

Global Biodiversity Information Facility

Dag Endresen - GBIF Norway

University of Oslo, Natural History Museum

GBIF Global Biodiversity Information Facility

Illustration: GBIF data portal

GBIF data mobilization seminar | 13th March 2023

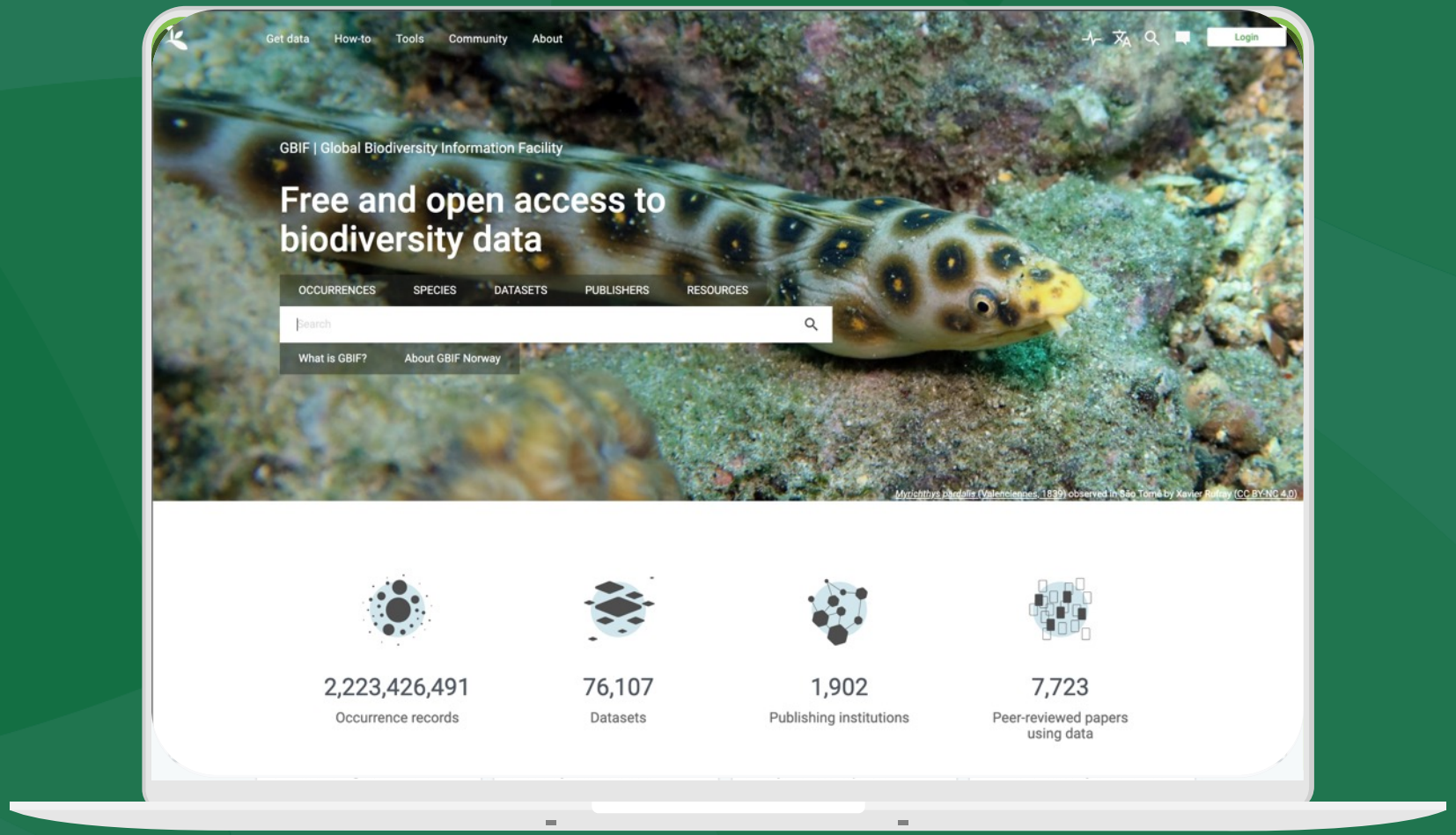
WHAT IS GBIF?

Intergovernmental network and research infrastructure

Provides anyone, anywhere, **free and open access to data** about all types of life on Earth

Voluntary collaboration through Memorandum of Understanding (MoU)

Participant nodes, Secretariat in Copenhagen, Denmark



https://www.gbif.org

GBIF

BY THE NUMBERS | 12TH MARCH 2023

GBIF

Species occurrence records

2 285 643 950

Datasets

83 101

Country Participants **66**

Organizational Participants **43**

Publishers

1 969

Peer-review papers using data

8 420

Average records downloaded per month (2022)

119.7 billion

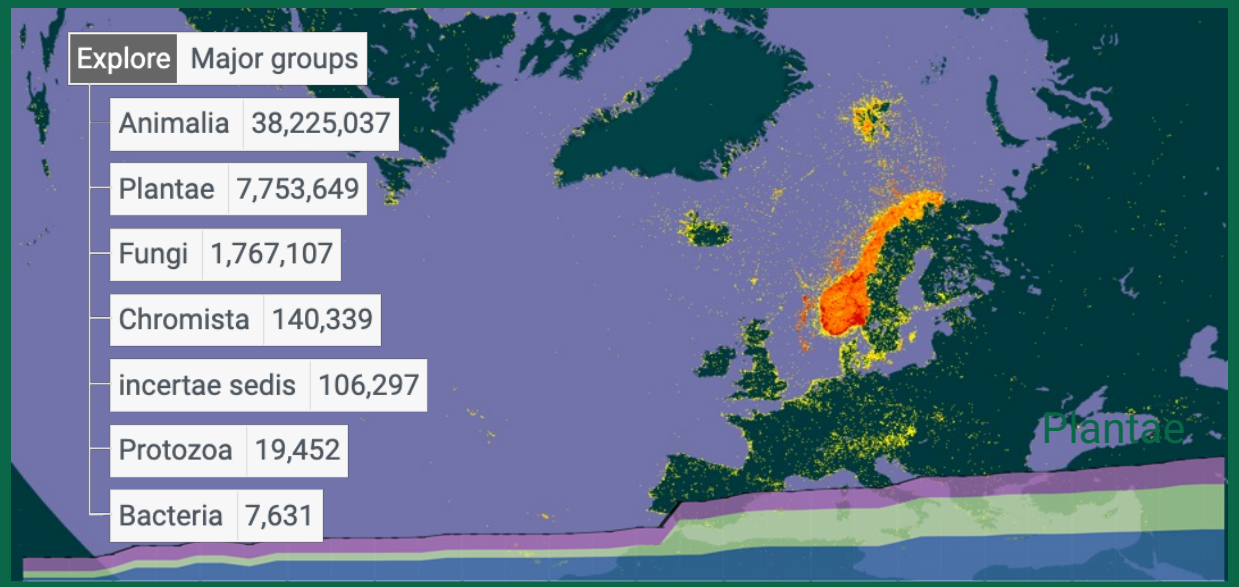
GBIF

BY THE NUMBERS | 12TH MARCH 2023 - NORWAY

GBIF.no

Species occurrence records *(published from)*

49 184 519



Datasets *(published from)*

394

Publishers (from Norway)

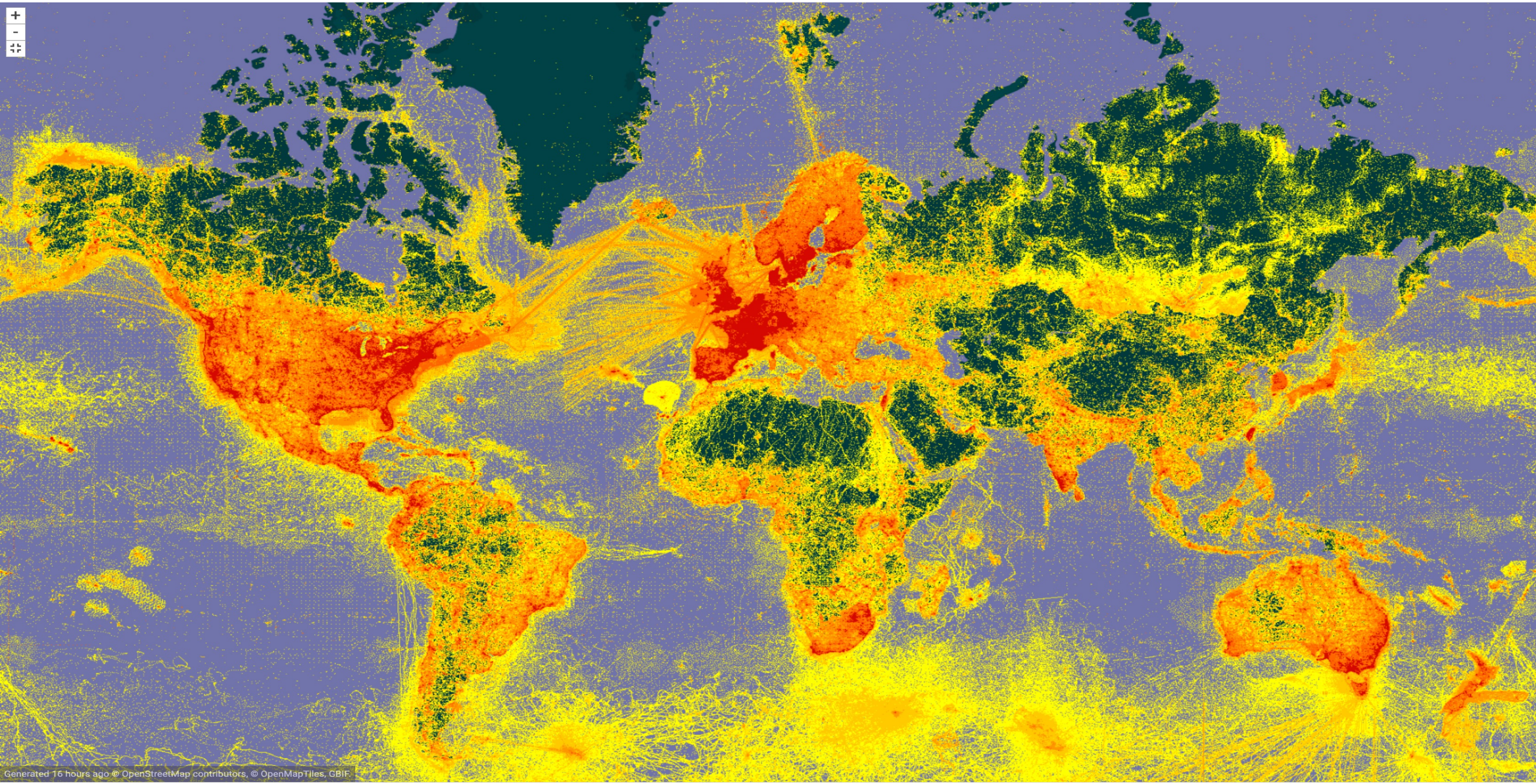
37

Peer-review papers using data (co-author from Norway

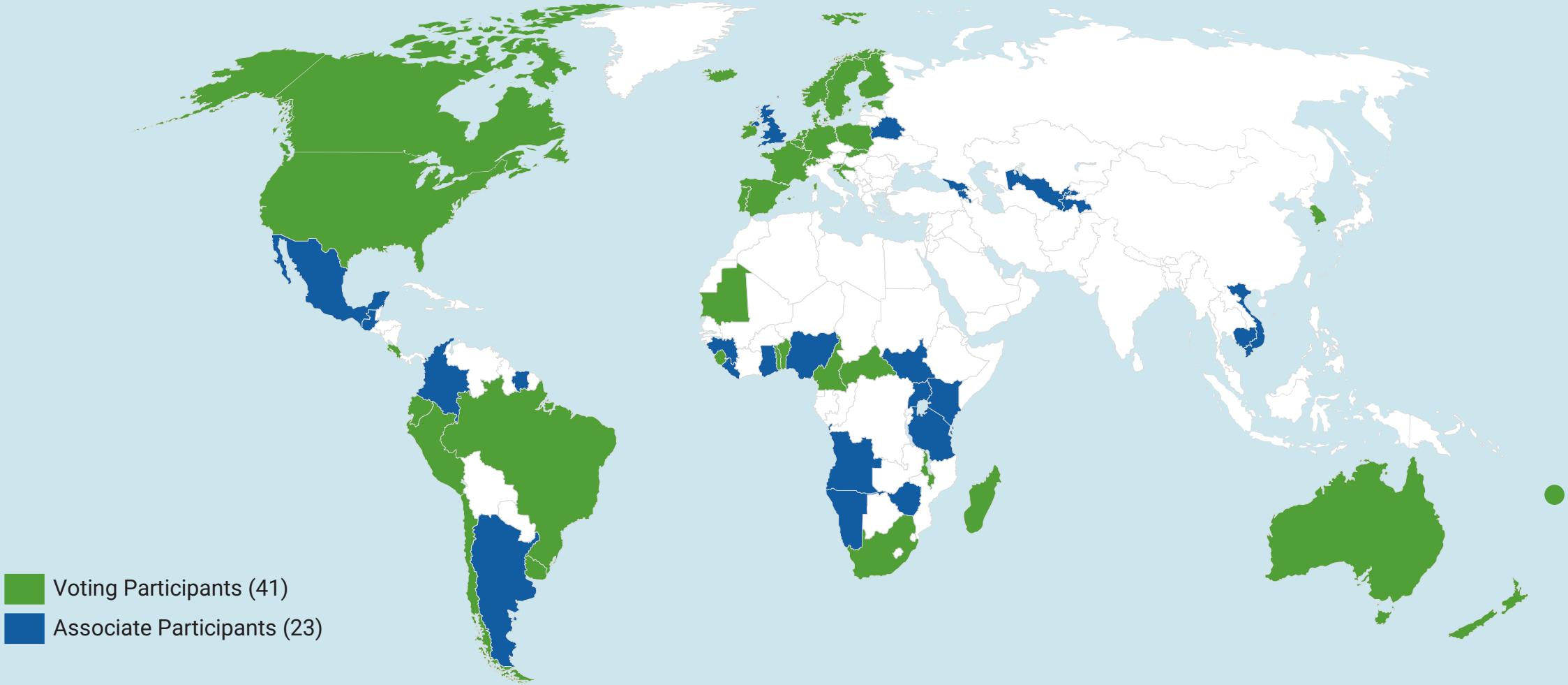
210

GBIF

DATA FROM THE GBIF NETWORK 12th March 2023



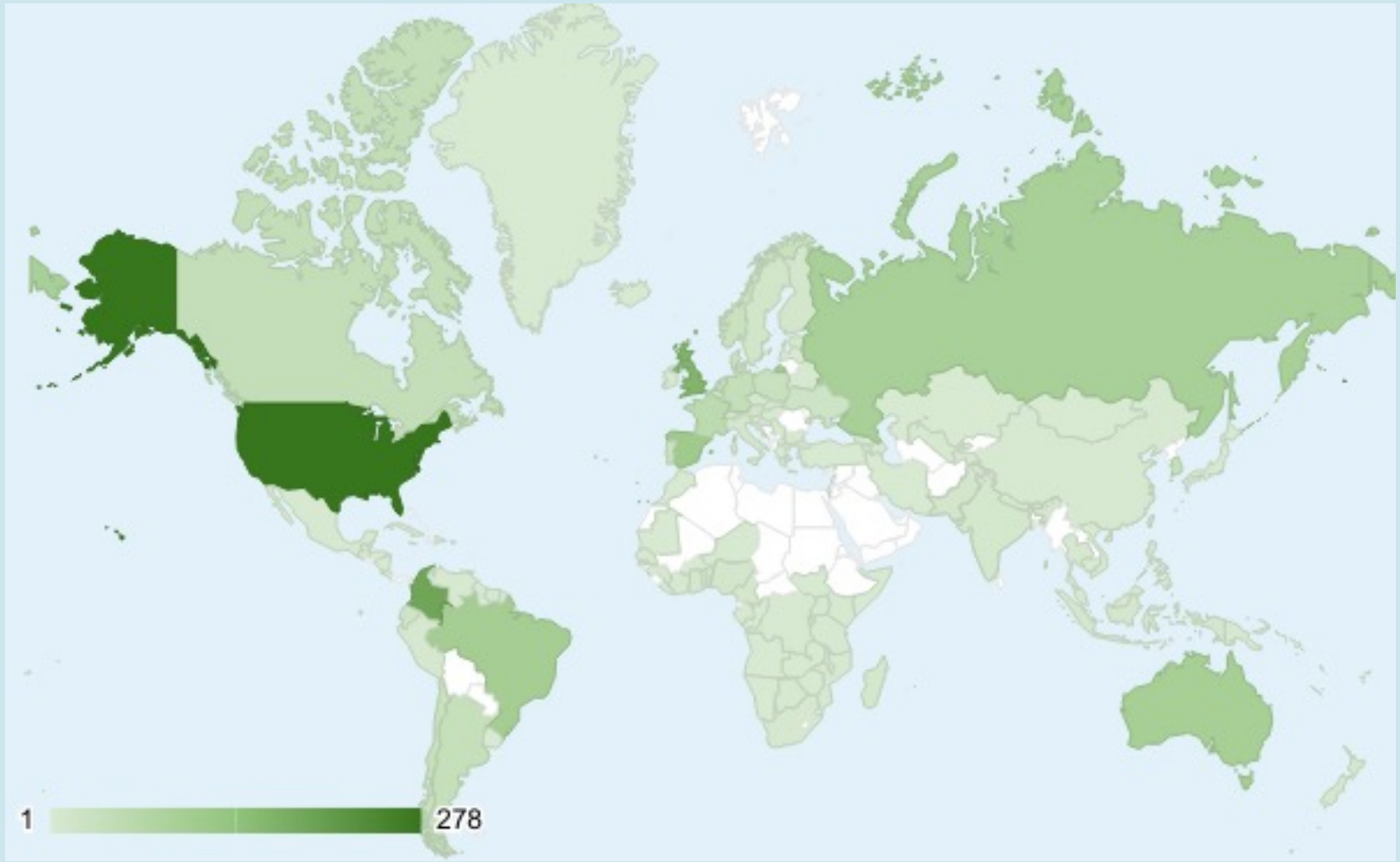
GBIF PARTICIPANT COUNTRIES 12th March 2023



https://www.gbif.org/the-gbif-network

GBIF

GBIF NETWORK OF DATA PUBLISHING INSTITUTIONS 12th March 2023



134 countries/areas with institutions sharing data through GBIF

Top 10 countries: number of data publishers		
1	United States	333
2	Colombia	199
3	United Kingdom	181
4	Spain	128
5	Russian Federation	127
6	Brazil	110
7	Australia	109
8	France	63
9	Canada	56
10	Netherlands	45
--	Norway	37

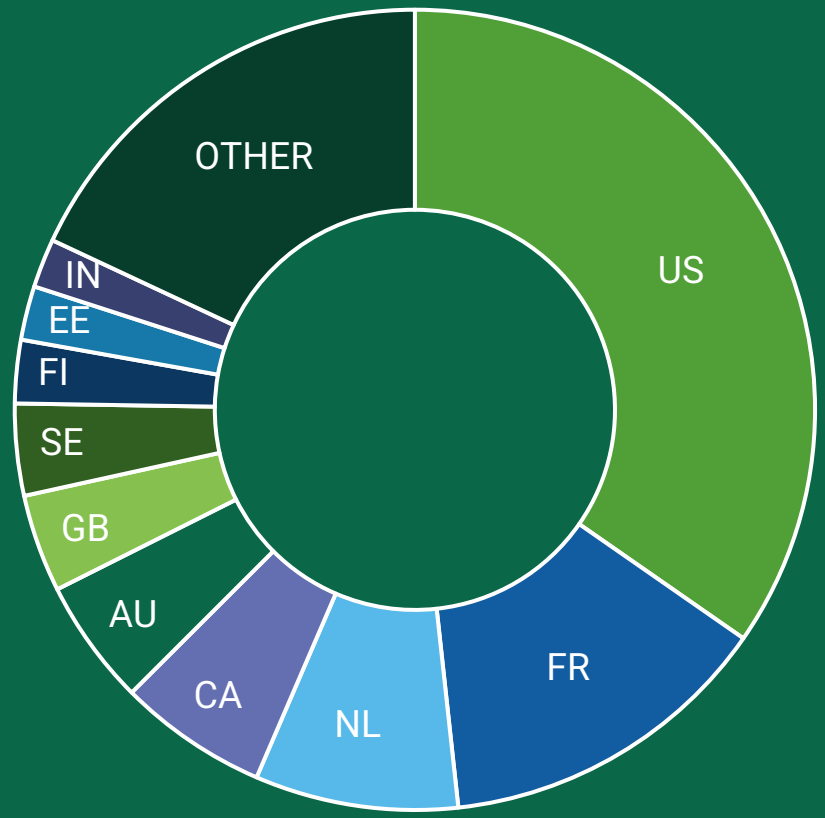
1969 data publishing institutions -- https://www.gbif.org/publisher/search

GBIF

OCCURRENCES PUBLISHED ANNUALLY BY COUNTRY 31st Dec 2022

		2022	2021 total	2021 rank
1	United States	123,152,046	145,520,357	1
2	France	40,243,552	1,229,583	24
3	Netherlands	30,330,743	2,609,622	17
4	Canada	22,328,115	24,595,830	2
5	Australia	18,787,368	6,676,589	4
6	United Kingdom	14,602,384	24,285,291	3
7	Sweden	13,698,583	5,466,063	7
8	Finland	9,447,541	499,507	37
9	Estonia	8,054,503	1,145,145	26
10	India	7,300,428	5,857,419	5
OTHER COUNTRIES + AREAS		66,653,115	54,341,744	
	TOTAL	369,598,387	252,569,971	

Occurrences published in 2022



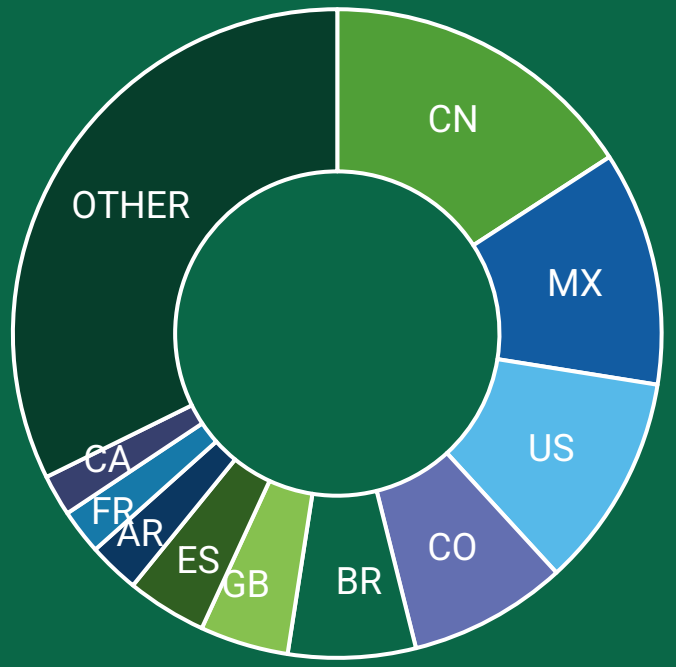
https://www.gbif.org/occurrence/search

GBIF

DATA DOWNLOAD REQUESTS 31st Dec 2022

		2022	2021 total	2021 rank
1	China	36,844	29,819	1
2	Mexico	27,172	21,725	3
3	United States	24,812	24,796	2
4	Colombia	18,392	16,414	4
5	Brazil	14,777	15,156	5
6	United Kingdom	10,256	11,605	6
7	Spain	9,280	9,777	7
8	Argentina	5,938	2,206	15
9	France	5,231	4,725	10
10	Canada	4,782	3,953	12
	OTHER COUNTRIES + AREAS	75,107	72,355	
	TOTAL	232,591	216,157	

Download requests by country, island or territory



156 countries, islands & territories where users have requested downloads (+5)

GBIF

GBIF.org TRAFFIC BY COUNTRY 31ST Dec 2022

Rank	Country/Territory	Visitors	Sessions	Effective visits*	% Effective visits	2021 rank	Pages / session
1	United States	187,349	368,757	193,524	10.32%	1	5.83
2	Colombia	99,317	247,478	124,135	6.62%	4	5.57
3	Mexico	106,410	208,920	99,007	5.28%	2	5.51
4	Brazil	88,111	187,383	100,063	5.34%	3	5.92
5	Spain	76,449	181,399	97,339	5.19%	5	6.86
6	China	41,158	167,951	116,709	6.22%	11	12.91
7	France	53,966	119,079	58,920	3.14%	6	5.58
8	United Kingdom	50,267	115,731	61,222	3.26%	8	6.37
9	Germany	49,357	114,235	57,323	3.06%	7	6.77
10	India	43,633	109,329	55,299	2.95%	9	6.05
	TOTAL	1,514,697	3,555,918	1,875,391	100.00%		6.56

SOURCE: Google Analytics. * Effective visits = total sessions - (total sessions *x* bounce rate)

