Accelerating the Discovery of Biocollections Data

Deborah Paul, GBIF Task Force Member, iDigBio Digitization and Workforce Training
Siro Masinde, Programme Officer for Content Mobilization

Entomological Collections Network Meeting (ECN) 2015, Minneapolis, MN, 14-15 NOVEMBER 2015

TASK FORCE OBJECTIVES, 2015 - 2016

1. Document best practices in data mobilization

2. Document successful business models for resource mobilization

3. Consult with other capacity building and content mobilization initiatives. e.g., iDigBio, ALA

4. Guidance in training and outreach materials

GBIF's TASK FORCE MEMBERS

1. Leonard Krishtalka, USA, Chair
2. Barbara Thiers, USA
3. Deborah Paul, USA
4. Eduardo Dalcin, Brazil
5. Masanori Nakae, Japan
6. Ian Owens, UK
7. Jean Ganglo, Benin
8. Marc Pignal, France

● Siro Masinde, GBIF's task force Coordinator

● Shari Ellis, GBIF task force Consultant & iDigBio External Evaluator

Photo by S. Masinde: GBIF task force meeting, 3 Nov 2015.

ACCELERATING DISCOVERY OF BIOCOLLECTIONS DATA

Metadata approach

- c. 90% of world's c. 3 billion specimens of NHCs are not digitized

- Metadata is critical to advancing digitization efforts, priorities and opportunities, and enables the community of users to search and find relevant collections

- **Opportunity to participate whether institution is digitizing or not**

SURVEY ON NATURAL HISTORY COLLECTIONS

Purpose is to determine:

- Digital readiness of the world's biocollections and their institutions

- Benefits to a collection / institution that digitization engenders

- Impediments to collection data digitization

- Make appropriate recommendations

Survey

https://ufl.qualtrics.com/SE/?SID=SV_byg2mgbtYdeF3bT

Please participate in survey, it takes 10 - 30 minutes

INSTITUTION, COLLECTION AND ROLE

Q1. Which of the following best describes your institution?

University-based

☐ Part of a primarily privately-funded university, e.g., Harvard's Museum of Comparative Zoology

☐ Part of a primarily public-funded university, e.g., University of Florida Natural History Museum; Zoological Museum, Univ. Copenhagen, Museu Nacional - UFRJ - Brazil, Bolus Herbarium (University of Cape Town)

Not university-based

☐ Primarily funded by national, state or city government, e.g., Smithsonian's National Museum of Natural History, Museum National d'Histoire Naturelle (Paris), Virginia Museum of Natural History, IVPP (People's Republic of China), Rio de Janeiro Botanical Garden, National Museums of Kenya

☐ Primarily privately-funded institution/museum, e.g., Carnegie Museum of Natural History, California Academy of Sciences (San Francisco), New York Botanical Garden

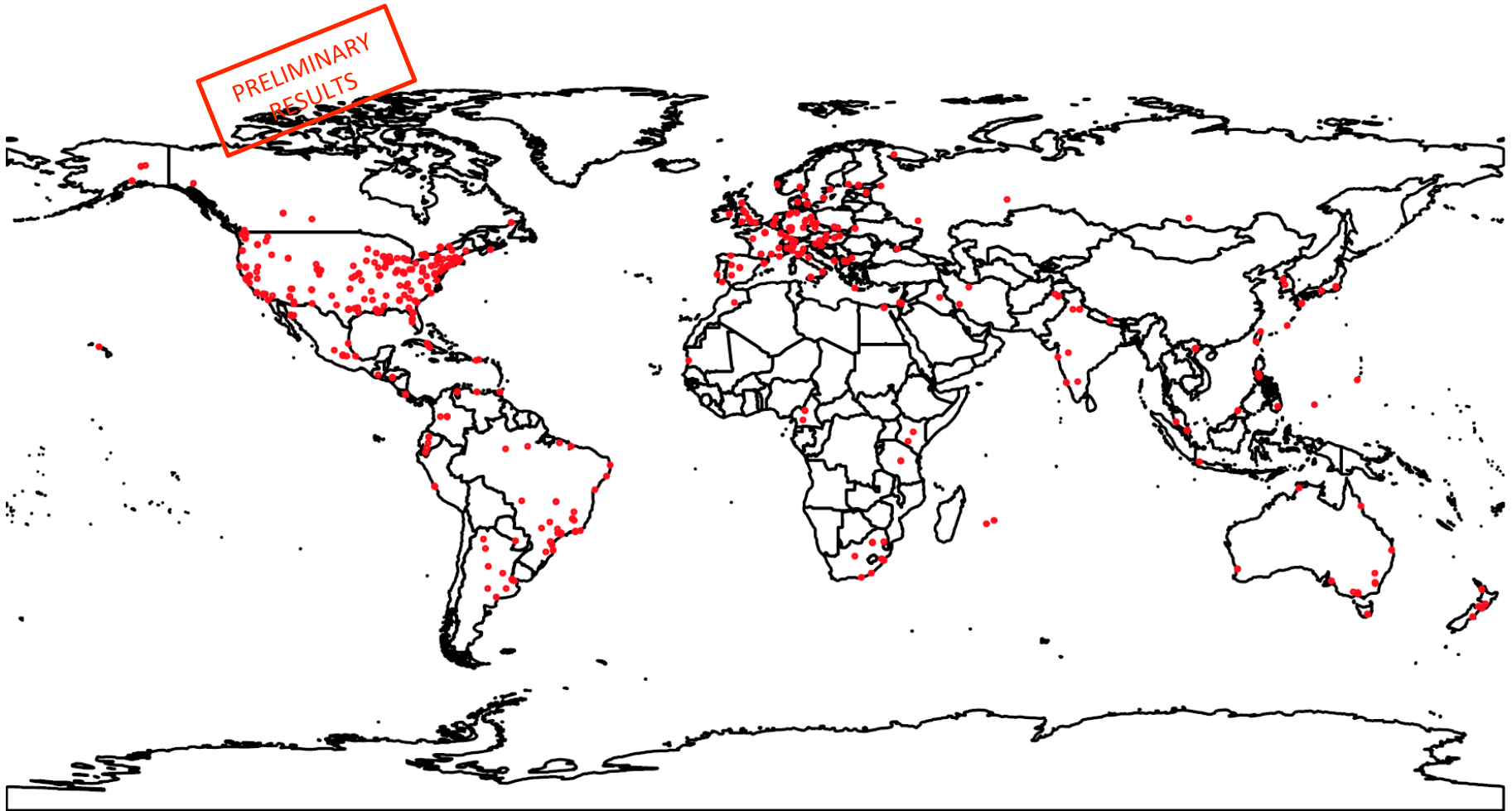☐ Individual Private collection, e.g., American Entomological Institute

