MorphoSource

Open Use 3D *Data Repository*

Doug M. Boyer / Assistant Prof. (Duke)
Julia M. Winchester / Postdoc (Duke)
Tim McGeary / Associate University Librarian (Duke)
Tim Ryan / Assoc. Prof (PSU)
Ed Gomes / Associate Dean of IT (Duke)
Gregg Gunnell / Director of Fossil Primates (Duke)
Seth Kaufman / CEO & Founder Whirligig Inc.

CAREER

BoyerLab

DUKE UNIVERSITY LIBRARIES

PennState

Duke LEMUR CENTER

MORPHO SOURCE

Started 2013

A place for 3D Data on vouchered specimens

Emulates a museum experience

www.morphosource.org

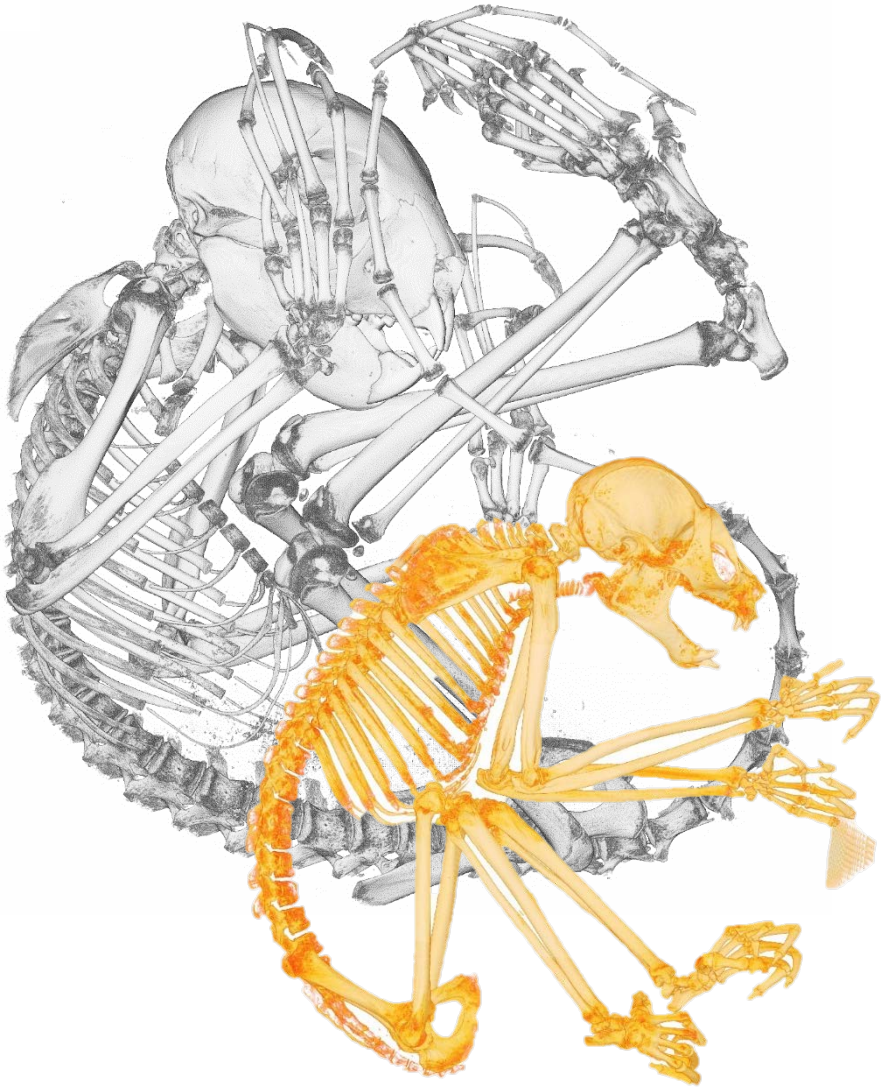
MORPHO SOURCE

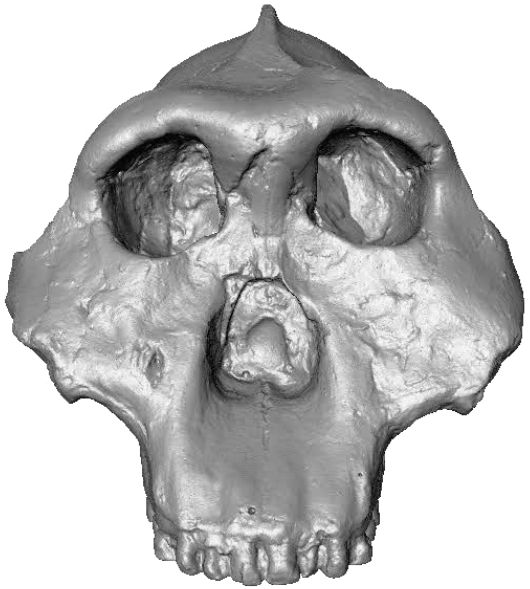
3D Data

Adhere to best-practice standards (Davies et al. 2016)

Volume
Modality – CT, MRI
Formats – tiff, dicom, jpeg, bmp

Surface
Modality – Laser, structured light, photogrammetry
Formats – ply, stl, obj



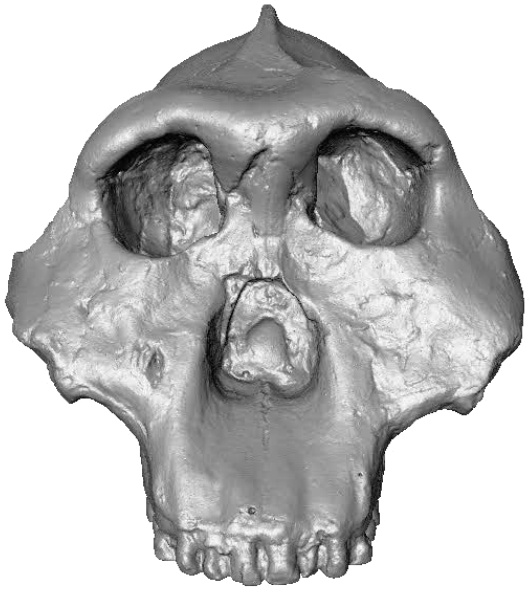


www.morphosource.org

MORPHO SOURCE

Outline

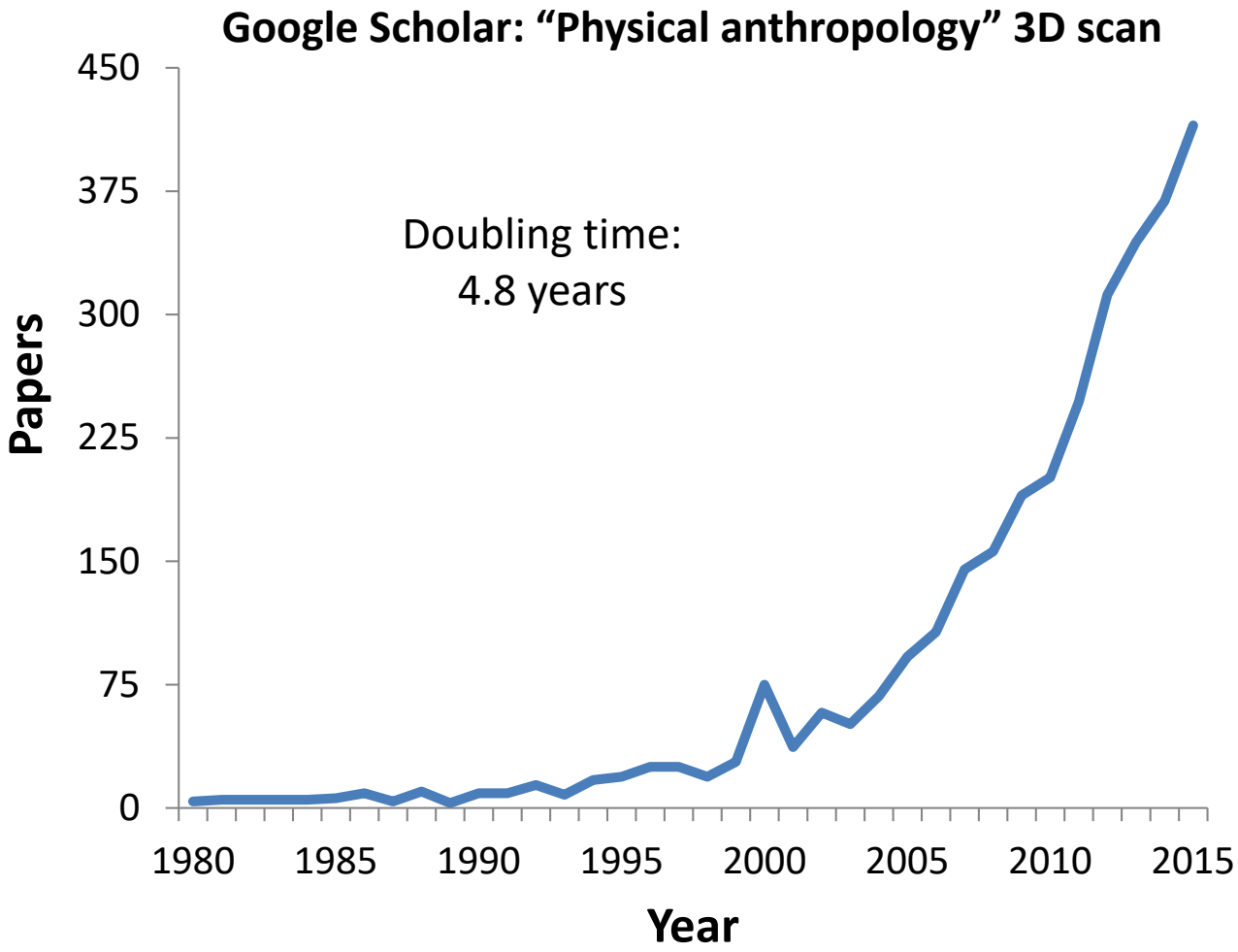
- Why do researchers need a repository for 3D data?
- How do researchers interact with MorphoSource?
- Sustainability and future growth



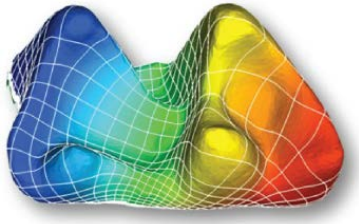
Why do we need a repository for 3D data?

