NATURAL HISTORY MUSEUM

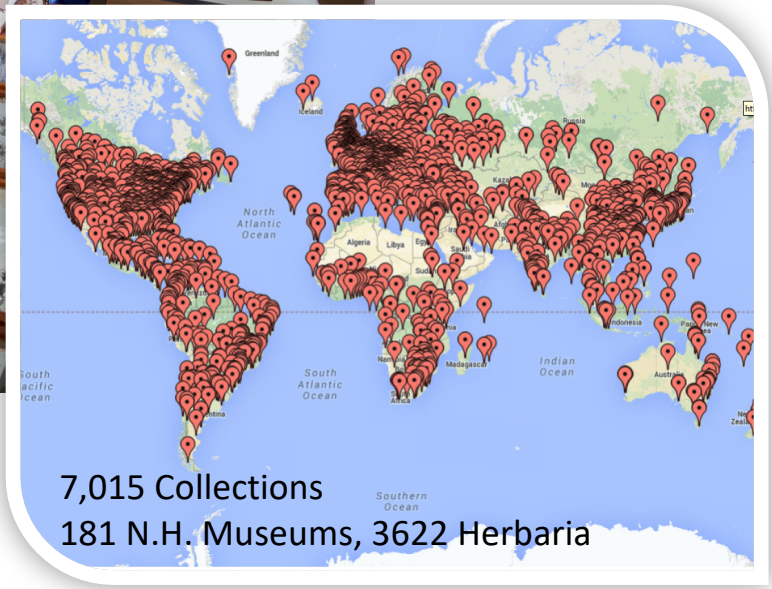
Digitisation, Open Data & Digital Partnerships

An overview for the Parliamentary and Scientific Committee

16 July 2018

NATURAL HISTORY MUSEUM

Collections



Data from http://grbio.org/

- Circa 80m specimens
- Globally c.1.5-3B
- Diverse & unique
- Digital Collections Programme (2014-2024)
- 5% with a digital record (from 2.5%)
- Part of several global initiatives

NATURAL HISTORY MUSEUM

Challenges

1. Mass Digitisation

Specimen image

2. Data Extraction

Transcription
(e.g. Zooniverse)

OCR
(e.g. reCAPTCHA)

Georeferencing
(Online mapping tools)

Image Recognition
(e.g. Image Search)

3. Using the Data

Linking to Archives and Literature
(e.g. Biodiversity Heritage Library)

Analytical Tools
(e.g. OpenRefine)

Data Visualisation
(e.g. CartoDB)

Search
(e.g. WikiData)

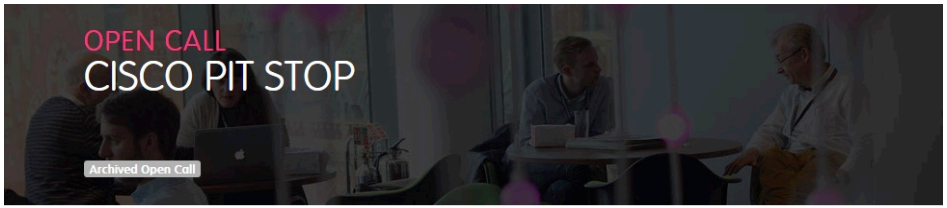
4. End Products

NHM in an App

Digital Exhibitions
(Virtual Tours)

