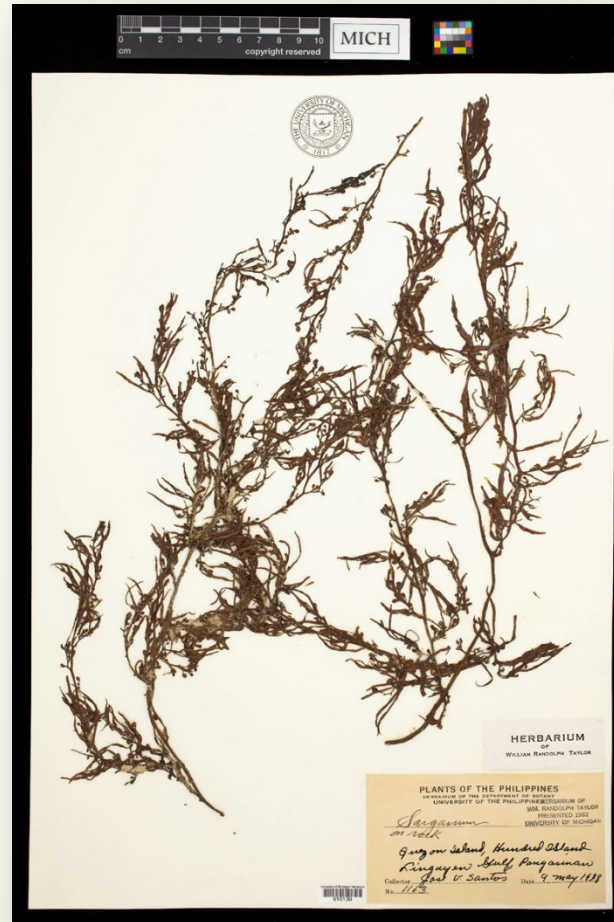
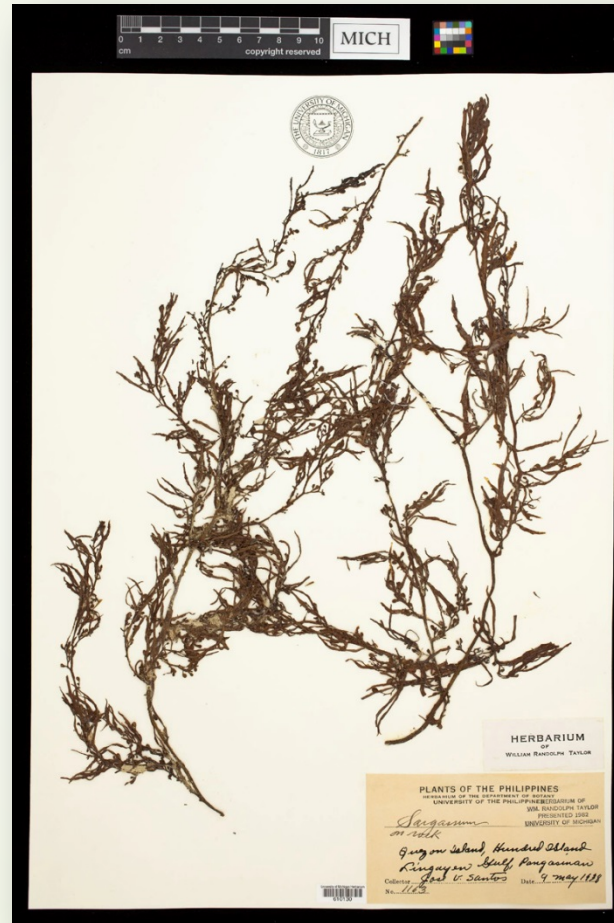


The Role of Small Herbaria in Large Digitization Projects



**Chris Neefus, Albion Hodgdon Herbarium (NHA)
University of New Hampshire, Durham, New Hampshire, USA**

The Role of Small Herbaria in Large Digitization Projects?



**Chris Neefus, Albion Hodgdon Herbarium (NHA)
University of New Hampshire, Durham, New Hampshire, USA**

Some Background

NIBA – Network Integrated Biological Alliance – a 2010 strategic plan for a 10-year effort to digitize and mobilize the scientific information associated with biological specimens held in U.S. research collections.



NSF ADBC – Advancing Digitization of Biodiversity Collections – a 10-year program to digitize biodiversity specimens stored in all U.S. Collections. Total program funding is \$10 million per year. Fund 3-4 new TCN projects each year.

TCN – Thematic Collection Network – Digitization projects funded by NSF ADBC are networks of institutions with a strategy for digitizing information that addresses a particular research theme, such as impacts of climate change on the biota of a region.



iDigBio - Integrated Digitized Biocollections - funded by NSF to act as a hub for the TCNs and other digitization projects. iDigBio is a national resource that integrates data, coordinates and promotes tool development, and fostering partnerships and innovations.

NSF ADBC Projects

2011

A Collections Digitization Framework for the 21st Century.

PI: Lawrence Page, University of Florida

InvertNet--An Integrative Platform for Research on Environmental Change, Species Discovery and Identification.