Q2. What is/are the primary role(s) you play at your institution? Please check all that apply.

☐ Director/President/CEO

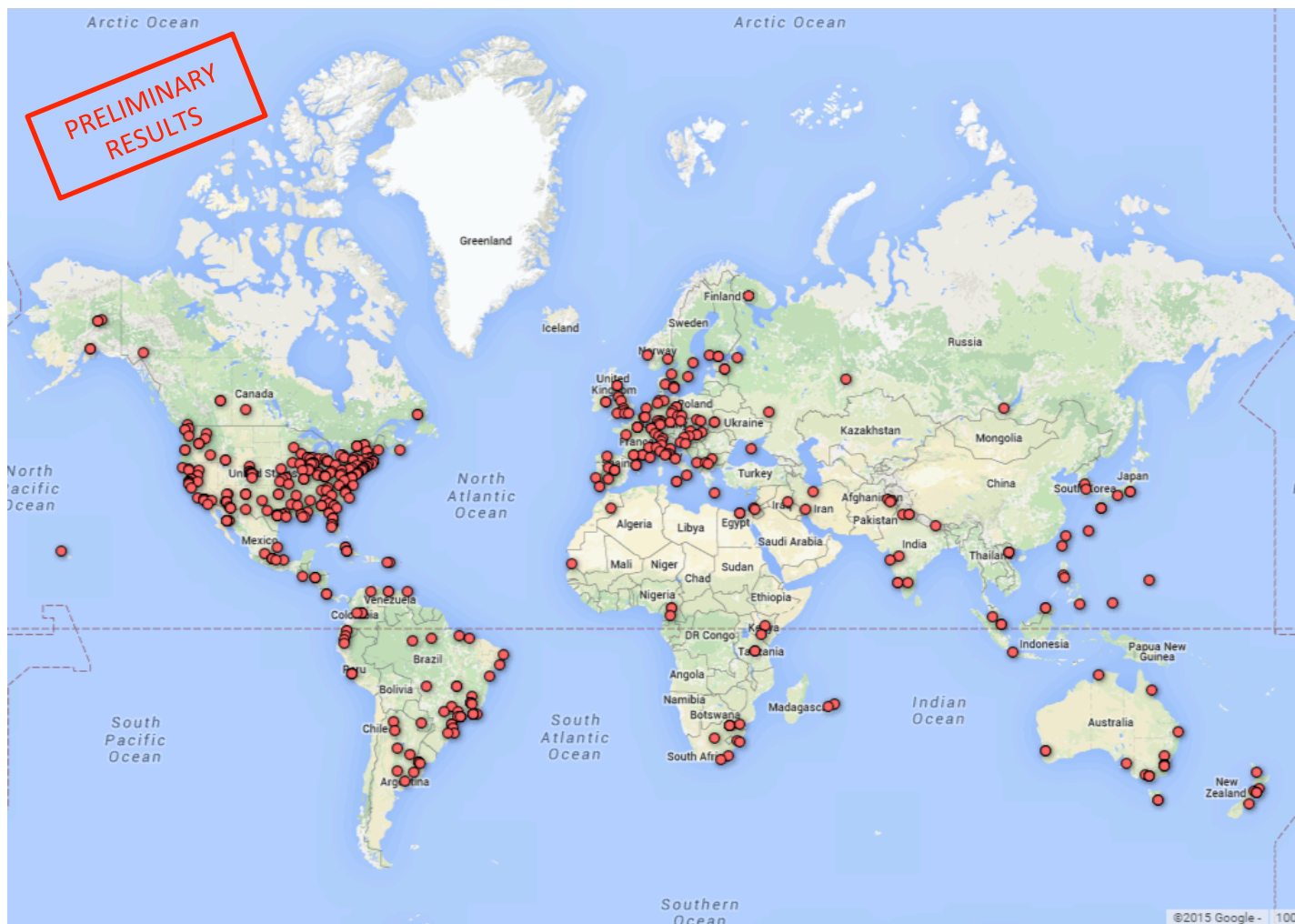☐ Head of Research and Collections

☐ Curator

GBIF

WORLDWIDE SURVEY *in-progress*

PRELIMINARY RESULTS

GBIF

PRELIMINARY SURVEY RESULTS
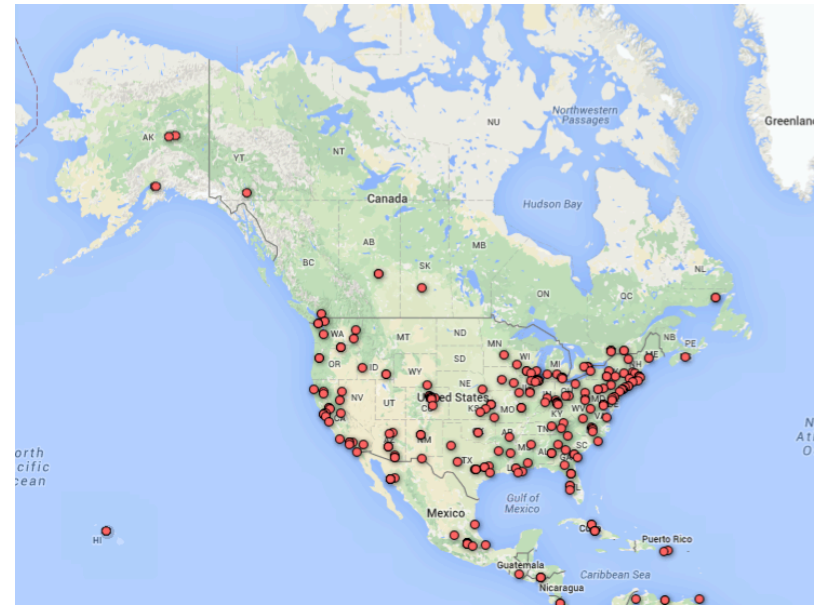

- > c. 600 responses
- **USA among best response rate**
- *Survey closes Nov 20, 2015 -- responses still coming in*