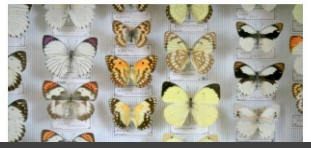
NATURAL HISTORY MUSEUM

Industry

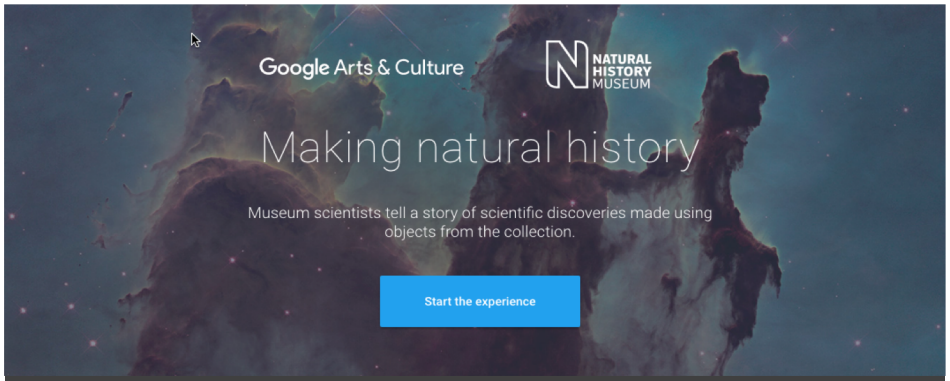


Cisco Pit Stop: Digitising the Natural History Museum's collections

The Digital Catapult wants to hear from any small or medium sized businesses interested in developing innovative approaches to digitising scientific collections for our Pit Stop with Cisco and the Natural History Museum London happening in February.



