Global Biodiversity Information Facility & the NHM MaNa research group

Dag Endresen

MANA \ Machine Readable Nature

GBIF Global Biodiversity Information Facility

Illustrasjon: GBIF data portal

UiO dScience visit to Tøyen, Oslo | 20.th June 2024

MACHINE-READABLE NATURE RESEARCH GROUP - MANA

Research group for Machine-readable Nature works with research data infrastructures at the Natural History Museum in Oslo

Including the Norwegian participant nodes of GBIF (OECD) and DiSSCo (ESFRI)

Our key research questions focus on machine-readable FAIR research data

Established in 2023

MANA

Machine Readable Nature

https://www.nhm.uio.no/english/research/groups/mana/index.html

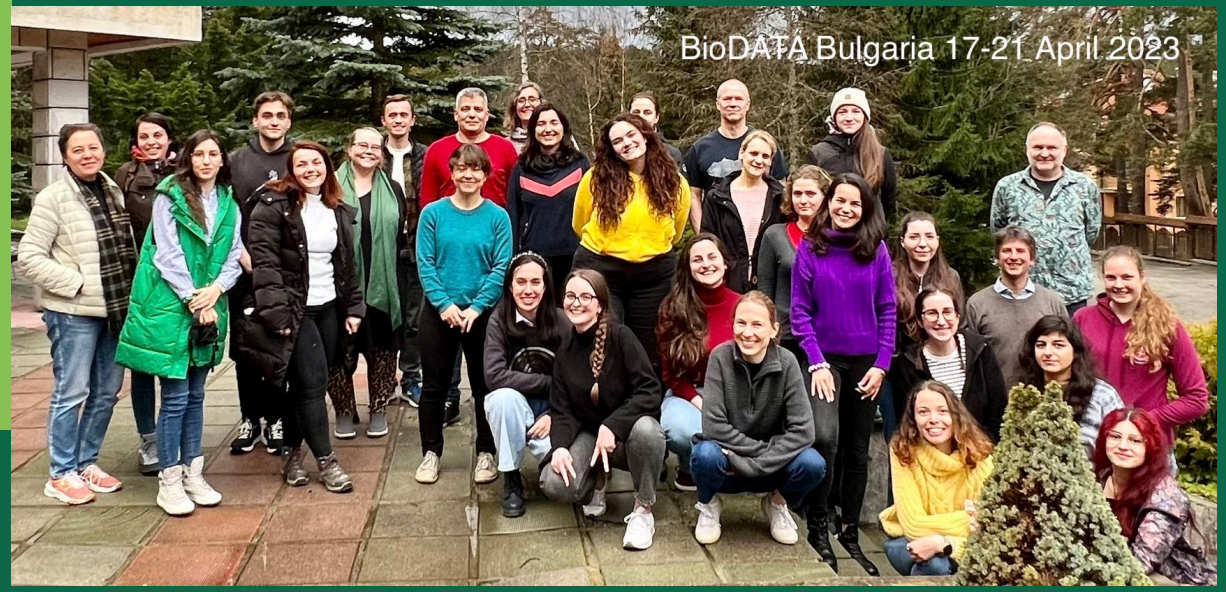
GBIF

MANA RESEARCH INFRASTRUCTURE PROJECTS

- GBIF Norway - National participant node of the Global Biodiversity Information Facility (since 2005)
- BioDT - Biodiversity Digital Twin, Horizon Europe (2022-2025)
- BioMedData - FAIR data management with the life sciences (2021-2024) Coordinated by ELIXIR Norway
- BioDATA Advanced - Accelerating biodiversity research through DNA barcodes, collection, and observation data (2021-2024)
- UiO core RI - BioData Helpdesk (2024)

STUDENT COURSES, RESEARCHER TRAINING AND WORKSHOPS

- Data publishing workshop in Zagreb, Croatia, May 2024
- Data mobilisation seminar in Tromsø, May 2024
- Data publishing workshop in Slovakia, May 2024
- Herbarium data mobilisation in Uzbekistan, April 2024
- Digital Twin hackathon in Oslo, January 2024
- BioDATA student course in South Africa, November 2023
- Open Science course, Hjerkinn,, Nov 2023, & Finse Nov 2022
- Marine biodiversity data for Nansen Legacy, Oslo Sept 2023, & Tromsø Nov 2022, (…)
- Data publication workshops in Bergen Feb 2023 (, etc …)
- GBIF Data and Stats Forum, NMBU, Ås, March 2023 (, …)
- BioMedData ELIXIR data management (2020-2024)
- Etc …



https://www.gbif.no/events/ | https://www.gbif.no/projects/

BIODIVERSITY DIGITAL TWIN (BIODT.EU)

GBIF and GBIF Norway (UiO) participate in the Horizon Europe development of digital twins for biodiversity (2022-2025).

