

An Advisory Report to the U.S. Federal Interagency Working Group on Scientific Collections

*Presentation to the US
Culture Collection Network
April 21, 2021*

*Scott E. Miller and
David E. Schindel
Smithsonian Institution*

IWGSC Background (1 of 2)

- IWGSC created in 2005, part of White House National S&T Council (NSTC)
- Reflects federal view of collections as S&T infrastructure
- 2009 IWGSC “Green Report” with 7 recommendations:
 - **Develop realistic cost projections**
 - **Improve Web-based information access**
 - **Document and share collection policies**
 - **Create online information clearinghouse**
 - Review legislation governing collections
 - Report periodically to NSTC. OSTP, OMB
 - Continue IWGSC activity beyond 2009



IWGSC Background (2 of 2)

- **January 2011:** America COMPETES Act of 2010 adopts three IWGSC recommendations
- **January 2013:** IWGSC recommends standard contents of collection policies
- **March 2014:** OSTP issues policy requirements for agency collections
- **October 2015:** IWGSC releases online Clearinghouse:
 - Registry of US Federal collections
 - Agency policies for their collections
- **December 2020:** Smithsonian publishes advisory report to IWGSC





Scientific Collections as Infrastructure

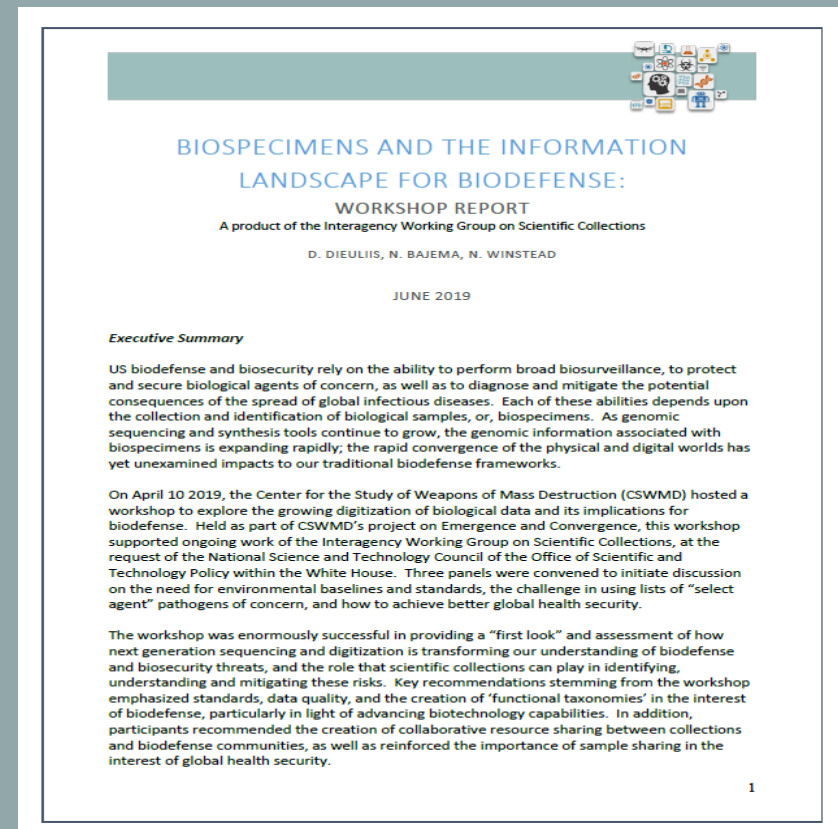
- Law and policy in the USA recognize scientific collections as distributed infrastructure
 - America Competes Act 2010 (42 USC 6624.104)
 - White House Office of Management and Budget memos
 - White House Office of Science and Technology Policy memos (especially 20 March 2014)
- “Scientific collections ... an essential base for developing scientific evidence and ... resource for scientific research, education, and resource management.”
- USG goal is “systematic improvement of the development, management, accessibility, and preservation of scientific collections ...”

