Seeding Sustainable Digitization

SCAN 2012-2016 (Active - PENs continue)

1. **10** Collections, **6** funded PEN Projects
2. Original focus on Southwest Ground-Dwelling Arthropods
3. **2,050,000** records to date (1.8X expected)
4. **65** non-ADBC funded collections, **1,040,293** digitized records
5. **32** pubs on SCAN or using data
6. **Current focus = North American Arthropods, 18 million records**

LepNet 2016-2020

1. **26** Collections, **4** PEN Proposals
2. Focus on North American Lepidoptera
3. **900k** records, **120k** images,
4. Ingesting Lepidoptera data from **160** North America collections
   > 2 million

SCAN & LepNet

1. Two flavors of homepage, but one database hosted by ACIS (UF)

2. Everyone contributes to taxonomy tables, data cleaning, data standards and development of added-value functions

Search Records

http://scan-bugs.org/portal/collections/index.php

Spatial Module

http://scan-bugs.org/portal/spatial/index.php

Statistics Summa

http://scan-bugs.org

nformation

SCAN Data Portal Statistics (www.scan-bugs)

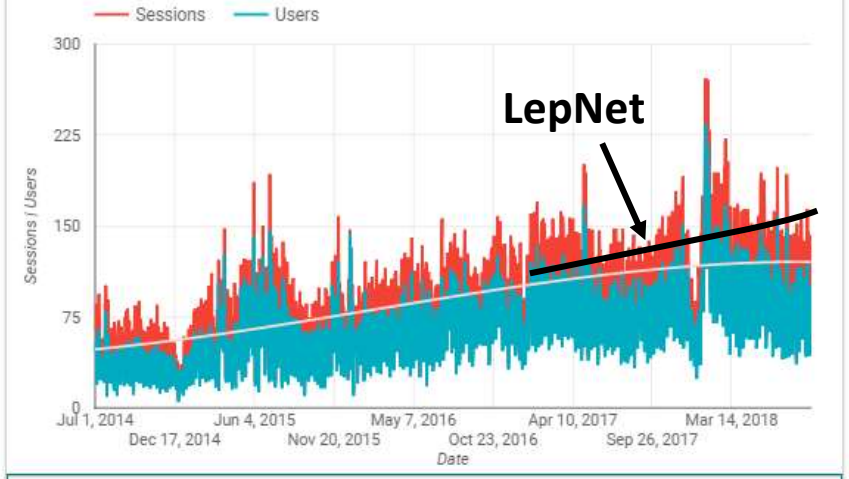
Data from Google Analytics

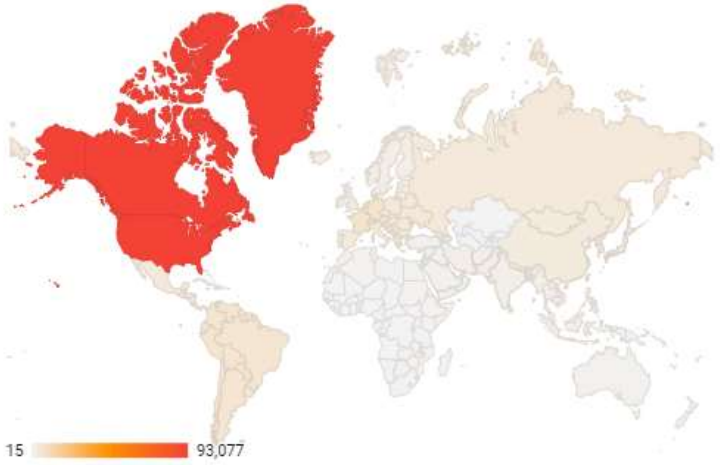
Google Analytics

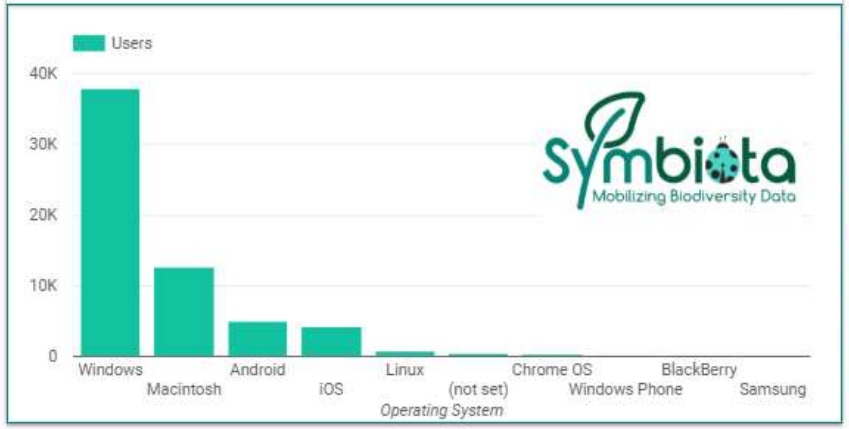
Users	Sessions	Pageviews	Bounce Rate	New Users	Pages / Session	Avg. Session Duration	Number of Sessions per User
62,111	135,822	476,992	49.96%	61,920	3.51	00:04:17	2.19