Including a digital twin for the crop wild relative genepool of grasspea (Lathyrus) and a digital twin for DNA-detected biodiversity in poorly known habitats.

Including identifying biodiversity infrastructure (ESFRI roadmap, EOSC) collaboration models.

22 partner institutions from 13 countries.

https://biodt.eu/use-cases

GBIF

BIODT PROJECT OBJECTIVES

LUMI

Objective 1

- Build and deploy pre-operational BioDT platform LUMI HPC in Finland, 4 use case groups

Objective 2

- Integration with RI platforms and workflows GBIF, eLTER, LifeWatch, DiSSCo

Objective 3

- Interoperability with European DT initiatives (including DestinE) and European Data Infrastructure

Species response to environmental change

- Biodiversity Dynamics
- Ecosystem Services

Dynamics and threats from and for species of policy concern

- Invasive Species

Genetically detected biodiversity

- Crop wild relatives and genetic resources for food security
- DNA detected biodiversity in cryptic habitats

Species interactions with each other and with humans

- Pollinators
- Disease Outbreaks

GBIF

GBIF

The Global Biodiversity Information Facility (GBIF) is an international network and data infrastructure providing open access to biodiversity data.

eLTER

eLTER

The Integrated European Long-Term Ecosystem (eLTER) focuses on critical zone and socio-ecological research.

LifeWatch ERIC

LifeWatch ERIC

LifeWatch ERIC is the e-Science European infrastructure for biodiversity & ecosystem research.

DiSSCo

DiSSCo
Distributed System of Scientific Collections

The Distributed System of Scientific Collections (DiSSCo) is a Research Infrastructure (RI) for Natural Science Collections.

