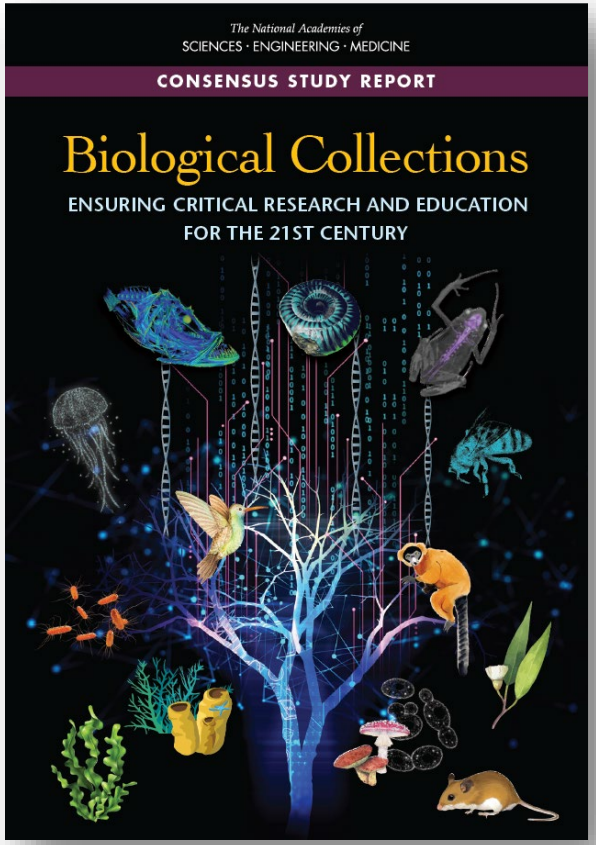
The National Academies of
SCIENCES • ENGINEERING • MEDICINE




# Biological Collections:

## Ensuring Critical Research and Education for the 21st Century

Report Highlights

September 25, 2020

Study Co-Chair

Shirley Pomponi, Florida Atlantic University

Board on Life Sciences

# Study Committee

**James P. Collins** (co-chair)
Virginia M. Ullman Professor of Natural History and the Environment
School of Life Sciences
Arizona State University

**Shirley A. Pomponi** (co-chair)
Research Professor
Harbor Branch Oceanographic Institute
Florida Atlantic University

**Andrew C. Bentley**
Collections Manager, Ichthyology
Biodiversity Institute
University of Kansas

**Rick E. Borchelt**
Director
Communications & Public Affairs
Office of Science
U.S. Department of Energy

**Kyria Boundy-Mills**
Professional Research Microbiologist
Curator, Phaff Yeast Culture Collection
University of California Davis

**Joseph A. Cook**
Regents Professor
Curator of Mammals
Museum of Southwestern Biology
University of New Mexico

**Lynn D. Dierking**
Sea Grant Professor
in Free-Choice STEM Learning
Director, Strategy & Partnerships
Institute for Learning Innovation
Oregon State University

**Scott V. Edwards**
Alexander Agassiz Professor
Organismic and Evolutionary Biology
Harvard University

**Manzour H. Hazbón**
Senior Scientist
Bacteriology
American Type Culture Collection

**Talia S. Karim**
Collection Manager
Invertebrate Paleontology
University of Colorado
Museum of Natural History

**George I. Matsumoto**
Senior Education and Research Specialist
Monterey Bay Aquarium
Research Institute

**Pamela S. Soltis**
Director, Biodiversity Institute
Distinguished Professor and Curator
University of Florida
Florida Museum of Natural History

**Barbara Thiers**
Vice President
Patricia K. Holmgren Director of the William and Lynda Steere Herbarium
The New York Botanical Garden

**Staff**

**Audrey Thévenon,** Project Director
**Keegan Sawyer,** Sr. Program Officer
**Jessica De Mouy,** Sr. Program Assistant

*The National Academies of*
SCIENCES • ENGINEERING • MEDICINE

# Statement of Task (abbreviated)

NSF recognizes the breadth of needs for maintaining biological collections exceeds the capabilities of any one federal agency.

NSF asked NASEM for guidance on questions regarding long-term sustainability, including operational structures, policies, and social cultures that could provide momentum to maintain and grow biological collections.

- Explore the contributions of biological collections of all sizes and institutional types to research and education.
- Envision future innovative ways in which biological collections can be used to advance science.
- Outline the critical challenges to and needs for use and maintenance of biological collections.
- Suggest a range of long-term strategies that could be used for their sustained support.