Cisco / Digital Catapult



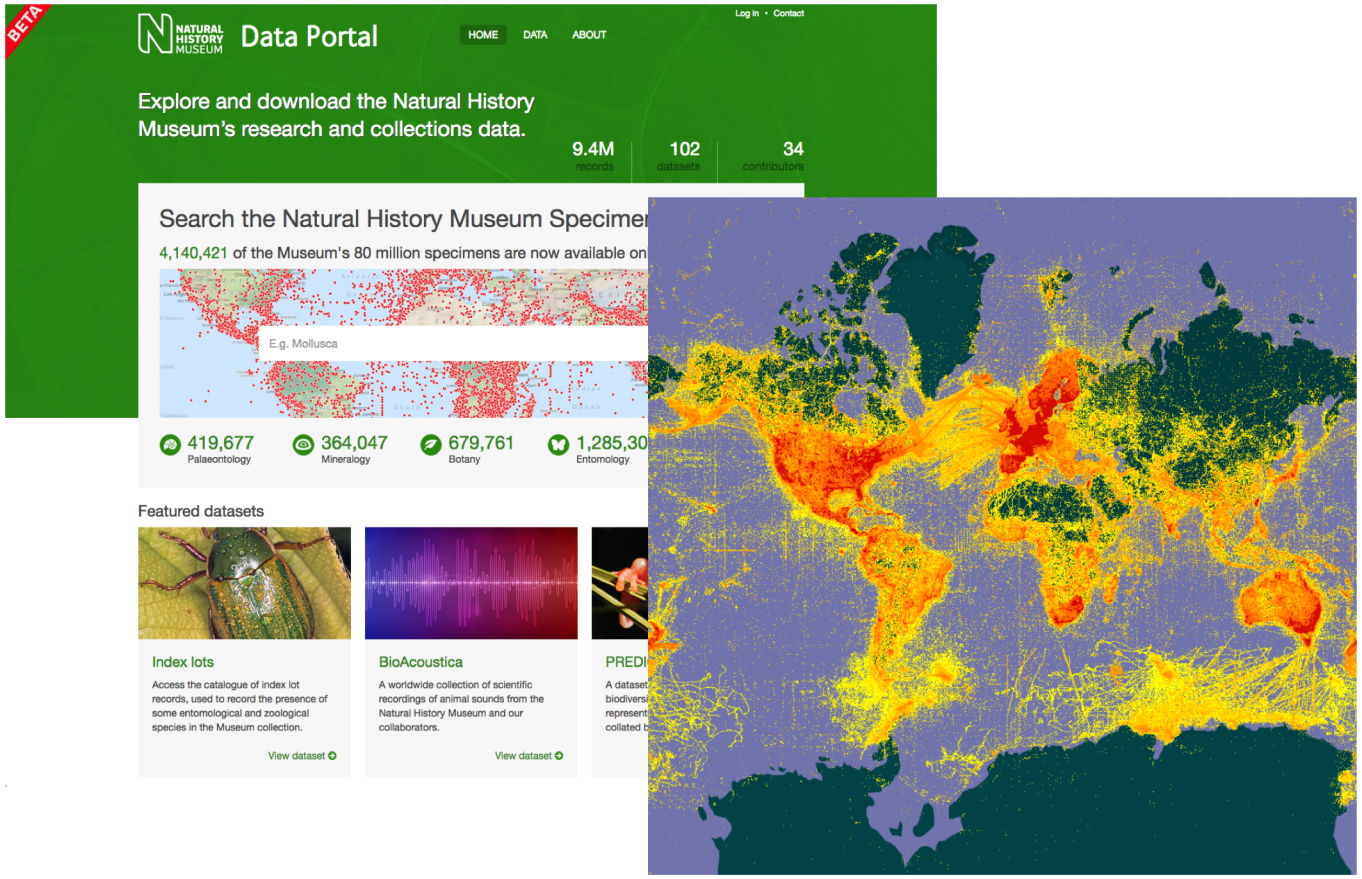
Google

Canon

Picturae

NATURAL HISTORY MUSEUM

Data Portal

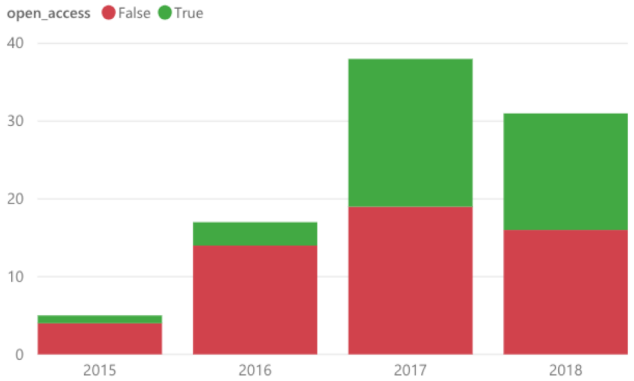


- NHM research & collections data
- Access, reuse & citation
- 9.4m records, 102 datasets since 2015
- Images, sound, video & 3D
- Default open licensing
- 13.42B records downloaded in 162k events, 95 papers

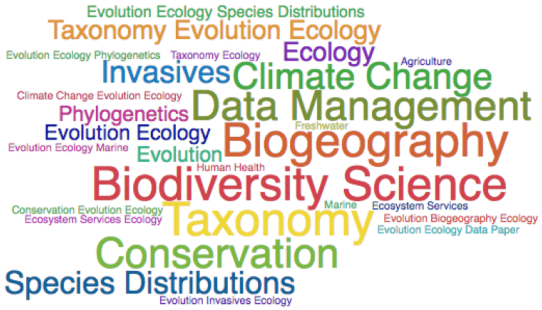
NATURAL HISTORY MUSEUM

Impact

Publications citing NHM data (GBIF only)



GBIF topic tags by frequency



91	81	38
Publications	Peer-reviewed	Open Access

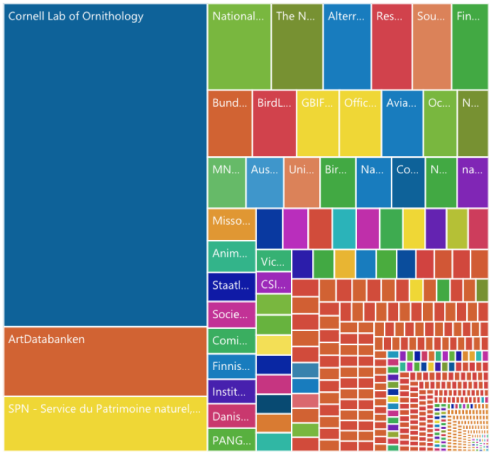
1,007.7M	3.7M	0.37%	2.54%
GBIF total occurrences	NHM occurrences	NHM % of total GBIF	NHM % preserved sp.

