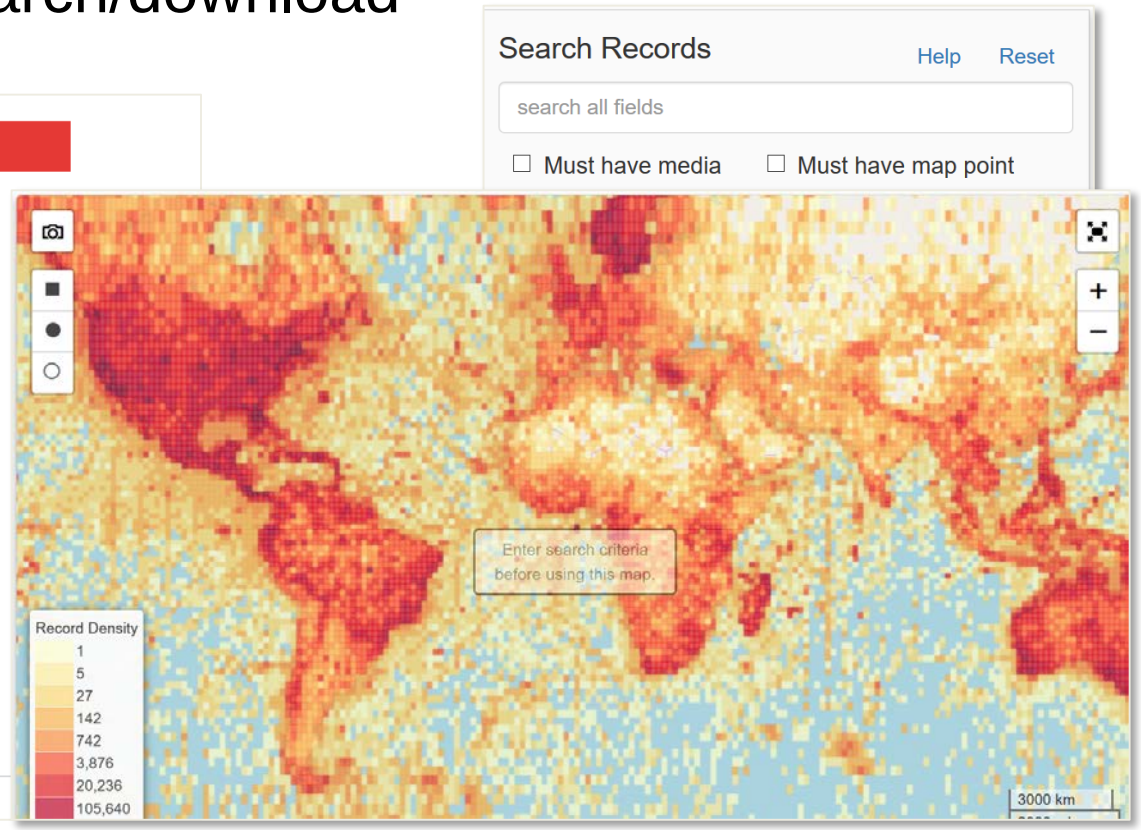
Media



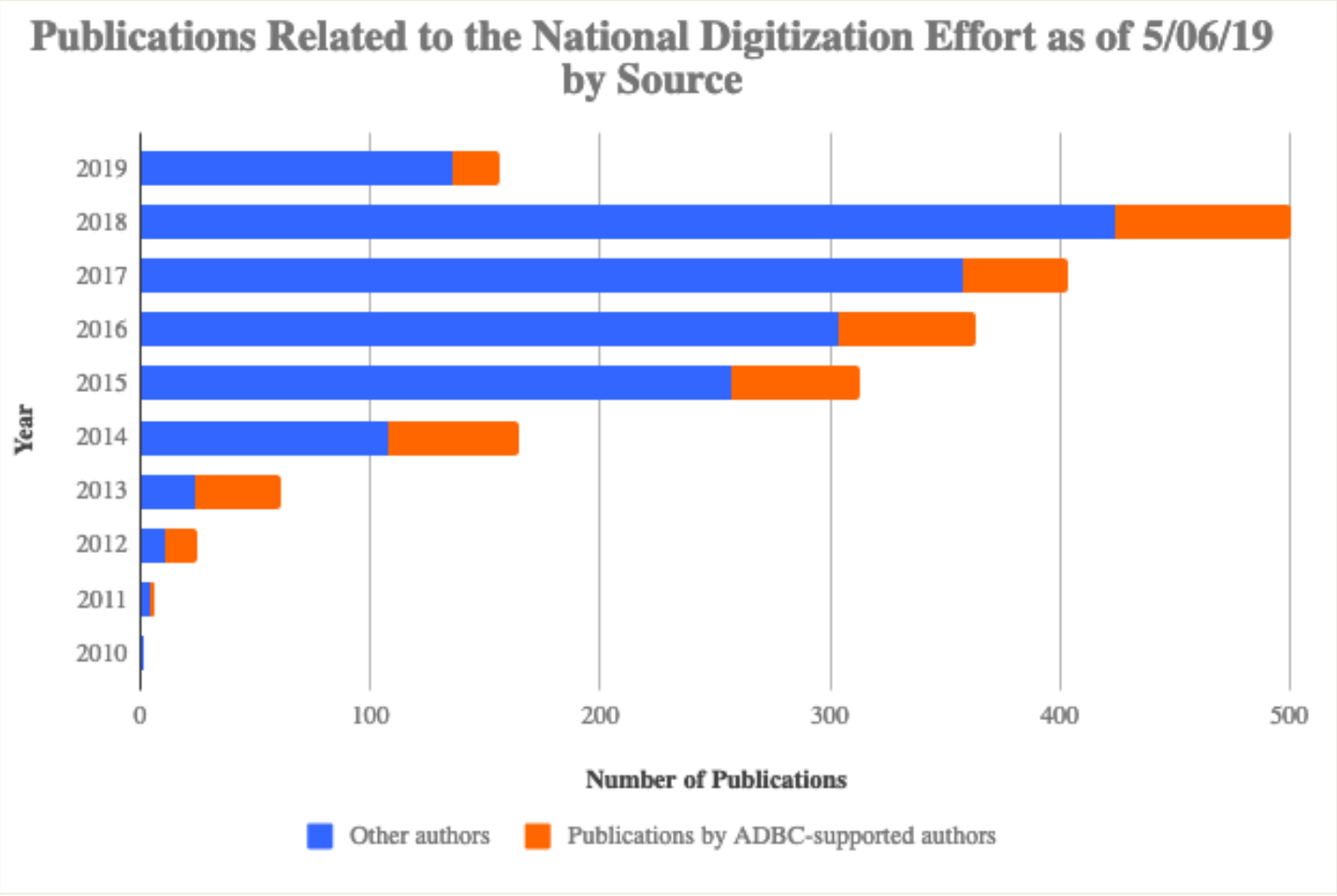
Everyone needs use metrics
aggregators – stakeholder metrics

searchers asked to share purpose of search/download





Everyone needs use metrics
publications increasing



Main Points

- Chapter One: Building a US Collections Resource
  - the world needs your collections metadata
  - what are the carrots needed to get this data?
- Chapter Two: iDigBio as an Aggregator
  - work with aggregators to get the metrics you need
  - help develop / discover new metrics
- Chapter Three: Everyone needs metrics
  - let's build a baseline, a benchmark resource
- Chapter Four:
  - your input needed (surveys, working groups, …)
  - **strategic plans require benchmarks**
  - **(Talanoa dialogue), where are we?, where do we want to go?, how do we get there?**

Got metrics?