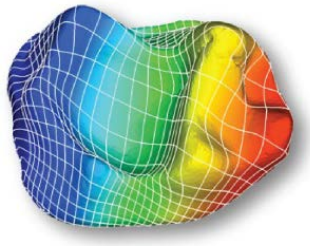
- 3D data are critical to research
- Transparency in science means access to these data

- Time consuming, costly, and risky to specimen
- Obligation to avoid taxing museums by rescanning specimens



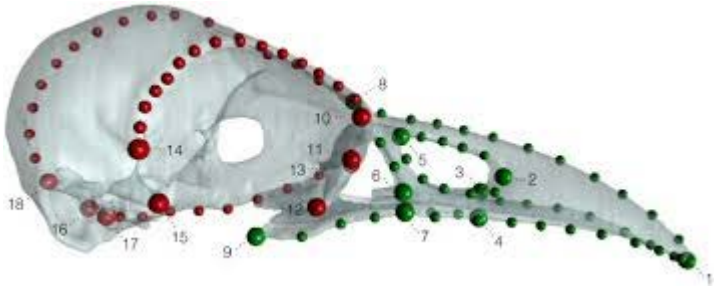
Why do we need a repository for 3D data?

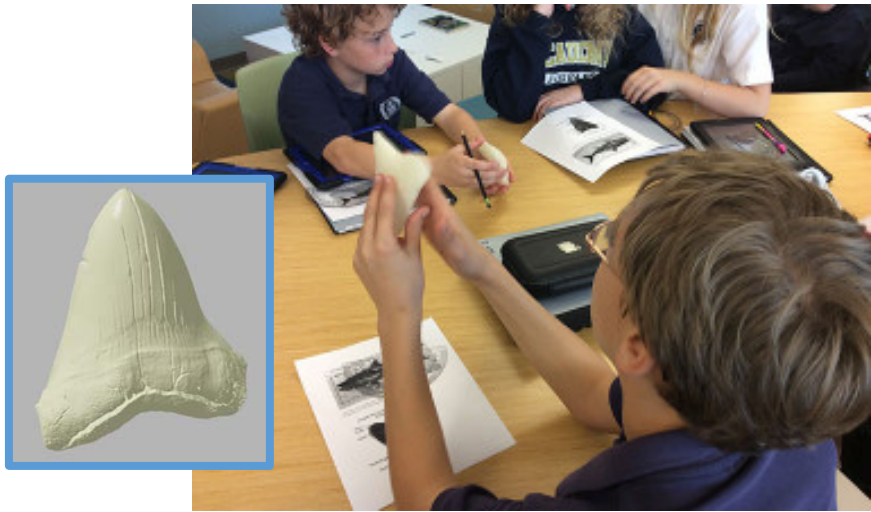




Benefits

- **Advance research** (quick access, bigger samples, more sophisticated analyses)
- **Support museums** (improved real & demonstrable value of collections, reduced demand for new scanning)
- **Empower science education** (access to previously off-limits collections)





Why not use generic archives?

- ‘publication packet’ organization
- Lack metadata specific to 3D formats and biological specimens
- Hinder accessibility, discoverability, and long term readability

Shouldn’t museums host data on their specimens?

ONLINE REPOSITORY of FOSSILS
MUSEUM OF PALEONTOLOGY
UNIVERSITY OF MICHIGAN

Smithsonian Institution
Digitization | 3D

They should…..

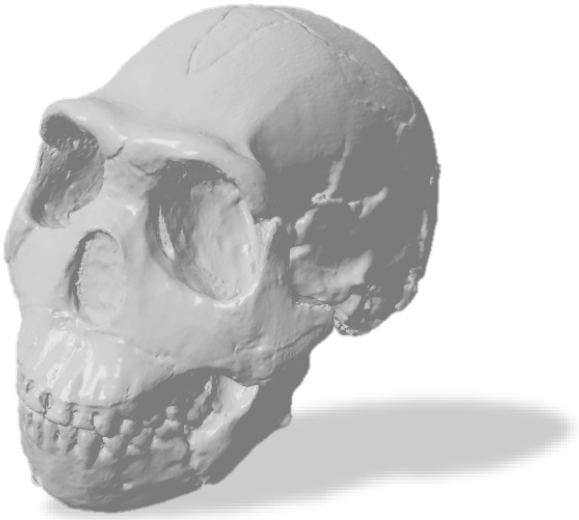
- Costly to start and manage a digital archive
- Dividing data among isolated platforms limits discoverability
- Researchers need control of when and how they publish raw data
- Redundancy is not necessarily a bad thing

www.morphosource.org

MORPHO SOURCE

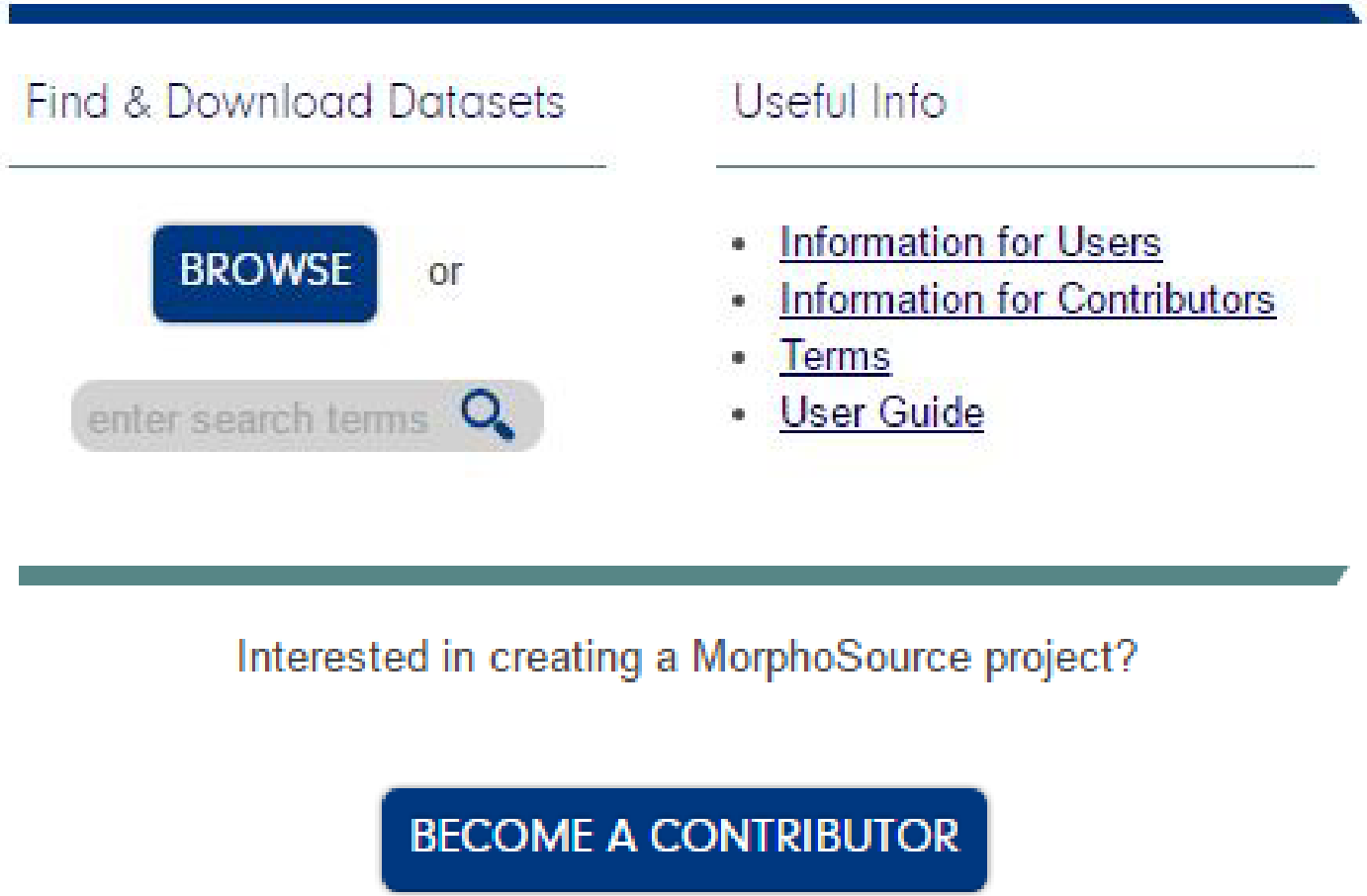
Outline

- Why do researchers need a repository for 3D data?
- **How do researchers interact with MorphoSource?**
- Sustainability and future growth