Default Data
Click to select your dat

Jul 1, 2014 - Aug 29, 2018







	Country	Sessions	Users
1.	United States	90,123	29,912
2.	Germany	2,069	1,646
3.	France	2,667	1,992
4.	Brazil	2,546	2,023
5.	United Kingdom	1,539	1,357
6.	Canada	2,940	1,966
7.	Japan	1,616	1,285
8.	Italy	1,507	1,093
9.	Netherlands	650	529
10.	Mexico	3,688	2,512

1 - 100 / 210

SCAN/LepNet Investment in Symbiota modules

Increasing research capacity of data

Symbiota
Mobilizing Biodiversity Data

Checklist & Keys
Specimen Entry
IPT Harvest
Genetic Link
OC Cr So
2002
2014
2015

eol Encyclopedia of Life

GEOLocate

API

FgBatch: Automated image ID service

Fieldguide

Fieldguide

Computer Vision Biodiversity

PhyloLink

GLOBI

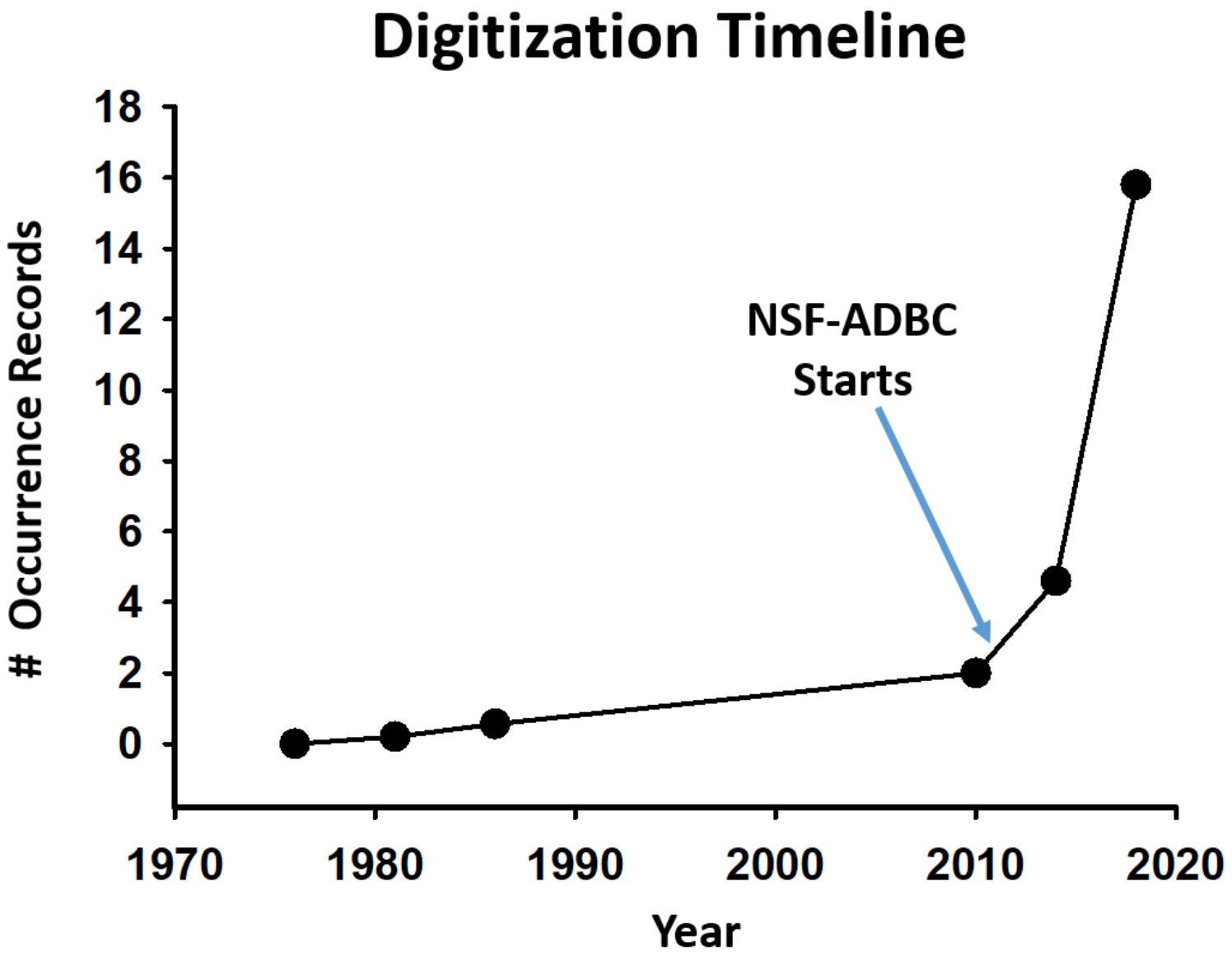
Biotic Associations

Counts!
R-Packages