Recent and current activities of IWGSC

- Now in NSTC Committee on Science and Technology Enterprise, reflecting role of collections as science infrastructure
- Continue to help federal agencies refine and update policies
- Work with ORCID on linking collections resources to data and publications, making them more accessible, information rich, and tracking impact
- Work with Global Biodiversity Information Facility (GBIF), new host for Global Registry of Scientific Collections
- Promote role of collections in the bioeconomy (e.g., NASEM study)
- Workshop at National Defense University on biosecurity perspective of collections, their genomic and other data in 2019
- Study on economic models for collections published 2020
- Revising “Green Report” for publication 2021

Biospecimens and the Information Landscape for Biodefense

- Workshop with National Defense University, April 2019
- “First look” assessment of how next generation sequencing and digitization is transforming our understanding biodefense and biosecurity threats, and the role that scientific collections can play in identifying, understanding and mitigating these risks.



Economic Study Group Report

- IWGSC and this report based on and addresses Federal collections, but relevant to non-Federal also
- Considers all scientific disciplines (“microbes to moon rocks”) and types of object-based collections
- Reflects input from scientists, economists, collections professionals, program policy specialists
- Describes methods for projecting costs and documenting benefits of long-term “institutional collections”
- Discusses cost recovery through user fees
- Addresses implications for evidence-based management and policies

Assumptions and Definitions

The Report assumes:

- Costs of collections means benefits minus operating and opportunity costs
- Agency missions include programmatic goals as well as legislative and regulatory mandates

Project collections are created for a specific project or goal, and are managed and used by that project

Institutional collections are preserved long-term and managed by collections professionals for future use

Non-renewable collections have unique objects that cannot be replaced; destructive sampling is a management issue

Renewable collections have objects that can be sampled destructively because replacements can be grown or manufactured

Six Services Provided

1. ***Accessioning*** material into collection
2. ***Preserving*** and maintaining contents
3. ***Documenting*** holdings and disseminating information
4. ***Providing access*** to qualified users
5. ***Data curation*** (error correction, adding metadata, linkage to publications and online data)
6. ***Education and outreach*** to increase public understanding

Operating Costs

- Institutional collections can offer some or all of six general services. Not applicable to project collections
- Per unit costs of operating collections vary by:
 - the number and amount of services provided,
 - the size and type of collection/preservation,
 - geographic location, and
 - agency budgeting method,
- The list of standard services provides a framework for (re)constructing operating budgets.
- Comparing operating costs across collections only makes sense in the context of the services provided.

Five Methods for Estimating Benefits

- 1. Technology/Knowledge transfer (“Value chains”)**
 - USDA National Plant Germplasm, ARS Culture Collection, CDC National Health and Nutrition Survey
- 2. Success Stories (“Winning lottery tickets”)**
 - Yellowstone National Park microbes and biotechnology
- 3. Option Value (“Insurance policies”)**
 - FDA Foodborne Bacteria Collections
- 4. Value added by users (“Co-investment”)**
 - USGS Core Research Center
- 5. Counter-factual Scenarios (“It’s a Wonderful Life”)**
 - Reference collections for USDA/APHIS border inspections and NIST Standard Reference Materials