# Biological Collections Covered in This Report

Non-living specimens include organisms preserved by scientists and naturally preserved remains, such as fossils, commonly referred to as natural history collections.

Living specimens include research and model organisms that are grown and maintained in genetic stock centers, germplasm repositories, or living biodiversity collections.

Focused on collections that receive, or are eligible to receive, support for infrastructure or digitization from NSF. This report does not cover biological collections owned by federal agencies

This report does not explicitly address living collections in zoos, aquaria, or botanical gardens; biobanks or repositories of human tissues; or anthropological and geological collections (excluding fossils).

# The Committee's Approach

- Committee's expertise, peer-reviewed literature, and information-gathering events.
- Identified areas of tension that stem from the scope of the study and that are inherent within the biological collections community:
  - Collections are diverse—taxonomically, organizationally, and in their missions and needs.
  - The communities surrounding collections operate largely independently of one another.
  - There are inherent differences between living stock collections and natural history collections.
- The report is intended to launch a national conversation about the future of biological collections.

# The Committee's Vision

For biological collections to survive and thrive…

*Provide long-term support for collections-based scientific research, instill a culture of stewardship for and access to biological specimens, build and grow biological collections to better represent global biodiversity in space and time, promote access to biological collections as important educational resources for the general public, and encourage the exchange of biological resources and knowledge.*