7.4bn	21M	0.29%
GBIF occ. downloaded	NHM downloads	NHM % of downloads

Data provider: All

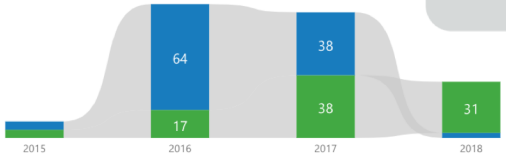
Source title: All

Occurrence count by data provider



Title	Records	NHM records	NHM record %
The collection of birds from Mozambique at t...	10,654	5,101	47.88 %
Spatial quantification of the world population ...	1,442	234	16.23 %
Digitization of museum collections holds the ...	80	10	12.50 %
First record of the Fitzingers False Coral Snake...	51	6	11.76 %
On the absence of the Green-tailed Trainbear...	947	95	10.03 %
Genus Glyphonycteris Thomas, 1896 (Mamma...	39	2	5.13 %
A New Species of Bachia Gray 1845 (Squamat...	1,447	59	4.08 %
Habitat diversity predicts orchid diversity in th...	14,205	494	3.48 %
Occurrence of Glaucus atlanticus in the Midrif...	31	1	3.23 %
Species Distribution Modeling of Deep Pelagi...	24,578	769	3.13 %
Diversity among peripheral populations: gene...	538	16	2.97 %

Papers + downstream citations by year of initial publication



6

H-index

Altmetric

iCollections

97

Phthiraptera collection

46

3D Cetacean Scanning

16

NATURAL HISTORY MUSEUM

Science

International Journal of Epidemiology, 2017, 1–10
doi: 10.1093/ije/dyw366
Original article

Original article

Spatial quantification of the world population potentially exposed to Zika virus

Alberto J. Alaniz,[1,2*] Antonella Bacigalupo[3] and Pedro E. Cattan[3]

Journal of Ecology — BRITISH ECOLOGICAL SOCIETY

Research Article

Climate change increases ecogeographic isolation between closely related plants

Karl J. Duffy, Hans Jacquemyn

First published: 22 June 2018 | https:/

Journal of Ecology — BRITISH ECOLOGICAL SOCIETY

RESEARCH ARTICLE

Effect of pollination strategy, phylogeny and distribution on pollination niches of Euro-Mediterranean orchids

Nina Joffard, François Massol, Matthias Grenié, Claudine Montgelard, Bertrand Schatz

First published: 29 May 2018 | https://doi.o

Journal of Ecology — BRITISH ECOLOGICAL SOCIETY

Standard Paper

Urban warming favours C_4 plants in temperate European cities

Grant A. Duffy, Steven L. Chown

First published: 11 August 2016 | https://doi.org/10.1111/1365-2745.12652 | Cited by: 1

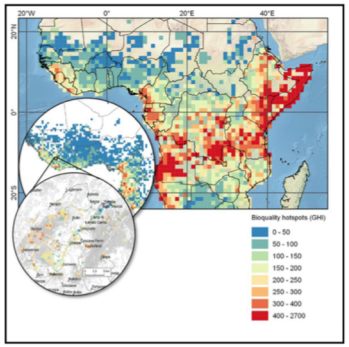
Read the full text › PDF TOOLS SHARE

Report

Current Biology

Bioquality Hotspots in the Tropical African Flora

Graphical Abstract



Authors
Cicely A.M. Marshall, Jan J. Wieringa, William D. Hawthorne

Correspondence
cicely.marshall@plants.ox.ac.uk

In Brief
Marshall et al. introduce a new conservation framework for tropical Africa. The authors use "big data" to integrate species-level conservation assessments into reliable minimum local estimates of global irreplaceability across the region, providing a framework for conservationists and researchers applicable at the local scale.

