

# Data citation and GBIF powered literature

Dmitry Schigel | Scientific officer



*Biodiversity data in montane and arid Eurasia  
Almaty, Kazakhstan*

*18-19 November 2024*

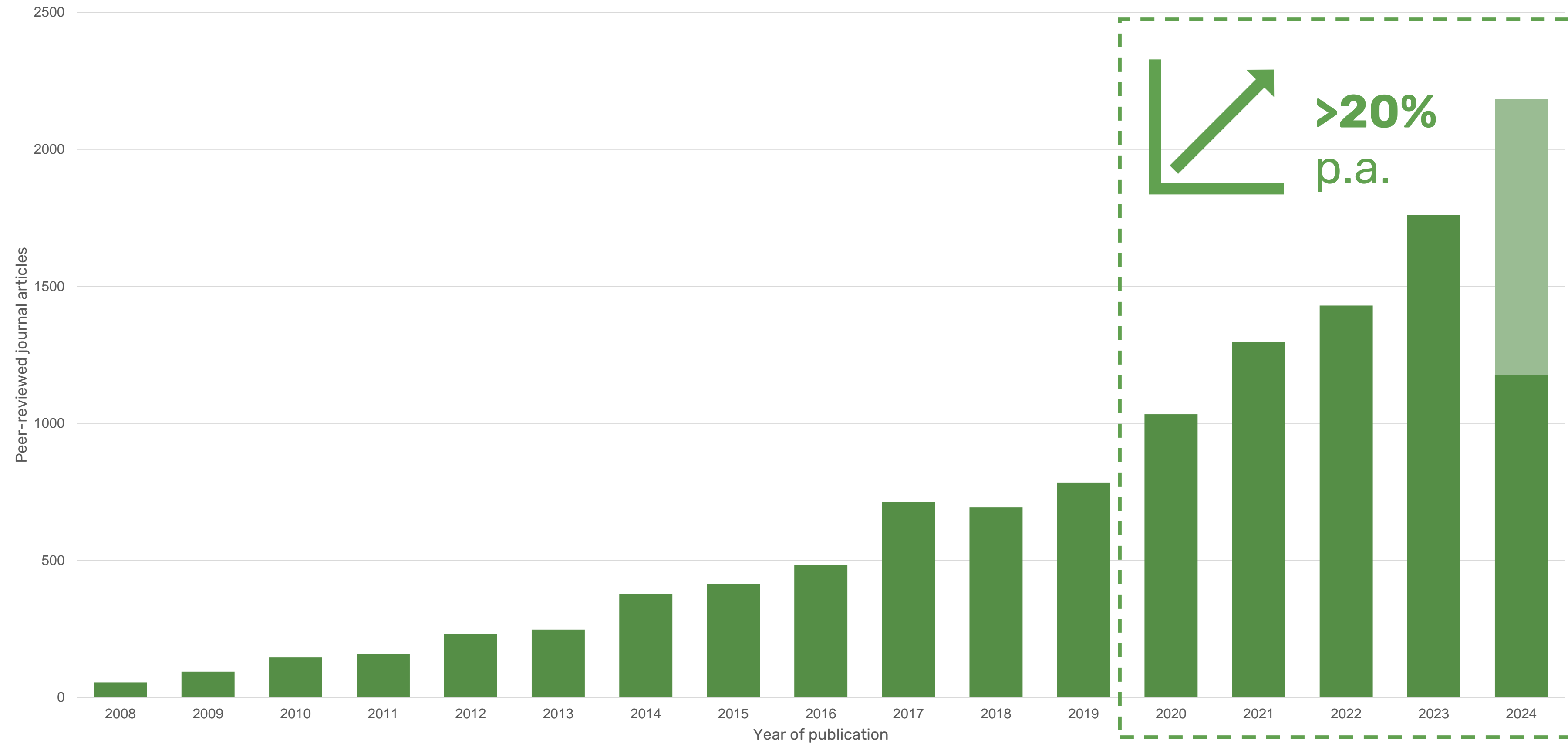


# Why track citations?


- Knowledge of where, when, how and by whom data is used
- Evidence of the impact
- Credit to data publishers
- Started in 2010
- Optimized, streamlined and automated since 2017

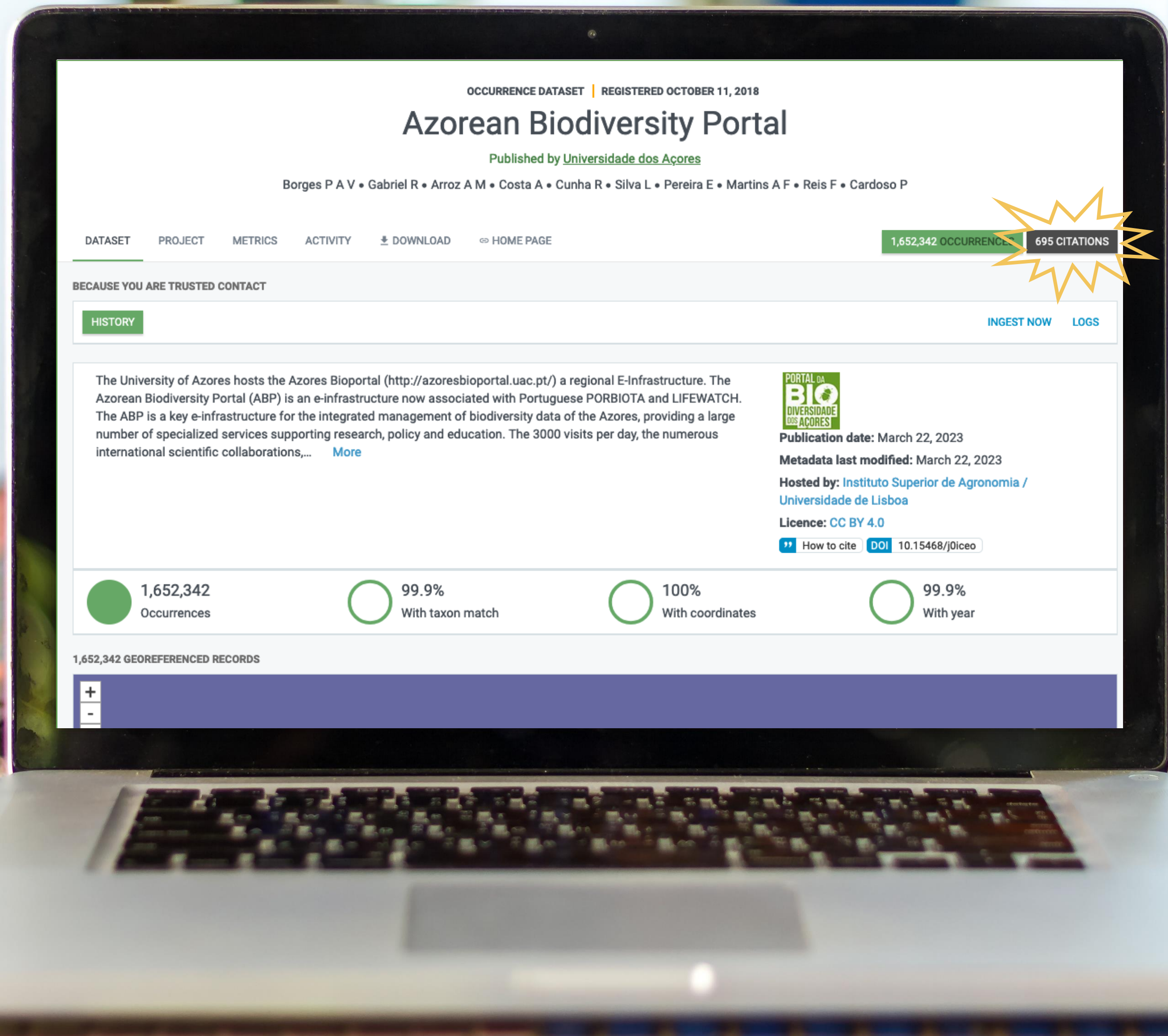


# Use of GBIF-mediated data in research



## Data citations: All about the DOIs

- **DOIs** for datasets
- **DOIs** for downloads
- **DOIs** for derived datasets
  
- Paper → (download)  dataset(s)
- Dataset citations





OCCURRENCE DATASET | REGISTERED OCTOBER 11, 2018

# Azorean Biodiversity Portal



Published by [Universidade dos Açores](#)

Borges P A V • Gabriel R • Arroz A M • Costa A • Cunha R • Silva L • Pereira E • Martins A F • Reis F • Cardoso P



DATASET PROJECT METRICS ACTIVITY  DOWNLOAD  HOME PAGE


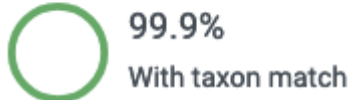


1,652,342 OCCURRENCES 695 CITATIONS

BECAUSE YOU ARE TRUSTED CONTACT



HISTORY  INGEST NOW  LOGS

The University of Azores hosts the Azores Bioportal (<http://azoresbioportal.uac.pt/>) a regional E-Infrastructure. The Azorean Biodiversity Portal (ABP) is an e-infrastructure now associated with Portuguese PORBIOTA and LIFEWATCH. The ABP is a key e-infrastructure for the integrated management of biodiversity data of the Azores, providing a large number of specialized services supporting research, policy and education. The 3000 visits per day, the numerous international scientific collaborations,... [More](#)

 **Publication date:** March 22, 2023  
**Metadata last modified:** March 22, 2023  
**Hosted by:** [Instituto Superior de Agronomia / Universidade de Lisboa](#)  
**Licence:** [CC BY 4.0](#)  
 How to cite [DOI 10.15468/jdiceo](#)

 1,652,342 Occurrences  99.9% With taxon match  100% With coordinates  99.9% With year

1,652,342 GEOREFERENCED RECORDS



# Data in 108,709 datasets: attribution, credit and affiliation

Get data How-to Tools Community About

OCCURRENCE DATASET | REGISTERED JULY 27, 2009

## Swiss National Bryophyte Databank

Published by [Swiss National Biodiversity Data and Information Centres – infospecies.ch](https://www.infospecies.ch)

Hofmann H • Cailliau A • Hartwig A

DATASET METRICS ACTIVITY DOWNLOAD HOME PAGE

236,552 OCCURRENCES 111 CITATIONS

This dataset is maintained by Swissbryophytes (National Data- and Information Center of Swiss Bryophytes, formerly "National Inventory of Swiss bryophytes", NISM). We are a member of InfoSpecies. The dataset includes records of Bryophytes (Anthocerotophyta, Bryophyta, Marchantiophyta) from Switzerland and the adjacent area. Data sources include official herbaria and private collections from a large network of volunteer collaborators, inventories (National Inventory of Swiss bryophytes NISM, Red ... [More](#))

**Swissbryophytes**

Publication date: March 8, 2024  
Metadata last modified: March 8, 2024  
Hosted by: GBIF Swiss Node  
Licence: CC BY 4.0  
How to cite DOI: 10.15468/ajkhha

236,552 Occurrences 100% With taxon match 100% With coordinates 100% With year

236,552 GEOREFERENCED RECORDS

1 Affiliation

2 Authorship

3 Data citations

4 DOI

Frontiers of Biogeography 2021, 13.04, e51146

**RESEARCH ARTICLE** Frontiers of Biogeography the scientific journal of the International Biogeography Society

### Climatic drivers of *Sphagnum* species distributions

Charles Campbell<sup>1,2\*</sup>, Gustaf Granath<sup>2</sup> and Håkan Rydin<sup>2</sup>

<sup>1</sup> Greensway AB, Ulls väg 24A, 756 51 Uppsala, Sweden; <sup>2</sup> Department of Ecology and Genetics, Evolutionary Biology Centre, Uppsala University, Norbyvägen 18D, SE-752 36 Uppsala, Sweden. \*Correspondence: Charles Campbell, charlescampbell@outlook.com

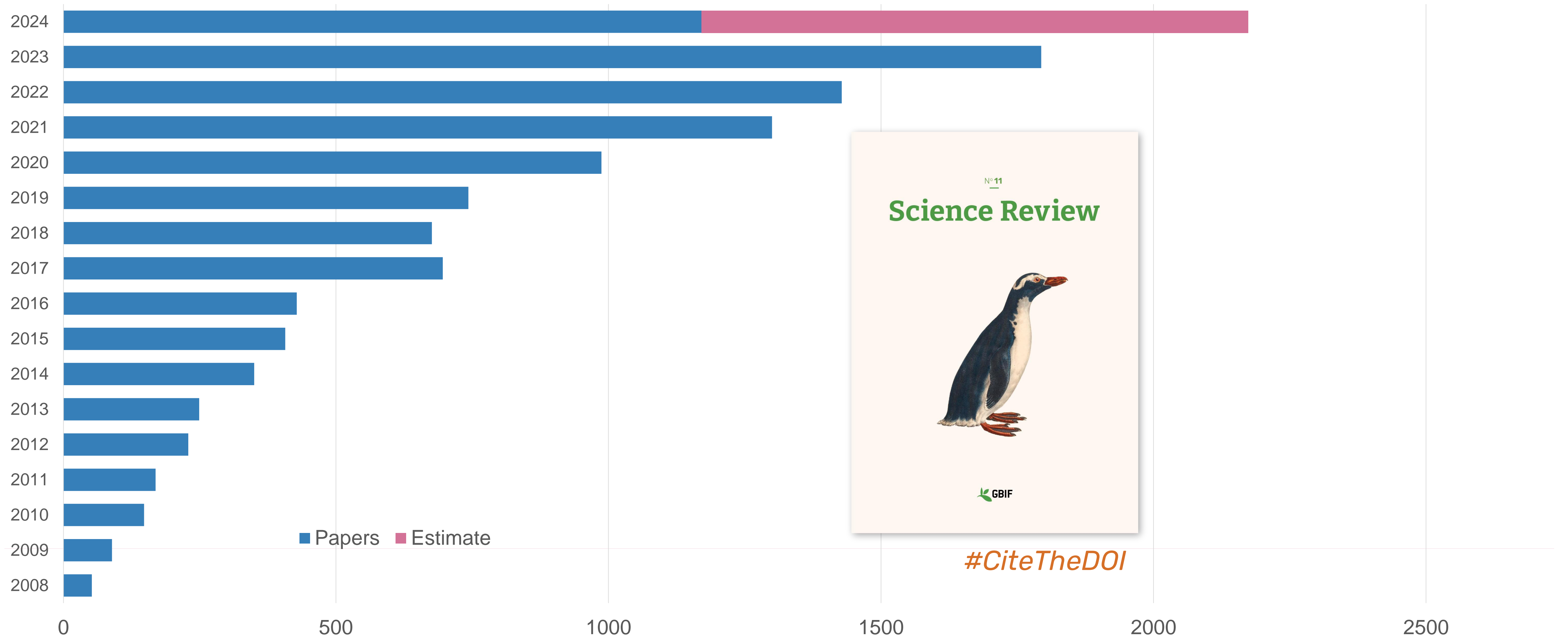
**Abstract**  
Peat mosses (genus *Sphagnum*) dominate most Northern mires and show distinct distributional limits in Europe despite having efficient dispersal and few dispersal barriers. This pattern indicates that *Sphagnum* species distributions are strongly linked to climate. *Sphagnum*-dominated mires have been the largest terrestrial carbon sinks in Europe over the last few millennia. Understanding the climatic drivers of *Sphagnum* species distributions is important for predicting the future functionality of peatlands. We used MaxEnt, with biologically relevant climatic variables, to model and clarify the current distributions of 45 *Sphagnum* species in Europe. We

**Highlights**

- Peat mosses (*Sphagnum*) form northern peatlands and species have different distributions across Europe.
- We model the climatic suitability for all European species using multiple databases and MaxEnt models.
- The climatic suitability for most species can be accurately modelled with mean annual temperature and water balance and their variation over the year.
- *Sphagnum* has its highest species richness in northwestern Europe.

Hofmann H, Kiebacher T, Moser T, Meier M (2021). Swiss National Bryophyte Databank. Swiss National Biodiversity Data and Information Centres – infospecies.ch. Occurrence dataset <https://doi.org/10.15468/ajkhha> accessed via GBIF.org on 2022-04-28.

# Peer-reviewed publications using GBIF-mediated data



# Data citations

## Ideal world

- All citations include a DOI, uniquely and specifically identifying the data used in a publication
- All data citations are included in the reference list of an article
- All data citations are machine-readable links between DOIs included in the DOI metadata, easily interpreted and used for quantifying citations
- 100% automatic 🤖

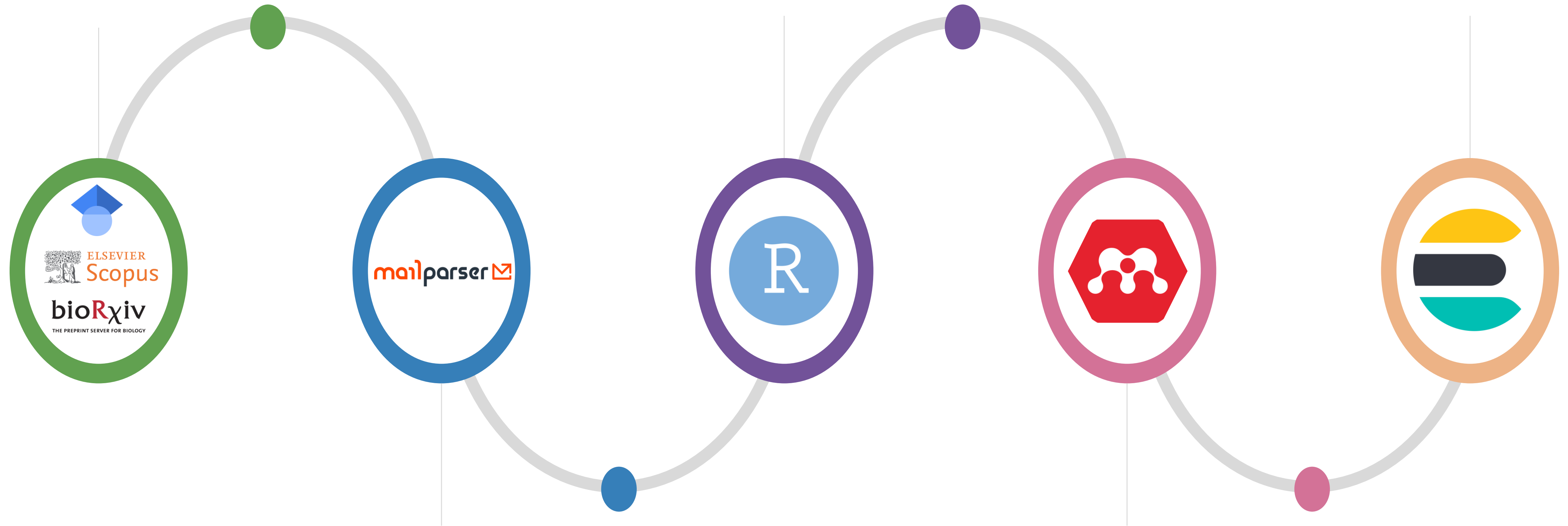
## Real world

- Far from all citations include a DOI, and some tend to cite more data than is actually used
- Some citations are included in the body text of papers, some in data availability statements, and some are even hidden in supplements
- Even when done right, not all citations become available as machine-readable links between papers and dataset
- Automatic? 😬

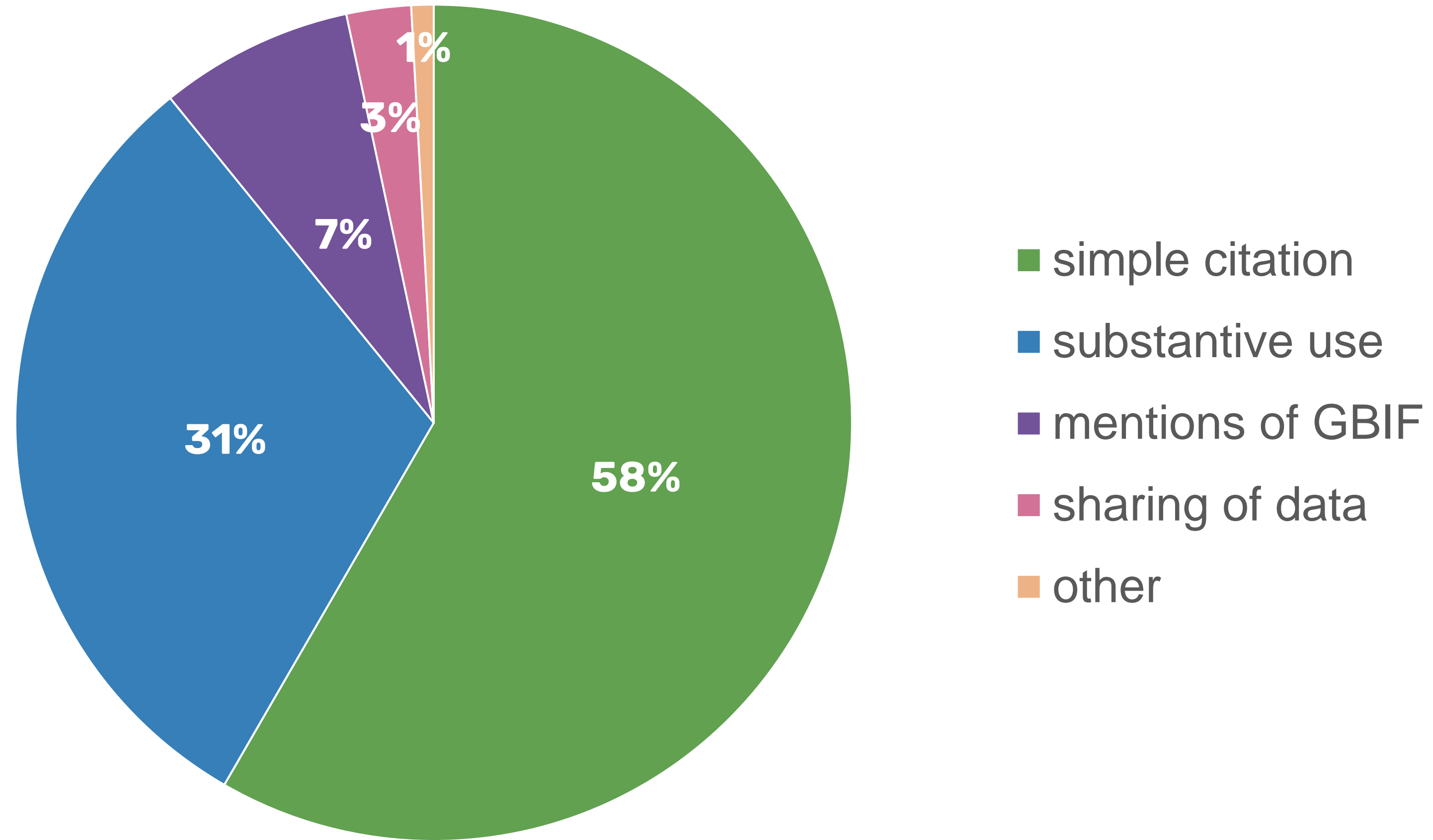




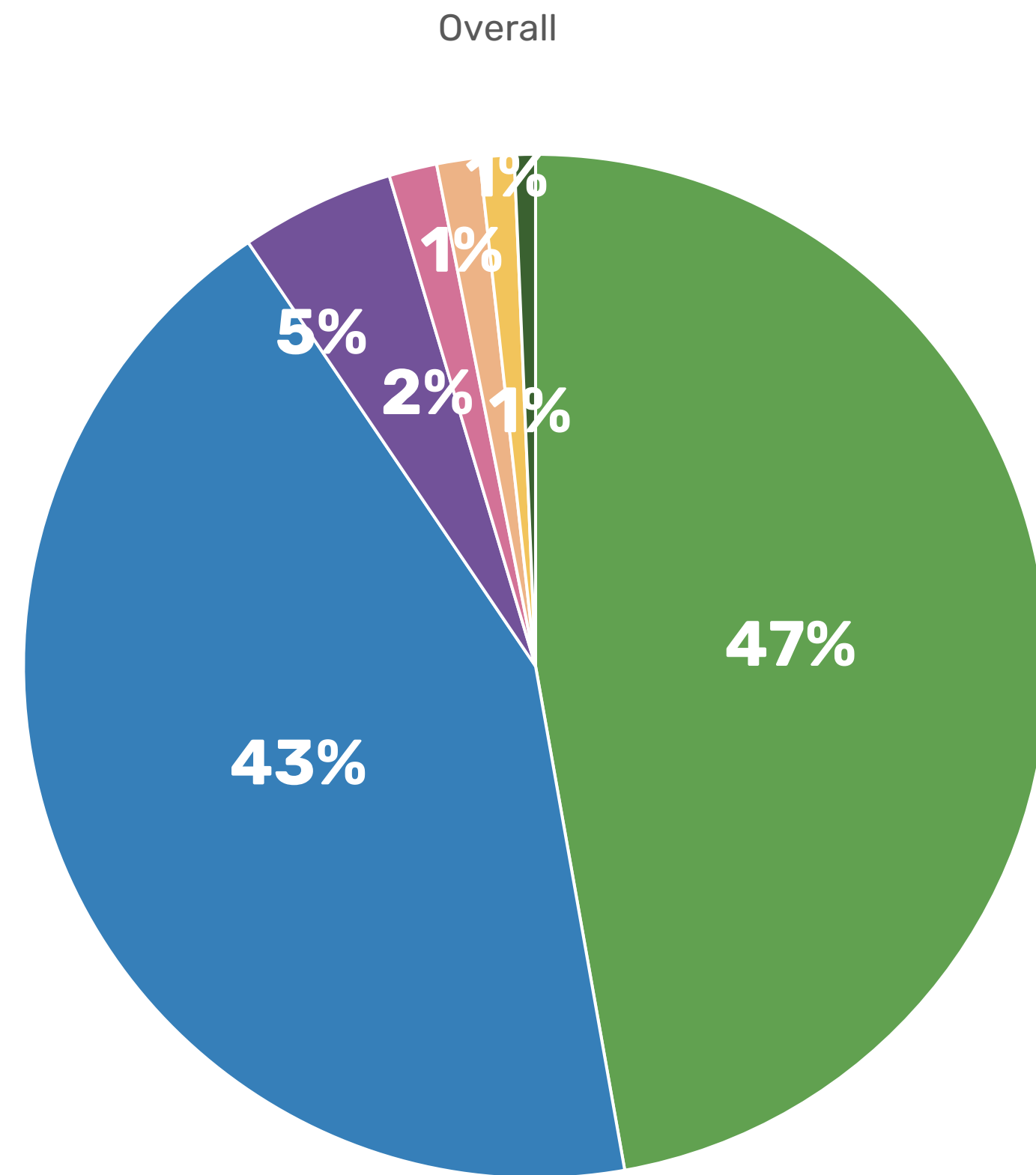
# Literature tracking in GBIF



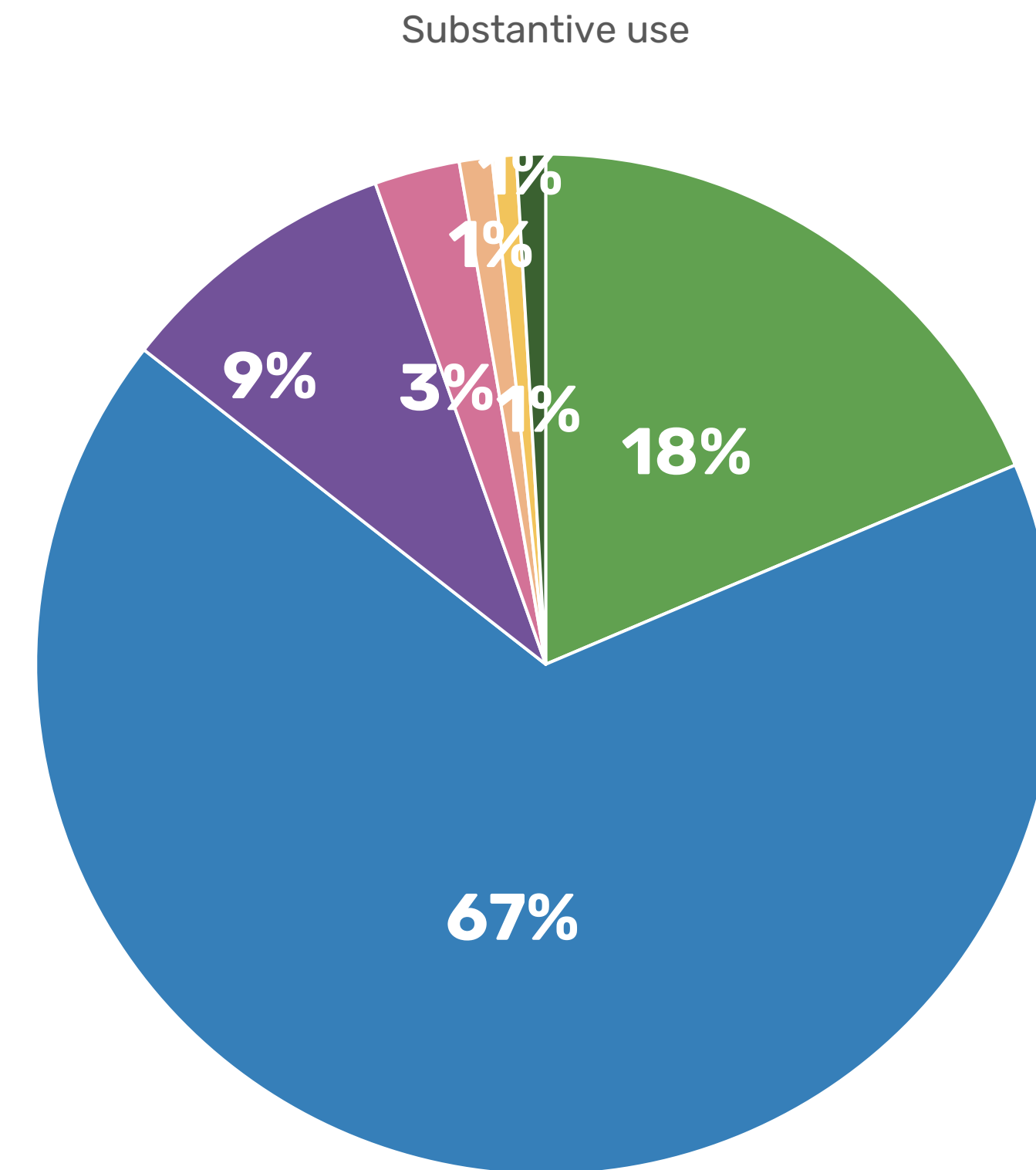
# Why are papers mentioning GBIF?