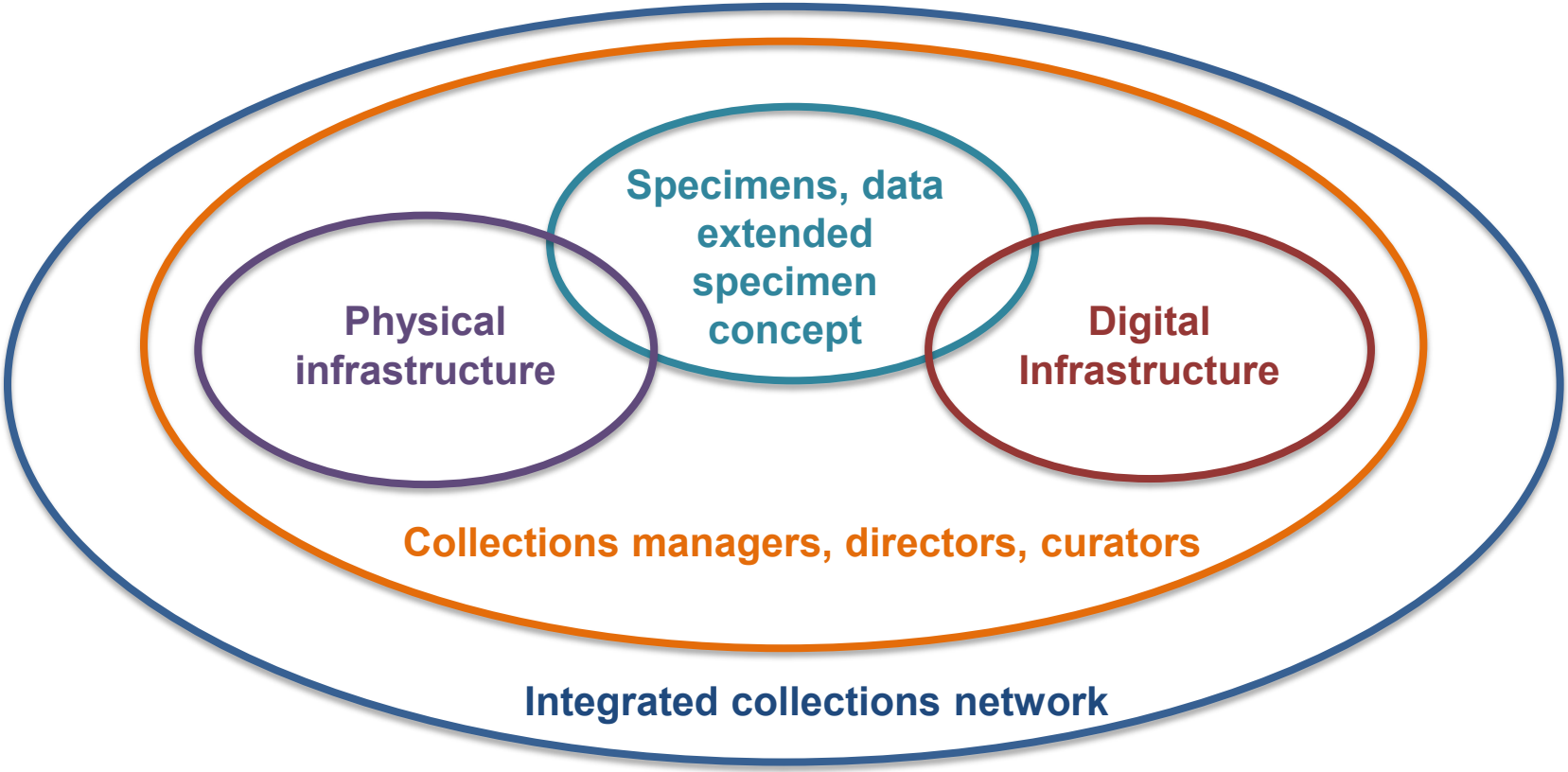
# Report Roadmap

- In chapters 2 and 3, the committee highlights ways in which biological collections contribute to science, education, and society.

- The committee recognizes that future success of biological collections depends on addressing four interrelated issues:
  1) Upgrade and maintain physical infrastructure and growth of collections (Chapter 4);
  2) Develop and maintain tools and processes needed to transform digital data to an easily accessible and integrated cyberinfrastructure (Chapter 5);
  3) Recruit, train, and support the workforce of the future (Chapter 6); and
  4) Ensure long-term financial sustainability (Chapter 7).

- Realizing the committee's vision will require enhanced communication and collaboration within the biological collections community and beyond (Chapter 8).

# Infrastructure has many intersecting components.



# Report Structure

- Each Chapter lays out main challenges for all collections and path forward for the biological collections community.

- Recommendations are offered to:
  - The leadership (directors, managers and curators) of biological collections
  - The biological collection community (professional societies, associations, coordination network etc..)
  - The NSF Directorate for Biological Sciences

# Chapter 4: Building and Maintaining a Robust Infrastructure

**Recommendation 4-1:** The leadership (directors, curators, and managers) of biological collections should assess and define the infrastructure needs of their individual facilities and develop comprehensive strategic plans in accordance with those needs and their strategic missions.

**Recommendation 4-2:** Biological collections should take advantage of existing training opportunities and collaborative platforms at the national and international levels.

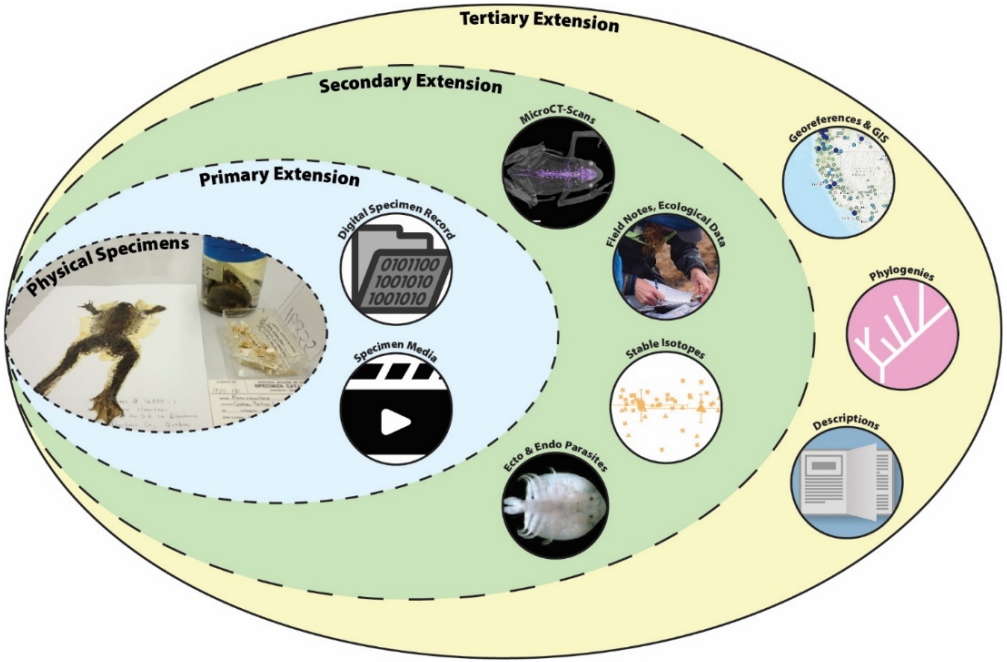
**Recommendation 4-3**: Professional societies, associations, and coordination networks should collaborate and combine efforts aimed at addressing community-level infrastructure needs of the nation's biological collections

**Recommendation 4-4:** The NSF Directorate for Biological Sciences should continue to provide funding support for biological collections infrastructure and expand endeavors to coordinate support within and beyond the Directorate.