NORTH AMERICA RESPONSES
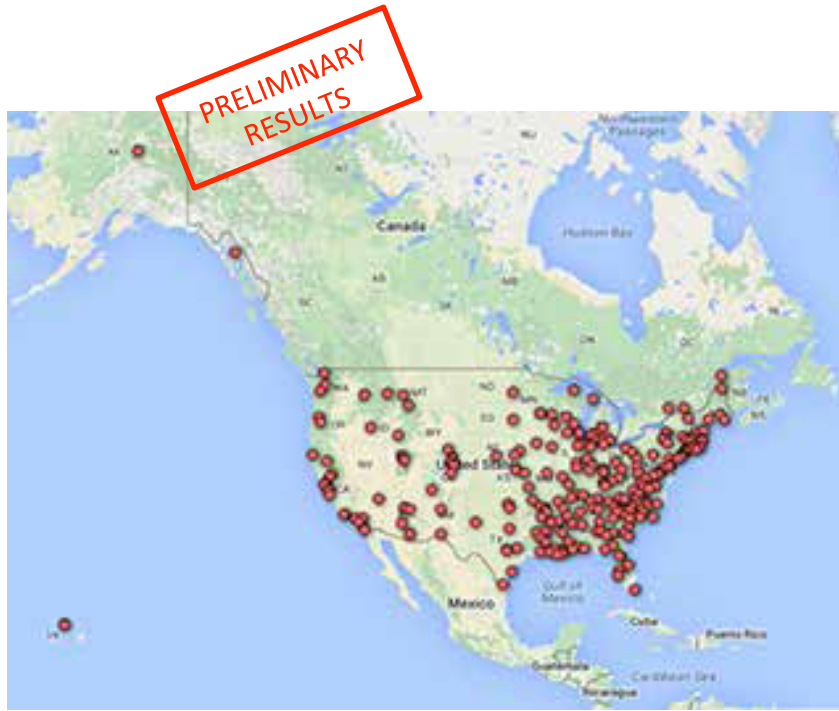
ADBC PROJECTS

Image taken from iDigBio Summit V program

SURVEY RESPONSES

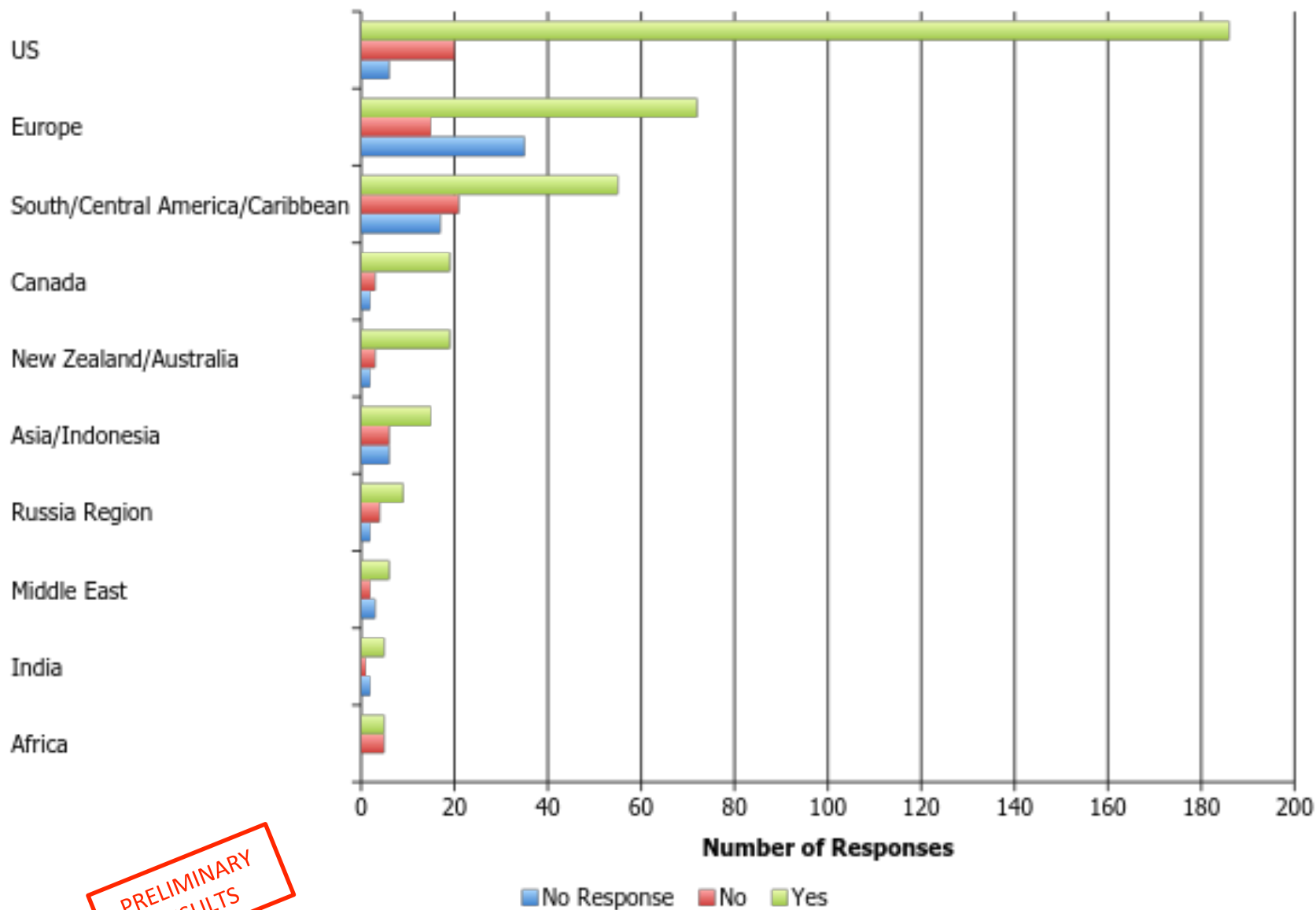## Number and Percent of Respondents from Each Collection Type

Vascular Plants	337	63%	Ichthyology	63	12%
Bryophytes	187	35%	Malacology	58	11%
Fungi & lichens	182	34%	Invertebrates, terrestrial	55	10%
Algae	143	27%	Marine Invertebrates	54	10%
Arthropods	116	22%	Paleobotany & Palynology	42	8%
Mammalogy	67	13%	Vertebrate Paleontology	41	8%
Ornithology	68	13%	Other	48	9%
Herpetology	63	12%	Invertebrate Paleontology	40	7%

PRELIMINARY RESULTS

GBIF

Percent of Collection Being Databased & Shared/Published				
Collection	#	Databased	#	Shared
Vascular Plants	n = 239	46%	n = 231	37%
Bryophytes	n = 141	37%	n = 132	28%
Fungi & lichens	n = 141	47%	n = 123	38%
Algae	n = 112	36%	n = 101	37%
Arthropods	n = 87	31%	n = 77	28%
Mammalogy	n = 58	67%	n = 27	47%
Ornithology	n = 54	65%	n = 51	30%
Herpetology	n = 52	65%	n = 49	34%

PRELIMINARY RESULTS

GBIF

Digitization by Region

Horizontal bar chart titled "Digitization by Region" showing Number of Responses by region for three categories: No Response (blue), No (red), Yes (green).

Regions (top to bottom): US, Europe, South/Central America/Caribbean, Canada, New Zealand/Australia, Asia/Indonesia, Russia Region, Middle East, India, Africa.

X-axis: Number of Responses (0 to 200).