State of US Entomology Collections: Digitization-Holdings

Sustainable Digitization Post-ADBC

North American Arthropod Occurrences



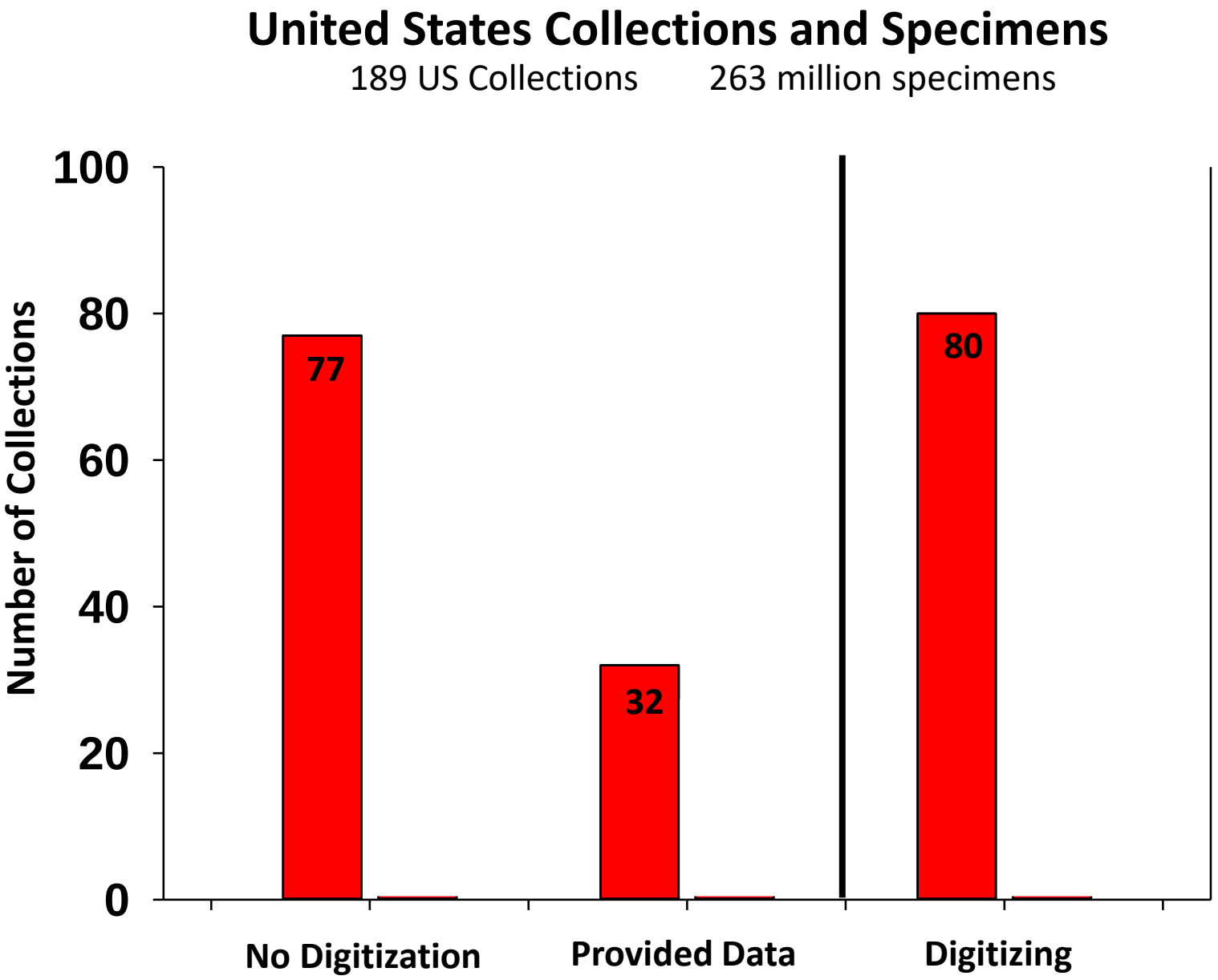
Conclusions

1. ADBC mobilized arthropod digitization in US
(Tri-Trophic, SCAN, InvertEBase, LepNet)

2. It was “Transformational”
   A. Data
   B. Culture

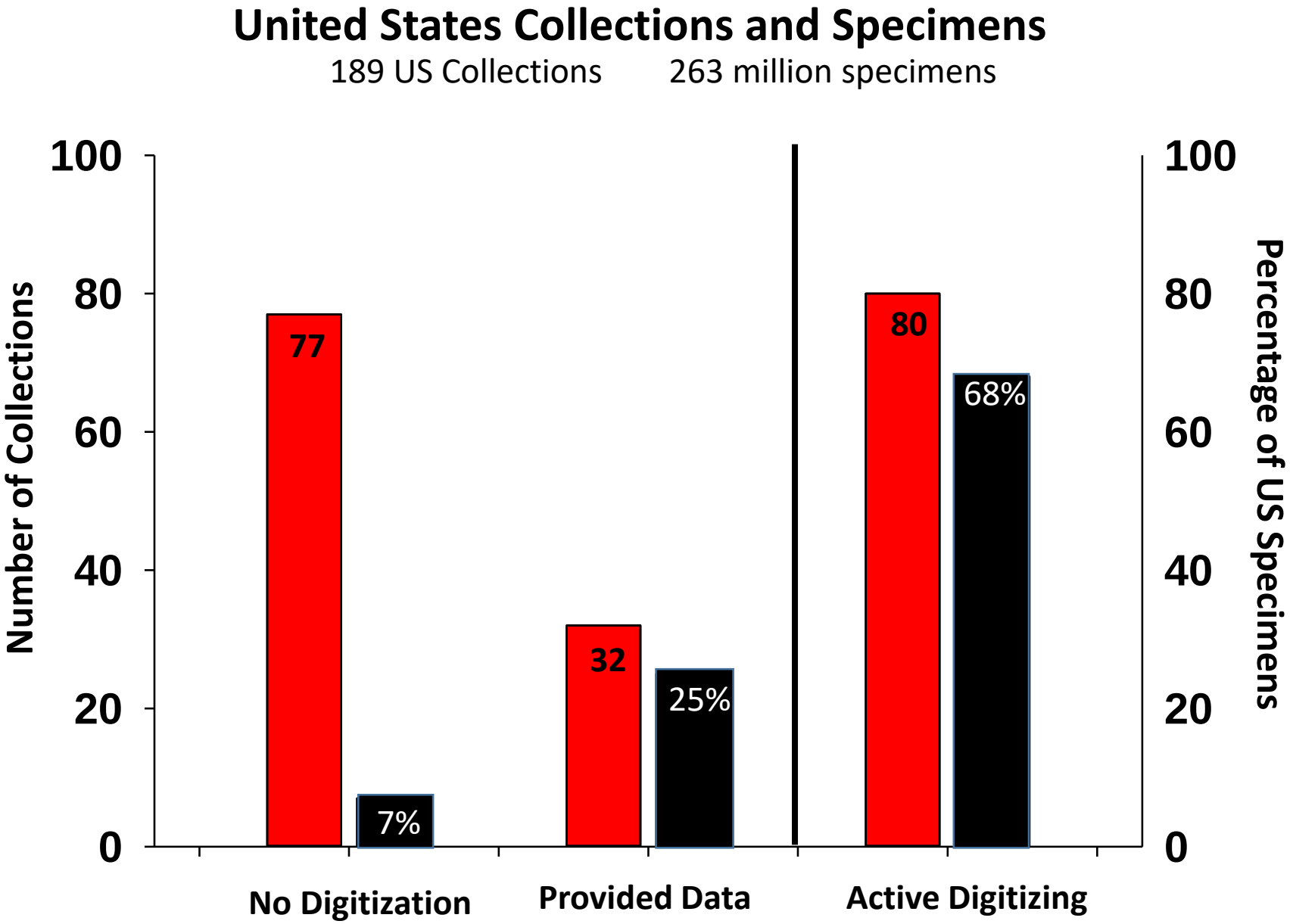
3. 6% of North American specimens digitized