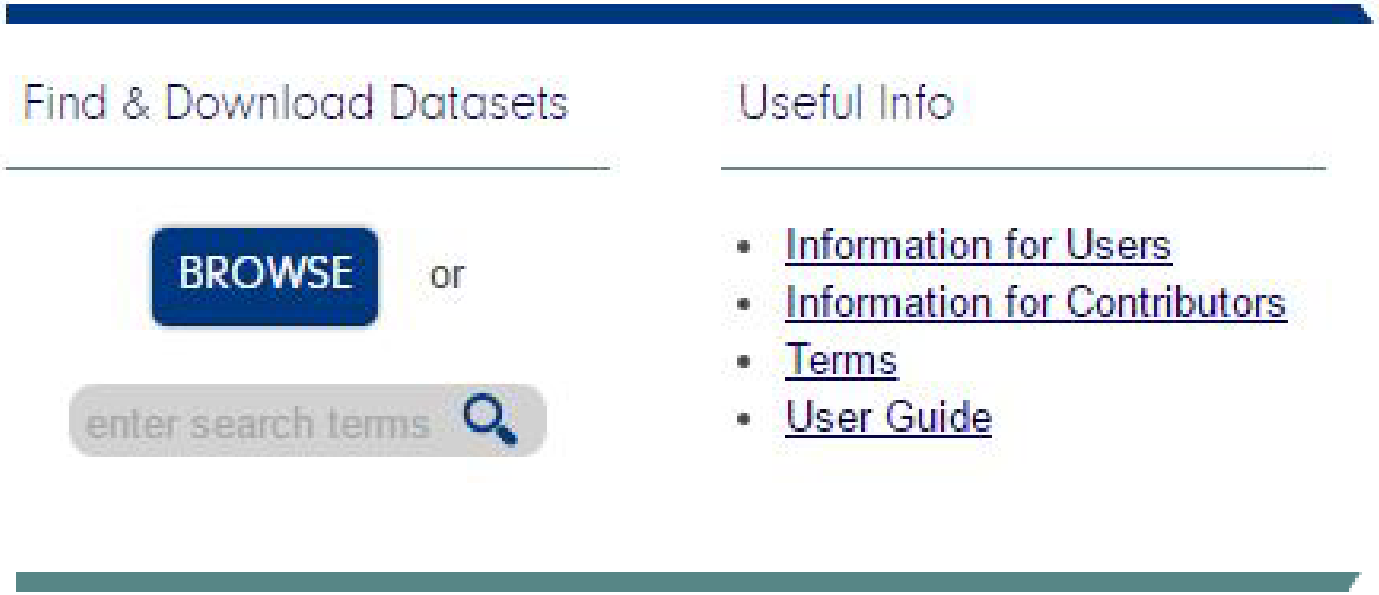
Homo naledi

Researchers contribute their datasets



www.morphosource.org

Researchers contribute their datasets



CREATE A MORPHOSOURCE PROJECT

Researchers contribute their datasets

www.morphosource.org

Researchers contribute their datasets

Project: AMNH Mammal Collection

BACK

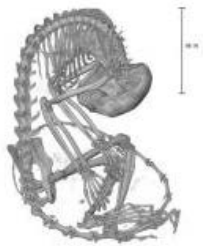
55 Project Specimens

NEW SPECIMEN

Group by: Specimen Number | Genus | Species

AMNH-M-174360, *Microcebus rufus*
iDigBio

AMNH-M-183285, *Callimico goeldii*
iDigBio



AMNH-M-183291, *Callimico goeldii*
iDigBio

AMNH-M-185280, *Callimico goeldii*
iDigBio

AMNH-M-185640, *Daubentonia madagascariensis*
iDigBio

AMNH-M-185643, *Daubentonia madagascariensis*
iDigBio

AMNH-M-187861, *Cynocephalus volans*
iDigBio

AMNH-M-188009, *Alouatta seniculus*

www.morphosource.org

Researchers can obtain data GUIDs (DOIs)

DNMNH-G-22149, *Procavia sp.*

M6014-5467
3D Mesh (Polygon File Format), 92.61 MB

Citation Elements

ADD

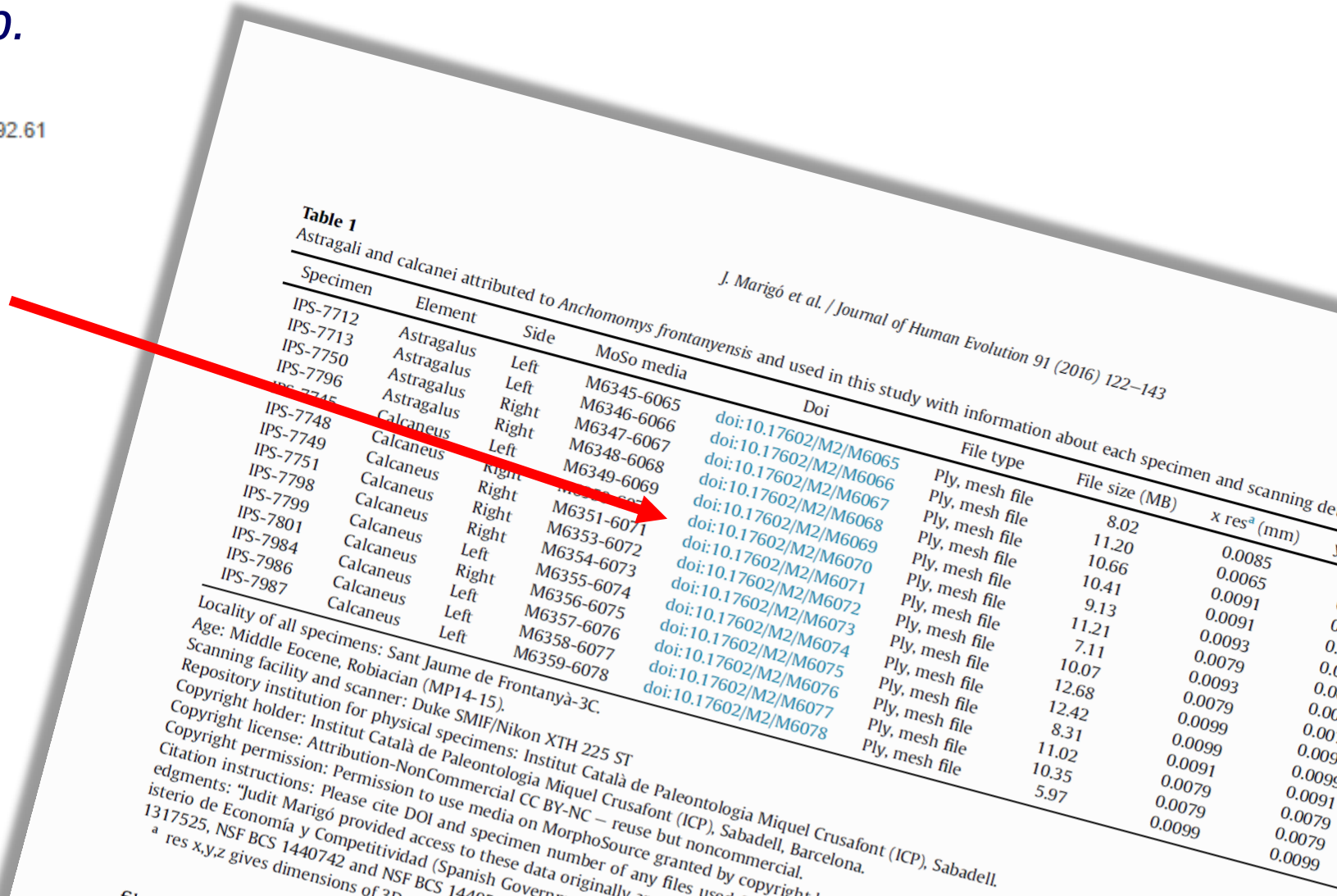


Table 1
Astragali and calcanei attributed to *Anchomomys frontanyensis* and used in this study with information about each specimen and scanning de

Specimen	Element	Side	MoSo media	Doi	File type	File size (MB)	x res[a] (mm)
IPS-7712	Astragalus	Left	M6345-6065	doi:10.17602/M2/M6065	Ply, mesh file	8.02	0.0085
IPS-7713	Astragalus	Left	M6346-6066	doi:10.17602/M2/M6066	Ply, mesh file	11.20	0.0065
IPS-7750	Astragalus	Right	M6347-6067	doi:10.17602/M2/M6067	Ply, mesh file	10.66	0.0091
IPS-7796	Astragalus	Right	M6348-6068	doi:10.17602/M2/M6068	Ply, mesh file	10.41	0.0091
[illegible]	Calcaneus	Left	M6349-6069	doi:10.17602/M2/M6069	Ply, mesh file	9.13	0.0093
IPS-7748	Calcaneus	Right	[illegible]	doi:10.17602/M2/M6070	Ply, mesh file	11.21	0.0079
IPS-7749	Calcaneus	Right	M6351-6071	doi:10.17602/M2/M6071	Ply, mesh file	7.11	0.0093
IPS-7751	Calcaneus	Right	M6353-6072	doi:10.17602/M2/M6072	Ply, mesh file	10.07	0.0079
IPS-7798	Calcaneus	Left	M6354-6073	doi:10.17602/M2/M6073	Ply, mesh file	12.68	0.0099
IPS-7799	Calcaneus	Right	M6355-6074	doi:10.17602/M2/M6074	Ply, mesh file	12.42	0.0099
IPS-7801	Calcaneus	Left	M6356-6075	doi:10.17602/M2/M6075	Ply, mesh file	8.31	0.0091
IPS-7984	Calcaneus	Left	M6357-6076	doi:10.17602/M2/M6076	Ply, mesh file	11.02	0.0079
IPS-7986	Calcaneus	Left	M6358-6077	doi:10.17602/M2/M6077	Ply, mesh file	10.35	0.0079
IPS-7987	Calcaneus	Left	M6359-6078	doi:10.17602/M2/M6078	Ply, mesh file	5.97	0.0099

Locality of all specimens: Sant Jaume de Frontanyà-3C.
Age: Middle Eocene, Robiacian (MP14-15).
Scanning facility and scanner: Duke SMIF/Nikon XTH 225 ST
Repository institution for physical specimens: Institut Català de Paleontologia Miquel Crusafont (ICP), Sabadell, Barcelona.
Copyright holder: Institut Català de Paleontologia Miquel Crusafont (ICP), Sabadell.
Copyright license: Attribution-NonCommercial CC BY-NC – reuse but noncommercial.
Copyright permission: Permission to use media on MorphoSource granted by copyright
Citation instructions: Please cite DOI and specimen number of any files
edgments: "Judit Marigó provided access to these data originally
isterio de Economía y Competitividad (Spanish Govern
1317525, NSF BCS 1440742 and NSF BCS
[a] res x,y,z gives dimensions of

And/or increase accessibility of collections

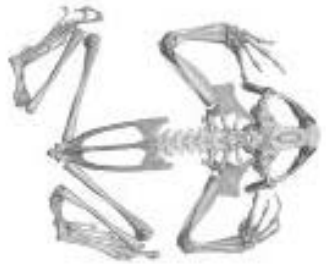
Project: Digitizing the Florida Museum of Natural History's Herpetology collections

