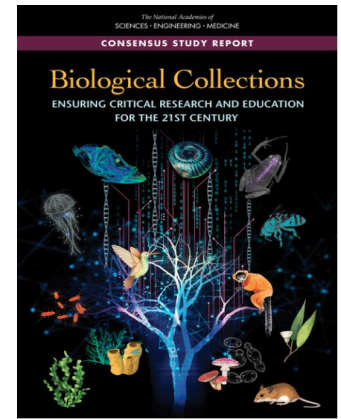


Forward-Thinking Discussion on Biological Collections



January 14, 2021



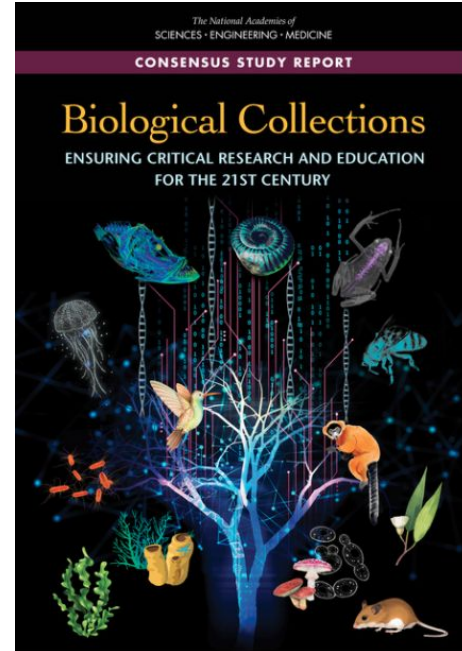
Biodiversity Collections Network (BCoN) Panelists

- John Bates, The Field Museum of Natural History
- Andrew Bentley, Biodiversity Institute, University of Kansas
- Linda S. Ford, Harvard University Museum of Comparative Zoology
- David Jennings, iDigBio, Florida Museum of Natural History
- Anna Monfils, Central Michigan University
- Barbara Thiers, The New York Botanical Garden
- Jen Zaspel, Milwaukee Public Museum



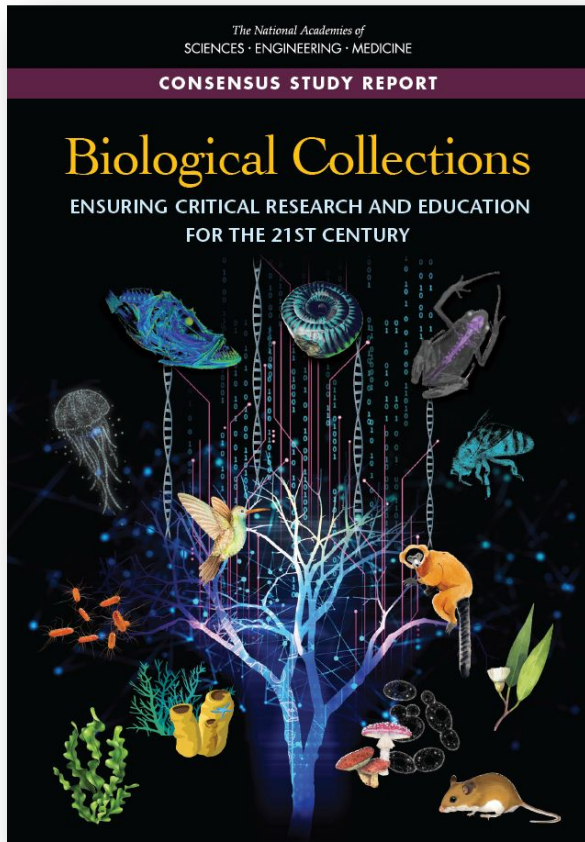
National Academies of Sciences, Engineering, and Medicine (NASEM) Study Panelists

- Andrew Bentley, Biodiversity Institute, University of Kansas
- Kyria Boundy-Mills, University of California, Davis
- James Collins, Arizona State University
- Manzour H. Hazbón
- Shirley Pomponi, Florida Atlantic University Harbor Branch Oceanographic Institute
- Barbara Thiers, The New York Botanical Garden



Program Overview

- Introduction
- Implementation Challenges
 - Need for Additional Collecting
 - Continued Digitization
 - Data Integration and Cyberinfrastructure
 - Infrastructure
 - Education and Workforce Training
- Break
- Case Studies
 - Parasite Tracker
 - Living Collections and COVID-19
- Legal Compliance in Biodiversity Collections
- Roles of Collections Organizations
- Conclusions
- Discussion



Biological Collections:

Ensuring Critical Research and Education for the 21st Century

Forward-Thinking Discussion on Biological Collections
January 14, 2021

Co-Chairs

Shirley Pomponi, Florida Atlantic University
James Collins, Arizona State University

Board on Life Sciences



The National
Academies of

SCIENCES
ENGINEERING
MEDICINE



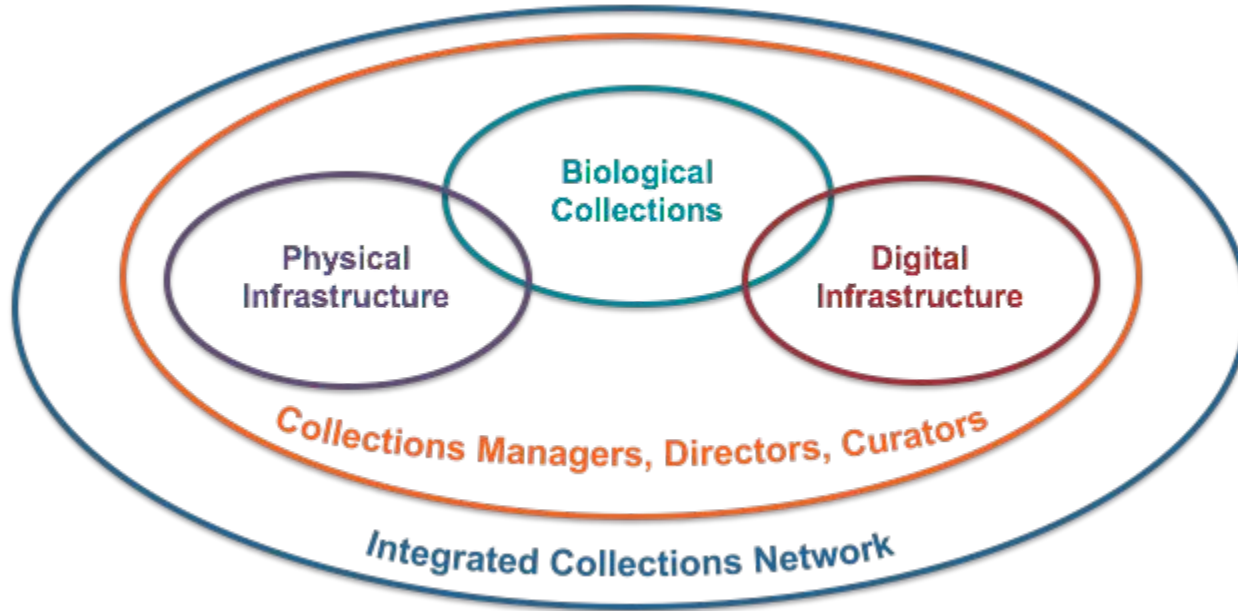
AMERICAN INSTITUTE OF
BIOLOGICAL SCIENCES



SPNHC
ADVANCING COLLECTIONS CARE



Infrastructure has many intersecting components.