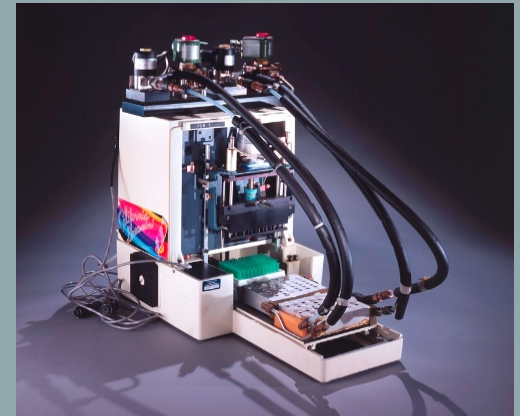
Approach 1. Tech/Knowledge Transfer

- **“Value chain”**: Something from a collection provides a starting point in wealth generation, e.g.:
 - Drug development
 - Bio-inspired design
- Collections are only one part of R&D value chain; how to partition benefits among parts?
- Hard to document process due to delays, multiple components in value chain
- USDA Tech Transfer Reports:
 - **National Plant Germplasm System**, used for plant breeding
 - **ARS Culture Collection**: Microbes with industrial users
- **CDC NHANES national health survey samples and data**
 - Serum, plasma and urine samples collected with health exams
 - Analyzed to establish statistical distributions of diseases
 - NHANES data used to calibrate “normal range” for new diagnostics



Approach 2. Success Stories

- **“Winning lottery ticket”**: Rare events in which collections play a pivotal role, often in unanticipated ways
- Enormous (but hard to calculate) socioeconomic impact
 - Biomedical and wildlife collections that help predict epidemics
 - Collections of agricultural pests that help prevent crop failure
- Impossible to predict occurrence or anticipated economic benefits
- Often doesn't reflect the normal, everyday activity of collections
- Impact of collections may be indirect, delayed, hard to trace causality
- **Yellowstone National Park** hot spring microbes and PCR



Approach 3. Option Value

- **“Insurance Policies”**: Anticipating and preparing for future emergencies
- Forward-looking counterpart to Success Stories. Data on the costs of past emergencies provide range of potential emergencies
- Requires that collections can demonstrate potential to avoid or mitigate the emergencies
- **USDA National Plant Germplasm System**
 - Collects, preserves, characterizes crop varieties and their relatives
 - Evaluates and develops and accessions for developing new food crops
 - Maintains seedbanks to re-establish farming after famine events
- **HHS/FDA Foodborne Pathogen collection**
 - Used for developing methods for detection of naturally occurring pathogens in seafood
 - State health agencies responsible for response to disease outbreaks use FDA’s samples to identify sources of outbreaks