# Types of literature tracked



- Reports
- Journal articles
- Preprints and working papers
- Theses and dissertations
- Books
- Conference papers
- Misc.



Nº 11

# Science Review



 GBIF





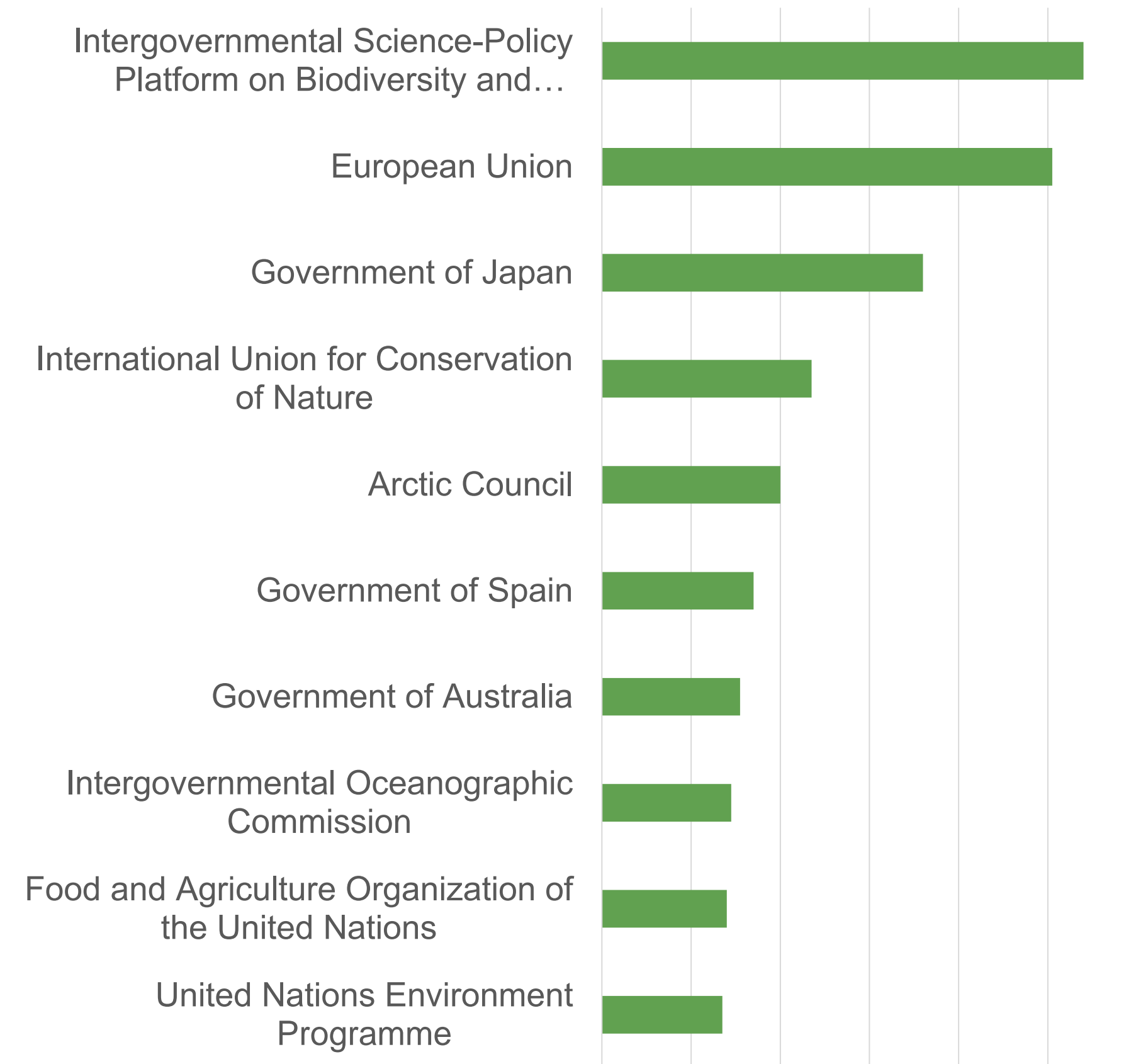
# GBIF and policy

More than just research articles

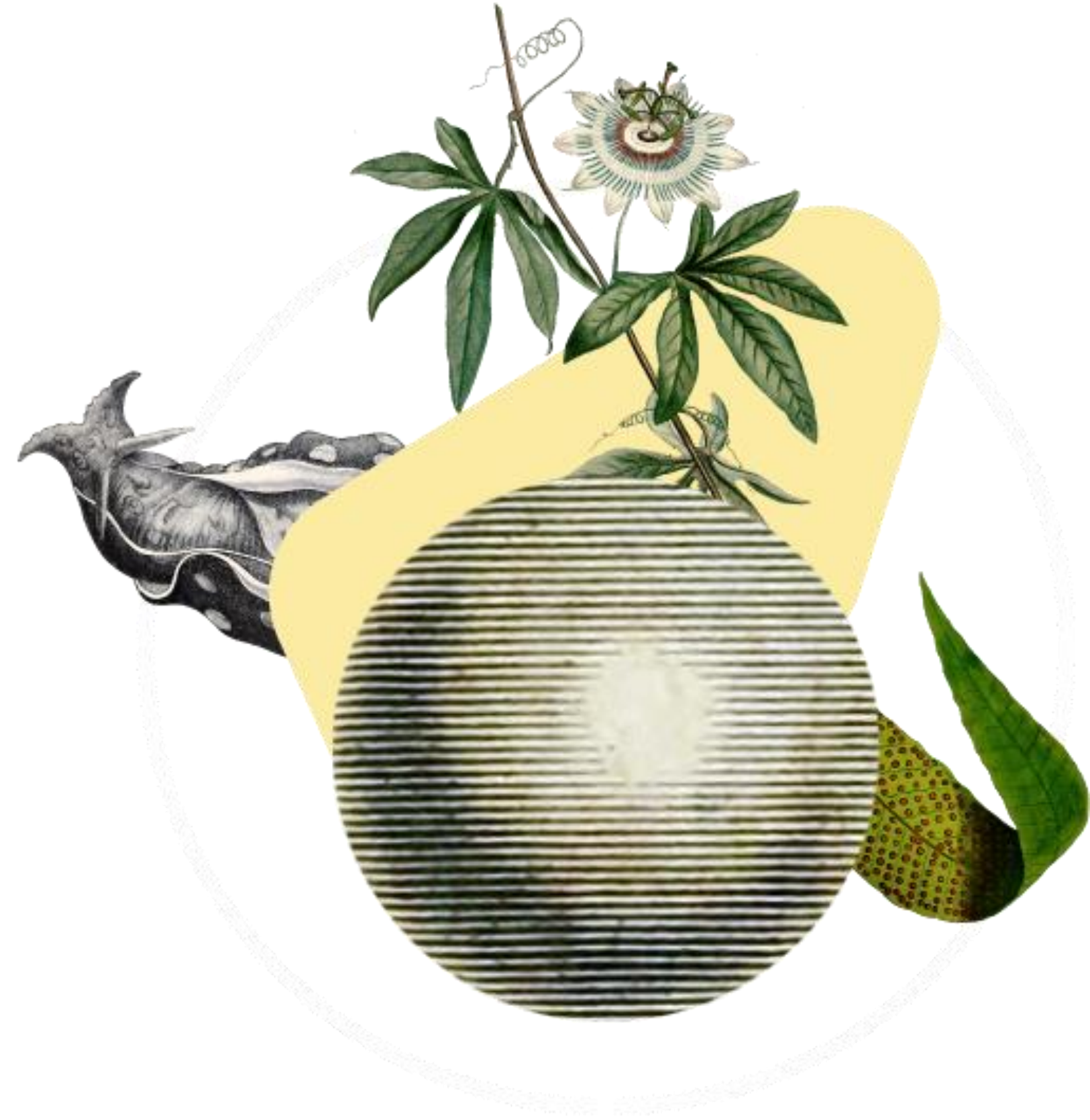
# Overton

The world's largest searchable index of policy documents, guidelines, think tank publications and working papers

- ~1,400 GBIF-relevant documents identified, published by 350 bodies including more than 100 national, regional and municipal governments
- Other top contributors include IPBES, IUCN, Arctic Council, IOC-UNESCO and FAO



## ~10 years of tracking data use and citations



- Manuscript in progress
- Process of tracking literature
- Compiling all the findings from the programme including new deep dives into download and citation practices, taxonomic and geographic focus and more
- Stay tuned!

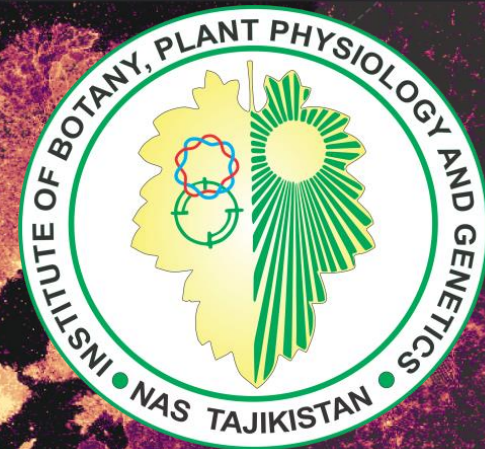


Dmitry Schigel  
dschigel@gbif.org

@dschigel







# Tajikistan National Report

Samariddin Barotov- Node manager for GBIF in  
Tajikistan

Almaty 2024



Illustration: GBIF data portal

# Tajikistan – Country Profile

- **Location** - The Republic of Tajikistan is an inland country located in the south-eastern part of Central Asia.
- **Population** - The population of Tajikistan as of January 1, 2023 is 10 million people.
- **Capital:** Dushanbe
- **Area:** 142,600 sq km
- **Languages:** Tajik, Russian, Uzbek, English and others



The biological diversity of Tajikistan today has more than 23,300 species of flora and fauna, and on average there are more than 164 species per thousand square kilometers of territory, which is ten times more than the world indicator.

Due to anthropogenic and other types of impact on nature, 226 plant species and 162 animal species are included in the Red Data Book of Tajikistan, which have become rare and are under threat of extinction.

The loss of agrobiodiversity in Tajikistan is especially negatively affected by the process of global climate change. Therefore, it is now necessary to take measures to preserve local biodiversity and increase the adaptive capacity communities to climate change.



Norway

# BioDATA

2018 - 2022



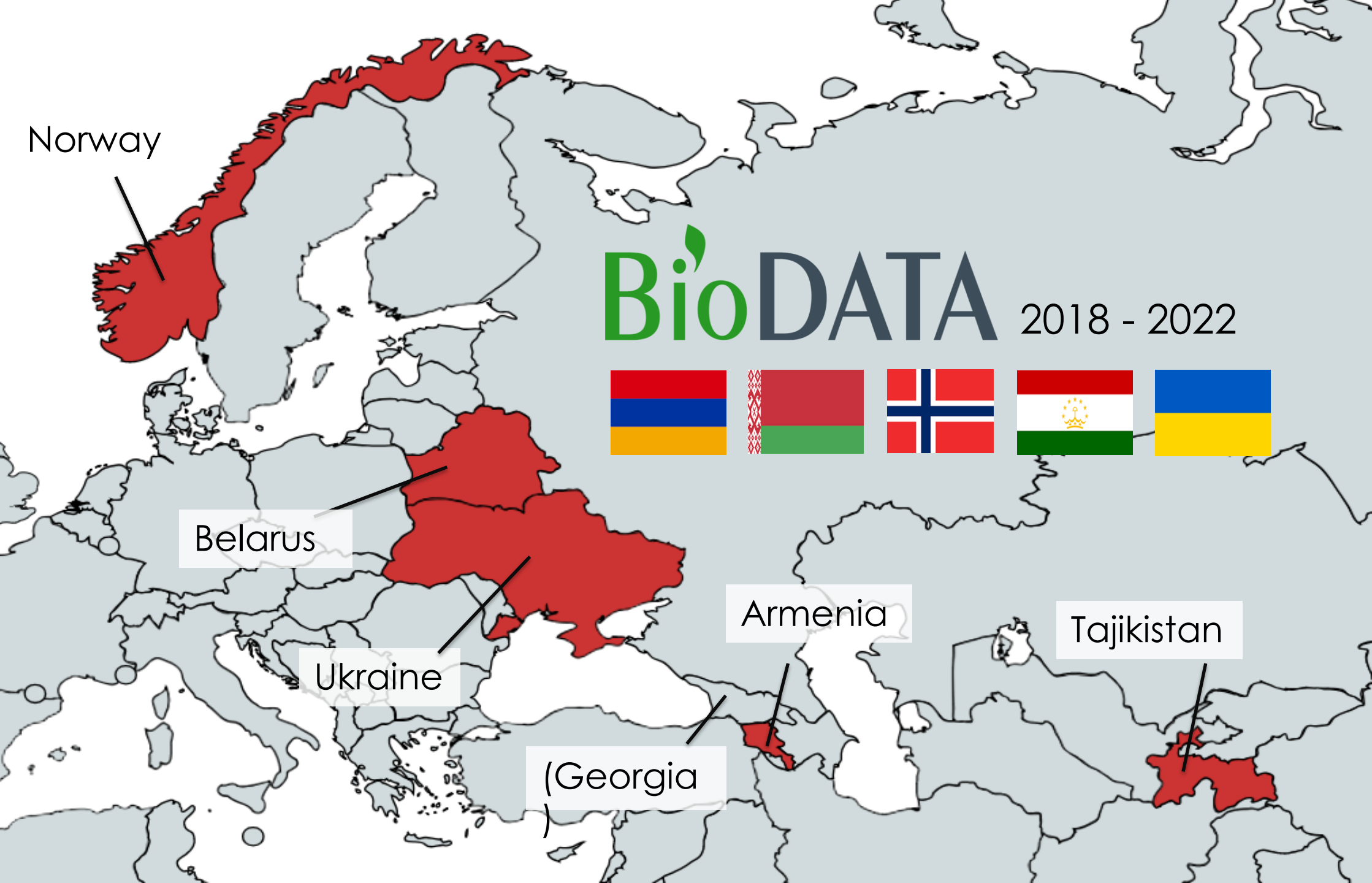
Belarus

Armenia

Tajikistan

Ukraine

(Georgia)





*Including additional students  
from Uzbekistan and Kyrgyzstan*



Regional training | June 2019 | Shambari, Tajikistan



# GBIF Tajikistan

# Kick-starting the biodiversity data publication process for Tajikistan ID:CESP2022-001

GBIF  
Norway  
main  
Partner

IBPPG  
TNAS

Khatlon  
SC  
TNAS

TPU

TNU

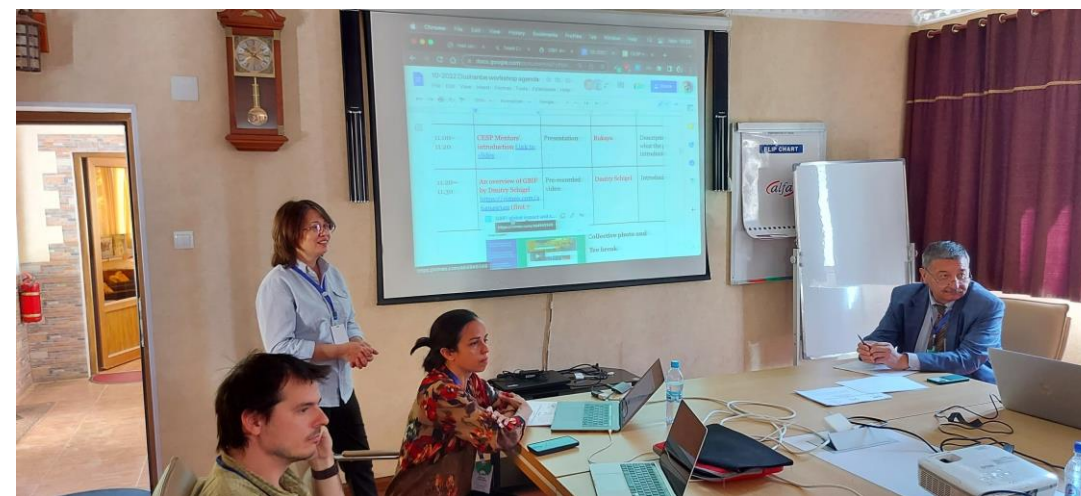
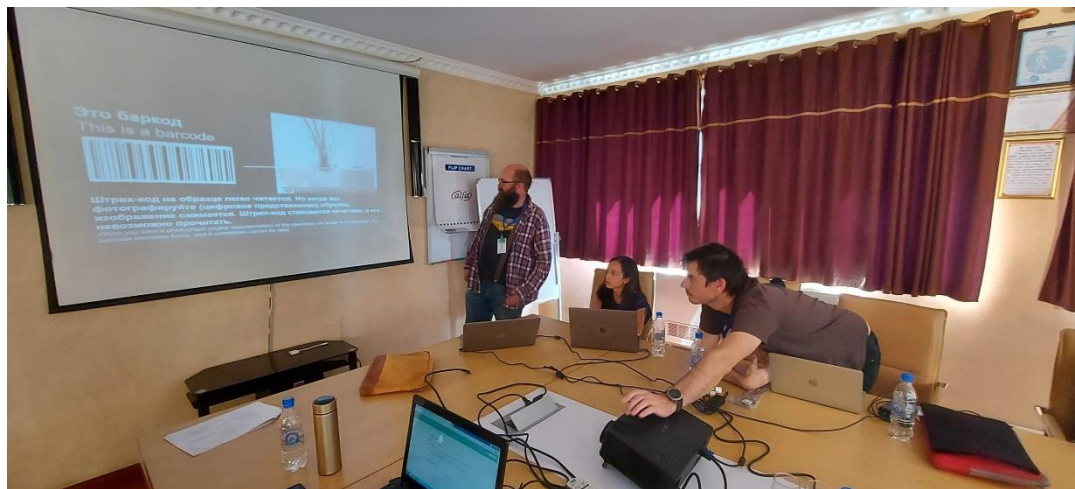
## Activity Detail Summary

Activity	Description	Start Date	End Date	Deliverable or Impact
Identifying and inventorying the organizations in Tajikistan that hold specimen collections	A list of potential data publication partners and contact details. All these institutions will be registered into GRSciColl and Wikidata.	1/9/2022	26/9/2022	Impact: This will build communication and support capacity within the Tajikistan GBIF node.
Engaging key regional organizations	COVID -19 restrictions permitting, we plan 1 physical workshop/conference, with optional digital participation	27/9/2022	29/9/2022	Increased awareness of the data publication process in Tajikistan, and data publication support capacity at the Tajik GBIF Node. We can follow that up with an offer of one-on-one digital assistance from GBIF Norway in conjunction with the data publisher and GBIF Tajikistan



Digitization of legacy biodiversity datasets in Tajikistan	A digitisation workflow. This will include the setup for technical equipment and specimen imaging	3/10/2022	30/12/2022	Impact: Greater capacity for specimen digitisation at the Herbarium, and more importantly greater knowledge + capacity at the Tajik node for collection digitisation and data publication”
GBIF data publication	Deliverable: Data publication of approximately 645 records of which some recordings are in a BRAHMS database. Data will be visible on gbif.org and on <a href="https://tajik.ipt.gbif.no">https://tajik.ipt.gbif.no</a>	3/1/2023	30/6/2023	Impact: Increased coverage of important biodiversity information for a region with limited published data

# Training workshop in Tajikistan



# PUBLISH FIRST

A novel, rapid digitisation/publishing workflow for herbaria low on staff time and technical resources

26 - 28 SEPT 2022: GBIF  CESP WORKSHOP



STEPS - 1: SCAN

2: Upload



A python app listens to create events on the bucket

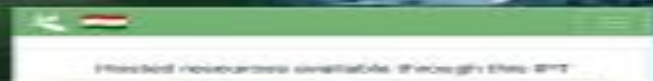


Google Vision API parses handwritten and printed text in image - some mistakes but generally good, even with handwritten Cyrillic



Google Translate API for the Russian translation (original and translation are stored verbatim in `desc:dynamicProperties`)

A simple parser splits out obvious label information (scientific name, altitude, etc). This is output into a `desc` stored on the bucket



IPT set up to publish automatically from source file on bucket



# Tajikistan

An associate participant from Europe and Central Asia  
Names of countries and areas are based on the [ISO 3166-1 standard](#)

## DATA FROM TAJIKISTAN

53,160	4	26	5
Published occurrences	<a href="#">Published datasets</a>	Countries and areas covered by data from Tajikistan	Publishers from Tajikistan

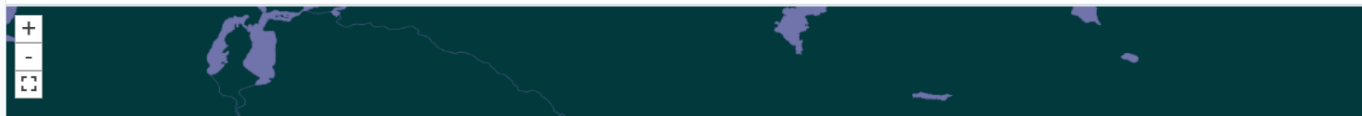


# Tajikistan

An associate participant from Europe and Central Asia  
Names of countries and areas are based on the [ISO 3166-1 standard](#)

## DATA ABOUT TAJIKISTAN

126,438	454	35	211
Occurrences	Datasets	Countries and areas contribute data	Publishers



# Achievements GBIF Tajikistan

Get data How-to Tools Community About

OCCURRENCE DATASET | REGISTERED OCTOBER 14, 2022

## The Herbarium Fund of the Institute of Botany, Plant Physiology and Genetics at the Tajikistan National Academy of Sciences - BRAHMS records

Published by [Institute of Botany, Plant Physiology and Genetics, National Academy of Sciences of Tajikistan](#)

Barotov S


DATASET PROJECT METRICS ACTIVITY [DOWNLOAD](#)

11,100 OCCURRENCES 2 CITATIONS

BECAUSE YOU ARE TRUSTED CONTACT

HISTORY [INGEST NOW](#) [LOGS](#)

Specimens from the Herbarium Fund of the Institute of Botany, Plant Physiology and Genetics at the Tajikistan National Academy of Sciences.



Project ID: CESP2022-001

Get data How-to Tools Community About

OCCURRENCE DATASET | REGISTERED OCTOBER 14, 2022

## The Herbarium Fund of the Institute of Botany, Plant Physiology and Genetics at the Tajikistan National Academy of Sciences - BRAHMS records

Published by [Institute of Botany, Plant Physiology and Genetics, National Academy of Sciences of Tajikistan](#)

Barotov S

DATASET METRICS ACTIVITY [DOWNLOAD](#)

11,100 OCCURRENCES 25 CITATIONS

BECAUSE YOU ARE TRUSTED CONTACT

HISTORY [INGEST NOW](#) [LOGS](#)

# Achievements GBIF Tajikistan

During the implementation project, we organized several virtual meetings about the evaluation and monitoring of the project with our main partners GBIF team Norway. About evaluation and monitoring local organization I personally visited to the Universities. During my 3h lecture, I explained for teachers and student's biological faculty about data publication in the GBIF one more again. From students and teachers were most interest about our project it talks about outputs and deliverables and capacity building young generations. Implementation project is going well efficiency that we connected 4 local organizations. The strength side that in our workshop participated around 27 young specialist from different of regions of Tajikistan.



# Achievements GBIF Tajikistan



TNU



TPU



# BEFORE



Animalia  
**41,661**  
occurrences



Plantae  
**11,491**  
occurrences



Fungi  
**0**  
occurrences



Archaea  
**0**  
occurrences



Bacteria  
**0**  
occurrences



Chromista  
**0**  
occurrences



Protozoa  
**0**  
occurrences



Viruses  
**0**  
occurrences



Incertae sedis  
**8**  
occurrences

# NOW



Animalia  
**67,979**  
occurrences



Plantae  
**54,255**  
occurrences



Fungi  
**1,994**  
occurrences



Archaea  
**0**  
occurrences



Bacteria  
**224**  
occurrences



Chromista  
**1,213**  
occurrences



Protozoa  
**4**  
occurrences



Viruses  
**150**  
occurrences



Incertae sedis  
**619**  
occurrences

## The Herbarium Fund of the Institute of Botany, Plant Physiology and Genetics at the Tajikistan National Academy of Sciences - BRAHMS records

Occurrence dataset

Specimens from the Herbarium Fund of the Institute of Botany, Plant Physiology and Genetics at the Tajikistan National Academy of Sciences.

Published by Institute of Botany, Plant Physiology and Genetics, National Academy of Sciences of Tajikistan

11 100 occurrences 12 citations

## The Herbarium of Tajik National University

Occurrence dataset

This dataset contains specimens from the herbarium at Tajik National University. Tajik National University was established by the Resolution of the Soviet of Ministries of the USSR 21st of March 1947,...



Published by Tajik National University

259 occurrences 3 citations

## Khatlon Scientific Center

Occurrence dataset

This dataset contains specimens from the herbarium at Khatlon Scientific Center. The specimens were imaged by herbarium staff, and published via an automatic process: 1) OCR text was gathered from th...



Published by Khatlon Scientific Center of the National Academy of Sciences of Tajikistan

140 occurrences 1 citation

## The Herbarium Fund of the Institute of Botany, Plant Physiology and Genetics at the Tajikistan National Academy of Sciences - BRAHMS records

Occurrence dataset

Specimens from the Herbarium Fund of the Institute of Botany, Plant Physiology and Genetics at the Tajikistan National Academy of Sciences.

Published by Institute of Botany, Plant Physiology and Genetics, National Academy of Sciences of Tajikistan

11 100 occurrences 25 citations

## The Herbarium of Tajik National University

Occurrence dataset

This dataset contains specimens from the herbarium at Tajik National University. Tajik National University was established by the Resolution of the Soviet of Ministries of the USSR 21st of March 1947,...



Published by Tajik National University

259 occurrences 7 citations

## Khatlon Scientific Center

Occurrence dataset

This dataset contains specimens from the herbarium at Khatlon Scientific Center. The specimens were imaged by herbarium staff, and published via an automatic process: 1) OCR text was gathered from th...



Published by Khatlon Scientific Center of the National Academy of Sciences of Tajikistan

140 occurrences 6 citations

Thank you!!  
Questions?



