iDigBio
Integrated Digitized Biocollections

Integrating Institutional Archives with Disciplinary Web Repositories

January 23-24, 2020

MorphoSource
3D Data Repository
DUKE UNIVERSITY

UF UNIVERSITY of FLORIDA
oVert

Acknowledgments

Funding, travel logistics & wiki page

Jillian Goodwin, Allie Blackwell, Gil Nelson (UF, iDigBio)

Venue & day of workshop logistics

Kathy Peterson, Mackenzie Shepard (Duke, iiD & Ev Anth)

Content & agenda

Doug Boyer & Julie Winchester (Duke, MorphoSource) / David Blackburn (UF, oVert lead PI) / Michelle Koo (Berkeley, CalPhotos) / Mike Webster (Cornell, Macaulay Sound Library)

Reception & Lemur/Fossil Center tours

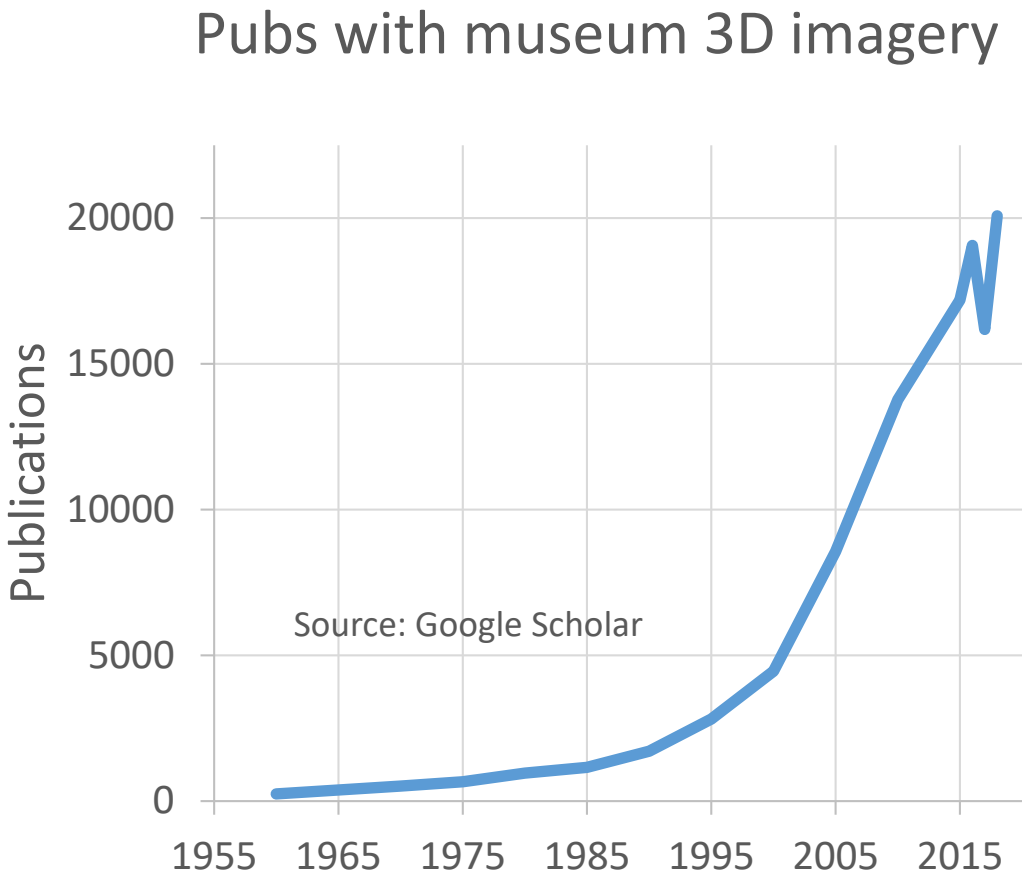
Dept. of Evolutionary Anthropology / Erin Ehmke, Greg Dye, Matt Borths (Center directors)

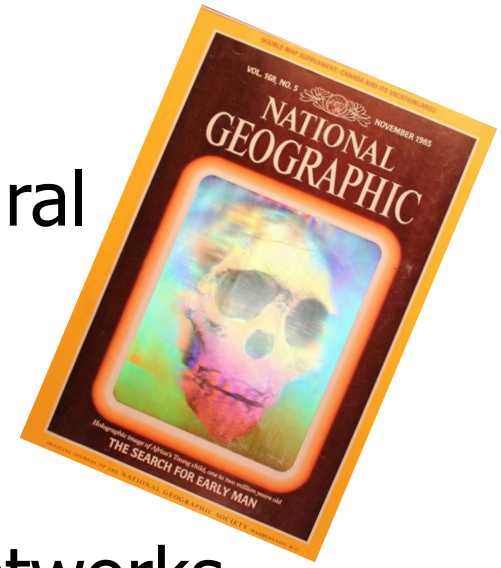
Things to note

- Restrooms right outside
- Students & faculty working outside (be considerate)
- Coffee snacks throughout day + lunch
- Reception at 5-ish
- Code of conduct that emphasizes inclusiveness and respect (in your packet)
- We will try hard to stay on schedule, but morning may creep into afternoon
- Please fillout video/presentation consent forms
- After lunch, we will be reconvening here briefly before splitting into our break-out groups

Motivation/Vision

Help realize the potential of digitized specimen representations





- Deepen insight into natural history & environmental phenomena
- Increase value to public
- Maximize effects with networks of human interpretable & machine readable data

Workshop Goals

(1) Survey current needs, workflows and trajectories of the represented informatic community