Conservation Biology

Contributed Paper

Effectiveness of protected areas for vertebrates based on taxonomic and phylogenetic diversity

Qing Quan,[1] Xianli Che,[1] Yongjie Wu,[2,3] Yuchun Wu,[1] Qiang Zhang,[1] Min Zhang,[1] and Fasheng Zou[1]*

[1]Guangdong Key Laboratory of Animal Conservation and Resource Utilization, Guangdong Public Laboratory of Wild Animal Conservation and Utilization, Guangdong Institute of Applied Biological Resources, Guangzhou 510260, China
[2]Key Laboratory of Bio-resources and Eco-environment of Ministry of Education, College of Life Sciences, Sichuan University, Chengdu 610065, China
[3]Department of Ecology and Evolution, University of Chicago, Chicago, IL 60637, U.S.A.

ECOSPHERE

Forecasting an invasive species' distribution with global distribution data, local data, and physiological information

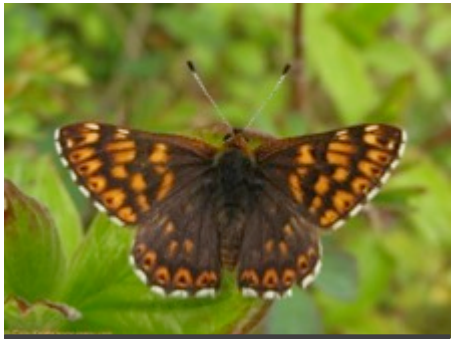
Catherine S. Jarnevich,[1,†] Nicholas E. Young,[2] Marian Talbert,[3] and Colin Talbert[1,3]

[1]U.S. Geological Survey Fort Collins Science Center, 2150 Centre Ave Bldg C, Fort Collins, Colorado 80526 USA
[2]Natural Resource Ecology Laboratory, Colorado State University, Fort Collins, Colorado 80523-1499 USA
[3]Department of Interior, North Central Climate Science Center, Colorado State University, Fort Collins, Colorado 80523 USA