Samariddin Barotov GBIF Node Manager for Tajikistan  
[barotov.ikai@mail.ru](mailto:barotov.ikai@mail.ru)  
[www.gbif.org](http://www.gbif.org)

## Data publishing with GBIF: concepts and tools

Laura Anne Russell | Training Officer

Presented and adapted by: Salza Palpurina | ECA  
support team



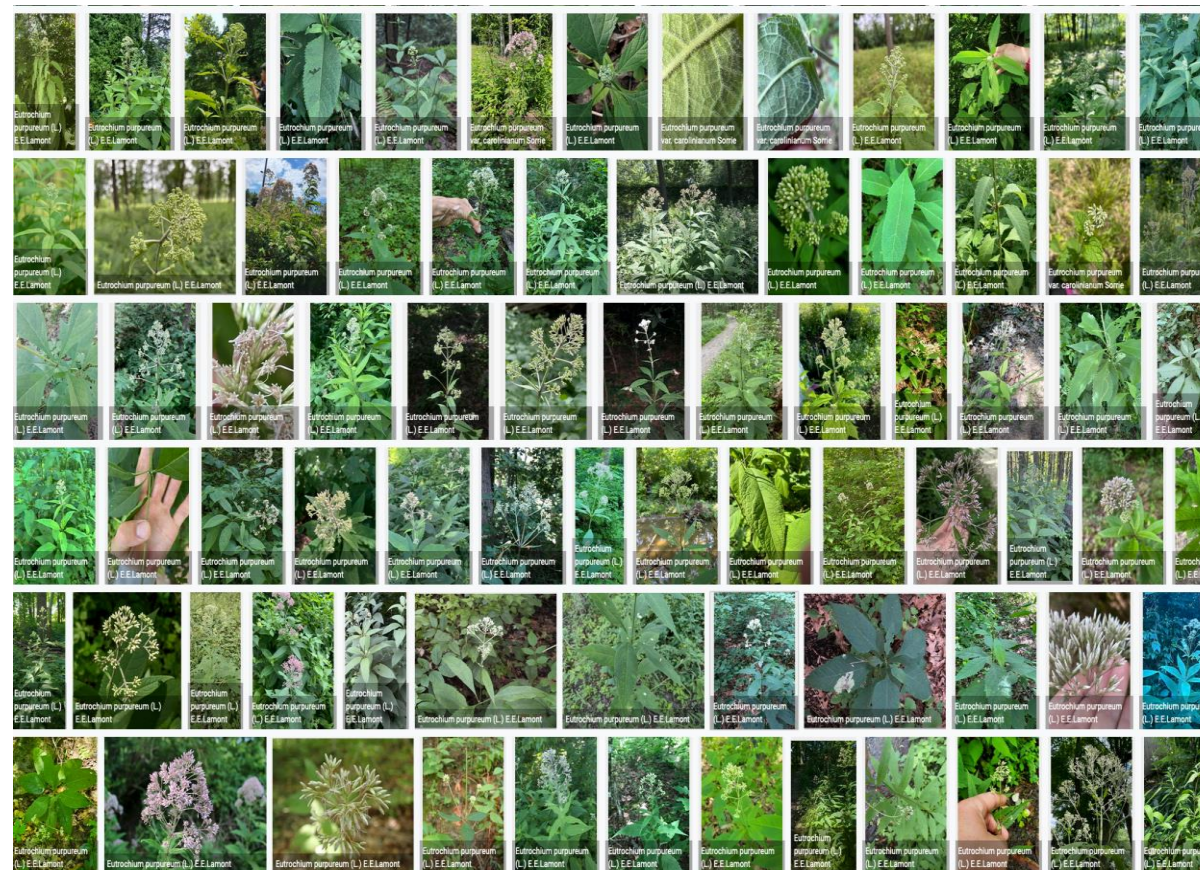
Species  
occurrence records

3,024,404,259

## Multimedia evidence

**210.7 million records** with taxonomically identified images

- 137.4 million human observations
- 63.8 million specimens
- 7.3 million material samples
- 1.4 million fossil specimens



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

# Incentives for publishing open-access biodiversity data

- contributes to global knowledge about biodiversity
- reveals new opportunities for collaboration
- gives visibility to publishing institutions
- reveals usage and citations of digitized data
- fulfills requirements to make data freely accessible



---

## BECOME A PUBLISHER

Agree to the terms

Complete the form

Wait for endorsement

Prepare data

Publish / register data with GBIF

---

## Become a publisher

*Organizations that want to share data through GBIF may register here to request endorsement as a data publisher*

As part of their preparations for publishing data through GBIF, prospective publishers must complete the online form below.

Your answers help us provide proper credit and attribution for the datasets you share. They also help users understand more about the provenance of data shared through the GBIF network.

Before GBIF indexes a new publisher's datasets, the institution must [receive endorsement as a data publisher](#) from one of the Participant nodes that coordinate activities of the [national and organizational Participants in the GBIF network](#). If your country is not yet a participant, GBIF will coordinate the endorsement request through the [GBIF Nodes Steering Group \(NSG\)](#).

The endorsement procedure aims to ensure that:

1. Published data are relevant to GBIF's scope and objectives
2. [Arrangements for data hosting](#) are stable and persistent
3. Data publishing and use are supported by strong national, regional and thematic engagement
4. Data are as open as possible and available for sharing and reuse
5. Data publishers can respond to feedback and improve data quality

# Publisher page



PUBLISHER | SINCE AUGUST 8, 2022

## Institute of Zoology of the Republic of Kazakhstan

838 OCCURRENCES 1 DATASET 11 CITATIONS

[ABOUT](#) [METRICS](#) [⇌ HOME PAGE](#)

**Description:** Mission: To develop Kazakhstan zoology science from animal world investigations - from Republic of Kazakhstan to around of the globe. Vision: Institute of Zoology RK - leading world class science organization of fundamental and practical zoology. Institute of Zoology RK main directions: Kazakhstan wildlife investigation; animal population dynamics changes from environmental impact studying; studying of evolution, phylogeny, taxonomy diversity of present and past wildlife; ecologically valuable zoology problem solving.



**Endorsed by:** [Participant Node Managers Committee](#)

**Administrative contact:** [Roman Jashenko](#)

**Country or area:** [Kazakhstan](#)

[Download activity report](#)



# SHARING YOUR DATA WITH GBIF.ORG

The image shows a Google Sheet on the left and a GBIF dataset page on the right. The Google Sheet contains a table with columns for 'Cat. Numb.', 'University', and 'Collector'. The GBIF page displays the dataset title, publication information, and a summary of 1,045,806 occurrences with various metrics like '99.9% With taxon match' and '100% With coordinates'. A map of Belgium shows the geographic distribution of the data points.

	A	B	C
1	Cat. Numb.	University	Collector
2	UWP:100217	University of Guatemala	Betancur J
3	UWP:100218	University of Guatemala	Betancur J
4	UWP:101378	University of Guatemala	Fonnegra R
5	UWP:101717	University of Guatemala	Betancur J
6	UWP:101737	University of Guatemala	Betancur J
7	UWP:102143	University of Guatemala	Betancur J
8	UWP:102144	University of Guatemala	Betancur J
9	UWP:102233	University of Guatemala	Vargas I
10	UWP:103108	University of Guatemala	Cardona F
11	UWP:104139	University of Guatemala	Fonnegra R
12	UWP:104512	University of Guatemala	Callejas R
13	UWP:105292	University of Guatemala	Acevedo P
14	UWP:106768	University of Guatemala	Vargas I

**Belgian IFBL Flora Checklists (1939-1971)**  
 OCCURRENCE DATASET | REGISTERED 4 MAY 2018  
 Published by Botanic Garden Meise  
 Wouter Van Landuyt • Nicolas Noe

1 045 806 OCCURRENCES | 9 CITATIONS

The data in IFBL 1, 2 & 3 covers all of the IFBL 1 km<sup>2</sup> flora checklists sampled between 1939 and 1971. About 10000 original lists corresponding with some 1 200 000 data representative of the former distribution of vascular plant species in Belgium, were digitised. The IFBL data is integrated in existing national and regional flora databases and will contribute to the realisation of regional Flora Atlases. The analysis of the digitised data will improve the possibilities to compare floral data ov... more

Metadata Last Modified: 5 March 2018  
 Data Last Changed: 9 September 2015  
 License: CC0 1.0  
 How to cite: DOI: 10.15468/vnibke

1 045 806 Occurrences | 99.9% With taxon match | 100% With coordinates | 91% With year

1 045 806 GEOREFERENCED RECORDS

Generated 14 hours ago © OpenStreetMap, GBIF

Any year | 1919 - 1990 | EXPLORE AREA

Description

Temporal

Geographic

Description

The data in IFBL 1, 2 & 3 covers all of the IFBL 1 km<sup>2</sup> flora checklists sampled between 1939 and 1971. About 10000 original lists corresponding with some 1 200 000 data representative of the former distribution of vascular plant species in Belgium were digitised. The IFBL data is integrated in existing national and regional flora databases and will contribute to the realisation of regional Flora Atlases. The analysis of the digitised data will improve the possibilities to compare floral data ov... more

FULL TITLE  
BOS Arthropod Collection of University of Oviedo (S  
events subset

*Dataset description, taxonomic/geographic/temporal scope, methodology*

01

### Dataset metadata

FLORA  
EUROPAEA  
VOLUME 1  
FLORA  
EUROPAEA  
VOLUME 1  
FLORA  
EUROPAEA

*List based on geography, taxonomy, theme (e.g. invasive, medicinal)*

02

### Species checklists

*Species occurrences based on observation/collection, dates, coordinates*

03

### Occurrence-only data

*Occurrences associated with survey/sample, protocol, effort, abundance*

04

### Sampling-event data

## DARWIN CORE

"List of fields and their definitions, as they relate to biodiversity data."

The image shows a laptop screen displaying the Darwin Core Occurrence page. The page is titled "Occurrence" and features a grid of field names in rounded rectangular boxes. Below the grid, there are sections for "Occurrence" and "occurrenceID", each with an identifier and a definition. A sidebar on the right, titled "On this page", lists various related terms and concepts.

**Occurrence** Class

Identifier <http://rs.tdwg.org/dwc/terms/Occurrence>

Definition An existence of a `dwc:Organism` at a particular place at a particular time.

Comments

Examples

- a wolf pack on the shore of Kluane Lake in 1988
- a virus in a plant leaf in the New York Botanical Garden at 15:29 on 2014-10-23
- a fungus in Central Park in the summer of 1929

**occurrenceID**

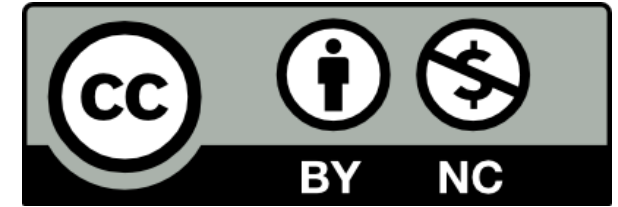
Identifier <http://rs.tdwg.org/dwc/terms/occurrenceID>

Definition An identifier for the `dwc:Occurrence` (as opposed to a particular digital record of the `dwc:Occurrence`). In the absence of a persistent global unique identifier, construct one from a combination of identifiers

**On this page**

- Record-level Occurrence
- Organism
- MaterialEntity
- MaterialSample
- Event
- Location
- GeologicalContext
- Identification
- Taxon
- MeasurementOrFact
- ResourceRelationship
- UseWithIRI
- LivingSpecimen
- PreservedSpecimen
- FossilSpecimen
- MaterialCitation
- HumanObservation
- MachineObservation
- Cite Darwin Core

# CREATIVE COMMONS WAIVER AND LICENSES



All data published and registered for use in the GBIF portal, must be issued with one of three Creative Commons licenses.

## CC0

For data made available for any use without any restrictions



## CC BY

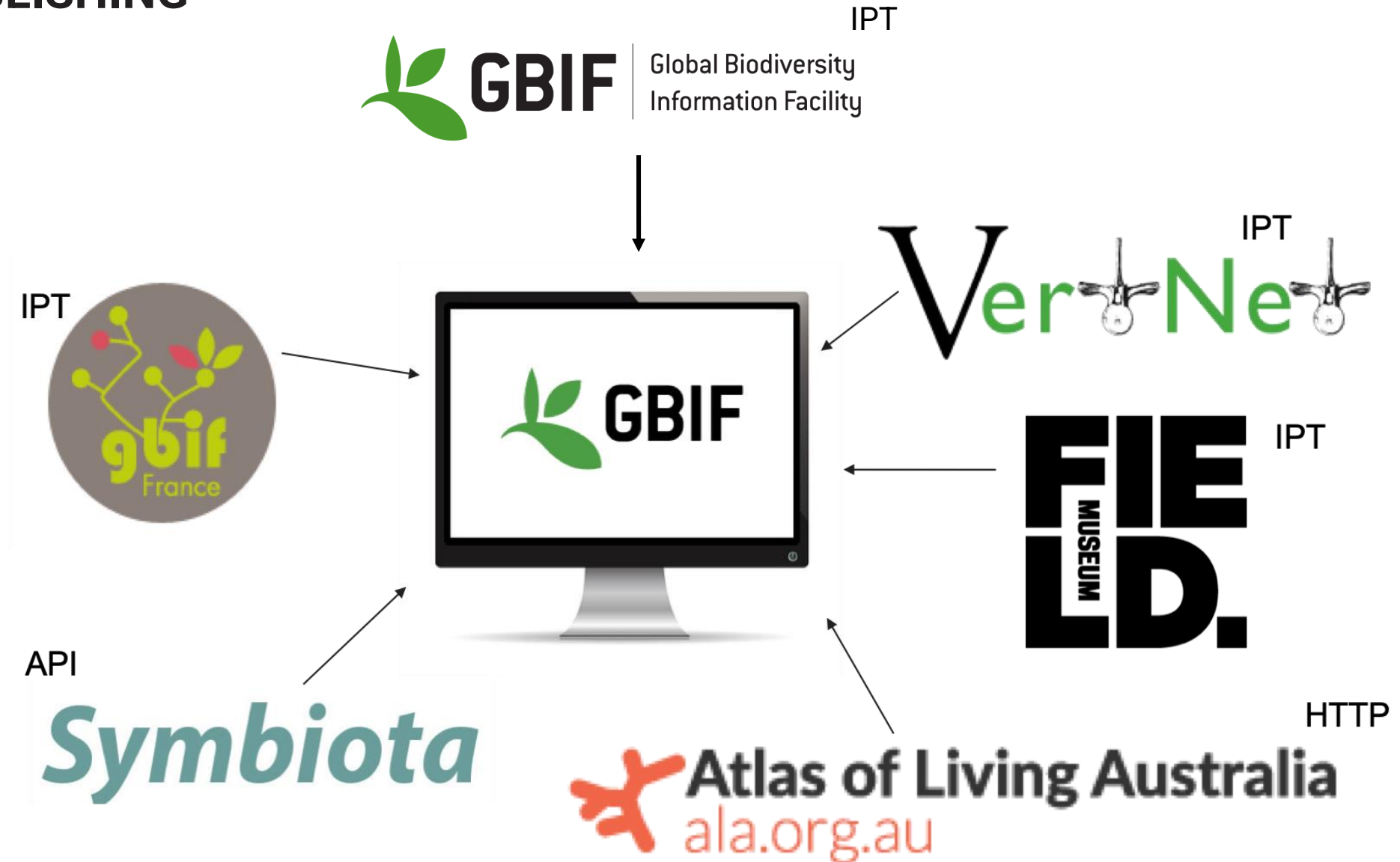
For data made available for any use with appropriate attribution



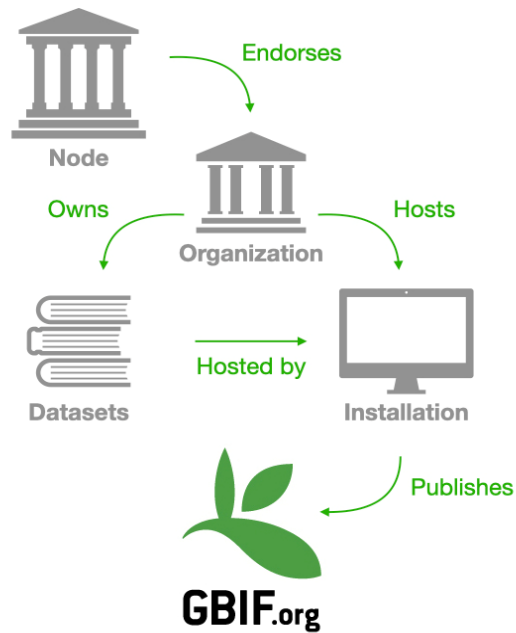
## CC BY-NC

For data made available for any non-commercial use with appropriate attribution

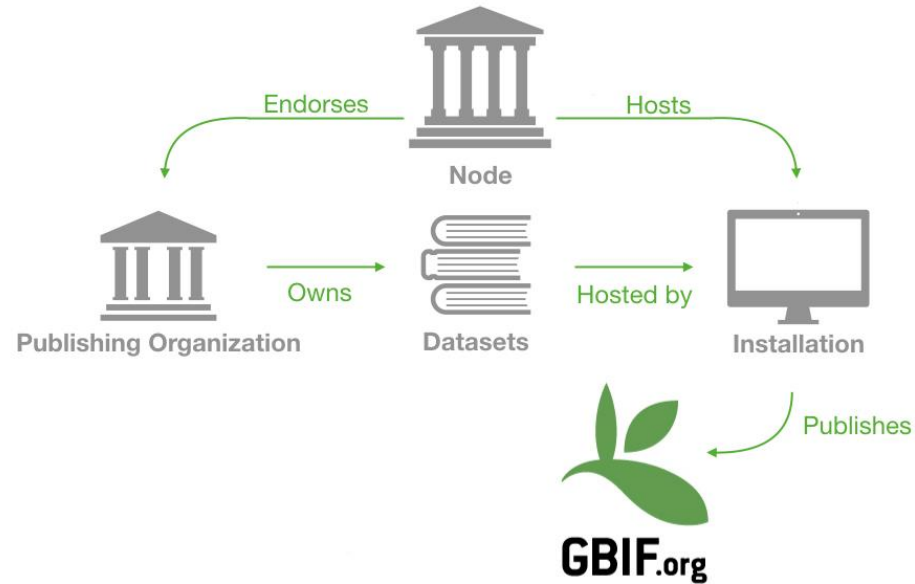
# DATA PUBLISHING



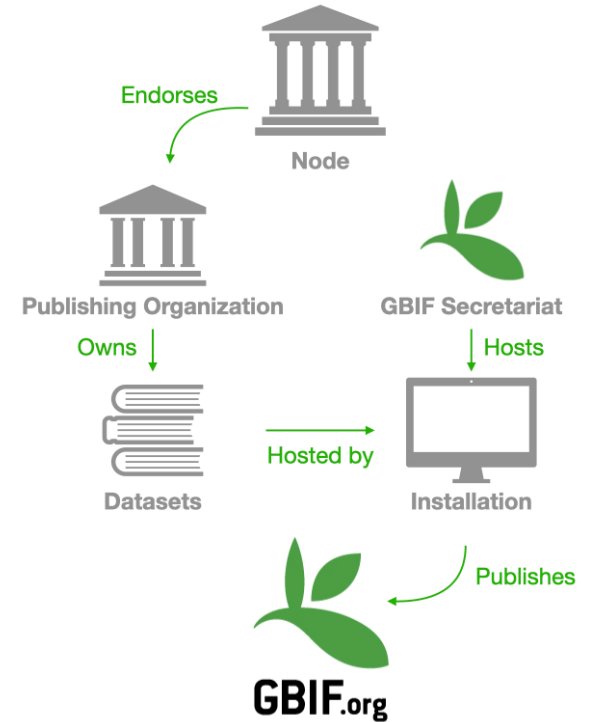
# Hosting models



Organization-hosted installation



Node-hosted installation



GBIF-hosted installation

Publishing organization ≠ Hosting installation

# A dataset published by Institute of Zoology of the Republic Kazakhstan but hosted by GBIF Secretariat

The screenshot shows the GBIF dataset page for 'Distribution of marsh frogs (Pelophylax ridibundus complex) in Kazakhstan'. The page features a green navigation bar at the top with links for 'Get data', 'How-to', 'Tools', 'Community', and 'About', along with a search icon, a language selector, and a 'Login' button. The main title is 'Distribution of marsh frogs (Pelophylax ridibundus complex) in Kazakhstan', with the publisher 'Institute of Zoology of the Republic of Kazakhstan' highlighted in a red box. Below the title, the authors are listed: Dujsebeyeva T., Kaptyonkina A., Arifulova I., Ualiyeva D., Akhmedenov K., Ivanov A., Khromov V., Krainyuk V., Sarzhanov F., Tarasovskaya N., Titov S., Timoshenko A., Ermakov O., Malakhov D., and Starikov S. The dataset statistics show 838 occurrences and 11 citations. The 'Hosted by' field is highlighted in a red box as 'GBIF Secretariat'. The page also includes a description of the data, publication and metadata dates, a license (CC BY 4.0), and a 'How to cite' section with a DOI (10.15468/et4dus). At the bottom, there are four circular progress indicators showing the dataset's completeness: 838 Occurrences (100%), 100% With taxon match, 99.6% With coordinates, and 94% With year. The total number of georeferenced records is 835.

Get data How-to Tools Community About Login

OCCURRENCE DATASET | REGISTERED OCTOBER 13, 2022

## Distribution of marsh frogs (*Pelophylax ridibundus* complex) in Kazakhstan

Published by [Institute of Zoology of the Republic of Kazakhstan](#)

Dujsebeyeva T • Kaptyonkina A • Arifulova I • Ualiyeva D • Akhmedenov K • Ivanov A • Khromov V • Krainyuk V • Sarzhanov F • Tarasovskaya N • Titov S • Timoshenko A • Ermakov O • Malakhov D • Starikov S

838 OCCURRENCES 11 CITATIONS

DATASET PROJECT METRICS ACTIVITY DOWNLOAD HOME PAGE

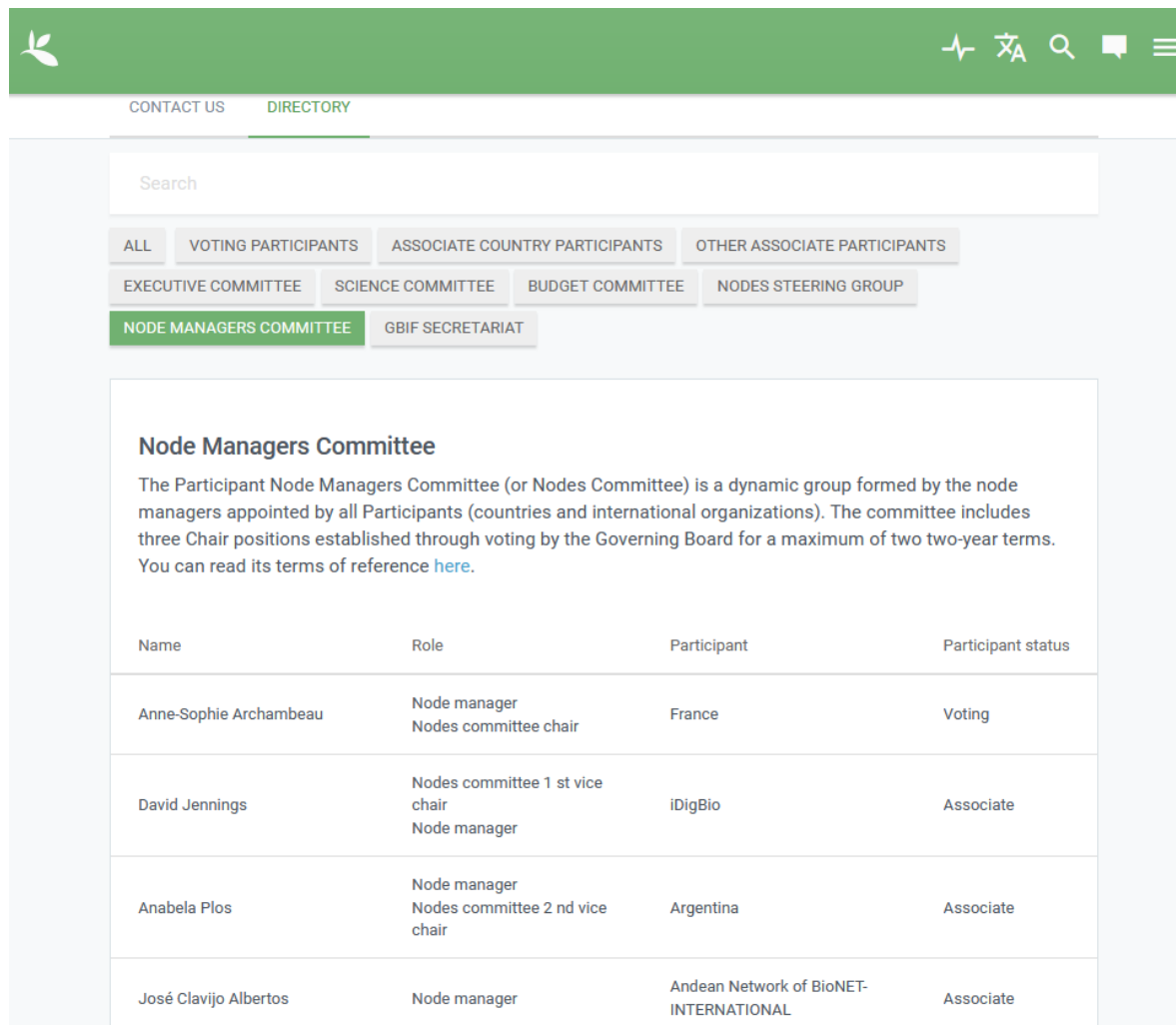
The presented data are the result of generalization and reconciliation of literary, museum and archival information on the distribution of lake frogs of the *P. ridibundus* complex in Kazakhstan, and new data were obtained during field work in 2021-2022. Based on the collected material, a database has been compiled for all the frog finds known today for the period from the end of the XX century to the present.

Publication date: April 9, 2023  
Metadata last modified: April 9, 2023  
Hosted by: [GBIF Secretariat](#)  
Licence: [CC BY 4.0](#)  
How to cite [DOI](#) 10.15468/et4dus

838 Occurrences 100% With taxon match 99.6% With coordinates 94% With year

835 GEOREFERENCED RECORDS

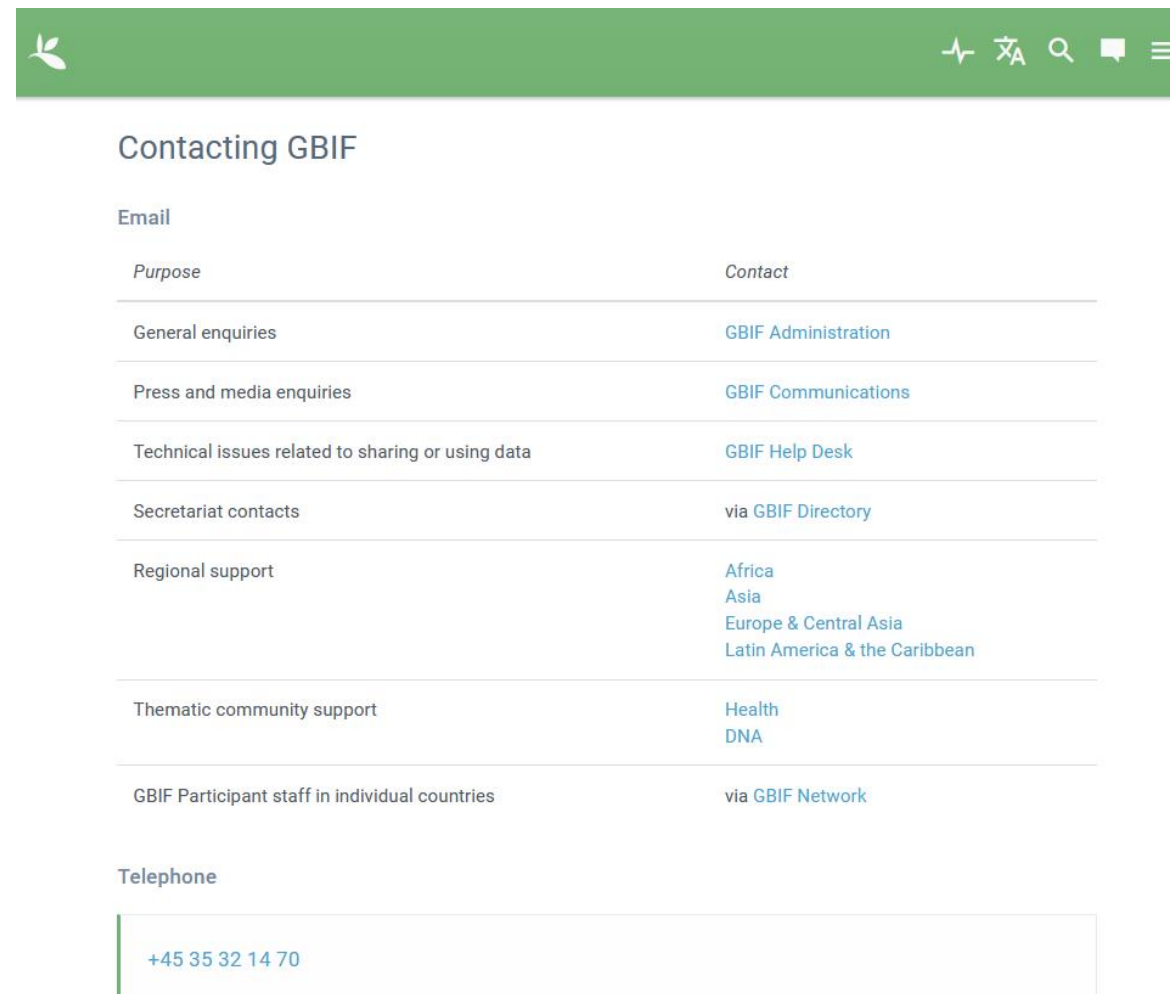
# Node or Regional support



The screenshot shows the GBIF Directory page for the Node Managers Committee. The page has a green header with a logo and navigation icons. Below the header, there are links for 'CONTACT US' and 'DIRECTORY'. A search bar is present, followed by a list of filter categories: ALL, VOTING PARTICIPANTS, ASSOCIATE COUNTRY PARTICIPANTS, OTHER ASSOCIATE PARTICIPANTS, EXECUTIVE COMMITTEE, SCIENCE COMMITTEE, BUDGET COMMITTEE, NODES STEERING GROUP, NODE MANAGERS COMMITTEE (highlighted), and GBIF SECRETARIAT. The main content area features a section titled 'Node Managers Committee' with a descriptive paragraph. Below this is a table listing the committee members.