Recommendations for Collaborative Action

- NSF should help establish a permanent national Action Center for Biological Collections to coordinate action and knowledge, resources, and data-sharing among the nation's biological collections as they strive to meet the complex and often unpredictable needs of science and society.
- NSF should lead efforts to develop a vision and strategy, such as a decadal survey, for targeted growth of the nation's biological collections, their infrastructure, and their ability to serve a broader range of users and scientific and educational needs.
- NSF should expand partnership capabilities more broadly across NSF, other federal agencies, international programs, and other sectors to maximize investments.

Extending U.S. Biodiversity Collections to Address National Challenges



<https://bcon.aibs.org/2019/01/16/community-input-requested-extending-u-s-biodiversity-collections-to-address-national-challenges/>

Released March 2019



EXTENDED SPECIMEN
NETWORK

The Extended Specimen Network: A Strategy to Enhance US Biodiversity Collections, Promote Research and Education

James Lendemer , Barbara Thiers, Anna K Monfils, Jennifer Zaspel, Elizabeth R Ellwood, Andrew Bentley, Katherine LeVan, John Bates, David Jennings, Dori Contreras ... [Show more](#)

BioScience, Volume 70, Issue 1, January 2020, Pages 23–30,
<https://doi.org/10.1093/biosci/biz140>

Published: 22 November 2019



Advertisement 1



BIODIVERSITY
COLLECTIONS NETWORK

The National
Academies of

SCIENCES
ENGINEERING
MEDICINE



AMERICAN INSTITUTE OF
BIOLOGICAL SCIENCES



SPNHC
ADVANCING COLLECTIONS CARE



EXTENDED SPECIMEN
NETWORK

Areas of Overlap in Recommendations in NASEM and BCoN Reports

1. Improve Infrastructure (for physical and digital collections)
2. Continue to add new collections, accommodating a wider array of associated data
3. Prioritize linkage of specimens and derived data
4. Expand accessibility of data to communities outside of the primary stakeholders
5. Enhance workforce training for managers of collections and data
6. Expand educational uses of collections
7. Establish a permanent action center to manage the data repository and support education and training initiatives

Implementation challenges

Lots of commonalities in recommendations of the two reports. Shared vision of the future of collections.

5 broad thematic areas or “pillars” required to implement an ***Extended Specimen Network*** concept

- Collecting
- Digitization
- Infrastructure
- Integration
- Workforce and Education

Ongoing efforts are putting some pieces in place, some still to be implemented and lots of challenges.

Collecting

- Continued collecting and collection growth - vouchering, temporal series
- Holistic collecting - not just target taxa, multi-disciplinary
- Associations and preparations - host:parasite, plant:pollinator, tissue:voucher
- Promote collaboration with collections to identify gaps - taxonomic, geographic and temporal
- Born digital to alleviate pressure on digitization
- Observations

PERSPECTIVE

The next generation of natural history collections

David E. Schindel^{1*}, Joseph A. Cook²

¹ Smithsonian Institution, Washington, D. C., United States of America, ² Museum of Southwestern Biology, University of New Mexico, Albuquerque, New Mexico, United States of America

*schindel@si.edu



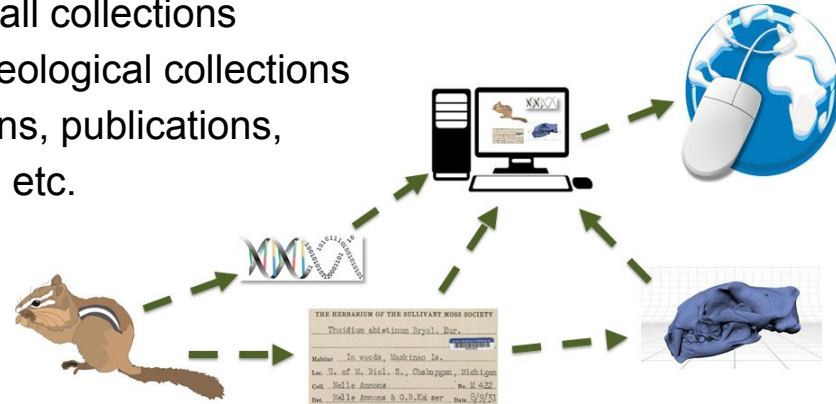
Collecting

Challenges:

- Permitting restrictions - international, regional and taxonomic
- Collecting priorities driven by research \$\$
- Little direct support for collections from research
- Specimen Management Plan
 - Link research \$\$ to collections maintenance
 - Support digitization, curation and long-term care
 - \$\$ amount per specimen (based on taxon)
 - Collaborative agreement during proposal phase identifying voucher repository
 - Reinforce value of voucher deposition

Continued digitization, incorporation of dark data and other collections

- Leverage past and ongoing investments in digitization efforts → estimate ~30% of specimens in collections are currently digitized
- Focus on hard to digitize collections → entomology, paleontology
- Find the “dark data” → project-based and small collections
- Integrate new collection types → living and geological collections
- Augment data → field notes, images, CT scans, publications, Genbank sequences, Isotope data, trait data, etc.
- Complete stub records → focus on data use

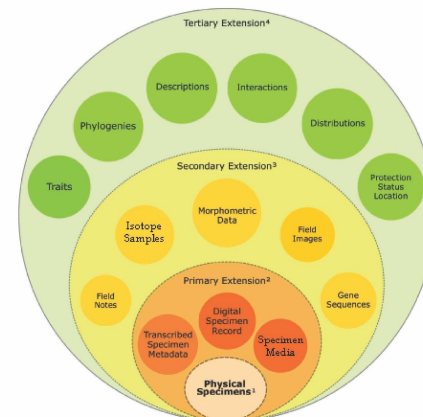
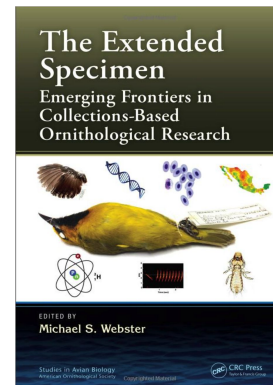


Challenges

- Data cleaning, standardization, and de-duplication
- Comprehensive registry of collections → where are they?
- Culture change → digitization as part of the normal process

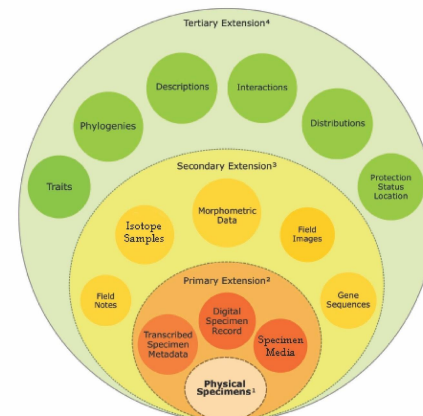
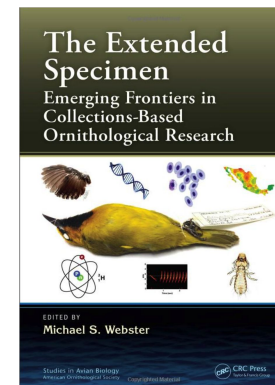
Integration and cyberinfrastructure

- Extended specimens rely on integration of data
- Allows us to answer important, complex questions
- Need integration at all levels - preparations, specimens, collections, outside data sources
- Need integration of products of research - publications, sequences, isotopes, CT scans, images
- Reliant on unique identifiers - objects, datasets, institutions/collections, people - and their use throughout the data lifecycle - everyone needs to play the game!!
- Reliant on published, complete, FAIR data
- Requires robust cyberinfrastructure to support digitization, storage and dissemination