(2) Identify common ground among institutions

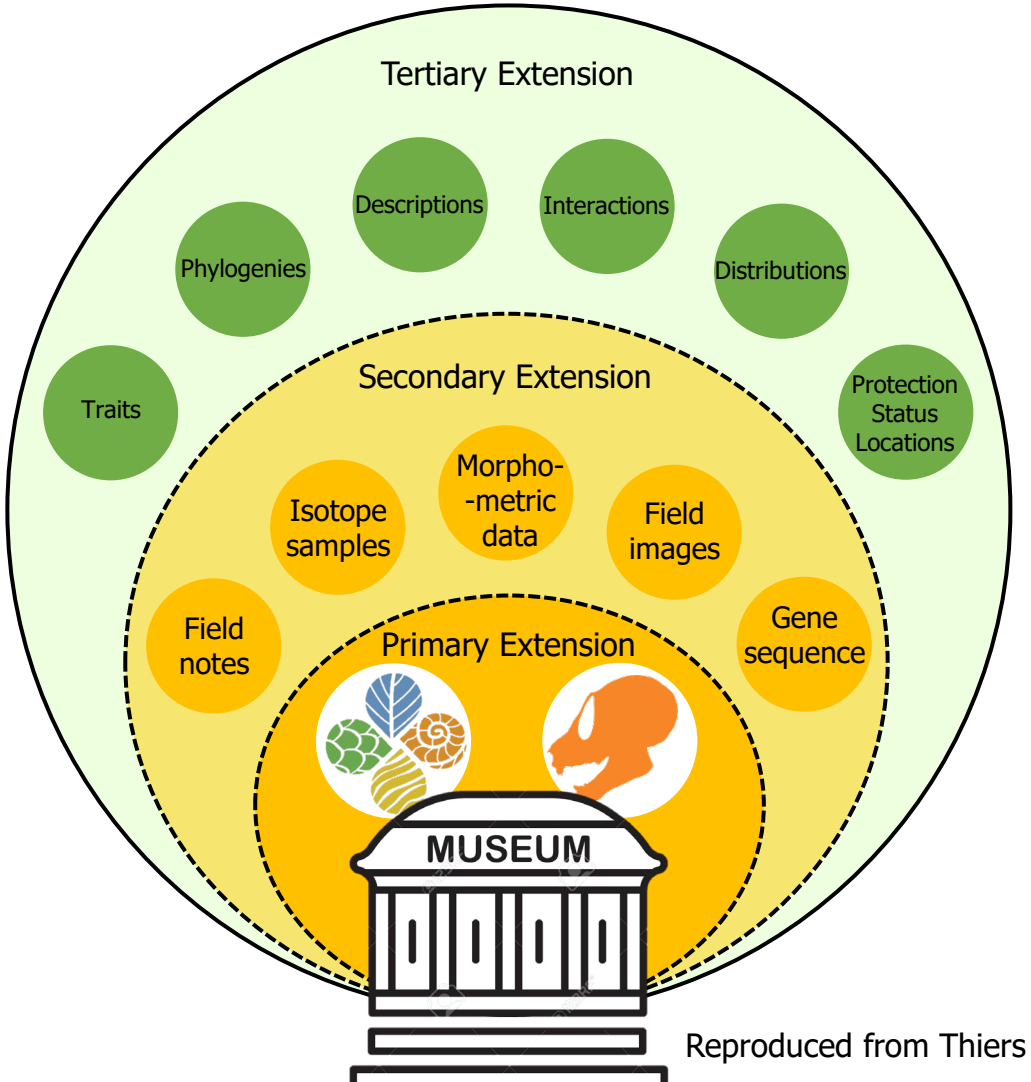
(3) Explore potential for unifying approaches where appropriate

(4) Assess the role that domain-specialized repositories can/should serve here

(5) Articulate an overarching plan representing a consolidated set of workflows that will accommodate the diverse requirements and limitations of represented institutions.