Name	Role	Participant	Participant status
Anne-Sophie Archaubeau	Node manager Nodes committee chair	France	Voting
David Jennings	Nodes committee 1 st vice chair Node manager	iDigBio	Associate
Anabela Plos	Node manager Nodes committee 2 nd vice chair	Argentina	Associate
José Clavijo Albertos	Node manager	Andean Network of BioNET-INTERNATIONAL	Associate

<https://www.gbif.org/contact-us/directory?group=nodesCommittee>



The screenshot shows the 'Contacting GBIF' page. It has a green header with a logo and navigation icons. The main content is titled 'Contacting GBIF' and is divided into 'Email' and 'Telephone' sections. The 'Email' section contains a table with two columns: 'Purpose' and 'Contact'. The 'Telephone' section contains a single entry for a phone number.

Purpose	Contact
General enquiries	<a href="#">GBIF Administration</a>
Press and media enquiries	<a href="#">GBIF Communications</a>
Technical issues related to sharing or using data	<a href="#">GBIF Help Desk</a>
Secretariat contacts	<a href="#">via GBIF Directory</a>
Regional support	<a href="#">Africa</a> <a href="#">Asia</a> <a href="#">Europe &amp; Central Asia</a> <a href="#">Latin America &amp; the Caribbean</a>
Thematic community support	<a href="#">Health</a> <a href="#">DNA</a>
GBIF Participant staff in individual countries	<a href="#">via GBIF Network</a>

**Telephone**

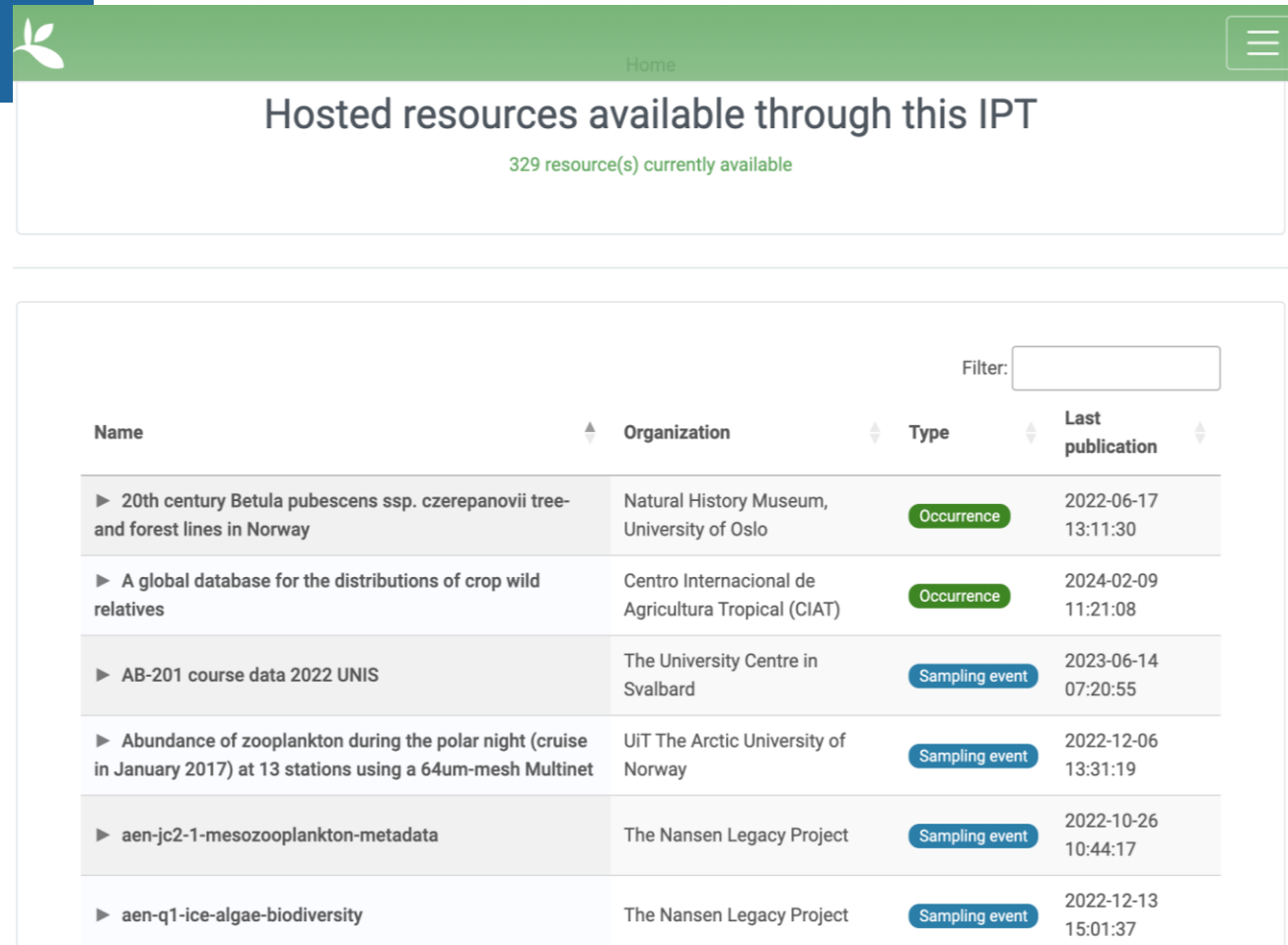
[+45 35 32 14 70](tel:+4535321470)

<https://www.gbif.org/contact-us>



# INTEGRATED PUBLISHING TOOLKIT

- Main (*but not only*) publishing tool for GBIF
- One IPT can host many datasets, on behalf of several institutions, while giving proper credit
- Test mode and production mode
- Multilingual – 7 languages
- Server-side software, needs a stable connection

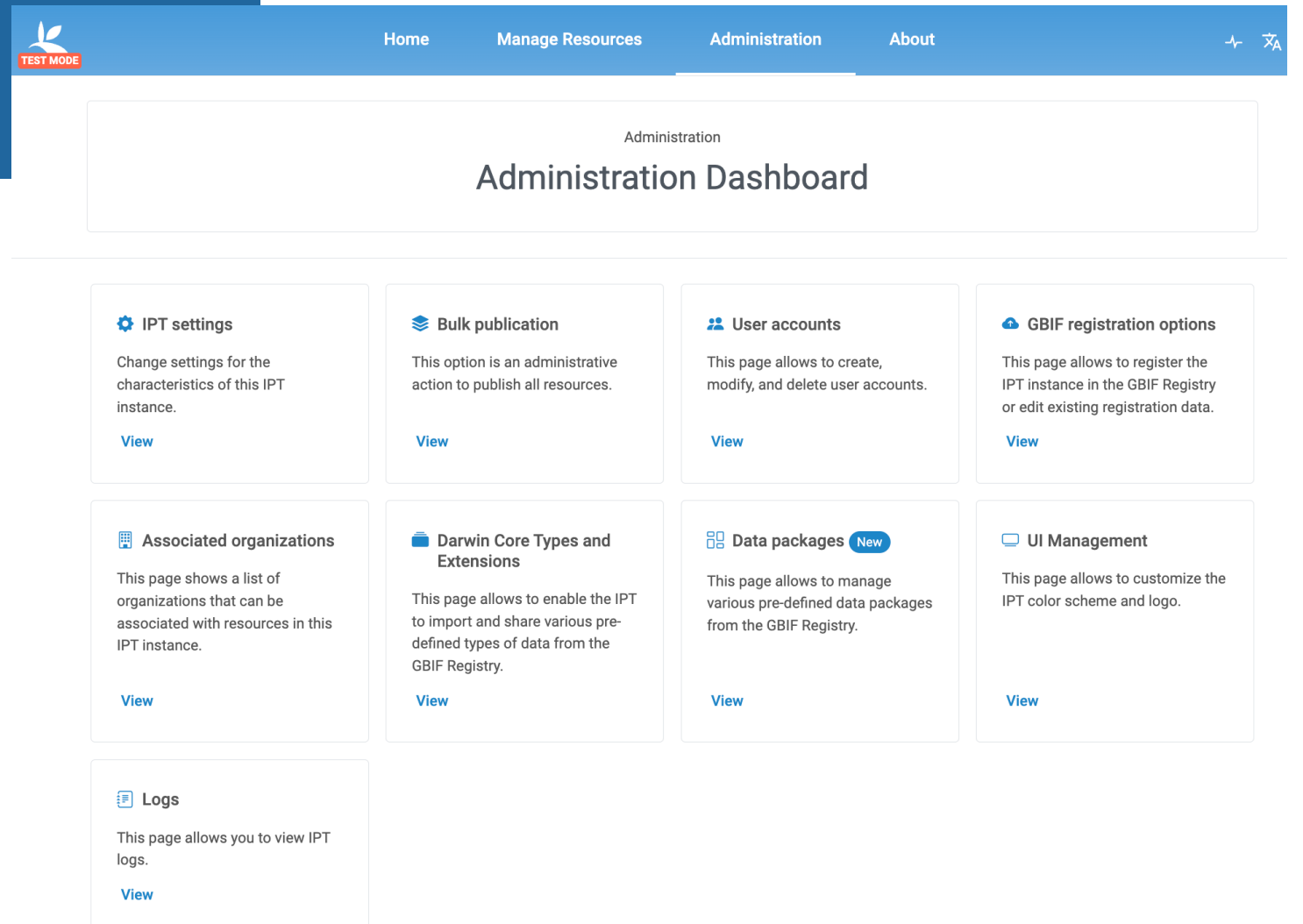


The screenshot displays the GBIF IPT interface. At the top, there is a green navigation bar with a leaf icon on the left and a 'Home' link on the right. Below the navigation bar, the main heading reads 'Hosted resources available through this IPT' with a subtext '329 resource(s) currently available'. A search filter box is located on the right side of the table. The table lists several datasets with columns for Name, Organization, Type, and Last publication. The 'Type' column uses colored buttons to indicate the resource type: green for 'Occurrence' and blue for 'Sampling event'.

Name	Organization	Type	Last publication
▶ 20th century <i>Betula pubescens</i> ssp. <i>czerepanovii</i> tree- and forest lines in Norway	Natural History Museum, University of Oslo	Occurrence	2022-06-17 13:11:30
▶ A global database for the distributions of crop wild relatives	Centro Internacional de Agricultura Tropical (CIAT)	Occurrence	2024-02-09 11:21:08
▶ AB-201 course data 2022 UNIS	The University Centre in Svalbard	Sampling event	2023-06-14 07:20:55
▶ Abundance of zooplankton during the polar night (cruise in January 2017) at 13 stations using a 64um-mesh Multinet	UiT The Arctic University of Norway	Sampling event	2022-12-06 13:31:19
▶ aen-jc2-1-mesozooplankton-metadata	The Nansen Legacy Project	Sampling event	2022-10-26 10:44:17
▶ aen-q1-ice-algae-biodiversity	The Nansen Legacy Project	Sampling event	2022-12-13 15:01:37

# IPT ADMINISTRATION

- An IPT requires an administrator to:
- Manage user accounts
- Manage publishing organizations
- Manage extensions

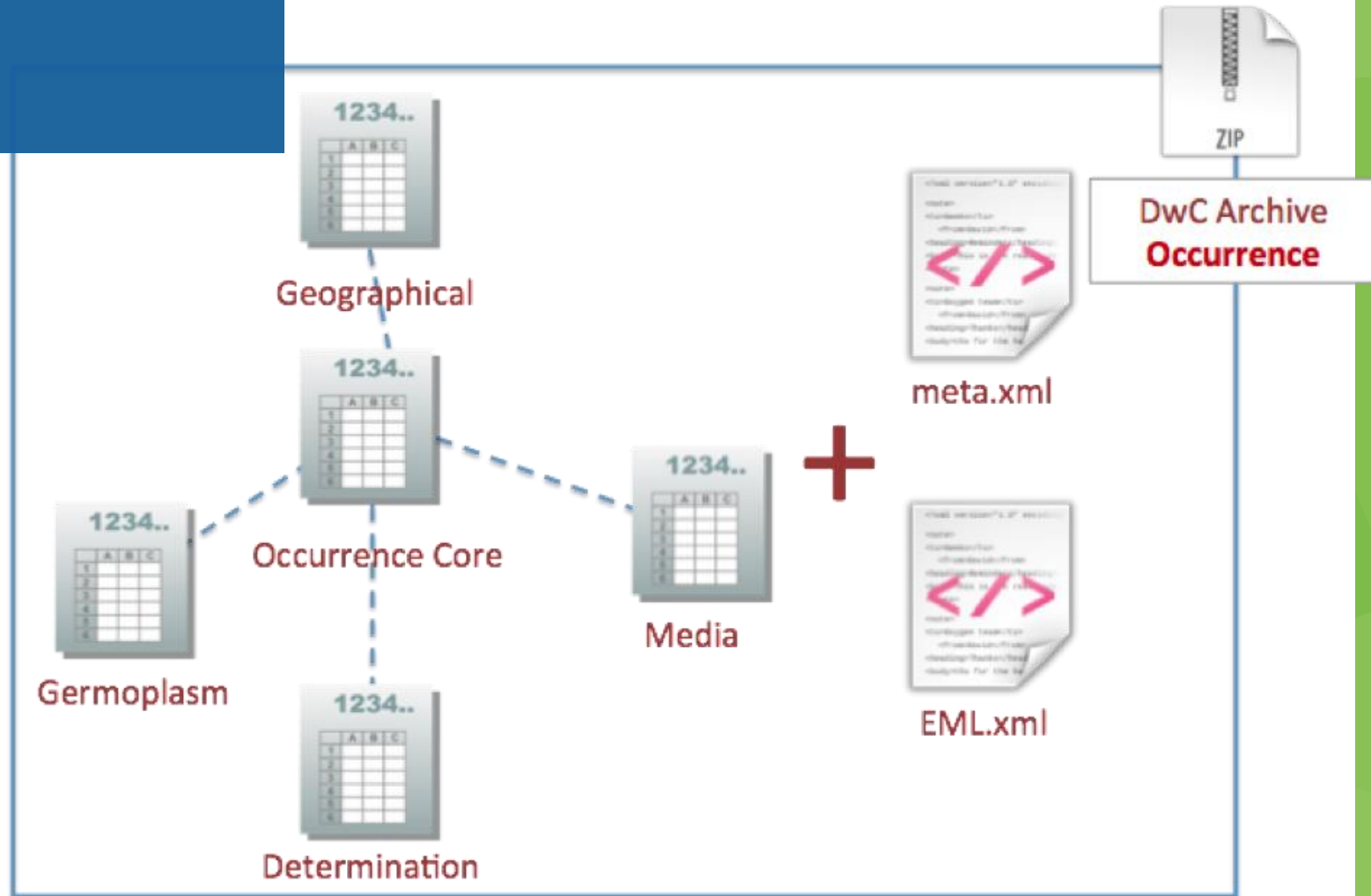


The screenshot displays the IPT Administration Dashboard. At the top, there is a navigation bar with a logo on the left, a 'TEST MODE' indicator, and menu items for 'Home', 'Manage Resources', 'Administration', and 'About'. The main content area is titled 'Administration Dashboard' and contains several interactive cards:

- IPT settings**: Change settings for the characteristics of this IPT instance. [View](#)
- Bulk publication**: This option is an administrative action to publish all resources. [View](#)
- User accounts**: This page allows to create, modify, and delete user accounts. [View](#)
- GBIF registration options**: This page allows to register the IPT instance in the GBIF Registry or edit existing registration data. [View](#)
- Associated organizations**: This page shows a list of organizations that can be associated with resources in this IPT instance. [View](#)
- Darwin Core Types and Extensions**: This page allows to enable the IPT to import and share various pre-defined types of data from the GBIF Registry. [View](#)
- Data packages** (New): This page allows to manage various pre-defined data packages from the GBIF Registry. [View](#)
- UI Management**: This page allows to customize the IPT color scheme and logo. [View](#)
- Logs**: This page allows you to view IPT logs. [View](#)

# DARWIN CORE ARCHIVE WITH EXTENSIONS

- Occurrence Core
  - Media extension
  - Determination extension
  - Germoplasm extension
- Meta file
- EML file



# EXTENSIONS IN THE IPT

Extensions can be added to a production or test IPT by an administrator.

Reach out to the TDWG and Darwin community members if you are having trouble finding an extension that suits your needs.

The screenshot shows the 'Administration' section of the IPT interface. The navigation bar includes 'Home', 'Manage Resources', 'Administration' (highlighted in green), and 'Ab'. The main content area lists several administrative functions, each with a green circular icon:

- IPT settings** (gear icon) with an information icon and a 'Publish all resources' button.
- Users accounts** (person icon).
- GBIF registration options** (cloud with up arrow icon).
- Organisations** (classical building icon).
- Core Types and Extensions** (stack of blocks icon), which is highlighted with a red rectangular border.
- Logs** (magnifying glass icon).

# DATA QUALITY REQUIREMENTS

Each core has a set of required and strongly recommended fields.

If the required information is not included, the dataset will not be indexed by GBIF.

## Darwin Core records

Term	Status
<a href="#">occurrenceID</a>	Required
<a href="#">basisOfRecord</a>	Required
<a href="#">scientificName</a>	Required
<a href="#">eventDate</a>	Required
<a href="#">countryCode</a>	Strongly recommended
<a href="#">taxonRank</a>	Strongly recommended
<a href="#">kingdom</a>	Strongly recommended
<a href="#">decimalLatitude &amp; decimalLongitude</a>	Strongly recommended
<a href="#">geodeticDatum</a>	Strongly recommended
<a href="#">coordinateUncertaintyInMeters</a>	Strongly recommended
<a href="#">individualCount, organismQuantity &amp; organismQuantityType</a>	Strongly recommended
<a href="#">informationWithheld</a>	Share if available
<a href="#">dataGeneralizations</a>	Share if available
<a href="#">eventTime</a>	Share if available
<a href="#">country</a>	Share if available

# Registered dataset Digital Object Identifier (doi)

OCCURRENCE DATASET | REGISTERED OCTOBER 13, 2022

## Distribution of marsh frogs (*Pelophylax ridibundus* complex) in Kazakhstan

Published by [Institute of Zoology of the Republic of Kazakhstan](#)

Dujsebayeva T • Kaptyonkina A • Arifulova I • Ualiyeva D • Akhmedenov K • Ivanov A • Khromov V • Krainyuk V • Sarzhanov F • Tarasovskaya N • Titov S • Timoshenko A • Ermakov O • Malakhov D • Starikov S

838 OCCURRENCES 11 CITATIONS

DATASET PROJECT METRICS ACTIVITY DOWNLOAD HOME PAGE

The presented data are the result of generalization and reconciliation of literary, museum and archival information on the distribution of lake frogs of the *P. ridibundus* complex in Kazakhstan, and new data were obtained during field work in 2021-2022. Based on the collected material, a database has been compiled for all the frog finds known today for the period from the end of the XX century to the present.

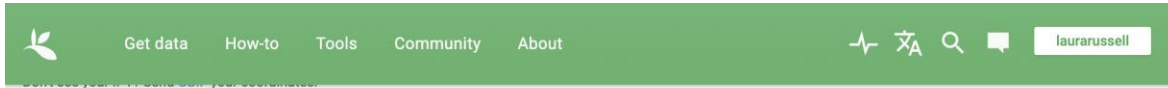
Publication date: April 9, 2023  
Metadata last modified: April 9, 2023  
Hosted by: GBIF Secretariat  
Licence: CC BY 4.0

How to cite DOI 10.15468/et4dus

838 Occurrences 100% With taxon match 99.6% With coordinates 94% With year

835 GEOREFERENCED RECORDS

# IPT references



## About the IPT

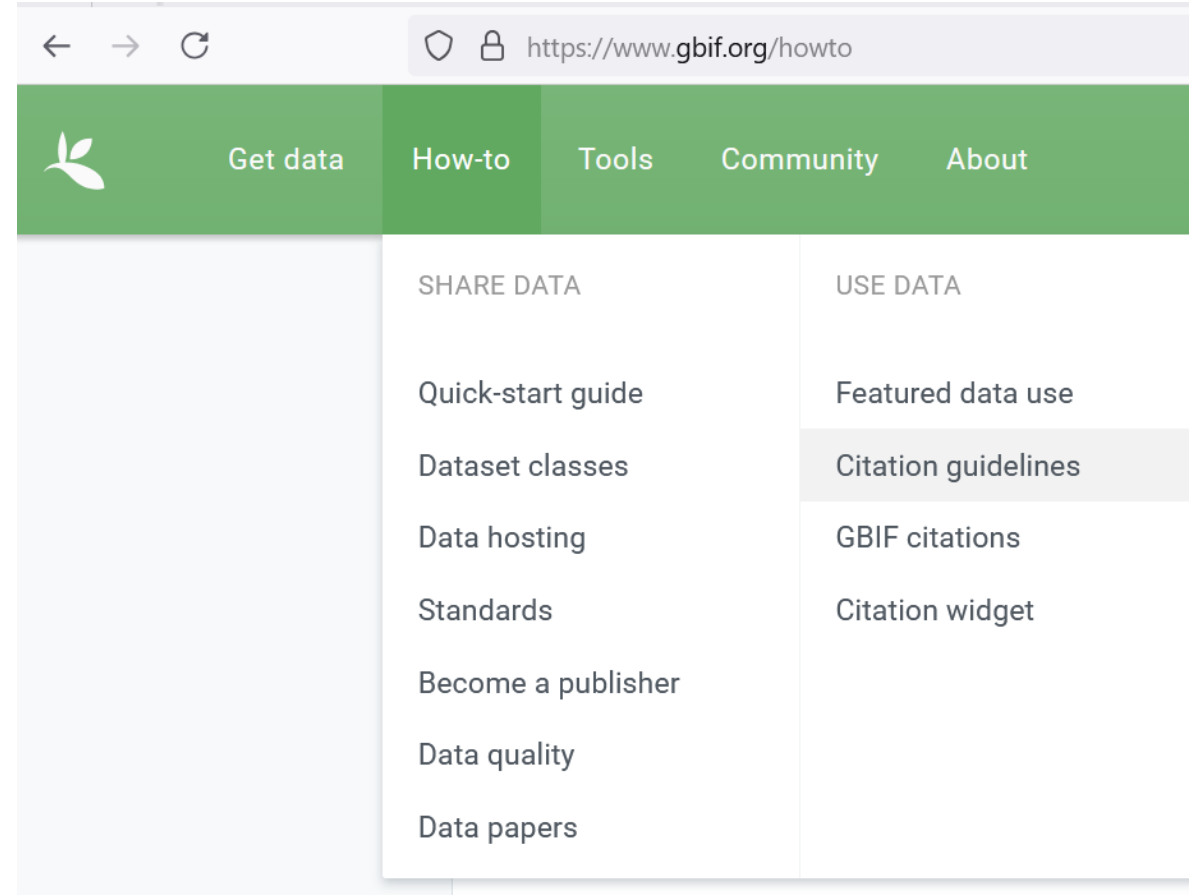
The Integrated Publishing Toolkit—commonly referred to as the IPT—is free open-source software developed by GBIF and used by organizations around the world to create and manage repositories for sharing biodiversity datasets.

GBIF currently supports three modes of use for the IPT.

- **Self-hosted**  
Dozens of organizations choose to [install](#) and [run](#) their own individual instances of the software. [Get started](#) if this "IPT classic" approach is what you came expecting.
- [National/thematic node installations](#)
- [Regional cloud-hosted installations](#)

## More information

- [Latest releases](#)
- [Installation methods](#)
- [Usage](#)
- [Data-hosting centres](#)
- [Development](#)
- [Translation](#)
- [IPT User Manual](#)



# Incentives for publishing open-access biodiversity data

- Standard formats → Integrated datasets → Improved research
- Supports different types of data
- **Different data sources:** monitoring, collections, eDNA, sensors
- **Free use of infrastructure**
- Effective citation tracking
- Properly credited for your work





## Incentives for publishing open-access biodiversity data

- Easier publishing of data papers
- Well established community of practice
  - Online training courses, manuals and guidelines
  - Community Forum
  - Technical support – by email and virtual calls
  - Data Use club

Community	About	
NETWORK	VOLUNTEERS	ACTIVITIES
Participant network	Mentors	Capacity enhancement
Nodes	Ambassadors	Programmes & projects
Publishers	Translators	Training and learning resources
Network contacts	Citizen scientists	Data Use Club
Community forum ↻		Living Atlases ↻
alliance for biodiversity knowledge ↻		



# Thank you!

ECA regional support team | [eca\\_support@gbif.org](mailto:eca_support@gbif.org)

Oleg Borodin | [oborodin@gbif.org](mailto:oborodin@gbif.org)

Salza Palpurina | [spalpurina@gbif.org](mailto:spalpurina@gbif.org)





pensoft.net

# Biodiversity data papers

Lyubomir Penev  
CEO & Founder, Pensoft  
Publishers  
[l.penev@pensoft.net](mailto:l.penev@pensoft.net)



# PENSOFT: Science Publisher & Technology Provider



Founded in 1992, based in Sofia, ca. 60 permanent staff

Publisher of 70 academic journals, books & conference materials

Pioneer in semantic publishing in biodiversity & ecology

Coordinator of the EU-funded project BiCIKL: Biodiversity Community Integrated Knowledge Library

# Open Access Journals in Biodiversity



# Why publish data?

- **Data authors, data managers and their institutions:** credit, citations, registration of priority and “data ownership”
- **Science managers:** measure impact of published data via citation metrics; proves the Open Science agenda
- **Publishers:** drive visits and citations of data to their journals
- **Data users:** facilitate data discovery, use and re-use
- **Data aggregators:** publication improves data quality!
- **Society in general:** multiplies the public investments in data collecting and maintenance

# Data publishing benefits also AI

To understand the complexity of past, recent and future changes in biodiversity and natural environments

the training and use of AI tools

should be based on

adequately curated, semantically structured and

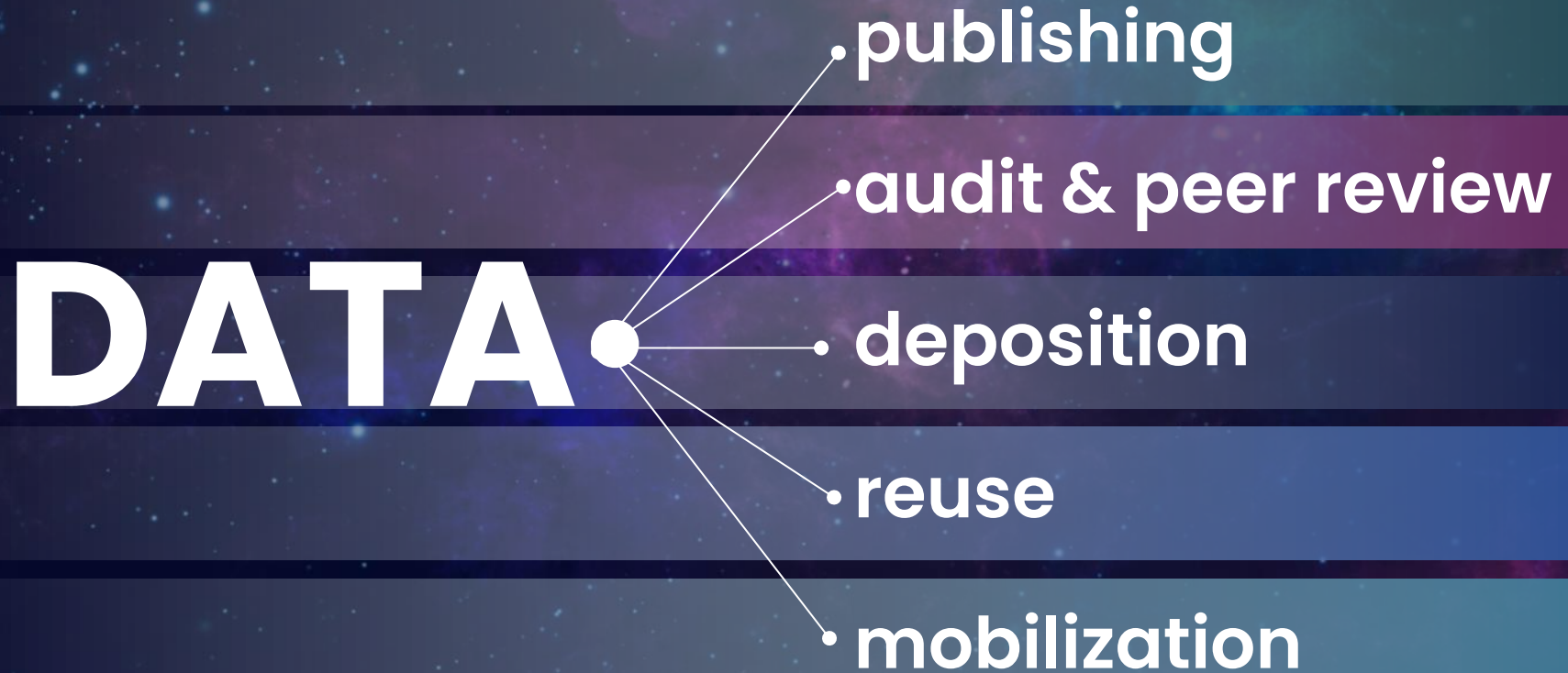
interlinked biodiversity data

# Means of data publishing

- Standalone data publishing (GBIF, GenBank, etc.)
- Data published together with a research article
  - within the article narrative (specimen records, tables)
  - supplementary file(s)
  - deposited in a repository and linked in the article
- Data published as **data papers**



# Advanced Open Data Publishing



# How to publish biodiversity data?

[Home](#)[Articles](#)[About](#)[About Pensoft](#)[Books](#)[E-Books](#)[Blog](#)[Journals](#)[My tasks](#)[Lyubomir Penev](#)[Guidelines](#)

Research Ideas and Outcomes 3: e12431

<https://doi.org/10.3897/rio.3.e12431> (28 Feb 2017)

Reviewed

v1



XML

PDF



## Strategies and guidelines for scholarly publishing of biodiversity data

▼ Lyubomir Penev, Daniel Mietchen, Vishwas Shravan Chavan, Gregor Hagedorn, Vincent Stuart Smith, David Shotton, Éamonn Ó Tuama, Viktor Senderov, Teodor Georgiev, Pavel Stoev, Quentin John Groom, David Remsen, Scott C. Edmunds

### Abstract ▲

The present paper describes policies and guidelines for scholarly publishing of biodiversity and biodiversity-related data, elaborated and updated during the Framework Program 7 EU BON project, on the basis of an earlier version published on Pensoft's website in 2011. The document discusses some general concepts, including a definition of datasets, incentives to publish data and licenses for data publishing. Further, it defines and compares several routes for data publishing, namely as (1) supplementary files to research articles, which may be made available directly by the publisher, or (2) published in a specialized open data repository with a link to it from the research article, or (3) as a data paper, i.e., a specific, stand-alone publication describing a particular dataset or a collection of datasets, or (4) integrated narrative and data publishing through online import/download of data into/from manuscripts, as provided by the Biodiversity Data Journal.