BACK

80 Project Specimens

Order by: Specimen number | Taxonomic name

UF-11978, *Sphenodon punctatus*



UF-H-100788, *Spea multiplicata*

UF-H-14110, *Sphenodon punctatus*

UF-H-43506, *Kinosternon hirtipes tarascense*

www.morphosource.org

Researchers control levels of access

Access

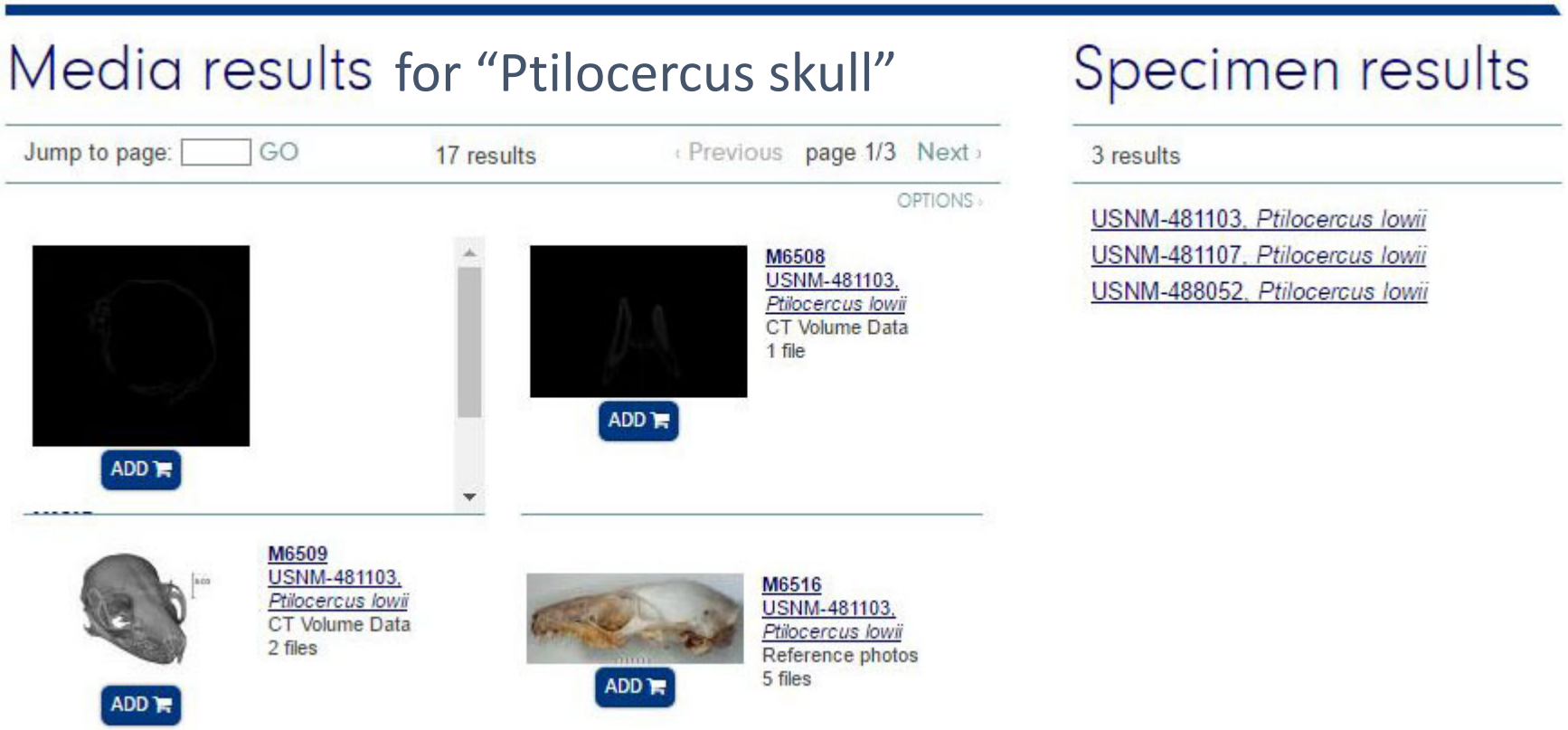
- On upload, data are hidden from public database.
- Data can be made visible with author-approved or open downloading.
- Reuse of shared data is regulated through cc licenses.

Anyone can get data from MorphoSource

Exploring specimens & media

Anyone can get data from MorphoSource

Exploring specimens & media

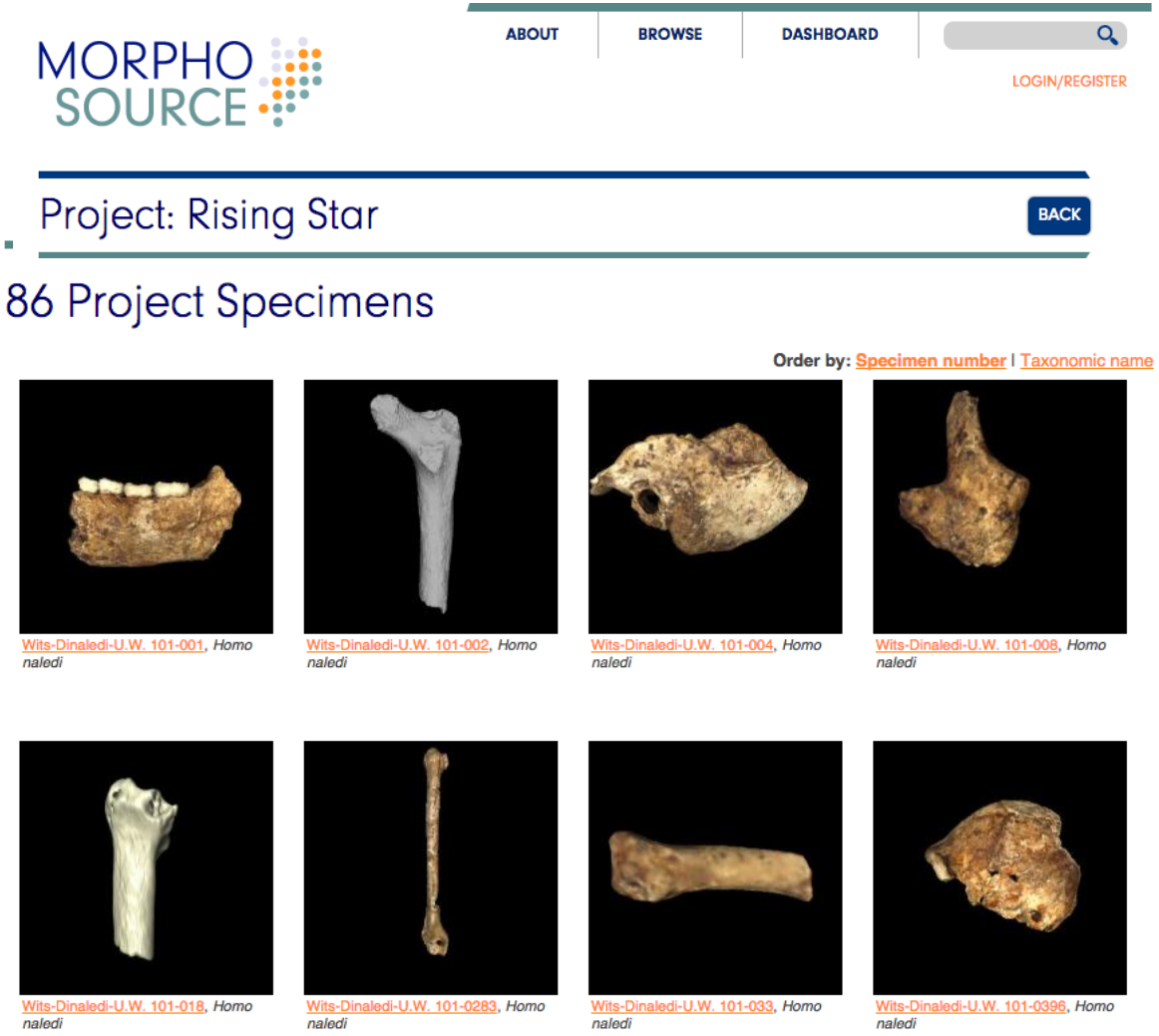


www.morphosource.org

Data on MorphoSource

FOSSIL HOMININS

- 91 specimens of newly described South African hominin
- *Homo naledi* (eLife 2015)



www.morphosource.org

Data on MorphoSource

PRIMATES

- ~1000 primate skulls
- 95% of all species
- (Nature Scientific Data, 2016)

Media results

Jump to page: GO 311 results ‹ Previous page 1/39 Next ›

OPTIONS ›

M2601
MCZ-25626,
Cercocebus torquatus
DICOM medical imaging data image series ZIP file
3.01 MB

M2602
MCZ-25630,
Cercocebus torquatus
DICOM medical imaging data image series ZIP file
2.59 MB

M2852
MCZ-62639,
Cercocebus torquatus
DICOM medical imaging data image series ZIP file
2.63 MB

M2853
MCZ-32625,
Cercocebus torquatus
DICOM medical imaging data image series ZIP file
3.47 MB

M2854
MCZ-32624,
Cercocebus torquatus
DICOM medical imaging data image series ZIP file
3.44 MB

M2860
MCZ-23195,
Cercocebus sp.
DICOM medical imaging data image series ZIP file
2.82 MB

M2861
MCZ-19982,
Cercocebus torquatus
DICOM medical imaging data image series ZIP file
3.52 MB

M2862
MCZ-19184,
Cercocebus torquatus
DICOM medical imaging data image series ZIP file
3.54 MB

Specimen results

308 results

MCZ-10131, *Saimiri oerstedii*
MCZ-10132, *Saimiri oerstedii*
MCZ-10133, *Saimiri oerstedii*
MCZ-10134, *Saimiri oerstedii*
MCZ-10138, *Ateles geoffroyi*
MCZ-12758, *Macaca fascicularis*
MCZ-14657, *Euoticus elegantulus*
MCZ-14659, *Galago alleni*
MCZ-14725, *Cercocebus albigena*
MCZ-15312, *Pan troglodytes troglodytes*
MCZ-15324, *Saguinus sp.*
MCZ-16075, *Galago senegalensis*
MCZ-16354, *Eulemur fulvus fulvus*
MCZ-16356, *Eulemur fulvus rufus*
MCZ-16370, *Eulemur fulvus rufus*
MCZ-16375, *Propithecus verreauxi verreauxi*
MCZ-16382, *Varecia variegata variegata*
MCZ-16390,
MCZ-16391, *Lemur catta*
MCZ-16392, *Lemur catta*
MCZ-16393, *Eulemur fulvus rufus*
MCZ-17548, *Perodicticus potto*
MCZ-17550, *Perodicticus potto*
MCZ-17589, *Galago alleni*
MCZ-17590, *Euoticus elegantulus*
MCZ-17591, *Euoticus elegantulus*
MCZ-17592, *Euoticus elegantulus*
MCZ-17593, *Euoticus elegantulus*
MCZ-18607, *Perodicticus potto*