Approach 4. Value Added by Users

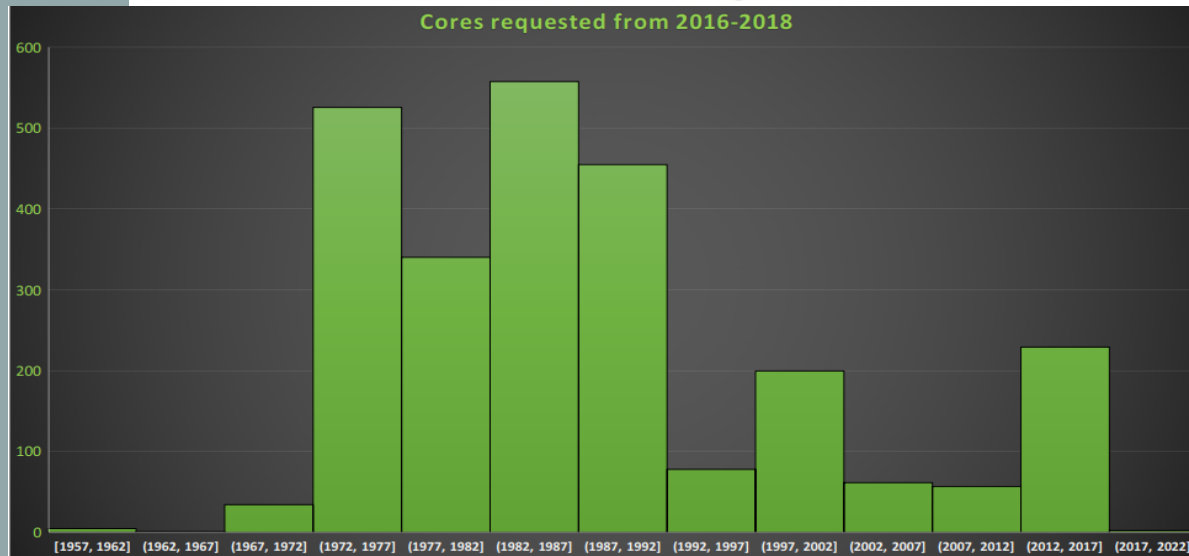
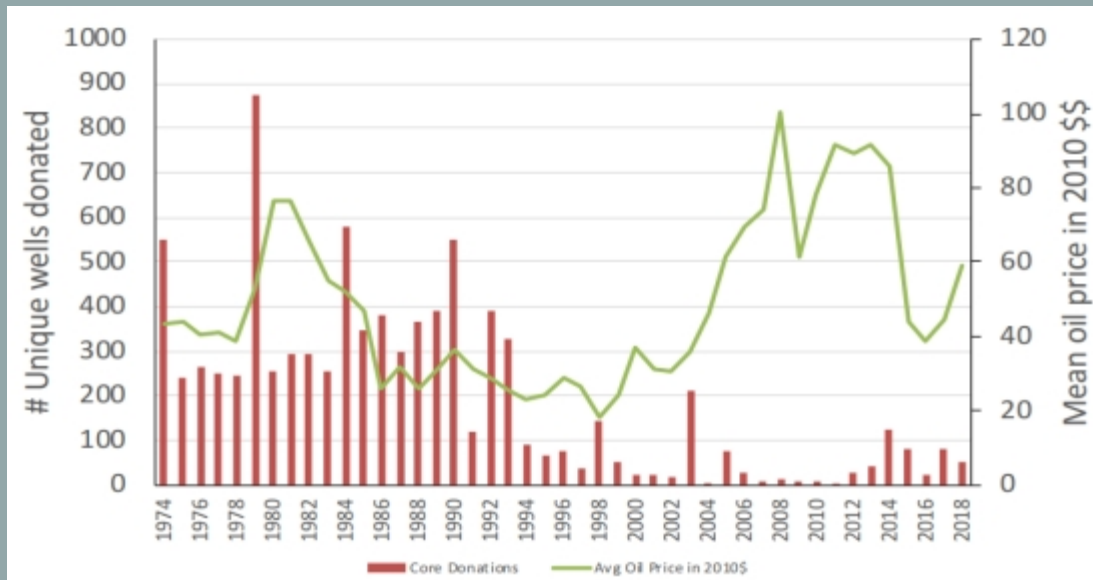
- **“Co-Investment”** by users makes the collection more valuable
- Results in greater: reliability; discoverability; diversity and volume of uses; cost of replacement through:
 - Data curation by users is an unpaid service
 - Analytical data and research publications linked to collection record
 - Preparations returned to collection (e.g., rock thin sections, DNA extracts)
- **USGS Core Research Center**
 - Rock cores from intramural research and donated by companies
 - No user fees, but users must provide analytical data, thin sections produced, publications within time limit
- CDC NHANES curates analytical results from users



Does User Interest Decline Over Time?