Integration and cyberinfrastructure

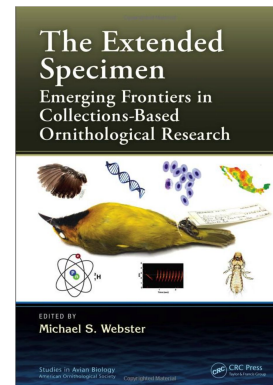
- Integration also important for collections attribution, annotation and legal compliance
 - Showing our worth to the scientific enterprise
 - Advocating for collections - funding
 - Community engagement and annotation
 - Nagoya protocol requirement compliance
- Some parts of cyberinfrastructure being put in place:
 - GBIF collections registry and others
 - GBIF clustering tool
 - DiSCCo openDS concept
 - IGSN sample registry
 - Pensoft ARPHA writing tool
- But there are still significant holes and hurdles



Integration and cyberinfrastructure

Challenges:

- Still no unifying identifier system - registry
- Data landscape is very complex and convoluted - simplification/streamlining needed
- Identify and fill in gaps in cyberinfrastructure
- Implement technology to circumvent current social needs
- Shift in our traditional publishing mechanism to a more transactional system
- Link our current highly distributed and disparate communities



Infrastructure



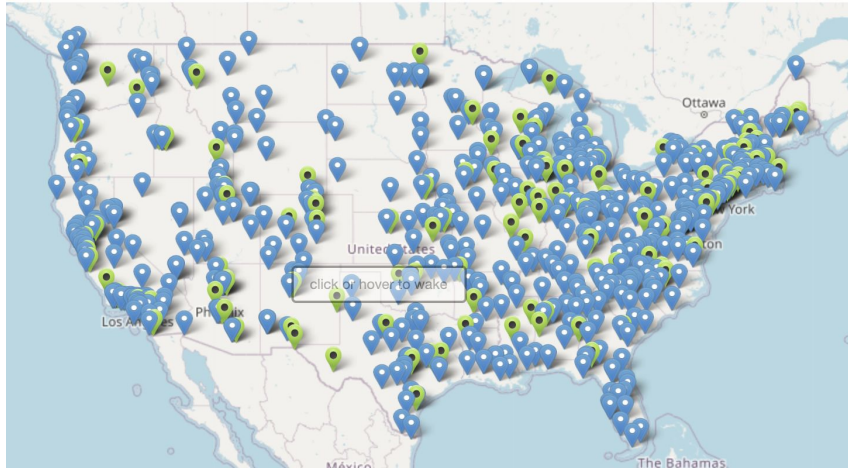
Institutional Infrastructure

Typical operational costs typically include funds for

- equipment and staff needed to maintain the collection
- adding new accessions
- making collections and data available to users by supporting in person visits and loans, and digitizing specimens and maintaining access to collections.



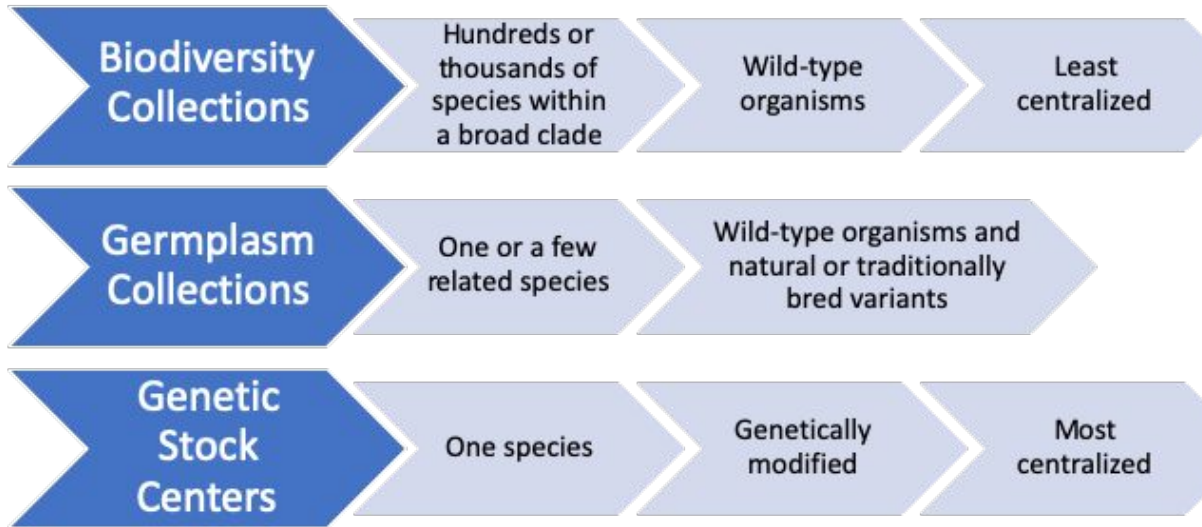
Community Infrastructure



OVERVIEW OF LIVING COLLECTIONS

Kyria Boundy-Mills

General categories of living collections



Examples

Zoos, aquaria, botanical gardens, arboretums, seed banks, microbes (Eubacteria, Archaea, fungi, nematodes, microalgae)

Crop and livestock germplasm collections, primate centers, plant pathogens, starter cultures, cell and tissue cultures

Genetically modified model organisms: viruses, bacteria, fungi, plants, invertebrates, vertebrates



E. coli type strain at ATCC

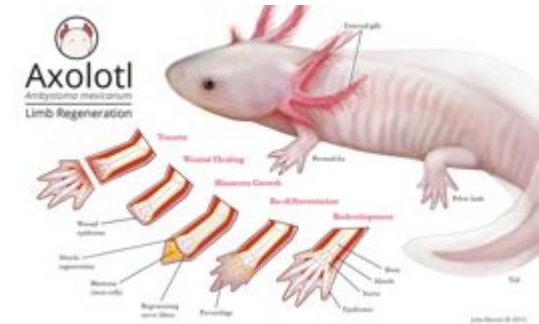
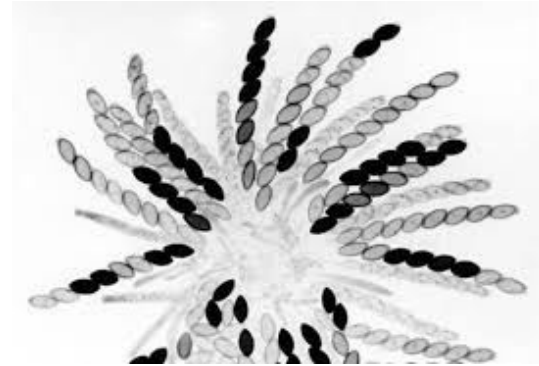


>90,000 wild isolates at ECRC



Notes on living stock collections

- Living specimens usually propagated and distributed to researchers
- Many **different types** of living collections because they are used in many **different areas** of basic and applied life science
- Different research fields may have one or more **centralized** repositories
 - **MMRRC**: Five facilities, 50,000+ lines of *Mus musculus* (lab mouse) for genetic, health research
 - **FGSC**: *Neurospora crassa* used for classical genetics because 8 spores are produced
 - **ABRC**: *Arabidopsis thaliana* used for plant genetics because it grows quickly and makes lots of seeds; three repositories in the world distribute hundreds of thousands of specimens annually; 9,000 research labs
 - **AGSC**: Axolotl *Ambystoma mexicanum* is used to study limb regeneration
 - **Five primate centers**: UC Davis includes Rhesus monkeys over 19 years old used for aging research
- **Less interaction** among different collections