GBIF

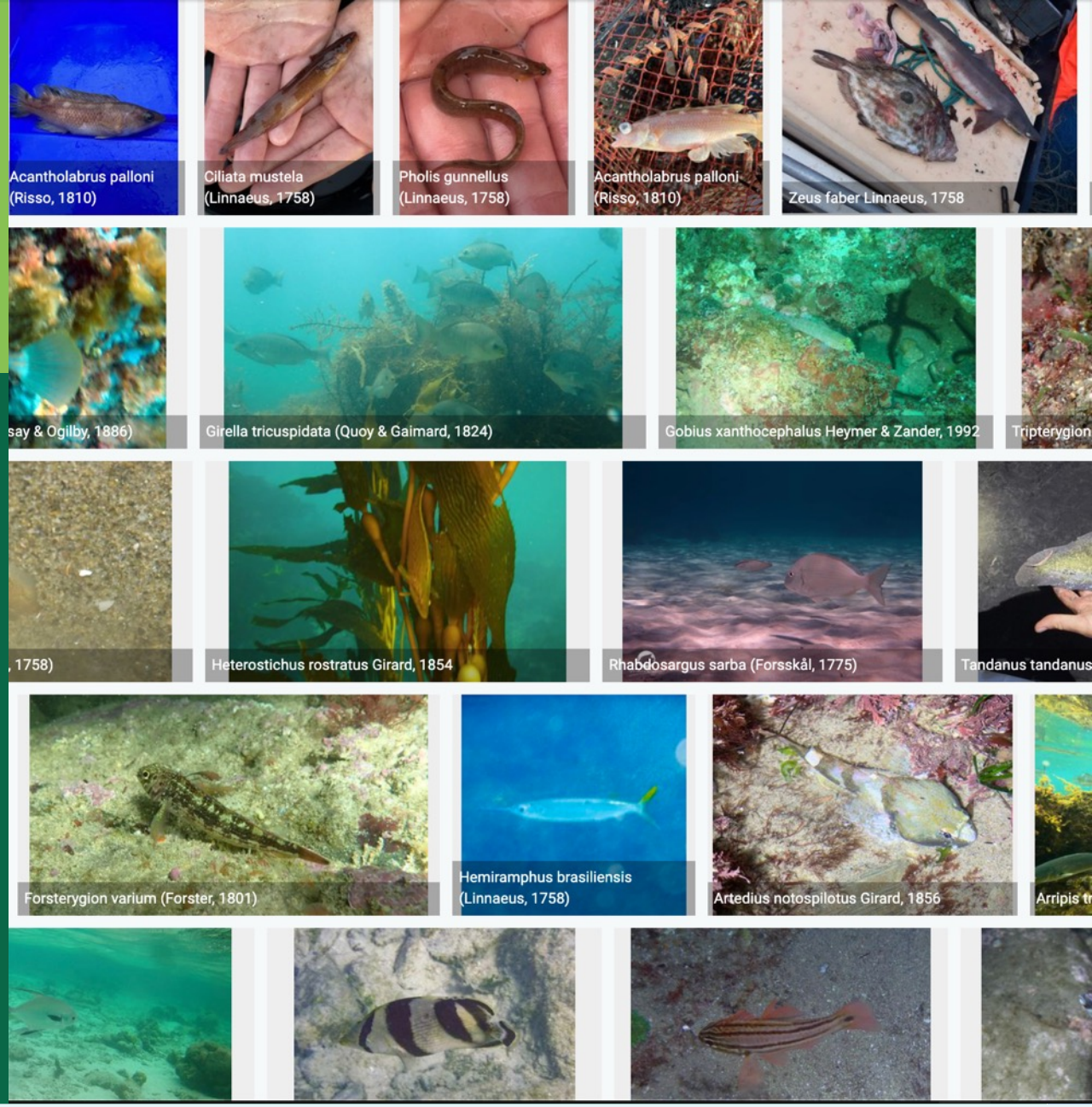
SPECIES OCCURRENCE RECORDS WITH MULTIMEDIA EVIDENCE

Status 12th March 2023

Taxonomically identified multimedia records:

- **133 million images**
- **1 million audio files**
- **4 558 video files**

- 53 million specimens
- 79 million human observations



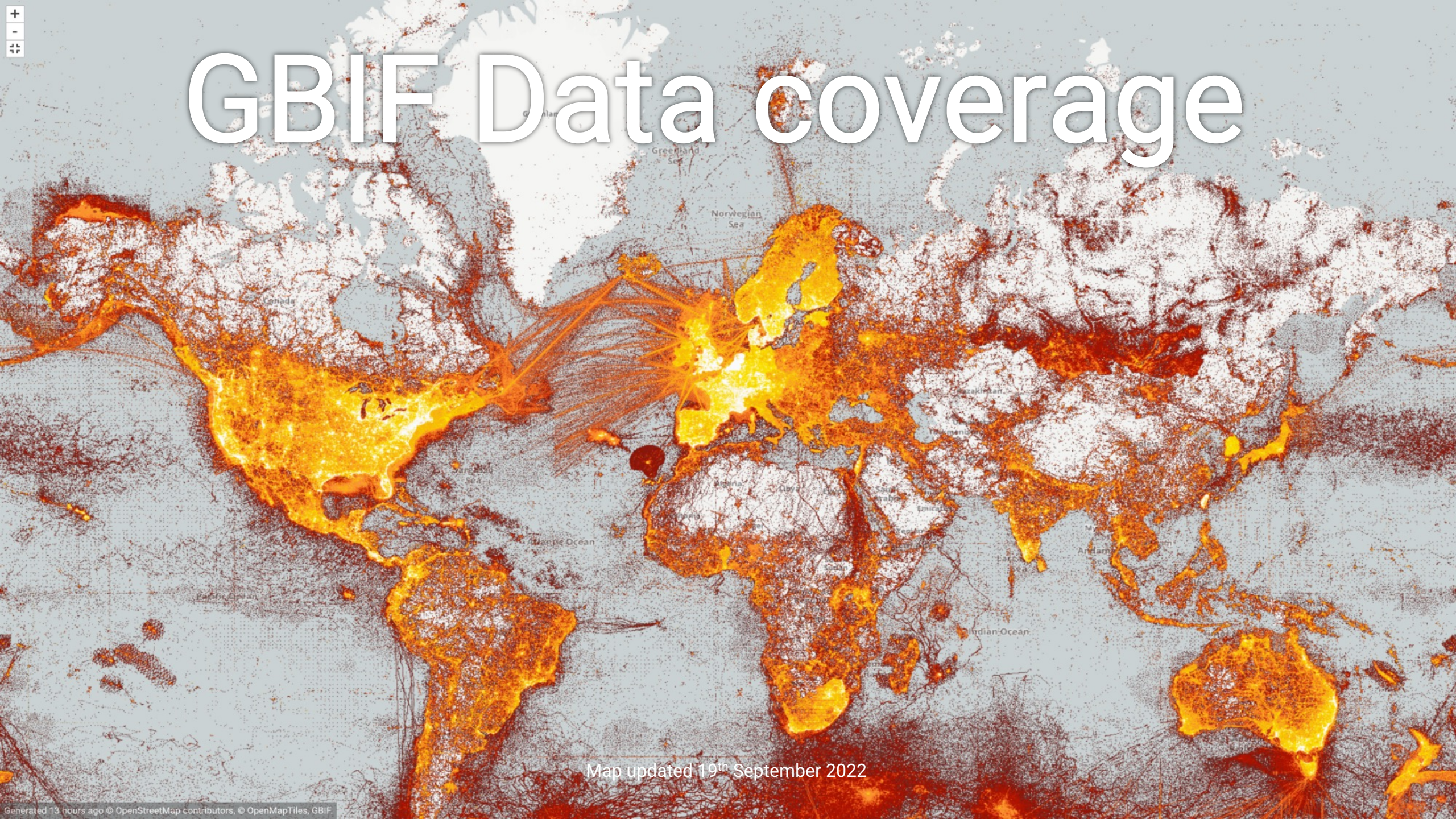
https://www.gbif.org/occurrence/gallery

GBIF



data coverage

GBIF Data coverage



Map updated 19th September 2022

GLOBAL BIODIVERSITY VS. DIGITALLY AVAILABLE DATA

1200 mill. animals

300 m plants

20 m fungi

16 m bacteria

0,04 m virus

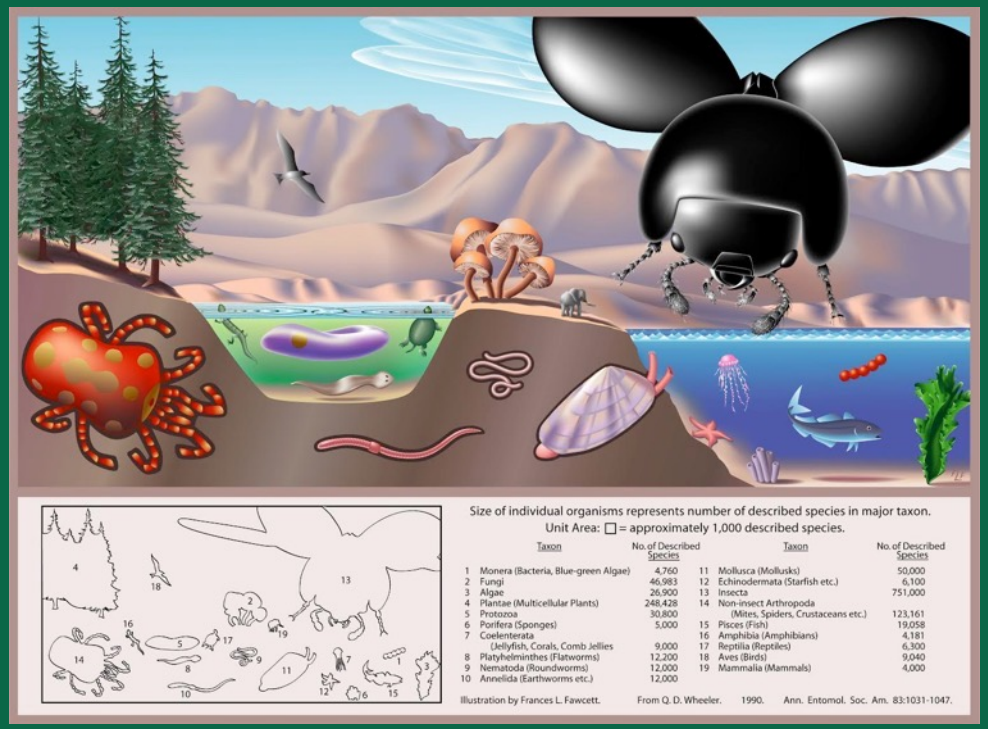
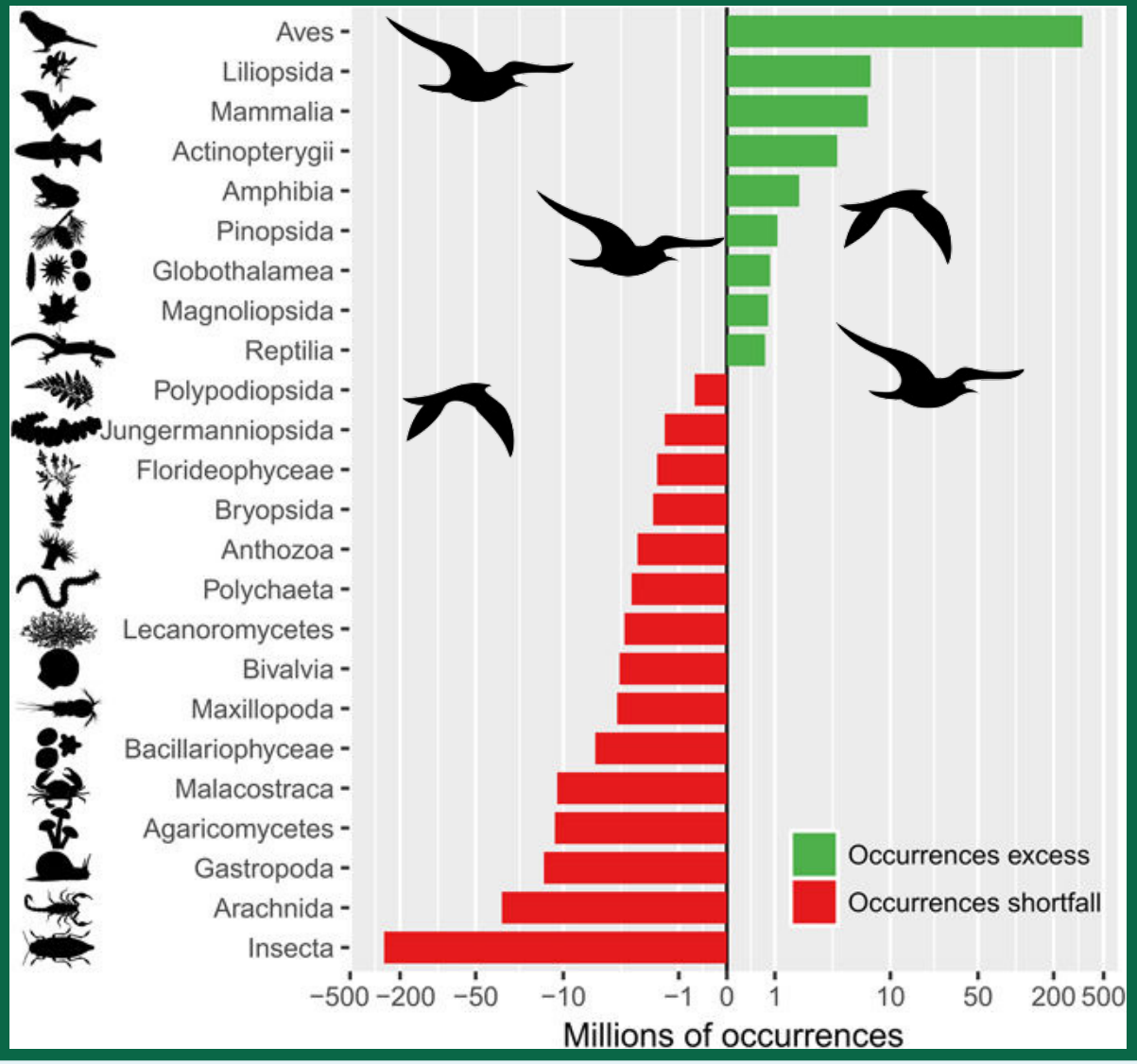


Image: FL Fawcett in Wheller Ann. Entomol. Soc. Am. 1990

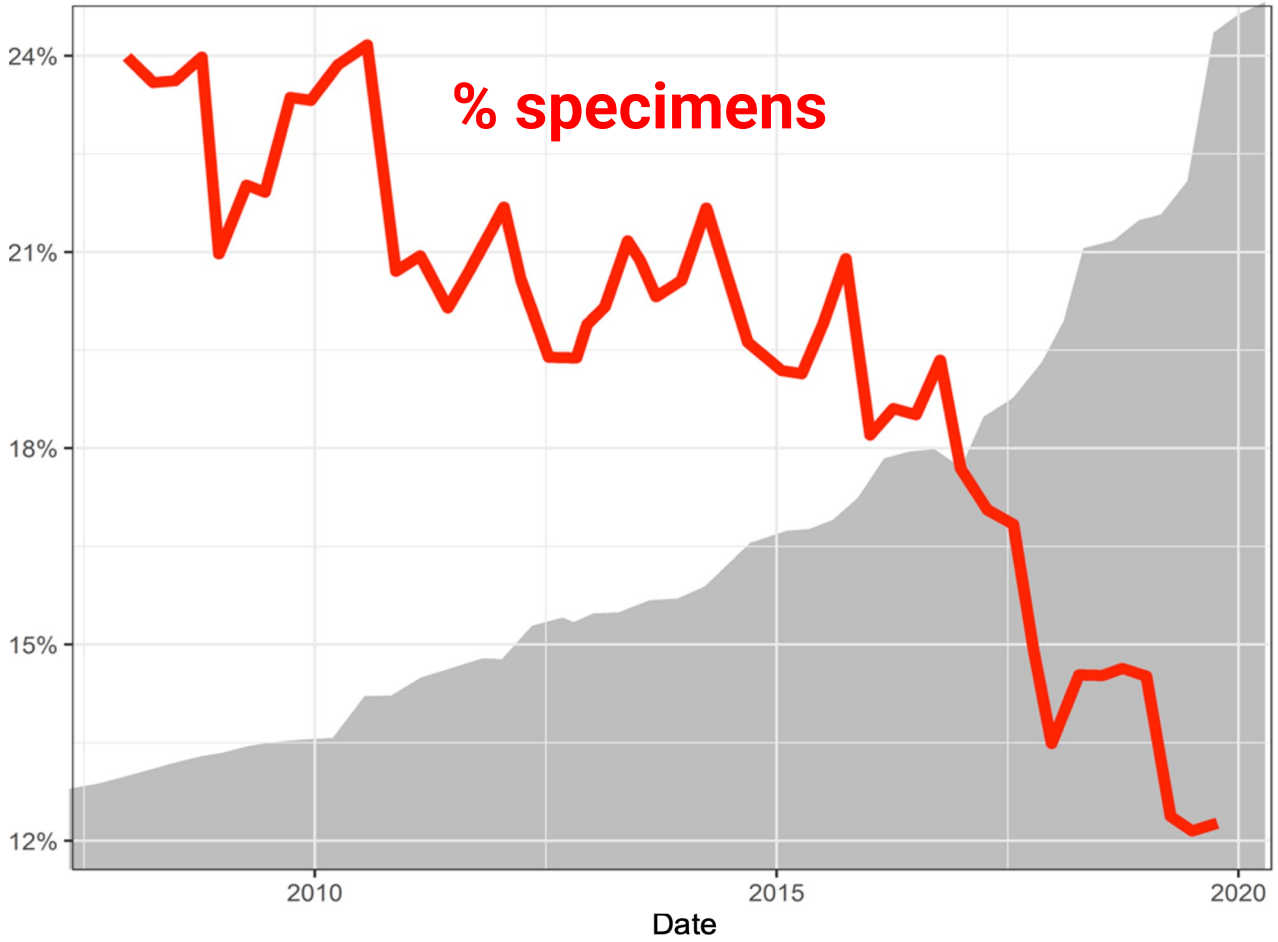


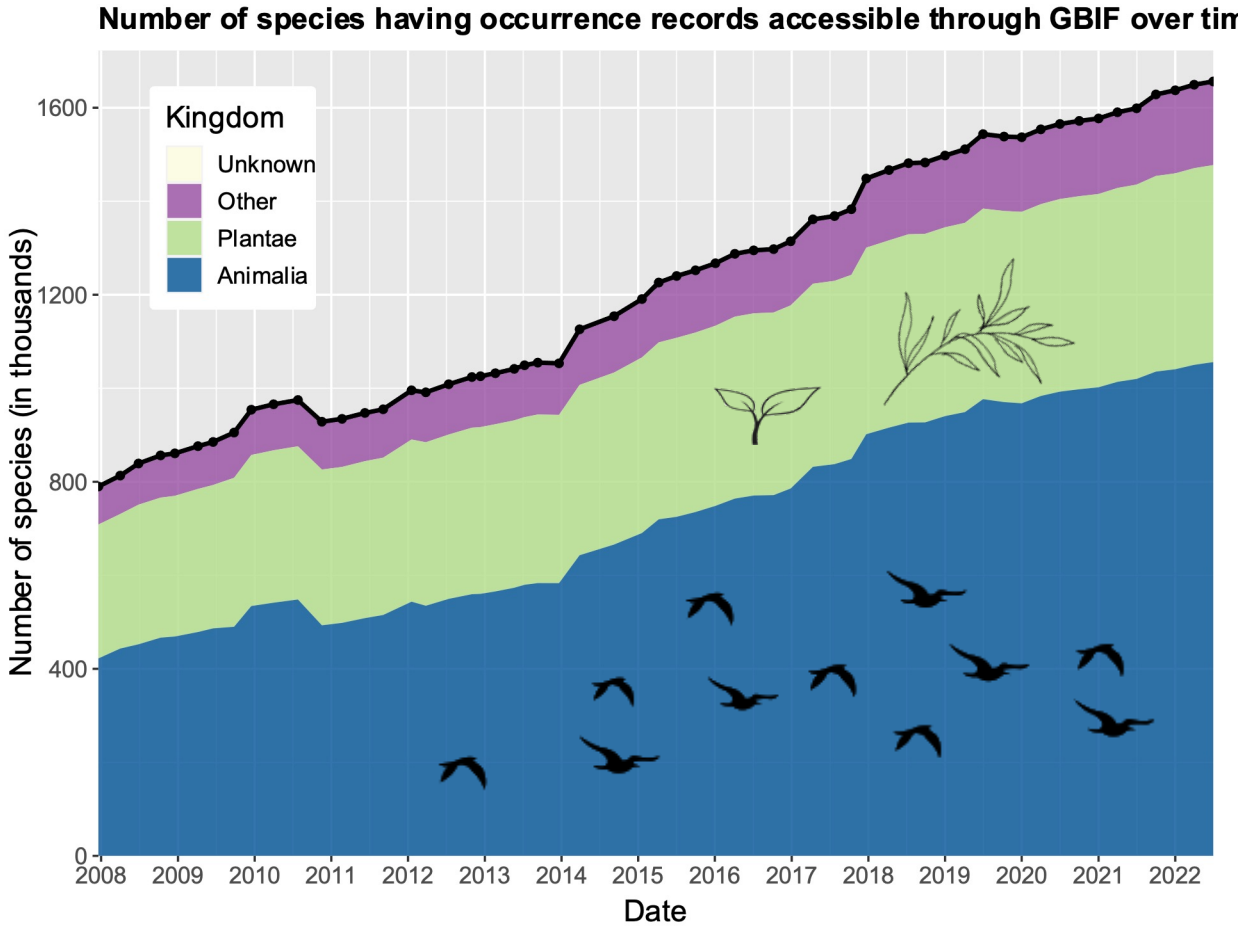
Troudet et al. Nature Scientific Reports 2017

Taxonomic bias towards birds and against insects and small organisms

GBIF

DATA TRENDS ON GBIF.org





https://www.gbif.org/analytics/global



Photo: Botany Collection, Algae, Smithsonian National Museum of Natural History Museum, by Chip Clark.

Very few museum specimens are digitized

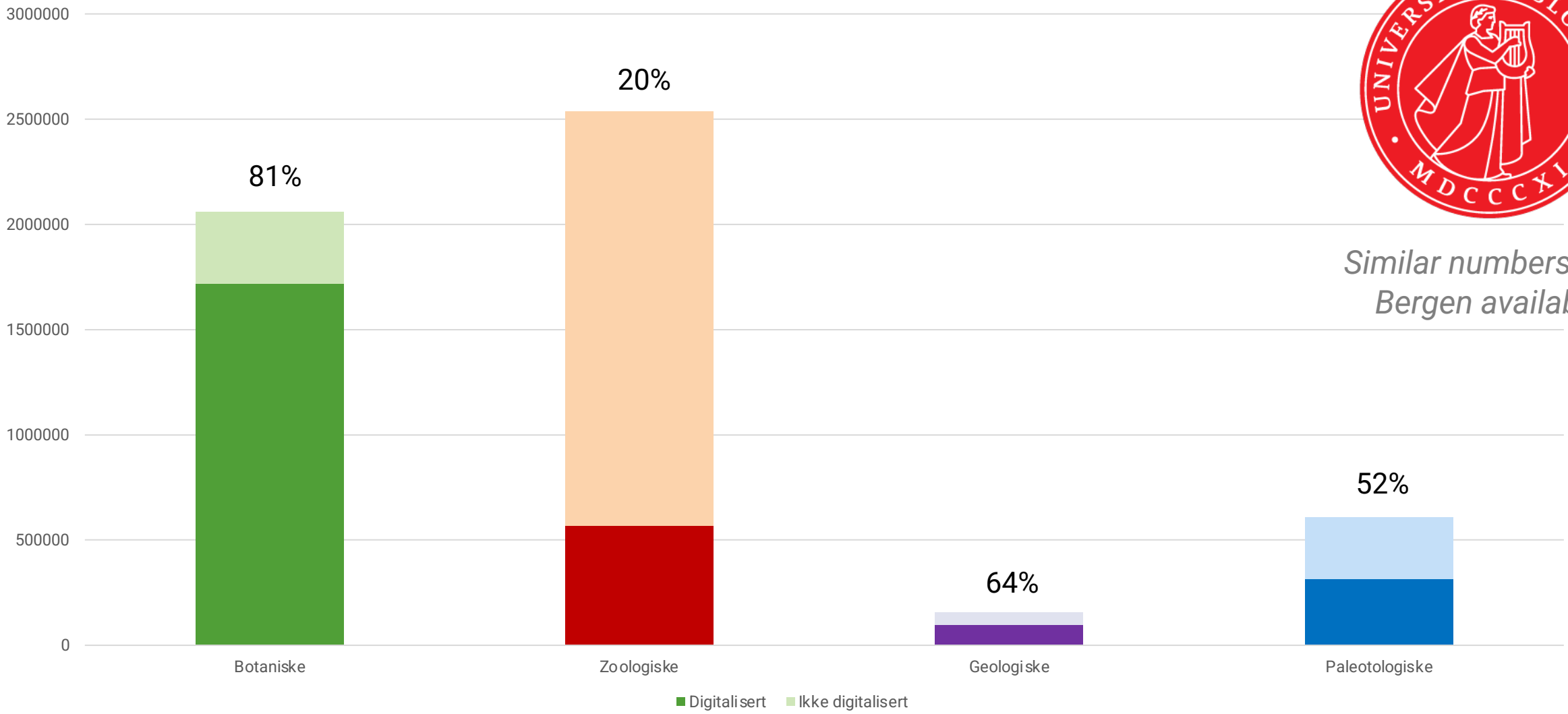
Natural history museum collections worldwide conserve an estimated **1.2 - 3 billion specimens**

(Ariño 2010; Duckworth *et al*. 1993)

GBIF publishes 2,3 billion records – including **278 million specimens**

approx. 10% coverage?