22 partner institutions from 13 countries

GBIF

GLOBAL BIODIVERSITY INFORMATION FACILITY

GBIF.ORG | GBIF.NO

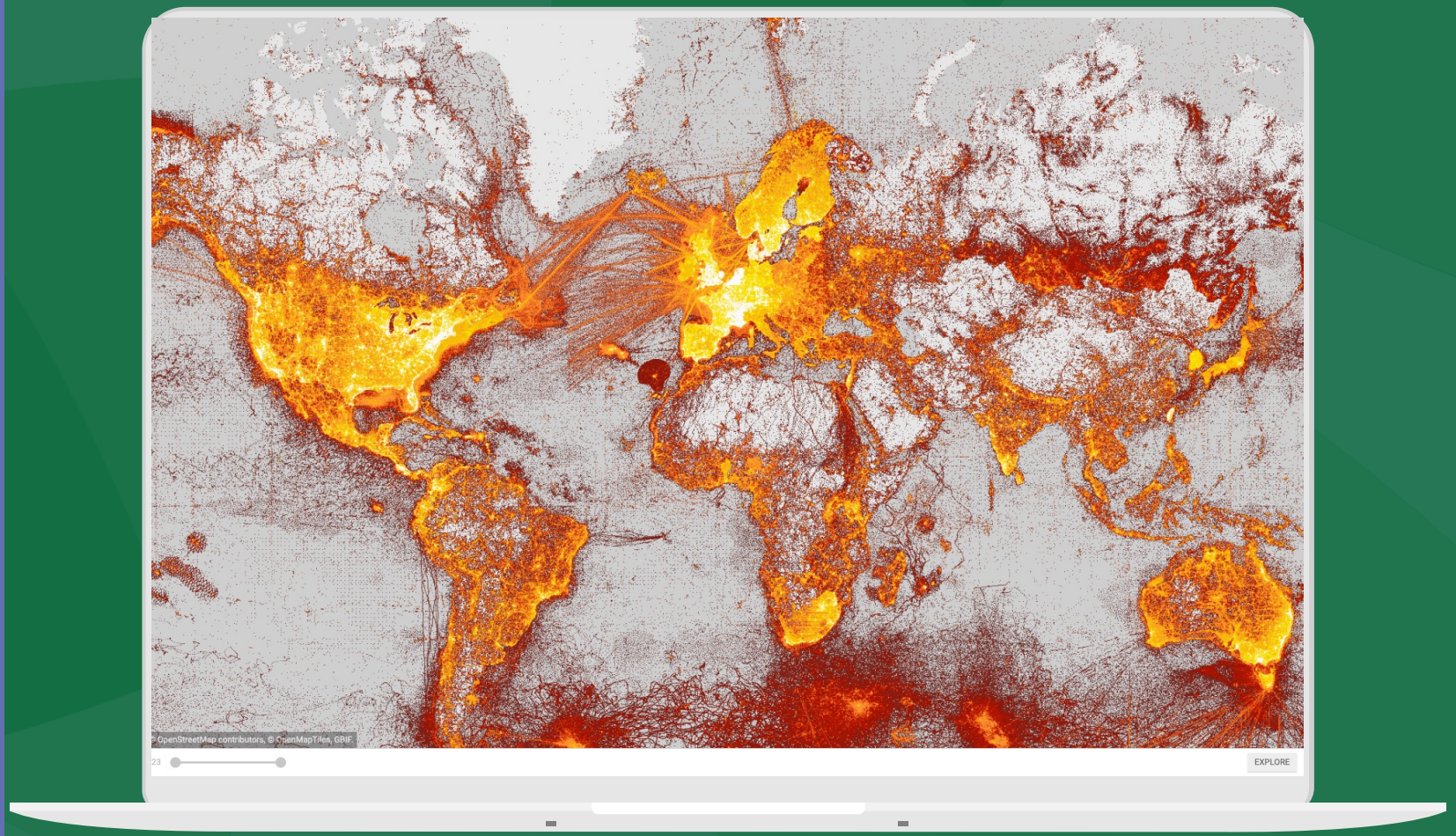
WHAT IS GBIF?