PI: Christopher Dietrich, University of Illinois, Urbana-Champaign

Plants, Herbivores and Parasitoids: A Model System for the Study of Tri-Trophic Associations.

PI: Randall T. Schuh, American Museum of Natural History

North American Lichens and Bryophytes: Sensitive Indicators of Environmental Quality and Change.

PI: Corinna Gries, University of Wisconsin, Madison

NSF ADBC Projects

2012

Mobilizing New England Vascular Plant Specimen Data to Track Environmental Changes.

PI: Patrick Sweeney, Yale University

Digitizing Fossils to Enable New Syntheses in Biogeography - Creating a PALEONICHES-TCN.

PI: Bruce Lieberman, University of Kansas

The Macrofungi Collection Consortium: Unlocking a Biodiversity Resource for Understanding Biotic Interactions, Nutrient Cycling and Human Affairs.

PI: Barbara Thiers, New York Botanical Garden

Southwest Collections of Arthropods Network (SCAN): A Model for Collections Digitization to Promote Taxonomic and Ecological Research.

PI: Neil Cobb, Northern Arizona University

NSF ADBC Projects

2013

Fossil Insect Collaborative: A Deep-Time Approach to Studying Diversification and Response to Environmental Change.

PI: Dena M. Smith, University of Colorado Boulder

Developing a Centralized Digital Archive of Vouchered Animal Communication Signals.

PI: Michael S. Webster, Cornell University

The Macroalgal Herbarium Consortium: Accessing 150 Years of Specimen Data to Understand Changes in the Marine/Aquatic Environment.

PI: Christopher D. Neefus, University of New Hampshire

ADBC Herbarium Projects

North American Lichens and Bryophytes: Sensitive Indicators of Environmental Quality and Change.

2.3 Million Specimens in 60 Collections

Mobilizing New England Vascular Plant Specimen Data to Track Environmental Changes.



1.3 Million Specimens in 15 Collections

The Macrofungi Collection Consortium: Unlocking a Biodiversity Resource for Understanding Biotic Interactions, Nutrient Cycling and Human Affairs.

1.4 Million Specimens in 24 Collections

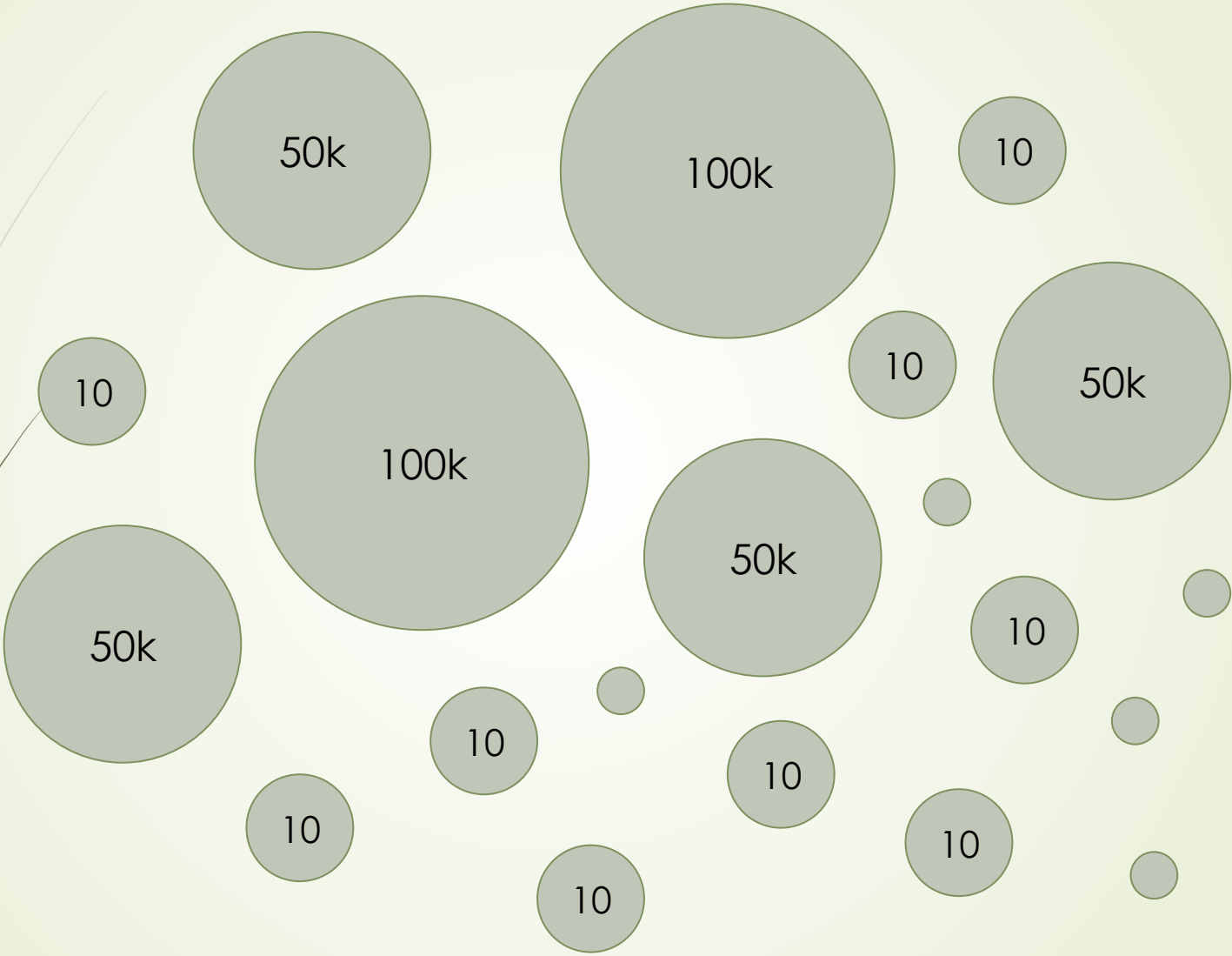
The Macroalgal Herbarium Consortium: Accessing 150 Years of Specimen Data to Understand Changes in the Marine/Aquatic Environment.