Data associated with living collections

- Online catalogs of specimens
- Data varies by collection and field of use
 - Genetic stock center: Phenotype, Genotype (no occurrence data)
 - Germplasm collection: Crop or livestock properties; physiological and biochemical properties; genetics
 - Biodiversity collection: Geographic source; physiological and genetic properties

Additional features of GENETIC STOCK CENTERS: model for Extended Specimen

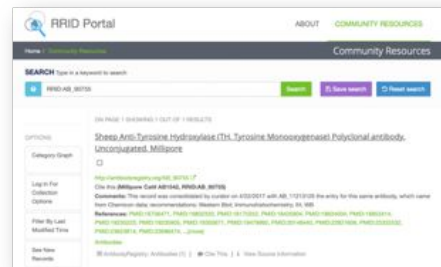
- An inventoried list of genes
- Curated, robust, data-supported gene annotation
- Tools for genetic manipulation
- Computational platform for systems-level analysis

Collection registries

Collection organization	Resource
WFCC	Global microbe collections
USCCN	US microbe collections
NCBI Biollections	Database of collections with specimens cited in GenBank accessions
Scientific societies	Collections committees
ISBER	Biobanks (human tissue specimens; human-associated microbes)
CGC	Crop germplasm collections at universities, USDA, companies
USDA-NLGRP	Cryopreserved seeds, semen, microbes from USDA, companies and universities
Seed Library Network	600+ seed libraries

Data aggregators

Data aggregator	Types of materials
GRIN Global	Crop collections
GCM	Global Catalog of Microorganisms
Genetic stock center database	Model organisms; genomes; gene sequences; “extended specimen” (products, tools, protocols)
MMRRC	Centralized lab mouse catalog
RRID	Research Resource Identifiers for repositories, antibodies, plasmids, organisms, cell lines, tools.



Decisions: living collections

What types of **collections** would be included?

What collections would be excluded, and why?

What types of **data** would be compatible? Useful?

•What types of data be excluded and why?

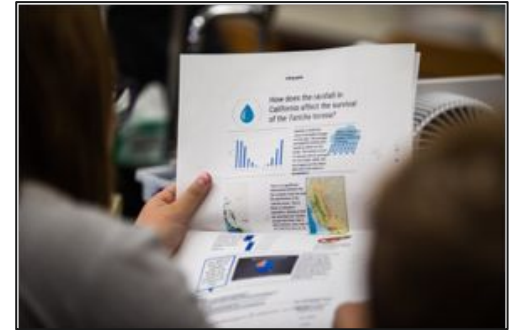
Education and Workforce Training



Education and Workforce Training - Next Steps

Recruiting, training, and supporting a diverse workforce and engaged citizenry for the future:

- Training current workforce/continuing education
- Training the next-generation of cyber collection managers
- Training/engaging new end users of biodiversity data
- Diversity, Equity and Inclusion



Education and Workforce Training - Challenges

Challenges to recruiting, training, and supporting a diverse workforce and engaged citizenry for the future:

- Cyber collections workforce needs assessment
- Coordinated training in biodiversity and data science at multiple levels
- Need for accessible database/interface design for E&O
- Community plan and funding to address Diversity, Equity, and Inclusion at multiple levels



After the break

- Case Studies
 - Terrestrial Parasite Tracker
 - Living collections as a critical resource to respond to pandemics
- Legal Compliance in Biodiversity Collections
- Roles of Collections Organizations
- Where to from here?
- Discussion

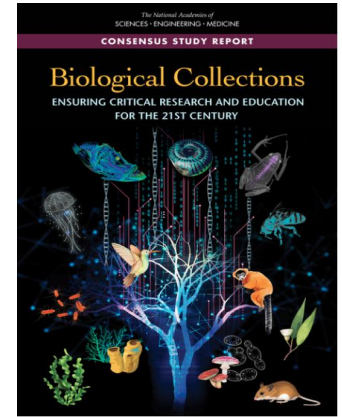
The program will resume after a short break.

Forward-Thinking Discussion on Biological Collections



January 14, 2021

The program will resume after a short break.



Next on the Program

- Case Studies
 - Terrestrial Parasite Tracker
 - Living collections as a critical resource to respond to pandemics
- Legal Compliance in Biodiversity Collections
- Roles of Collections Organizations
- Where to from here?
- Discussion

Case Study: Terrestrial Parasite Tracker TCN

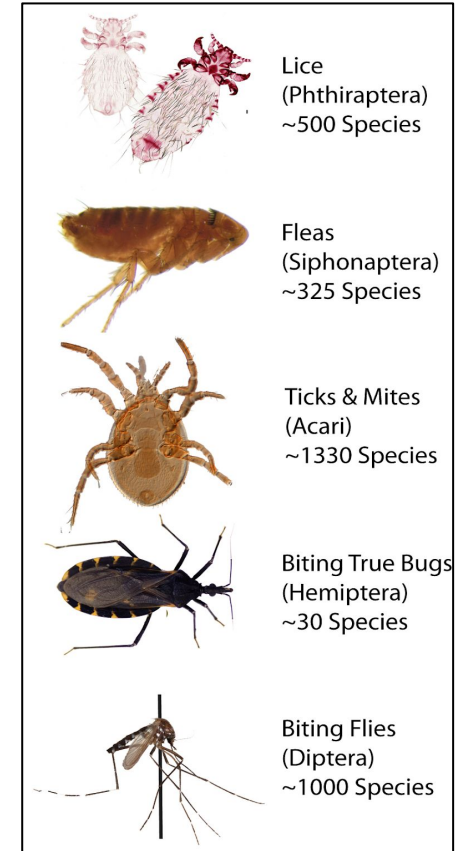
Recall: 5 broad thematic areas or “pillars” required to implement an **Extended Specimen Network** concept

- Collecting
- Digitization
- Infrastructure
- Integration
- Workforce and Education

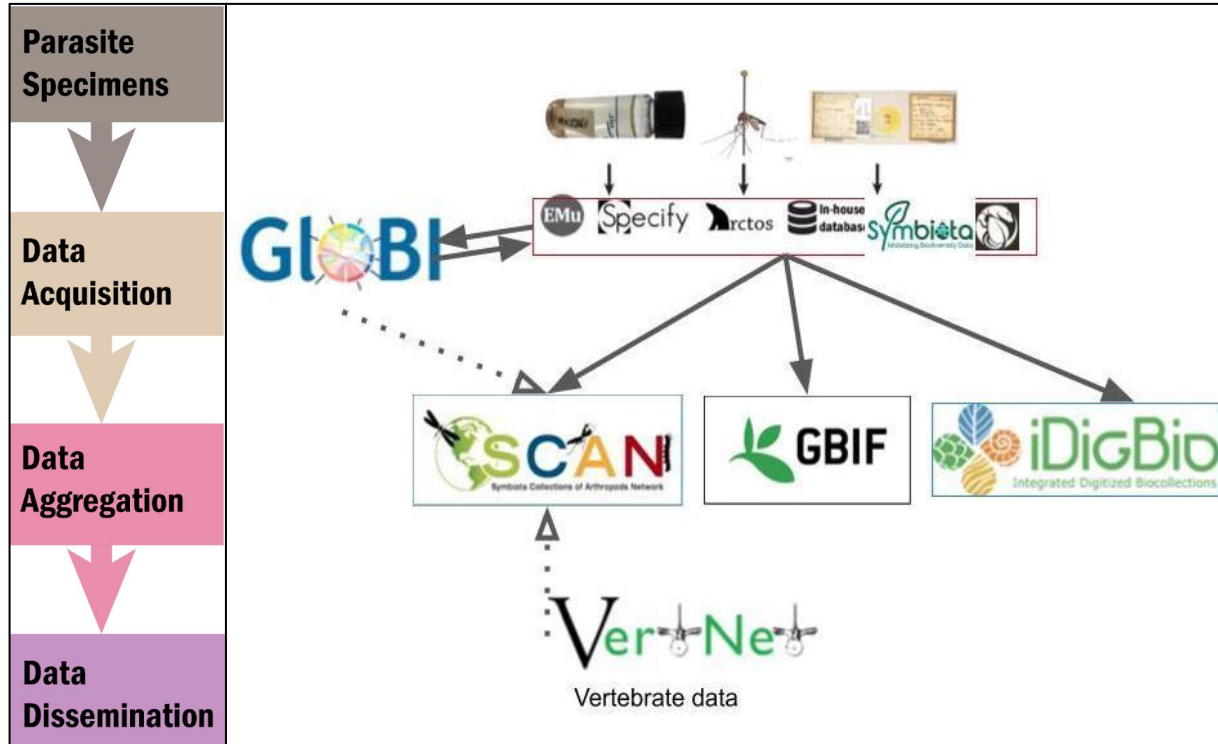
Ongoing efforts are putting some pieces in place, some still to be implemented and **lots of challenges**.