[Contents](#)[Article info](#)[Citation](#)[Metrics](#)[Reviews](#) <sup>2</sup>[Related](#)[Figs](#)[Tabs](#)[Refs](#)[Cited](#)[Article metadata](#)[Data Publishing in a Nutshell](#)[Introduction](#)[What Is a Dataset](#)[Why Publish Data](#)[How to Publish Data](#)[How to Cite Data](#)[Data Publishing Policies](#)[General Policies for Biodiversity data](#)[Data Publishing Licenses](#)[Data Deposition in Open Repositories](#)[General Information](#)[Taxonomy](#)[Species-by-Occurrence and Sample-Based data](#)



What is a “Data paper”?

# The Data paper is:

- A standard, already widely accepted type of scholarly article
- It does not analyse data, it describes data
- Extended “metadata description” of the data
- A key element of Open Science

# Data Paper concept for biodiversity (2011)



BMC Bioinformatics

[Home](#) [About](#) [Articles](#) [Submission Guidelines](#)

- Abstract
- Background
- The data paper
- Discussion
- Conclusions
- Declarations
- References

Volume 12 Supplement 15

[Data publishing framework for primary biodiversity data](#)

Research | [Open Access](#)

## The data paper: a mechanism to incentivize data publishing in biodiversity science

[Vishwas Chavan](#) <sup>†</sup>✉ and [Lyubomir Penev](#) <sup>†</sup>

<sup>†</sup>Contributed equally

*BMC Bioinformatics* 2011 12 (Suppl 15) :S2

<https://doi.org/10.1186/1471-2105-12-S15-S2>

© Chavan and Penev; licensee BioMed Central Ltd. 2011

Published: 15 December 2011

Download PDF

Export citations ▼

Metrics

Article accesses: 16537

Citations: 67 [more information](#)

Altmetric Attention Score: 76



Share This Article

# Examples of data papers

Data Paper

Biodiversity Data Journal 12: e135019

<https://doi.org/10.3897/BDJ.12.e135019> (18 Oct 2024)

XML

PDF



## Occurrences of **Neuroptera** and **Raphidioptera** in some regions in European Russia

▼ [Alexander Ruchin](#), [Vladimir Makarkin](#), [Mikhail Esin](#), [Leonid Egorov](#), [Oleg Artaev](#), [Evgeniy Lobachev](#), [Sergey Lukiyonov](#), [Vasilii Anikin](#), [Anatoliy Khapugin](#), [Gennadiy Semishin](#)

### Abstract ▲

### Background

The document presents an extensive set of data on the occurrence of **Neuroptera** and **Raphidioptera** in some regions of European Russia. The results of our own research, as well as scientific collections, have been processed. The data were collected in 17 regions. In our own research, we used different ways to obtain information, which allowed us to collect extensive material for the dataset. This dataset provides valuable information about the biodiversity of **Neuroptera** and **Raphidioptera**, the abundance of each taxon collected and the time of taxon collections.

### New information

Our dataset contains up-to-date information on the occurrence of **Neuroptera** and **Raphidioptera** in

Contents

Article Info

Cite

Metrics

Comment

Related

Figs

Tabs

Taxa

Refs

Cited

Nanopubs

[Article metadata](#)

[Introduction](#)

[General description](#)

[Sampling methods](#)

[Geographic coverage](#)

[Taxonomic coverage](#)

[Temporal coverage](#)

[Usage licence](#)

[Data resources](#)

[Additional information](#)

[Observers](#)

[Acknowledgements](#)

[References](#)

Subscribe to email alerts for current Article's categories

Email

Sign up

Data Paper

Biodiversity Data Journal 12: e117169

<https://doi.org/10.3897/BDJ.12.e117169> (16 May 2024)

XML

PDF



Contents

Article Info

Cite

Metrics

Comment

Related

Figs

Taxa

Data

Refs

Cited

Nanopubs

# The InBIO Barcoding Initiative Database: DNA barcodes of Portuguese moths

▼ Sónia Ferreira, Martin F. V. Corley, João Nunes, Jorge Rosete, Sasha Vasconcelos, Vanessa A. Mata, Joana Veríssimo, Teresa L. Silva, Pedro Sousa, Rui Andrade, José Manuel Grosso-Silva, Catarina J. Pinho, Cátia Chaves, Filipa MS Martins, Joana Pinto, Pamela Puppo, Antonio Muñoz-Mérida, John Archer, Joana Pauperio, Pedro Beja

## Abstract ▲

## Background

The InBIO Barcoding Initiative (IBI) Dataset - DS-IBILP08 contains records of 2350 specimens of moths (*Lepidoptera* species that do not belong to the superfamily *Papilionoidea*). All specimens have been morphologically identified to species or subspecies level and represent 1158 species in total. The species of this dataset correspond to about 42% of mainland Portuguese *Lepidoptera* species. All specimens were collected in mainland Portugal between 2001 and 2022. All DNA extracts and over 96% of the specimens are deposited in the IBI collection at CIBIO, Research Center in Biodiversity and Genetic Resources.

## New information

The authors enabled "The InBIO Barcoding Initiative Database: DNA barcodes of Portuguese moths"

Article metadata  
Introduction  
General description  
Project description  
Sampling methods  
Geographic coverage  
Taxonomic coverage  
Temporal coverage  
Collection data  
Usage licence  
Data resources  
Acknowledgements  
References  
Supplementary files

Subscribe to email alerts for current Article's categories



Data Paper

Biodiversity Data Journal 5: e11764

<https://doi.org/10.3897/BDJ.5.e11764> (21 Mar 2017)

XML

PDF



Contents

Article Info

Cite

Metrics

Comment

Related

Tabs

Refs

Cited

Nanopubs

Crossref (296)  
Cited-by Linking

Scopus (317)



Europe PM

M. Aranda, G. Peralta, J. Montes, F.J. Gracia, G.S. Fivash, T.J. Bouma, D. van der Wal (2022)

**Salt marsh fragmentation in a mesotidal estuary: Implications for medium to long-term management.** *Science of The Total Environment* 846: 157410.

DOI: [10.1016/j.scitotenv.2022.157410](https://doi.org/10.1016/j.scitotenv.2022.157410)

Daniel M. Alongi (2020)

**Carbon Balance in Salt Marsh and Mangrove Ecosystems: A Global Synthesis.** *Journal of Marine Science and Engineering* 8: 767.

DOI: [10.3390/jmse8100767](https://doi.org/10.3390/jmse8100767)

Ralph J. M. Temmink, Leon P. M. Lamers, Christine Angelini, Tjeerd J. Bouma, Christian Fritz, Johan van de Koppel, Robin Lexmond, Max Rietkerk, Brian R. Silliman, Hans Joosten, Tjisse van der Heide (2022)

**Recovering wetland biogeomorphic feedbacks to restore the world's biotic carbon hotspots.** *Science* 376: .

DOI: [10.1126/science.abn1479](https://doi.org/10.1126/science.abn1479)

Subscribe to email alerts for current Article's categories

Sign up

# A global map of saltmarshes

▼ Chris J Mcowen, Lauren V Weatherdon, Jan-Willem Van Bochove, Emma Sullivan, Simon Blyth, Christoph Zockler, Damon Stanwell-Smith, Naomi Kingston, Corinne S Martin, Mark Spalding, Steven Fletcher

## Abstract ▲

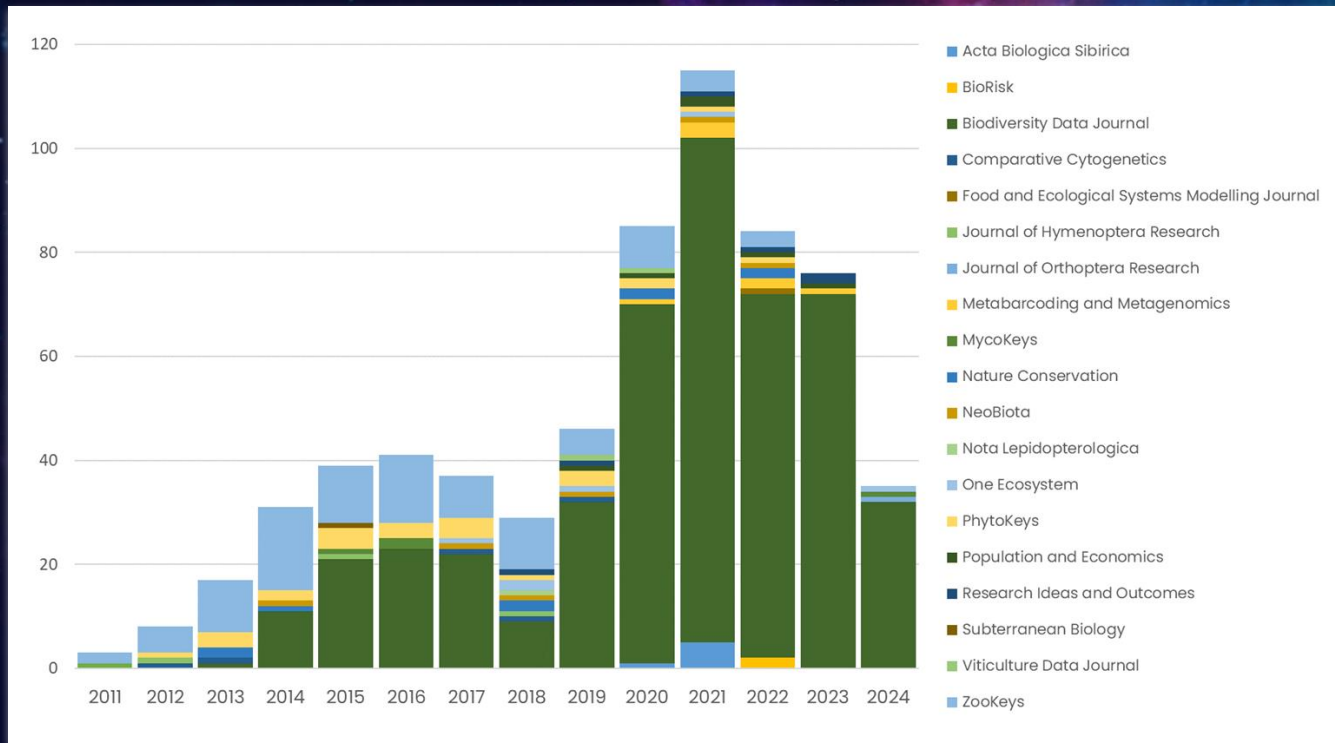
## Background

Saltmarshes are extremely valuable but often overlooked ecosystems, contributing to livelihoods locally and globally through the associated ecosystem services they provide, including fish production, carbon storage and coastal protection. Despite their importance, knowledge of the current spatial distribution (occurrence and extent) of saltmarshes is incomplete. In light of increasing anthropogenic and environmental pressures on coastal ecosystems, global data on the occurrence and extent of saltmarshes are needed to draw attention to these critical ecosystems and to the benefits they generate for people. Such data can support resource management, strengthen decision-making and facilitate tracking of progress towards global conservation targets set by multilateral environmental agreements, such as the Aichi Biodiversity Targets of the United Nations' (UN's) Strategic Plan for Biodiversity 2011-2020, the Sustainable Development Goals of the UN's 2030 Agenda for Sustainable Development and the Ramsar Convention.

## New information



# Data papers in Pensoft's journals



More than

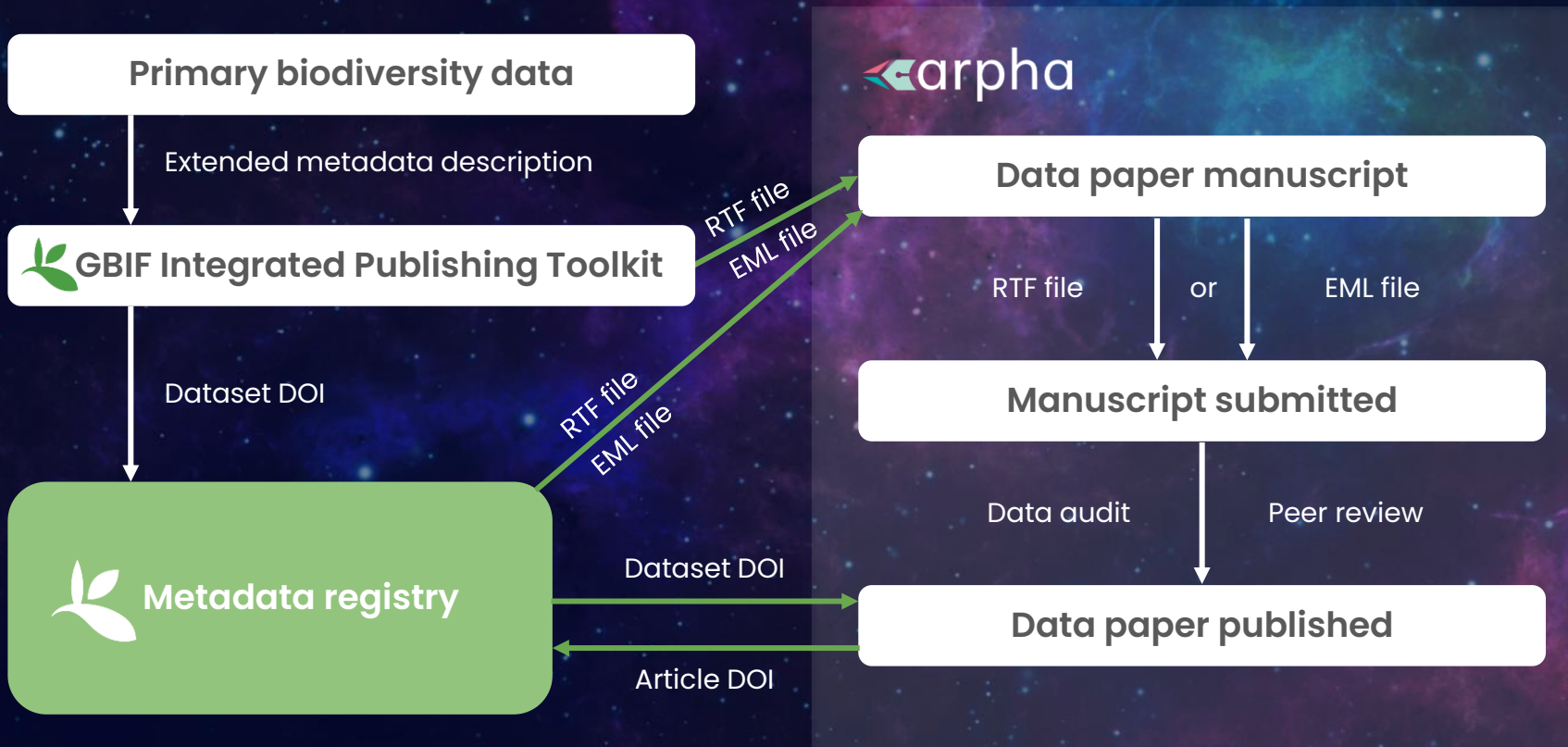
**700**

Data papers  
published  
in Pensoft's  
journals

# Publishing GBIF data papers with Pensoft

Improving data quality

# Data Paper workflows



# Download EML or RTF metadata file

The image shows two overlapping screenshots of the IPT (Institutional Publishing Tool) website. The top screenshot displays a checklist titled "A checklist to the wasps of Peru (Hymenoptera, Aculeata)", published by ZooKeys on 17 February 2011. Below the text, there are several download buttons: GBIF, DwC-A, EML, RTF, and Versions. The EML and RTF buttons are highlighted with a red box. The bottom screenshot shows a dataset page titled "A dataset of bird inventory records at Cloudbridge Nature Reserve, Talamanca Mountains, Costa Rica, between March 2016 and May 2020". It includes a navigation menu with "DATASET", "PROJECT", "METRICS", "ACTIVITY", and "DOWNLOAD". A dropdown menu is open over the "DOWNLOAD" button, showing options for "GBIF annotated archive", "Source archive", "Darwin Core Archive", and "GBIF annotated metadata EML". The "GBIF annotated metadata EML" option is highlighted with a red box. Below the dropdown, there are statistics for 40,263 occurrences, all with 100% match for taxon, coordinates, and year. A "How to cite" section provides a DOI: 10.15468/73nqqr.

# Upload the GBIF EML file and create a data paper manuscript

## Data Paper (Biosciences)

- Authors
- Contributors
- ▼ Article metadata
  - Title
  - Abstract & Keywords
  - Classifications
  - Funder
  - Nanopublications

### Introduction

- General description
- Project description
- Sampling methods
- Geographic coverage
- Taxonomic coverage

- ▼ Traits coverage
  - Data coverage of traits
- Temporal coverage
- Collection data
- Usage licence
- Data resources
- Additional information
- Acknowledgements
- Author contributions

Biodiversity Data Journal : Data Paper (Biosciences)



## A checklist to the wasps of Peru (Hymenoptera, Aculeata)

Teodor Georgiev

Corresponding author: Teodor Georgiev ([preprint@pensoft.net](mailto:preprint@pensoft.net))

© Teodor Georgiev



Citation:

### Abstract

#### Background

The first checklist to the 225 genera and 1169 reported species-group taxa of aculeate wasps of Peru is presented. The list is based on a literature survey and examination of Peruvian entomological collections and include locality references for each taxon. Bibliographic references for the identification of families, genera, and species are provided when available. The occurrence data are published in addition as a downloadable file (doi: 10.3897/zookeys.15.196.app.2.ds, doi: 10.3897/zookeys.15.196.app.3.ds, and 10.3897/zookeys.15.196.app.4.ds) and were uploaded onto GBIF infrastructure simultaneously with the publication process. The following new combinations are proposed: *Ancistroceroides cirifer* (Zavattari, 1912), *Ancistrocerus epicus* (Zavattari, 1912), and *Stenodynerus corallineipes* (Zavattari, 1912).

### Keywords

Aculeata, Biodiversity, Gazetteer, Hymenoptera, Sampling coverage, Peru

### Introduction

# Data audit & peer review

Improving data quality

# Author-performed data check



A peer-reviewed open-access journal  
**Biodiversity  
Data Journal**  
Making your data count! ISSN 1314-2828 (online)



Full Text: Author Title

[Submit manuscript](#)[My drafts](#)[My tasks](#)[About](#)[Articles](#)[Topical collections](#)[Nanopublications](#)[Author Guidelines](#)[Editorial Team](#)[Contacts](#)

## About

[Author Guidelines](#)[Data Publishing Guidelines](#)[Data Quality Checklist and Recommendations](#)[FAIR Data Checklist](#)[Linked Data Table for Primary Biodiversity Data](#)[Data Review Guidelines](#)[Omics Data Papers](#)[What is "omics" data?](#)[Where do I deposit my omics data and metadata?](#)

## CHECKLIST

### Characters

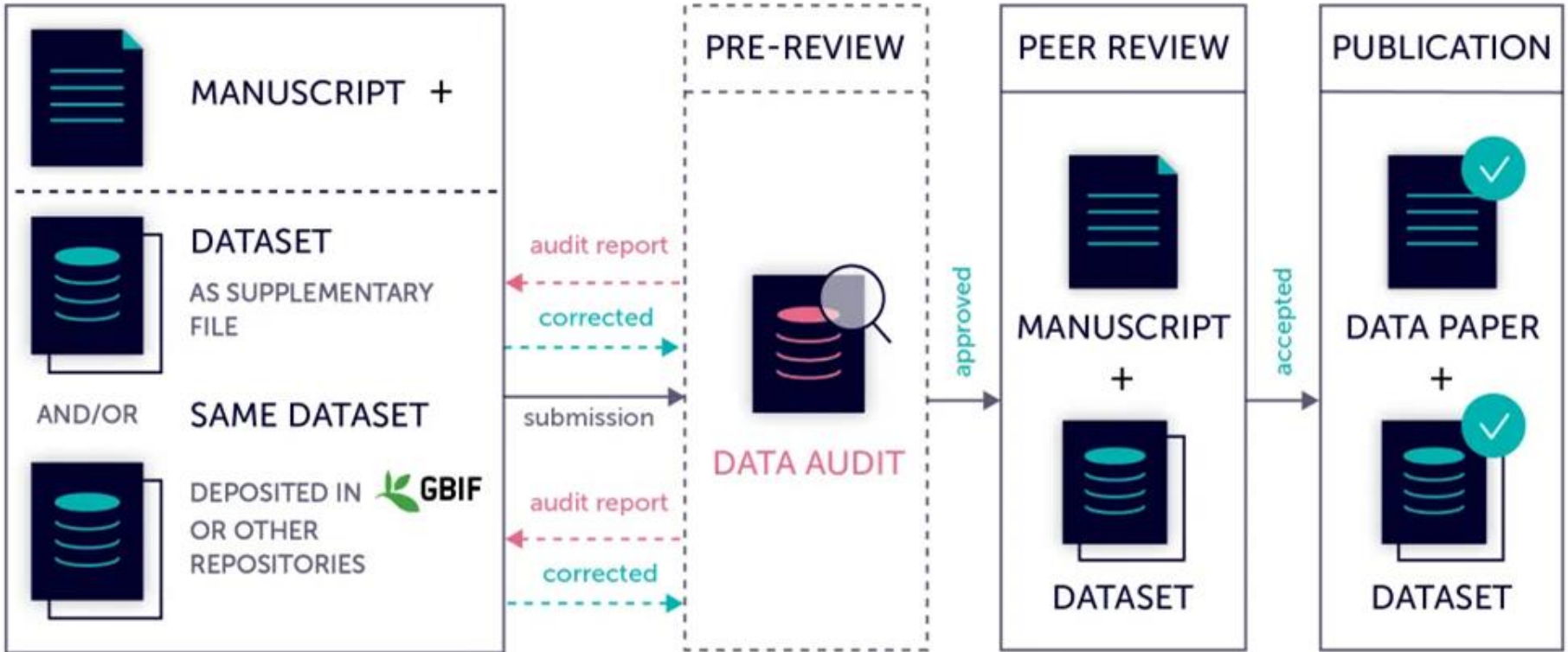
- The dataset is UTF-8 encoded
- The only characters used that are not numbers, letters or standard punctuation, are tabs and whitespaces
- Each character has only one encoding in the dataset
- No line breaks within data items
- No field-separating character within data items (tab-separated data preferred)
- No "?" or replacement characters in place of valid characters
- No Windows carriage returns
- No leading, trailing, duplicated or unnecessary whitespaces in individual data items

### Records

- No broken records, i.e. records with too few or too many fields
- No blank records



# Editor-performed data check





# Data audit report: list of errors



- inappropriately used fields
- non-compliance with the Darwin Core recommendations
- pseudo-duplicates
- data inconsistency

(7) *municipality* has "\_" for EC42F49A-68D5-4504-8F9E-0010859712A1.

(8) *locality* needs cleaning for the many pseudo-duplicates, e.g.

```
2 casco urbano, avda. del Brillante, nº 187, carril de la Huerta de los Arcos
7 casco urbano, avda. del Brillante, nº 187, carril Huerta de los Arcos
7 casco urbano, avda. del Brillante, nº 187, Carril Huerta los Arcos
```

and the many unnecessarily quoted entries, e.g.

```
"casa ""Rompealbardas""""
""""Villa Carmen""", ""El Calvario""""
```

Also, *locality* is "\_" for CC465E40-9868-4B01-8D2B-5CB9AC747674 and 8547AA0D-682B-4848-B31F-0399427D51FA

(9) *decimalLatitude* errors:

```
1 30S266977.44
1 37,91560°
1 40.9449°
1 41.9425N
```

Also, several entries have too many significant figures and should be rounded off, e.g. "37.0233172796695"

# Data audit report: recommendations



Data audit for technical evaluation of

## Vascular plants dataset of the COFC herbarium (University of Cordoba, Spain) (associated GBIF dataset)

Downloaded on 2019-06-19 from <https://www.gbif.org/dataset/837c0162-f762-11e1-a439-00145eb45e9a>

Dr Robert Mesibov (robert.mesibov@gmail.com; <https://www.datafix.com>)  
2019-06-20

### About this evaluation

Pensoft does a technical evaluation of the dataset (or datasets) referred to in the data paper problems, the data paper manuscript is referred to reviewers. If the dataset has major problems the dataset has been corrected.

To see what features of a dataset are checked in a technical evaluation, please go to <https://zookeys.pensoft.net/about#DataQualityChecklistandRecommendations>

Please note that Pensoft does not check the details of the *content* of a dataset, for example scientific name, or whether the correct latitude/longitude is given for a locality.

**Recommendation.** The dataset associated with the manuscript has been processed by GBIF and the data paper could go on to review. However, there are many data problems in the GBIF upload, and I recommend to the authors that these problems be fixed and the data re-uploaded to GBIF for processing. The problems are detailed below by Darwin Core field in the field order in the dataset.

Many of the problems are not trivial and are causing data loss. For example, the 4360-86C7-5C9D483D6DAE is "30S266977.44". GBIF has rejected the location (<https://www.gbif.org/occurrence/2235670578>).

**Recommendation.** The dataset associated with the manuscript has been processed by GBIF and the data paper could go on to review. However, there are many data problems in the GBIF upload, and I recommend to the authors that these problems be fixed and the data re-uploaded to GBIF for processing. The problems are detailed below by Darwin Core field in the field order in the dataset.

Many of the problems are not trivial and are causing data loss. For example, the decimalLatitude in FF92A873-601C-4360-86C7-5C9D483D6DAE is "30S266977.44". GBIF has rejected the location as "Coordinate invalid" (<https://www.gbif.org/occurrence/2235670578>).

In addition, the number of records in the dataset uploaded to GBIF (**verbatim.txt** in the user download) does not always agree with the number of records given in the data paper:

Family	Data paper	verbatim.txt
Asteraceae	8625	8625
Fabaceae	7929	7929
Poaceae	6324	5513 (as stated in paper, without 811 <i>Festuca</i> records)
Lamiaceae	3105	3087
Caryophyllaceae	2156	2156
Plantaginaceae	2023	750
Brassicaceae	1851	1852
Apiaceae	1707	1707
Ranunculaceae	1319	1318
Boraginaceae	1111	1289

Genus	Data paper	verbatim.txt
Centaurea	1538	1537
Trifolium	1406	1406
Euphorbia	708	750

# Data published in GBIF by the journal

Taxonomy & Inventories

Biodiversity Data Journal 12: e124006  
<https://doi.org/10.3897/BDJ.12.e124006> (25 Apr 2024)

XML

PDF



## A new species of *Otacilia* Thorell, 1897 (Araneae, Phrurolithidae) from Yintiaoling National Nature Reserve, Chongqing, China

▼ Changbin Zheng, Yannan Mu

Abstract ▲

Background

Phrurolithidae is a family of spiders with 395 species belonging to 26 genera distributed worldwide, of which 205 species belong to 17 genera was recorded in China.

New information

Contents Article Info Cite Metrics Comment Related

Figs Map Taxa **Data** Refs Cited Nanopubs

💡 Tables and Figures, if present, can be downloaded from the article.

📄 [Explore this dataset on GBIF](#)

📄 [Download all occurrences as Darwin Core Archive](#)

📄 [Download all treatments as Darwin Core Archive](#)

# Data visualised on GBIF



[Get data](#) [How-to](#) [Tools](#) [Community](#) [About](#)

[Home](#) [Search](#) [Help](#) [Login](#)

TREATMENT ARTICLE | REGISTERED MAY 3, 2024

## A new species of *Otacilia* Thorell, 1897 (Araneae, Phrurolithidae) from Yintiaoling National Nature Reserve, Chongqing, China

Mediated by [Biodiversity Data Journal](#)

Mu Y

[DATASET](#) [TAXONOMY](#) [METRICS](#) [ACTIVITY](#) [DOWNLOAD](#)

1 MATERIAL EXAMINED 1 RECORD

Phrurolithidae is a family of spiders with 395 species belonging to 26 genera distributed worldwide, of which 205 species belong to 17 genera was recorded in China. A new species of the genus *Otacilia* Thorell, 1897 is described from Yintiaoling National Nature Reserve, Chongqing, China. Diagnosis, morphological description, photos of the habitus and genitalia of the new species are provided.