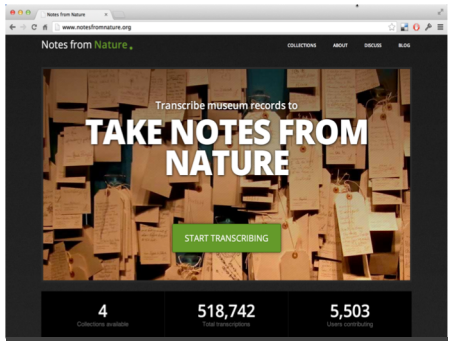
Circa 25% UK users

NATURAL HISTORY MUSEUM

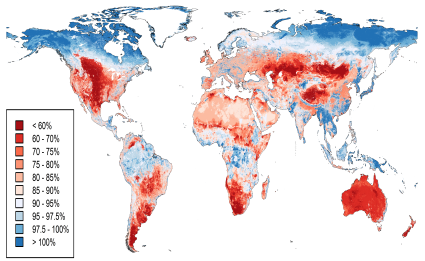
Opportunities



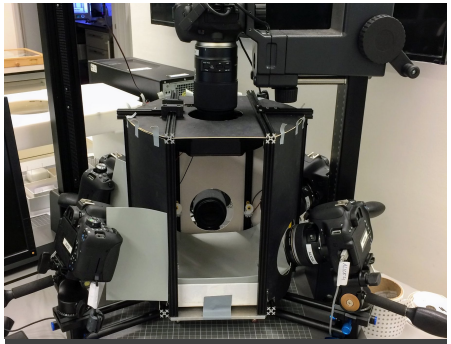
Environmental Change



Citizen science

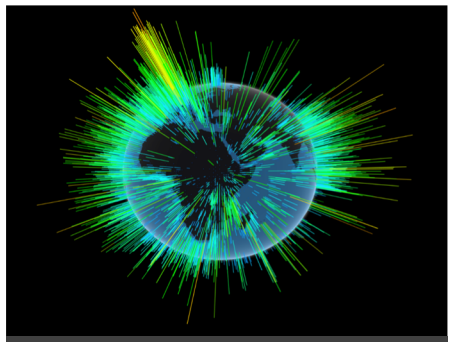


Measuring biodiversity loss

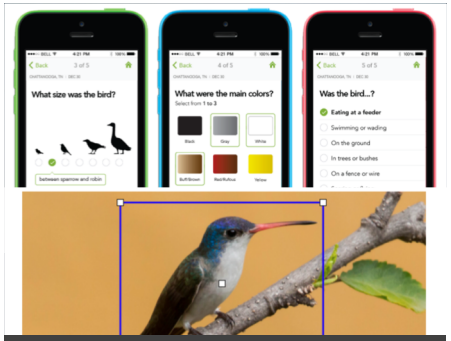


Digitisation workflows

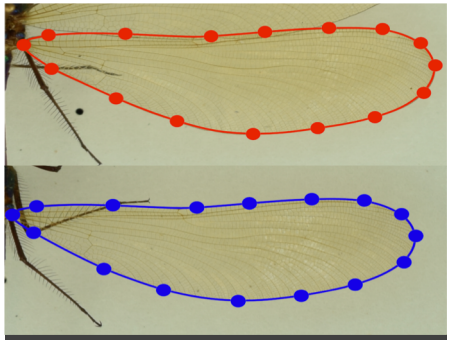
Measuring hyper diversity



Data visualisation



Species recognition