1.2 Million Specimens in 49 Collections

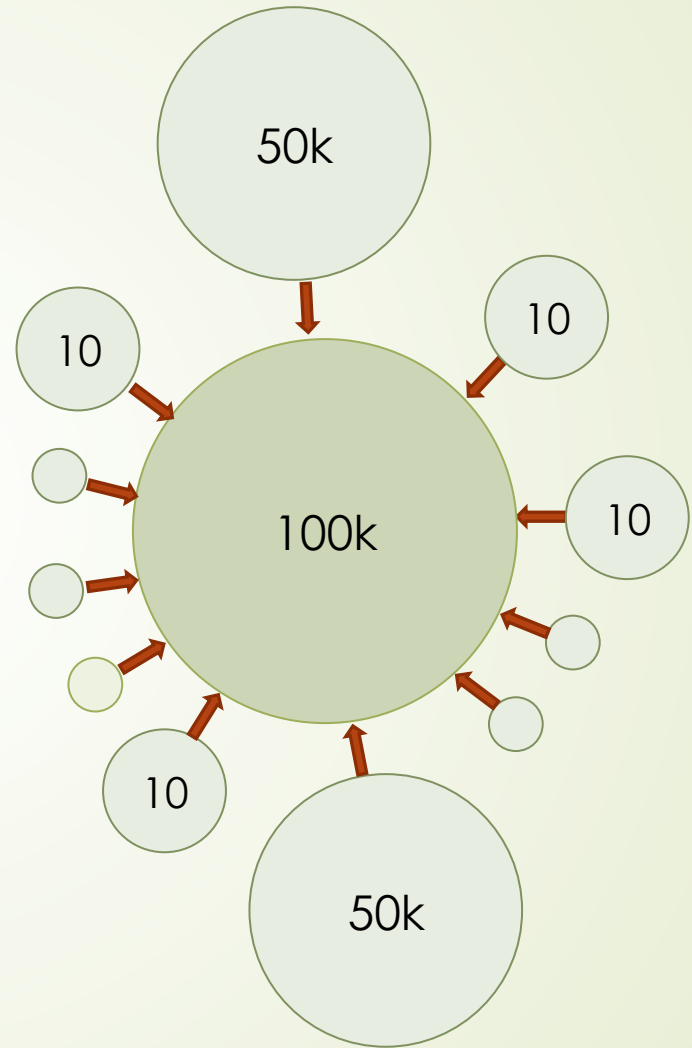
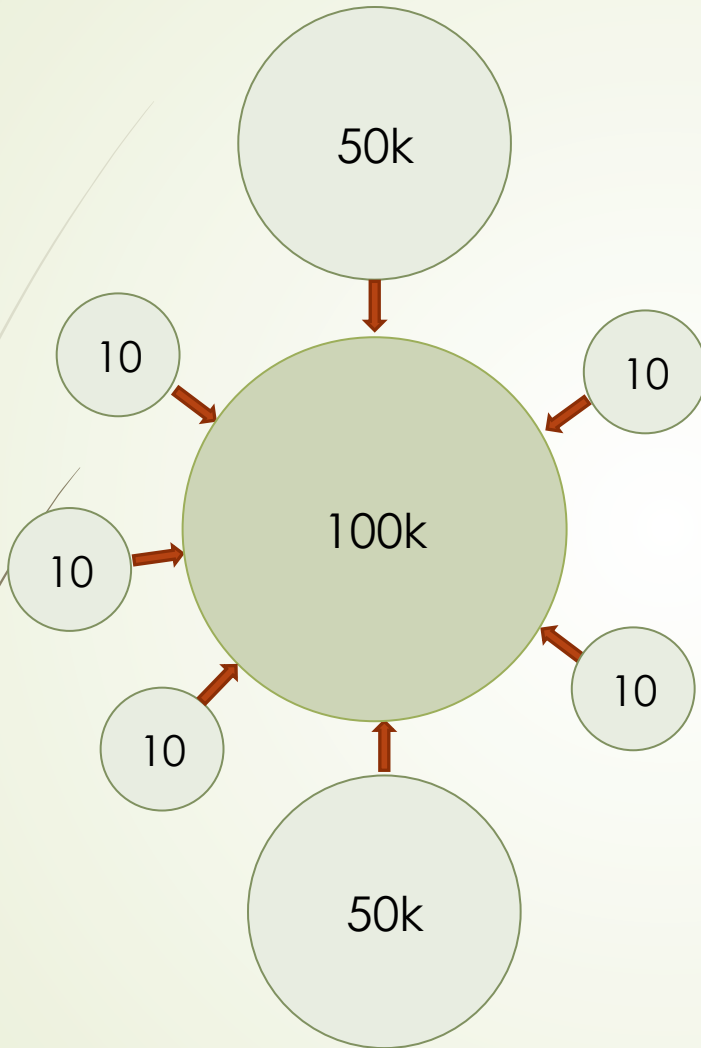


What is the Role of
Small Herbaria
in Projects Like These?