Case Study: Terrestrial Parasite Tracker TCN

- **Project Title:** Digitization TCN: Digitizing collections to trace parasite-host associations and predict the spread of vector-borne disease
- **Project Start Date:** September 1st 2019
- **Project Period:** 3 years
- **Participating Institutions:** 28
- **Co-PIs/Leads on Subs:** 34
- **Participants:** 120+
- **Primary Objective:** Aggregate arthropod parasite collections to build an accessible, comprehensive database of parasite-host associations and vector distributions
- **Outcome:** Provide needed baseline information for research and management of the ecological interactions among parasites, pathogens, and their hosts in North America



Case Study: Terrestrial Parasite Tracker TCN challenges are opportunities!



- Collecting
- Digitization
- Infrastructure
- Integration
- Workforce and Education



Initial steps towards development of a Global Health and Biodiversity Data working group --

- Development of a working group focused on biological collections and associated specimen data relevant to global health and biodiversity issues
- Emerged out of a TPT workshop held at the FMNH last February where PIs gathered with stakeholders representing other projects and agencies related to medically-important arthropods
- Interested in integration of specimen data produced by projects funded both within and outside of NSF
- Group has identified key issues that could be addressed with additional resources
- In recent months, the group has met 3 times to discuss a possible collaboration with the NMNH/Walter Reed Biosystematics Unit
- Future directions for this group include submission of RCN application to NSF, and development of collaborative project with Vectorbase

Living collections as a critical resource to respond to pandemics

WHAT IS THE ROLE OF A LIVING CULTURE COLLECTION?



■ Culture collections provide *qualified* biomaterials to the community.

- Pure
- Viable
- Preserved
- ID Confirmed
- QC
- Storage
- Biosafety
- Information
 - Website
 - Searchable
 - Description
 - Ordering
- Compliance
 - Registration
 - Traceability
 - Documentation
 - Permits
- Shipping
 - Cold chain
 - Domestic/International
 - Packaging
 - Import/export permits

2

Credit: Liquid Nitrogen Tanks at the American Type Culture Collection

LIVING COLLECTIONS DEVELOP CONTROLS AND STANDARDS

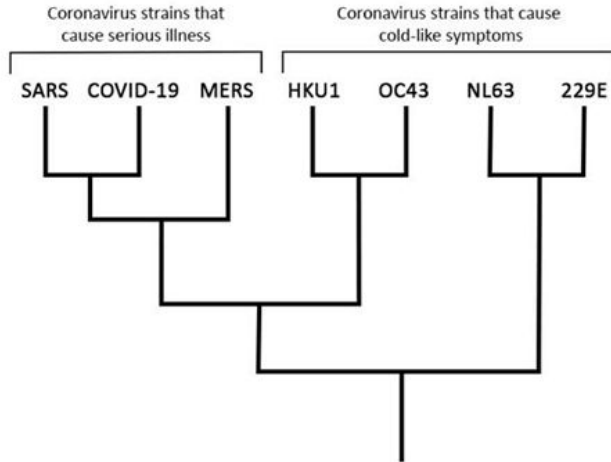


- Authenticate isolates
- Adapt strains to growth in new hosts/cell lines
- Develop qualified controls for:
 - Diagnostic assays
 - Clinical/preclinical studies
- Provide near neighbors for inclusivity and exclusivity panels
- Produce derivatives essential for researchers
 - Inactivated materials
 - Antigens
 - Antibodies
 - Nucleic acids (genomic and synthetic)
- Produce and distribute diagnostic assays

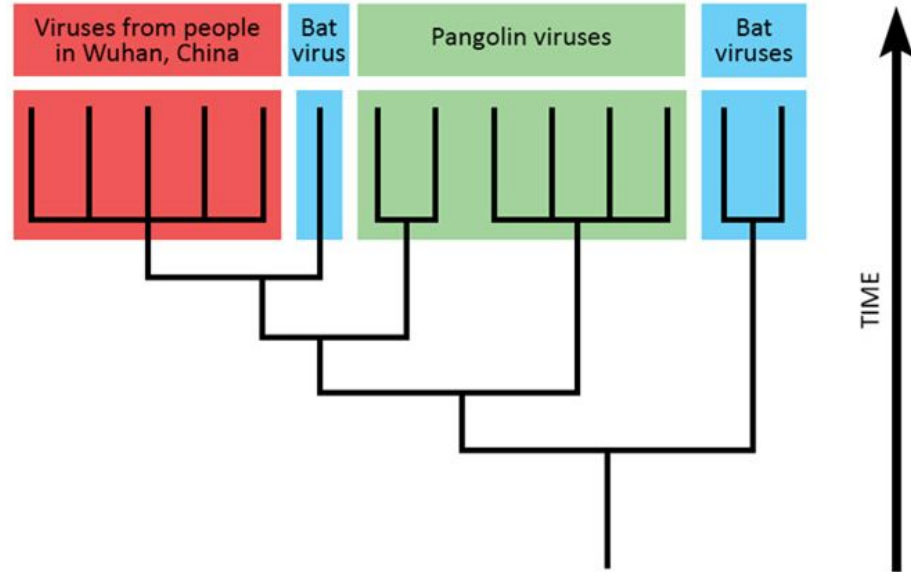
Credit: Pixabay/CC0 Public Domain

LIVING COLLECTIONS HAVE BIOMATERIALS FROM DIVERSE SOURCES

- Historical value
- Geographical origins
- Clinical/environmental
- Human/animal/healthy/sick



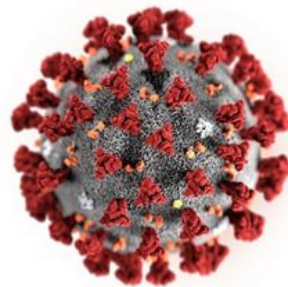
CORONAVIRUS EVOLUTIONARY TREE



Source: https://evolution.berkeley.edu/evolibrary/news/200305_coronavirus
https://evolution.berkeley.edu/evolibrary/news/200409_coronavirus

LIVING COLLECTION'S RESPONSE TO COVID-19:

- Developed a wide range of qualified materials
 - Viral isolates
 - Genomic and synthetic nucleic acids, plasmids
 - Inactivated materials
 - Monoclonal and polyclonal antibodies
 - Serum/plasma
 - Proteins, peptide arrays
- Qualified controls and materials to support
 - Basic research
 - Diagnostic assay development
 - Diagnostic equipment validation
 - Vaccine development
 - Development and validation of novel therapeutics
- Distributed materials globally
- Published scientific papers