Intergovernmental network and research infrastructure

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

Distributed research data infrastructure based on **national participant NODES**

Secretariat in Copenhagen, Denmark

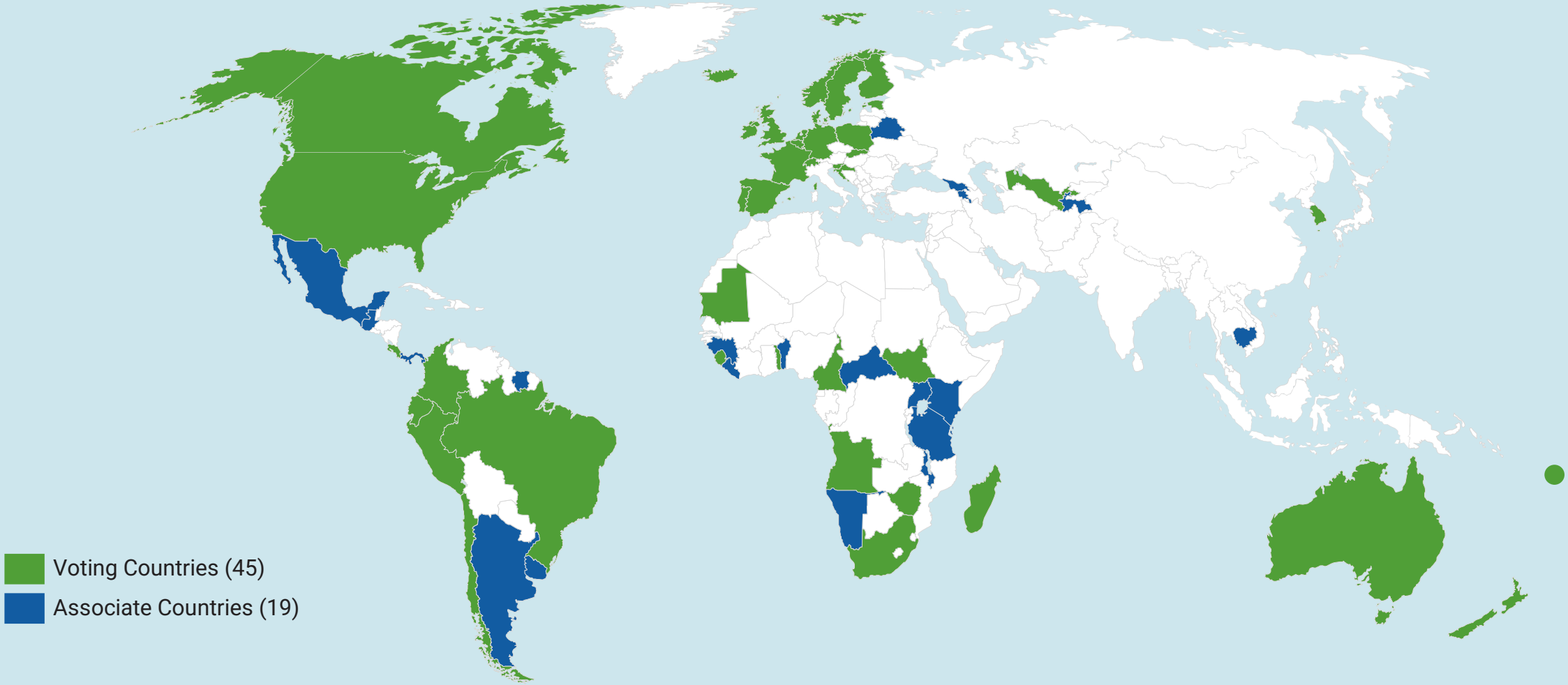


https://www.gbif.org

GBIF

GBIF PARTICIPANT COUNTRIES 20th June 2024

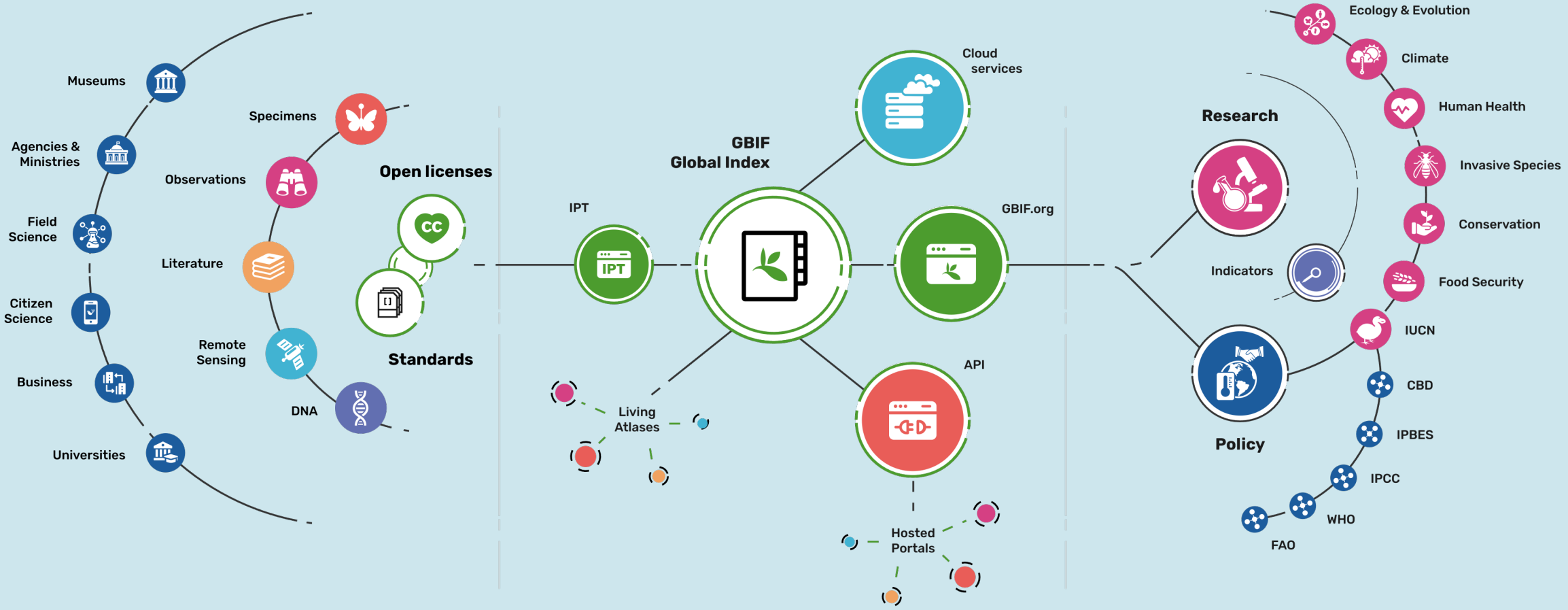
GBIF Norway (hosted by UiO NHMO)
GBIF member since 2005 (KD decision, RCN)