Participant scope

Core/Foundation of Extended Specimen Network

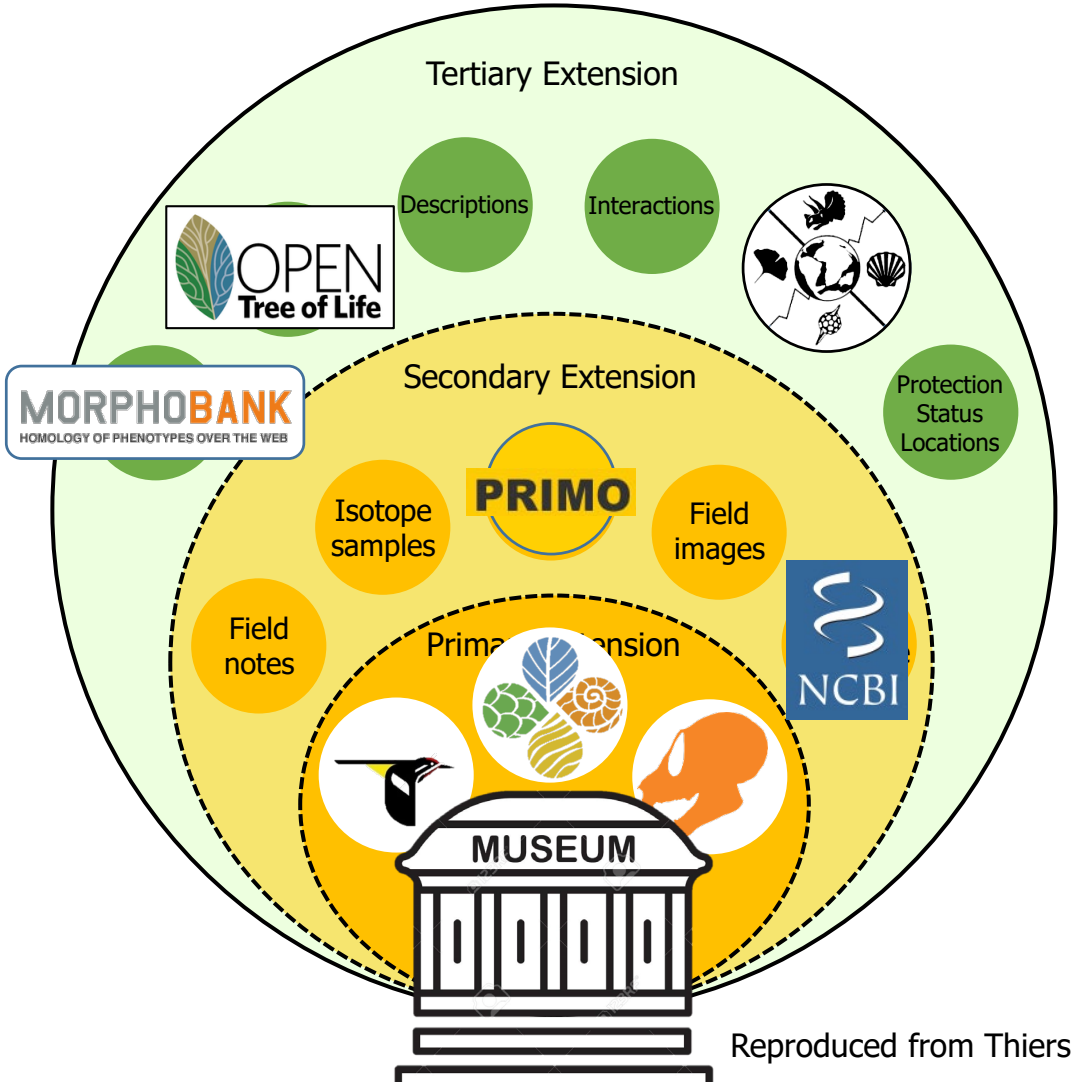


Reproduced from Thiers et al. (2018)

Participant scope

Core/Foundation of Extended Specimen Network

Other resources are critical, but live at higher abstraction level with additional challenges



Reproduced from Thiers et al. (2018)

Participants

16 institutions, ~24 departments/programs

- American Mus. of Natural History
- Berkeley Mus. of Vertebrate Zoology
- Berkeley Mus. of Paleontology (UCMP)
- CalPhotos (Berkeley)
- Cornell (Macaulay Sound Library)
- Duke Lemur Center, Fossil Div.
- Duke Library Research Data Repository
- Duke MorphoSource
- Field Mus. of Natural History
- Florida Mus. of Natural History (iDigBio)
- Harvard Mus. of Comparative Zoology
- Indiana Univ. (IUScholarworks)
- Indiana Geological Survey
- Smithsonian Mus. of Natural History
- Specify (Univ. of Kansas)
- Texas Advanced Computing Center (TACC)
- Texas A&M Univ.
- Univ. of Kansas
- Univ. of Michigan Mus. of Zoology
- Univ. of Michigan Mus. of Paleontology
- Univ. of Michigan Deep Blue Data
- Univ. of Wyoming
- Virginia Tech
- Yale Peabody Mus.

Agenda/format

Day 1 symposium

- Domain specialized repositories
- Institutional repositories & archiving
- Museum integration examples
- Ownership issues
- MorphoSource perspectives

Breakout groups

- *Session 1:* Evenly mixed groups tackle same topics
- *Session 2:* Interest groups tackle different topics

Day 2 symposium

- Aspirations for broad integration
- Examples of D.A.M.
- Sustainability for third-party resources

Breakout groups

- *Session 3:* Stakeholder groups tackle same topics
- *Session 4:* Whole group discussion about next steps

Continuing the convo

Upcoming workshops

- *March 2020:* Future of Fluid Preserved Specimen Imaging
- *June 2020:* Digital Data 2020 (4th annual), Harnessing the Data Rev.
- *September 2020:* Biodiversity Summit 2020

Why MorphoSource?

In 2000's

- Morphology increasingly digital
- No clear protocol, requirements, or resources for archiving/sharing (neither museums nor aggregators were ready)
- Poor data transparency

What was needed...

A web-accessible platform that...

1) allows researchers to submit image data on specimens
2) integrates the data with museum-provisioned/maintained records
3) incentivizes use

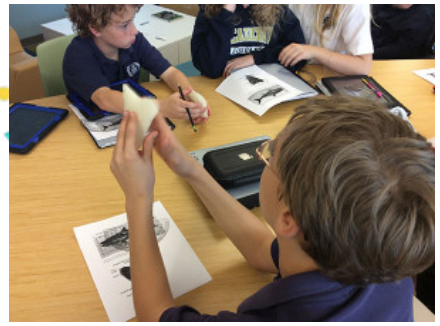
MorphoSource

Launched in 2013

1) Archive & share 3D (and 2D) specimen data supporting published research.
2) Manage ownership & access according to museum policies
3) Track use

MorphoSource

Milestones



Homo naledi

2013: launched with Duke funds

2014: began hosting outreach collections

2015: became official repo for *H. naledi*

2017: Integration with iDigBio records

2017-18: three NSF awards & additional Duke funding commitment / integration with Duke Library Digital Repo.

2018-19: began serving as repo. for multiple NSF projects

overt

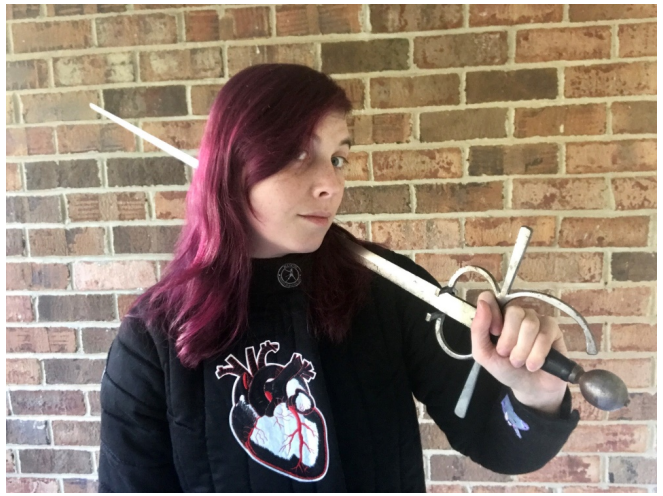
UTCT DigiMorph

MorphoSource

“the team”

Doug Boyer

Tim McGeary



Julie Winchester

Will Sexton

David C.-S.