# The “Extended Specimen” Concept

Biological collections provide a natural platform for data integration, particularly when *extended specimens* are available to stimulate diverse sets of questions.

The *extended specimen* concept can only be realized and used for maximum effect if dark data are digitized, aggregated, and made accessible through robust cyberinfrastructure.



# Chapter 5: Generating, Integrating, and Accessing Digital Data

## CHALLENGES

- Dark data (undigitized collections)
- Inefficient data pipeline
- Variability in data quality and format
- Limitations affecting data usability
- Inadequate methods for data integration and attribution
- Limited mechanisms to support a cyberinfrastructure that promotes collaboration

## PATHS FORWARD

- Innovative approaches to reducing dark data
- Increase data visibility
- Tools to improve data quality
- Promote integration and attribution
- Develop a national cyberinfrastructure
- Connect data to promote coordination and collaboration

# Conclusions

- Harnessing the opportunity for data-driven discoveries and transdisciplinary collaboration will depend on a continuing effort to digitize new and existing biological collections.
- National and global portals and catalogues have made important contributions to the biological collections community by providing a platform to access, exchange and share data and promote standardization and consistency.
- Integration of specimen data with non-biological data sources will require implementation of a network of cyberinfrastructure resources not yet realized.
- Without a permanent national cyberinfrastructure that supports expanded digitization of dark data, improvement in data quality, collections—both physical and digital—will continue to be underused.

# Recommendations for Individual Biological Collections

**Recommendation 5-1:** The leadership (directors, curators, and managers) of biological collections should provide the necessary mechanisms for staff to keep pace with advances in digitization and data management through training in digitization techniques and publishing of standardized quality data that can be efficiently integrated into portals.

# Recommendations for the Biological Collections Community

**Recommendation 5-2:** Professional societies should initiate and cultivate opportunities for research collaborations within the biological collections community. These collaborations should include working with the computer and data sciences communities to promote the development and implementation of tools to build the cyberinfrastructure (e.g., data storage, annotation, integration, and accessibility) and expand use to a broader range of stakeholders.

# Recommendations for the National Science Foundation

**Recommendation 5-3:** The NSF Directorate for Biological Sciences should continue to provide funding for digitization of biological collections and cyberinfrastructure to support both living and natural history collections. Specifically, NSF should:

- partner with other Directorates within NSF (e.g., physicists, chemists, computer sciences, and education) and beyond (HHS, USDA, FDA, DOI, NOAA, NASA, DOE, etc.);
- establish ongoing mechanisms for the biological collections community to meet, develop best practices, and work towards goals such as establishing and implementing unique identifiers, clear workflows, and standardized data pipelines; and
- promote and fund the development of cyberinfrastructure, tools, and technology to effect the efficient multi-layer integration of data and collections attribution.

# Chapter 6: Cultivating a Highly Skilled Workforce

**Recommendation 6-1:** The leadership of individual collections, host institutions, relevant professional societies, and collections funders should collaborate to develop and strengthen the workforce pipeline through community-level action on issues such as critical skills, workforce analysis, diversity, equity, and inclusion, education and training coherence and alternative staffing models.

**Recommendation 6-2**: As part of its programmatic endeavors to promote a robust biological infrastructure, the NSF Directorate for Biological Sciences should support initiatives that focus explicitly on systemic, systematic, and thoughtful development of the biological collections workforce pipeline. In partnership with other Directorates, such a programmatic focus should encompass future (e.g., students and postdocs) and existing collections personnel (e.g., early-career and senior curators and collections managers), predicated on maintenance and growth of biological collections infrastructure to meet diverse needs of societal import.

# Chapter 7: Securing Financial Sustainability

**Recommendation 7.1:** The leadership (directors, managers and curators) of biological collections should work with business strategists and communication experts to develop business models for financial sustainability and infrastructure of biological collections.

**Recommendation 7.2:** Professional societies should develop extensive networked training and platforms for sharing best practices for financial management and planning and business models for collections of all sizes and types. This could be an ongoing activity centered at a national biological collections center and should include both natural history and living collections together.

**Recommendation 7.3:** The NSF Directorate for Biological Sciences should continue to provide stable, long-term funding to support investigators who rely on biological collections for research and education.