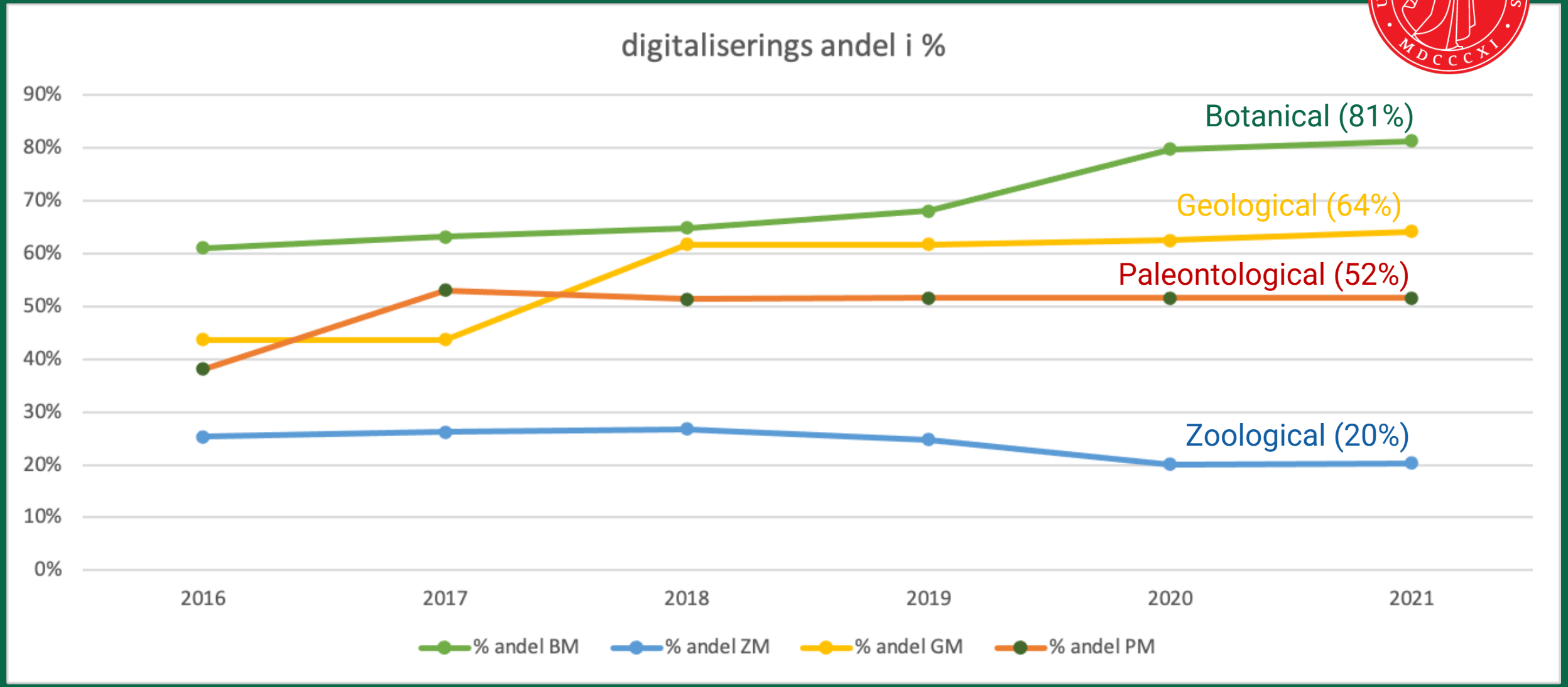
Specimens at UiO NHM Oslo (approx 52% digitized)



Similar numbers for Bergen available?

https://wiki.uio.no/nhm/skf/best-practices/index.php/Samlingstall | https://www.gbif.org/grscicoll/

SPECIMENS AT UIO NHM OSLO (PROGRESS PROPORTION DIGITIZED)



https://wiki.uio.no/nhm/skf/best-practices/index.php/Samlingstall | https://www.gbif.org/grscicoll/



digitization

Catalog number: O-L-000014 https://purl.org/nhmuio/id/41d9cbb4-4590-4265-8079-ca44d46d27c3

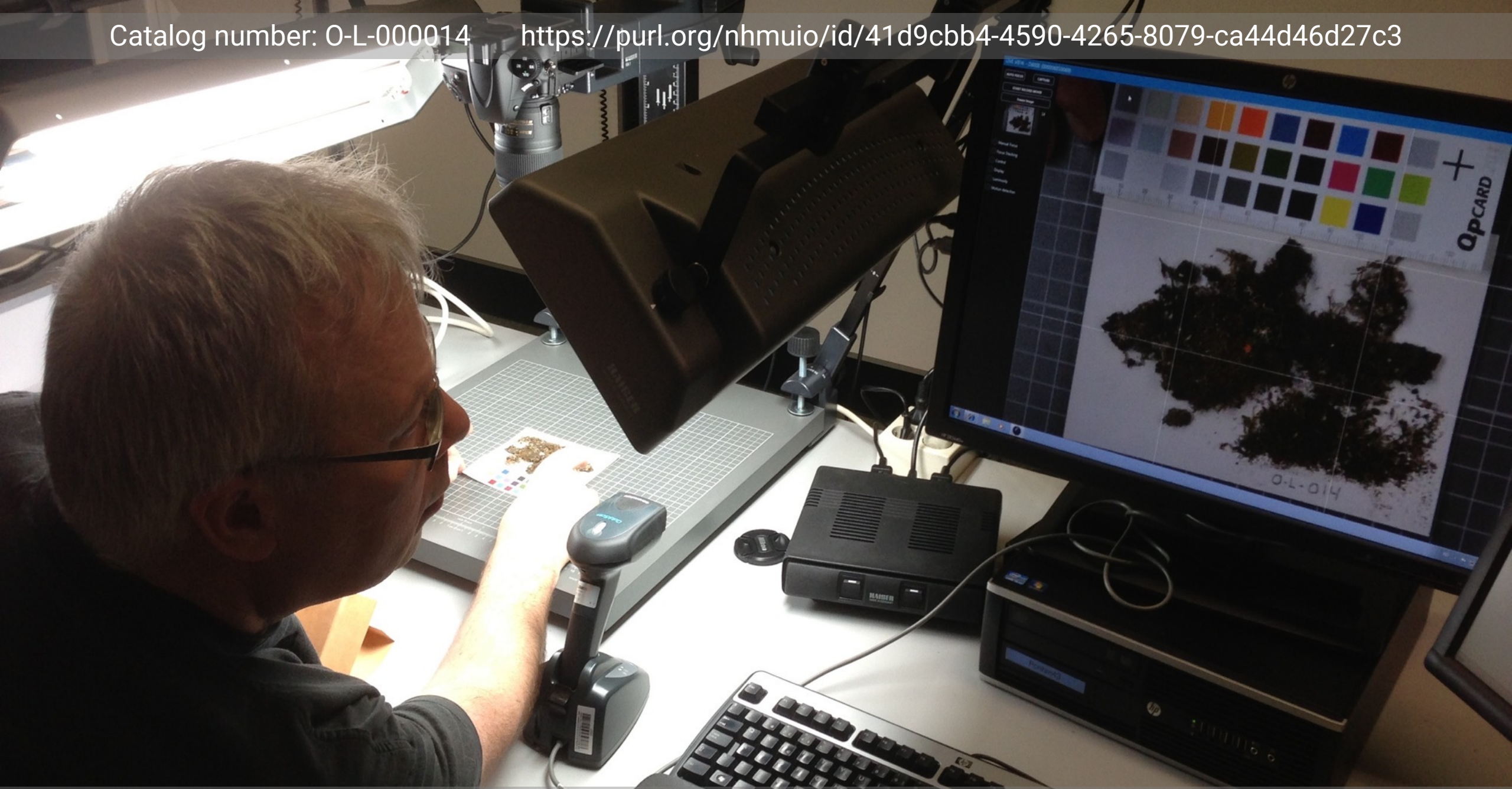


Photo: Curator Einar Timdal making images of type specimens in the UiO NHM Oslo lichen herbarium in 2013 - CC BY Dag Endresen.

https://gbif.no/dugnad/

HERBARIUM O

Herb. Univers. Osloensis

Heterodermia diademata (Tayl.) Awas.

KENYA
Rift Valley Prov. Uasin Gishu Distr.
5 km NW of Timboroa summit, on
moist brush by lake, and fence posts.
0-04 N 35-32 E 2650 m
2/1973 2K19/138 Hildur Krog

TLC (..../..):
zeorin
atranorin
Hildur Krog 19.



O-L-185347 Sign up or Sign in! (4 transcriptions)

Scientific name:	ɜrmia diademata (Tayl.) Awas.
Country:	Kenya
Province:	RiftValey Prov.
District:	Uasin Gishu Distr.
Locality:	5 km NW of Timbora summit, on moist brush by lake, and fence posts.
Habitat:	moist brush by lake
Position:	0-04n 35-32e
	Show map...
Elevation:	2650 m
Collector:	Hildur Krog
Collector no:	2K19/138
Date:	Day 2 1973

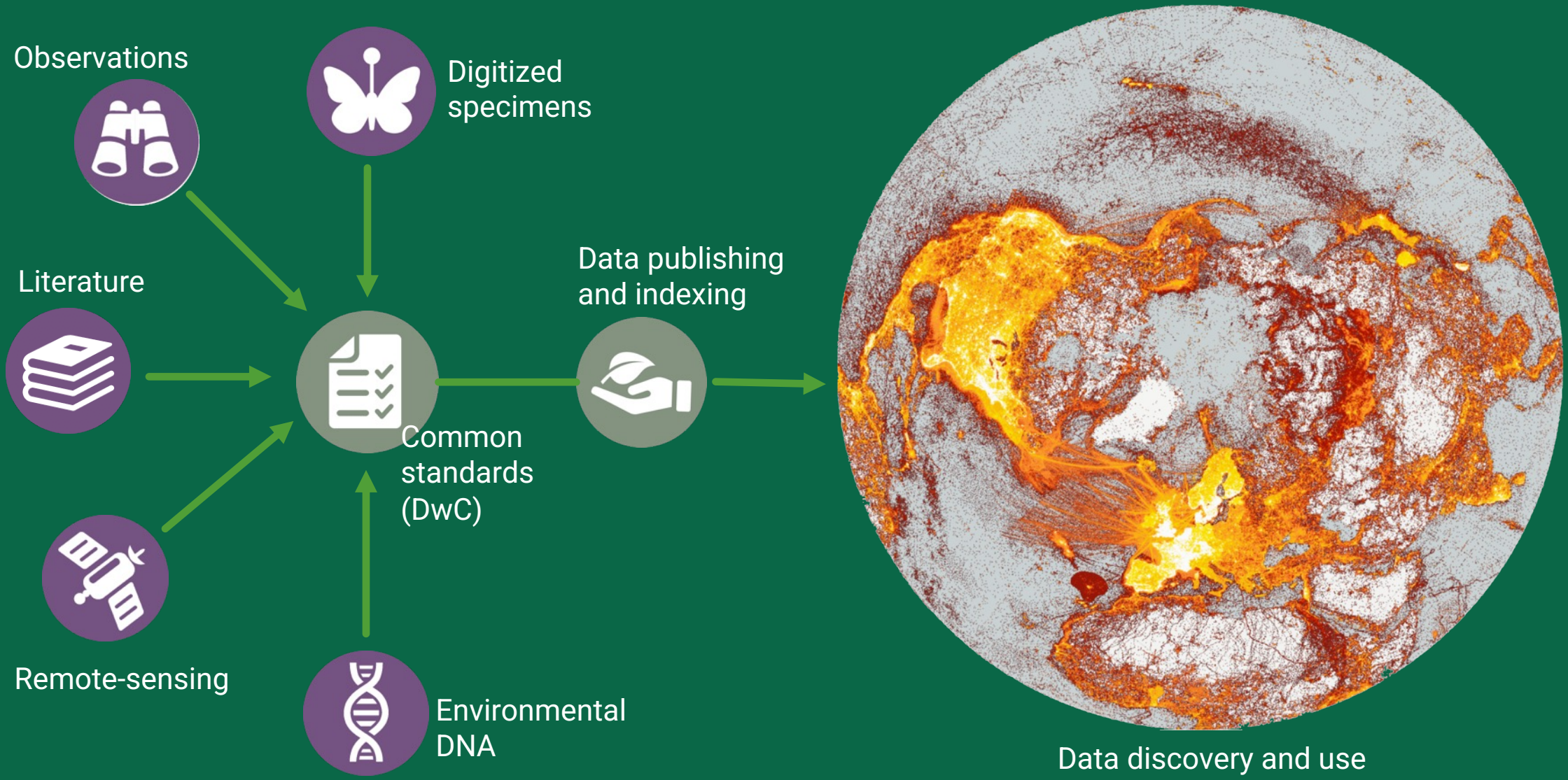
Skip this record... Next record...

Show help...



biodiversity data types

A WINDOW ON EVIDENCE ABOUT WHERE SPECIES HAVE LIVED, AND WHEN



https://www.gbif.org/occurrence/search

GBIF

DATA RICHNESS LEVELS SUPPORTED BY GBIF

Dataset description,
taxonomic/geographic/temporal scope



Dataset metadata

List of taxa
regional or thematic (e.g. invasive, medicinal)



Species checklists

Species occurrences
dates, coordinates, basis of record



Species occurrence data

Species occurrences and sampling events
dates, coordinates, sampling effort / protocol, abundance

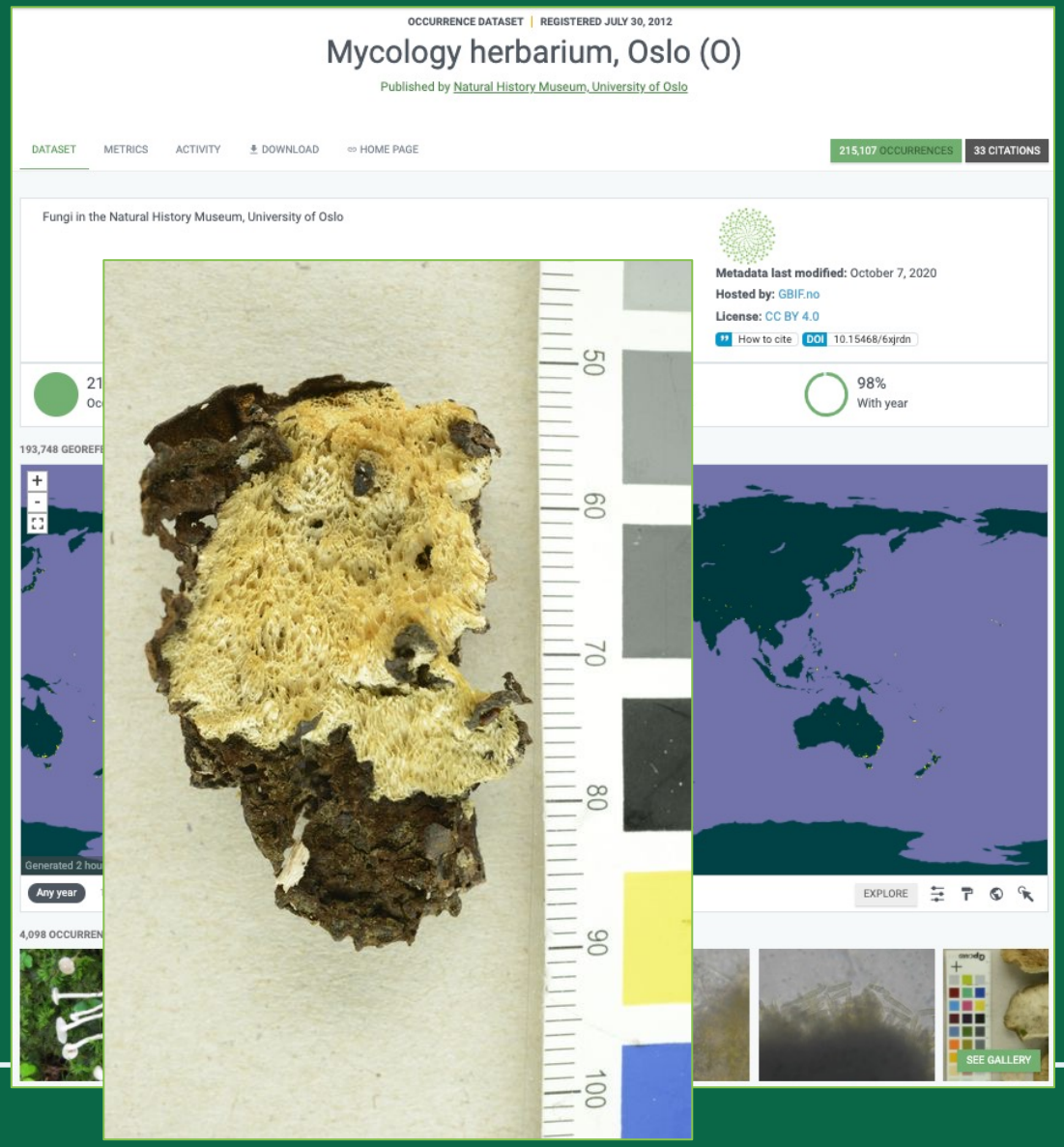


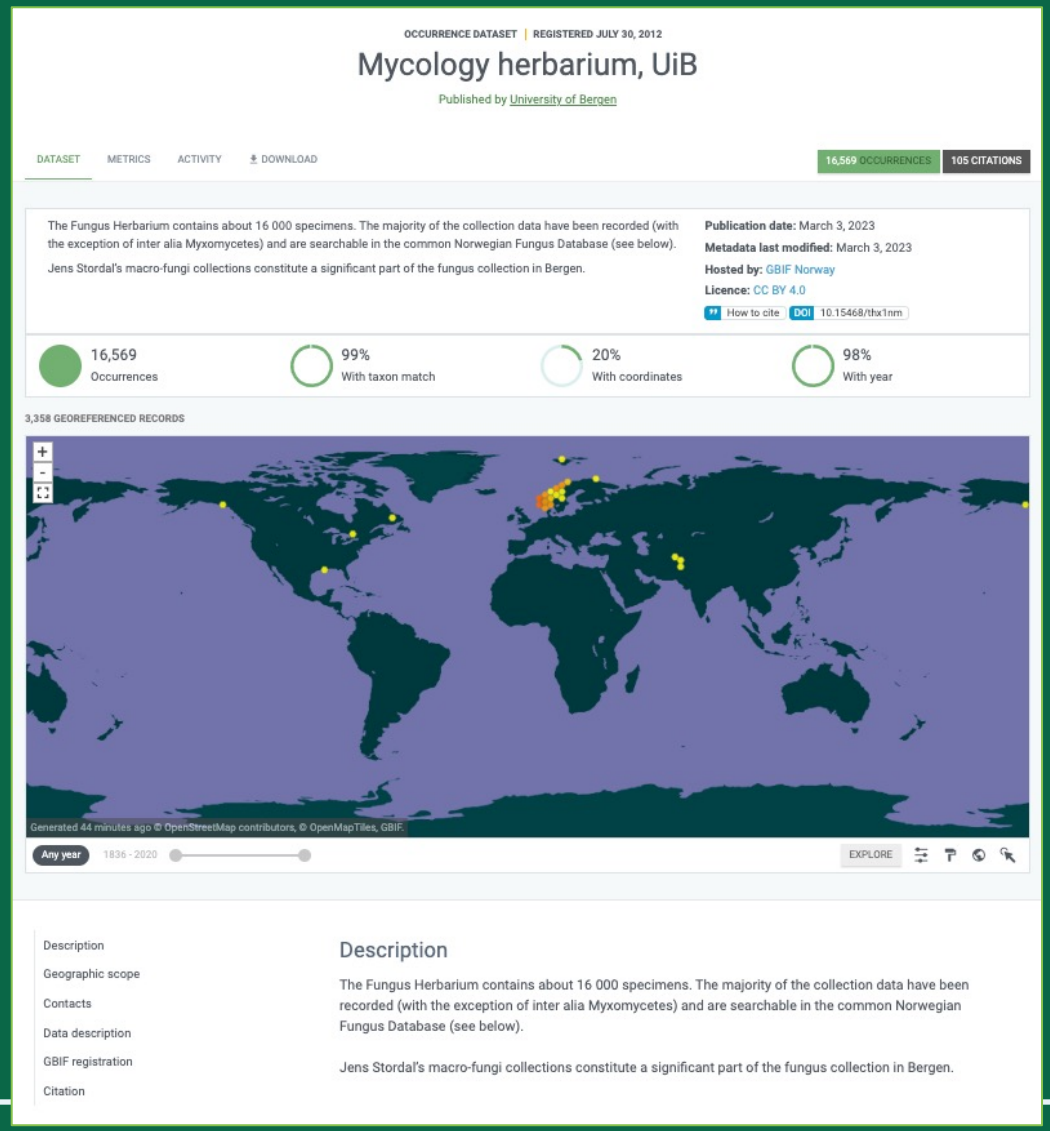
Sampling-event data

https://www.gbif.org/dataset-classes

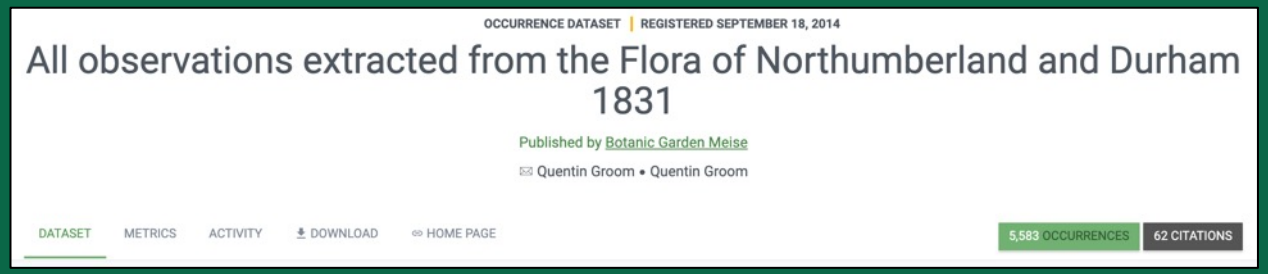
GBIF

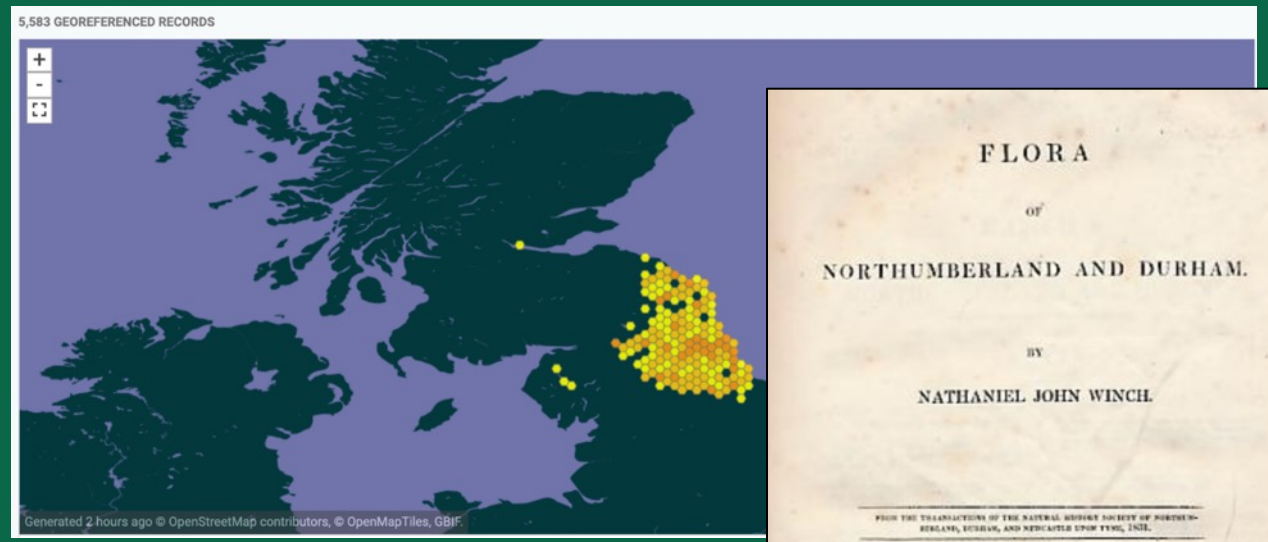
SOURCES OF DATA IN GBIF: DIGITIZED MUSEUM COLLECTION SPECIMENS





SOURCES OF DATA IN GBIF: TAXONOMIC LITERATURE, OLD AND NEW





FLORA

OF

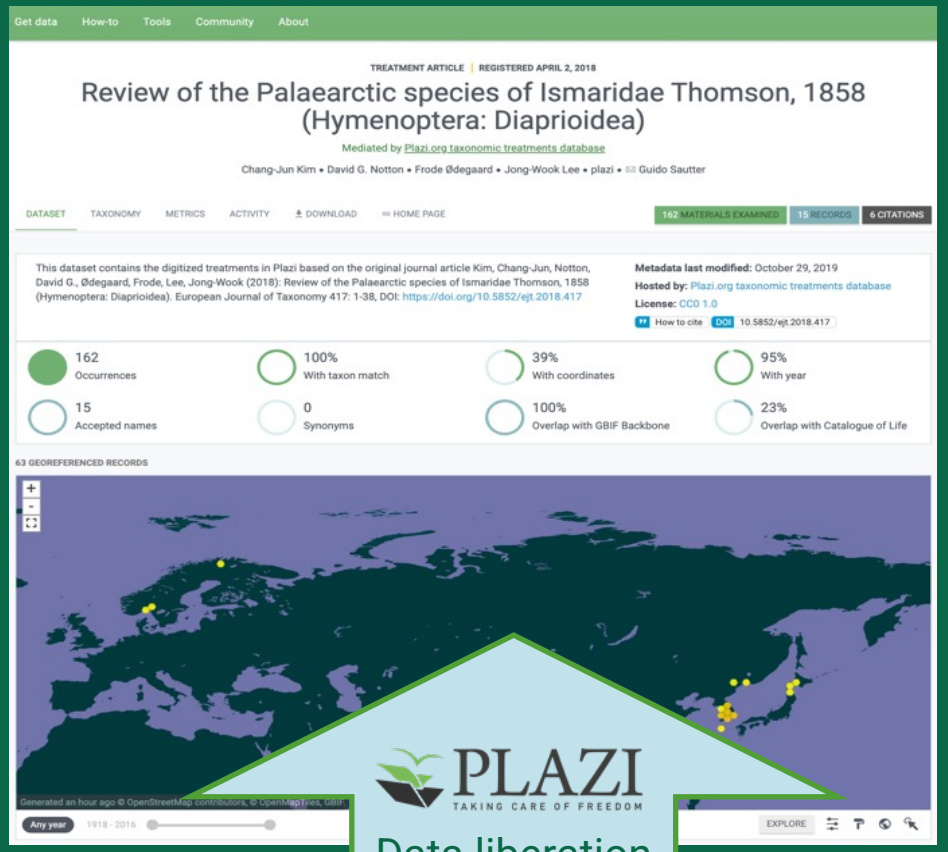
NORTHUMBERLAND AND DURHAM.