**Publication date:** April 24, 2024

**Metadata last modified:** May 3, 2024

**Hosted by:** [Biodiversity Data Journal](#)

**Licence:** CC BY 4.0

[How to cite](#) [DOI](#) 10.3897/bdj.12.e124006



1 Occurrences



100% With taxon match



100% With coordinates



100% With year



1 Accepted names



0 Synonyms



85% Overlap with GBIF Backbone



71% Overlap with Catalogue of Life

1 GEOREFERENCED RECORD



# Taxon descriptions and specimen records on GBIF



Get data How-to Tools Community About



## Classification

Select a species

You are browsing: A new species of *Otacilia* Thorell, 1897 (Araneae, Phrurolithidae) from Yintiaoling National Nature Reserve, Chongqing, China

Kingdom Animalia  
Phylum Arthropoda  
Class Arachnida  
Order Araneae  
Family Phrurolithidae  
Genus *Otacilia*  
Species *Otacilia wuxi* Zheng & Mu

No children

SPECIES | ACCEPTED

## *Otacilia wuxi* Zheng & Mu

In: Zheng C, Mu Y (2024) A new species of *Otacilia* Thorell, 1897 (Araneae, Phrurolithidae) from Yintiaoling National Nature Reserve, Chongqing, China. Biodiversity Data Journal 12: e124006. <https://doi.org/10.3897/BDJ.12.e124006>

Mediated through: Biodiversity Data Journal

TREATMENT VERSATIM

SOURCE

Male: total length 5.01, carapace 2.19 long, 1.91 wide; abdomen 2.59 long, 1.67 wide. Eye sizes and interdistances: AME 0.14, ALE 0.15, PME 0.11, PLE 0.13, AME-AME 0.04, AME-ALE 0.02, PME-PME 0.16, PME-PLA 0.09, ALE-PLA 0.15. MOA 0.37 long, anterior width 0.30, posterior width 0.42. Clypeal height 0.19. Chelicerae with three promarginal and eight retromarginal teeth. Measurements of legs: I 8.86 (2.34+3.44+2.00+1.08), II 7.05 (1.87+2.59+1.55+1.04), III 6.18 (1.70+1.93+1.58+0.97), IV 9.46 (2.52+2.95+2.62+1.37). Spinatio: tibia I pv 8 rv 8, tibia II pv 7 rv 7, metatarsus I pv 4 rv 4, metatarsus II pv 3 rv 3. Legs yellow. Carapace yellow, with several indistinct shapes resembling flowing water droplets beside fovea. Abdomen yellow, with a small, thin dorsal scutum and irregular black pattern anterior and four black chevron stripes posterior (Fig. 1A).

**Palp.** Femoral apophysis high, located at middle part of femur, well-developed (Fig. 1C and D). Dorsal tibial apophysis long and large, strongly curved as semi-elliptic, base wide, tapering from middle to tip (Fig. 1C and D); prolateral tibial apophysis distinct (Fig. 1B). Tegulum bean-shaped, wider than cymbium; tegular apophysis semicircular. Conductor membranous (Fig. 1B). Sperm duct obvious, tapering from retrolateral of tegulum to embolus. Embolus long, needle-like, strongly curved retrolaterally from basal part (Fig. 1B).

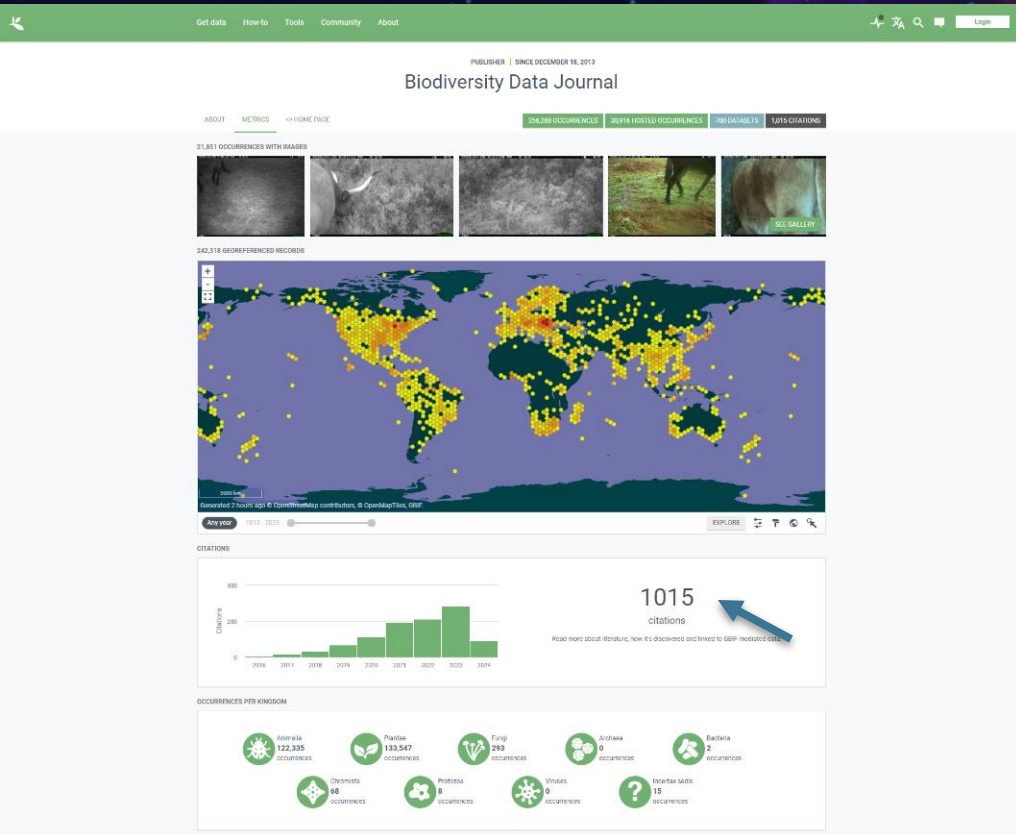
Female: unknown.

## FIGURES



Figure 1  
*Otacilia wuxi* sp. nov., male, holotype. **A** habitus; **B** left palp, ventral view; **C** same, prolateral view; **D** same, retrolateral view. Abbreviations: C—conductor; DTA—dorsal tibial apophysis; E—embolus; FA—femoral

# Datasets from the Biodiversity Data Journal



# Import of occurrence records from GBIF into manuscripts

## Taxonomy & Inventories

Authors

Contributors

### Article metadata

Title

Abstract & Keywords

Classifications

Funder

Nanopublications

Introduction

Materials and methods

Data resources

### Taxon treatments

Cyclamen hederifolium

Taxon name

External Links

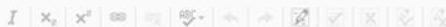
Nomenclature

Materials

Treatment sections

Checklists

Identification keys



Corresponding author: Teodor Georgiev (preprint@pensoft.net)

© Teodor Georgiev

Citation:



## Taxon treatment

*Cyclamen hederifolium* Aiton

Material [Download as CSV](#)

a. scientificName: *Cyclamen hederifolium* Aiton; taxonConceptID: NBNSYS0000003933; taxonomicStatus: ACCEPTED; taxonID: NBNSYS0000003933; kingdom: Plantae; phylum: Tracheophyta; class: Magnoliopsida; order: Ericales; family: Primulaceae; taxonRank: SPECIES; vernacularName: Sowbread; genus: Cyclamen; specificEpithet: hederifolium; continent: EUROPE; country: United Kingdom of Great Britain and Northern Ireland; countryCode: GB; stateProvince: England; locality: Copse Hill, Brighton and Hove, Brighton BN1 5EG, UK; decimalLatitude: 50.858986; decimalLongitude: -0.155622; geodeticDatum: WGS84; coordinateUncertaintyInMeters: 50; eventID: 24740609; eventDate: 2024-01-13; startDayOfYear: 13; endDayOfYear: 13; year: 2024; month: 1; day: 13; individualCount: 1; occurrenceDetails: <http://api.gbif.org/v1/occurrence/4594204304>; recordedBy: Rudling, M. Milly; occurrenceStatus: PRESENT; collectionCode: iNaturalist | UK and IOM data; basisOfRecord: HUMAN\_OBSERVATION; occurrenceID: D4572408-6EA4-5F3B-9801-947D19864F23

# Parsing occurrence records using AI

Search sections...

Title

Abstract

Keywords

> Introduction

> Materials and methods

✓ Data resources

No subsection

✓ Taxon treatments

└─ Taxon

└─ External Links

└─ Nomenclature

└─ > Materials

## Data resources

Enter section content

## Taxon treatments

### Materials [Download as CSV](#) or [XLSX](#)

#### Holotype :

a. catalogNumber: HBUMM08381-spec. 1; recordedByID: Chen, Tian; locationID: Zhangjiadi, Yunhe County, Lishui, Zhejiang Province; locality: around oaks in remote forest; verbatimElevation: c. 820 m a.s.l.; decimalLatitude: 27.974; decimalLongitude: 119.379; eventDate: 2019-08; basisOfRecord: PreservedSpecimen;

#### Paratype :

a. catalogNumber: HBUMM08381-spec. 2; recordedByID: Chen, Tian; locationID: Zhangjiadi, Yunhe County, Lishui, Zhejiang Province; locality: around oaks in remote forest; verbatimElevation: c. 820 m a.s.l.; decimalLatitude: 27.974; decimalLongitude: 119.379; eventDate: 2019-08; basisOfRecord: PreservedSpecimen;

b. catalogNumber: HBUMM08370-spec. 1; recordedByID: Ye, Shi-Han; locationID: Mihougu, Fengyangshan, Longquan County, Lishui, Zhejiang Province; locality: Mihougu; verbatimElevation: 1100 m a.s.l.; decimalLatitude: 7.897; decimalLongitude: 119.159; eventDate: 2019-08-26; basisOfRecord: PreservedSpecimen;



# Data mobilization



Filling the gaps

# Special GBIF Data papers collections in BDJ



## Biota of Russia

(2020-2021)

Papers published: **59**

Total pages: **1204**

Unique views: **130627**

Total views: **202481**



## Biota of Northern Eurasia

(2022-2023)

Papers published: **11**

Total pages: **202**

Unique views: **16991**

Total views: **25287**



## Soil biodiversity

(2023-2024)

Papers published: **3**

Total pages: **86**

Unique views: **4009**

Total views: **5644**



Collection of data papers  
on the biota of  
Central Asia?

**More than  
a decade  
together  
GBIF &  
Pensoft**

Incentivizing data publishing

Improving data quality

Gathering more data

Saving time, minimizing errors

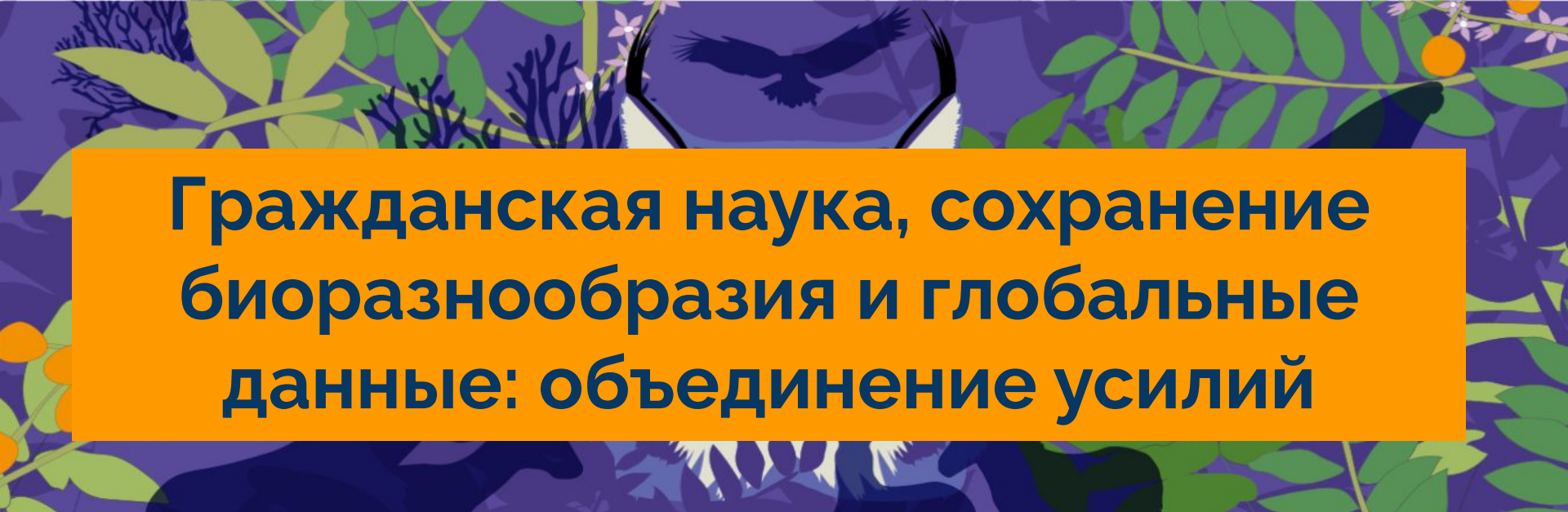
Filling the gaps

# Share your data through GBIF!

## Publish its metadata as data papers!



Lyubomir Penev,  
[l.penev@pensoft.net](mailto:l.penev@pensoft.net)  
Pensoft Publishers  
[pensoft.net](http://pensoft.net)



# Гражданская наука, сохранение биоразнообразия и глобальные данные: объединение усилий

**Александр Дубынин**

**Институт ботаники и фитоинтродукции**

**МЕЖДУНАРОДНЫЙ СИМПОЗИУМ**

**«ДАННЫЕ О БИОРАЗНООБРАЗИИ ГОР И ПУСТЫНЬ  
ЕВРАЗИИ», 18–19 НОЯБРЯ 2024 г., АЛМАТЫ, КАЗАХСТАН**

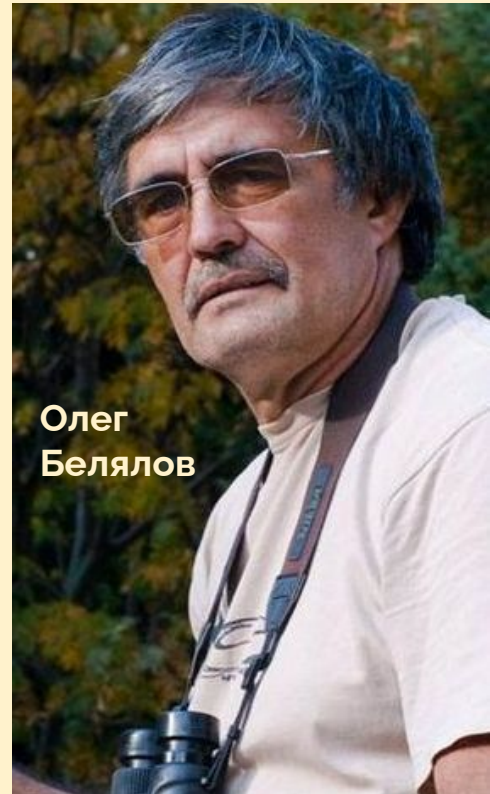


# Необходимые пояснения

**Гражданская наука** (citizen science, community science), или научное волонтерство — концепция проведения научных исследований с привлечением широкого круга научных волонтеров.

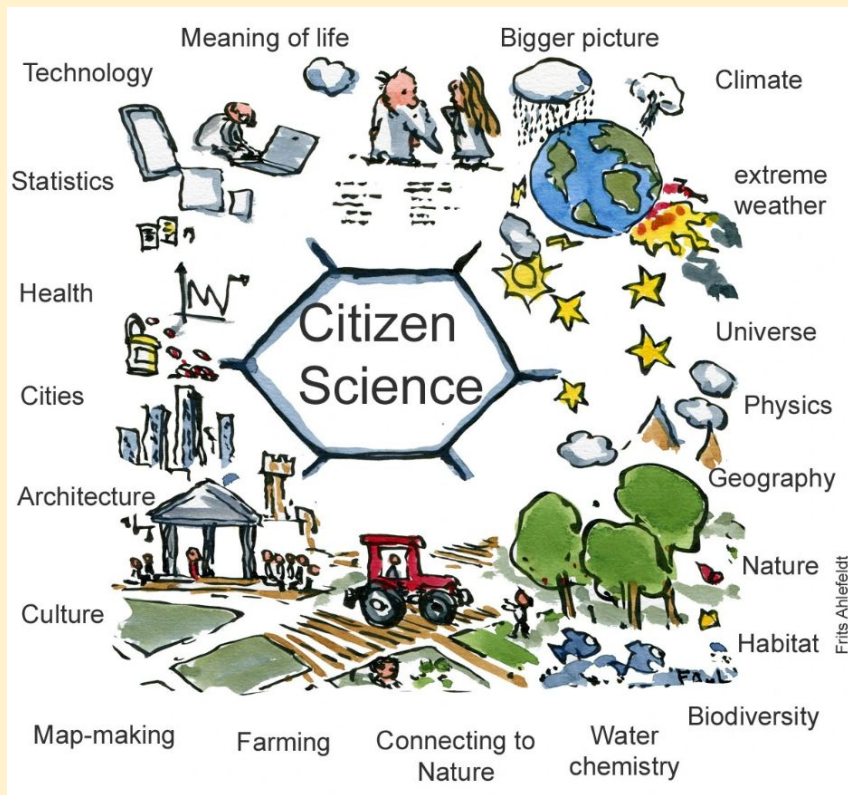
**"Citizen scientist" — это исследователь-любитель, человек, который вносит вклад в научные исследования на добровольной основе.**

**Биоразнообразие** — разнообразие жизни во всех её проявлениях, а также показатель сложности биологической системы, разнокачественности её компонентов. Выделяют генетическое разнообразие (разнообразие генов и их вариантов — аллелей), видовое разнообразие (разнообразие видов в экосистемах) и экосистемное разнообразие (разнообразие самих экосистем).



Олег  
Белялов

# Разнообразие гражданской науки



астрономия

технологии

статистика

города

архитектура

культура

карты

фермерство

связь с природой

химия воды

биоразнообразие

жилище

география

физика

Вселенная

климат

экстремальная

погода





## ● Zooniverse

- **Более 100 активных проектов** в областях астрономии, экологии, гуманитарных наук и др.
- **Сообщество**: свыше **2 миллионов** зарегистрированных пользователей по всему миру.

## ● SciStarter

- **Более 3 тысяч проектов** гражданской науки, доступных по всему миру.
- **Особенности**: поиск проектов по теме, местоположению, навыкам и времени участия.

## ● eBird

- **Свыше 1 миллиарда** записей наблюдений птиц.
- **Участники**: сотни тысяч орнитологов-любителей и профессионалов.

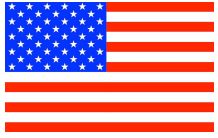
## ● Foldit

- **Тысячи пользователей** помогают решать задачи по сворачиванию белков через онлайн-игру.
- **Научные достижения**: вклад в открытие новых белковых структур, важных для медицины.

## ● Globe at Night

- **Более 200 тысяч наблюдений** уровня светового загрязнения со всего мира.
- **Цель**: понимание и снижение воздействия светового загрязнения на ночное небо и экосистемы.





**США: 502 проекта**

[About](#) ▼ [Catalog](#) ▼ [Toolkit](#) ▼

[Blog](#)

Helping federal agencies accelerate innovation through public participation.

[Learn more](#)



### Explore Projects

This searchable database provides a government-wide listing of citizen science and crowdsourcing projects



### Join Us

There are two primary groups within the federal government working collaboratively to advance the use of

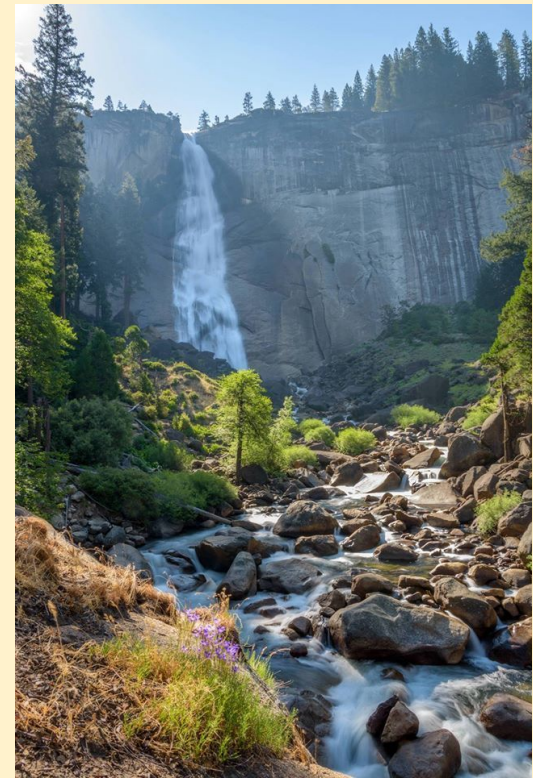


### Plan Your Projects

The Toolkit provides five basic process steps for planning, designing and carrying out a crowdsourcing or

# Факт 1. Вклад научных волонтеров в США сопоставим с госфинансированием науки

- ★ **Элинор Дженкинс Теобальд и ее коллеги** провели количественный обзор гражданской науки, связанной с биоразнообразием, чтобы определить, **могут ли данные, собранные в рамках этих проектов, эффективно использоваться, и как они в настоящее время используются в исследованиях биоразнообразия.**
- ★ **388 проектов**, в основном реализуемых в США, 1,3 миллиона добровольцев, их вклад **2,5 миллиардов долларов в натуральной форме ежегодно.**
- ★ **Вывод:** укрепление связей между профессиональными и непрофессиональными участниками научного процесса позволит лучше использовать этот **большой ресурс данных для понимания и устранения последствий глобальных изменений для биоразнообразия.**



## Факт 2. Проекты гражданской науки вносят большой вклад в образование

- ★ **Мария Питер и ее коллеги** исследовали роль проектов CS как формата образования в области **окружающей среды и устойчивого развития**, а также в естественнонаучном образовании.
- ★ Онлайн-опрос **1160 добровольцев**, участников CS в **63 проектах BDCS** в Европе, Австралии и Новой Зеландии.
- ★ **Эффект участия**: **высокий потенциал!** (1) **увеличение знаний, самооффективности, интереса и мотивации** оказалось более выраженным в отношении окружающей среды, а не науки, (2) высокий **прирост навыков сбора данных**, (3) **здоровье и благополучие, удовольствие, чувство удовлетворения, усиление связи с людьми и природой**, а также **более пессимистичный взгляд на будущее окружающей среды** (пессимист — это хорошо информированный оптимист!).



## Факт 3. Приоритет в сборе данных: распределение видов и оценка численности

- ★ **Марк Чандлер и его коллеги** оценили проекты CS на предмет их вклада в основные переменные биоразнообразия (**Essential Biodiversity Variables/EBV**).
- ★ **Вывод:** существующие проекты CS предоставляют значительные по охвату данные о распределении видов и численности популяций, фенологии, первичной и вторичной продуктивности экосистем.
- ★ **Мониторинг птиц, чешуекрылых и растений проводится повсеместно**, в то время как другие таксоны вызывают большой интерес только в отдельных регионах (грибы, амфибии, рептилии и таксоны коралловых рифов). Большинство проектов CS находятся в Европе, Северной Америке, Южной Африке, Индии и Австралии (**Северная и Центральная Азия, Южная Америка отстают**).



# Global Biodiversity Information Facility (GBIF)

- ★ **Крупнейший агрегатор данных** о биоразнообразии планеты
- ★ Это **международная сеть и инфраструктура** данных, финансируемая правительствами стран мира и направленная на предоставление любому человеку в любом месте открытого доступа к данным обо всех видах жизни на Земле.
- ★ GBIF объединяет **разнообразные источники данных с помощью стандартов данных, включая Darwin Core**.
- ★ В настоящий момент более 2000 организаций опубликовали на сайте GBIF информацию о более чем **3,01 миллиардов записей**
- ★ Сайт GBIF <http://gbif.org>



# TOP10 поставщиков сведений в GBIF

- ★ **Cornell Lab of Ornithology (eBird) 1 512 428 698**
- ★ UMS PatriNat (OFB-CNRS-MNHN), Paris 169 819 392
- ★ SLU Artdatabanken, Sweden 108 088 266
- ★ **iNaturalist.org 101 233 851**
- ★ Observation.org 101 144 945
- ★ Finnish Biodiversity Information Facility 46 348 955
- ★ The Norwegian Biodiversity Information Centre (NBIC) 35 137 107
- ★ Danish Ornithological Society 34 077 388
- ★ Butterfly Conservation 31 255 557
- ★ United States Geological Survey 30 994 402

The logo for eBird, featuring the word "eBird" in a green serif font.

**RG 136 657 864**

# GBIF и iNaturalist

- ★ В GBIF попадают не все записи из iNaturalist, а только имеющие «исследовательский уровень» и один из трех типов открытых лицензий (**CCo, CC BY, CC BY-NC**).
- ★ В январе 2023 года платформа iNaturalist стала крупнейшим поставщиком данных по сосудистым растениям мира в GBIF.
- ★ Тройка лидеров поставщиков GBIF-данных по сосудистым растениям из iNaturalist на 8 января 2023 г.: США (9 993 265 наблюдений), **Россия (1 959 783)** и Канада (1 821 636).







# История iNaturalist: Беркли у истоков

- ★ **2008** iNaturalist.org — диплом студентов **Беркли Нейта Эгрин, Джессики Клайн и Кен-ичи Уеда**. Позднее к ним присоединился веб-разработчик **Син Макгрегор**
- ★ **2011** **Кен-ичи Уеда** начал сотрудничество со **Скотом Лоури**, из Стэнфорда, своим преподавателем в Беркли. Сейчас они на равных директора iNaturalist
- ★ **24.04.2014** iNaturalist вошел в состав **Калифорнийской Академии наук**
- ★ **2014** на iNaturalist **1 000 000** наблюдений
- ★ **2017** iNaturalist получил статус «совместной инициативы» Калифорнийской Академии наук и National Geographic (США)
- ★ **2023** iNaturalist стал **НКО** со статусом 501 с(3)
- ★ **Сентябрь 2024** 29<sup>th</sup> Heinz Awards вручена Скоту Лоури и Кен-ичи Уеда
- ★ **14.11.2024** на iNaturalist **218 623 832** наблюдений



«Самое красивое место  
в Интернете»  
*The New York Times*

### Автор находки

Вы добавляете находки из своего аккаунта, пожалуйста загружайте только свои фотографии.



### Предмет наблюдения

Определение организма до группы или более точно до вида. С фотографией можно совсем не заполнять это поле и дождаться помощи сообщества.



### Место наблюдения

Находка должна иметь координату и оценку ее точности. Координату можно скрыть, если необходимо.



### Дата наблюдения

Записывайте дату находки, а не дату ее загрузки в яНатуралист.



### Доказательство наблюдения

Добавляя фотографию или запись звука к наблюдению, вы даете возможность сообществу с ее помощью помочь или улучшить определение организма. Фотографии лучшего качества и с разных ракурсов помогут сделать это более качественно.