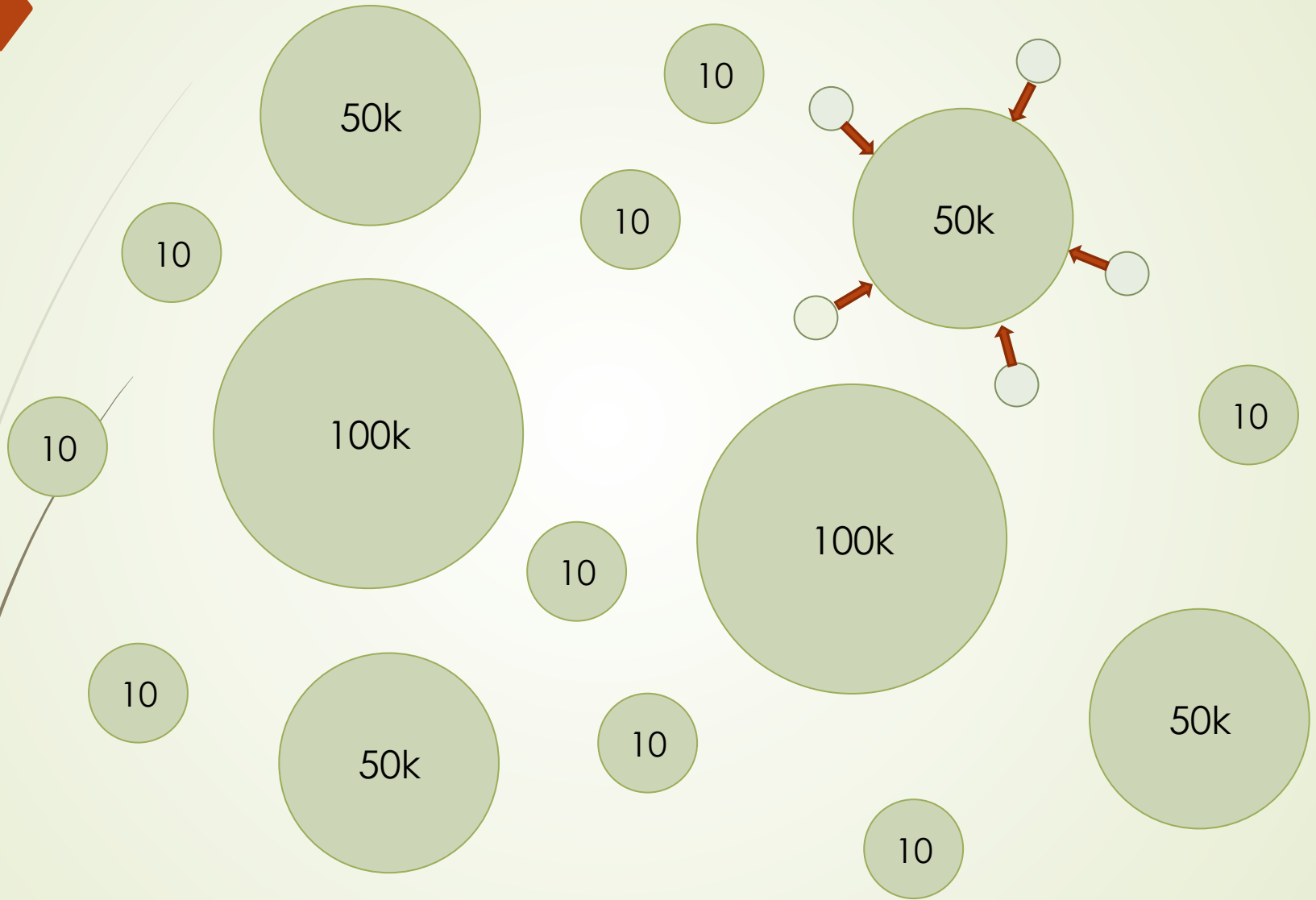
Hypothetical Herbarium TCN



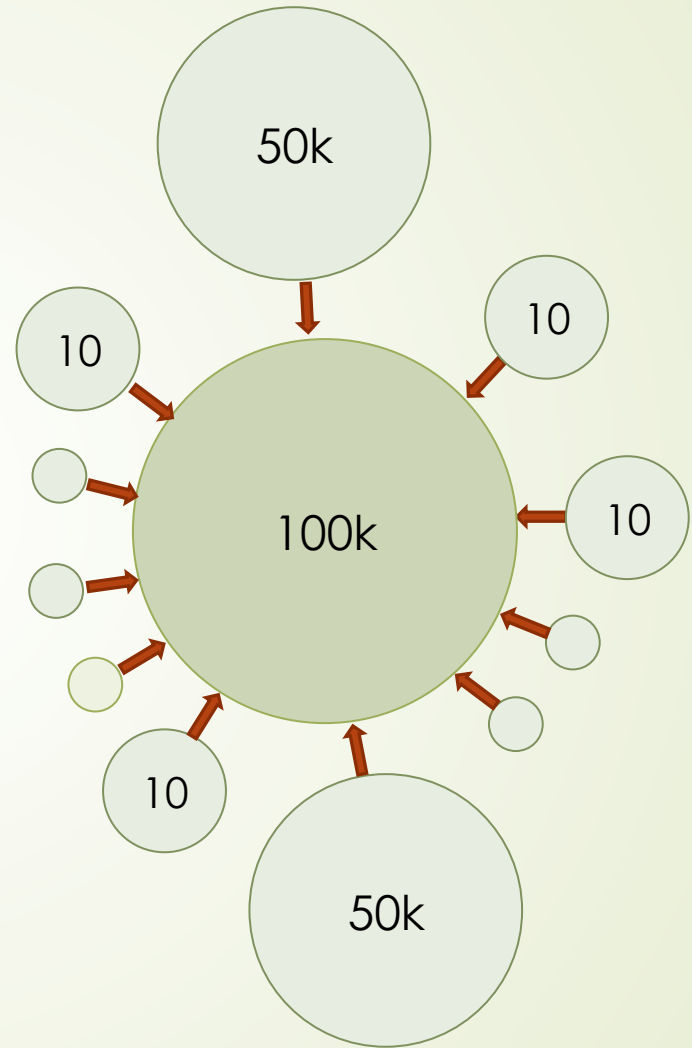