BY

NATHANIEL JOHN WINCH.

NEWCASTLE:
PRINTED BY T. AND J. HODGSON, UNION-STREET.
1811.

BHL
Biodiversity Heritage Library



PLAZI
TAKING CARE OF FREEDOM
Data liberation

ACS Publications
MOST TRUSTED. MOST CITED. MOST READ.
BioMed Central
The Open Access Publisher
ELSEVIER
NPF FOUNDATION
npg nature publishing group
PLOS
aap PSP
stm
Taylor & Francis Group an informa business
WILEY
Wolters Kluwer

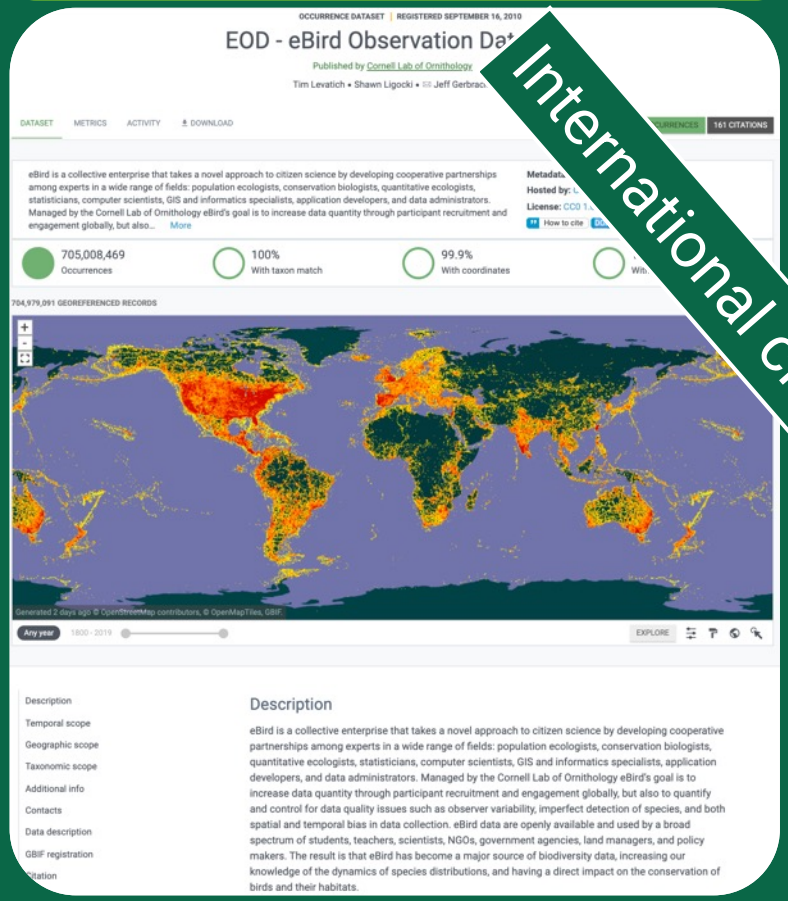
GBIF

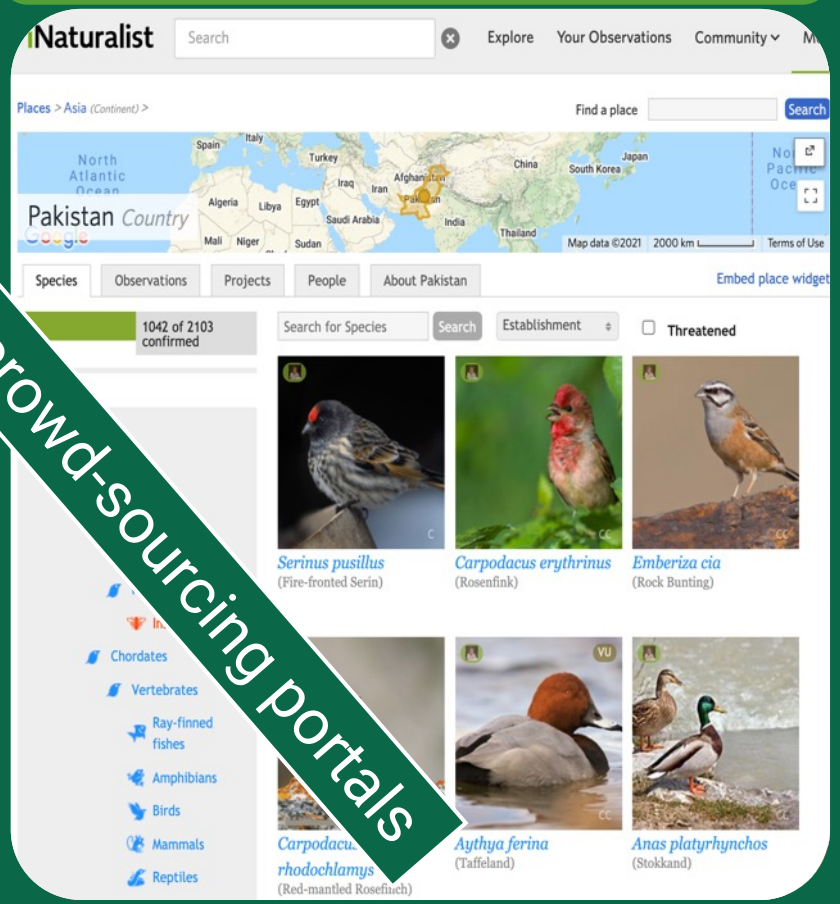
SOURCES OF DATA IN GBIF: CITIZEN SCIENCE OBSERVATIONS

eBird

iNaturalist

ARTSDATABANKEN





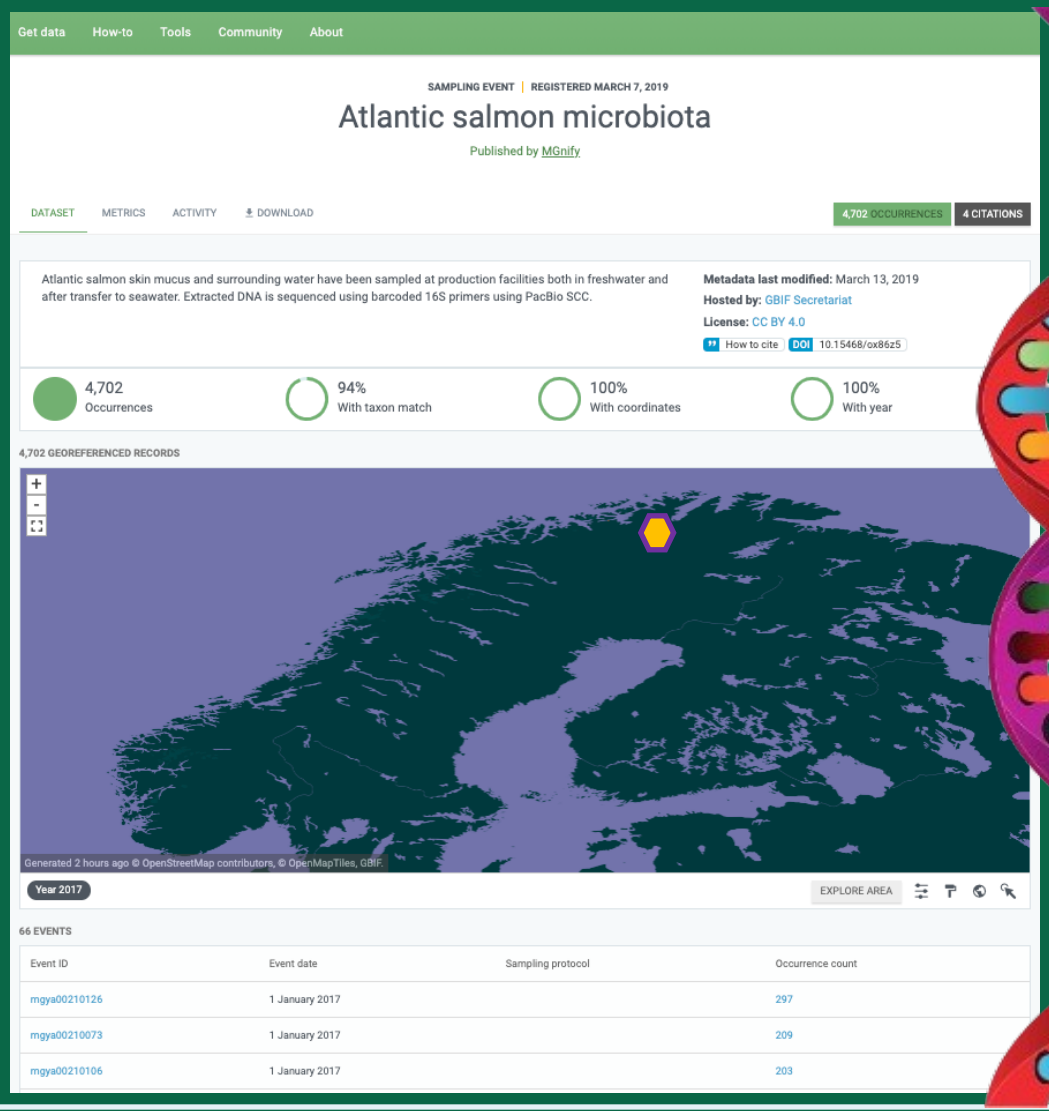


International crowd-sourcing portals

National crowd-sourcing portals

GBIF

SOURCES OF DATA IN GBIF: DNA SEQUENCE-DERIVED OCCURRENCE DATA

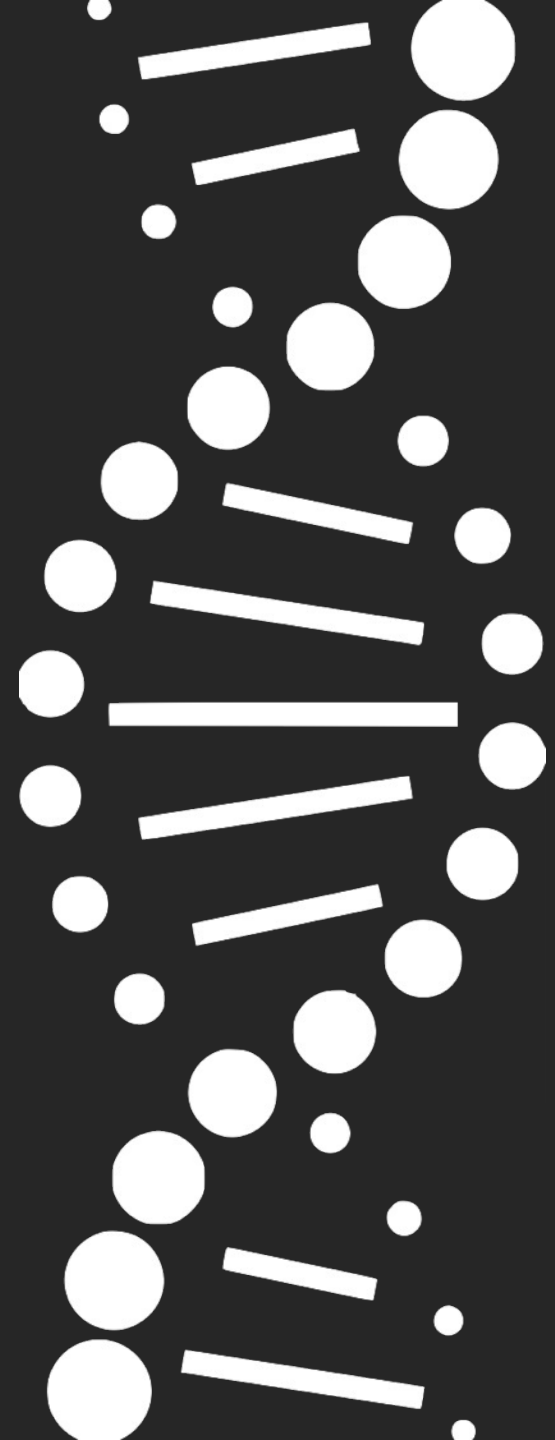






MGnify -- https://www.gbif.org/publisher/ab733144-7043-4e88-bd4f-fca7bf858880

molecular DNA data



NEW GBIF GUIDE: PUBLISHING **SEQUENCE-DERIVED DATA** THROUGH BIODIVERSITY DISCOVERY PLATFORMS

- Authors from Australia, **Norway**, Sweden, Denmark, UNITE, and GBIFS
- Based on practical mapping and data publishing experiences
- Cross-platform
- About 40 pages long ”cookbook”
  - **Introduction** – refresh your ”data culinary” knowledge
  - **Categorization** – what ”data ingredients” you got to publish?
  - **Mapping** – choose and follow the ”recipe”
  - **Visuals** – clarity and guidelines
  - Future prospects
  - Resources: glossary, links, references

Based on Darwin Core and MIxS data standards



Publishing DNA-derived data through biodiversity data platforms

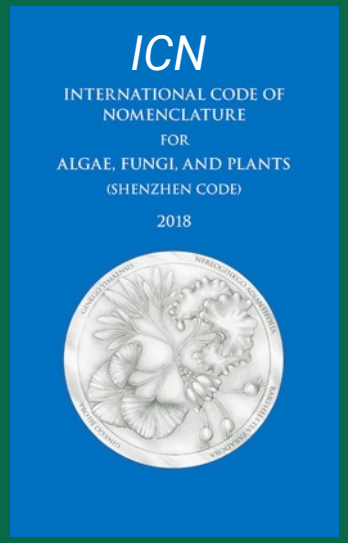
Anders F. Andersson · Andrew Bissett · Anders G. Finstad · Frode Fossøy · Marie Grosjean · Michael Hope · Thomas S. Jeppesen · Urmas Kõljalg · Daniel Lundin · R. Henrik Nilsson · Maria Prager · Cecilie Svenningsen · Dmitry Schigel – Version 3027b16, 2022-08-05 14:26:56 UTC

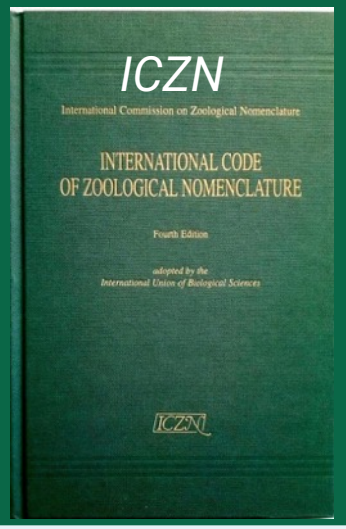
This document is also available in PDF format and in other languages: français.

https://doi.org/10.35035/doc-vf1a-nr22

GBIF

LATIN NAMES ARE RULED BY THE CODES





Domain (Eukarya)

Kingdom (Animalia)

Phylum (Chordata)

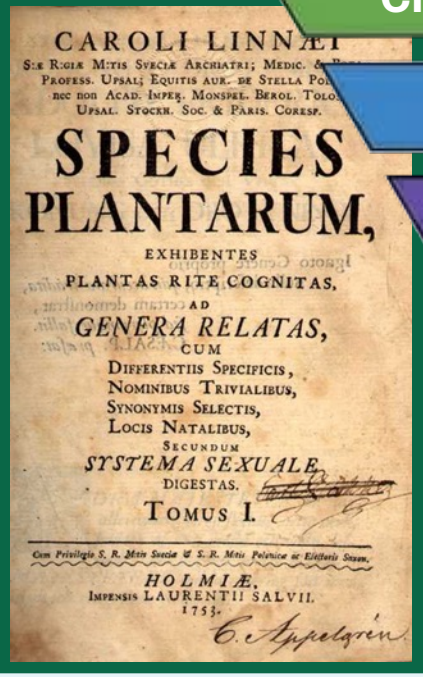
Class (Mammalia)

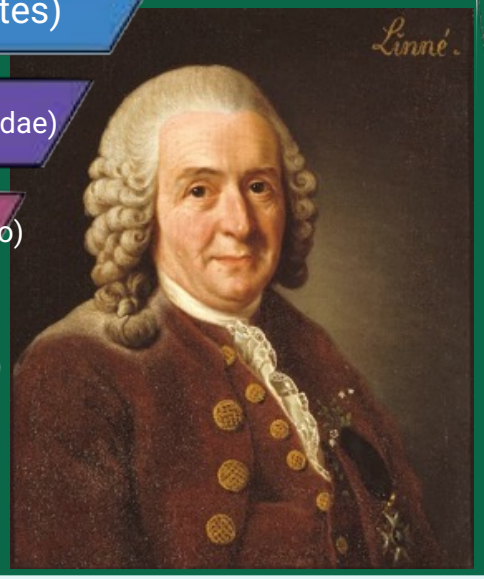
Order (Primates)

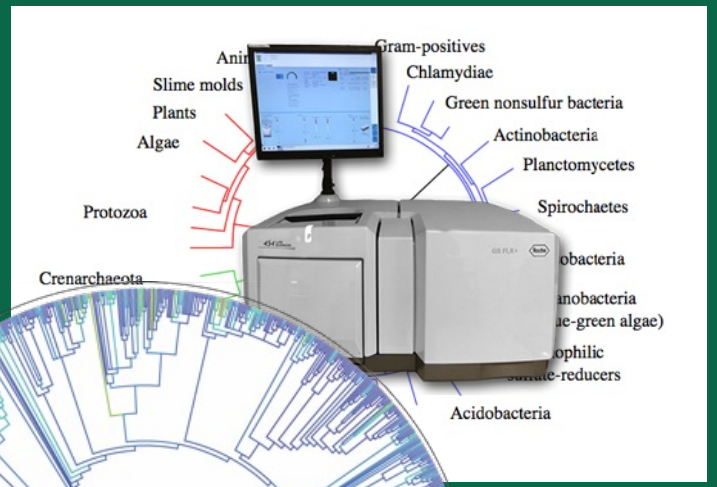
Family (Hominidae)

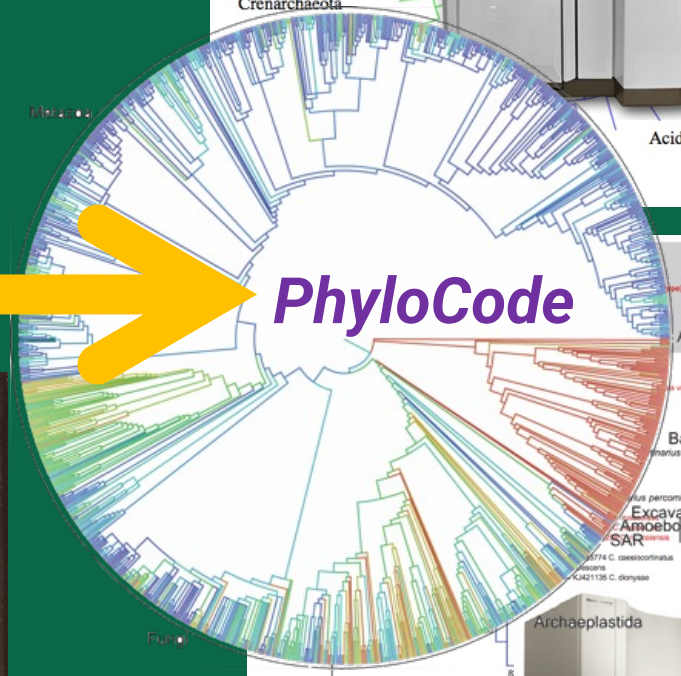
Genus (Homo)

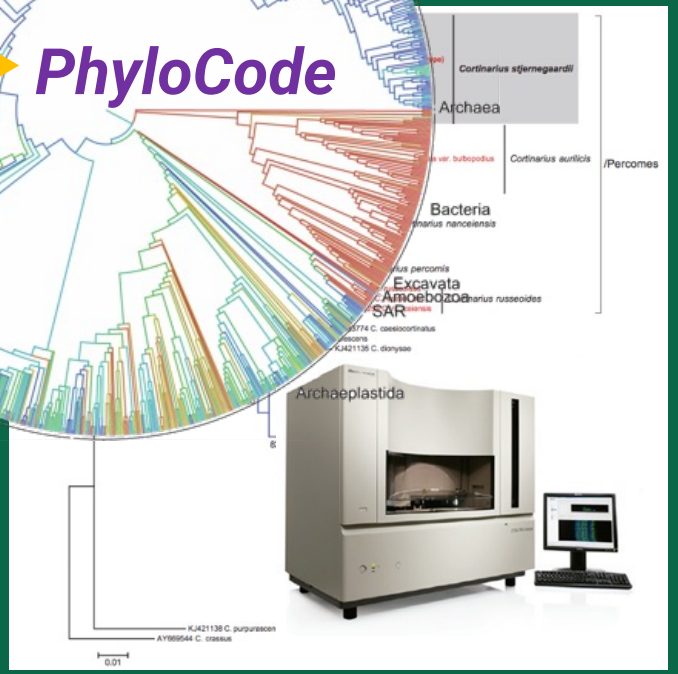
Species
(*Homo sapiens*)


CAROLI LINNÆI
SPECIES PLANTARUM,
EXHIBENTES
PLANTAS RITE COGNITAS,
AD
GENERA RELATAS,
CUM
DIFFERENTIIS SPECIFICIS,
NOMINIBUS TRIVIALIBUS,
SYNONYMIS SELECTIS,
LOCIS NATALIBUS,
SECUNDUM
SYSTEMA SEXUALE
DIGESTAS.
TOMUS I.
HOLMIÆ,
IMPENSIS LAURENTII SALVII.
1753.







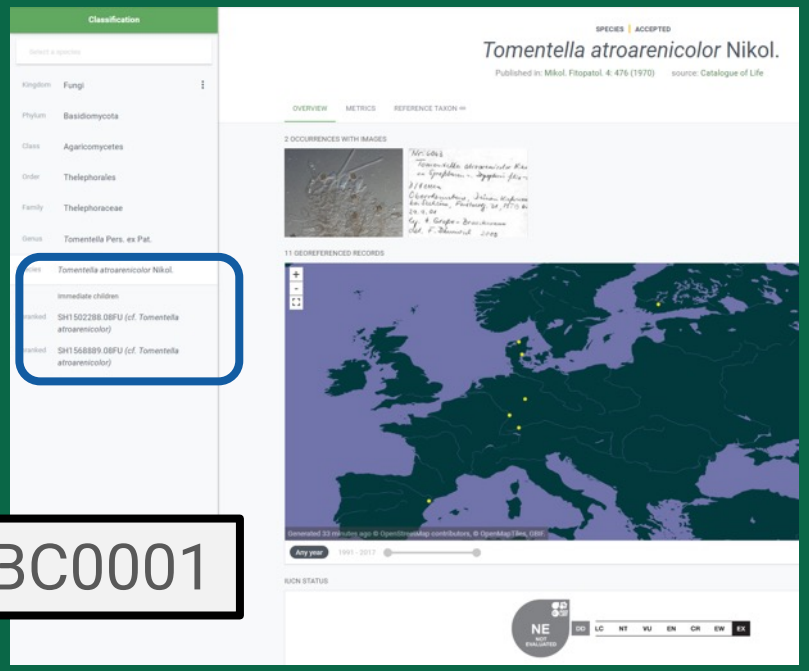


From codes to OTUs (operational taxonomic units)

GBIF

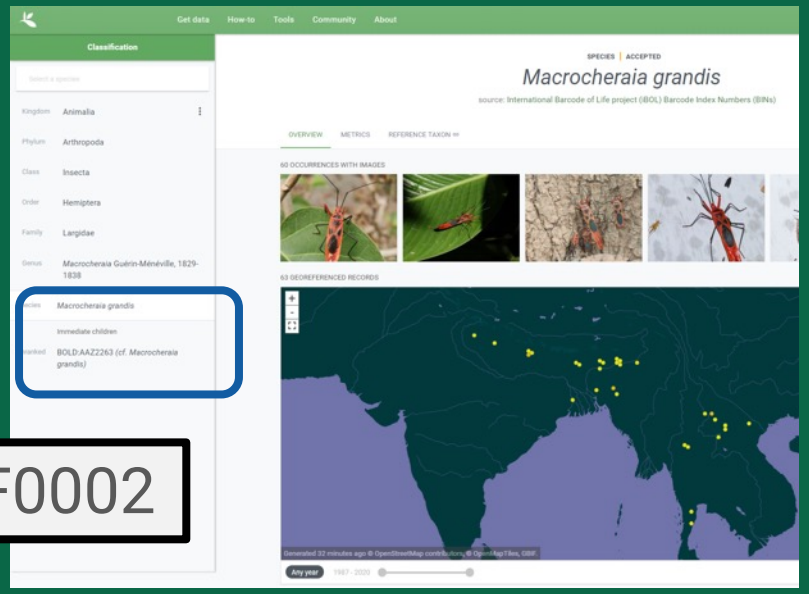


Species DNA Barcode



SH ABC0001

GBIF backbone taxonomy



BIN DEF0002

OTU = Operational Taxonomic Unit

OTU = **SH**
Species hypothesis numbers [DOI]

OTU = **BIN**
Barcode identification number

Catalogue of Life



- The most complete authoritative list of the world's species - maintained by hundreds of global taxonomists - probably includes just over 80% of the world's known species.
  - CoL 2022-10-20 doi:10.48580/d4t2 (approx. 2,3 M accepted species names)
  - CoL available in GBIF at doi:10.15468/rffz4x (4,7 M names, incl. synonyms)

Species 2000 ITIS GBIF alliance for biodiversity knowledge

WoRMS
World Register of Marine Species

World Register of Marine Species (WoRMS) provide an authoritative and comprehensive list of names of marine organisms

- WoRMS 2022-11-07 (241 387 accepted species names; 599 878 names)
- WoRMS available in GBIF at doi:10.14284/170 (719 524 accepted names)



LifeWatch BELGIUM

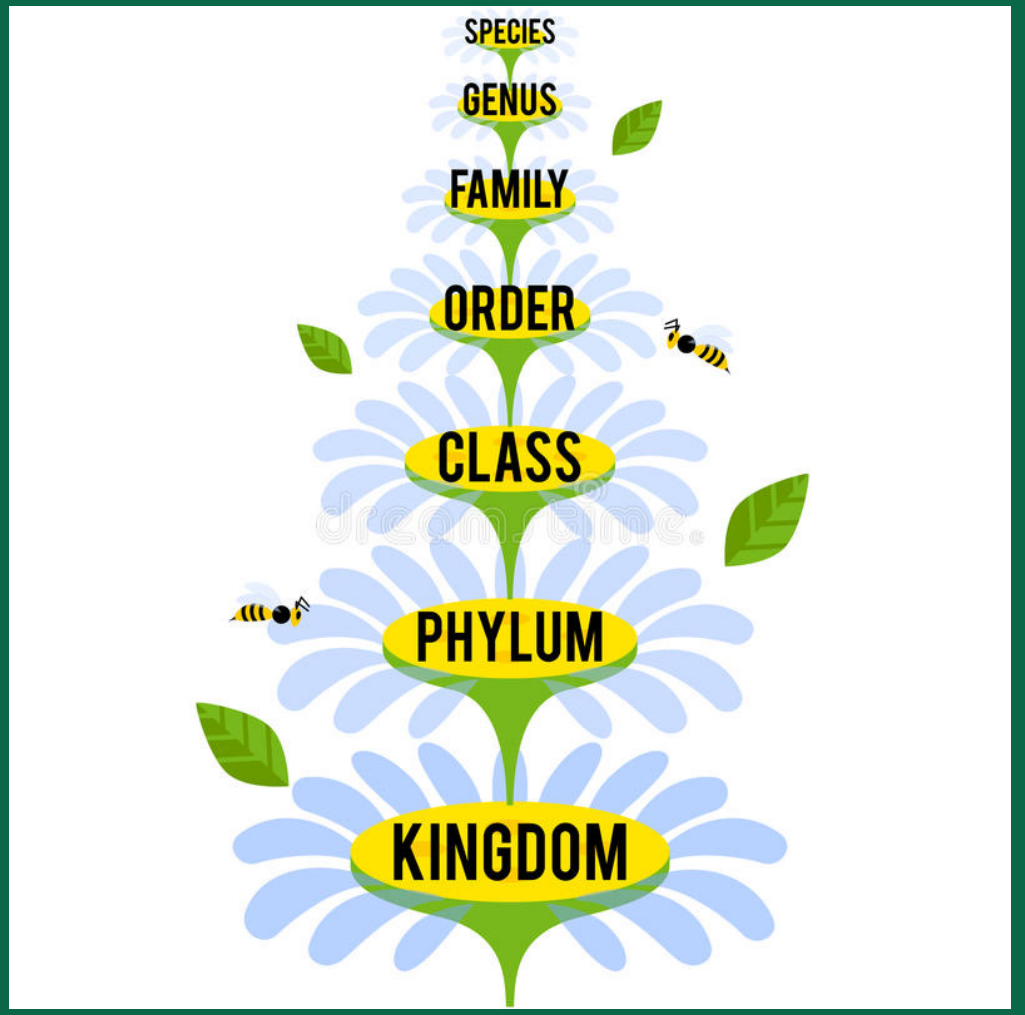
SeaLifeBase

FishBase

GBIF

ChecklistBank

- The Catalogue of Life (COL) support the publication and curation of checklists and provide a platform for their consistent discovery, use and citation.
- The taxonomic community can publish a checklist to ChecklistBank.
- ChecklistBank generates a standardised interpretation, and all datasets can be accessed via the ChecklistBank API.