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

GBIF

PROVIDING BIODIVERSITY EVIDENCE FOR RESEARCH AND POLICY



GBIF NORWAY SERVICES: PUBLISH DATASETS WITH GBIF

National GBIF nodes offer support with professional data preparation & publication (data stewardship).

The GBIF community including the national GBIF nodes offer **training and capacity building** for open and FAIR data management.

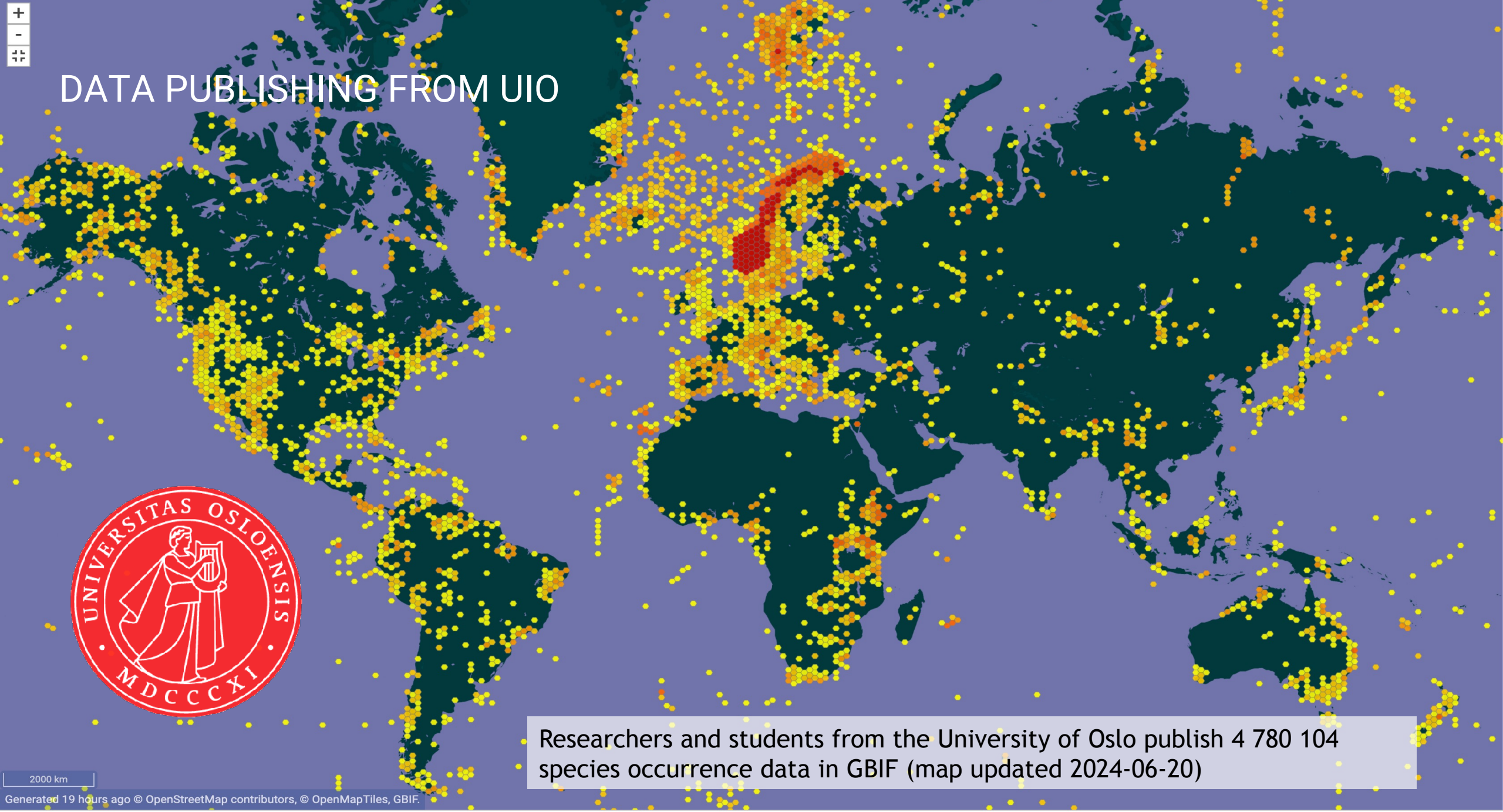
GBIF support **researchers and students** with FAIR biodiversity data publishing services.