# A framework for collaboration and innovation is needed.

"We are drowning in information, while starving for wisdom. The world will henceforth be run by synthesizers, people able to put together the right information at the right time, think critically about it, and make important choices wisely."

*E.O. Wilson, 1998*

# Chapter 8: Taking Collaborative Action

## The Time to Act is Now

- Biological collections provide the temporal, spatial, and taxonomic samples needed to document effects of changes in the biodiversity of natural and managed ecosystems.
- More than ever, biological collections have an energized community ready to step up to meet basic biological and societal challenges.
- Many of the recommendations are beyond the capability of individual collections to implement.
- Several research communities have established central hubs, multi-tiered networks, associations, or synthesis centers, funded through NSF grants or other federal and state support.
- The biological collection community could leverage the organizational structure of centers and networks as a model to establish an Action Center for Biological Collections.

# A National Action Center for Biological Collection

- The biological collections community needs an inclusive, integrated platform to strengthen the position of biological collections as a unified scientific infrastructure for the nation over the next decade and beyond.
- A national collections-focused action center dedicated to the support and use of biological collections could fill this need.
- Its mission would focus on all biological collections and offer a collaborative platform to provide actionable and lasting solutions for the collection community at large.

# National Decadal Survey

- Many scientific communities work together to set priorities for research and building infrastructure needs to accomplish those priorities.
- A biological collections-focused decadal survey would establish a set of priorities that could only be accomplished with a concerted effort of the community, rather than one biological collection.
- A decadal survey for the biological collections community needs to involve the natural history and living stocks collections communities.
- The collections community needs to make stronger connections with disciplines not traditionally associated with biological collections. Cross-directorate participation in a decadal survey would help to strengthen these connections.

# Chapter 8: Taking Collaborative Action

**Recommendations 8-1:** 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.

**Recommendations 8-2:** 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.

**Recommendations 8-3:** NSF should expand partnership capabilities more broadly across NSF, other federal agencies, international programs, and other sectors to maximize investments in support of Recommendations 8-1 and 8-2.

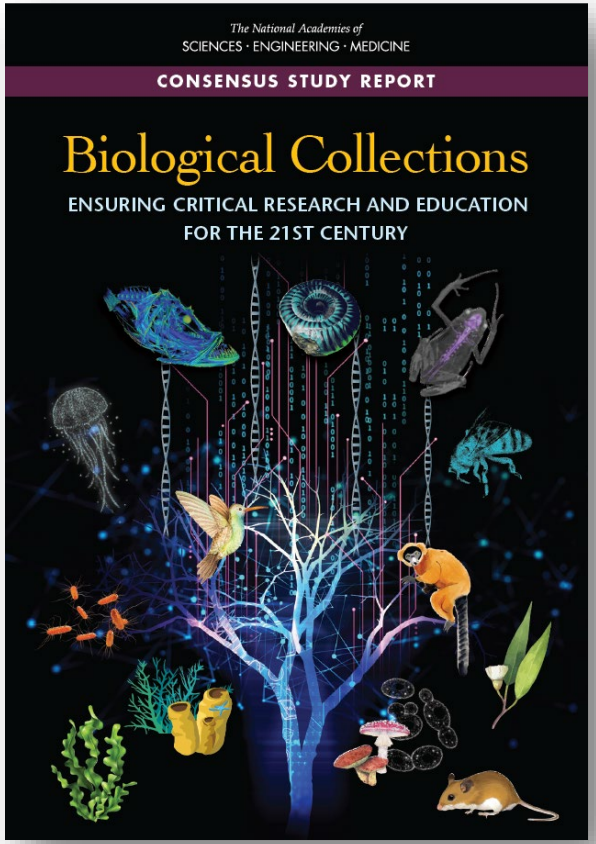
# Acknowledgements

This study was conducted at the request and with the support of the National Science Foundation

- NASEM Staff: Audrey Thévenon, Keegan Sawyer, and Jessica De Mouy
- Invited speakers
- Members of the public
- Report reviewers
- Committee members

The National Academies of
SCIENCES • ENGINEERING • MEDICINE

The National Academies of
SCIENCES · ENGINEERING · MEDICINE




## Where to find information about the report

Visit https://nationalacademies.org/biological-collections

- PDF of the report for free download
- Report highlights
- Webinar recordings
- Briefing slideshow and recording

#NASEMBioCollectionsReport

**Board on Life Sciences**

The National Academies of SCIENCES ENGINEERING MEDICINE

# Biological Collections

## ENSURING CRITICAL RESEARCH AND EDUCATION FOR THE 21ST CENTURY

## Questions?