Catalogue of Life

GBIF

GBIF

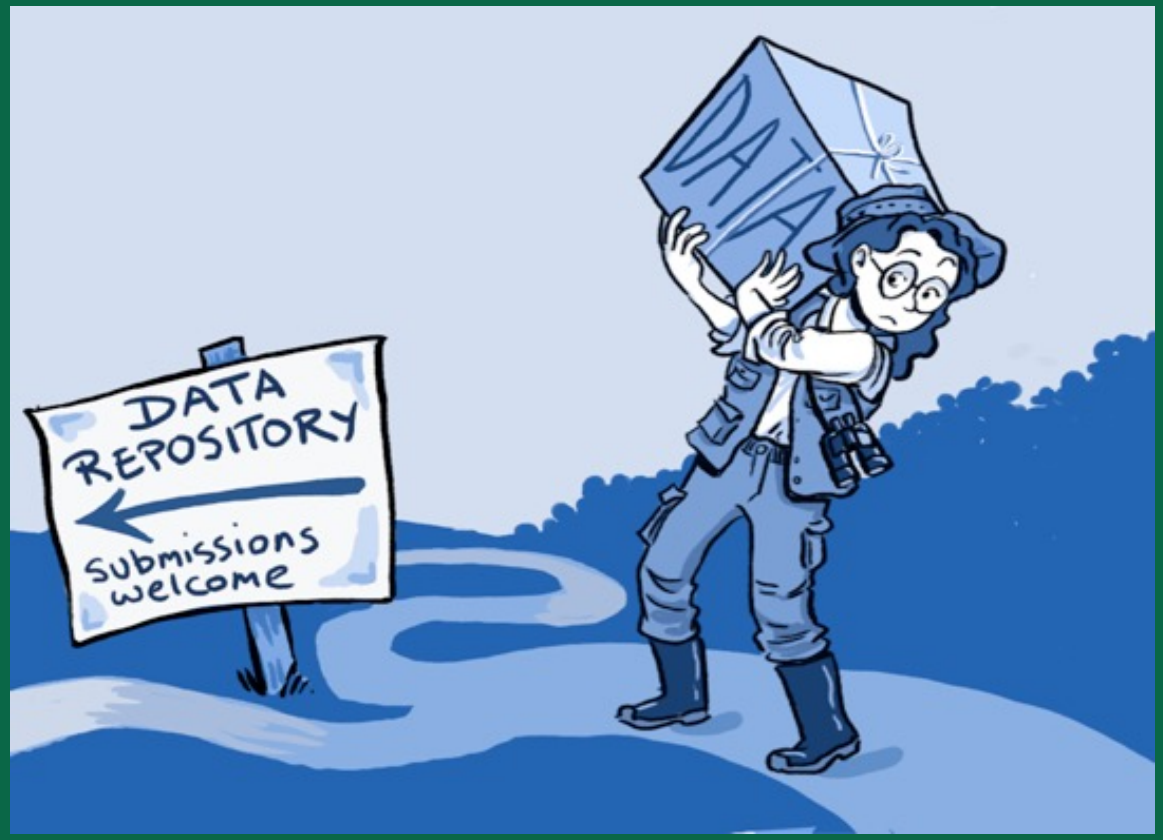
data publishing



PUBLISH YOUR DATASETS WITH GBIF

- *Step 1*: digitize collections & herbaria
- *Step 2*: register for endorsement in GBIF
- *Step 3*: convert to Darwin Core format
- *Step 4*: publish from national GBIF node

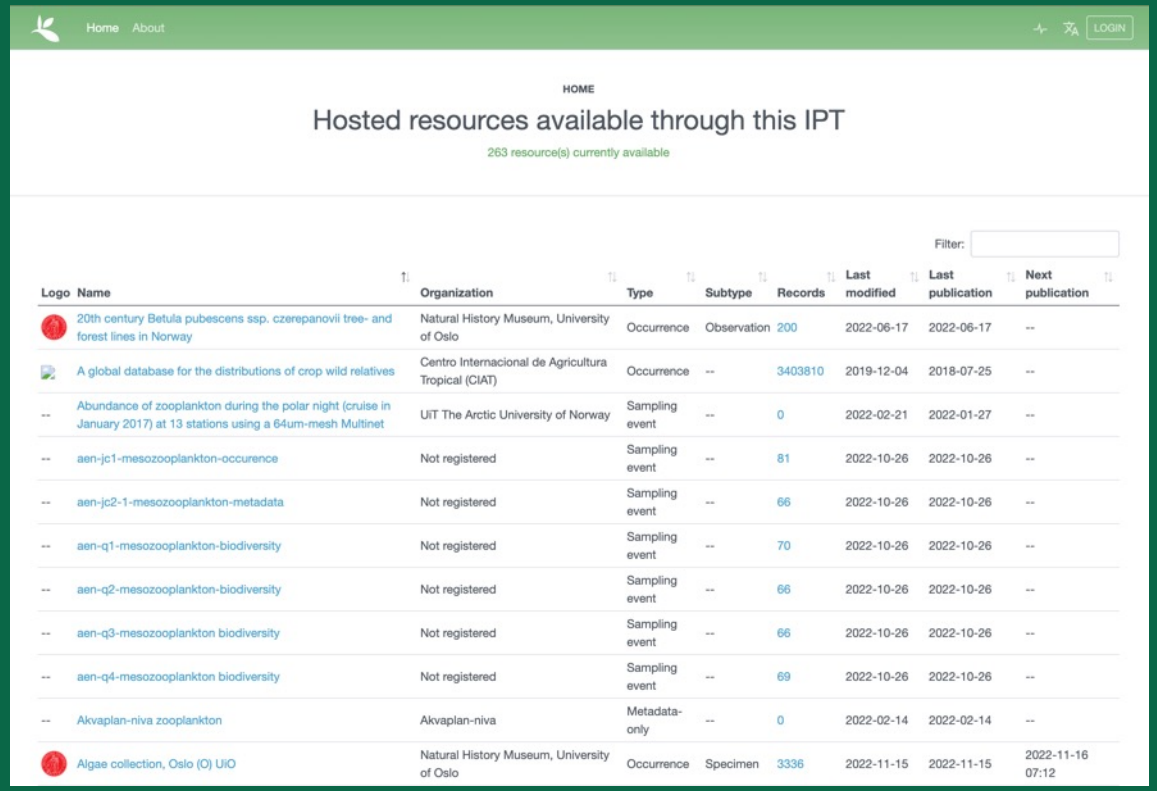
- *Alternative*: publish from regional GBIF cloud data repository - cloud.gbif.org/eca
- *Alternative*: Many citizen science data platforms publish data in GBIF



IPT DEMO

- Integrated data Publishing Toolkit (IPT)
- https://ipt.gbif.org/manual/en/ipt/2.6/dwca-guide

- https://ipt.gbif.no/ (Oslo)
- https://gbif.imr.no/ipt/ (IMR, Bergen)



GBIF

data use



GBIF IS A “FAIR” & OPEN BIODIVERSITY DATA INFRASTRUCTURE



GBIF: MULTIPLE-PURPOSE DATA PUBLISHING SERVICES



Bio-Collections
& ecology datasets

GBIF

Darwin Core

Environment directive reporting

Invasive Species

IUCN

IUCN RED LIST

Research data portals

elixir

BOLD SYSTEMS

GBIF portal

Convention on Biological Diversity

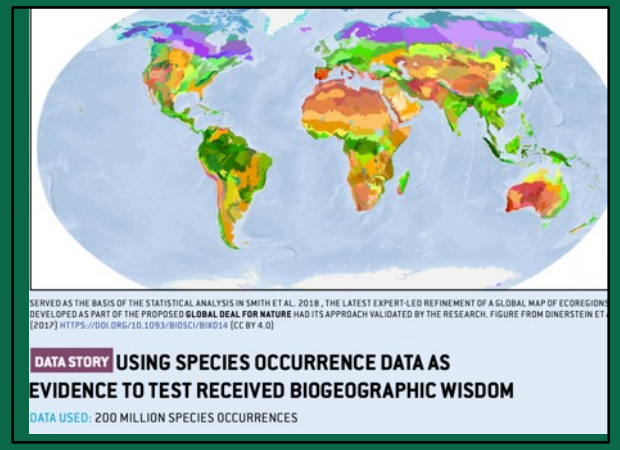
UNEP

FAO

ipbes
Science and Policy for People and Nature

GBIF

A DATA RESOURCE TO SUPPORT RESEARCH AND SUSTAINABLE DEVELOPMENT



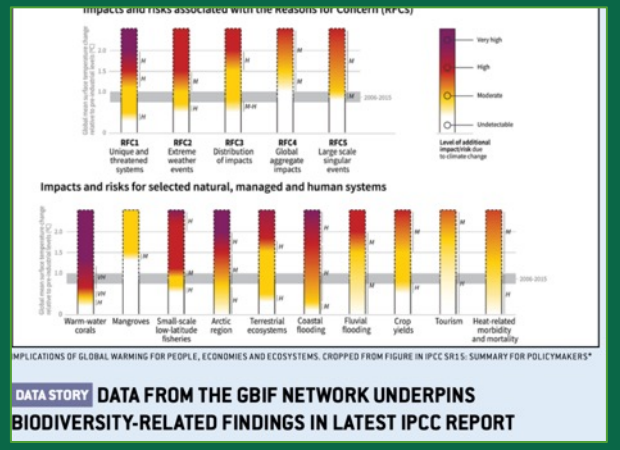
Conservation

- Protected areas
- Threatened species
- Invasive species risk



Food Security

- Crop wild relatives
- *In situ*, *ex situ* conservation of genetic diversity
- Fisheries planning



Climate change

- Modelling impacts on species ranges
- Adaptation strategies
- Mitigation benefits, risks



Human health

- Disease risk based on occurrence of vectors, hosts, reservoirs
- Medicinal plants
- Hazards e.g. snakebite

https://www.gbif.org/science-review

GBIF

POLICY LINKS: AICHI TARGETSUPPORTING BIODIVERSITY INDICATORS



- Trend in invasive alien species introductions (through Global Register of Introduced and Invasive Species)



- Species Protection Index
- Protected Area Representativeness Index



- Comprehensiveness of conservation of socioeconomically/culturally valuable species
- Agrobiodiversity Index
- Crop Wild Relative Index



- Growth in species occurrence records accessible through GBIF
- Species Status Information Index

https://www.cbd.int/cooperation/csp/gbif.shtml | https://www.cbd.int/csp/survey/GBIF.pdf

GBIF

POLICY LINKS

Convention on Biological Diversity

SUSTAINABLE DEVELOPMENT GOALS

COUNTRY PROFILES

Norway



Norway

Convention

Party since: 1993-12-29

By: Ratification

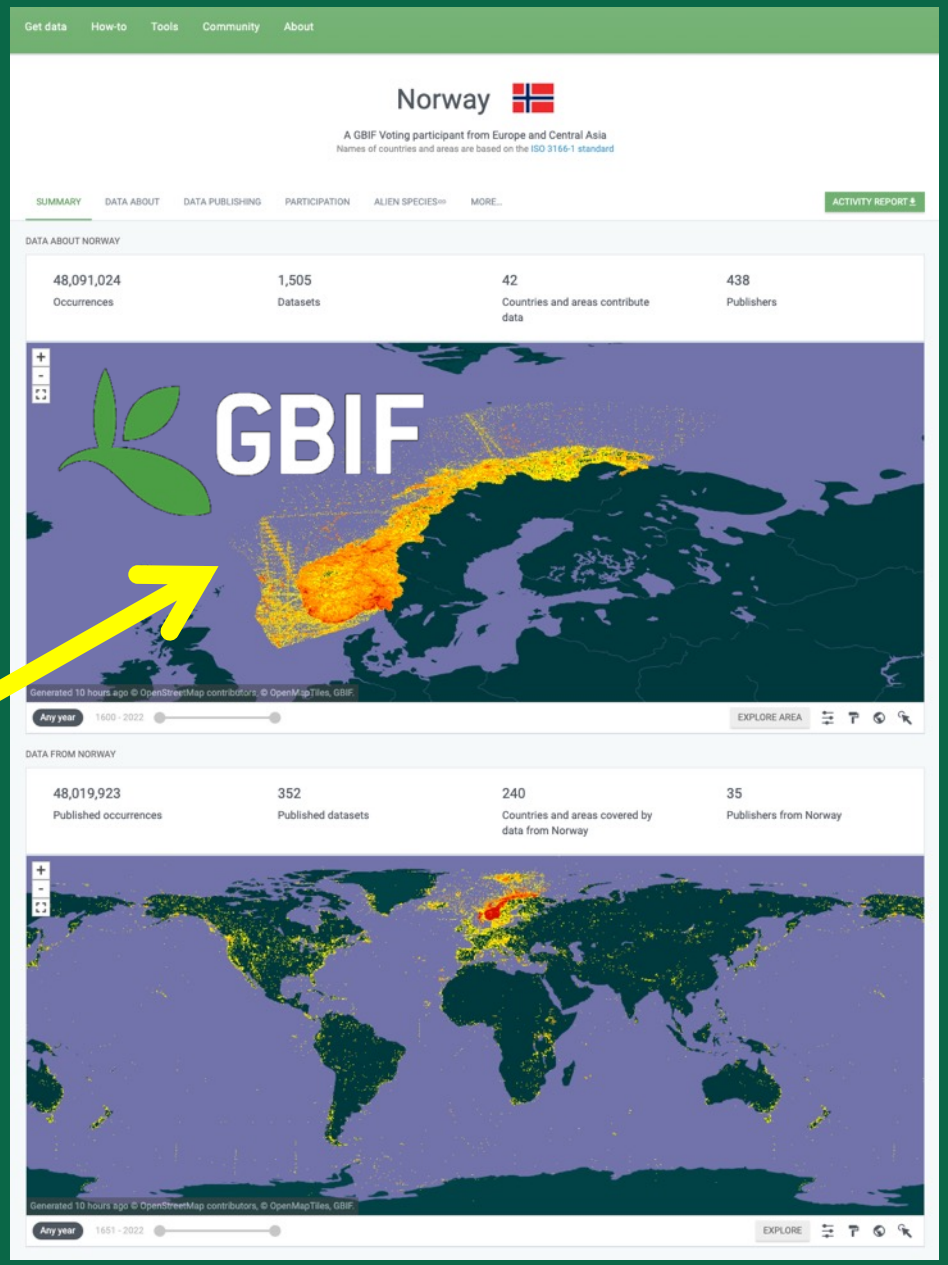
Cartagena Protocol

Party since: 2003-09-11

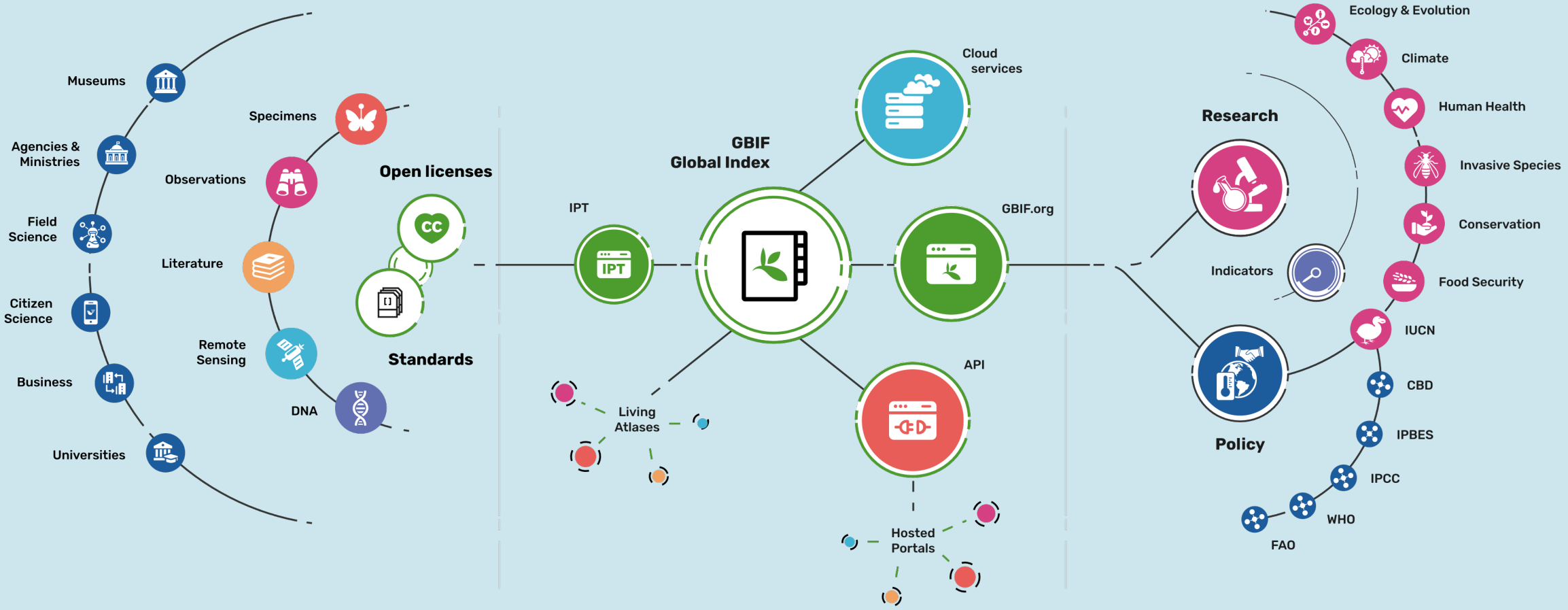
By: Ratification

Further Information

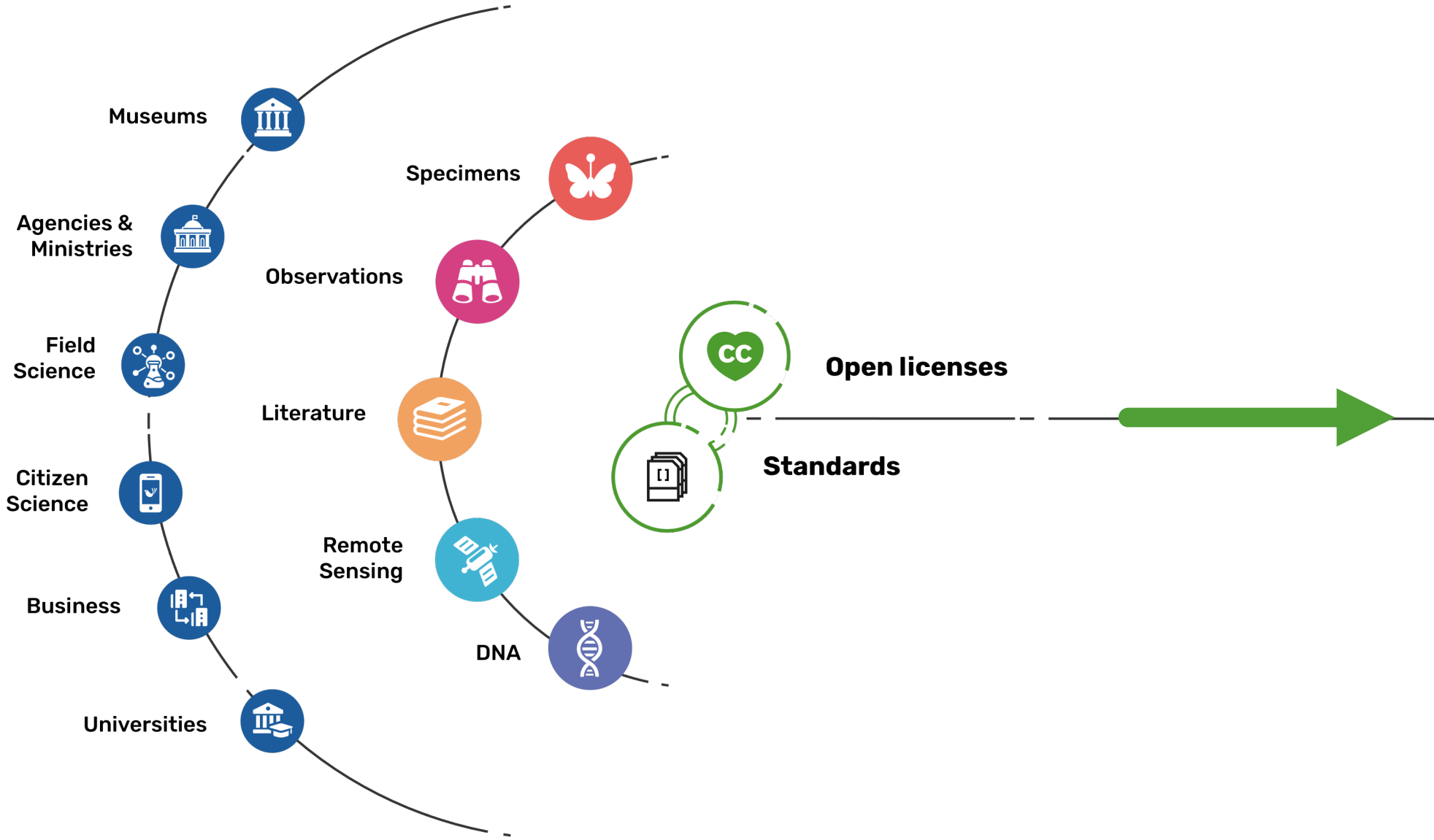
- InforMEA Profile
- World Heritage Sites
- GBIF Data
- WTO Statistics
- Law and Environment
- Invasive Alien Species
- Related Websites