GOVERNMENT



LIVING COLLECTIONS



<https://www.beiresources.org/>
<https://www.internationalreagentresource.org/Home.aspx>
<https://www.atcc.org/>

IRR is funded by CDC. BEI Resources is funded by NIAID

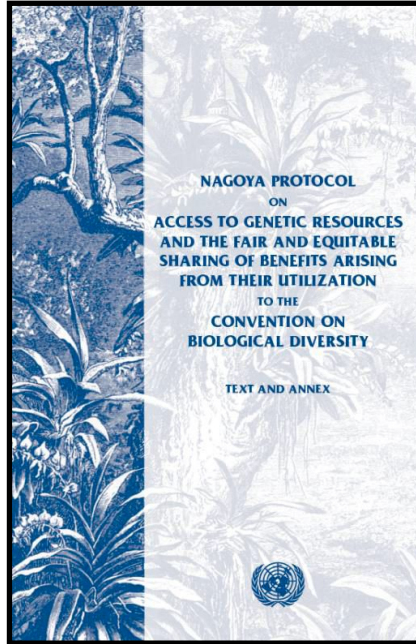
Legal Compliance in Biodiversity Collections

➔ *How concepts discussed -- benefit collections regarding legal compliance issues?*

- Legal issues are substantial concern for biological collections
 - Increased visibility due to successful digitization
 - Much material in U.S. biodiversity collections originates from other countries and requires legal documentation
- New legal landscape for international collections and research
 - Increased scrutiny of collection movement (shipments: what, when, where and how)
 - Nagoya Protocol implemented by the UN Convention on Biological Diversity



Legal Compliance: New Landscape



- Nagoya Protocol on Access and Benefit-Sharing (ABS)
 - Legally binding supplementary agreement to Convention on Biological Diversity
 - Came into force on 12 October 2014
 - Outlines obligations for genetic resources and associated traditional knowledge:
 - o Access regulated
 - o Benefits from utilization shared
 - o Compliance measures required

Even non-signatory countries must comply with domestic ABS legislation!

Legal Compliance: New Obligations for Biological Collections

- Track all documentation associated with specimens
- Share and document benefits as negotiated with providing countries
- Keep records of all instances of use
- Allow access to documentation



Nagoya requires legal compliance for entire life cycle of a specimen

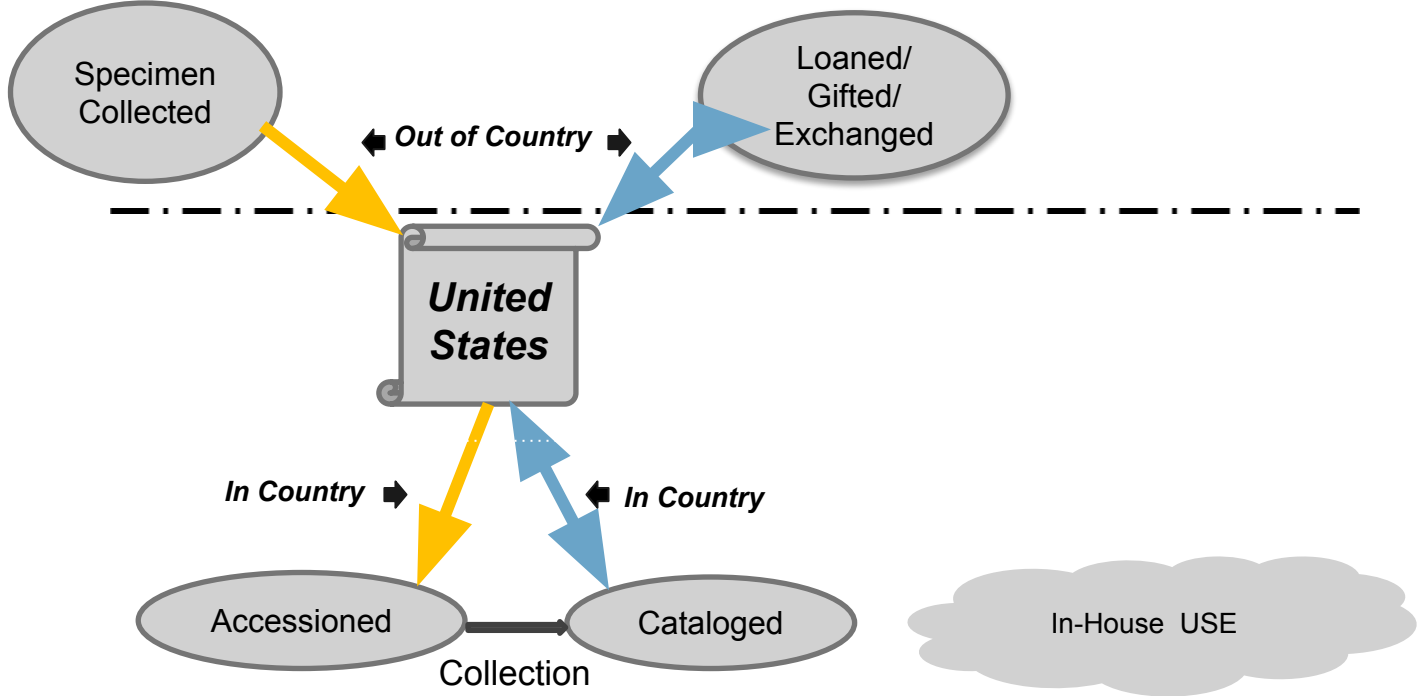
- Maintain all legal documents for each specimen throughout its entire research/ use life
- Track all activities and transactions from point of collection, and in perpetuity