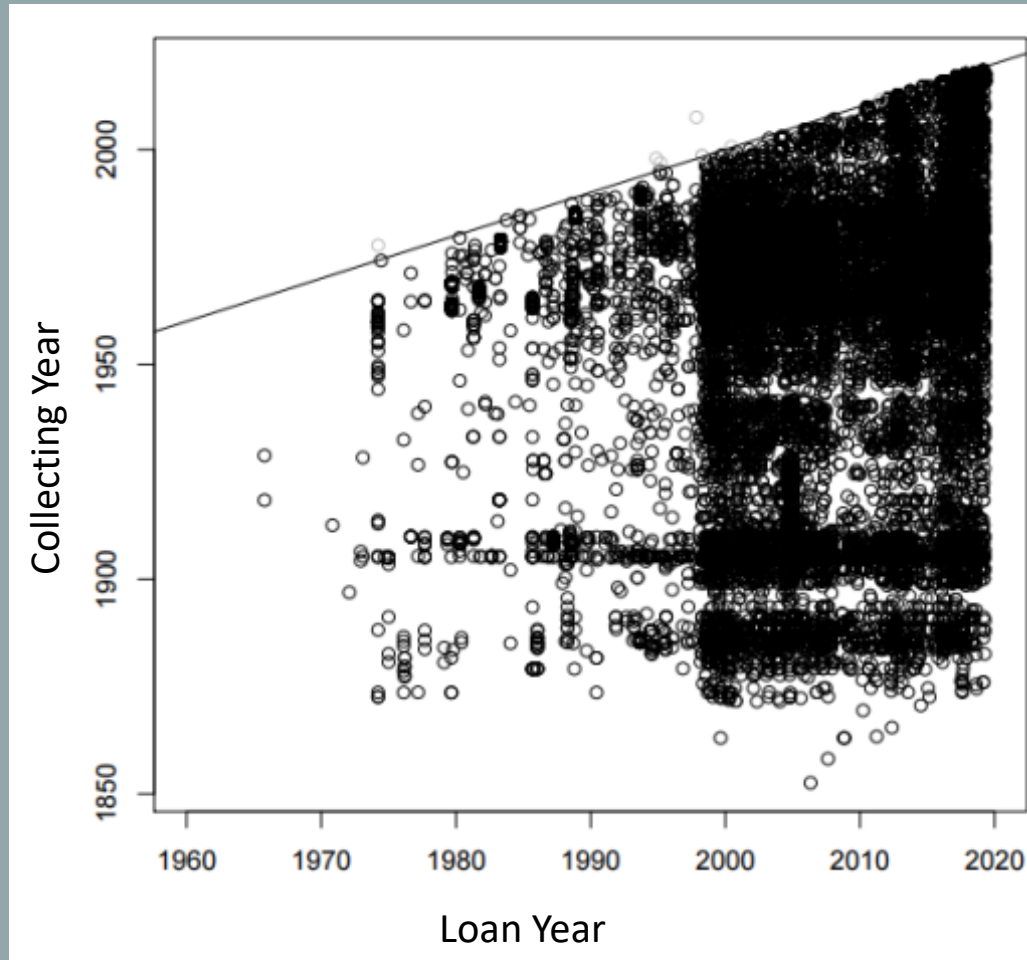
- Compared distributions of accession dates and user access requests
- Datasets from:
 - USGS Core Research Center
 - Smithsonian National Museum of Natural History
- **Patterns indicate that objects retain user interest over time**
- **Users can add value through co-investment in objects regardless of their collection date**

User Requests versus Accession Dates



Source: Lindsay Powers,
US Geological Survey

Users Requests versus Collection Dates



Data from National Museum of Natural History
Invertebrate Zoology Department

Approach 5. Counter-factual Scenarios

■ “It’s a Wonderful Life”

- What would users do if a collection didn’t exist?
- What would users be willing to pay to have it?



■ NIST Standard Reference Materials

- Enables companies to meet regulatory standards
- Survey of users’ estimates of cost and delay to create or find an alternative
- Paperwork Reduction Act limits sample size, reliability of results

■ USDA Plant Protection and Quarantine

- Reference collections used for border inspection of plant imports for pests, invasive species
- Of imports with insects, fungi, etc., which are safe or can be treated?
- Calculate value of all imports in which benign or treatable insects, fungi, etc. were found



Method	Principal Advantages	Principal Disadvantages
Technology/ Knowledge Transfer	<ul style="list-style-type: none"> Based on tangible outcomes, often monetary Usually connected to normal collections-based work Can be expressed in quantitative terms (e.g., benefit-cost ratios) 	<ul style="list-style-type: none"> Difficult to connect use of collection to ultimate outcome (delays, other contributors to process) Sometimes serendipitous
Success Stories	<ul style="list-style-type: none"> Can be dramatic, high value Easily understood 	<ul style="list-style-type: none"> Based on rare events that can't be predicted Can be serendipitous and unrelated to normal collections-based work
Option Value	<ul style="list-style-type: none"> Can be dramatic, high value Connects to historical events, easily understood 	<ul style="list-style-type: none"> Based on probability of future use, not past performance
Value Added by Users	<ul style="list-style-type: none"> Based on normal collection activities Highlights patterns of collection use Can be expressed in quantitative terms (e.g., rates of return) 	<ul style="list-style-type: none"> Requires cooperation of users Requires data curation Uses narrow definition of "value" (i.e., value to users, not others)
Counterfactual Scenarios	<ul style="list-style-type: none"> Highlights unique role of collections Based on customer feedback and/or performance data Can be expressed in quantitative terms (e.g., rates of return) 	<ul style="list-style-type: none"> Customer surveys can be expensive, labor-intensive Limitations on federal surveys (Paperwork Reduction Act) Distrust of survey results