Ed Gomes

Jocelyn Triplett

Simon Choy

Ryan Bauman

Jim Coble

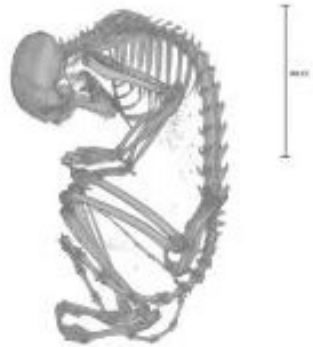
MorphoSource

Current holdings

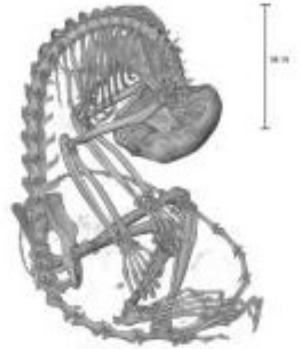
- 26,400 specimens
  - 390 museums & collections

NHMUK-pal-pv m 1345, *Adapis parisiensis*

Created 9/7/2017 at 22:52
Modified 9/7/2017 at 22:52
iDigBio



AMNH-M-183285, *Callimico goeldii*
iDigBio

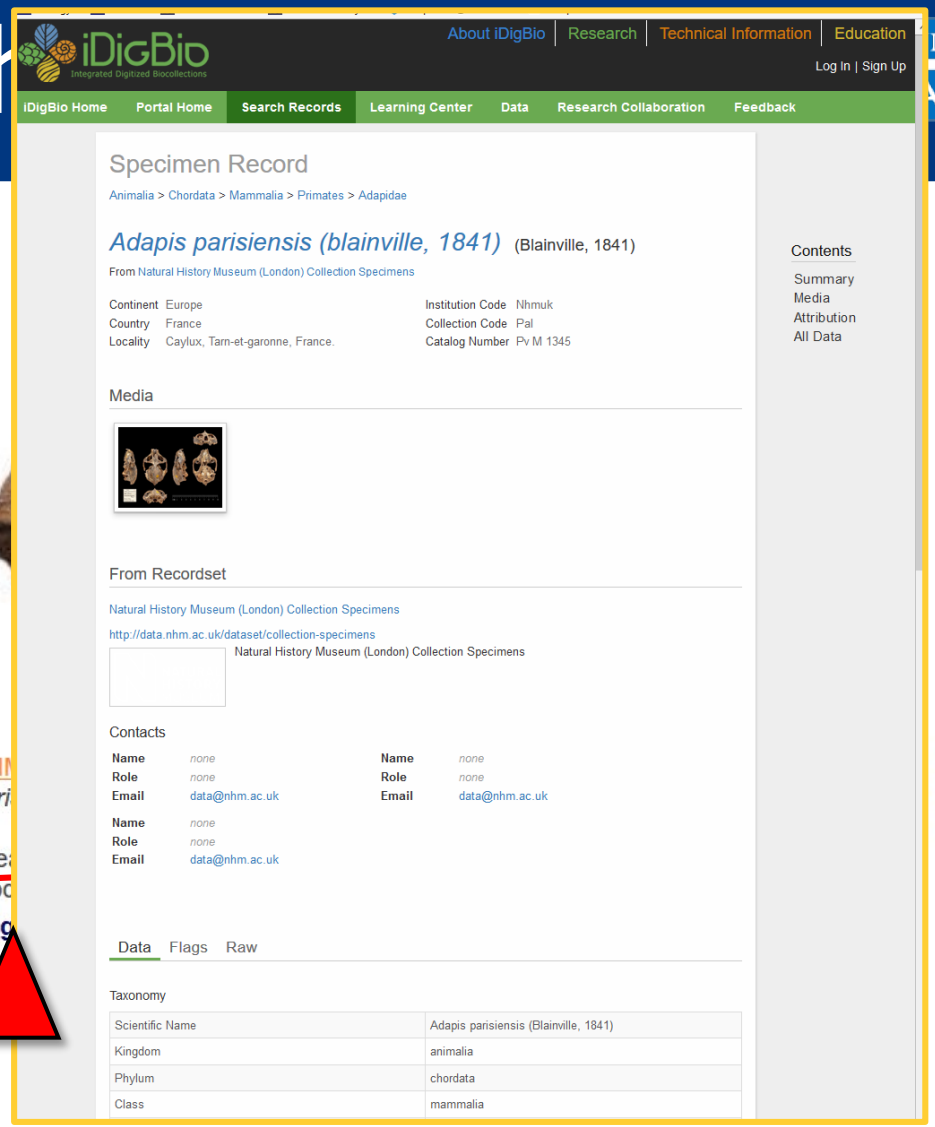


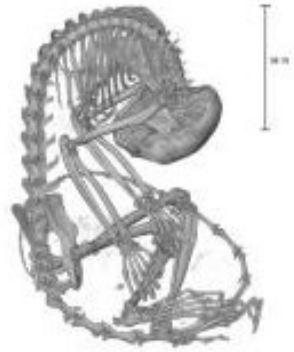
AMNH-M-183291, *Callimico goeldii*
iDigBio