Legend: ■ No Response ■ No ■ Yes

PRELIMINARY RESULTS

GBIF

MOTIVATION FOR DIGITIZING

Percent of Respondents
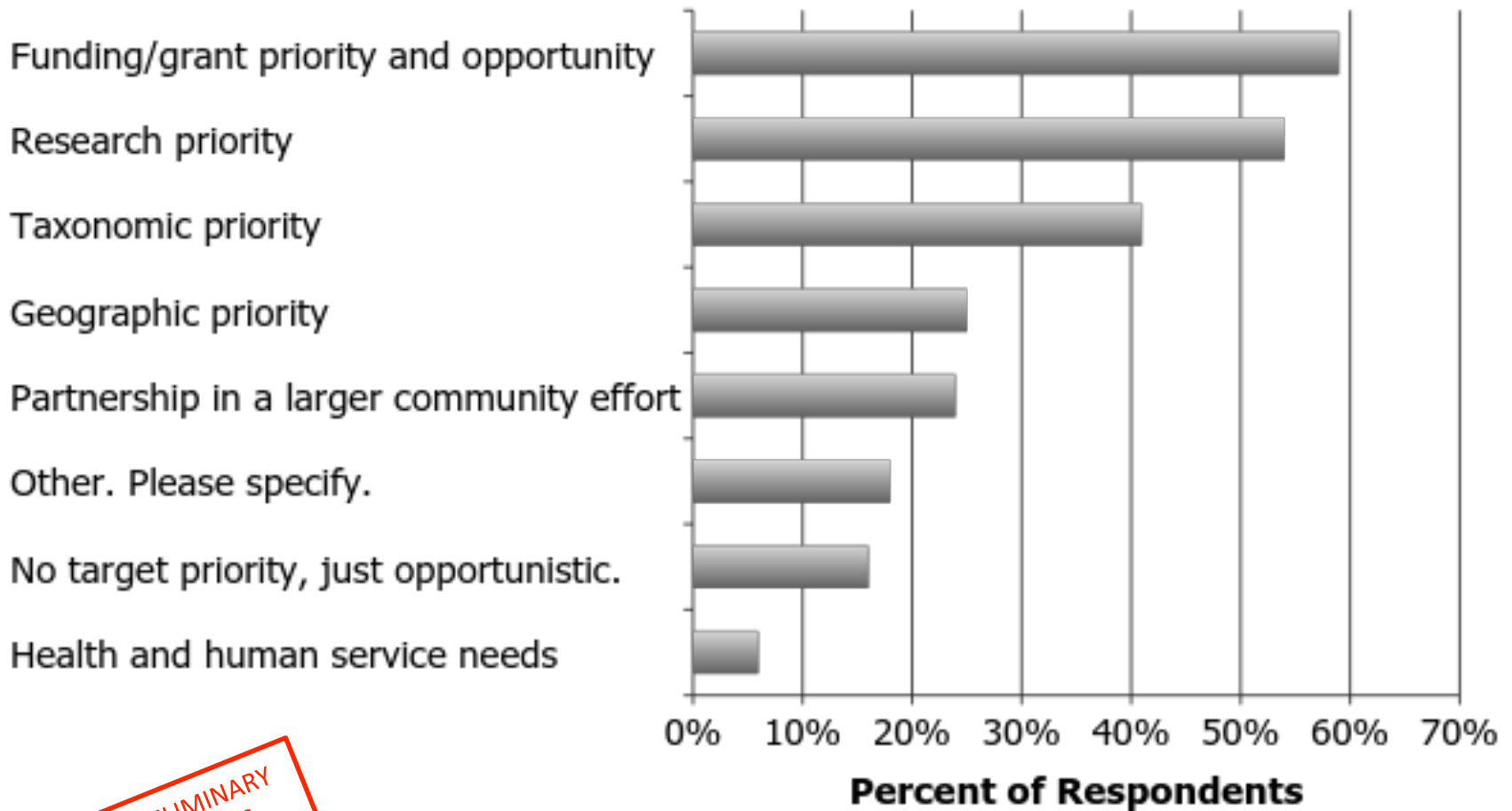
PRELIMINARY RESULTS