4. How do we continue this beyond ADBC?



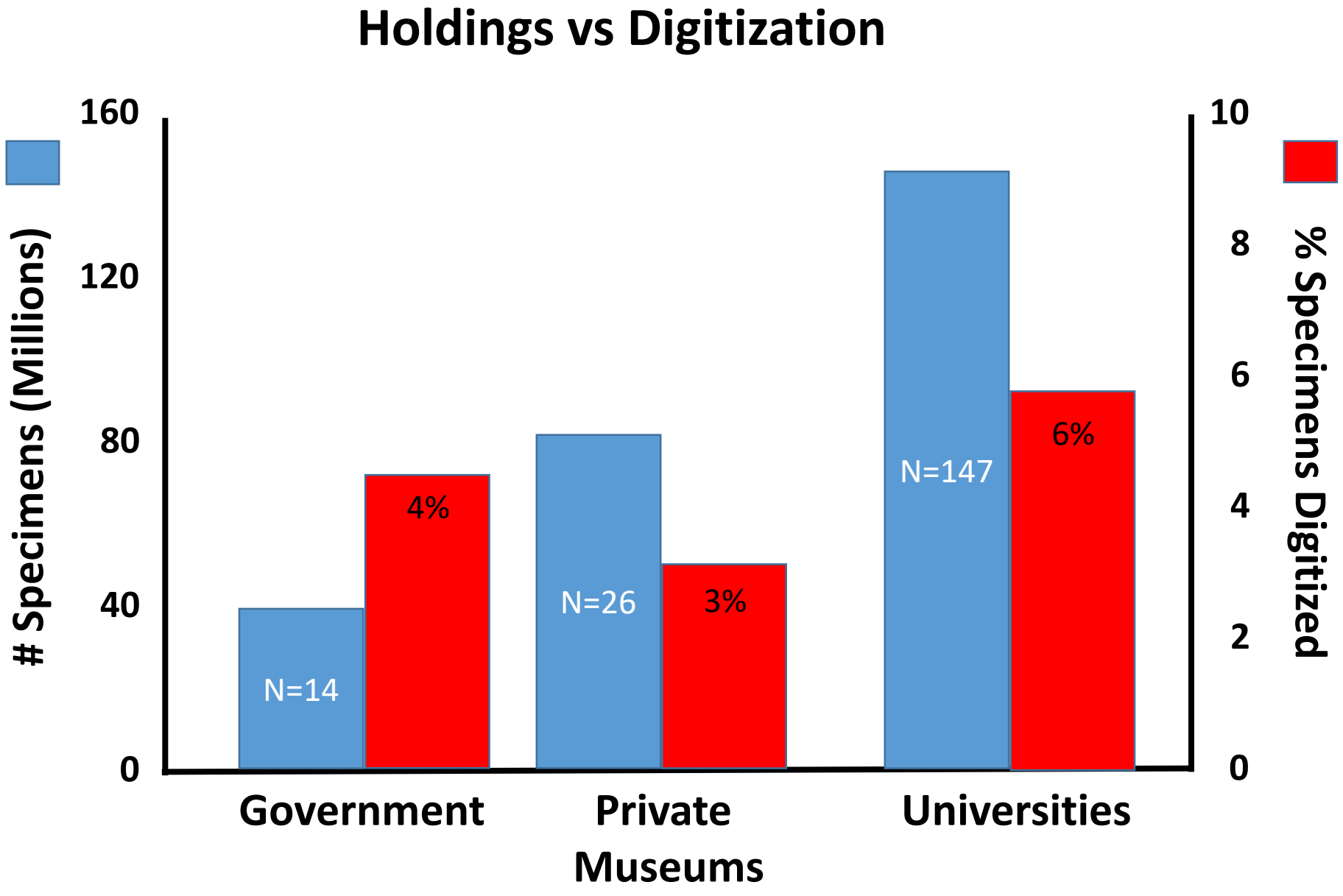
Conclusions

1. **80** collections active digitizing
2. **32** collections provided data
3. **77** collections not digitizing



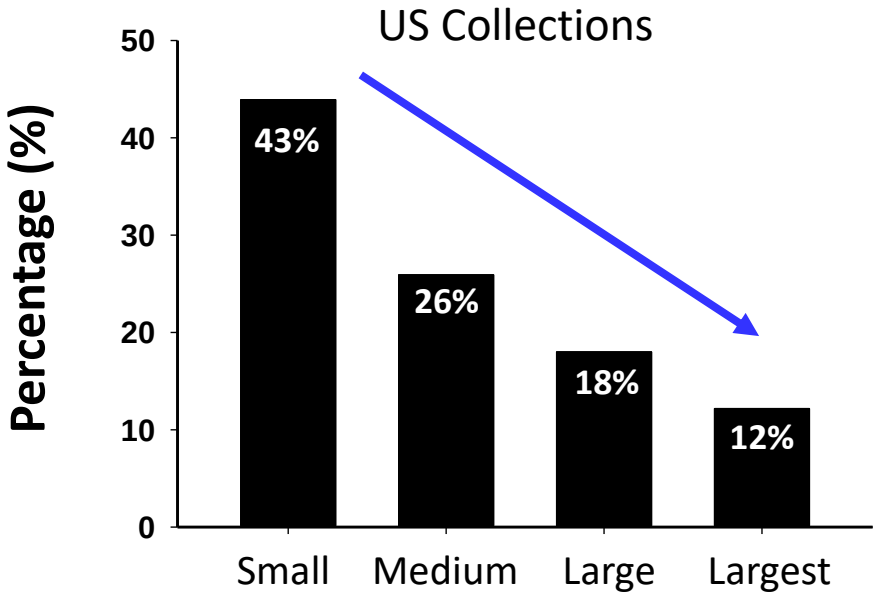
Conclusions

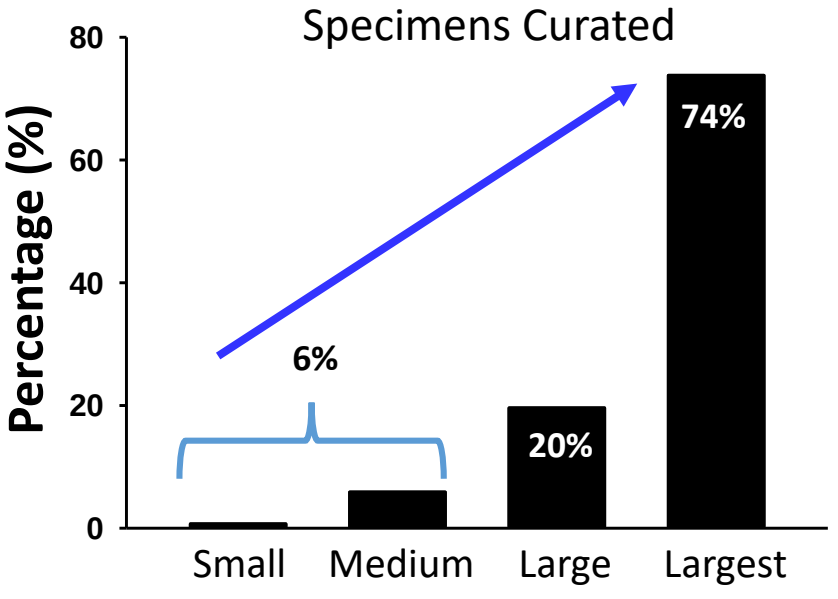
1. **80** collections active digitizing
2. **32** collections provided data
3. **77** collections not digitizing
4. Most specimens are held in collections actively digitizing!