Let's share them to get the numbers and recognition we need

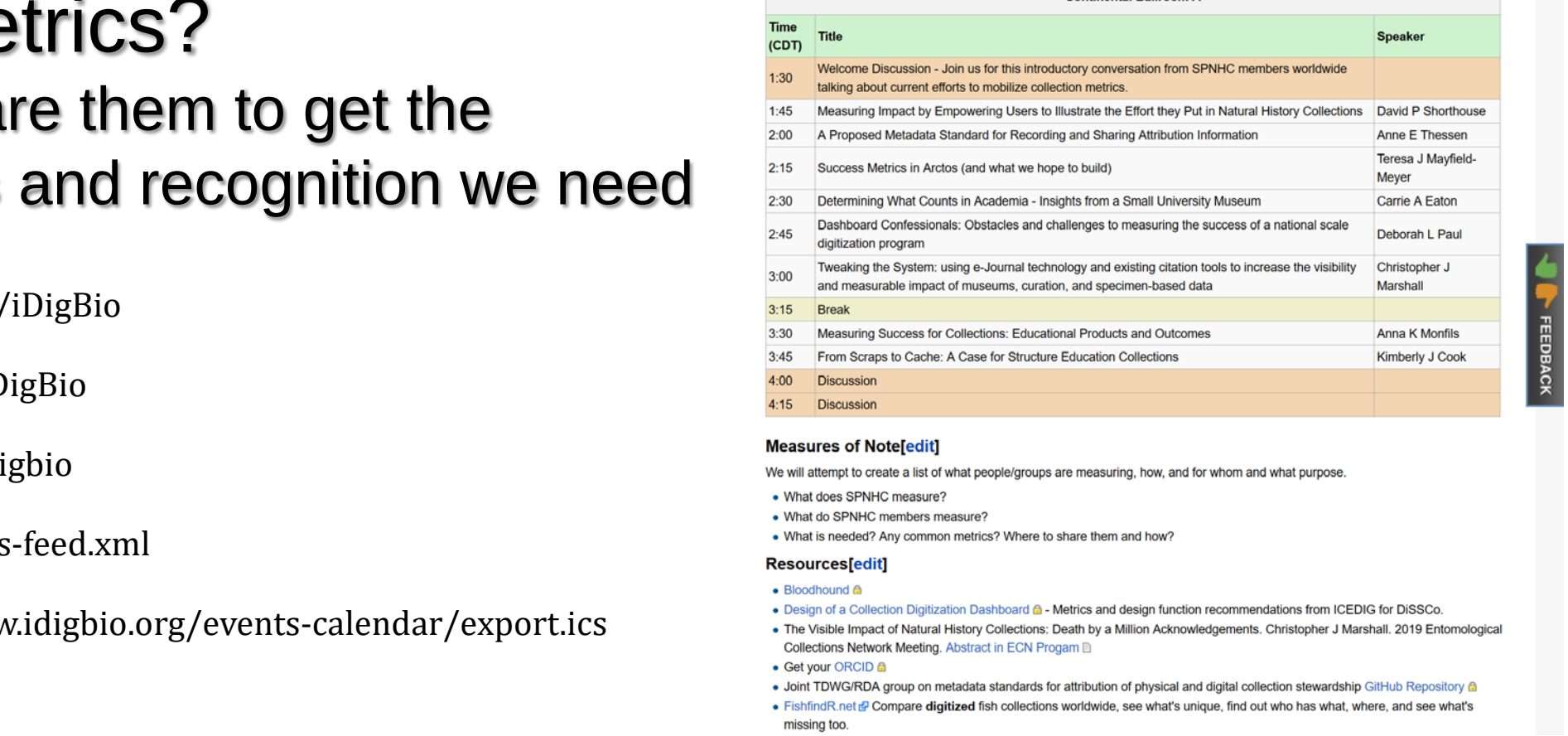
facebook.com/iDigBio

twitter.com/iDigBio

vimeo.com/idigbio

idigbio.org/rss-feed.xml

webcal://www.idigbio.org/events-calendar/export.ics



Symposium Schedule[edit]

Collecting Measures of Success: metrics for collections - a community conversation Continental Ballroom A		
Time (CDT)	**Title**	**Speaker**
1:30	Welcome Discussion - Join us for this introductory conversation from SPNHC members worldwide talking about current efforts to mobilize collection metrics.	
1:45	Measuring Impact by Empowering Users to Illustrate the Effort they Put in Natural History Collections	David P Shorthouse
2:00	A Proposed Metadata Standard for Recording and Sharing Attribution Information	Anne E Thessen
2:15	Success Metrics in Arctos (and what we hope to build)	Teresa J Mayfield-Meyer
2:30	Determining What Counts in Academia - Insights from a Small University Museum	Carrie A Eaton
2:45	Dashboard Confessionals: Obstacles and challenges to measuring the success of a national scale digitization program	Deborah L Paul
3:00	Tweaking the System: using e-Journal technology and existing citation tools to increase the visibility and measurable impact of museums, curation, and specimen-based data	Christopher J Marshall
3:15	Break	
3:30	Measuring Success for Collections: Educational Products and Outcomes	Anna K Monfils
3:45	From Scraps to Cache: A Case for Structure Education Collections	Kimberly J Cook
4:00	Discussion	
4:15	Discussion	

Measures of Note[edit]

We will attempt to create a list of what people/groups are measuring, how, and for whom and what purpose.

- What does SPNHC measure?
- What do SPNHC members measure?
- What is needed? Any common metrics? Where to share them and how?

Resources[edit]

- Bloodhound
- Design of a Collection Digitization Dashboard - Metrics and design function recommendations from ICEDIG for DiSSCo.
- The Visible Impact of Natural History Collections: Death by a Million Acknowledgements. Christopher J Marshall. 2019 Entomological Collections Network Meeting. Abstract in ECN Progam
- Get your ORCID
- Joint TDWG/RDA group on metadata standards for attribution of physical and digital collection stewardship GitHub Repository
- FishfindR.net Compare **digitized** fish collections worldwide, see what's unique, find out who has what, where, and see what's missing too.

FEEDBACK

iDigBio is funded by a grant from the National Science Foundation's Advancing Digitization of Biodiversity Collections Program (Cooperative Agreement EF-1115210). 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. All images used with permission or are free from copyright.