Morph

Current holdings

- 26,400 specimens
  - 390 museums & collections
- 89,000 datasets





AMNH-M-183291, *Callimico goeldii*
iDigBio

Museum-integrated

Morpho

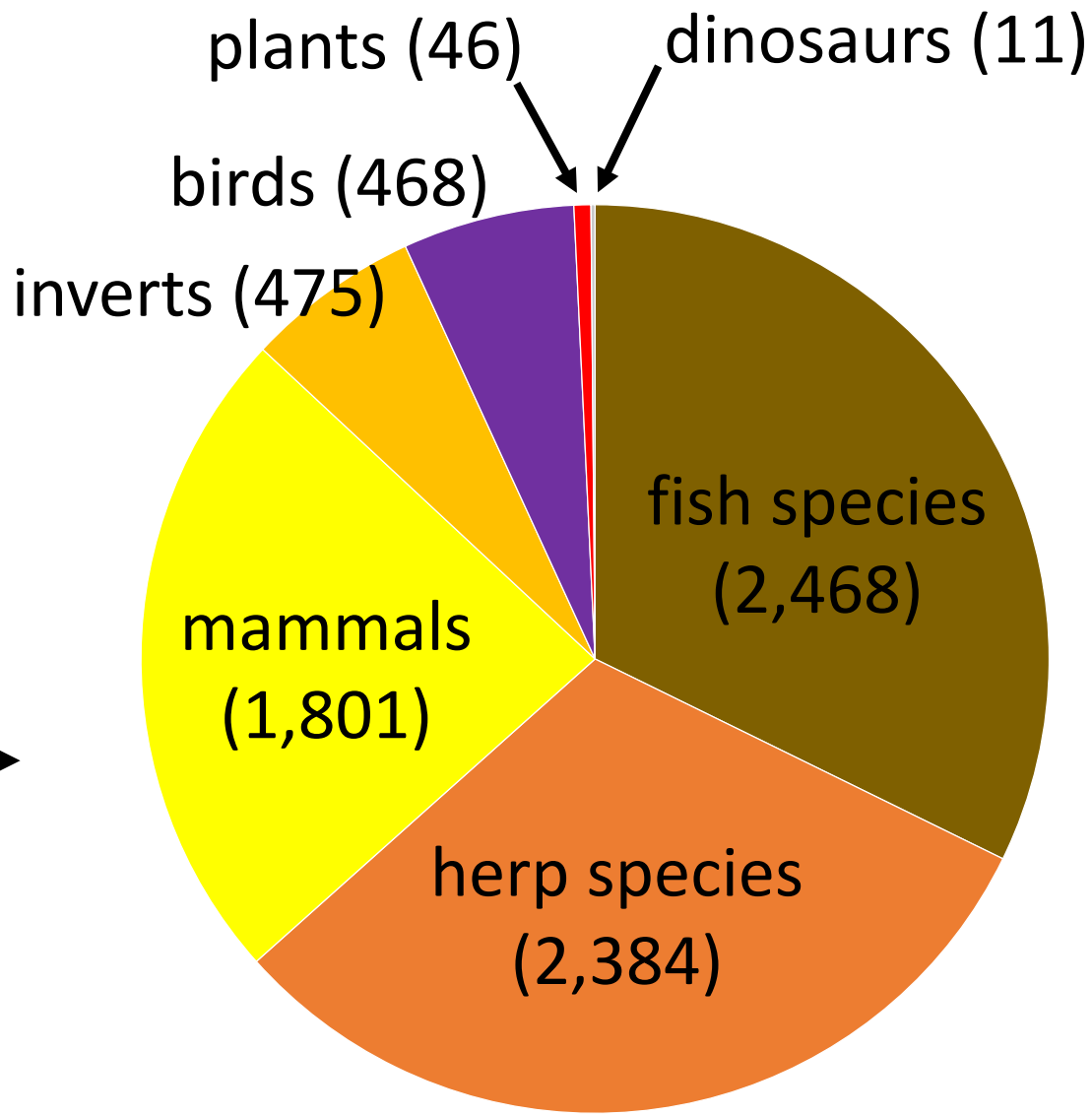
Current holdings

- 26,400 specimens
  - 390 museums & collections
- 89,000 datasets

MorphoSource

Current holdings

- 26,400 specimens
  - 390 museums & collections
- 89,000 datasets
- **11,300 species**
  - 3 kingdoms
  - 5 animal phyla