Legal Compliance: Challenges in Tracking Documentation

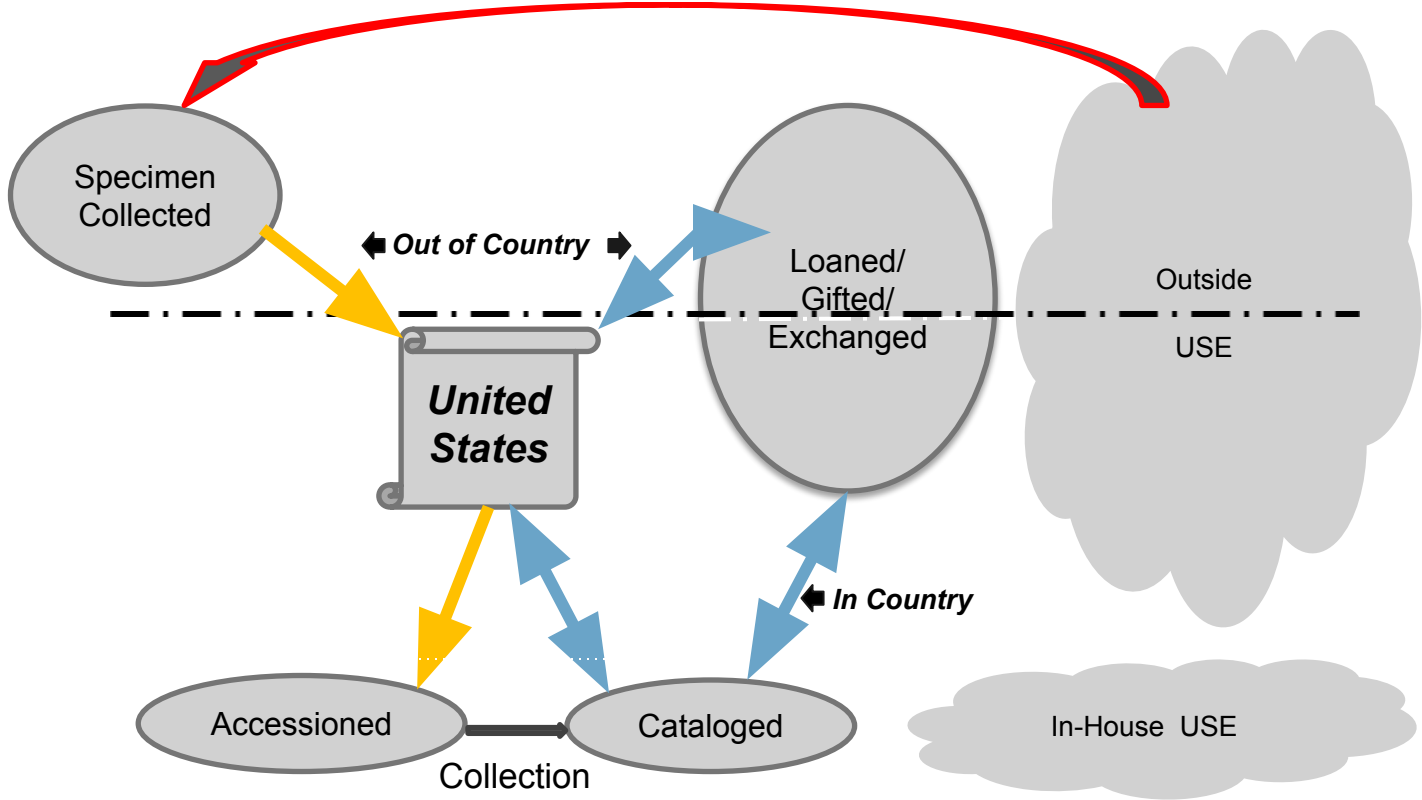
- Legal documents are not all the same type
 - Such as permits, certificates, agreements, contracts, memorandum of understanding, memorandum of cooperation, rights, etc.
- Temporal component differs among legal documents
 - Use may be one time, multiple times or range of time
 - Upstream permission agreements, historically done by researchers
 - Downstream future use, responsibility of collections
 - Need for flexibility, upstream agreement may change through time



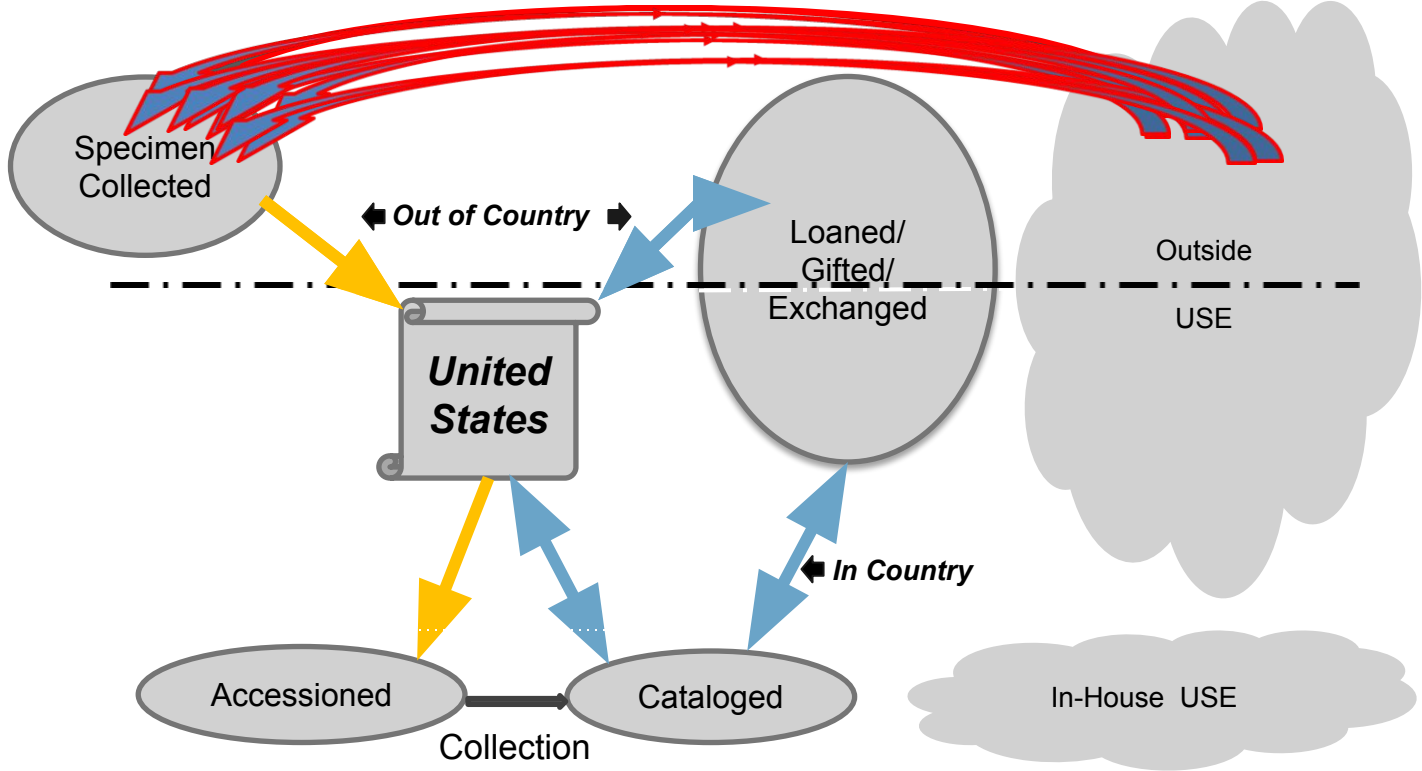
Legal Compliance: Life Cycle of a Specimen



Legal Compliance: New ABS Obligations for Collections



Legal Compliance: New ABS Obligations for Collections



Legal Compliance: Conclusions for Collections

- Institutions must increase transparency of their collections
- Digital solutions for tracking compliance needed for easy retrieval:
 - Compliance documentation
(e.g., permits, certificates, agreements)
 - Transactions
(e.g., accessions, loans, borrows, gifts, exchanges)
 - Products of use
(e.g., publications, sequence accessions, media)
 - Metadata of benefit-sharing agreements linked to individual specimens
(e.g., allowed uses determine whether negotiation is required)