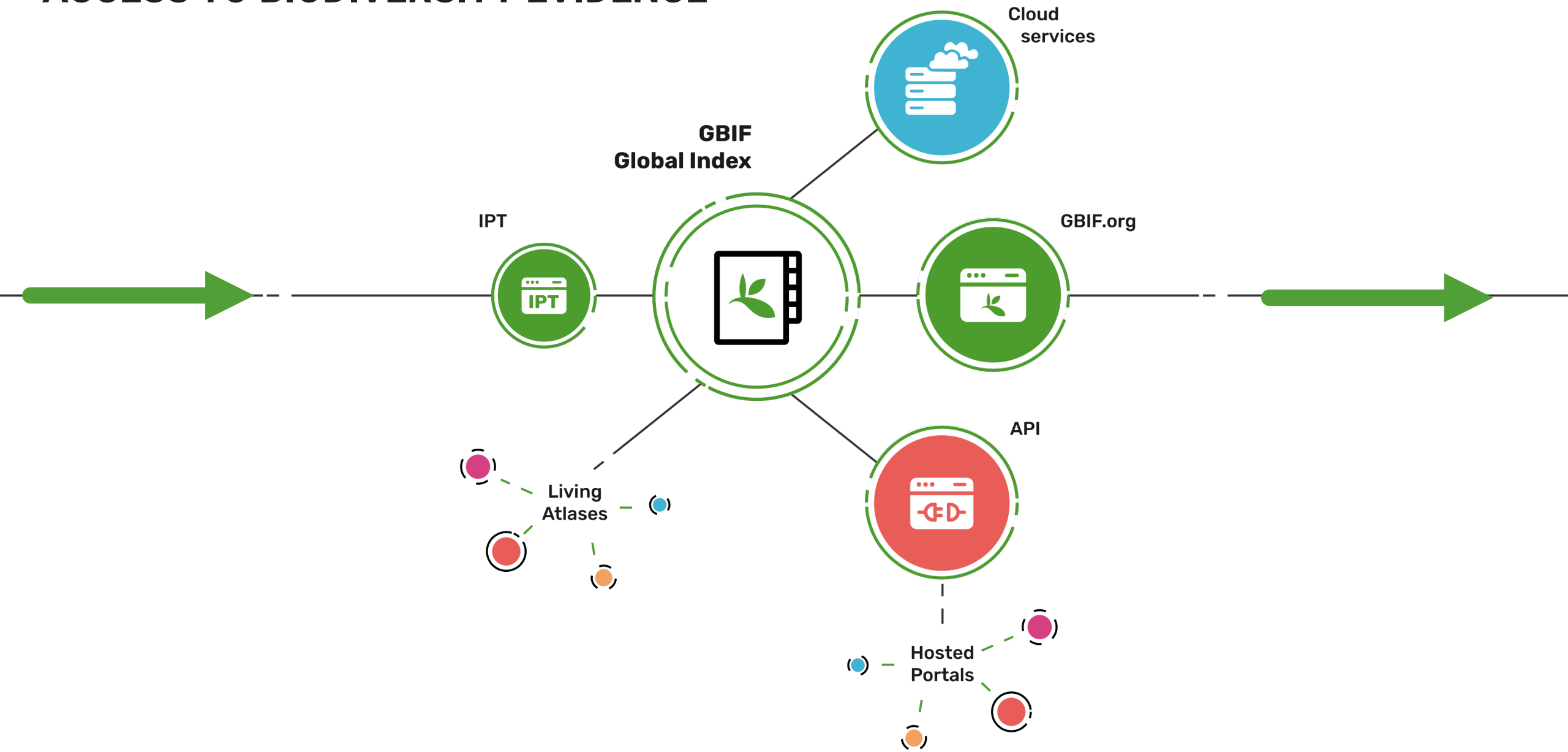
PROVIDING BIODIVERSITY EVIDENCE FOR RESEARCH AND POLICY



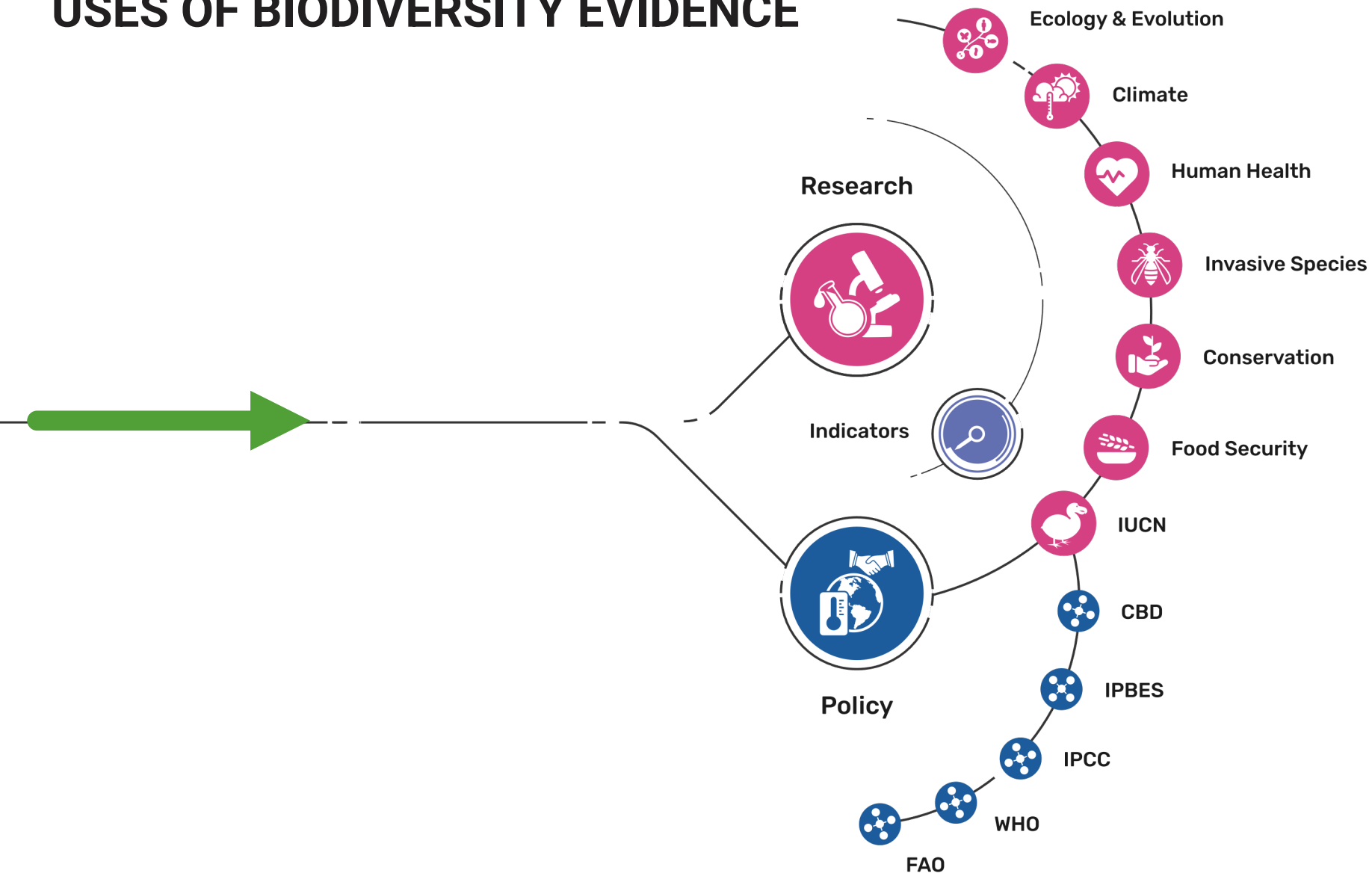
SOURCES OF BIODIVERSITY EVIDENCE



ACCESS TO BIODIVERSITY EVIDENCE



USES OF BIODIVERSITY EVIDENCE



WHY SHOULD MUSEUMS CARE ABOUT OPEN SCIENCE?





OPPORTUNITIES

- Enables new research methodologies that were not possible before.
- Skills for open research and open data are in increasing demand!
- Funding opportunities.
- GBIF brings new benefits and opportunities for our museums.

WHY APPROACH OPEN SCIENCE IN MUSEUMS?

- We are in the middle of an ongoing **paradigm shift** in scientific practice (*and impact metrics*).
- Natural History Museums will need to develop **different approaches**, than they needed in the past – *to remain relevant*.
- Society is gaining Big Data maturity and will **expect new services** from museum collections.
- The open science wave is moving **fast**!



DATA CITATION - A NEW CURRENCY OF SCIENCE

- **Peer-reviewed scholarly papers** in high impact journals still maintain considerable weight for impact metrics.
- A **movement** is under way to **build similar status for** open data, open metadata, open material samples, and other **open access scientific research products**…





DECLARATION ON RESEARCH ASSESSMENT



- DORA recognizes the need to improve the ways in which the outputs of scholarly research are evaluated.
- Worldwide movement covering all scholarly disciplines and all key stakeholders including funders, publishers, professional societies, institutions (universities), and researchers.

- Developed in 2012 in San Francisco
- To date (2022-11-07), 19 674 individuals and 2 636 organizations in 159 countries have signed DORA.
- The Research Council of Norway (RCN) signed DORA in May 2018 [link]

INCENTIVE FOR DATA REUSE

To incentivize the sharing of useful data, the scientific enterprise needs a well-defined ***system that links individuals with reuse of data sets they generate***

Pierce *et al*. Credit data generators for data reuse, *Nature* 6 June 2019



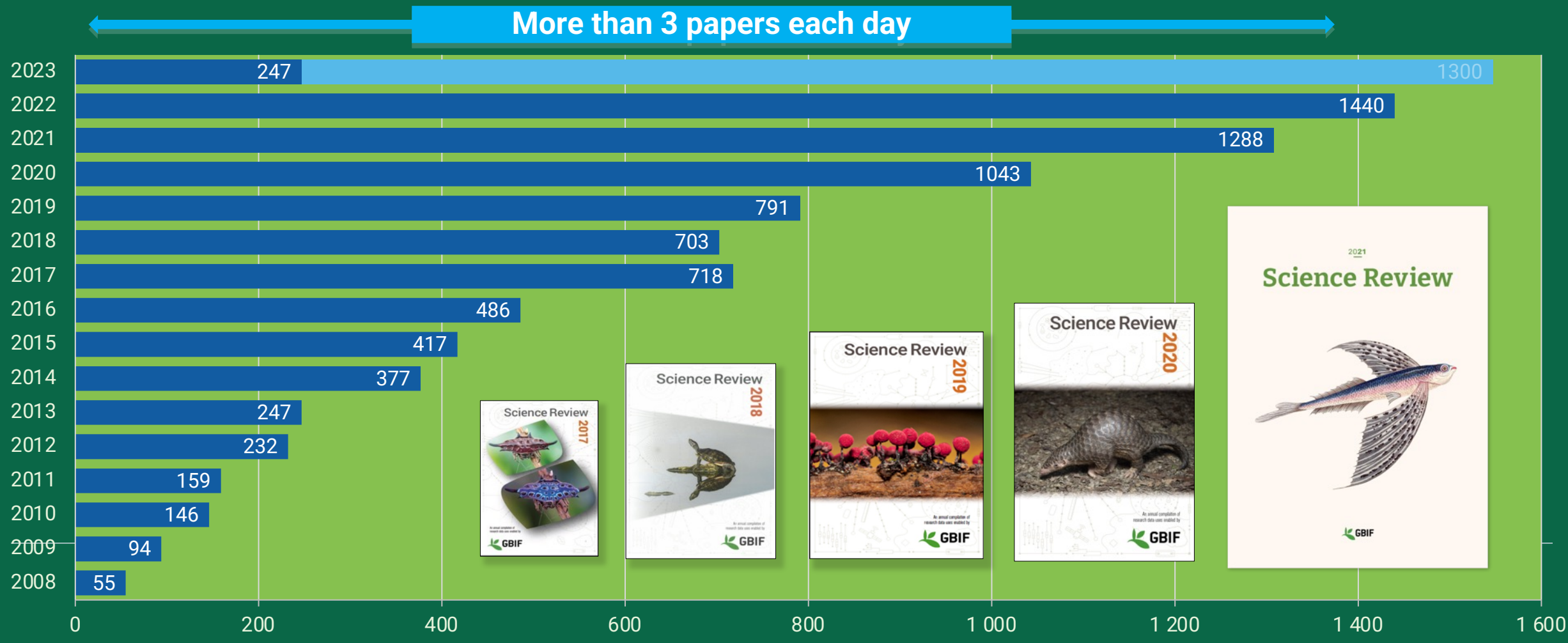


#CiteTheDOI

GBIF started issuing DOIs on 3 February 2015



PEER-REVIEWED PUBLICATIONS USING GBIF-MEDIATED DATA



https://www.gbif.org/resource/search?contentType=literature&literatureType=journal&relevance=GBIF_USED&peerReview=true

#CiteTheDOI

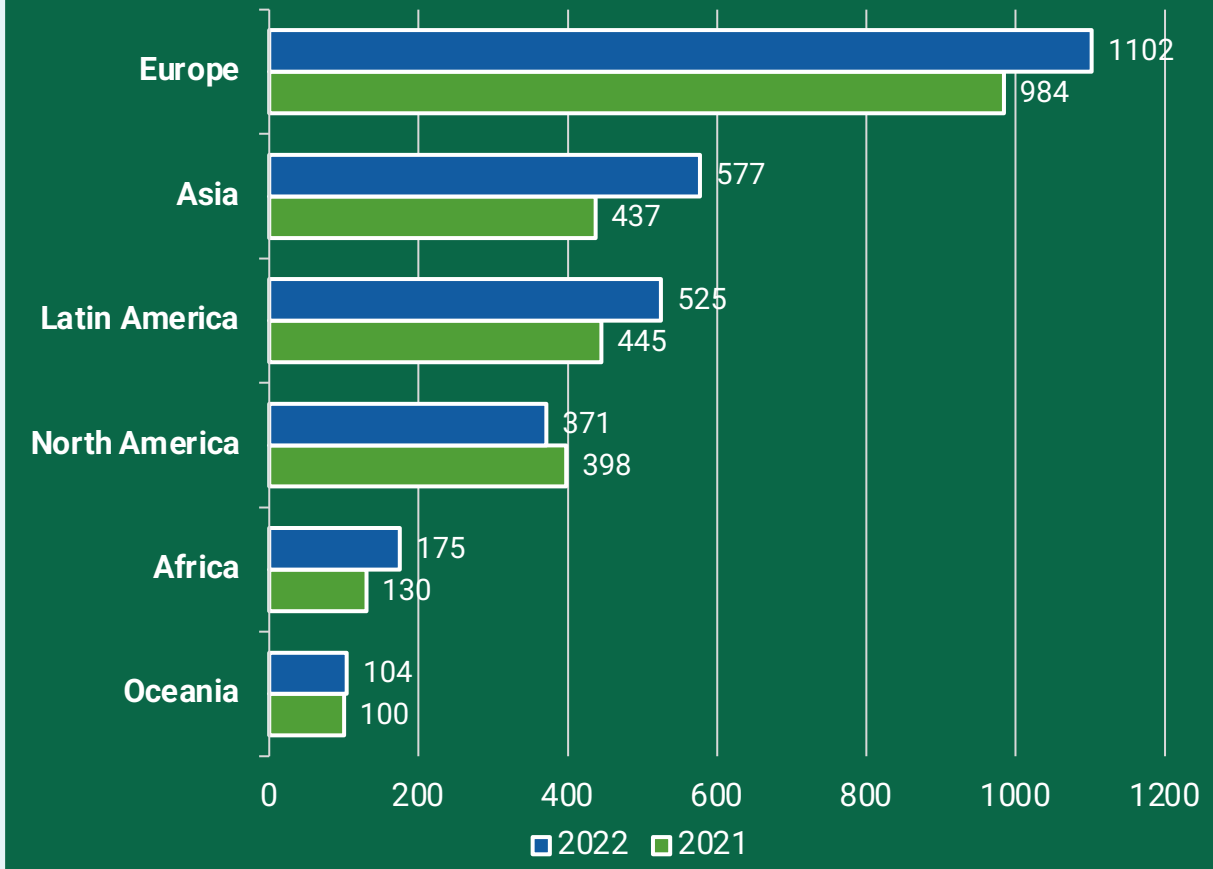
GBIF

DATA USE IN PEER-REVIEWED JOURNALS 31 Dec 2022

Peer-reviewed uses by country

End of Year		*2022 total*	*2021 total*	*2021 rank*
1	China	309	227 331	2
2	United States	306	331	1
3	Brazil	167	142	3
4	United Kingdom	149	140	4
5	Mexico	146	130	6
6	Germany	126	131	5
7	Spain	105	102	7
8	Australia	78	70	9
8	France	78	78	8
10	Italy	66	59	11

Peer-reviewed uses by region



https://www.gbif.org/resource/search?contentType=literature&literatureType=journal&relevance=GBIF_USED&peerReview=true

GBIF

HOW TO CITE DATA MEDIATED BY GBIF?

1. **Download data** from GBIF.org
2. and receive recommended citation with a **download DOI**
3. **Cite the DOI** in published research or other work

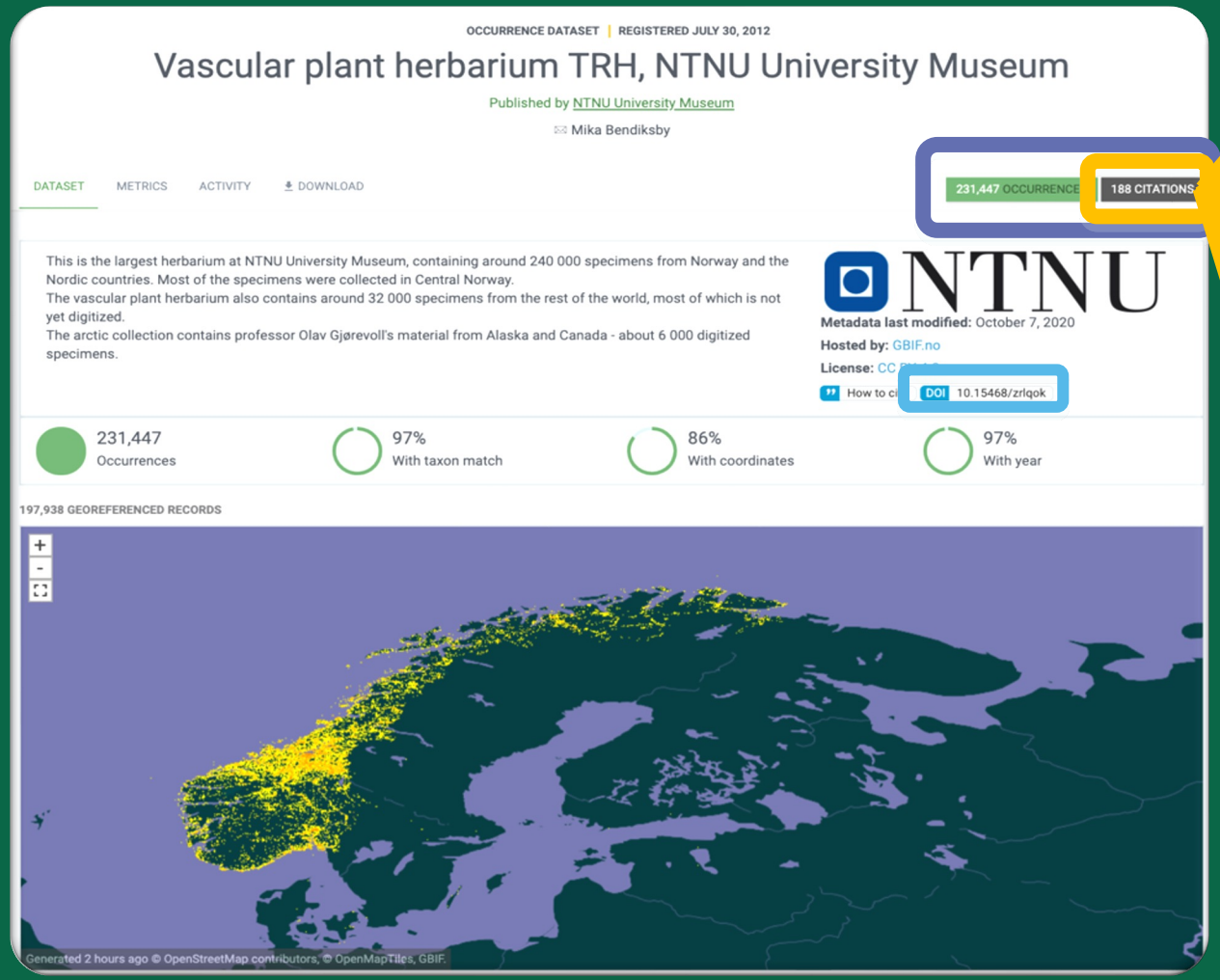
Example: GBIF.org (9 November 2021) GBIF Occurrence Download **https//doi.org/10.15468/dl.xxxxxx**

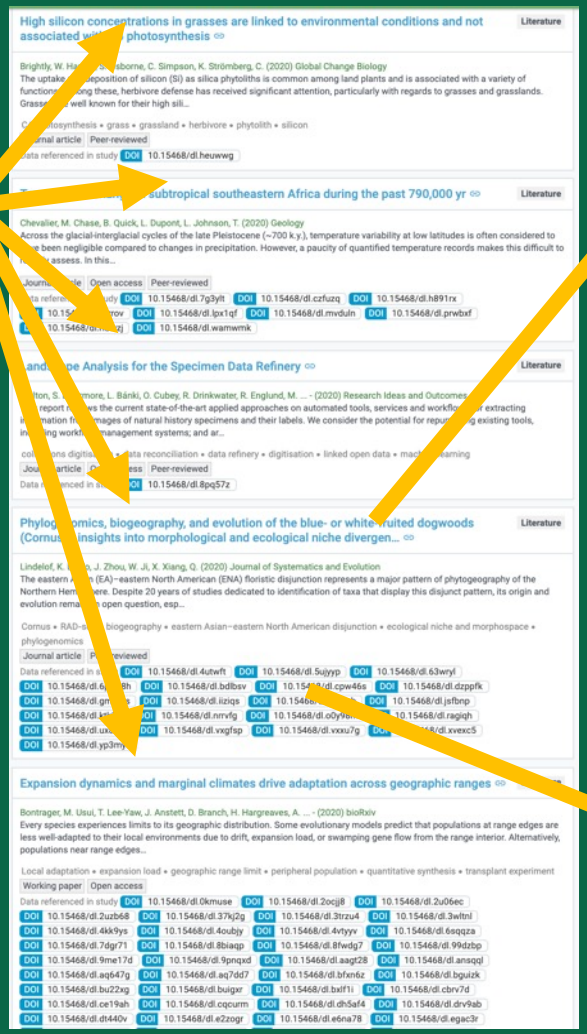
DOI BASED DATA CITATION AT GBIF.ORG -- #CITETHEDOI