IMPORTANCE OF METADATA

PRELIMINARY RESULTS

Number of Respondents

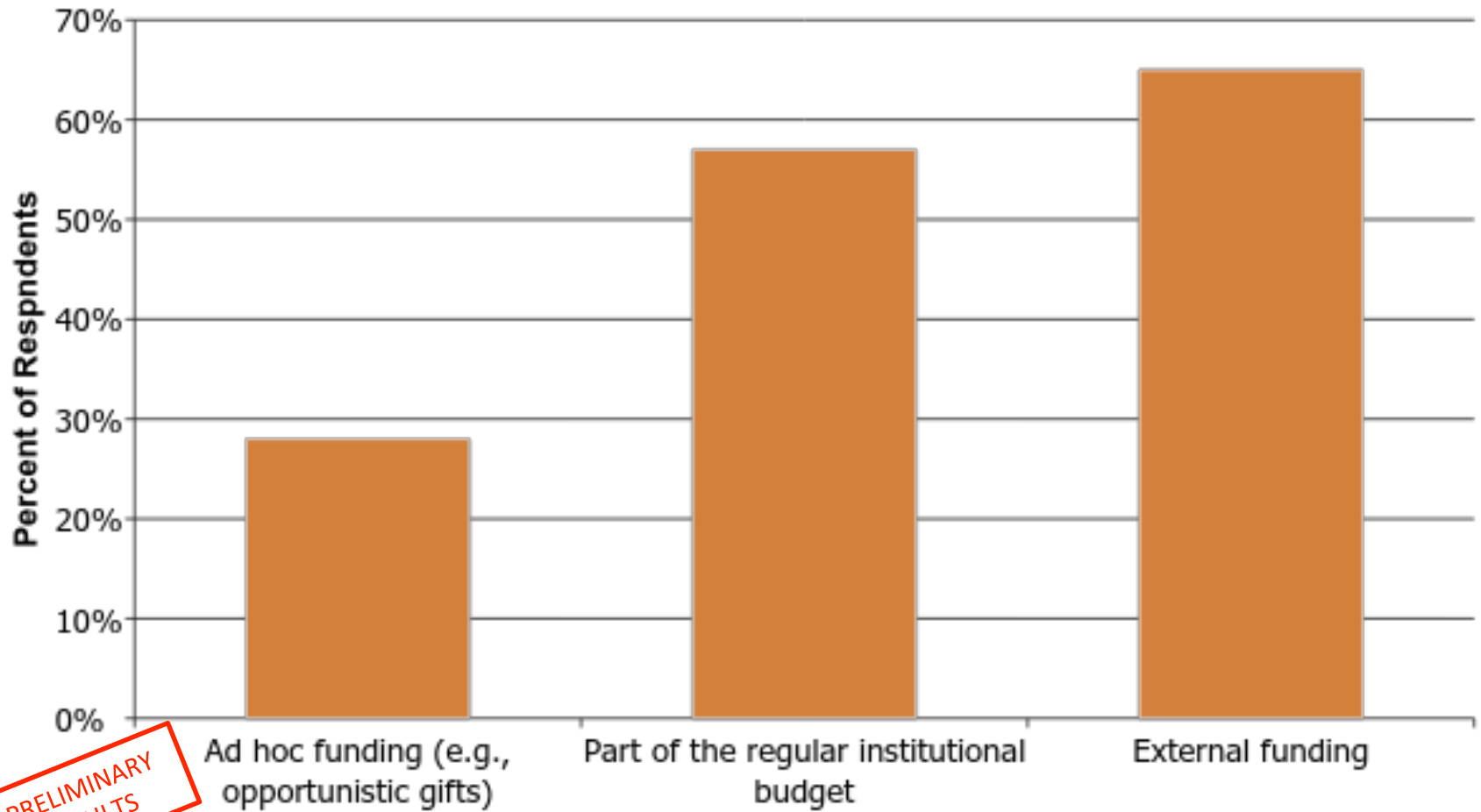■ I think it is critical we provide this information.   □ We are currently providing this information.