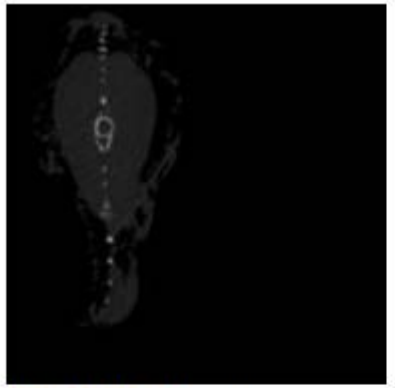
www.morphosource.org

Data on MorphoSource

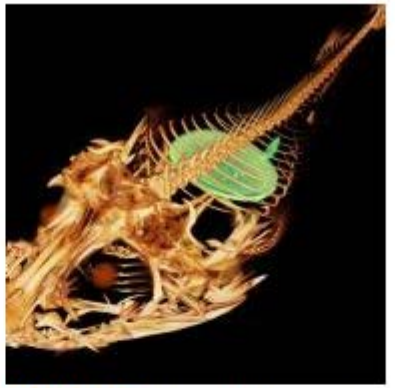
FISH

- Sharks
- Rays
- Teleosts

BMNH-2001.1.21.58-71, *Gastromyzon borneensis*



OSU-013876, *Cottus asper*



ROM-50997, *Chorisochismus dentex*

TCWC-11247.01, *Mystus armatus*

TCWC-13988.05, *Corydoras sp.*

TCWC-16994.04, *Cyprinella lepida*

TCWC-UnCat

UW-118012, *Artediellus camchaticus*

www.morphosource.org

Data on MorphoSource

INVERTEBRATES

- Beetles, Bees, Centipedes, Molluscs

Specimen Media

M1642, New centipede source images
unknown image series ZIP file, 1.86 GB

Specimen Information

Specimen: CBSS-CHP517-1, *Eupolybothrus cavernicolus*, Unvouchered, Female
The Croatian Biospeleological Society (CBSS) is non-governmental, non-profit organization acknowledged, registered and supported by Ministry of Science, Education and Sport that has been working successfully since its foundation in 1996. It is an organization for biospeleology that deals with research and conservation of subterranean fauna and its habitats on the whole territory of the Republic of Croatia with its members being both scientists, experts and cavers.

Locality: country: Croatia; stateProvince: Knin; locality: NP Krka, village Kistanje, Hydroelectric power plant Miljacka, cave Miljacka II; verbatimElevation: 115 m;
44.000306, 16.016250

Institution: Croatian Biospeleological Society (CBSS), Zagreb, Croatia

www.morphosource.org

Educational Collections

25,000 views, 2,200 downloads

Leveraging other work

- See talk by **Winchester et al.** today

VertNet

PBDB

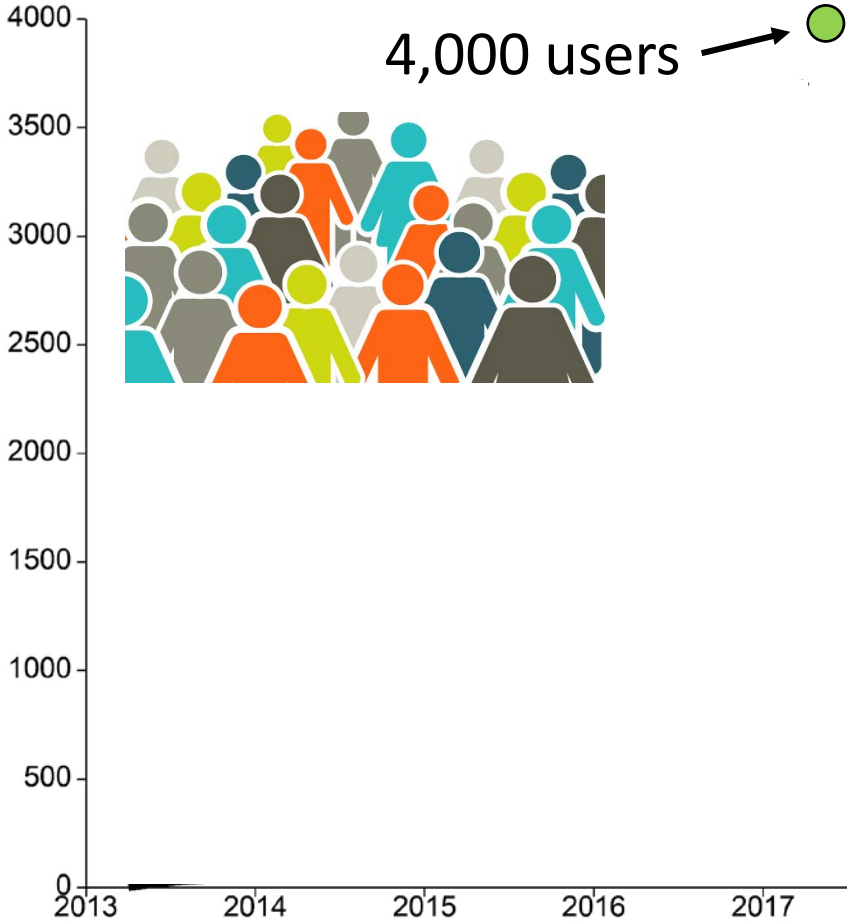
iDigBio
Integrated Digitized Biocollections

GBIF
Global Biodiversity Information Facility

GRBio
GLOBAL REGISTRY OF BIODIVERSITY REPOSITORIES

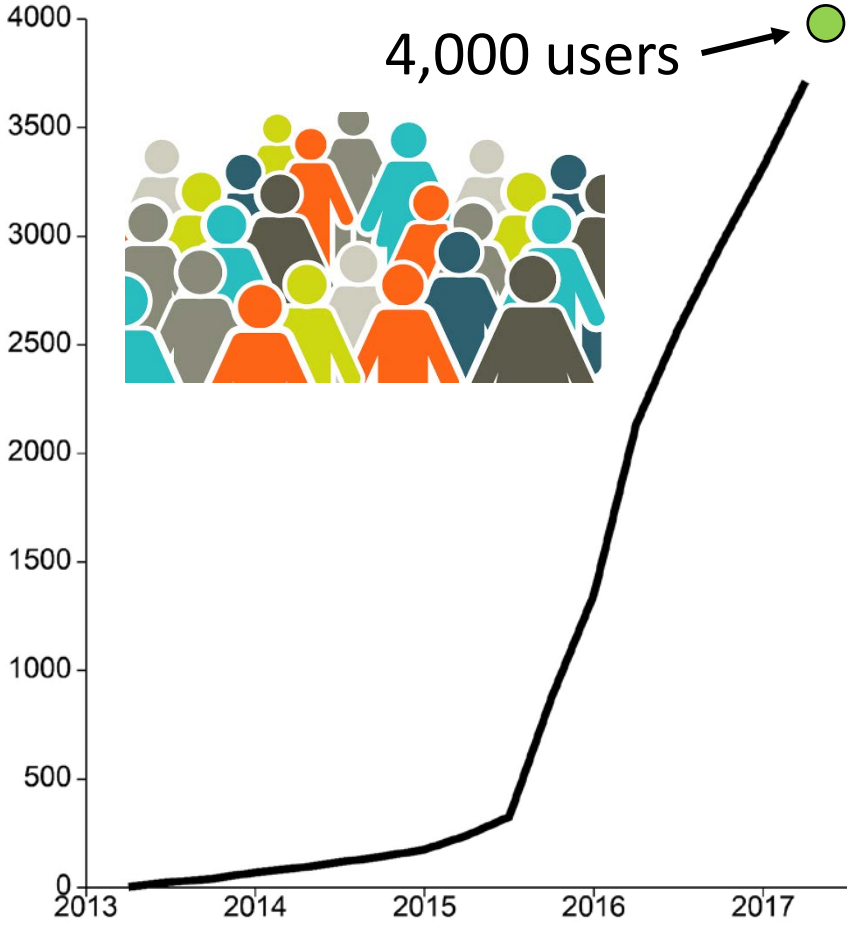
Is MorphoSource being used?

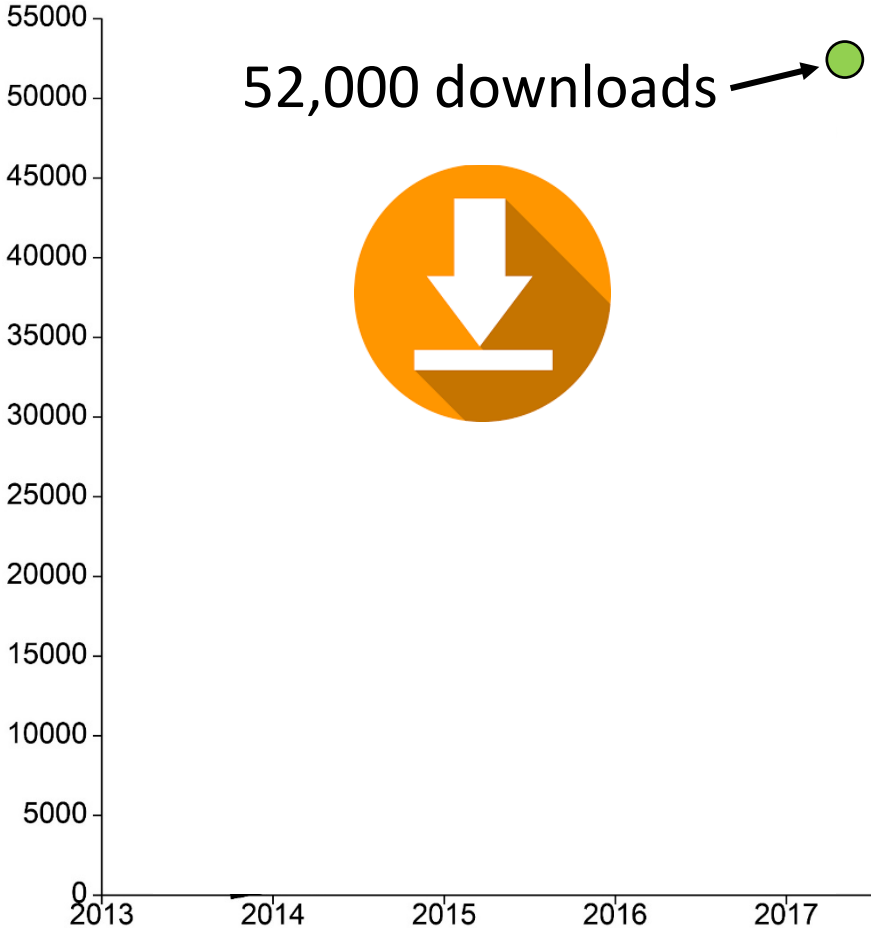
User/use growth since 2013 (quarterly data)



Is MorphoSource being used?