# iNaturalist как технологический лидер: нейронная сеть для определения вида



**v.2.17**

**1145** таксонов добавлено  
9.11.2024

Year	Month	No. taxa
2022	Jan	50
2022	Feb	55
2022	Mar	60
2022	Apr	65
2022	May	70
2022	Jun	75
2022	Jul	80
2022	Aug	85
2022	Sep	90
2022	Oct	95
2022	Nov	100
2023	Jan	105
2023	Feb	110
2023	Mar	115
2023	Apr	120
2023	May	125
2023	Jun	130
2023	Jul	135
2023	Aug	140
2023	Sep	145
2023	Oct	150
2023	Nov	155
2024	Jan	160
2024	Feb	165
2024	Mar	170
2024	Apr	175
2024	May	180
2024	Jun	185
2024	Jul	190
2024	Aug	195
2024	Sep	200
2024	Oct	205
2024	Nov	210

- ★ ИИ определяет **95 903** таксонов
- ★ ИИ натренирован на **>27 млн** фото
- ★ Надо загрузить **100** фото вида для ИИ

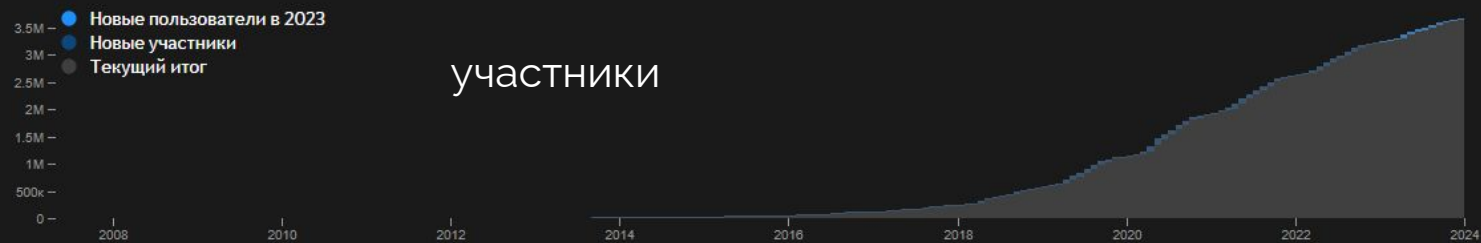
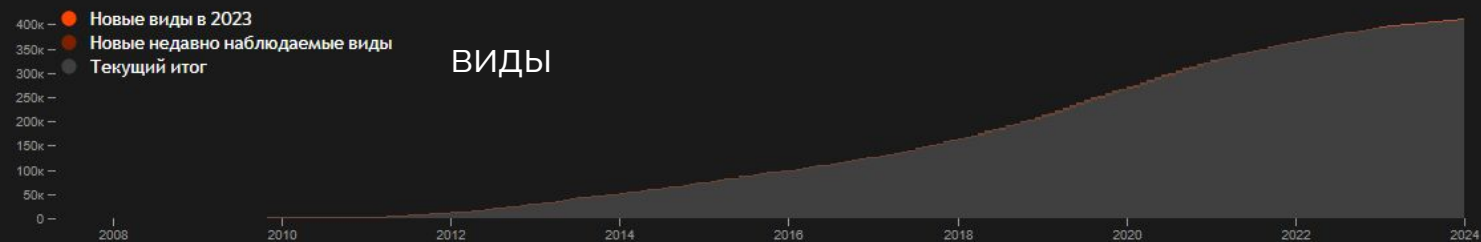
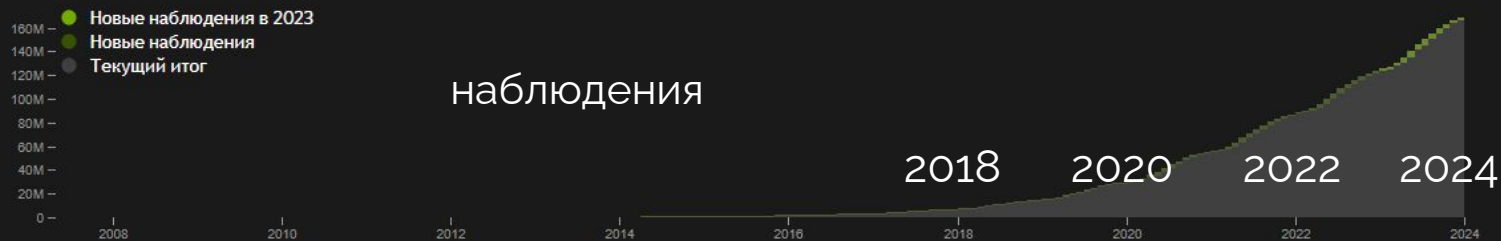
**Computer Vision Model**

**Included**

The current Computer Vision Model knows about this taxon, so it might be included in automated suggestions with the "Visually Similar" label.



## Рост

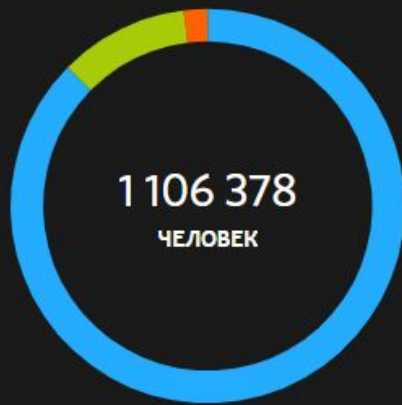
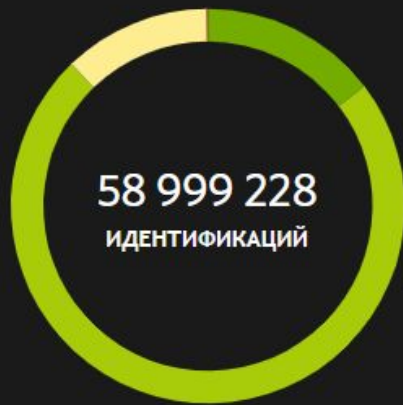
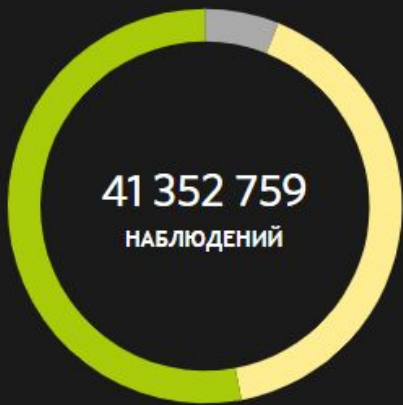


2023

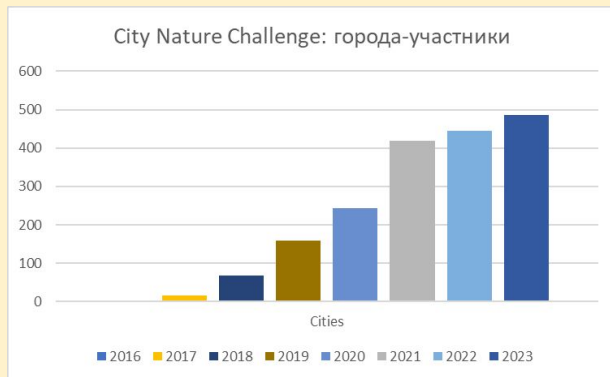


INATURALIST

2023



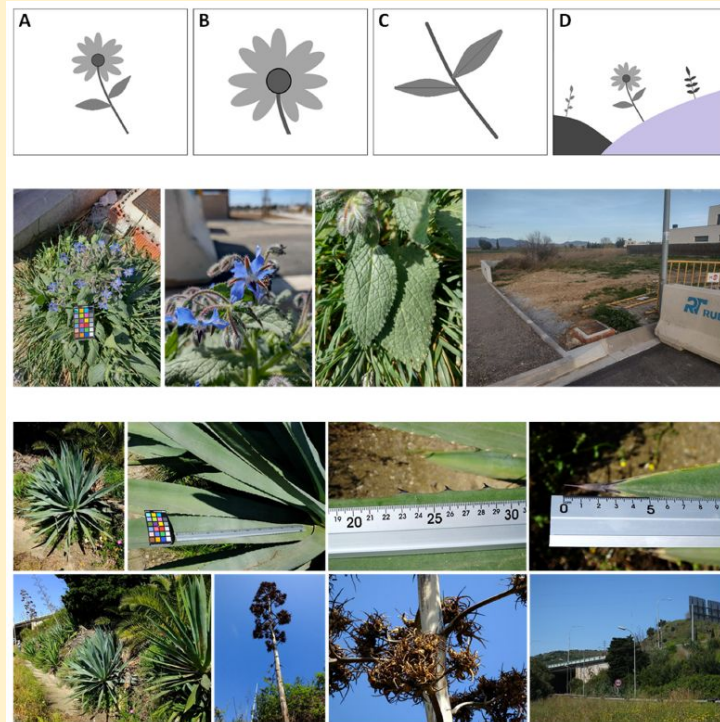
# City Nature Challenge и другие биоблици



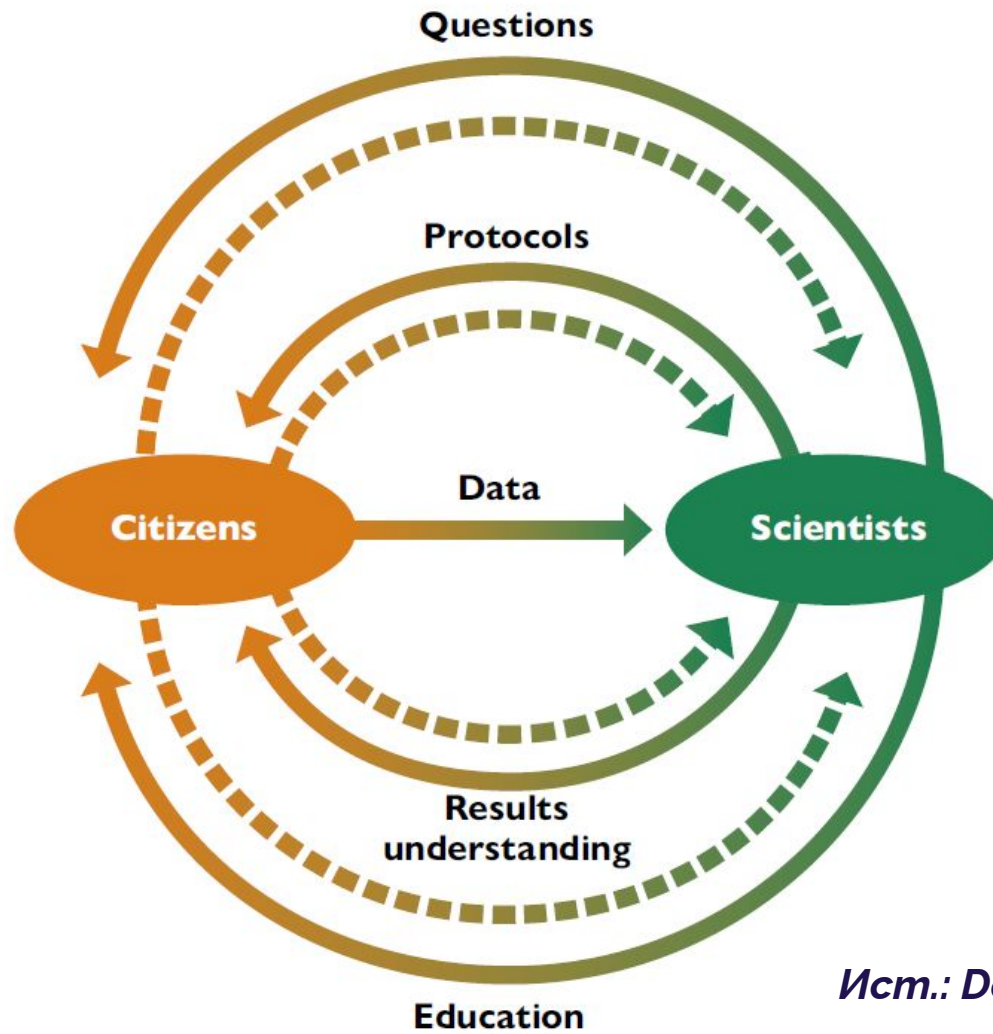
# Трудности использования данных iNaturalist



- ★ Фотографии низкого качества, связанные с наблюдениями
- ★ Ошибки таксономической идентификации
- ★ Отсутствие различий между наблюдениями за видами в неволе и в дикой природе
- ★ Отсутствие точности в географических координатах
- ★ Дублированные записи
- ★ Таксоны, которые не могут быть включены в таксономический справочник растений iNaturalist



[Дополнительные материалы \(видео\)](#)



*Исм.: Devictor et al. (2010)*



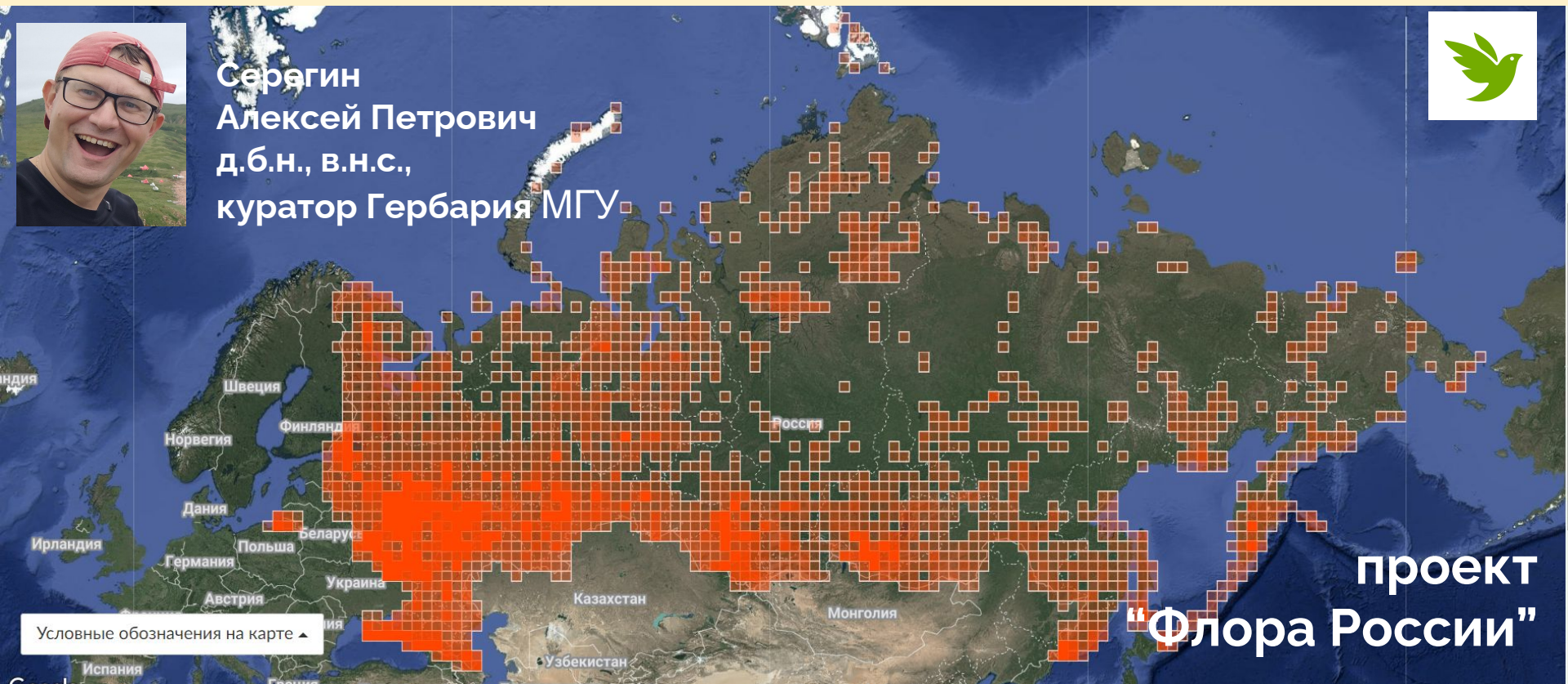
# Десять принципов гражданской науки

1. **Гражданская наука активно привлекает людей к участию в научных исследованиях**, в которых они играют значимую роль.
2. Проекты гражданской науки нацелены на **получение настоящих научных результатов**.
3. Участие в гражданской науке **приносит пользу** как профессиональным, так и гражданским ученым.
4. **Гражданские исследователи могут участвовать на всех этапах научного процесса**, если они этого хотят.
5. **Проекты гражданской науки предоставляют участникам обратную связь о том, как используются их данные**.
6. **Гражданская наука имеет более широкие возможности для вовлечения общества**.
7. **Данные проектов гражданской науки становятся доступными общественности, а результаты публикуются в открытом доступе**, если это возможно.
8. **Вклад гражданских исследователей (научных волонтеров) признается в результатах и публикациях проектов**.
9. **Программы гражданской науки оцениваются по их научным достижениям, качеству данных и влиянию на общество**.
10. **Руководители проектов учитывают правовые и этические аспекты**, включая вопросы авторства, конфиденциальности и охраны окружающей среды.





Серягин  
Алексей Петрович  
д.б.н., в.н.с.,  
куратор Гербария МГУ



Весь мир

RG

4 049 436  
НАБЛЮДЕНИЙ

8 431  
ВИД



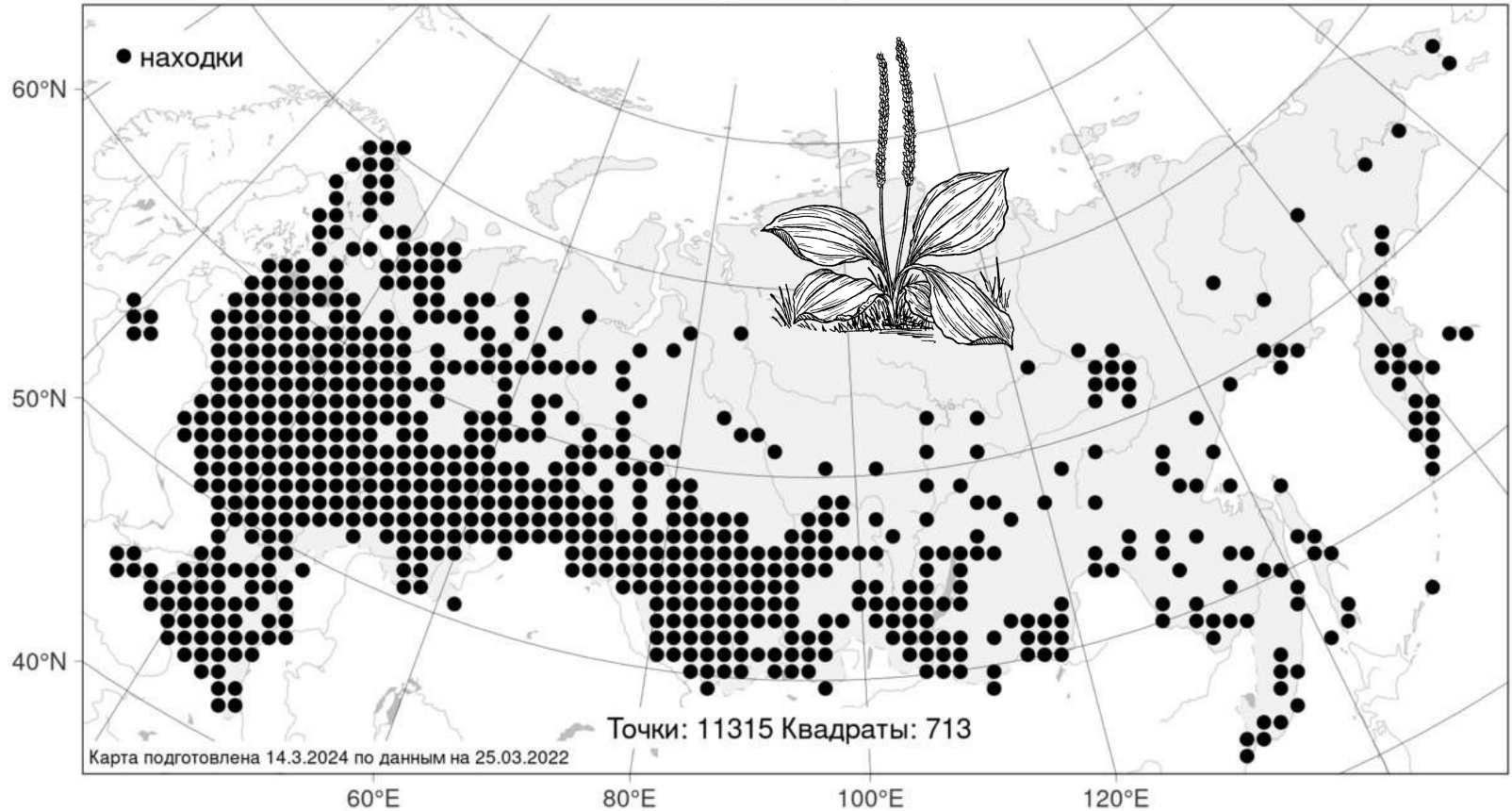
11 429  
ЭКСПЕРТОВ



28 740  
НАБЛЮДАТЕЛЕЙ



# Plantago major





# Помогатор Цифрового гербария МГУ

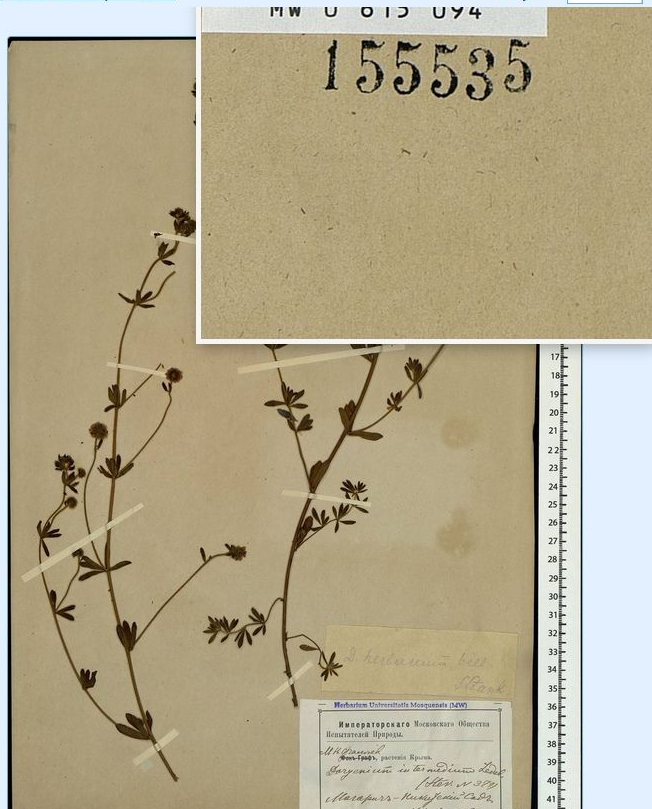
*Dorycnium intermedium* - Пятилистник средний [MW0615094]

[Пройти обучение повторно](#) [Открытая версия](#) [Полная карточка](#) [Сообщить об ошибке](#)

[Полное изображение](#)

Масштаб лупы **3X**

[← Вернуться к предыдущему](#)



Назаровский номер

Неочевидно

Пропустить

Ещё до войны наш куратор Михаил Иванович Назаров конторским нумератором проштамповал свыше 263 тыс. образцов, которые находились в тот момент в фондах Гербария МГУ. Особенно важно было подвести итоги бурного роста гербария в 1930-е гг. По итогам этой работы мы впервые точно узнали, сколько образцов находится в коллекции: по-видимому, образец [MW0585730](#) из Южной Африки стал последним, на котором Назаров поставил номер (263306). Это произошло, вероятно, в конце 1940 г. или в самом начале 1941 г.

С тех пор коллекция Гербария МГУ выросла вчетверо. В этой миссии мы проверяем все образцы, собранные до 1940 г. включительно, в поисках «назаровских номеров» – аккуратных чёрных оттисков в правом верхнем углу каждого гербарного листа. В этой миссии три варианта ответа: окошко для ввода номера, «Неочевидно» и «Пропустить».

Сложности связаны с тем, что после переезда Гербария МГУ в 1953 г. в новое здание на Ленинских горах некоторые образцы не помещались по ширине в новые шкафы. Некоторые листы с краёв были подрезаны, в том числе затронув последние цифры «назаровских номеров». Кроме того, часть номеров может быть случайно закрыта растениями, этикетками, а иногда и штрихкодами.

Формат ввода (с заглавным латинским «X» вместо закрытых / обрезанных цифр):

**89098** (целый номер или потеря цифры неочевидна);

**X78906** (первая цифра закрыта);

**17890X** (последняя цифра на краю листа обрезана).

Как установить, что какие-то цифры обрезаны? Образцы, которые попадают в миссию, идут по порядку родов системы Энглера и алфавиту видов. Этот порядок использовался в коллекции и при Назарове, что помогает восстанавливать частично утраченные номера. Так, если у вас идут образцы с оттисками «22340», «22338», «22346», то номер на краю листа «2234», скорее всего, просто обрезан и его нужно вводить как «2234X».

Если номер отсутствует, то в этом случае в окошке ставится прочерк (-).

## Рейтинги

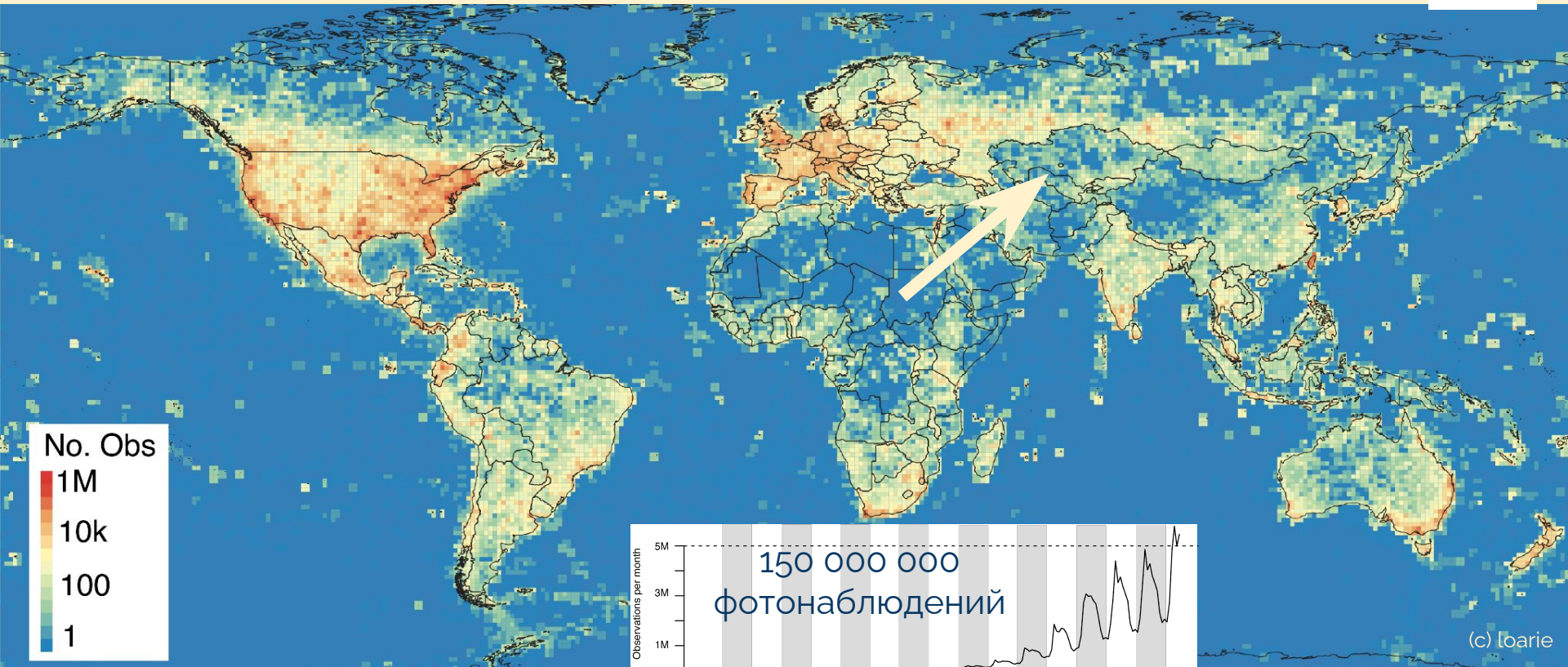
### Помогатор

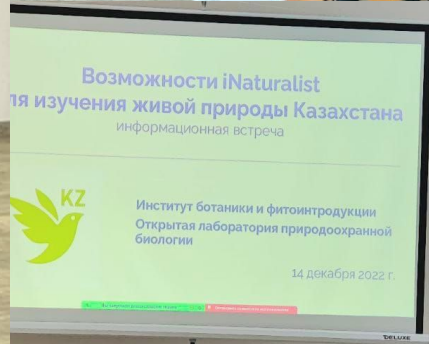
Ответо: 51190  
Правильно: 28986  
Не подтверждено: 21124  
Неверно: 678  
Пропущено: 406

### Назаровский номер

Всего образцов: 1307  
Доступно: 663  
Ответо: 3043  
Правильно: 2562  
Не подтверждено: 420  
Неверно: 60  
Пропущено: 1

# Глобальная статистика iNaturalist: точки





**Семинар о возможностях iNaturalist  
в декабре 2022 года**



## ПОЛОЖЕНИЕ регионального исследовательского конкурса «Биоразнообразие Семиречья: уязвимые виды, белые пятна»

Участникам конкурса предлагается найти и сфотографировать как можно большее количество растений, животных и грибов, особое внимание обратив на редкие и находящиеся под угрозой исчезновения виды, занесенные в Красную книгу Республики Казахстан, принять участие в инвентаризации флоры и фауны особо охраняемых природных территорий Семиречья (в границах Алматинской и Жетысуских областей) и малонарушенных природных территорий Семиречья, входящих в состав "горячей точки" биоразнообразия "Горы Центральной Азии", в том числе в "белых пятнах" — местах, которые мало посещаются научными работниками и натуралистами.