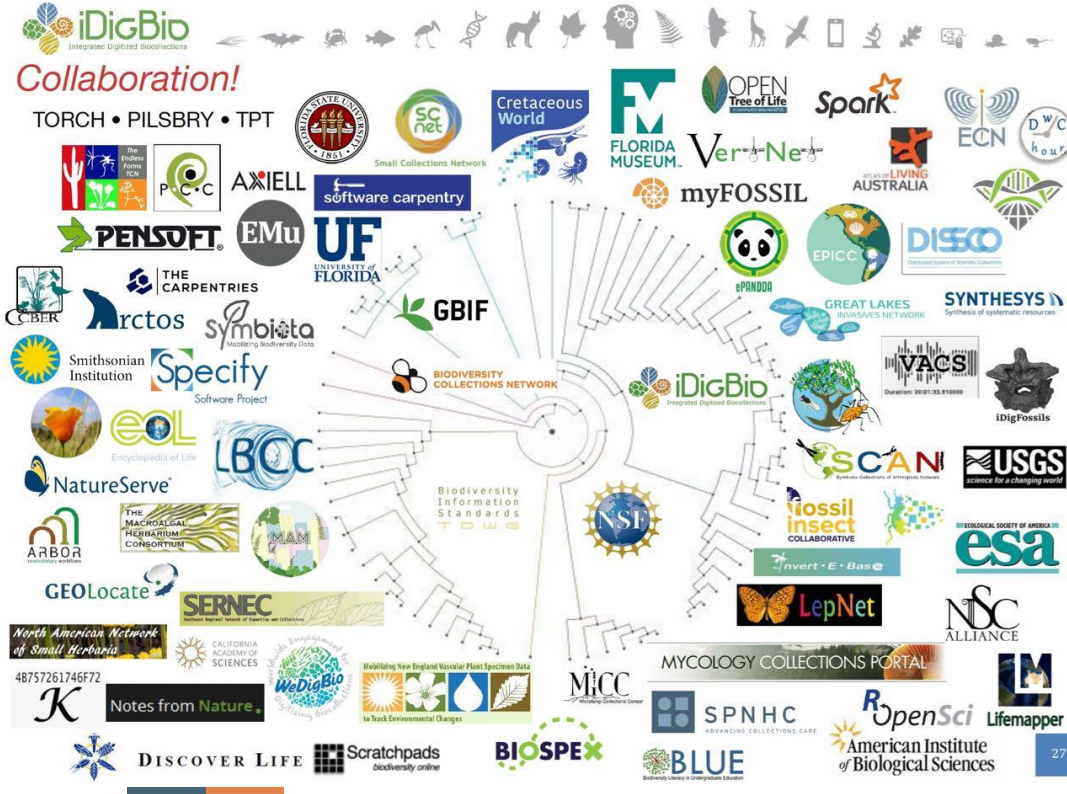
Legal Compliance: How to Address this Complexity?

- Biodiversity collections find individual solutions for compliance and tracking
- Address on global scale by networked provenance and existing use agreements
 - Data structures and systems allowing for storage, query, and retrieval of permissions and rights metadata
 - Increased transparency for community (e.g., collections, researchers, providing countries)

➔ *Latter: One example of an **extended specimen network** that would enhance biodiversity collections for research and education*



Roles of Collections Organizations in implementing NAS recommendations, ESN



Everyone has a role to play

- Collections
- Researchers
- Educators
- Aggregators
- **Collections-based societies**
- Other societies
- Funding agencies
- Technology
- Publishers

Natural History collections going forward

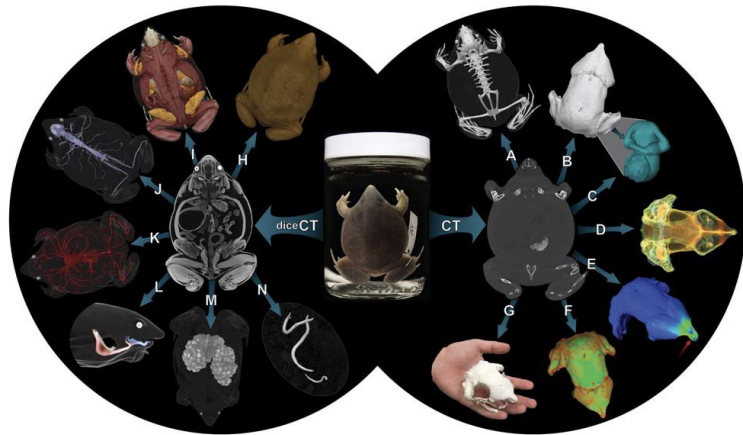
- Our community is a diverse network of networks leading back to individual collections (small to big) both nationally and internationally
- We are working with aggregators e.g, iDigBio, GBIF, DISSCo and others internationally



Natural History collections going forward

- Our collections have never been more accessed than today because of digitization

Quarterly downloads of Morphosource 3-D specimen data



overt

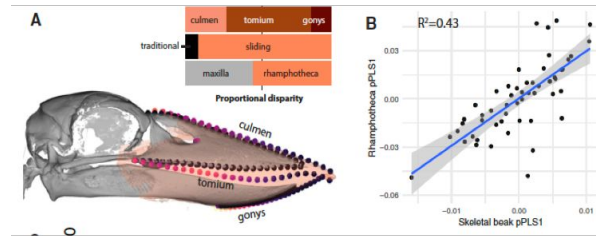


Natural History collections going forward

- The novel science possible is revolutionary

Morphological innovation and biomechanical diversity in plunge-diving birds

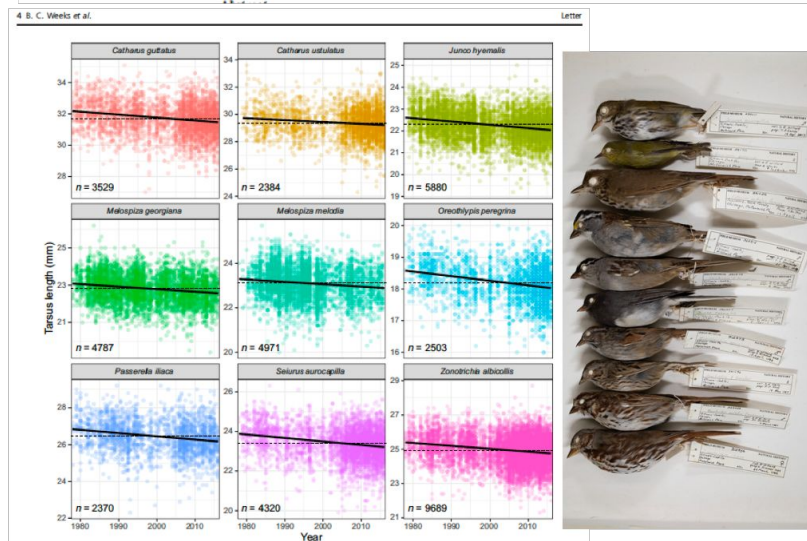
Chad M. Eliason,^{1,2} Lorian Straker,³ Sunghwan Jung,⁴ and Shannon J. Hackett^{1,5}



ECOLOGY LETTERS

Ecology Letters, (2019) doi: 10.1111/ele.13434

LETTER Shared morphological consequences of global warming in North American migratory birds



Natural History collections going forward

- Our community is a diverse network of networks leading back to individual collections (small to big)
- We need, as a community, to look out for one another and these networks can help
- Our collections have never been more accessed than today because of digitization
- The novel science possible is revolutionary
- We have a national blueprint in these reports, now we have to implement these recommendations
- Coordination comes from societies that include NSCA, SPNHC, and AIBS.

Converging Digital and Extended Specimens: Towards a global specification for data integration

Global Consultation, February 2021

See [alliance for biodiversity knowledge](https://allianceforbio.org) (allianceforbio.org) blog for details

Topics of the consultation will include:

- Digitizing/mobilizing FAIR data for specimens
- Extending, enriching and integrating data
- Annotating specimens and related data
- Crediting and attributing tasks like data and material curation
- Analysing/mining specimen data for novel applications

Where to from here?

- Series of workshops to outline the implementation challenges to recommendations of ESN and NASEM Reports, and suggest actions needed to overcome these challenges
- Propose a new RCN to coordinate implementation
- Engage international community in collaboration
- Promote and coordinate the development of projects designed to implement aspects of the plan
- Maintain dialog with NSF, other relevant government agencies to maximize funding opportunities.

***Most immediate need:* Expressions of support and especially interest in participating in this process. We want people from collections all types, large and small, as well as data experts, educators and data users, especially users who fall outside of the systematic community.**

Discussion

Questions?