GBIF
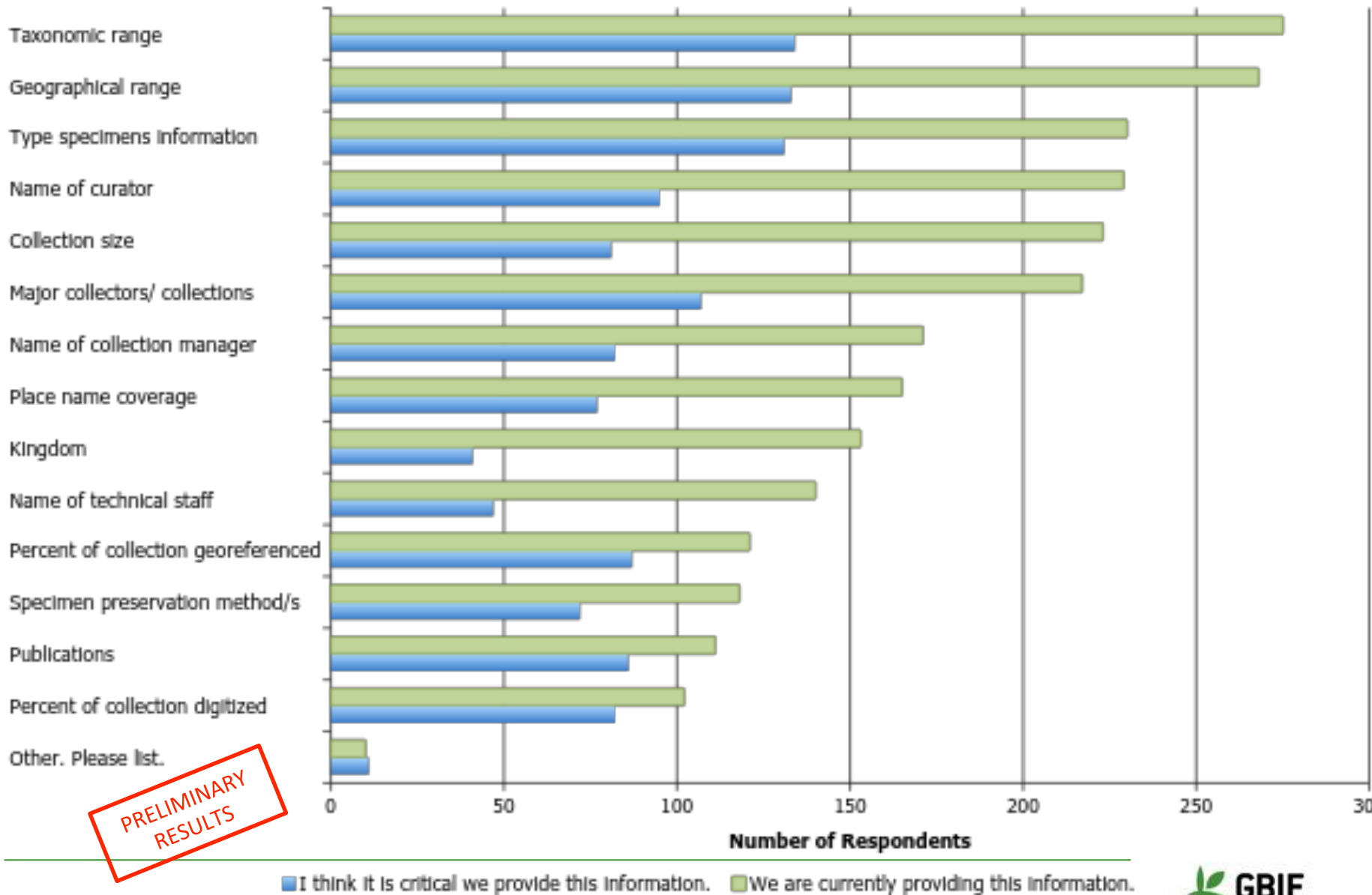
WHERE METADATA IS SHARED – 60 LOCATIONS
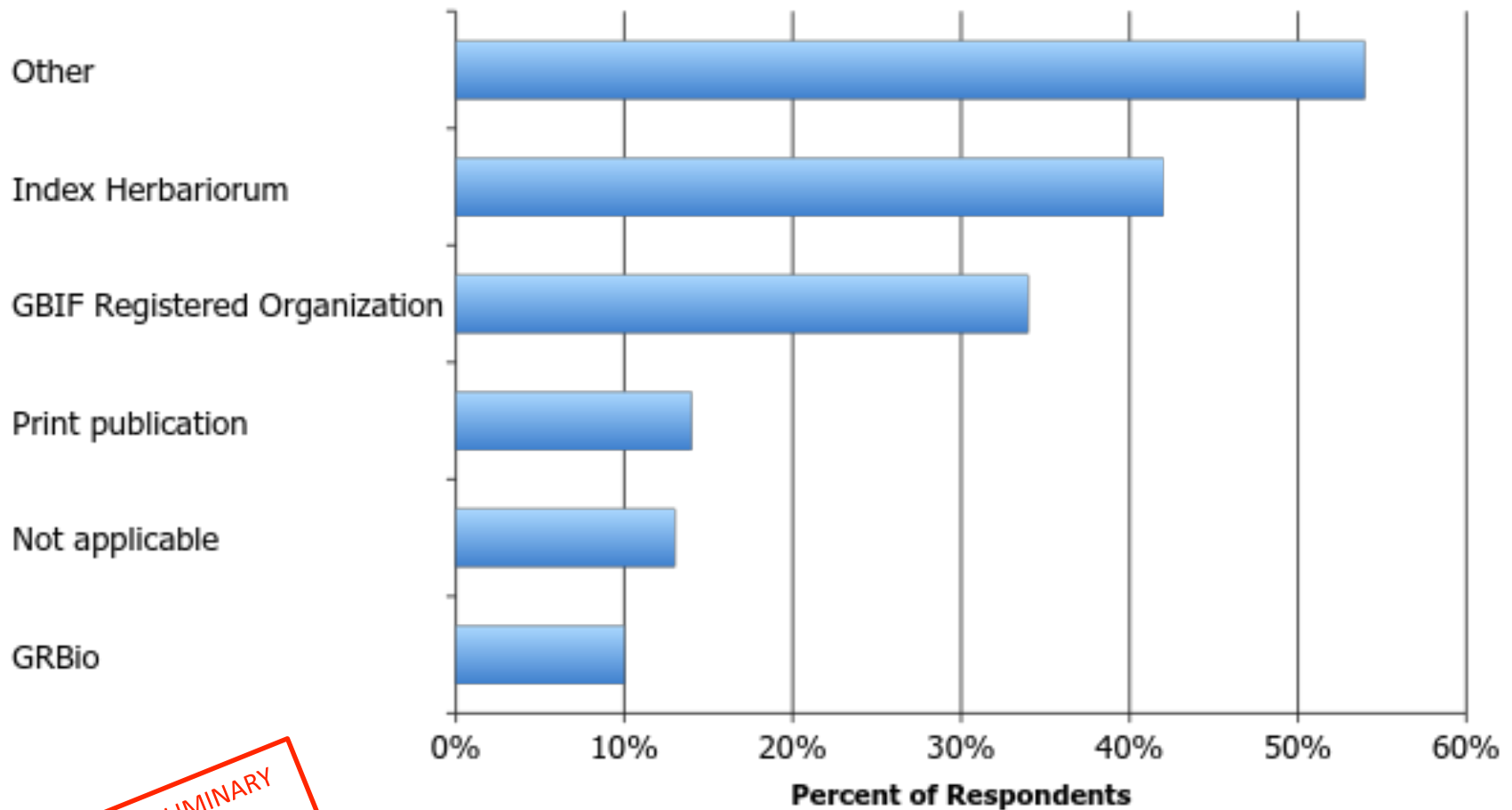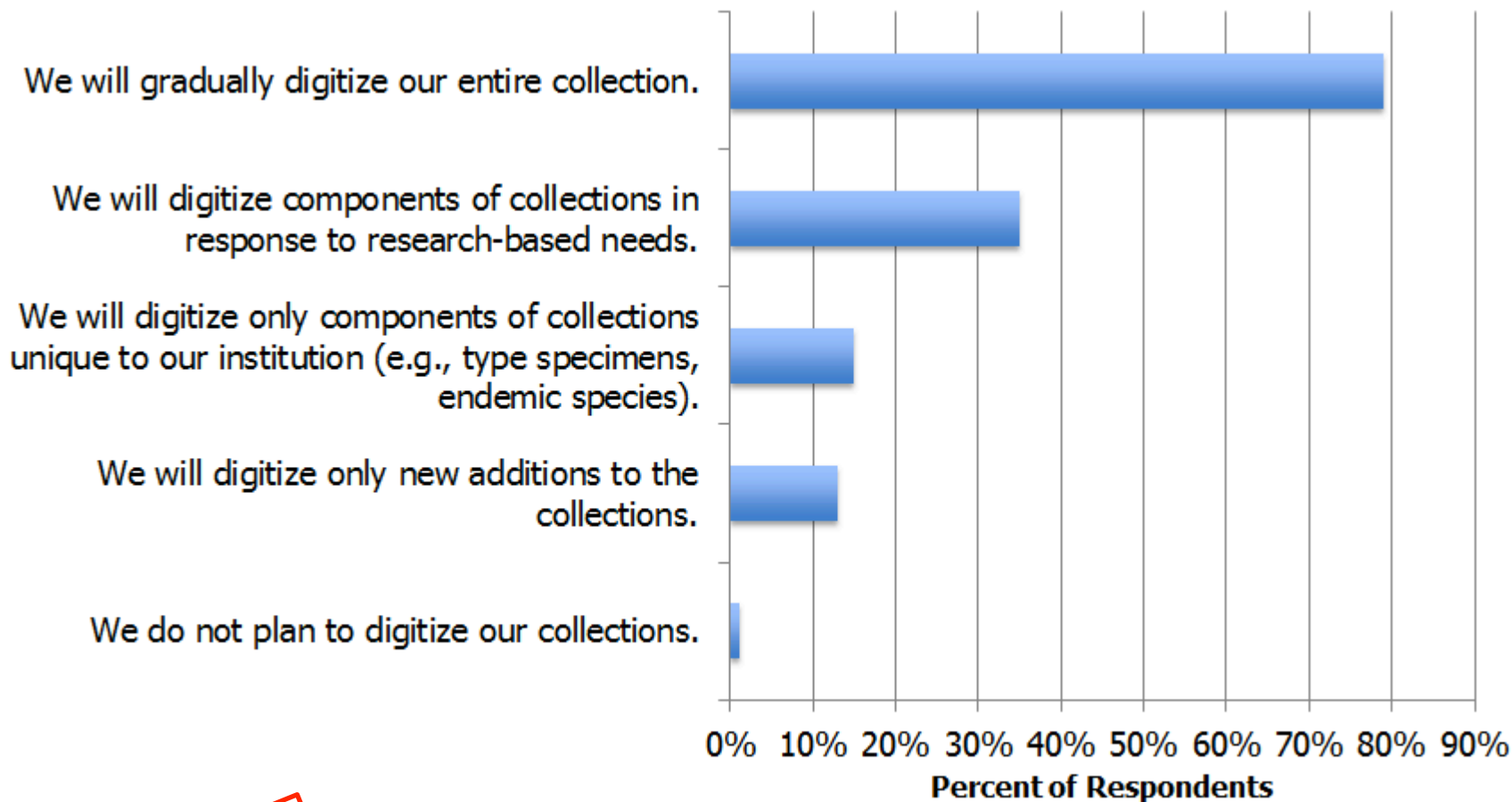

PRELIMINARY RESULTS

SOURCES OF FUNDING

PRELIMINARY RESULTS

GBIF

LONG TERM PLANS

TOP 6 REASONS FOR NOT DIGITIZING

1. *Funding*/resources not available
2. Lack of *time* among personnel
3. Limited expertise among personnel
4. Not a priority of the individual in charge of the collection
5. Insufficient information on digitization process and protocols
6. Size of task is overwhelming

PRELIMINARY RESULTS

GBIF

NEXT STEPS

Task Force recommending that the

DATA → KNOWLEDGE → APPLICATION

value chain framework be used to mobilize resources and prioritize digitization

GBIF

THANK YOU | GRACIAS

Deborah Paul @idbdeb
dpaul@fsu.edu

Siro Masinde
smasinde@gbif.org

Twitter	@gbif
Facebook	gbifnews
Research uses	gbif.org/mendeley
LinkedIn Group	linkedin.com/grp/home?gid=55171
Github	github.com/gbif