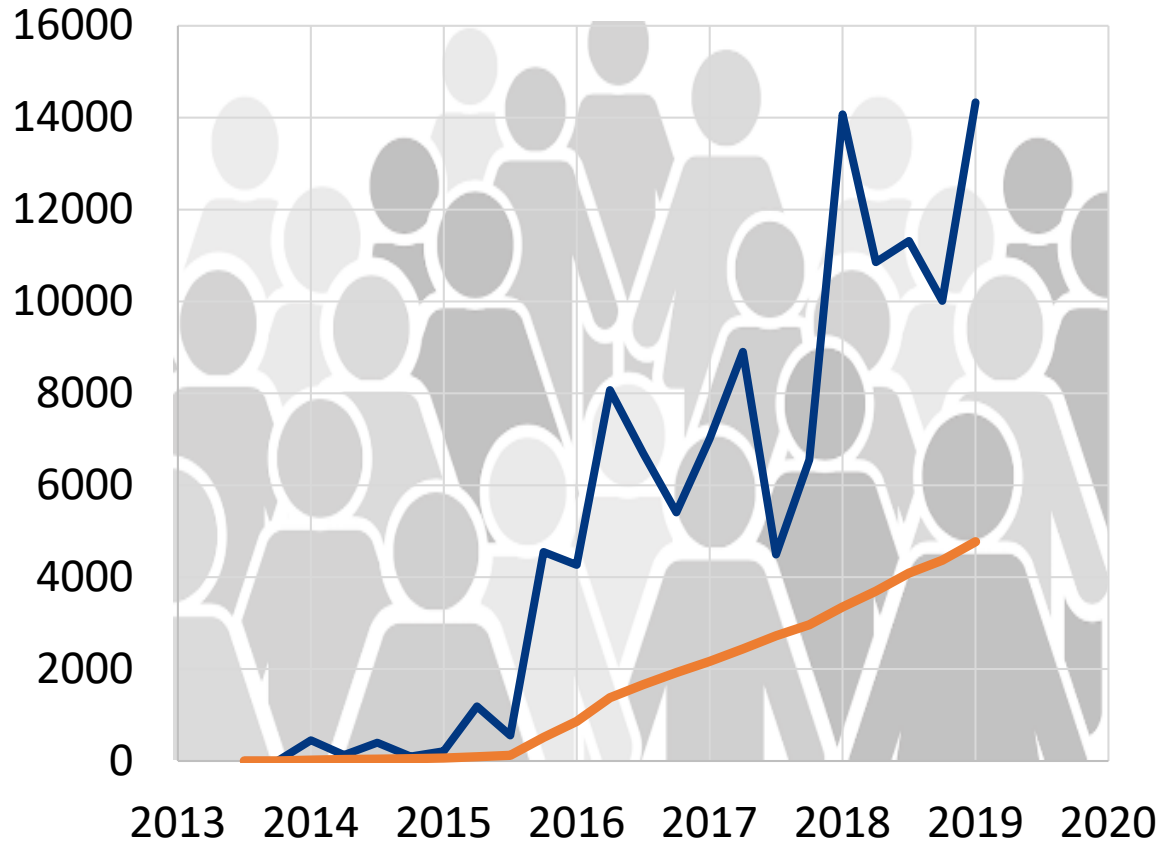


MorphoSource

User community / use

- 1,300 contributors
- 9,900 users
- Increasing use rate
- **181,000 downloads**
- 3,500,000 views

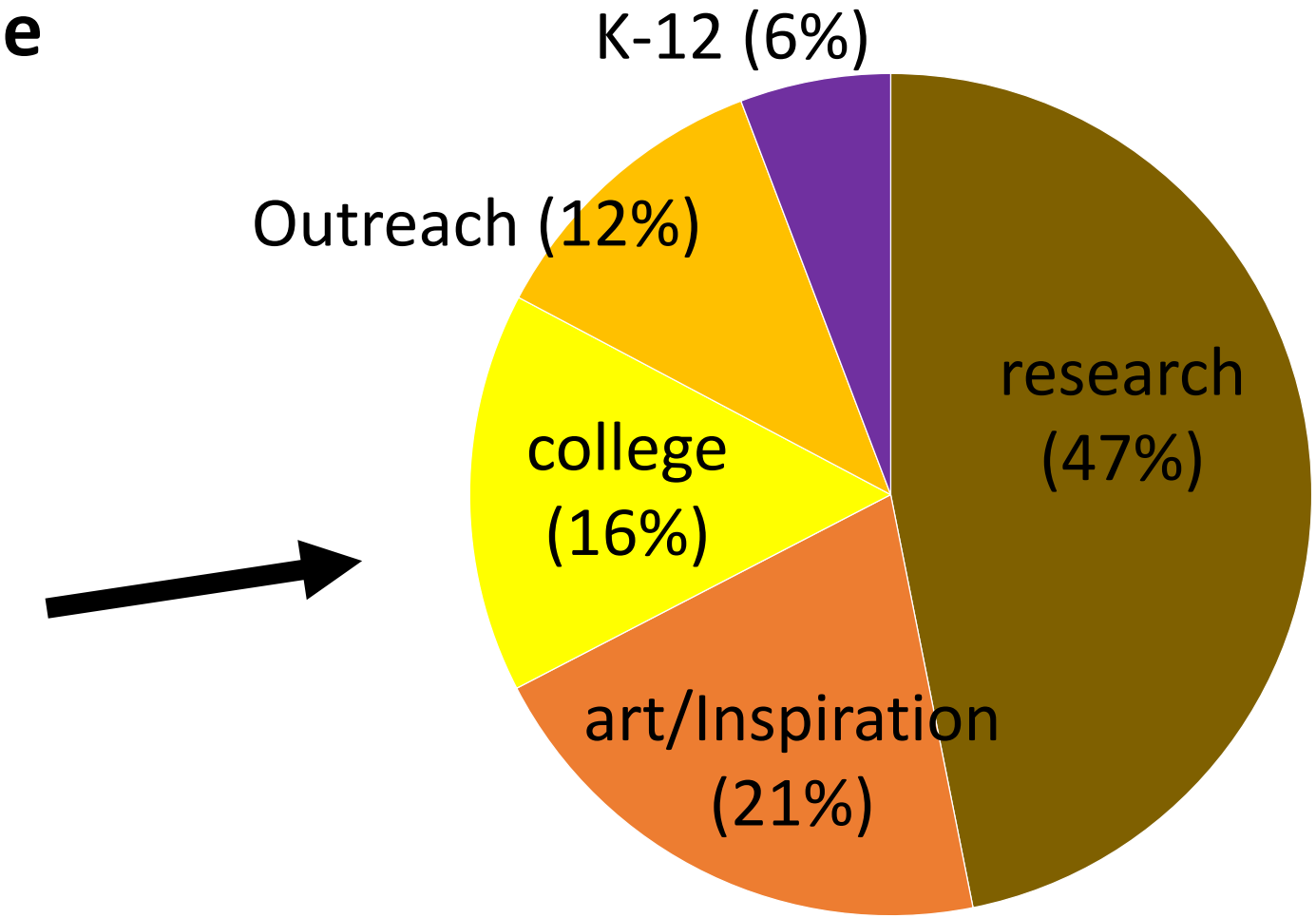
Quarterly Downloads & Downloaders



MorphoSource

User community / use

- 1,300 contributors
- 9,900 users
- Increasing use rate
- **181,000 downloads**
- 3,500,000 views