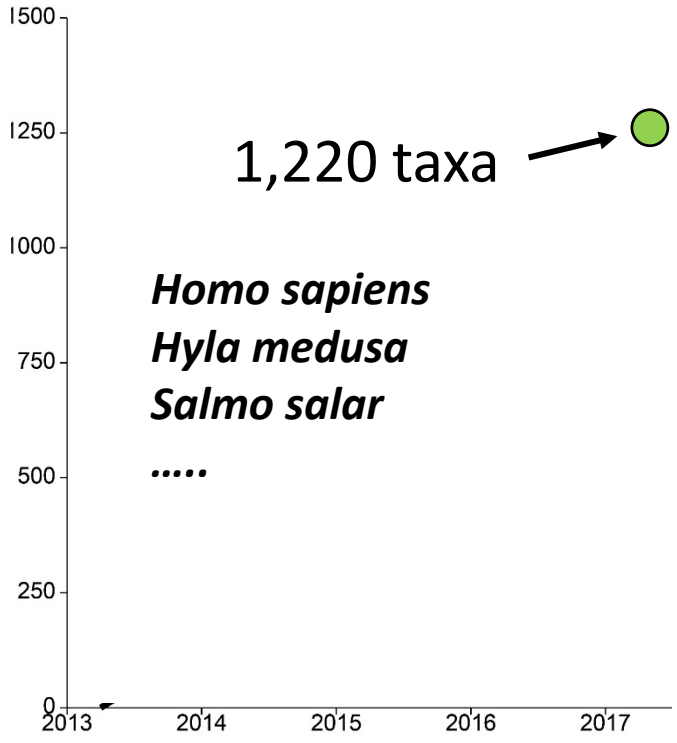
User/use growth since 2013 (quarterly data)





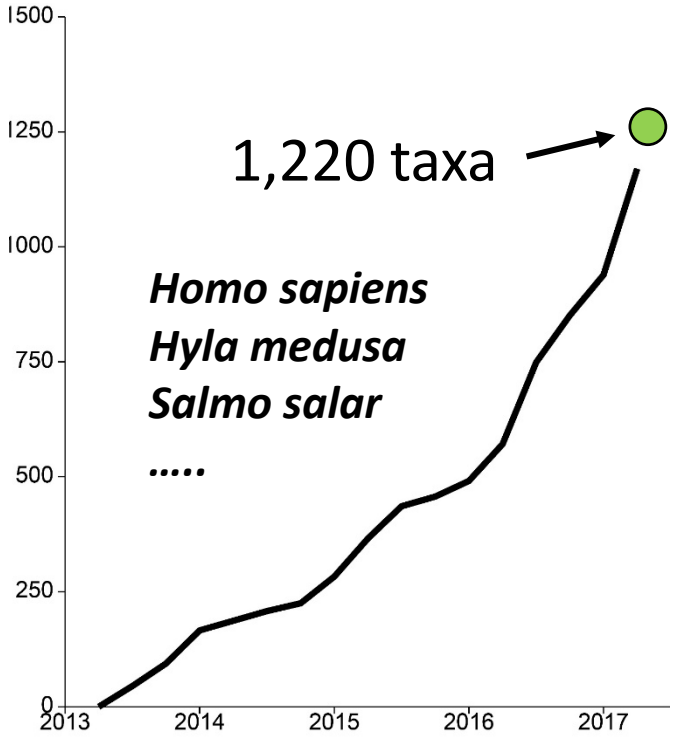
Is MorphoSource being used?

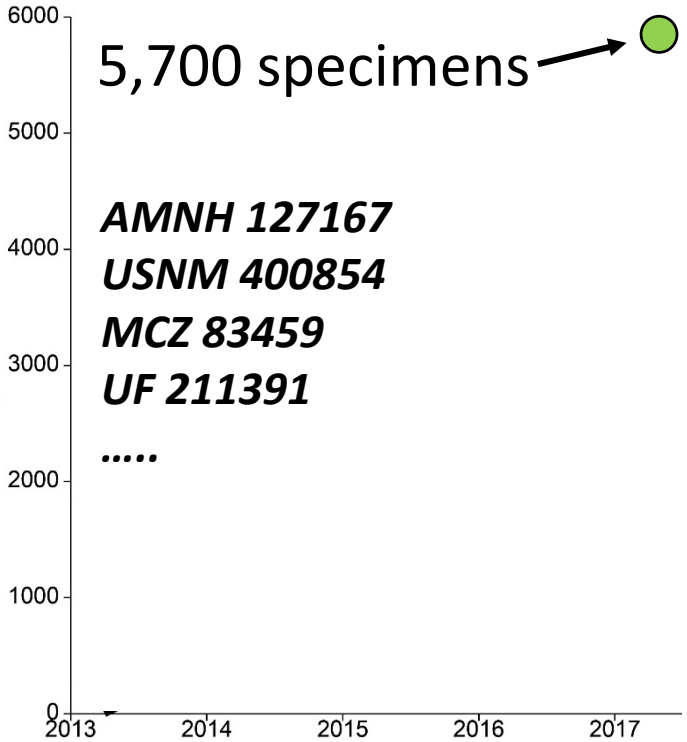
Growth of holdings since 2013 (quarterly data)



Is MorphoSource being used?

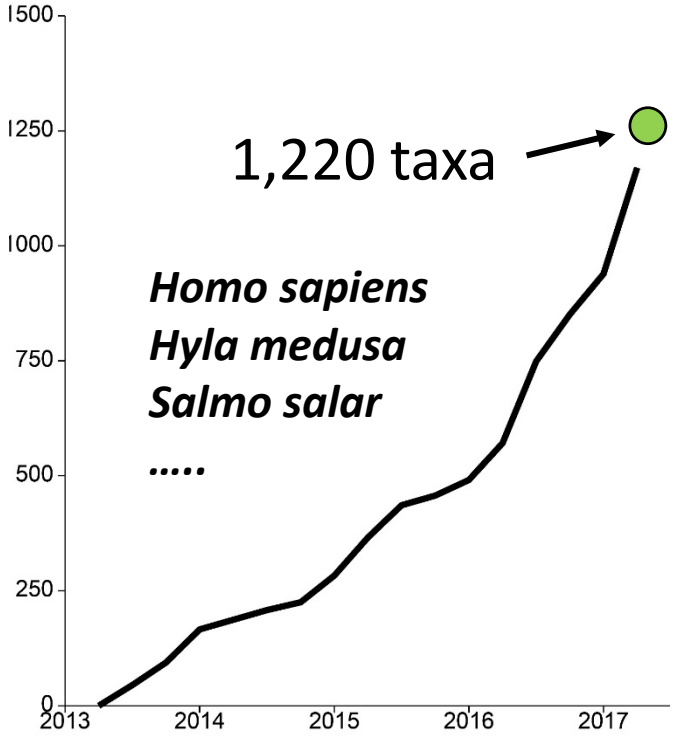
Growth of holdings since 2013 (quarterly data)

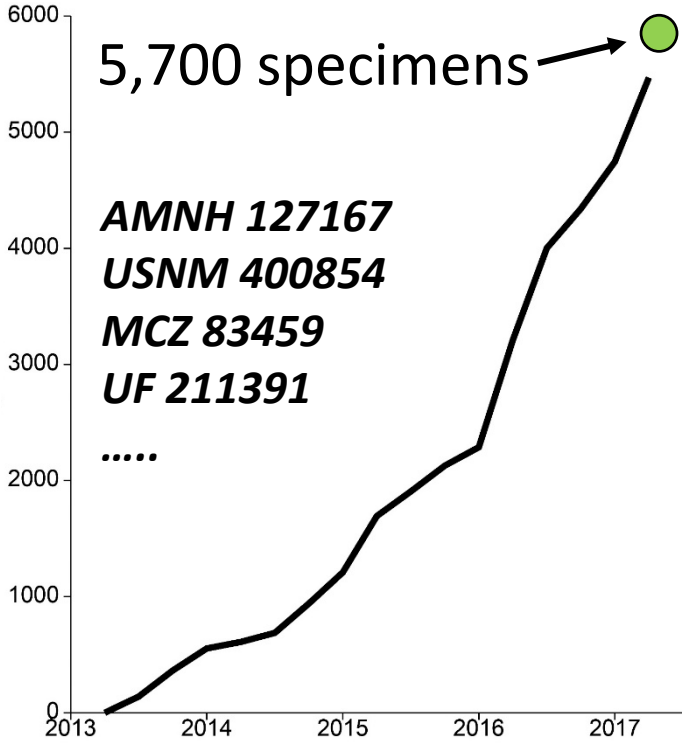


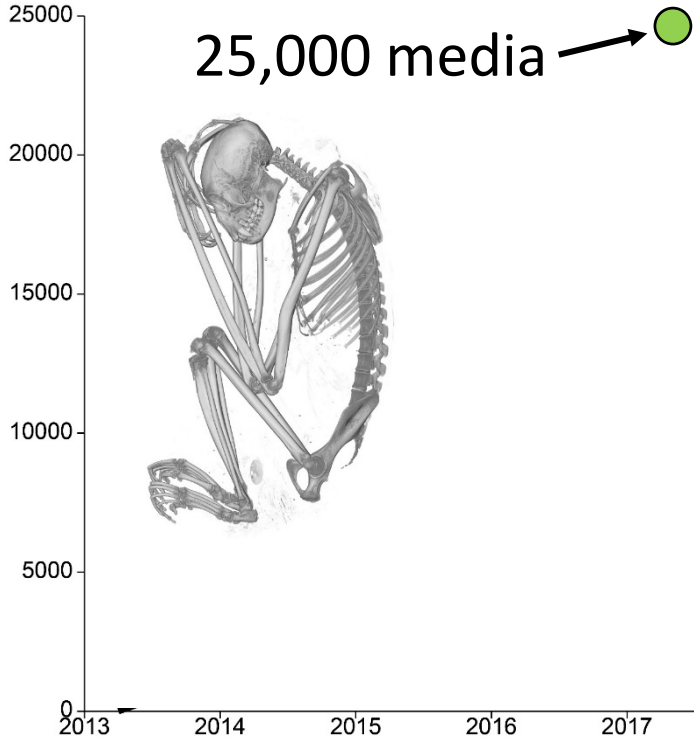


Is MorphoSource being used?