dataset

citations

papers





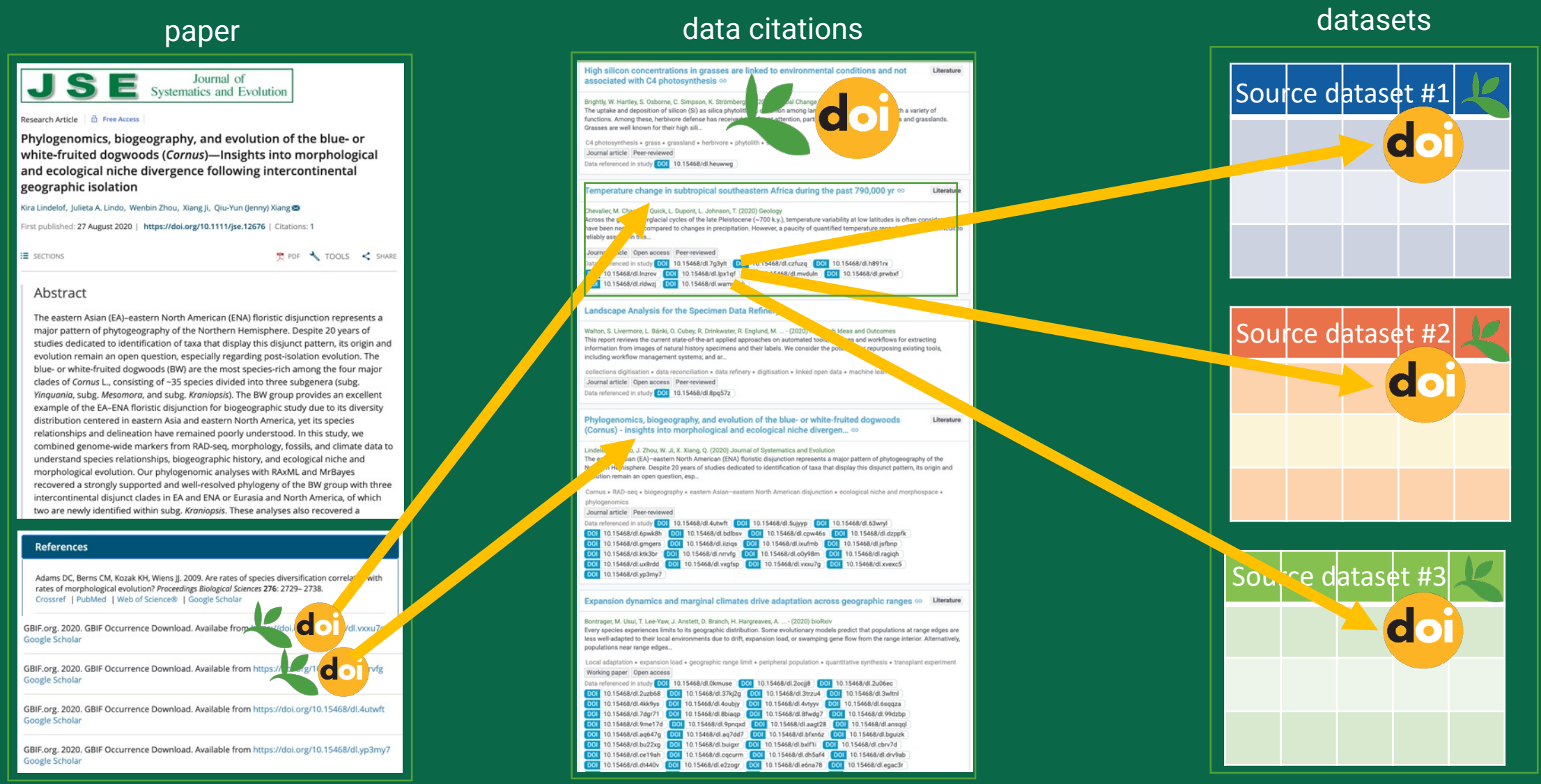


NTNU Vascular plants: https://doi.org/10.15468/zrlqok

#CiteTheDOI

GBIF

DOI BASED DATA CITATION AT GBIF.ORG -- #CITETHEDOI



#CiteTheDOI GBIF









Publish datasets in GBIF

Dataset DOIs

Source dataset #1

Source dataset #2

Source dataset #3

Filter & download

Download DOI

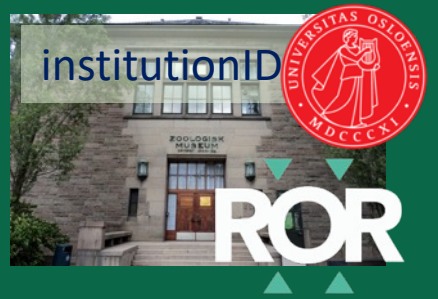
GBIF download

Process & archive

Final state of data

Analyze & publish

Bibliographic DOI









Publish datasets in GBIF

Dataset DOIs

Source dataset #1

Source dataset #2

Source dataset #3

Filter & download

Download DOI

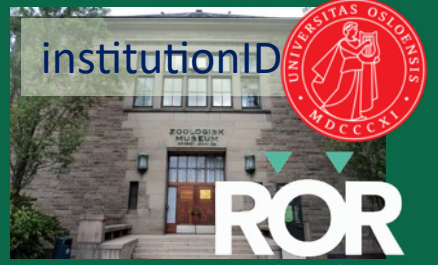
GBIF download

Process & archive

Final state of data

Analyze & publish

Bibliographic DOI









Publish datasets in GBIF

Dataset DOIs

Source dataset #1 doi
UUID
UUID

Source dataset #2 doi
UUID

Source dataset #3 doi
UUID

Filter & download

Download DOI

GBIF download doi
UUID
UUID
UUID
UUID

Process & archive

Final state of data
UUID
zenodo doi

Analyze & publish

Bibliographic DOI











Dataset DOIs

Source dataset #1 doi

Source dataset #2 doi

Source dataset #3 doi
iD ROR UUID WIKIDATA

Publish datasets in GBIF

Filter & download

Download DOI

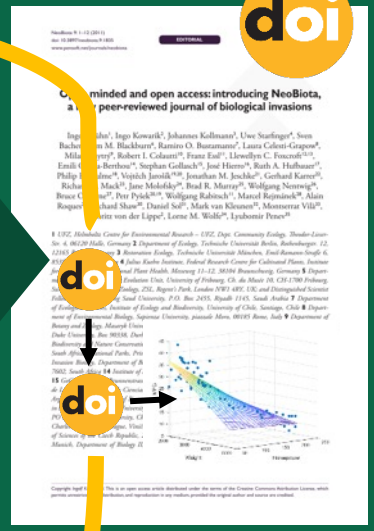
GBIF download doi
ROR WIKIDATA UUID iD

Process & archive

Final state of data
UUID iD
zenodo doi

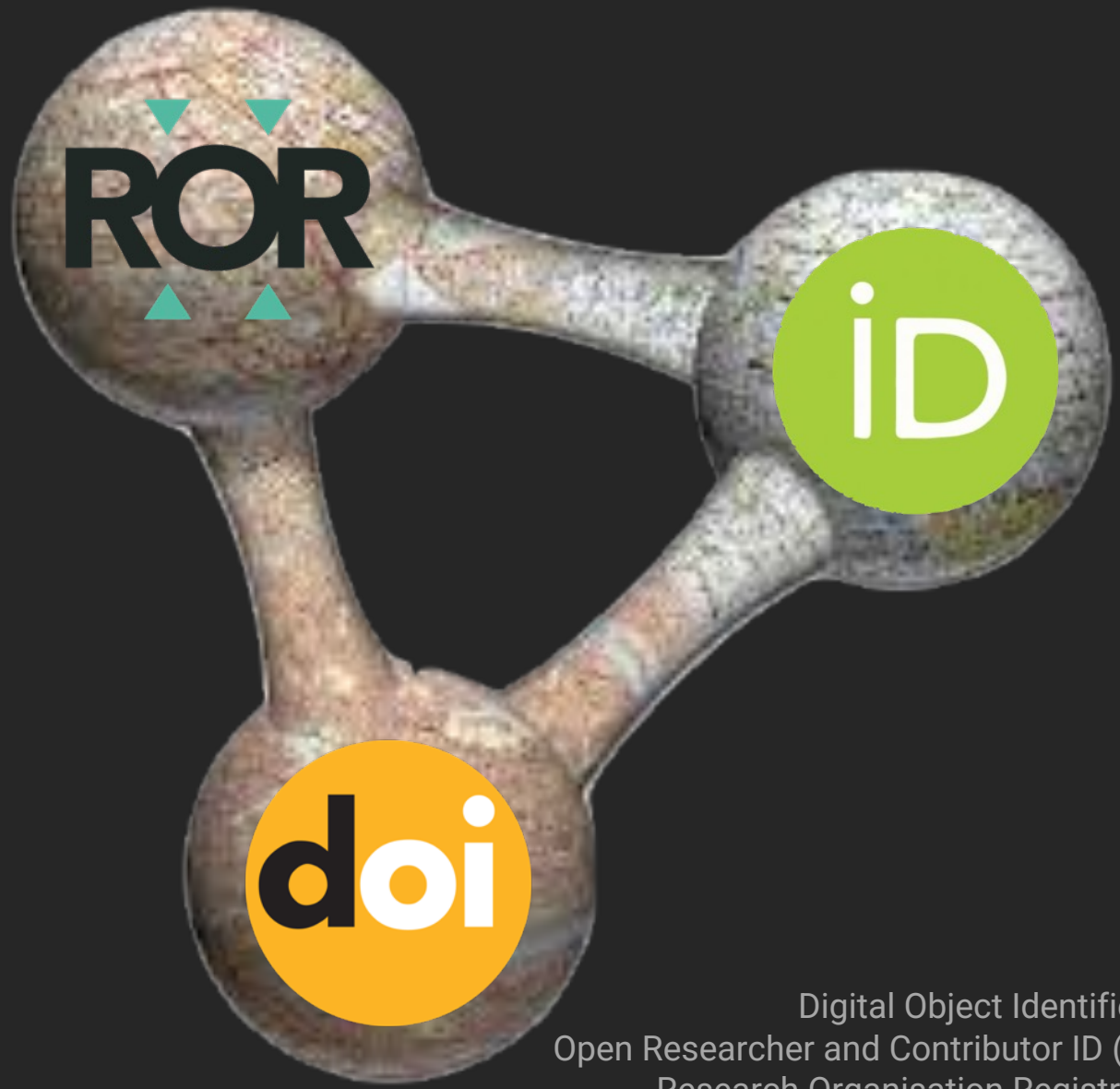
Analyze & publish

Bibliographic DOI



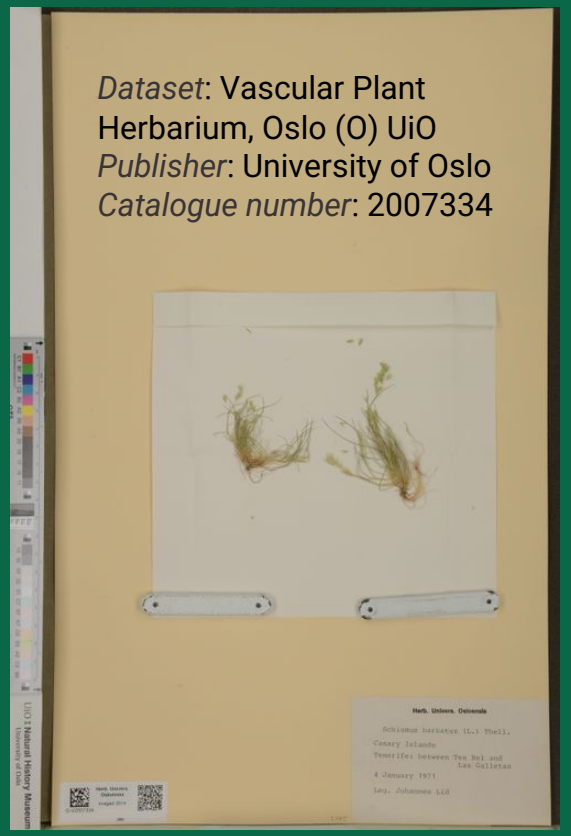
ROR for institutions
ORCID for people
DOI for datasets
GRSciColl UUID for collections

will enable the linking of museum collection specimens to scientific litterature and scientific actors (authors, curators, etc)



Digital Object Identifier (DOI)
Open Researcher and Contributor ID (ORCID)
Research Organisation Registry (ROR)
Universally Unique Identifier (UUID)

GBIF & IDENTIFIERS



Dataset: Vascular Plant Herbarium, Oslo (O) UiO
Publisher: University of Oslo
Catalogue number: 2007334

https://www.gbif.org/occurrence/1095052193

Get data How-to Tools Community About dag.endresen

Record

Basis of record	Preserved specimen
Collection code	V
Dataset name	Macaronesia
Institution code	O University of Oslo
Institution ID	https://ror.org/01xtthb56 University of Oslo

Occurrence

Catalogue number	2007334
Occurrence ID	urn:catalog:O:V:2007334
Occurrence status	PRESENT
Other catalogue numbers	0574816d-3d99-41b8-b3b8-c6035de0e929
Recorded by	Johannes Lid
Recorded by ID	https://www.wikidata.org/wiki/Q94522

Event

Day	4
Month	1
Year	1971
Event date	1971-01-04T00:00:00

Identification

Date identified	1971-01-04T00:00:00
Identified by	Johannes Lid
Identified by ID	https://www.wikidata.org/wiki/Q94522

Taxon

Kingdom	Plantae
Phylum	Tracheophyta
Class	Liliopsida
Order	Poales
Family	Poaceae
Genus	Schismus
Specific epithet	barbatus
Infraspecific epithet	
Scientific name	Schismus barbatus (L.) Thell.
Rank	Species
Taxonomic status	Accepted

Location

Country or area	Spain
Country code	ES
Locality	Tenerife:Tenerife: between Ten Bel and Las Galletas.

Other

Identifier	urn:catalog:O:V:2007334
Record licence	http://creativeco.../by/4.0/legalcode
Modified	2015-05-04T00:00:00.000+00:00

MACHINE-READABILITY REQUIRES PERSISTENT IDENTIFIERS

The purpose of identifiers is
… to name things
… making it possible to refer to them



- To uniquely identify something it needs a **persistent identifier**, a PID.
- A Persistent Identifier is **globally unique**, **persistent**, and **resolvable**“.
- A PID is resolvable when it allows both **human and machine** users to access an object or its representation, and its *Kernel* Information.
- **Kernel Information** is a structured record that contains information (metadata) about the referred object, such as a pointer to the location where the data for the object can be found.



Global Registry of Scientific Collections

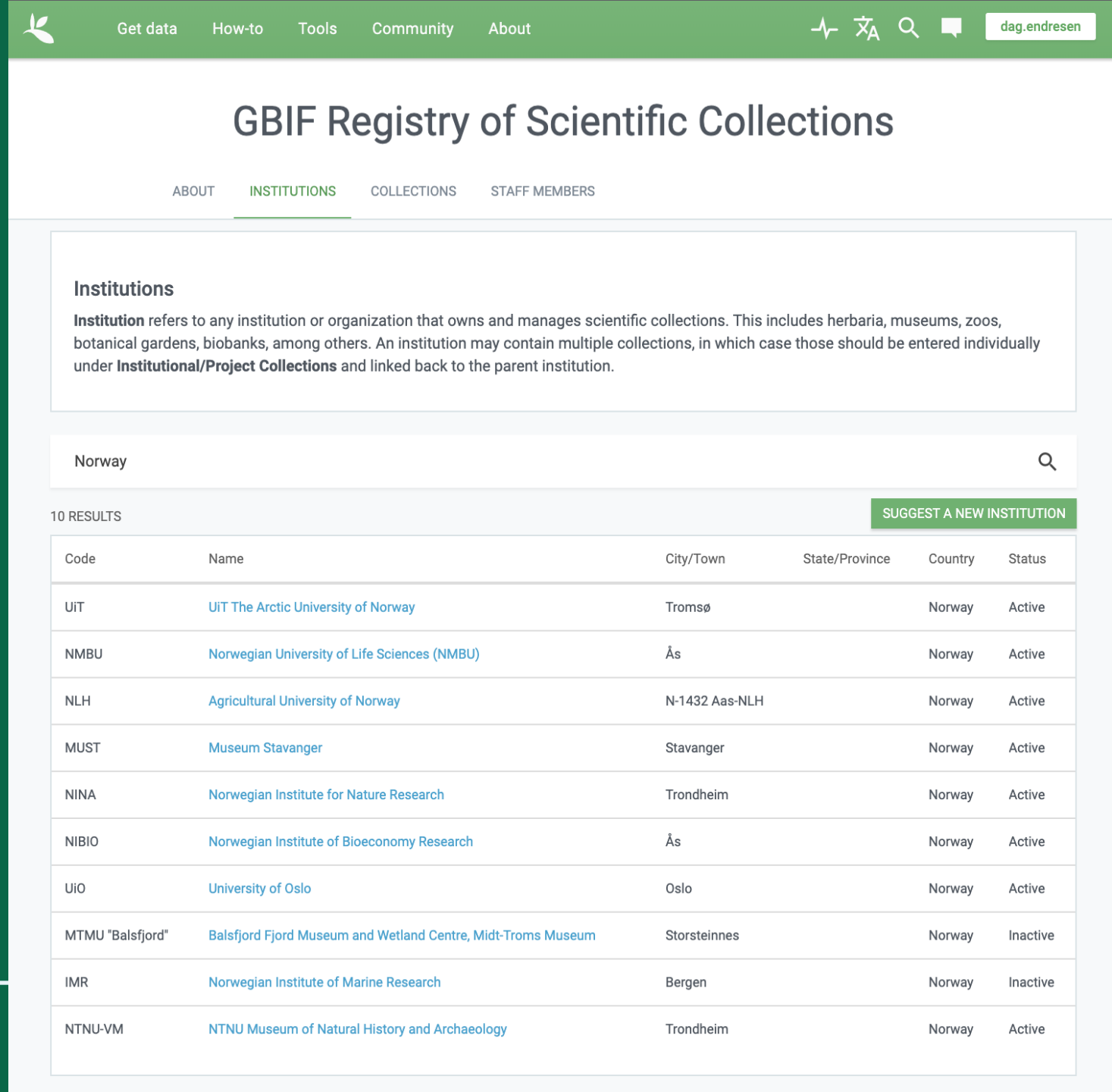
GRSciColl was established at Smithsonian in 2013
Hosting of GRSciColl was transferred to GBIF in 2019

GRSCICOLL

- The Global Registry of Scientific Collections (GRSciColl) was a community-curated clearing house of colletions information developed by the Consortium of the Barcode of Life (CBOL) – launched in 2013.

- Hosting the GRSciColl was transferred to GBIF in 2018 and the upgraded portal came back online in 2019.

https://www.gbif.org/grscicoll/institution/search?q=Norway

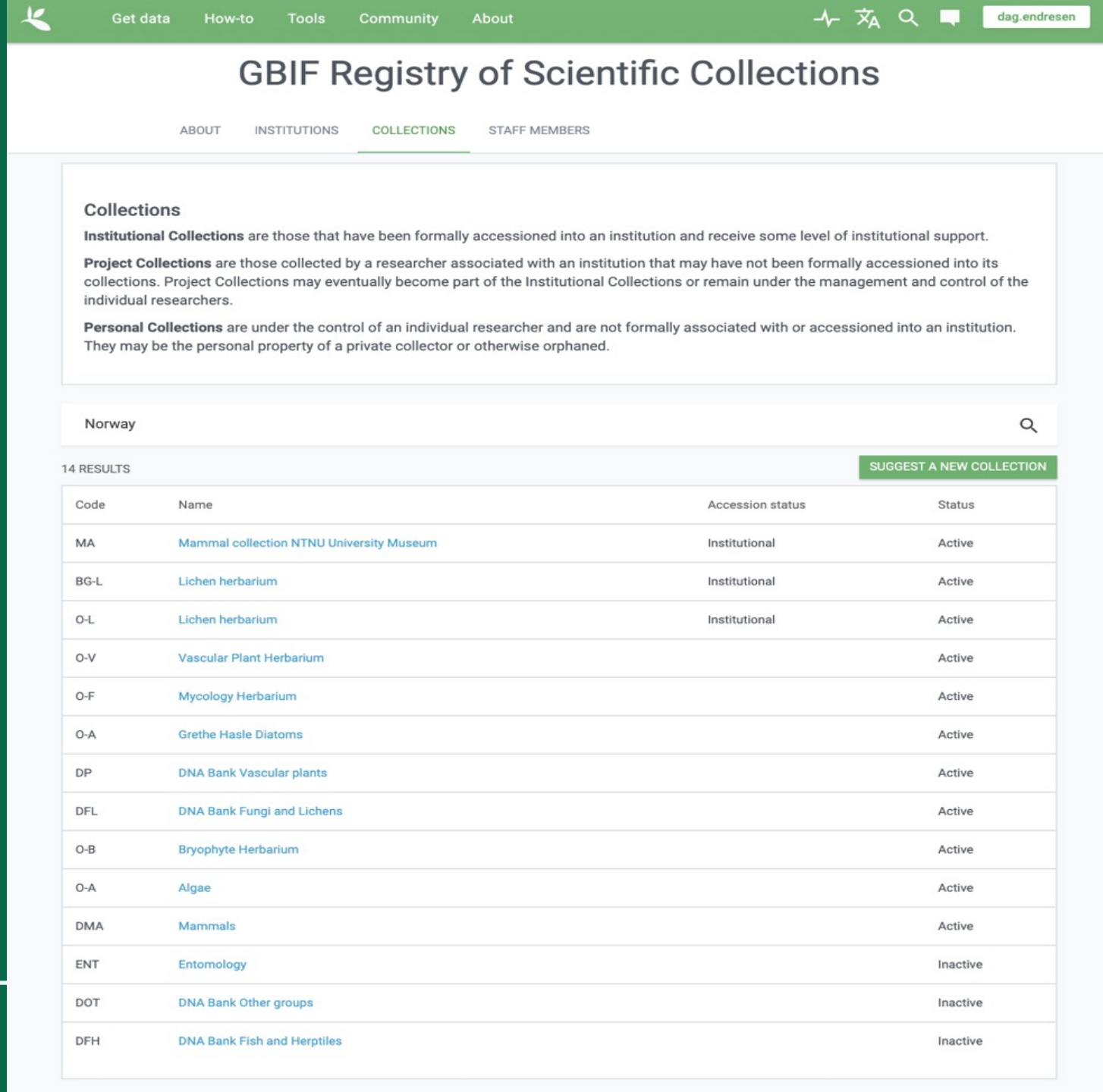


GRSCICOLL

- 8 117 collection-holding institutions
- 7 010 scientific collections
- 17 000 collection staff members

- https://www.gbif.org/grscicoll

https://www.gbif.org/grscicoll/collection/search?q=norway

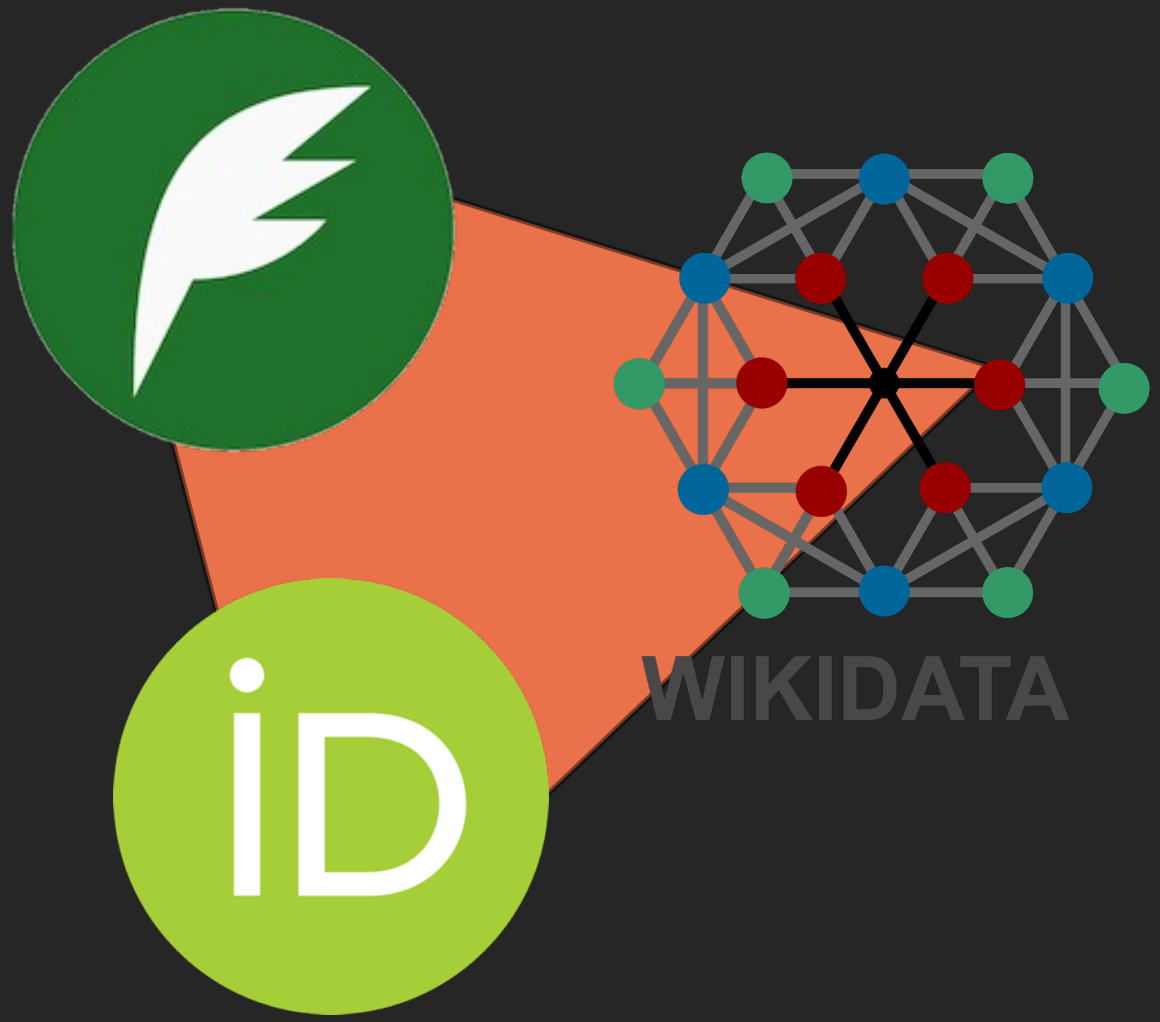


INSTITUTION IDENTIFIERS FOR SOME OF THE MUSEUMS IN NORWAY

Museums	ROR ID	Wikidata	GRSciColl	GBIF publisher institution
Universitetet I Agder (UiA) Agder naturmuseum og botaniske hage	03x297z98 --	Q3375341 Q3375341	UiA KMN	826d1920-7f5c-4091-a84e-668aa2e35b61 --
Universitetet I Bergen University Museum of Bergen Institutt for naturhistorie	03zga2b32 -- --	Q204457 Q301787 Q11990981	UiB -- --	3f3967bf-ecc7-4455-ba89-4e0ab6d6fd3c -- --
Helgeland Museum Rana Museum	02gyhy076 --	Q11057676 Q11997066	Helgeland --	a030b53d-7ab0-41dc-8ca7-77f65d5c8157 --
Midt-Troms Museum Balsfjord Fjordmuseum og Våtmarkssenter	-- --	Q12327078 Q105533121	-- MTMU Bjalsfjord	-- 9e8e7946-cd17-4c58-81c1-dc8bef359360
Museum Stavanger (MUST)	03bq5ar94	Q19382034	MUST	ecc5cd9e-2d25-4b8d-89c8-a0711eee813b
Universitetet I Oslo (UiO) Naturhistorisk museum i Oslo (NHMO) Botanisk museum (Oslo herbarium - O)	01xtthb56 -- --	Q486156 Q1840963 Q2036576	UiO NHMO O	f314b0b0-e3dc-11d9-8d81-b8a03c50a862 -- --
Randsfjordmuseene AS	--	Q11997108	--	--
Universitetet I Tromsø (UiT) Tromsø Museum, Universitetsmuseet	00wge5k78 --	Q279724 Q1686510	UiT TROM	689b40c4-ff31-4cd0-83a5-a7a828f1cd92 --
Varanger Museum	--	Q12009007	--	--
Norges Teknisk Naturvitenskapelige Universitet (NTNU) Vitenskapsmuseet	05xg72x27 --	Q314536 Q1770886	TRH NTNU-VM	a8144f37-5ff7-4137-9400-94b5b2ea4ec4 --