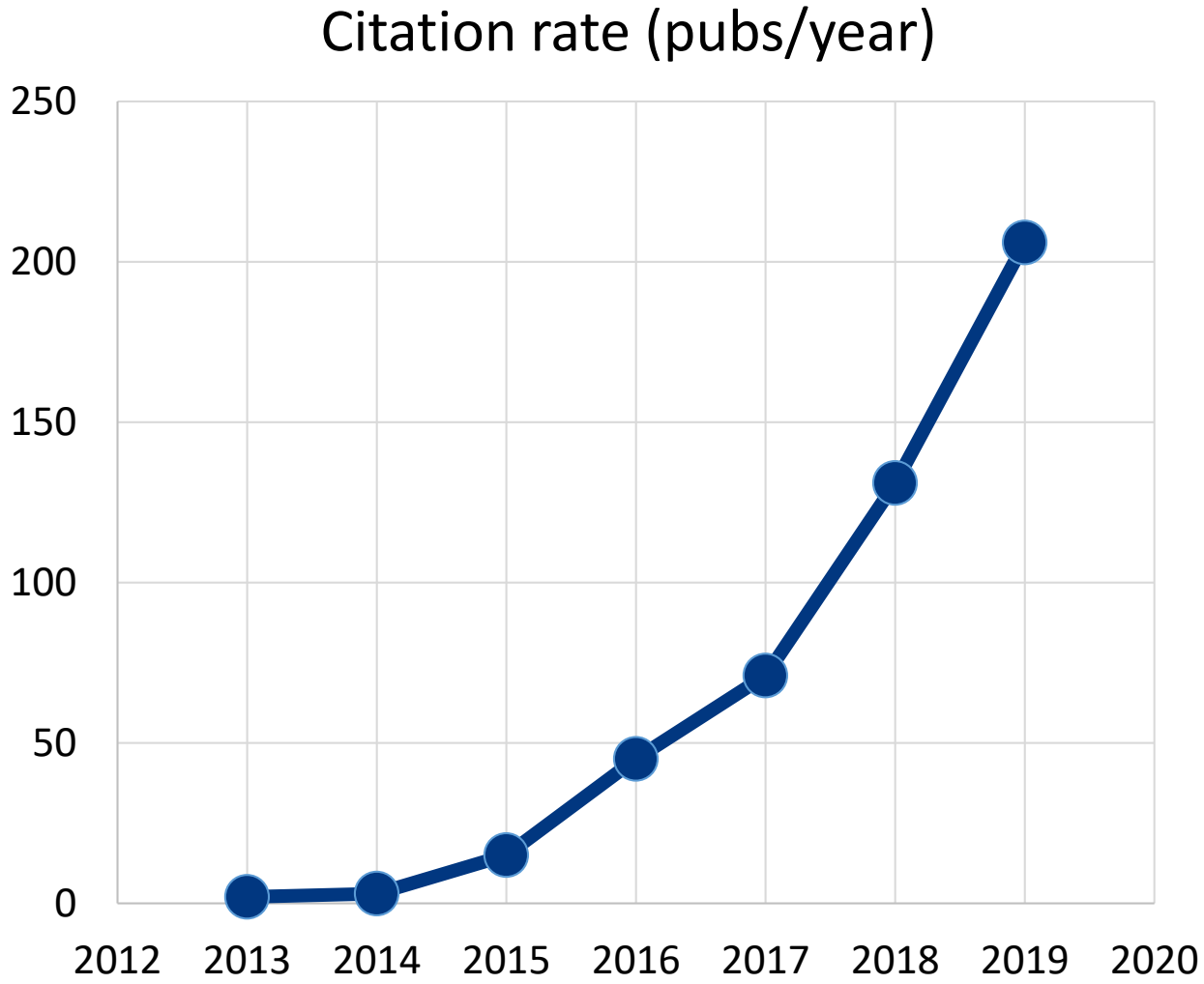


MorphoSource

Use in publications

- Growing presence
- 473 citations
- 123 journals
- 1,300 authors

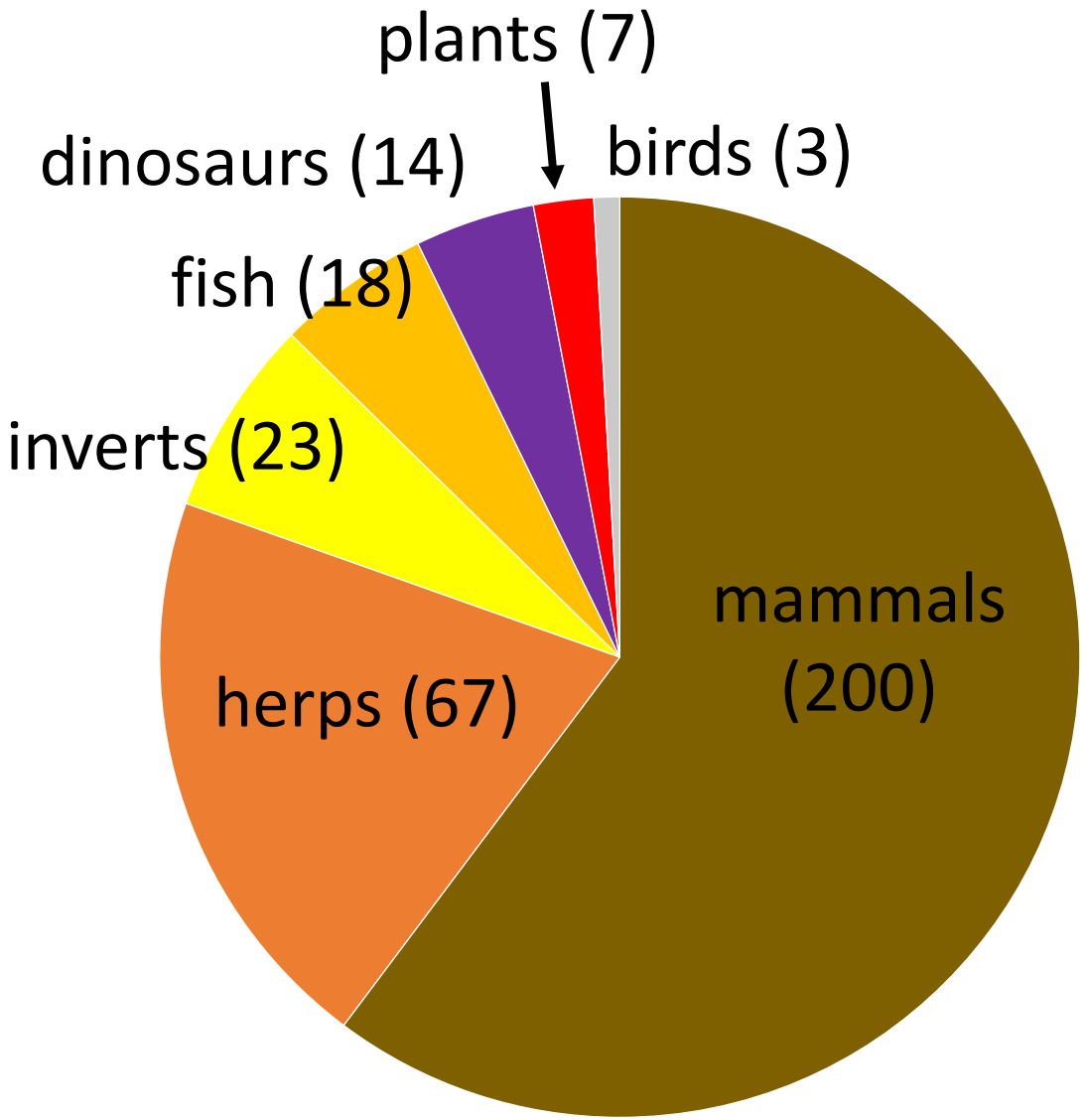
@MorphoSource



MorphoSource

Use in publications

- Growing presence
- 473 citations
- 123 journals
- 1,300 authors
- **Pubs by taxonomy**

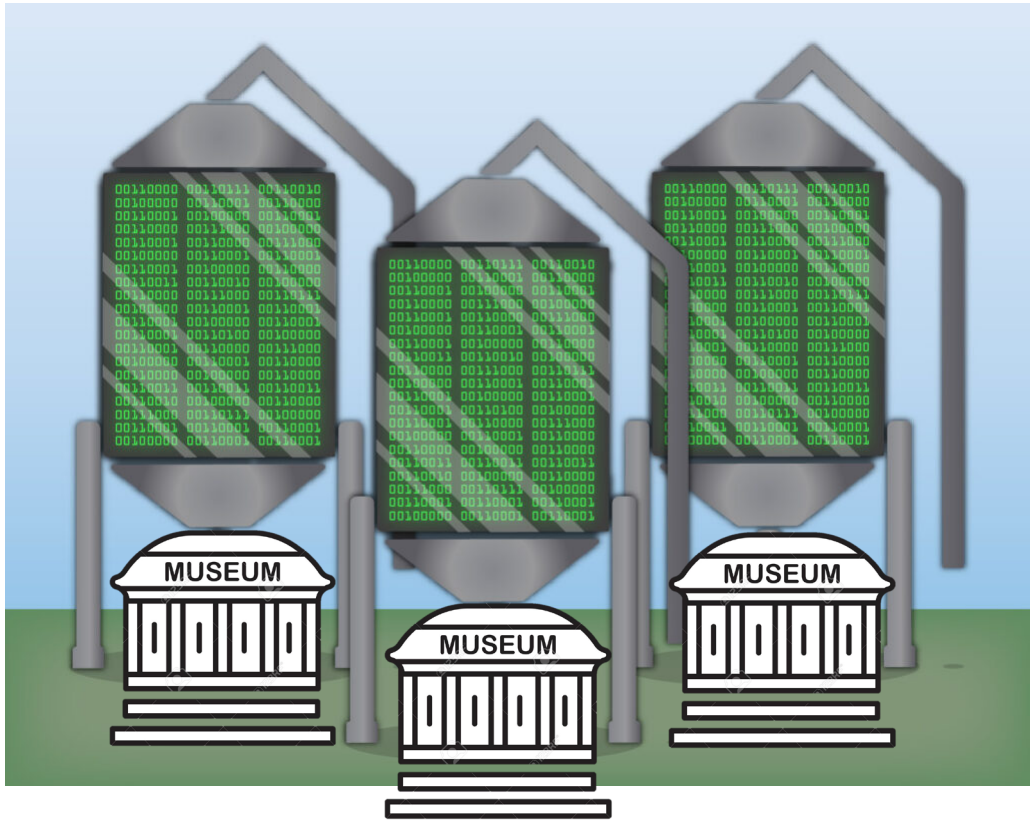


@MorphoSource

Why integration?

Current moment

- Proliferation of institutional resources to create & host data
- Aspirations to digitize very large amounts of data
- More intentional ownership policies by museums



Implications

- MorphoSource discoverability & access outweighed by simplicity of institutional resources

Why integration?

Use provisioned resources
(via institutions)
&
have powerful discovery too
(via domain-specialists)

Why integration?

Use provisioned resources
(via institutions)
&
have powerful discovery too
(via domain-specialists)