Growth of holdings since 2013 (quarterly data)

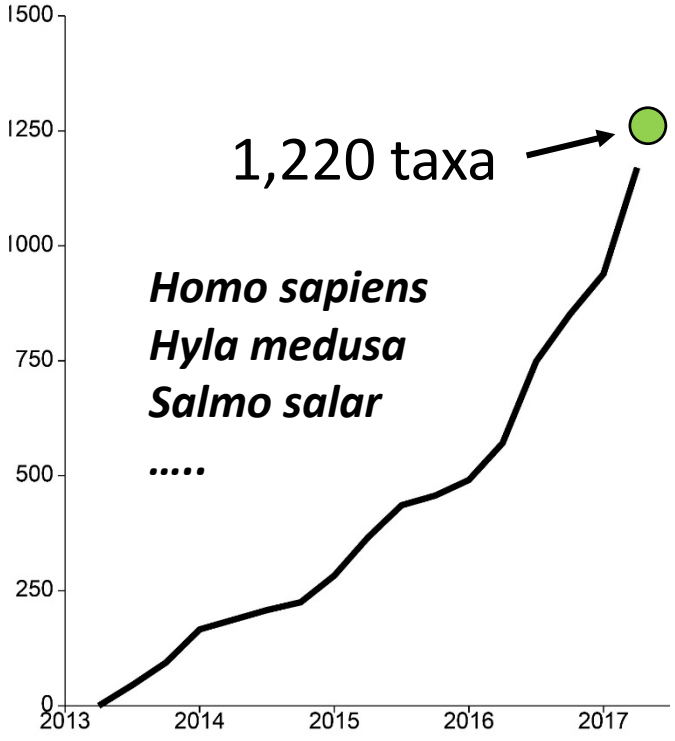


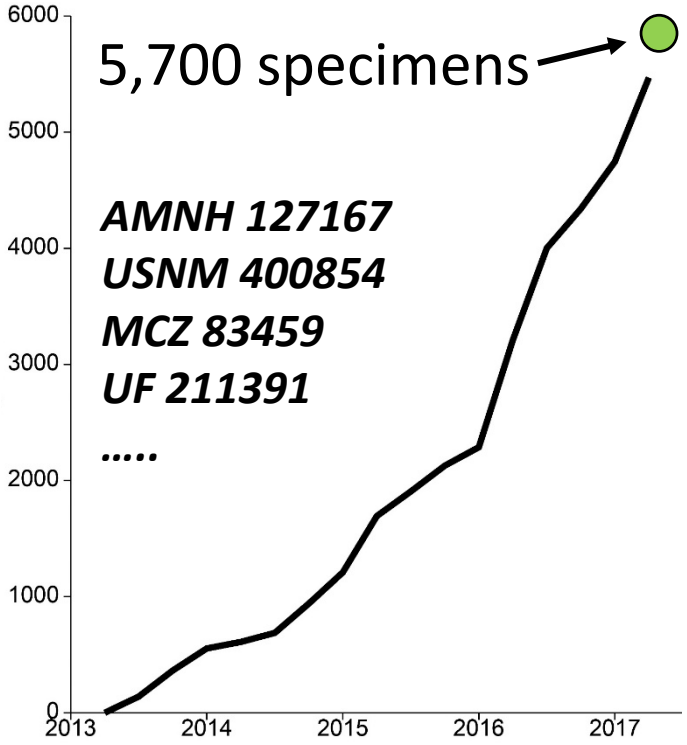


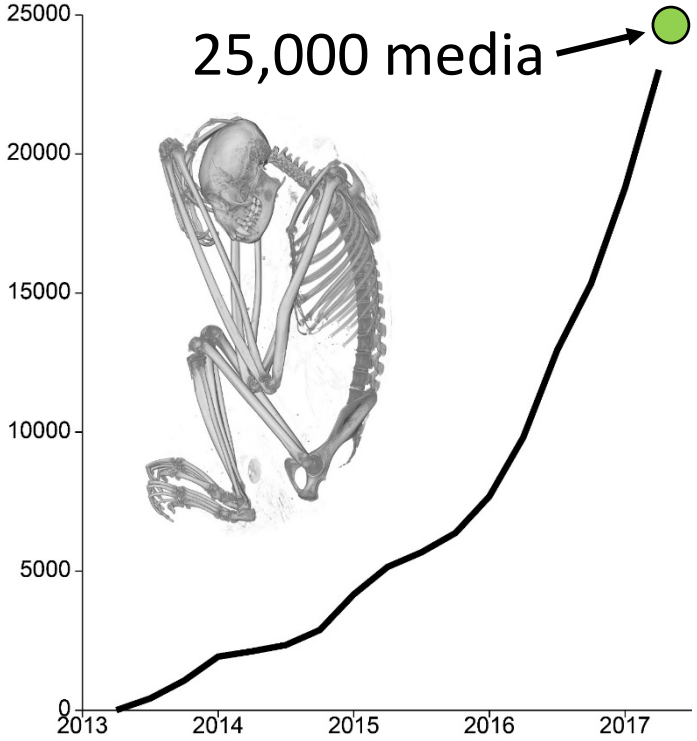


Is MorphoSource being used?

Growth of holdings since 2013 (quarterly data)



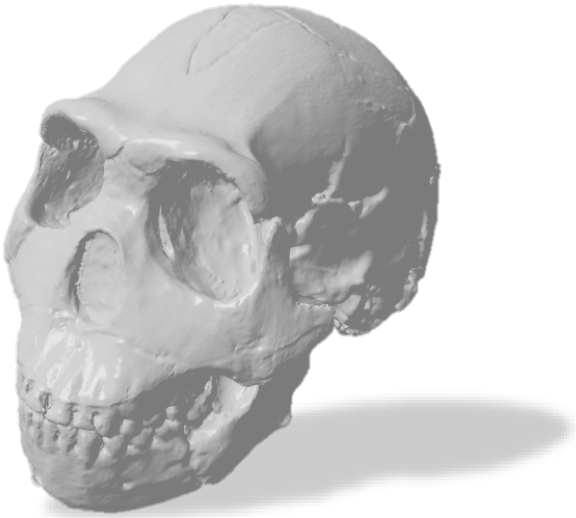




MORPHO SOURCE

Outline

- Why do researchers need a repository for 3D data?
- How do researchers interact with MorphoSource?
- **Sustainability and future growth**



Homo naledi

Growth

Storage

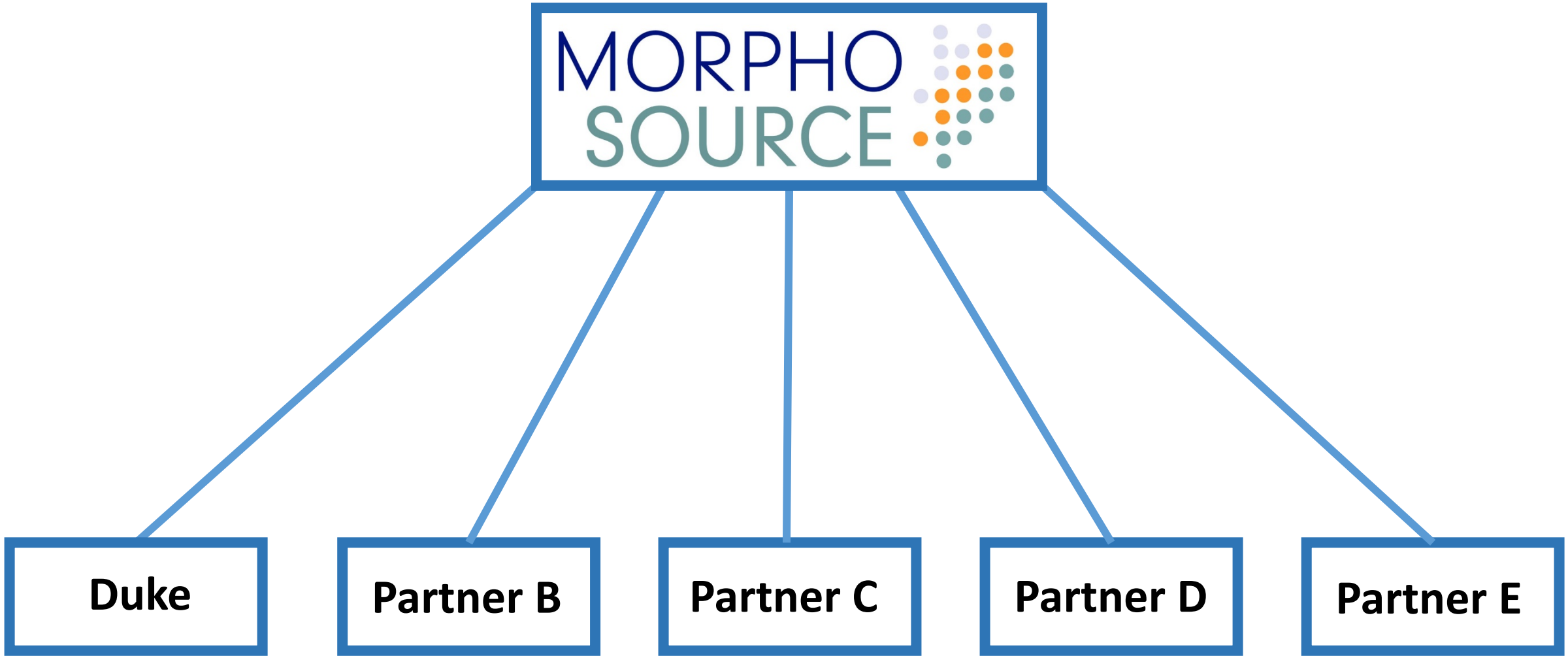
- MorphoSource occupies 15Tb with 25,000 datasets
- Cleared for 10Tb/year growth by Duke
- But one project alone (oVert) could add 30Tb in 5 yrs
- Plan???