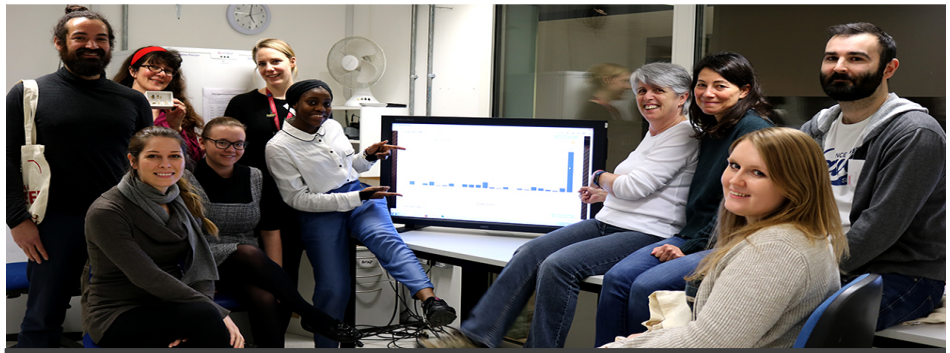
Computer vision

NATURAL HISTORY MUSEUM

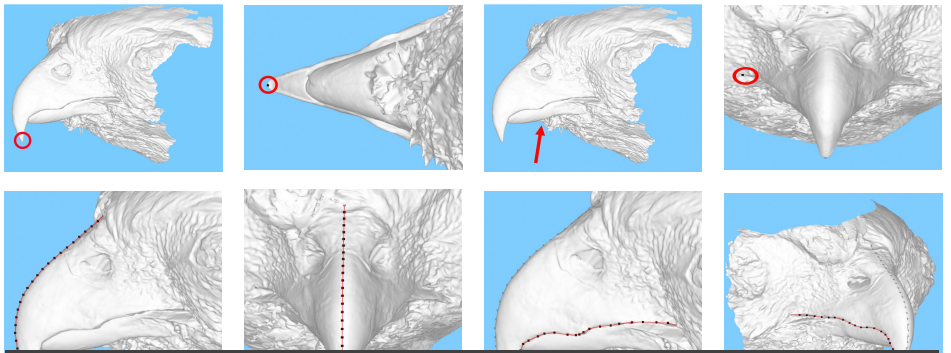
Outreach

Wired

Project Plumage



Visiteering

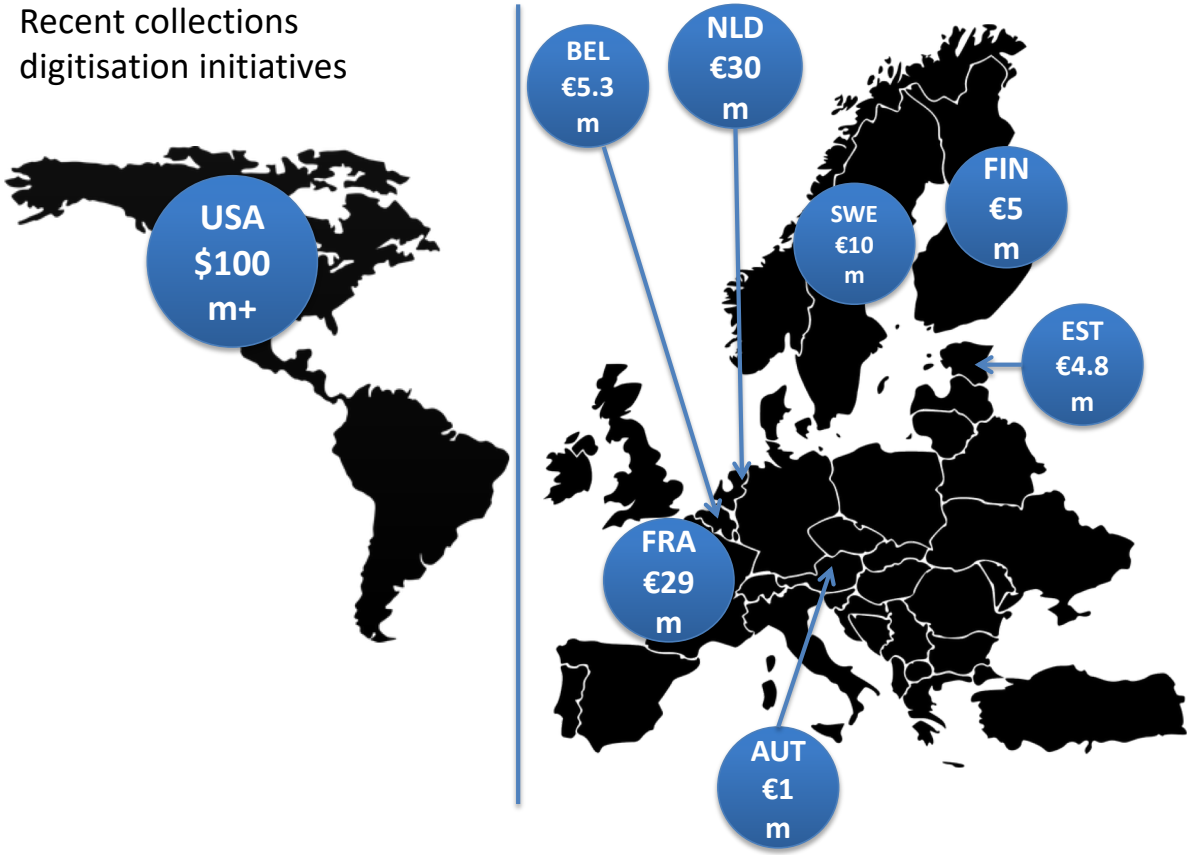


Mark-My-Bird

NATURAL HISTORY MUSEUM

International

Recent collections digitisation initiatives



SYNTHESYS
Synthesis of systematic resources

DISSCO
Distributed System of Scientific Collections

114 National Facilities
21 Countries

ESFRI
European Strategy Forum on Research Infrastructures