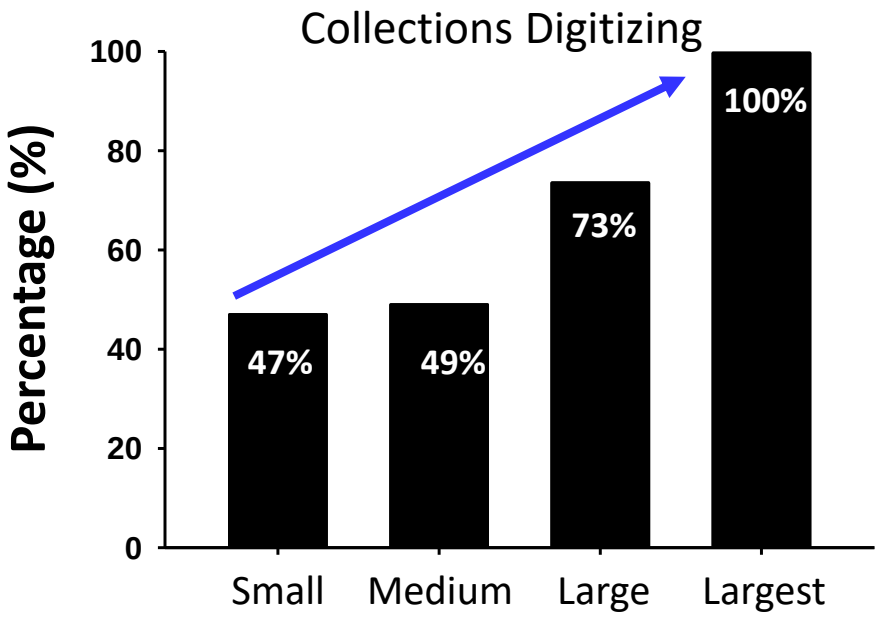


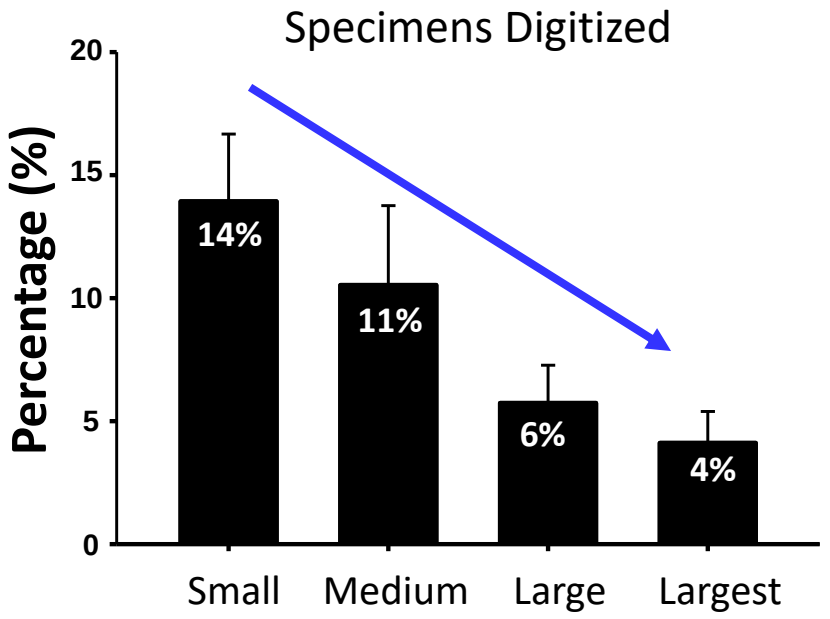
Conclusions

1. Universities house the most specimens and digitize the most.









Conclusion

1. Need to increase productivity in larger collections.

Collection Size Classes

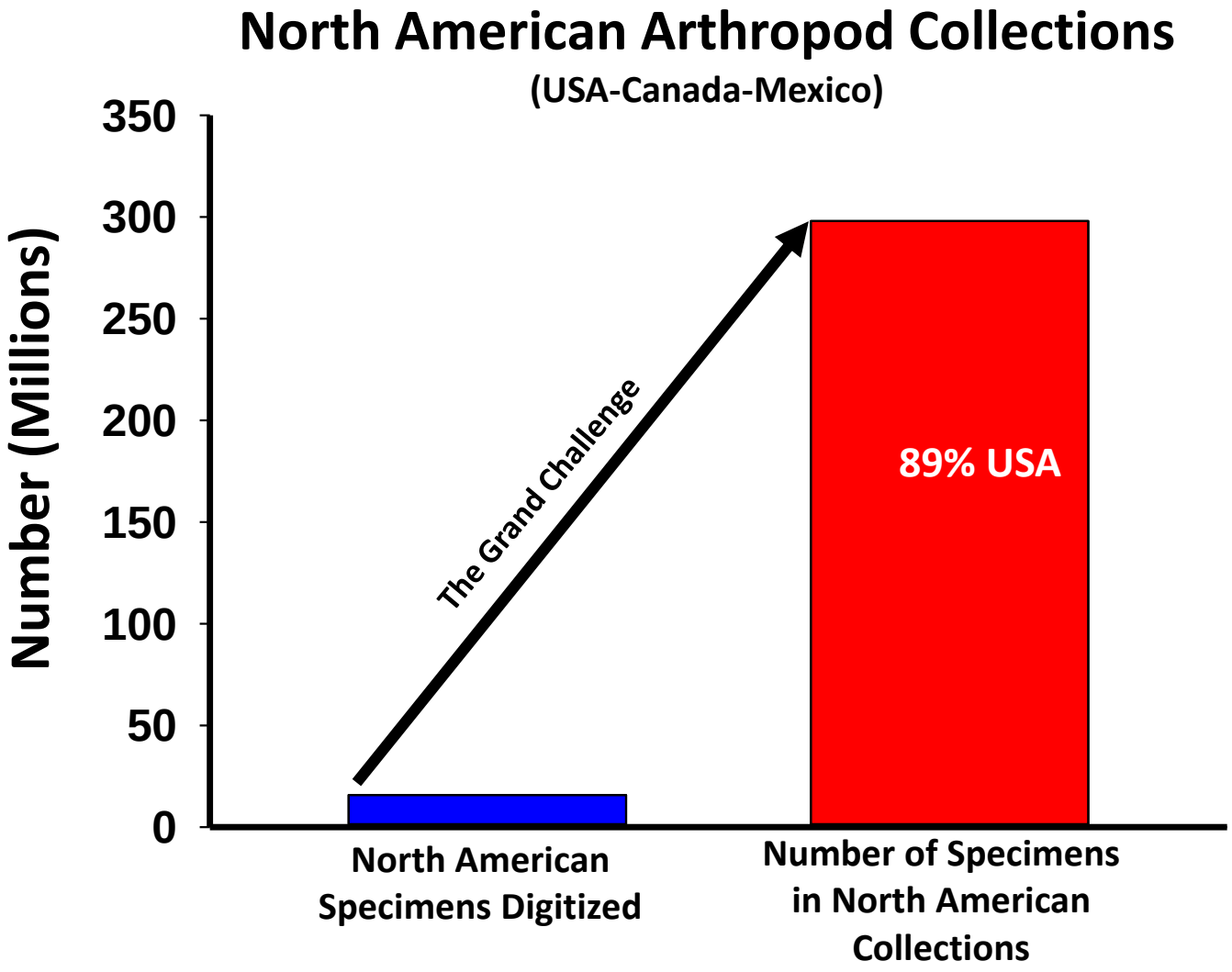
Largest: 3-35 million specimens