And with biodiversity **data reuse** and data citation.

https://www.gbif.org/publishing-data/ | https://www.gbif.org/publisher/f314b0b0-e3dc-11d9-8d81-b8a03c50a862/metrics

GBIF

DATA PUBLISHING FROM UIO



Researchers and students from the University of Oslo publish 4 780 104 species occurrence data in GBIF (map updated 2024-06-20)

GBIF-MEDIATED BIODIVERSITY DATA FROM RESEARCHERS AT UIO

Researchers and students from the University of Oslo publish 4 780 104 species occurrence data points in GBIF (updated 2024-06-20)

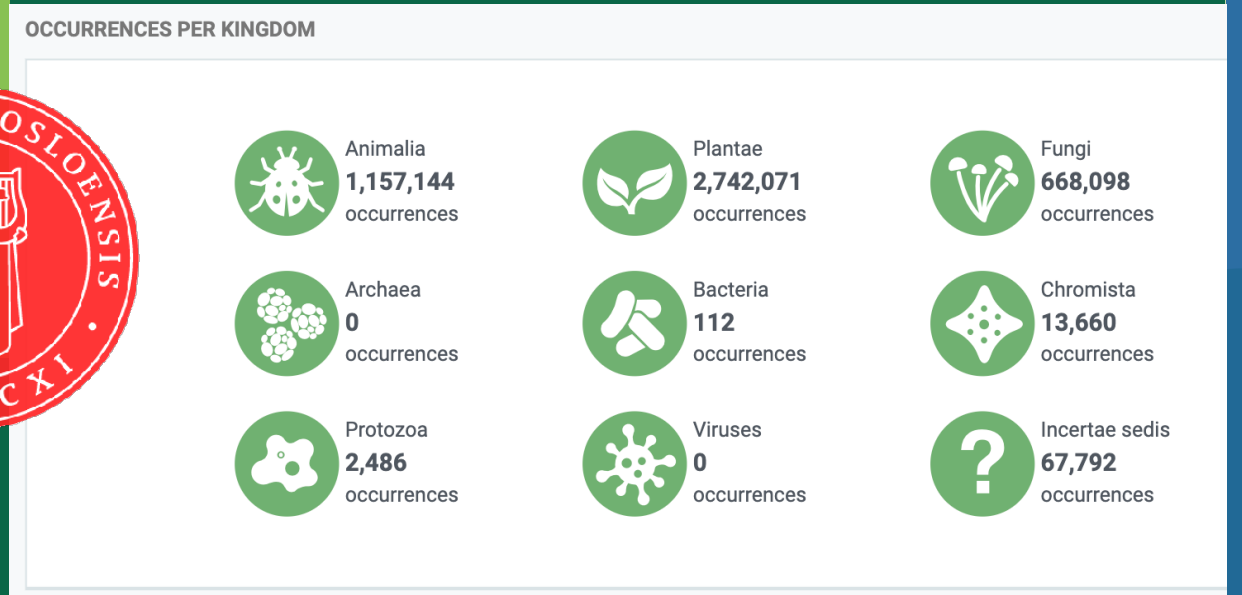
These data points have been reused and cited in 2 198 peer-reviewed scientific publications.