Growth

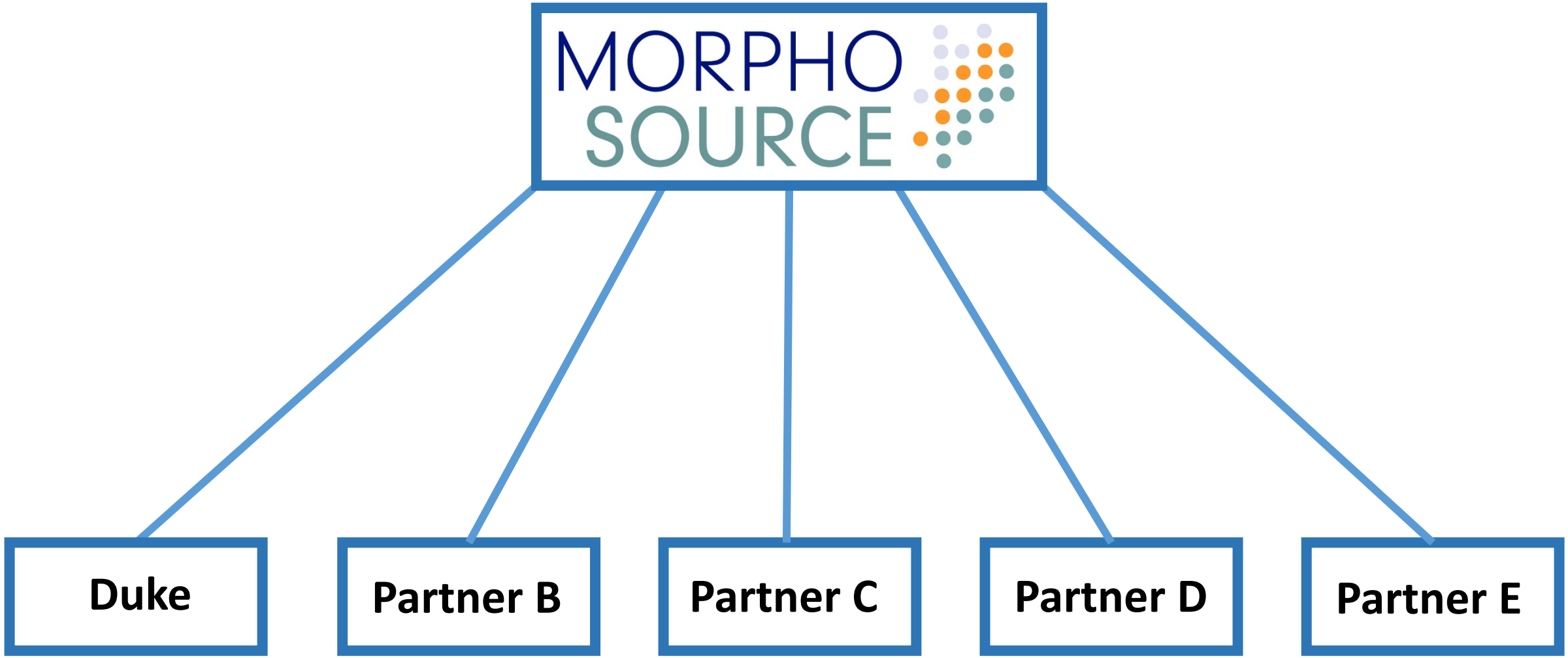
Many hands make light work

Grow through addition of consortium members who provide a finite storage commitment

Growth



How?



How?

Leverage state-of-the-art open access digital repository platforms

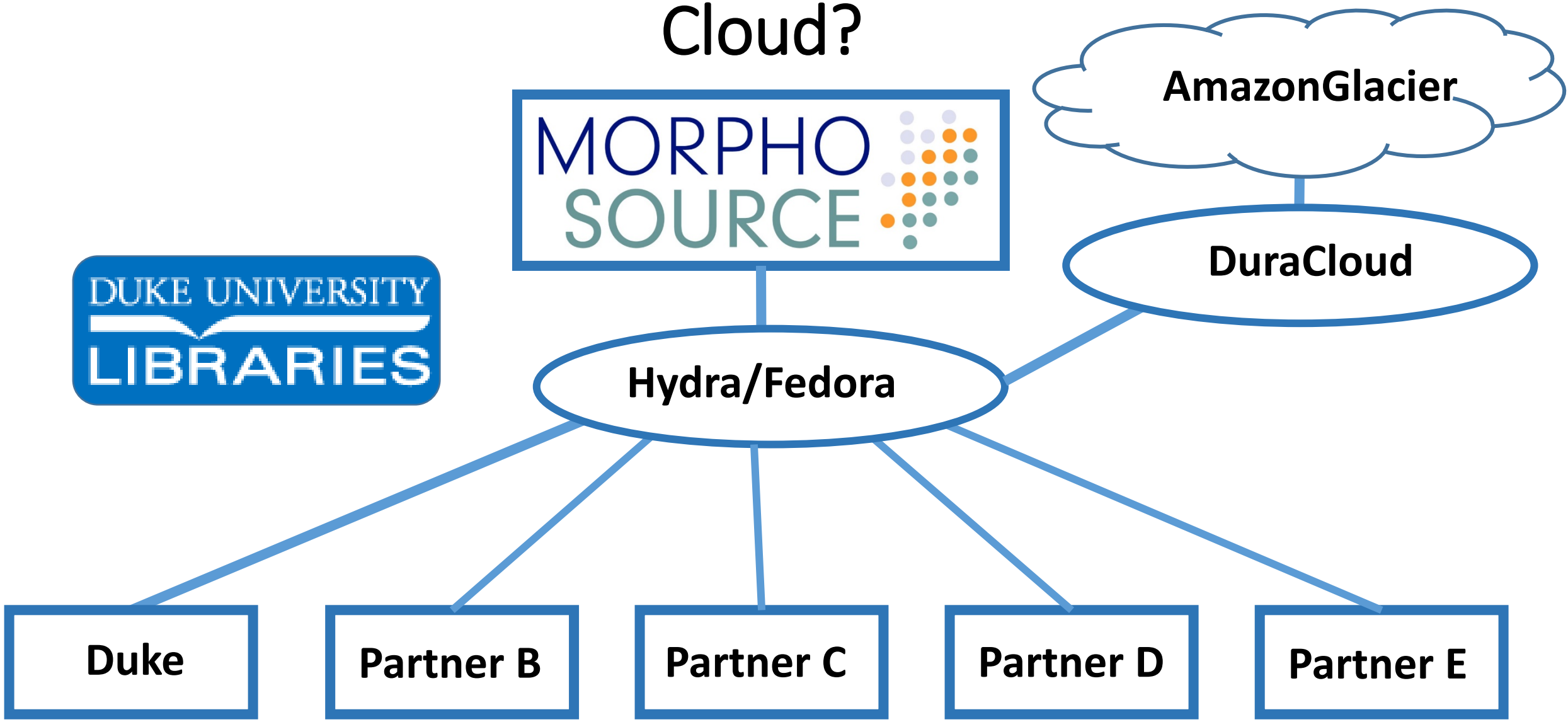
Duke | **Partner B** | **Partner C** | **Partner D** | **Partner E**

How?

Leverage state-of-the-art open access digital repository platforms

Duke	Partner B	Partner C	Partner D	Partner E

Cloud?


AmazonGlacier
MORPHO SOURCE
DuraCloud
DUKE UNIVERSITY LIBRARIES
Hydra/Fedora
Duke
Partner B
Partner C
Partner D
Partner E

The open source parts and their functions

Fedora (Fedorarepository.org)

- Digital Asset Management platform
- Allows integration of multiple data nodes

Hydra (projecthydra.org)

- Provides the database structure for interacting with Fedora

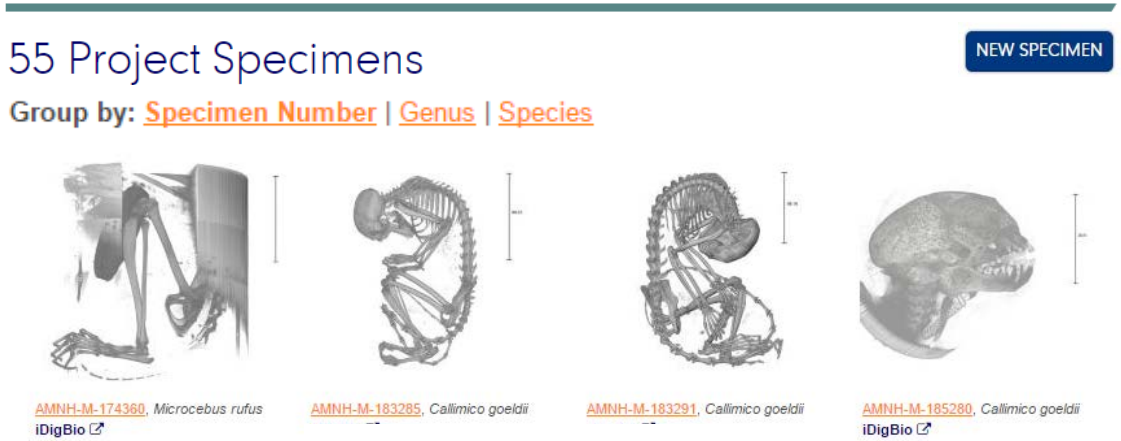
DuraCloud (duracloud.org)

- Provides bitrot prevention
- Manages full redundant cloud copy through Amazon Glacier
- Provides integration with Archivematica

MORPHO SOURCE

Summary

- **3D data are now vital** and maximizing their accessibility and discoverability will be transformative
- **MorphoSource archives 3D data on vouchered specimens** for any researcher and serves it in a setting that emulates a museum collection.
- **An integrated effort is needed** to manage preservation of 3D digital data from museums.



Acknowledgments

- Gil Nelson, iDigBio, conference organizers

For support & funding of MorphoSource Development

- Duke University Trinity College of Arts & Sciences (major funder so far)
- Duke Shared Materials Instrumentation Facility
- Duke Biology IT Center

For discussion leading to development of concepts

- Jukka Jernvall, Alistair Evans, Gudrun Evans, David Blackburn, Edward Stanley, James Thostenson, Jocelyn Sessa, Maureen O'Leary, Mark Uhen, Gabriel Yapuncich

For work developing and populating MorphoSource

- Maria Passarotti (Whirl-i-gig, Inc.)
- Alex Thompson, Kevin Love, Dan Stoner (iDigBio)
- Technicians & students: Mackenzie Nieto-Aguilar, Darbi Griffith, Ana Galvez, Miles Schaeffer, Mercedes Zapata-Garcia, Shane Daly, Sunghoon Liu, Ksenia Sokolova, Anne Driscoll, Kevin Vo, Annie Lott, Callie Crawford, and many more