GBIF

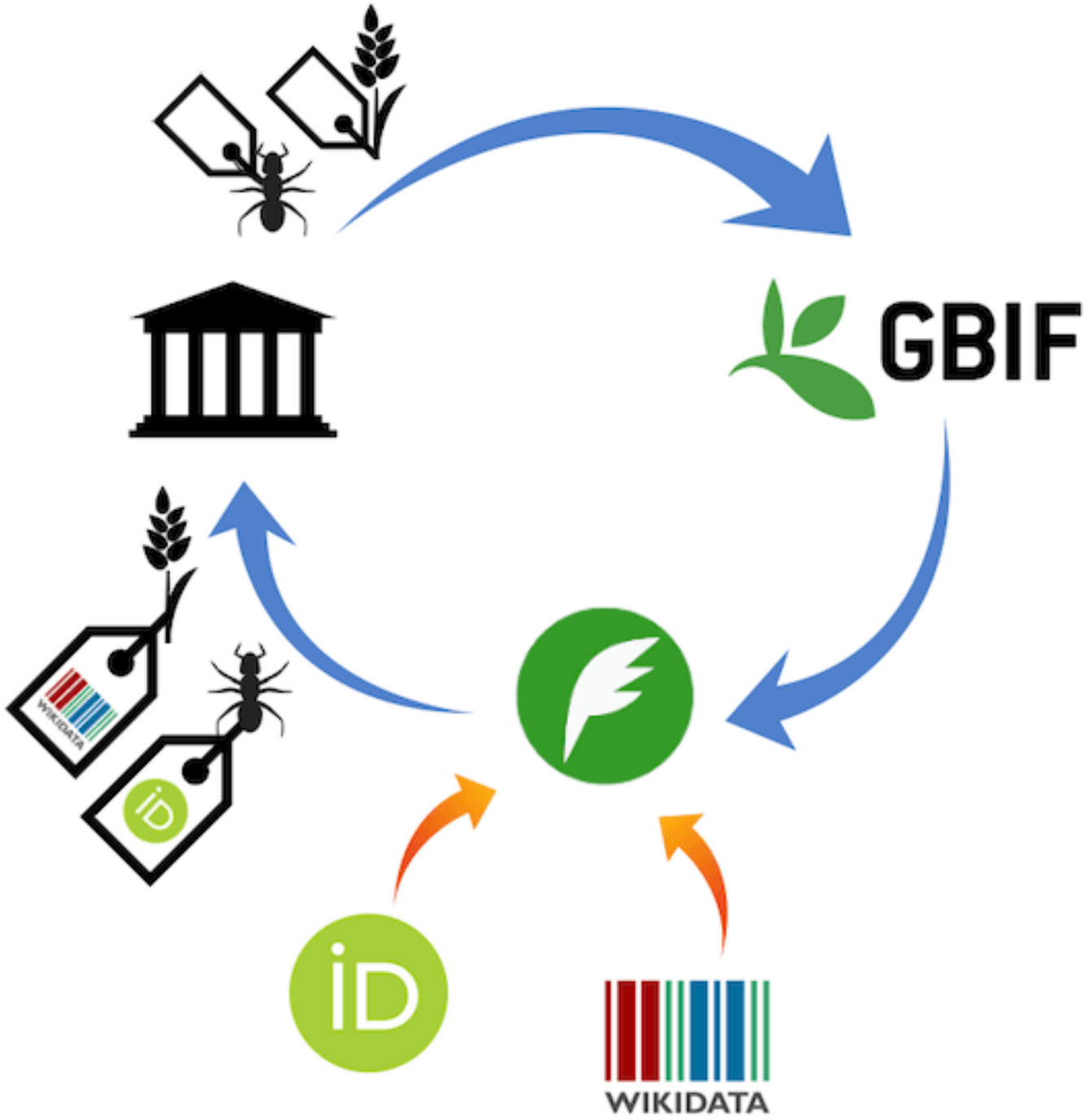
new tools



BIONOMIA

Link natural history specimens
to the world's collectors.

https://bionomia.net



Occurrences 3

Recorded by

Recorded by ID

https://www.wikidata.org/wiki/Q94522

Search

CLEAR

Identified by

Identified by ID

https://www.wikidata.org/wiki/Q94522

Search

CLEAR

Record number

Occurrence id

Organism id

Publishing country or area

Elevation

Depth

Locality

Water body

State province

Repatriated

Is in cluster

DWCA extension

SEARCH OCCURRENCES | 24,923 RESULTS

TABLE | GALLERY | MAP | TAXONOMY | METRICS | DOWNLOAD

Scientific name	Country or area	Coordinates	Month & year	Basis of record	Dataset
Sphaerophorus fragilis (L.) Pers.	Norway	60.6N, 6.7E	1975 August	Preserved specimen	Lichen herbarium, Oslo (O) UiO
Atriplex 1758	Spain		1971 January		o (O) UiO
Schismus barbatus (L.) Thell.	Spain		1971 January		o (O) UiO
Amaranthus L.	Spain		1971 January		o (O) UiO
Setaria Beauv.	Spain	28.3N, 16.5W	1971 January		o (O) UiO
Amaranthus L.	Spain		1971 January		o (O) UiO
Spergularia (Pers.) J. & K.Presl	Spain		1971 January		o (O) UiO
Salix phylicifoli			ust		o (O) UiO
Beta pate			uary		o (O) UiO
Bromus rigidus			uary		o (O) UiO
Asparagus past			uary		o (O) UiO
Cheilanthes ma			uary		o (O) UiO
Spergula fallax			uary		o (O) UiO
Schismus barb			uary		o (O) UiO
Asphodelus fist			uary		o (O) UiO
Polypodium serratum Milde	Spain	28.2N, 15.6W	1970 January		o (O) UiO
Sporobolus indicus R.Br.	Spain	28.1N, 15.5W	1970 January		o (O) UiO
Trisetaria pumila (Desf.) Maire	Spain	28.1N, 15.5W	1970 January		o (O) UiO
Schismus barbatus (L.) Thell.	Spain		1970 January	Preserved specimen	Vascular Plant Herbarium, Oslo (O) UiO

Occurrence ID	**urn:catalog:O:V:2007334**
Catalogue number	2007334
Material Sample ID	**urn:uuid:0574816d-3d99-41b8-b3b8-c6035de0e929**
Other catalogue numbers	
Event date	1971-01-04
Recorded by	Johannes Lid
Recorded by ID	http://www.wikidata.org/entity/**Q94522**
Date identified	1971-01-04T00:00:00
Identified by	Johannes Lid
Identified by ID	http://www.wikidata.org/entity/**Q94522**



Findable Accessible Interoperable Reusable



FAIR data is about ***machine-readable*** data

*researchers & museums need to do more than simply post their data on the web for it to be **re-usable**.*

FAIR DATA PRINCIPLES

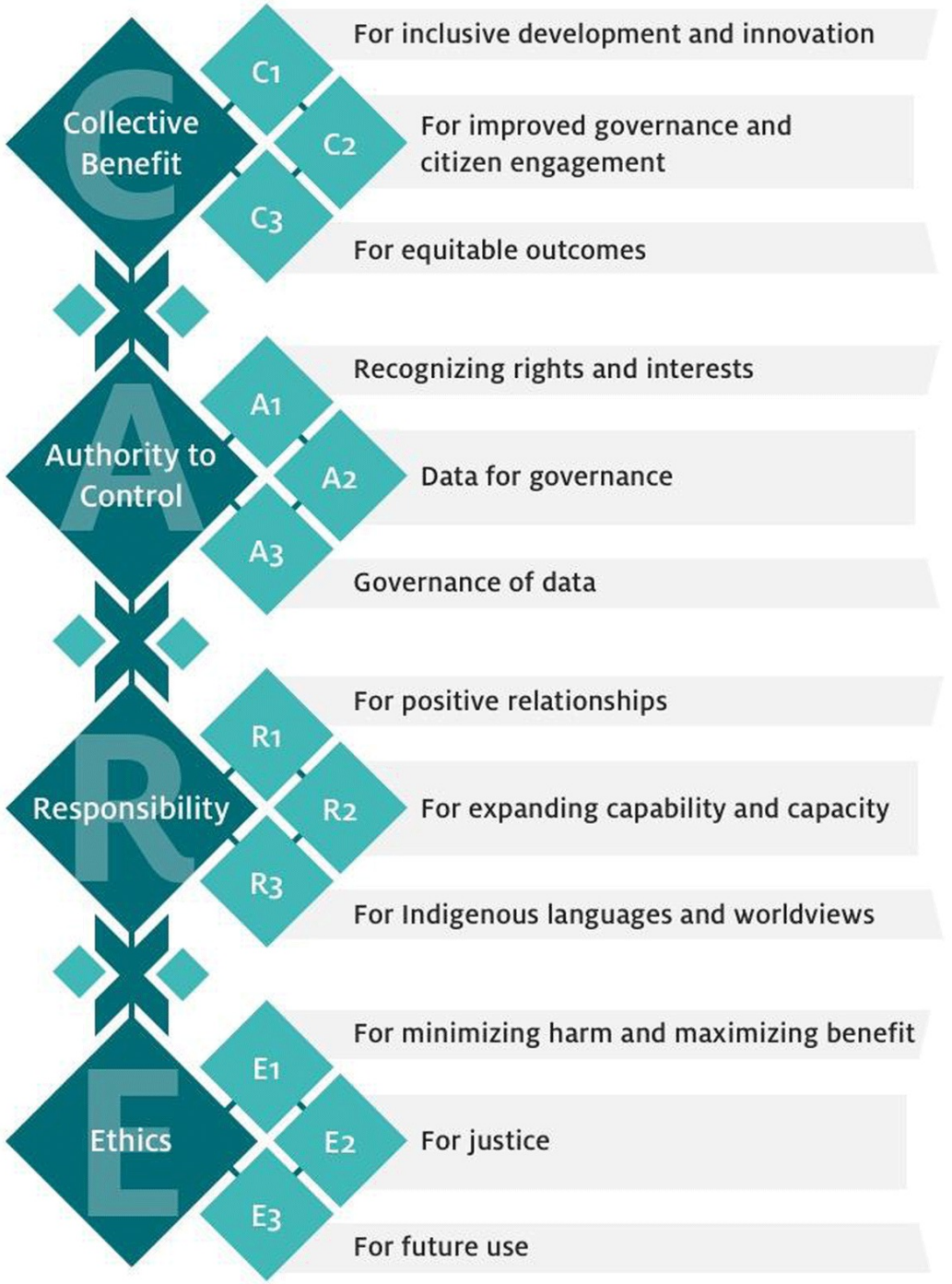



Be
FAIR
Findable
Accessible
Interoperable
Reusable
and
CARE
Collective Benefit
Authority to Control
Responsibility
Ethics
https://www.gida-global.org/care

CARE DATA PRINCIPLES

- CARE is an acronym which stands for:
  - **Collective** Benefit,
  - **Authority** to Control,
  - **Responsibility**,
  - **Ethics**
- CARE was created by the International Indigenous Data Sovereignty Interest Group, a group that is a part of the Research Data Alliance.
- Resolve Indigenous Peoples' rights to and interests in their data across the data lifecycle.
- Building on **FAIR** data principles while ensuring indigenous interests.

https://www.gida-global.org/care

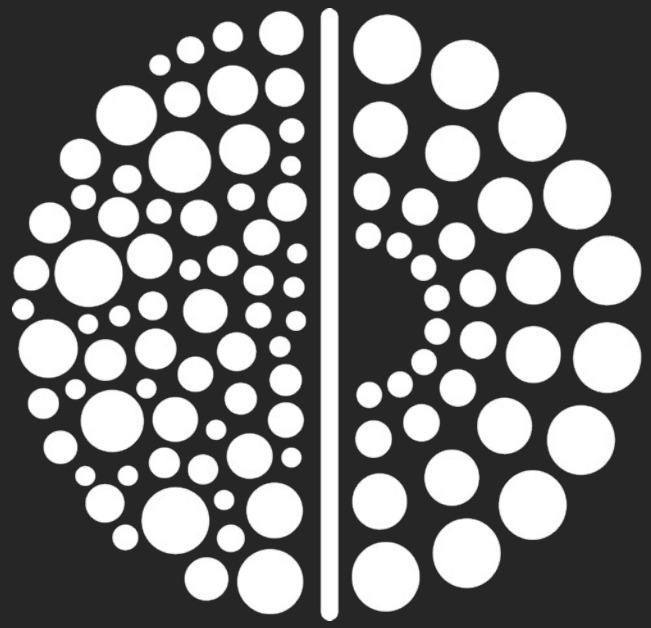




Biodiversity Informatics

Open Science

new possibilities for novel curiosity-driven research



Biodiversity Information Standards

TDWG

Darwin Core

DARWIN CORE

- Darwin Core is a standard maintained by the TDWG Darwin Core Maintenance Interest Group. It includes a **glossary of terms** (properties, elements, fields, columns, attributes, or concepts) intended to facilitate the sharing of information about biological diversity by providing **identifiers**, **labels**, and **definitions**. Darwin Core is primarily based on taxa, their occurrence in nature as documented by observations, specimens, samples, and related information.

- Quick reference guide at https://dwc.tdwg.org/terms/
- Darwin Core Archive (text guide) at https://dwc.tdwg.org/text/
- Detailed list of DwC terms at https://dwc.tdwg.org/list/
- DwC is maintained in a GitHub repository : https://github.com/tdwg/dwc
- Darwin Core Q&A at: https://github.com/tdwg/dwc-qa



Biodiversity Information Standards

TDWG

DARWIN CORE CLASSES

- Record-level
- Occurrence *
- Organism
- MaterialSample
- Event *
- Location
- GeologicalContext
- Identification
- Taxon *

- MeasurementOrFact
- ResourceRelationship

Asterix (*) are supported as “Cores” in GBIF
https://rs.gbif.org/core/

Most of the other classes available as “Extensions” in GBIF, meaning additional properties to the ”Core” records”
https://rs.gbif.org/extension/



Biodiversity Information Standards

TDWG

https://dwc.tdwg.org/terms/

GBIF

DARWIN CORE

TDWG Home Terms Guides

Darwin Core quick reference guide

This document is intended to be an easy-to-read reference of the currently (as of 2021-07-15) recommended terms maintained as part of the Darwin Core standard and is maintained by the Darwin Core Maintenance Group.

This page is not part of the standard, but combines the normative term names and definitions with the non-normative comments and examples that are meant to help people to use the terms consistently. Capitalized terms, such as Occurrence, are Darwin Core classes, which are special category terms used to group sets of terms for convenience. Comprehensive metadata for current and obsolete terms in human readable form are found in the document List of Darwin Core terms.

Additional files with just the current term names and a file with the full term history can be found in the Darwin Core repository.

For inquiries about how to use Darwin Core, either enter an issue in the Darwin Core Questions & Answers site or enter an issue in the alternative form, which will have the same effect. See the bottom of this document for how to cite Darwin Core.

Record-level

This category contains terms that are generic in that they might apply to any type of record in a dataset.

type | modified | language | license | rightsHolder | accessRights | bibliographicCitation | references | institutionID | collectionID | datasetID | institutionCode | collectionCode | datasetName | ownerInstitutionCode | basisOfRecord | informationWithheld | dataGeneralizations | dynamicProperties

type	Property
Identifier	http://purl.org/dc/elements/1.1/type
Definition	The nature or genre of the resource.
Comments	Must be populated with a value from the DCMI type vocabulary (http://dublincore.org/documents/2010/10/11/dcmi-type-vocabulary/).
Examples	`StillImage`, `MovingImage`, `Sound`, `PhysicalObject`, `Event`, `Text`

Record-level
Occurrence
Organism
MaterialSample
Event
Location
GeologicalContext
Identification
Taxon
MeasurementOrFact
ResourceRelationship
UseWithIRI

LivingSpecimen
PreservedSpecimen
FossilSpecimen
MaterialCitation
HumanObservation
MachineObservation

Biodiversity Information Standards TDWG

https://dwc.tdwg.org/terms/

GBIF

OCCURRENCEID

TDWG Home Terms Guides Namespace policy

Occurrence

occurrenceID | catalogNumber | recordNumber | recordedBy | individualCount | organismQuantity | organismQuantityType | sex | lifeStage | reproductiveCondition | behavior | establishmentMeans | occurrenceStatus | preparations | disposition | associatedMedia | associatedReferences | associatedSequences | associatedTaxa | otherCatalogNumbers | occurrenceRemarks

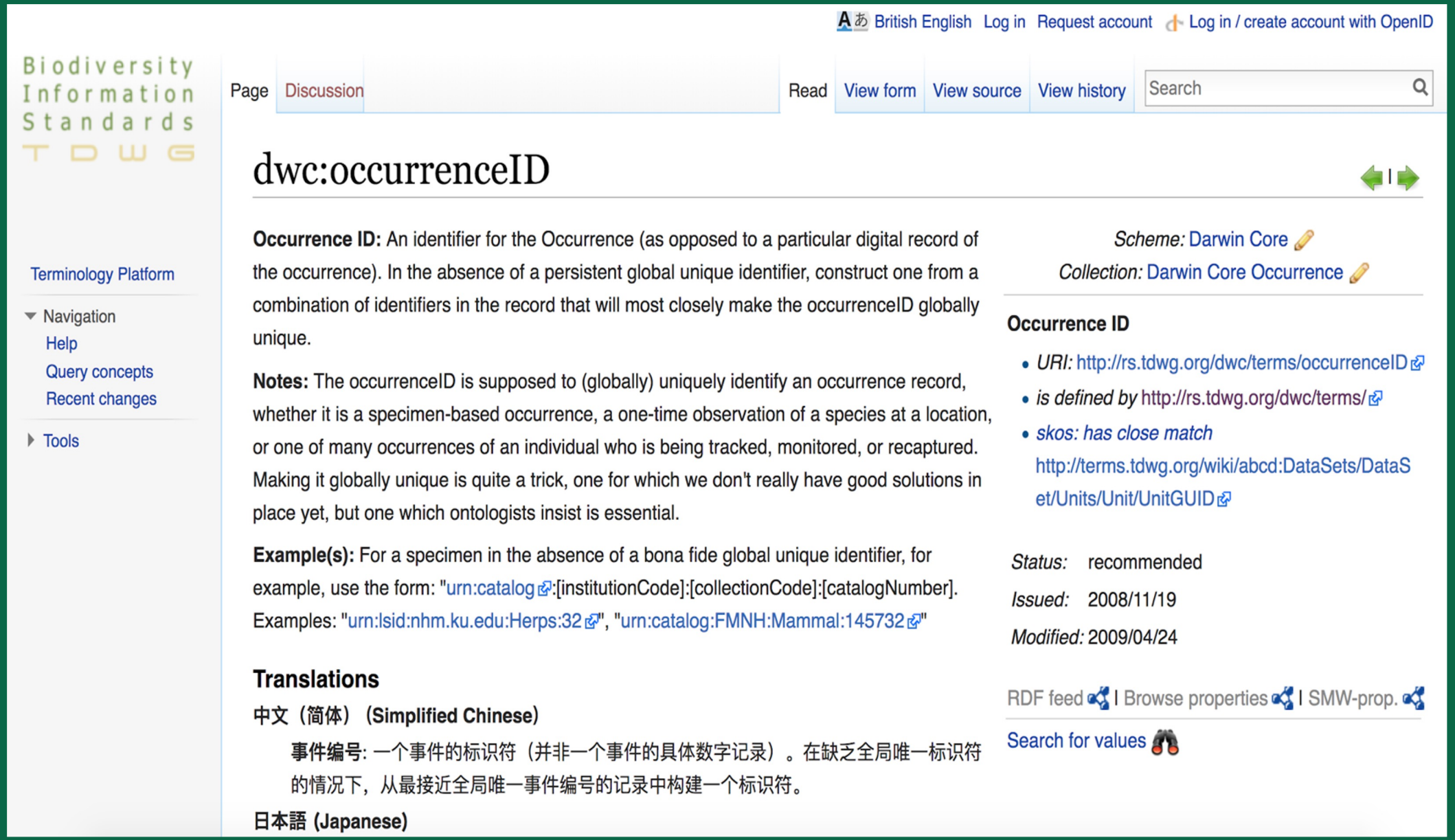
occurrenceID	Property
Identifier	http://rs.tdwg.org/dwc/terms/occurrenceID
Definition	An identifier for the Occurrence (as opposed to a particular digital record of the occurrence). In the absence of a persistent global unique identifier, construct one from a combination of identifiers in the record that will most closely make the occurrenceID globally unique.
Comments	Recommended best practice is to use a persistent, globally unique identifier.
Examples	http://arctos.database.museum/guid/MSB:Mamm:233627, 000866d2-c177-4648-a200-ead4007051b9, urn:catalog:UWBM:Bird:89776

Record-level
Occurrence
Organism
MaterialSample
Event
Location
GeologicalContext
Identification
Taxon
MeasurementOrFact
ResourceRelationship
UseWithIRI

LivingSpecimen
PreservedSpecimen
FossilSpecimen
HumanObservation
MachineObservation

GBIF

TDWG TERMS WIKI



https://terms.tdwg.org/wiki/dwc:occurrenceID

GBIF

DARWIN CORE ARCHIVE

Biodiversity Information Standards TDWG



EML
metadata



Core

- Occurrence
- Event
- Taxon







Extensions

- Identifications
- Multimedia resources
- MeasurementOrFact
- ResourceRelationship
- DNA-derived (MIxS, GGBN)
- … and more

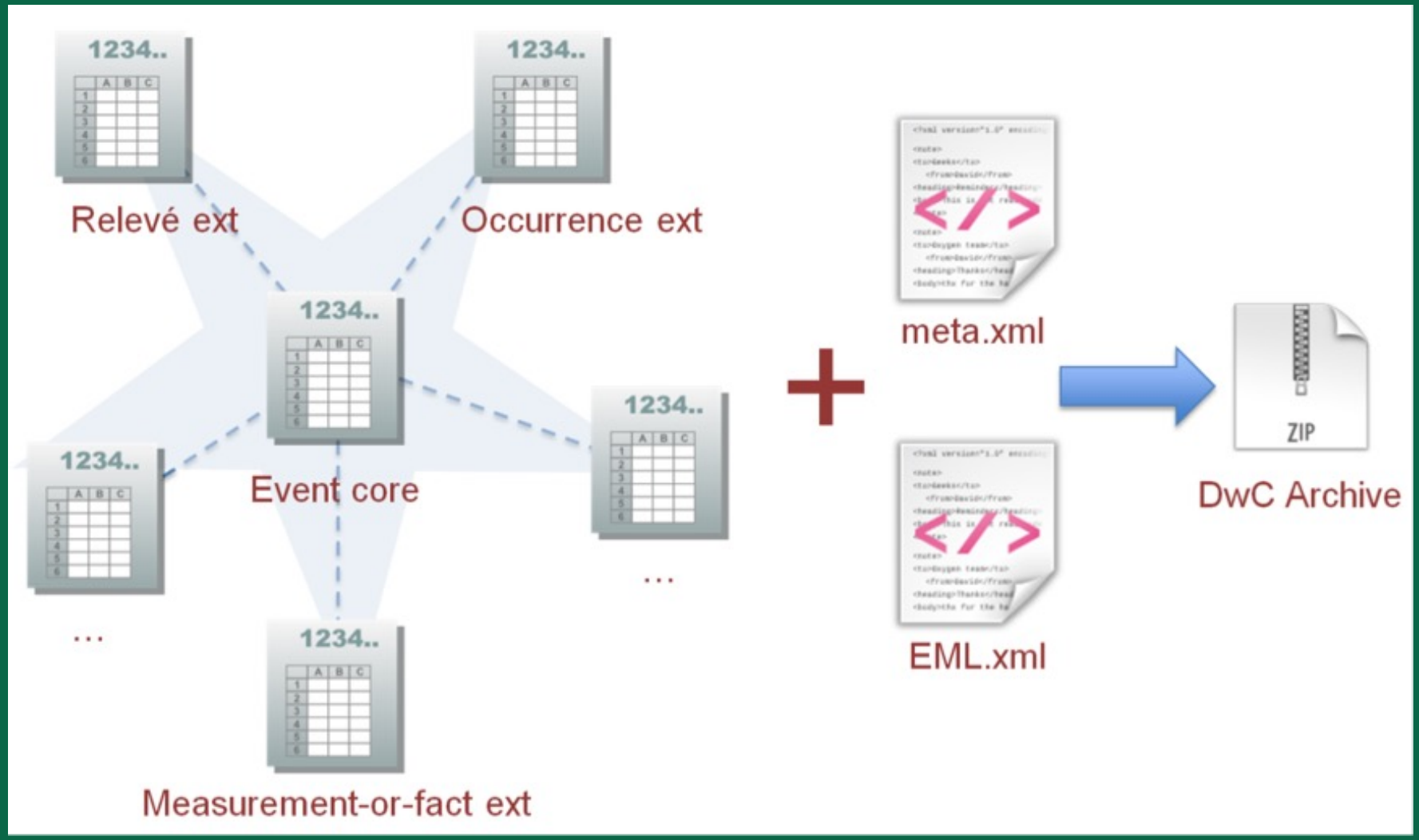


Darwin
Core
Archive

https://dwc.tdwg.org/text/ | https://rs.gbif.org/core/ | https://rs.gbif.org/extension/ |

GBIF

DARWIN CORE ARCHIVE – STAR SCHEMA



https://ipt.gbif.org/manual/en/ipt/2.6/dwca-guide

GBIF

GBIF DATA QUALITY REQUIREMENTS

The items listed constitute the minimum formal requirements for publishing an occurrence dataset. GBIF.org will not accept a dataset without these terms and will not index the records.

While these items are mandatory for publishing the dataset at all, they are only the starting point. The usefulness of the published data will still be severely limited unless additional information is supplied.

https://www.gbif.org/data-quality-requirements-occurrences

Darwin Core records

Term	Status
occurrenceID	Required
basisOfRecord	Required
scientificName	Required
eventDate	Required
countryCode	Strongly recommended
taxonRank	Strongly recommended
kingdom	Strongly recommended
decimalLatitude & decimalLongitude	Strongly recommended
geodeticDatum	Strongly recommended
coordinateUncertaintyInMeters	Strongly recommended
individualCount, organismQuantity & organismQuantityType	Strongly recommended
informationWithheld	Share if available
dataGeneralizations	Share if available
eventTime	Share if available
country	Share if available

DISCUSSION – DARWIN CORE

- What types of data can you currently publish in GBIF?
- What GBIF DwC-A Core type would you choose?
    - Organismal species occurrences *in situ*?
    - Natural History organismal collection specimens?
    - Natural History mineralolgy collection specimens?
    - Natural History paleontology collection specimens?
    - Cultural history antropogenic artefact specimens?
    - Undigitized natural history collection metadata?
    - Invasive species distribution occurrences?
    - Species monitoring data?

- What is your ***favourite*** Darwin Core class? – and Darwin Core term?
- Any ***least*** favourite class & term?
- Any particular class, term, or concept you ***miss*** in Darwin Core?



Biodiversity Information Standards

TDWG

https://dwc.tdwg.org/terms/

GBIF

THANK YOU

GBIF Norway

Dag Endresen | Node Manager

helpdesk@gbif.no