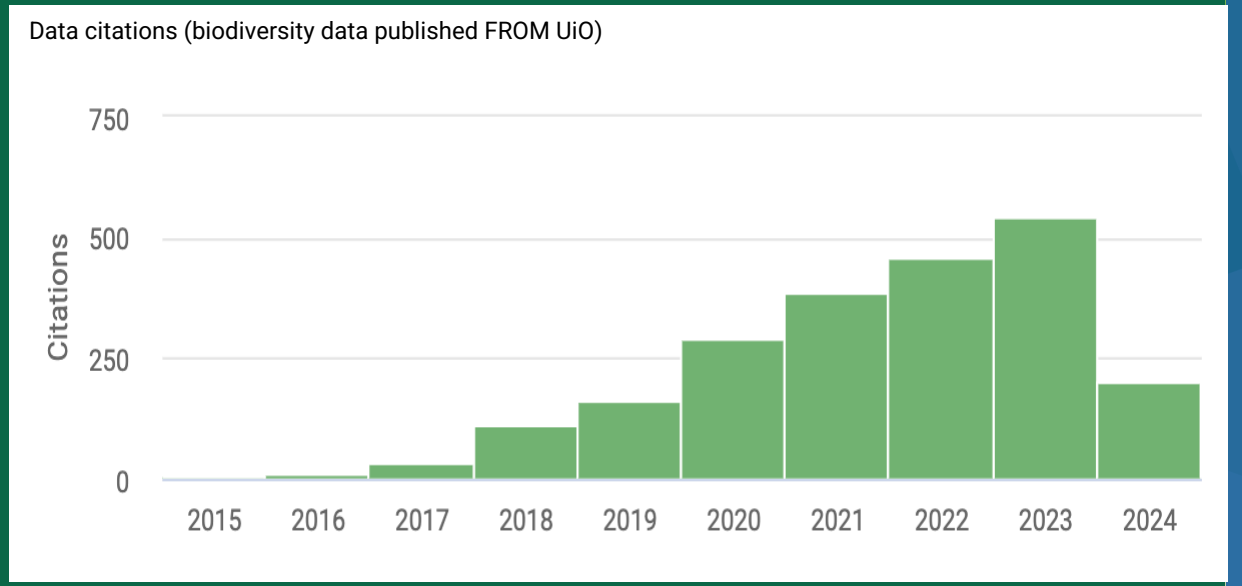
(*202 so far in 2024; 541 in 2023; 457 in 2022*)

In total for all datasets published from Norway

51 957 479 species occurrences published

447 datasets, 41 institutions





https://www.gbif.org/publisher/f314b0b0-e3dc-11d9-8d81-b8a03c50a862/metrics | https://analytics-files.gbif.org/country/NO/GBIF_CountryReport_NO.pdf

GBIF NORWAY END-USERS: GBIF-MEDIATED BIODIVERSITY DATA

GBIF provides anyone, anywhere, open access to data about all types of life on Earth.

Researchers and students from Norway has published 770 peer-reviewed papers citing the reuse of GBIF-mediated data.

(28 papers so far in 2024; 140 in 2023; and 137 in 2022)

https://www.gbif.org/resource/search?contentType=literature&countriesOfResearcher=NO&peerReview=true

GBIF

BIODIVERSITY INFORMATION STANDARDS (TDWG)

Biodiversity Information Standards (TDWG) is a non-profit association working for open standards for exchange of biodiversity data.

Originally named Taxonomic Databases Working Group (TDWG) and established in September 1985 (38 years ago).

GBIF Norway is an active member of TDWG.

TDWG organizes an annual conference for its members.

The TDWG 2026 conference will be in Oslo and hosted by GBIF Norway and UiO MaNa.

dScience is invited to take part in developing the 2026 conference.

AI LLM DIGITISATION PIPELINE

Prototype AI LLM herbarium specimen digitisation pipeline developed and tested for collections in Tajikistan (Sept 2022) and Uzbekistan (April 2024)

Another small-gant proof-of-concept project under development with Poland, Netherlands, Slovakia, Ukraine, Croatia, Belgium, Italy, and the DiSSCo ESFRI (concept note submitted, proposal August 2024)

Invitation to dScience to explore possibilities for a Horizon Europe proposal (lead by UiO?)

Google Cloud Vision API

ChatGPT

GBIF

GBIF

THANK YOU

Dag Endresen

GBIF Nodeleder for Norge

UiO Naturhistorisk museum

Forskningsgruppe for Maskin-lesbar Natur (MaNa)