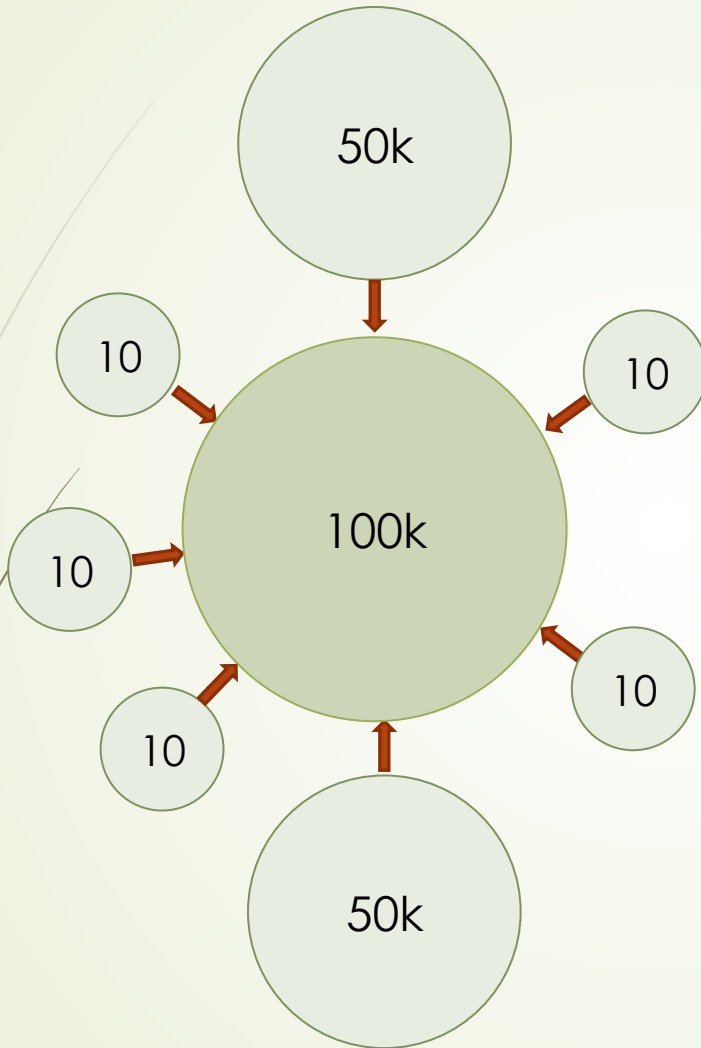
Centralized Digitization Model