Findings and Recommendations (1 of 2)

- The services offered by a collection determine the benefits generated, e.g.,
 - Preserving and maintaining objects extends their useful life
 - Providing user access and data curation expand the pool of potential users
 - Data curation creates a “virtuous cycle” that improves user access which leads to greater co-investment by users
 - Education and Outreach increases awareness, appreciation, and public support
- Accessioning and preserving are basic, can limit support available for other services
- The framework of interlocking services, costs, and benefits can facilitate evidence-based decisions about balance of services provided

Findings and Recommendations (2 of 2)

- Agencies have a choice of several methods for estimating and documenting benefits generated by their collections
 - Reflect agency/collection mission
 - Consider preferences of stakeholder audience
 - Surveys vs. program data?
 - Qualitative vs. quantitative evidence?
 - Retrospective vs. prospective impact?
 - Counterfactual Scenarios can rely on user surveys but their use by Federal collections may be limited by Paperwork Reduction Act

Acknowledgments (1 of 2)

The Co-chairs of the IWGSC wish to acknowledge the members of the **Economic Study Group** for their contributions to this report:

Chair: David E. Schindel, SI

Reed Beaman, NSF

Brad Bowzard, CDC

Vanessa Burrows, FDA

Jeffrey DeGrasse, FDA

Kevin Hackett, ARS

Marianne Henderson, NIH

Susan Lukacs, CDC

Gerry McQuillan, CDC

Scott Miller, SI

Tom Moreland, FS

OSTP Liaison: Ian Watson

Gerry McQuillan, CDC

Scott Miller, SI

Tom Moreland, FS

Emily Pindilli, USGS

Cassidy R. Sugimoto, NSF

John P. Swann, FDA

Michael Walsh, NIST

Ellen Wann, NIH

Paul Zankowski, USDA

The members of the IWGSC, representing their respective departments and agencies, provided valuable reviews and suggestions during development of this report.

Acknowledgments (2 of 2)

The Co-chairs also thank the following presenters who provided valuable input to the Economic Study Group

Lindsay Powers, USGS

Erik Lichtenberg and **Lars Olson**, Univ. Maryland

Douglas Gollin, Oxford Univ.

Jessica Jones, HHS/FDA Gulf Coast Seafood Laboratory

Peter Bretting, **USDA/ARS Crop Production and Protection**

Todd Ward, USDA/ARS Culture Collection

Kelly Day-Rubenstein and **Paul Heisey**, USDA/Economic Research Service

Abhi Rao, HHS/NIH

The Co-chairs gratefully acknowledge the following individuals who provided support to the Economic Study Group:

Eileen Graham, SI

Keith Crane, **Lauren Bartels**, and **Thomas Olszewski**, IDA Science and Technology Policy Institute

Deborah Paul, Univ. Florida

Additional information:

- IWGSC Clearinghouse: <https://iwgsc.nal.usda.gov>
- IWGSC Green Report:
https://iwgsc.nal.usda.gov/sites/default/files/IWGSC_GreenReport_FINAL_2009.pdf
- **Report webpage:**
<https://iwgsc.nal.usda.gov/economic-analyses-federal-scientific-collections>
 - Download the report
 - Short introductory video
 - Standard PowerPoint presentation
 - Press release

Questions?