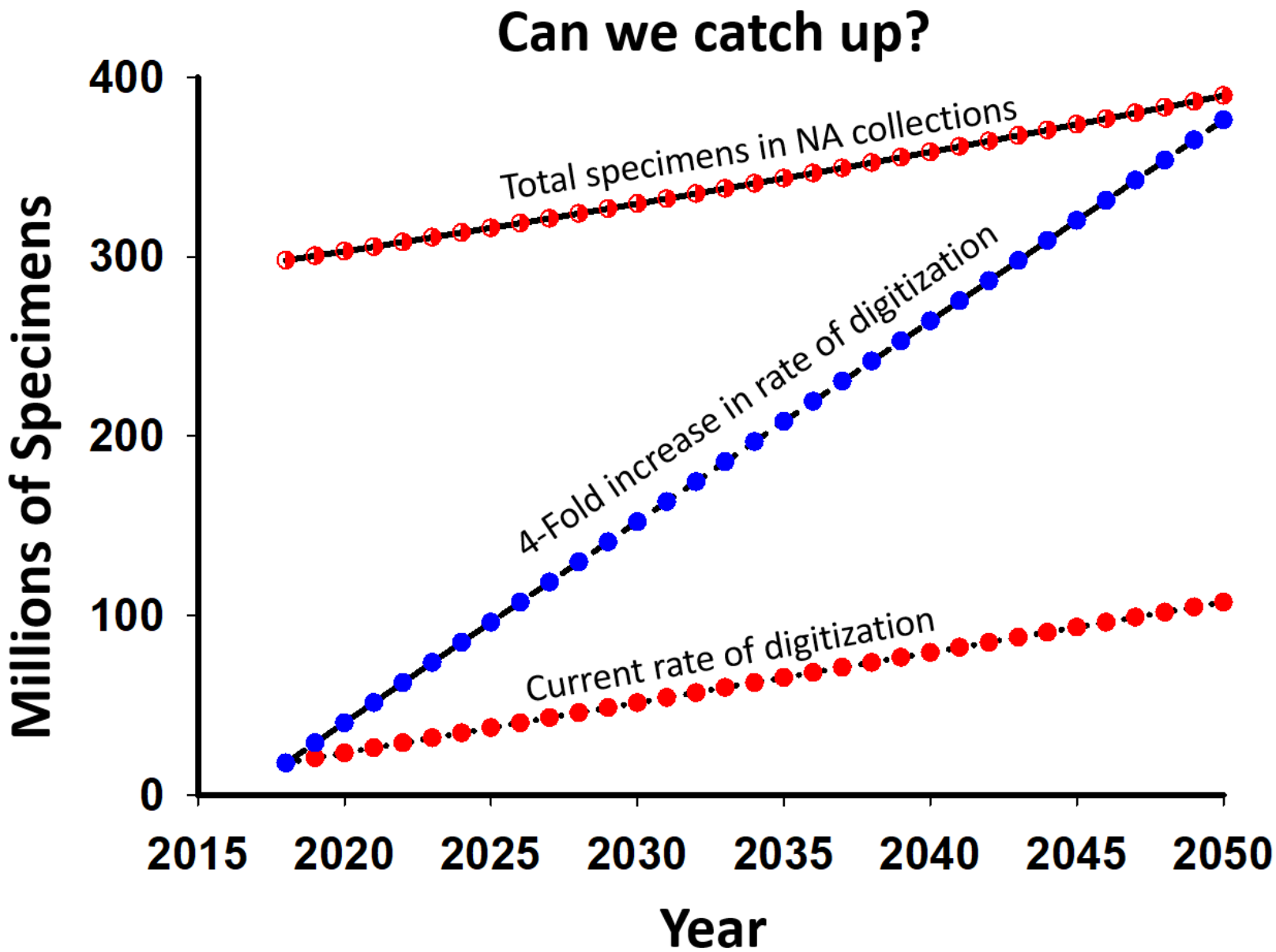
Large: 1-3 million specimens

Medium: 100k-1 million specimens

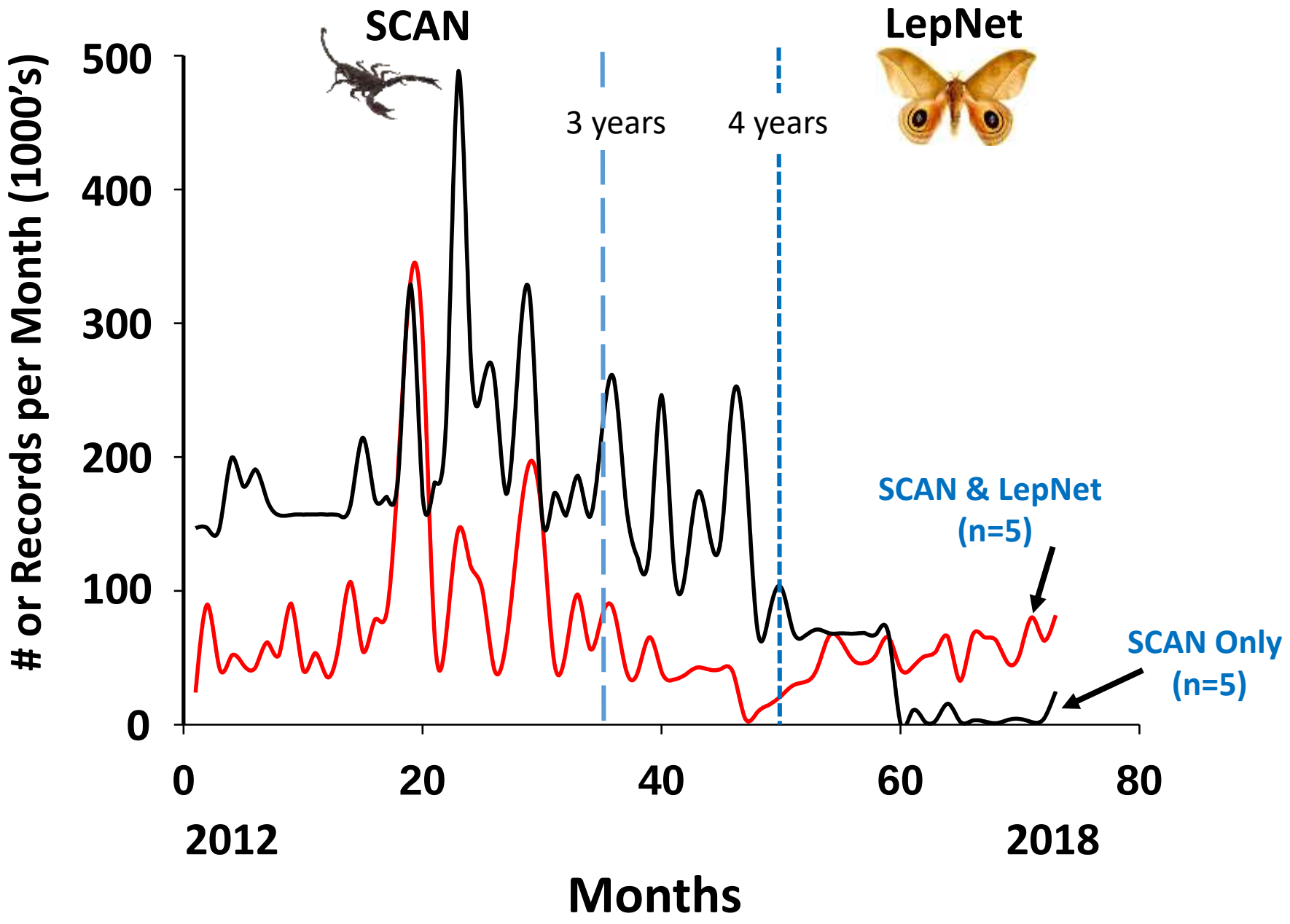
Small: Less than 100k specimens

Can we catch up by 2050?




Can we catch up?
Millions of Specimens
400
300
200
100
0
2015
2020
2025
2030
2035
2040
2045
2050
Year
Total specimens in NA collections
4-Fold increase in rate of digitization
Current rate of digitization

SUSTAINABILITY: 10 Original SCAN Museums



1. SCAN & LepNet are successful key initiatives : But sustainable digitization is still a ?

2. Understanding where we need to improve now will be key to “sustainable digitization”