Distributed Digitization Model



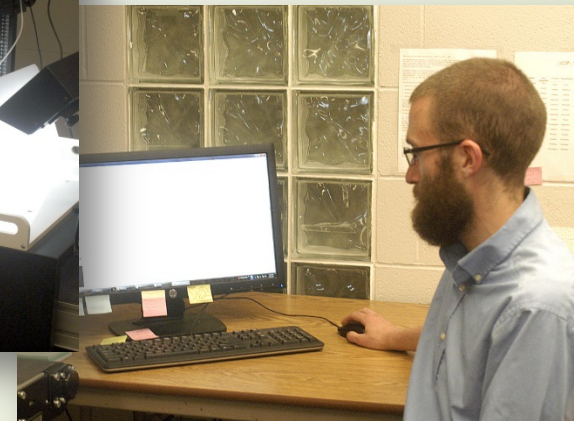
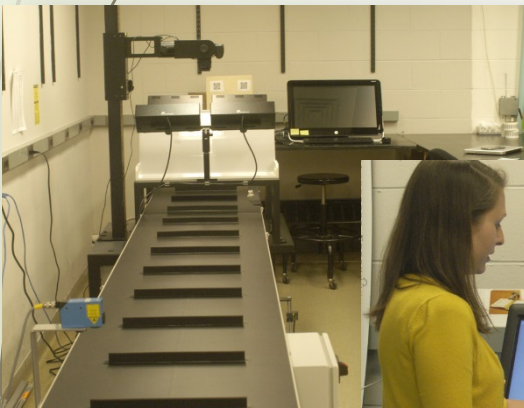
Centralized Digitization Model



Centralized Model Considerations

From the Project Directors' Perspective

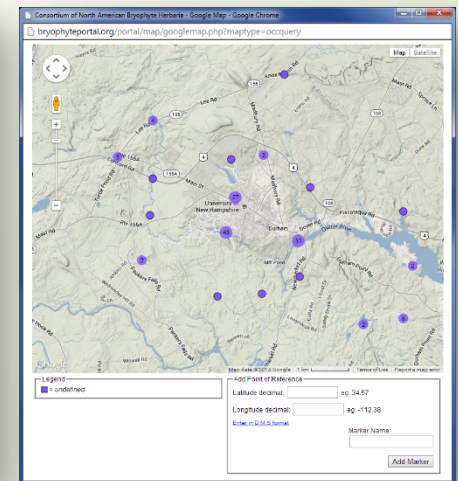
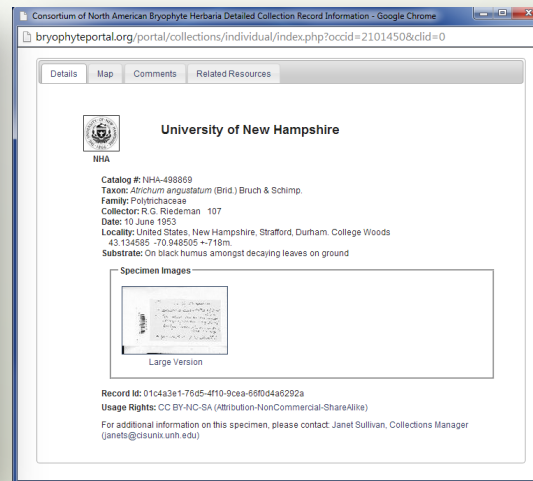
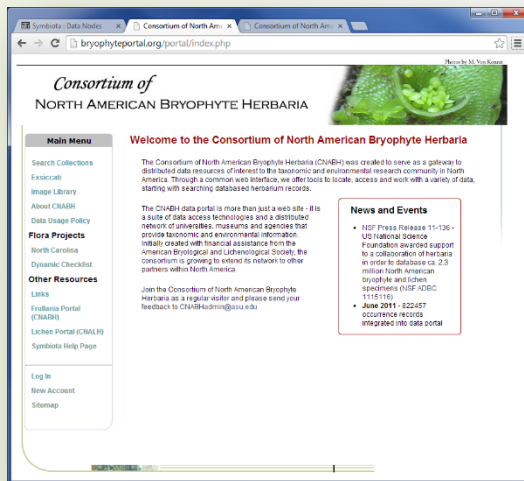
- Simplified Project Administration
 - Proposal Preparation – Budgets, C.V.s, Current & Pending, etc.
 - Progress Reports
- Full-time project staff
 - Fewer people to train and supervise
 - Proficiency improves with experience
 - Sense of project “ownership”
- Custom made high through-put equipment



Centralized Model Considerations

From the Small Herbarium Perspective

- Collection is digitized “for free”
 - Minimal time commitment
 - Little or no financial cost
- Digitized collection accessible via project portal
 - Increased collection utilization
 - Data integration with other related collections
 - Access to tools for georeferencing, mapping, taxonomic updates
- Increased collaboration with other herbaria



Centralized Model Considerations

Major Budget Item

Equipment - Design and construct 2 high through-put systems and custom control software.

Research & Development	\$	80,000
Construct 2 systems @ \$35,000 each	\$	70,000

Technicians - 1 full time and 1 half time technician at each of the Digitization Centers for 2 years @ \$40,000 per FTE plus benefits

\$60k x 1.5 FTE x 2 DCs x 2 yrs	\$	360,000
---------------------------------	----	---------

Training - Travel costs to send 2 technicians to 2 training sessions
2 people x 2 sessions x \$2,000

\$	8,000
----	-------

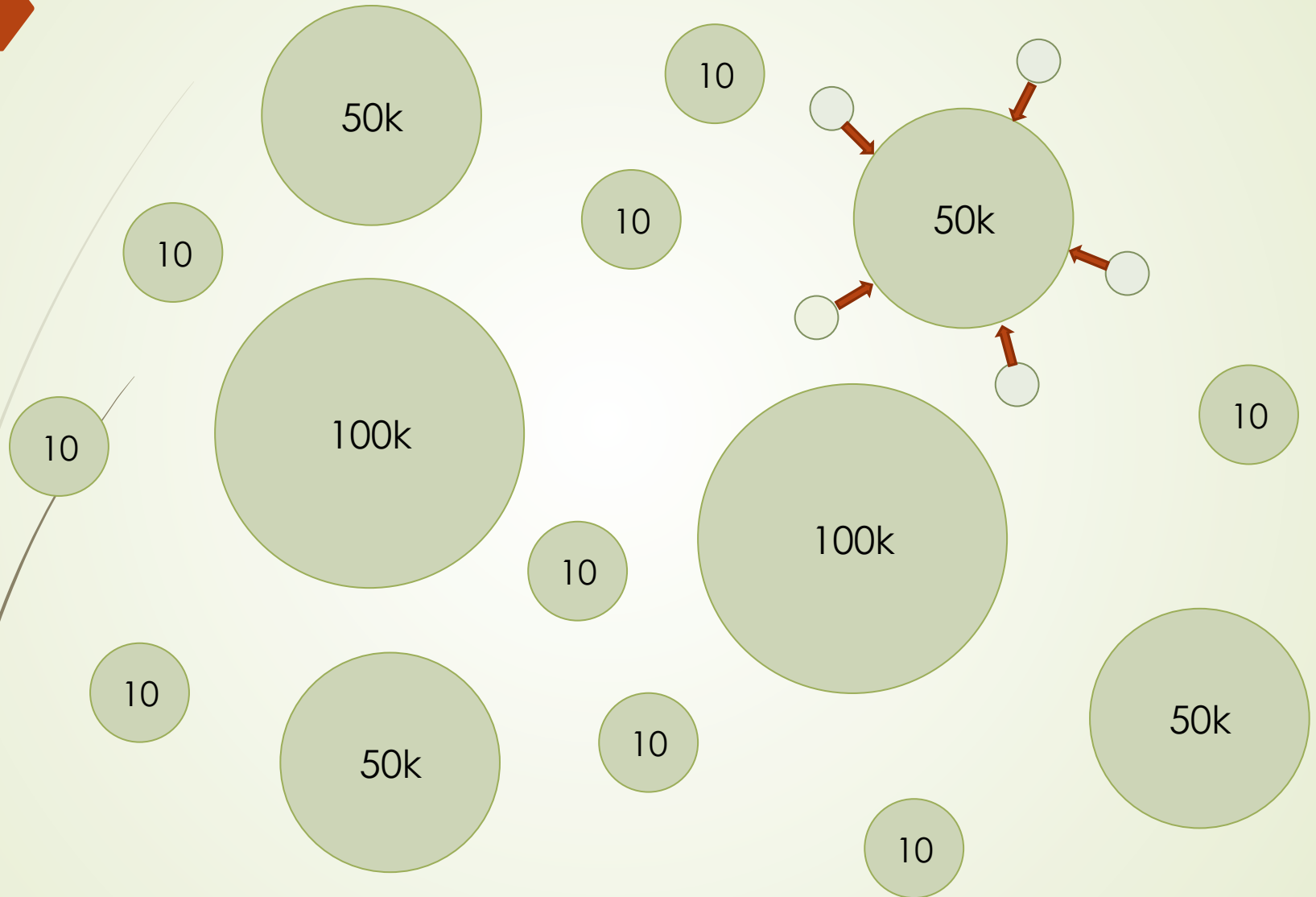
Shipping - 300,000 specimens to and from a Digitization Center with 200 specimens per carton

1500 cartons x \$25 x 2 ways	\$	<u>75,000</u>
------------------------------	----	---------------

Total \$ 593,000

*Based on through-put of 500 specimens per day per Digitization Center the project will take 500 days to complete. With a 5-day work-week, that's **2 years**.*

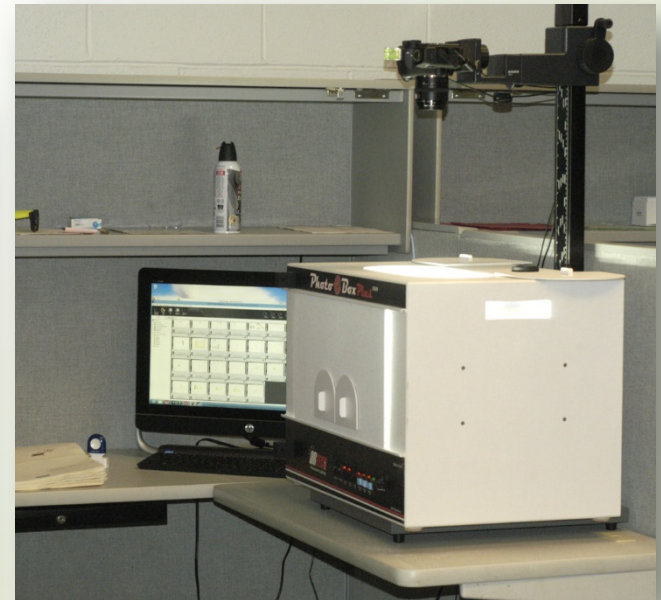
Distributed Digitization Model



Distributed Model Considerations

From the Project Directors' Perspective

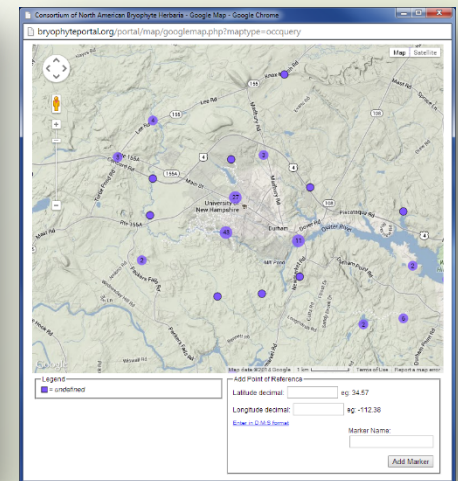
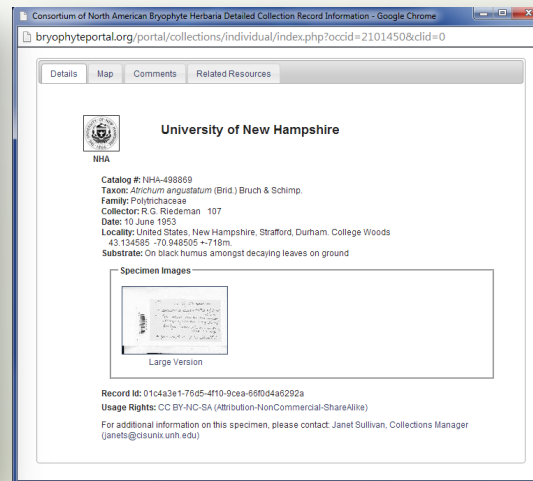
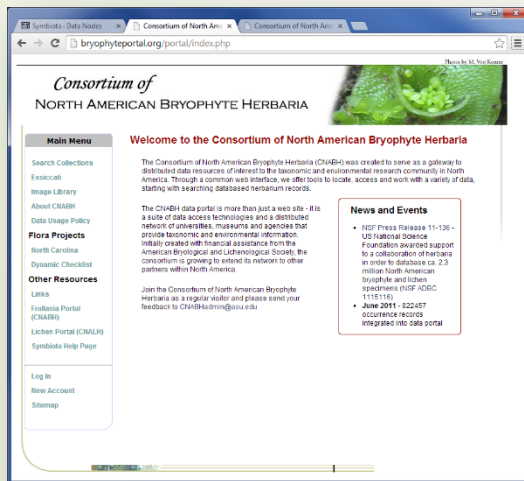
- Complex Project Administration
 - Proposal Preparation – 15 Budgets, 15+ C.V.s, 15+ Current & Pending, etc.
 - 15 Progress Reports
 - Larger project meetings
- Part-time project staff, lots of students
 - Lots of people to train and supervise
 - High turnover rate for student help
- 15 off-the-shelf imaging stations



Distributed Model Considerations

From the Small Herbarium Perspective

- Collection is digitized in-house
 - Significant time commitment
- Digitized collection accessible via project portal
 - Increased collection utilization
 - Data integration with other related collections
 - Access to tools for georeferencing, mapping, taxonomic updates
- Greatly increased collaboration with other herbaria



Distributed Model Considerations

Distributed Model

Equipment - Using off-the-shelf components.

15 Imaging Stations x \$6,000 \$ 90,000

Part-time Technicians - The 2 larger collections will each have 1 part time (75%) technician for 1.5 years @ \$30,000 per year plus benefits

\$45k x 1 technician x 2 DCs x 1.5 yrs \$ 135,000

Students/Interns - Digitization at the 13 other Digitizing Institutions will be done by students or interns supervised by existing herbarium staff. Students are paid hourly and work 10-20 hours per week.

\$9 per hour x 26,000 hours \$ 234,000

Existing Herbarium Staff - At herbaria without technicians, existing staff will oversee the project, train and supervise students. One month of salary per year will be covered by the project.

4 herbaria x 2 years x \$6,000 \$ 48,000

9 herbaria x 1 year x \$6,000 \$ 54,000

Training - Travel costs to send 2 technicians and 13 staff members to training sessions


15 people x \$2,000 \$ 30,000

Shipping - 10,000 specimens to and from a Digitization Center with 200 specimens per carton

50 cartons x \$25 x 2 ways \$ 2,500

Total \$ 593,500

Based on through-put of 15 specimens per student per hour. Smaller collections (10k) will take 6 months based on a 2 students working 15 hours/week. The 50k collections will take 1.5 years with 3 students. The 100k collections will take 1.5 years with 1 technician working 30 hours/week plus 3 students.

- 
- The cost is the same
 - The end product is the same
 - The administration is far more complicated

Why would anyone choose the
Distributed Model?

Distributed Model



- 30-40 students trained
- 15 staff members trained
- 15 herbaria have equipment and experience to digitize additional collections
- Students and staff from 15 institutions have a sense of ownership in the project
- Administration in 15 institutions have a greater appreciation for the value of their herbarium collections and personnel



Comments?