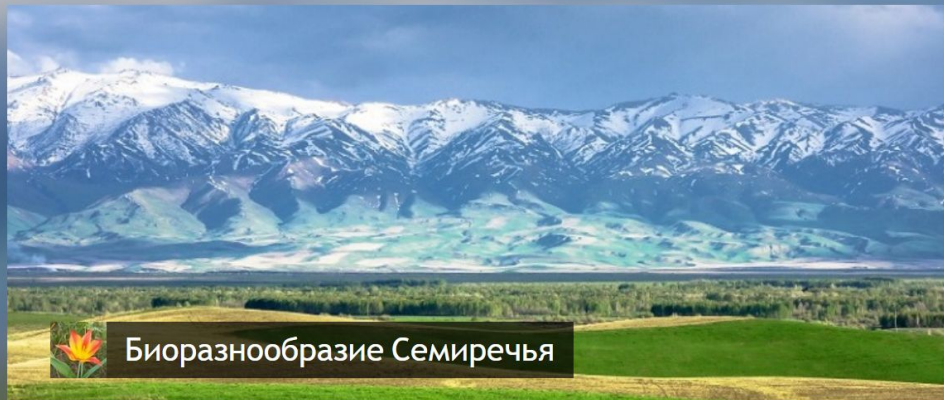
Информация, собранная участниками конкурса, размещается на платформе iNaturalist.org. После определения видового названия и получения исследовательского статуса (верификации) фотонаблюдения автоматически переносятся в мировой агрегатор данных о биоразнообразии — GBIF (Global

# Четыре приоритета 2023 года

- виды, занесенные в Красную книгу Республики Казахстан и как угрожаемые в различной степени в Красный список МСОП
- малоисследованные районы Семиречья (“белые пятна”)
- биологическое разнообразие ООПТ Алматинской и Жетысуской областей,
- биологическое разнообразие малонарушенных природных территорий “горячей точки” биоразнообразия “Горы Центральной Азии”







## Биоразнообразие Семиречья

### Подробнее

Участники 47

В рамках проекта аккумулируются фотонаблюдения растений, животных и грибов Семиречья (в границах Алматинской и Жетысуской областей Республики Казахстан).

[Читать далее >](#)

Ваше членство

Редактировать

Журнал проекта

**47**  
участников  
конкурса

Обзор

40 757  
НАБЛЮДЕНИЙ

3 951  
ВИД

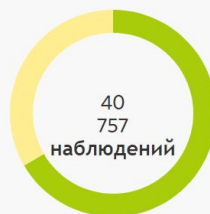
2 136  
ЭКСПЕРТОВ

693  
НАБЛЮДАТЕЛЕЙ

Статистика

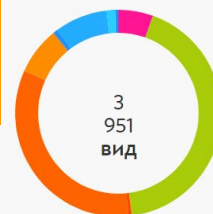
## Research grade. 2023

- **10 388** НАБЛЮДЕНИЙ
- **1 684** ВИДОВ
- **778** ЭКСПЕРТОВ
- **226** НАБЛЮДАТЕЛЕЙ

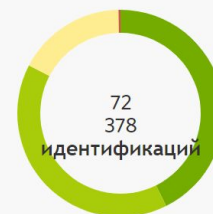


- Исследовательский уровень
- Требуется идентификация
- Обыкновенное

**66,58%**  
RG

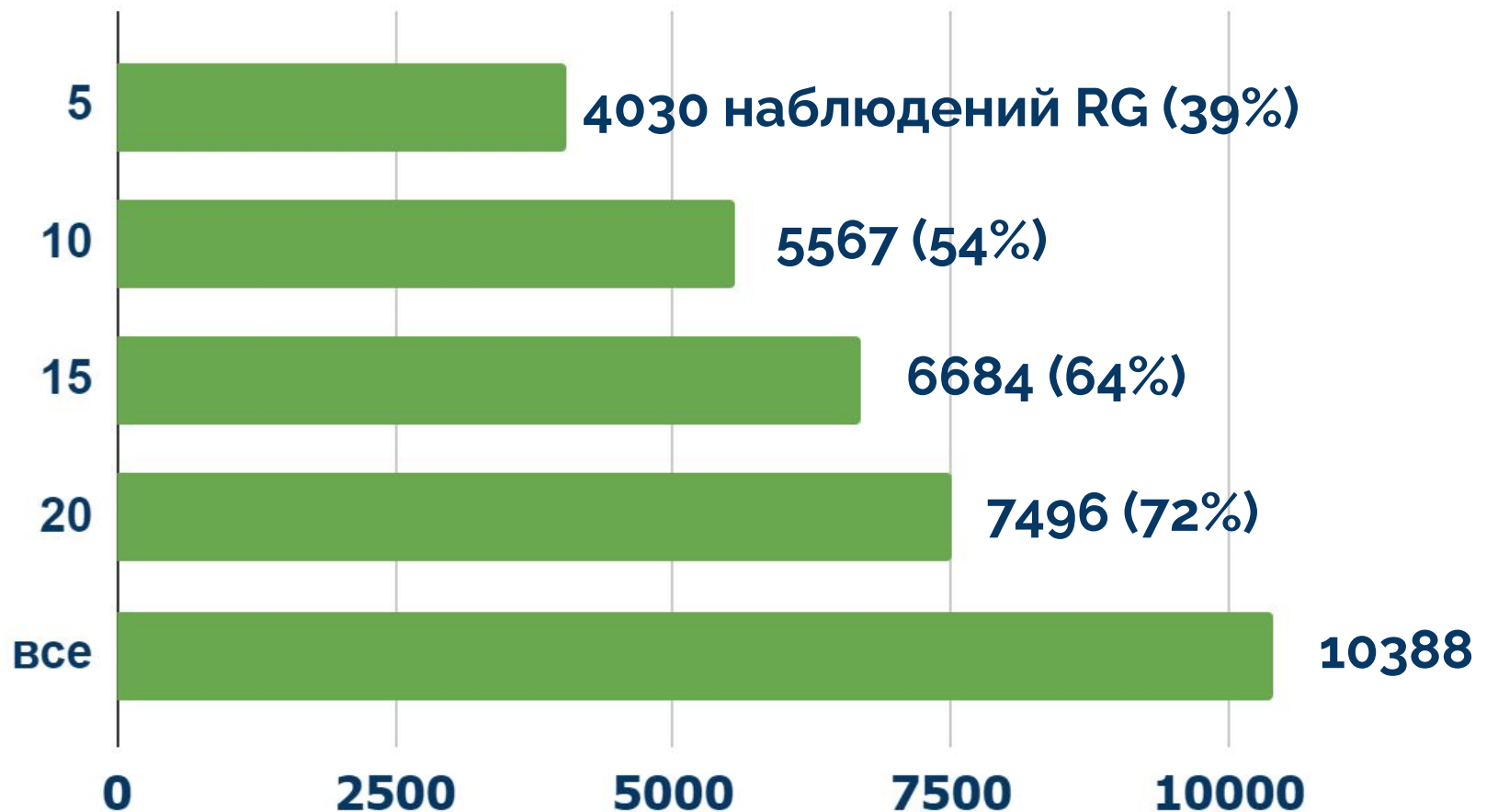


- Неизвестно
- Простейшие
- Грибы
- Растения
- Хромисты
- Моллюски
- Насекомые
- Паукообразные
- Лучепёрые Рыбы
- Земноводные
- Пресмыкающиеся
- Птицы
- Млекопитающие
- Другие Животные



- Улучшение
- Поддерживающее мнение
- Ведущий
- Независимое мнение

## Биоразнообразие Семиречья: вклад наблюдателей



% роста в 2023

Включая США

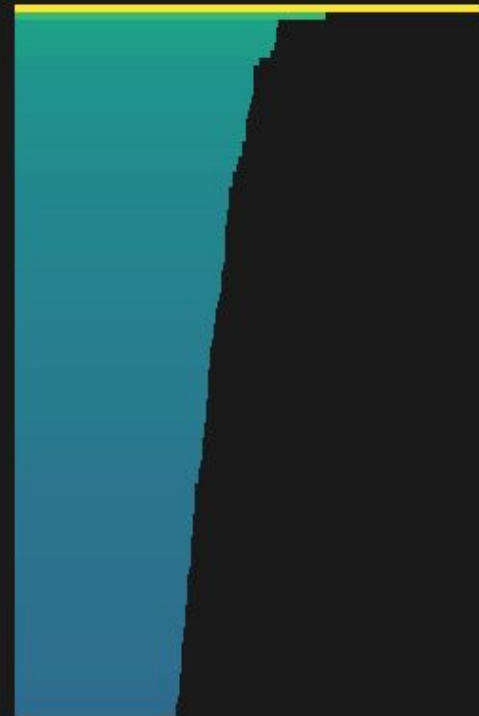
Шкала:  Линейная

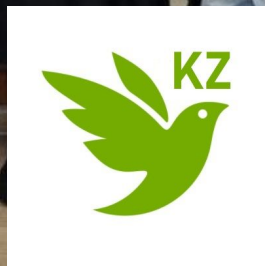
Логарифмическая

Казахстан (232,80%)

Казахстан  
на 20-м месте

2023: рост по странам (% к 2022)





# The 2024 CNC takes place in 2 parts



1

**April 26 – April  
29, 2024**

Taking pictures of wild plants  
and animals.

2

**April 30 – May 5,  
2024**

Identifying what was found.





ГЛАВНЫЙ  
БОТАНИЧЕСКИЙ  
САД



ОТКРЫТАЯ  
ЛАБОРАТОРИЯ  
ПРИРОДООХРАННОЙ  
БИОЛОГИИ



BEGONIA  
DESIGN STUDIO

# Знакомство с iNaturalist

- ✓ Советы новичкам
- ✓ City Nature Challenge 2024
- ✓ Включайся!

Алматы, 2024






# 1 Снять растение целиком

 **Ракурс:**  
сбоку,  
а не сверху

 **Свет:**  
тень, а не  
яркое солнце

 **Фон:**  
однотонный,  
без деталей,  
можно черный  
крафт



О науке без звериной серьезности

ЭВРИКА  
журнал

ГЛАВНАЯ

JUST WOW

SCI&ART

ЗАДАЧКИ

КНИГИ

НАУЧНОЕ КАФЕ

ОПЫТЫ

СООБЩЕСТВО

ЦИТАТА

## City Nature Challenge и iNaturalist: самое важное

📅 Апрель, 19, 2024 👤 Dubynin

В связи с грядущей всемирной акцией City Nature Challenge мы решили, что стоит ответить на наиболее волнующие вопросы ее потенциальных участников.

Итак, поехали.

### Что такое iNaturalist?

**iNaturalist — современная интернет-платформа для сбора и организации наблюдений дикой природы, которая объединяет энтузиастов её изучения.**

Привлекая любителей природы разного профиля, от туристов до экологов, сайт способствует расширению знаний о местном биоразнообразии и стимулирует дальнейшее исследование

 iNaturalist

Среднеазиатская черепаха  
(*Argyromys horsfieldi*)



Поиск



### НЕДАВНИЕ ЗАПИСИ

Мысли в День Земли

City Nature Challenge и iNaturalist: самое важное

Человек планеты Николай Полунин

Караканские тайнобрачники Николая Лащинского

Зеленый кот, история НГУ и бореализм: о чем будут говорить на Ночи научных историй в Академгородке

### АРХИВЫ





# 2024 Results



## City Nature Challenge



**2,436,844 Observations**



**65,682+ Species**

Including 3,940+ rare/endangered/threatened species



**83,528 Participants**

### Congratulations to these Cities!

- Most observations:** La Paz, Bolivia – 165,839  
 Monterrey, México – 81,727  
 San Antonio, TX, USA – 64,728
- 
- Most species:** La Paz, Bolivia – 5,352  
 Hong Kong, China – 4,775  
 Graz, Austria – 4,448
- 
- Most participants:** La Paz, Bolivia – 3,593  
 Monterrey, México – 2,576  
 San Francisco, CA, USA – 2,552

### Global Results Through the Years

	2016	2017	2018	2019	2020	2021	2022	2023	2024
<b>Cities</b>	2	16	68	159	244	419	445	482	690
<b>Countries</b>	1	1	17	28	40	44	47	46	51
<b>Observations</b>	19.8K	125K	441K	963K	815K	1.27M	1.7M	1.9M	2.4M
<b>Species</b>	2.5K	8.6K	18K	31K	32.6K	45.3K	50.1K	57.2K	65.7K
<b>Observers</b>	1K	4K	17K	32K	41K	52K	67.2K	66.4K	83.5K

K = thousand. M = million

# Interesting Finds from Around the World

## Naked Sea Butterfly

(*Clione limacina*)  
Greater Victoria, BC, Canada  
found by *Steph Brulot-Sawchyn*  
Not often observed on iNaturalist, sea butterflies are free-swimming pelagic sea snails!



## Dingy Purplewing

(*Eunica monima*)  
South Florida, USA  
found by *Liz T (@lt422)*  
A rare butterfly for Florida, & the only observation of this species for this year's CNC.



## Tibicina sp. Cicada

Pamplona / Iruña, Spain  
found by *Txalen Galina-Gaiton*  
Likely emerged from a nearby hole.



## Smooth Snake

(*Coronella austriaca*)  
Graz, Austria  
found by *@die\_fallenden*  
Affected by habitat degradation & destruction.



## Malus sieversii

Almaty, Kazakhstan  
found by *Kсения T.*  
Wild apple species native to Kazakhstan, in decline due to habitat loss.



## Lizard Goby

(*Rhinogobius flumineus*)  
Tokyo, Japan  
found by *@daichi2023*  
Endemic to Japan, found in fast-flowing mountain streams and small rivers.



## Corallorhiza macrantha

México City, México  
found by *Anayeli Guzmán*  
Coralroot orchid species found only in Mexico, it is parasitic and does not photosynthesize!



## Baird's Tapir

(*Tapirus bairdii*)  
Petén, Guatemala  
found by *@josebarrientos*  
Captured on a camera trap for CNC, this species is endangered due to loss of habitat.



## Blue Sea Star

(*Phataria unifasciata*)  
Galápagos, Ecuador  
found by *Billy Bensted-Smith*  
The color blue is rare in nature, and this echinoderm adds to the biodiversity of reefs they inhabit.



## Dynastes neptunus ssp. neptunus

Cali, Colombia, found by *Humberto Calero Mejia*  
This male beetle with striking ornaments is considered a vulnerable species and is rarely observed in the area.



## Lycosa erythrognatha

Adrianópolis, Brazil  
found by *Matheus Freitas*  
A cool wolf spider and the only one known to feed on a local species of toad.



## Heterosphenia indica



Kottat, Kerala, India  
found by *Athul Sankar C*  
First record of this moth on iNaturalist, previously only known from a single specimen collected in 1925 – a rediscovery after 99 years!

## Indri (Indri indri)



Madagascar  
found by *@kadae*  
This is one of the largest living lemurs, and is a critically endangered species.

## Anemone Stinkhorn (*Aseroe rubra*)

Greater Plettenberg Bay, South Africa, found by *@neehumbal*  
Very interesting, and very stinky, anemone-like mushroom!



## Taiwan Thrush

(*Turdus niveiceps*)  
Central Taiwan  
found by *Shao Qi*  
A rare bird species endemic to Taiwan, found in its central mountain forests. The only one seen in this year's CNC!



## Pygmy Greenhood

(*Pterostylis clivosa*)  
Geelong, VI, Australia  
found by *Peter Crowcroft*  
This small orchid with a green hood is locally endangered in Victoria, Australia.



## Katipō spider

(*Latrodectus katipo*)  
Ōtautahi, Aotearoa (Christchurch, New Zealand)  
found by *Dustin (@dustinlamont)*  
This endemic Māori-named spider is classified as seriously declining.



Malus sieversii  
Яблоня Сиверса

@ksu57



# BIG ALMATY



## CNC 2024: Big Almaty

АПР. 26, 2024 - АПР. 29, 2024

### Подробнее

Участники 41

Присоединяйтесь к участию в уникальной всемирной акции City Nature Challenge! Это наш шанс погрузиться в изучение удивительного мира растений и животных нашего города, сделать важный вклад в науку и помочь в сохранении биоразнообразия.

[Читать далее >](#)

Ваше членство

Редактировать

Журнал проекта

Обзор

4 869  
НАБЛЮДЕНИЙ

791  
ВИД

158  
ЭКСПЕРТОВ

57  
НАБЛЮДАТЕЛЕЙ

**Статистика**

Весь мир

# RG

2 918  
НАБЛЮДЕНИЙ

477  
ВИДОВ



135  
ЭКСПЕРТОВ




44  
НАБЛЮДАТЕЛЕЙ




59, 6% RG

# Красная книга: наблюдения CNS 2024

Almaty Agglomeration 

286  
НАБЛЮДЕНИЙ 

23  
ВИДОВ

17  
ЭКСПЕРТОВ 

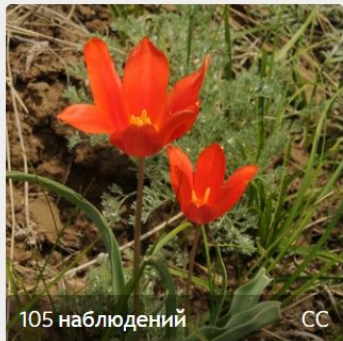
21  
НАБЛЮДАТЕЛЬ 



110 наблюдений

CC

*Iris alberti*  
(Ирис Альберта)



105 наблюдений

CC

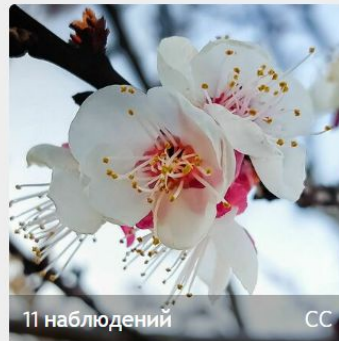
*Tulipa ostrowskiana*  
(Тюльпан Островского)



16 наблюдений

CC

*Malus sieversii*  
(Яблоня Сиверса)



11 наблюдений

CC

*Prunus armeniaca*  
(Абрикос Обыкновенный)



10 наблюдений

CC

*Gymnospermium altaicum*  
(Голосемянник Алтайский)



## Красная книга Казахстана

## Подробнее

Участники 28

Здесь накапливаются сведения о видах растений, животных, грибов и лишайников, занесенных в Красную книгу Республики Казахстан (список 2006 года).

[Читать далее >](#)

Ваше членство

Редактировать

Журнал проекта

Весь мир

5 733  
НАБЛЮДЕНИЯ

221  
ВИД

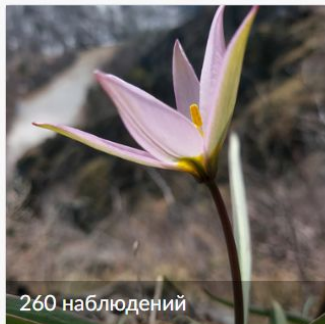
681  
ЭКСПЕРТ

482  
НАБЛЮДАТЕЛЯ



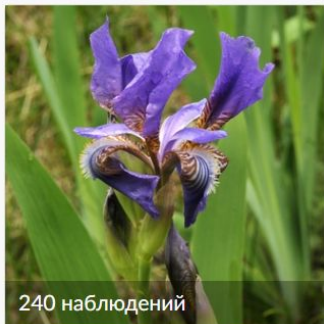
278 наблюдений

*Aquila nipalensis*  
Степной Орёл



260 наблюдений

*Tulipa patens*  
Тюльпан Поникающий



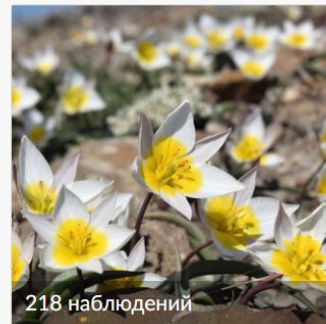
240 наблюдений

*Iris alberti*  
Ирис Альберта



229 наблюдений

*Tulipa regelii*  
Тюльпан Регеля



218 наблюдений

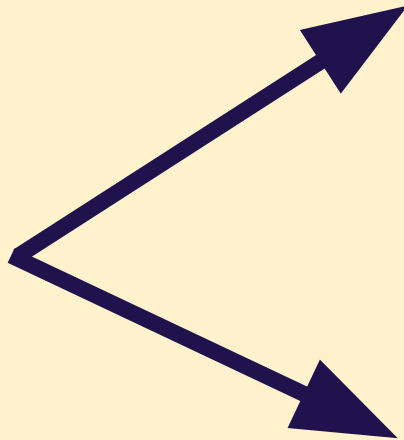
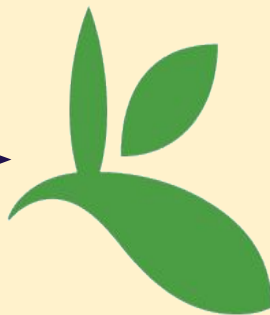
*Tulipa biflora*  
Тюльпан Двухцветковый

# Global Biodiversity Framework 2023-2030 (2050)

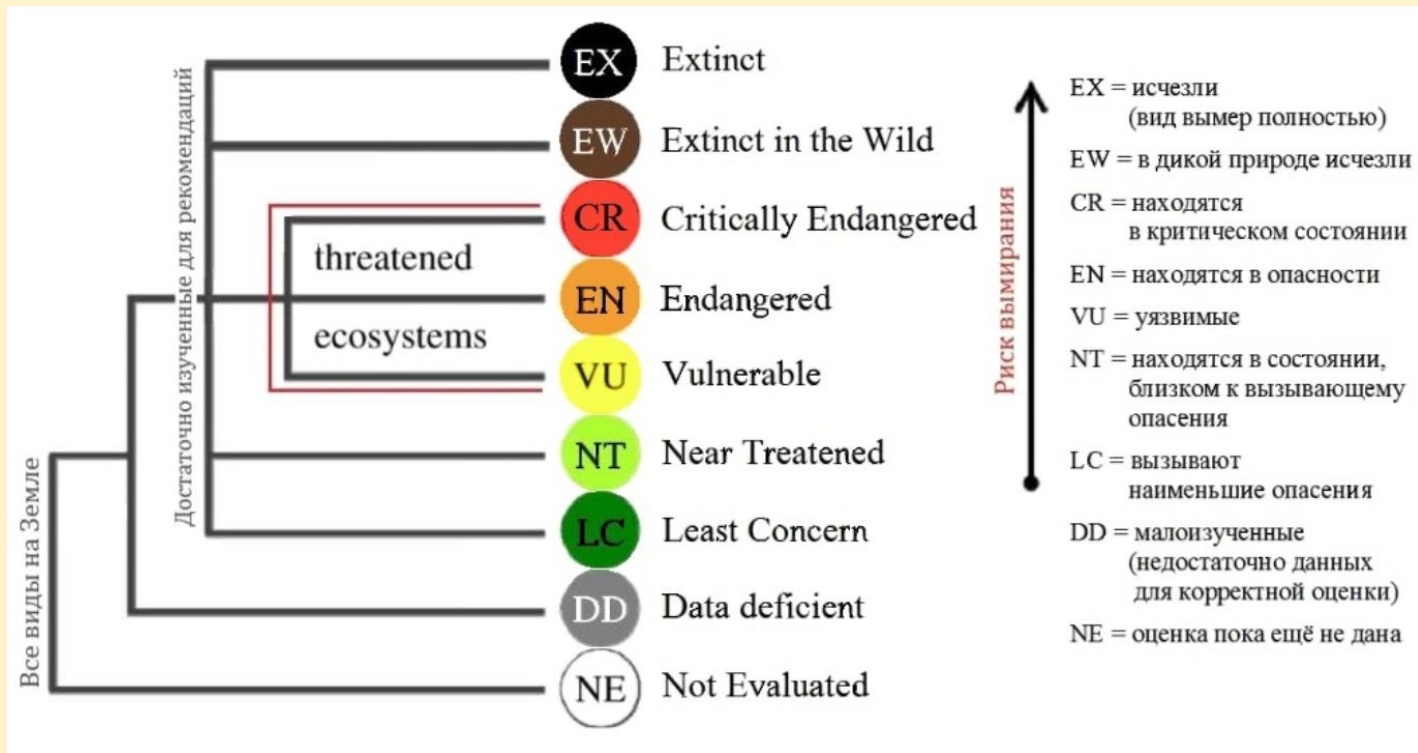
- ★ Декабрь 2022 года — **Куньмин-Монреальская глобальная рамочная программа по биоразнообразию** (Global Biodiversity Framework/GBF) Конвенции о биологическом разнообразии.
- ★ GBF содержит **четыре глобальные цели и 23 задачи**, в том числе **увеличение площади охраняемых природных территорий до 30% к 2030 году (задача «30 на 30»)**. Эти задачи невозможно решить без изменения отношения к проблеме утраты биоразнообразия и поведения и без деятельного участия значительного (гораздо большего, чем сейчас) числа жителей планеты.

- ★ **Проекты гражданской науки (citizen science/CS)** предоставляют для этого уникальные возможности для сбора новых сведений о видах и параметрах экосистем, мониторинга биоразнообразия, экологического образования и природоохранного просвещения.

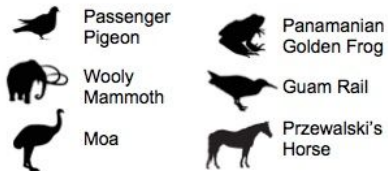




# Красный список МСОП (IUCN Red List)







**Extinct**



**Extinct in the Wild**

Out of all species that have ever lived, 99.9% of them are now extinct. The IUCN has documented 830 species known to have gone extinct since 500 AD.

Currently the IUCN has 69 species listed as extinct in the wild. These species have been extirpated from their natural range and now exist only in captivity.



**Threatened (encompasses CR, EN, and VU)**



**Critically Endangered**



**Endangered**



**Vulnerable**

*To be listed as threatened, a species must fit one of the following qualifications*

A projected population decline of greater than \_\_\_\_ over the next 10 years or three generations

**80%**      **50%**      **30%**

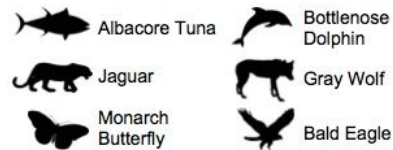
A global range of less than \_\_\_\_ km<sup>2</sup>

**100**      **5,000**      **20,000**

A stable global population size of less than \_\_\_\_ individuals

**50**      **250**      **1000**

\*Two other categories exist, although they have more complicated thresholds: small population size AND decreasing population, or quantitative scientific analysis that shows a high likelihood of extinction.



**Near Threatened**



**Least Concern**

Near threatened species are in danger of becoming threatened. Species with least concern listings may still have declining populations or face threats, but not to the level required for a threatened listing.

Out of the RedList's 77,000 listed species, 35,000 are listed as least concern and 5,000 are listed as near threatened.



For species that have yet to be sufficiently evaluated, two additional classifications exist: data deficient (DD) and not evaluated (NE).



# Оценка угрозы исчезновения вида по критериям МСОП





**RED LIST**  
**CERTIFICATE**  
OF COMPLETION OF TRAINING  
IN THE COURSE OF  
*Ирина Буряева*  
Specialist of Social Work

GOOD MORNING ANGEL

**RED LIST**  
**CERTIFICATE**  
OF COMPLETION OF TRAINING  
IN THE COURSE OF  
*Галина Зарина*  
Specialist of Social Work

**RED LIST**  
**CERTIFICATE**  
OF COMPLETION OF TRAINING  
IN THE COURSE OF  
*Александр Шибанов*  
Specialist of Social Work

**RED LIST**  
**CERTIFICATE**  
OF COMPLETION OF TRAINING  
IN THE COURSE OF  
*Султана Рахматова*  
Specialist of Social Work

**RED LIST**  
**CERTIFICATE**  
OF COMPLETION OF TRAINING  
IN THE COURSE OF  
*Ирина Громова*  
Specialist of Social Work

**RED LIST**  
**CERTIFICATE**  
OF COMPLETION OF TRAINING  
IN THE COURSE OF  
*Ирина Анабиева*  
Specialist of Social Work

# Расширенный образец / Extended Specimen

- **Физический образец.** Сам объект, собранный в природе и хранящийся в коллекции.
- **Морфологические данные.** Подробные описания внешнего вида, структуры и измерений.
- **Генетические данные.** ДНК-сиквенсы, геномы, метагеномные данные.
- **Экологические данные.** Информация о месте и условиях сбора, включая географические координаты, тип экосистемы, климатические условия и сопутствующие виды.
- **Изображения и мультимедийные материалы.** Фотографии высокого разрешения, видео- и аудиозаписи, микроскопические изображения.
- **Метаданные о сборе.** Дата и время сбора, имя коллектора, используемые методы сбора и обработки.
- **Связанные публикации и данные.** Научные статьи, отчеты, связанные исследования и дополнительные данные, полученные из образца.
- **Информация о хранении и доступе.** Условия хранения образца, его местонахождение в коллекции, доступность для дальнейших исследований.

# В качестве заключения

**Глобальные данные являются неотъемлемой частью усилий по сохранению биоразнообразия.** Они обеспечивают:

- **Информированное принятие решений.** Позволяют законодателям, сотрудникам министерств, научным работникам и специалистам по охране природы основывать свои действия на актуальной и точной информации.
- **Эффективное распределение ресурсов.** Помогают направлять ограниченные ресурсы туда, где они наиболее необходимы.
- **Сотрудничество.** Способствуют обмену знаниями и совместным действиям между странами и организациями, профессиональными исследователями и неправительственным сектором.

**Роль проектов гражданской науки в области биоразнообразия в этих процессах продолжит расти.**

**СПАСИБО ЗА ВНИМАНИЕ!**



## **Александр Дубынин**

н.с. лаб. геоботаники  
Института ботаники и  
фитоинтродукции,  
рук. международной  
экспертной группы  
"Открытая  
лаборатория  
природоохранной  
биологии"

[adubynin@yandex.ru](mailto:adubynin@yandex.ru)

+7 747 821 53 16



# Accessing and using GBIF- the essentials

Slides by: Chihjen Ko | Asia Regional Support Team  
Oleg Borodin | ECA Regional Support Team

Presented by: Oleg Borodin | ECA Regional Support Team



---

## Primary biodiversity data

### GBIF dataset classes

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

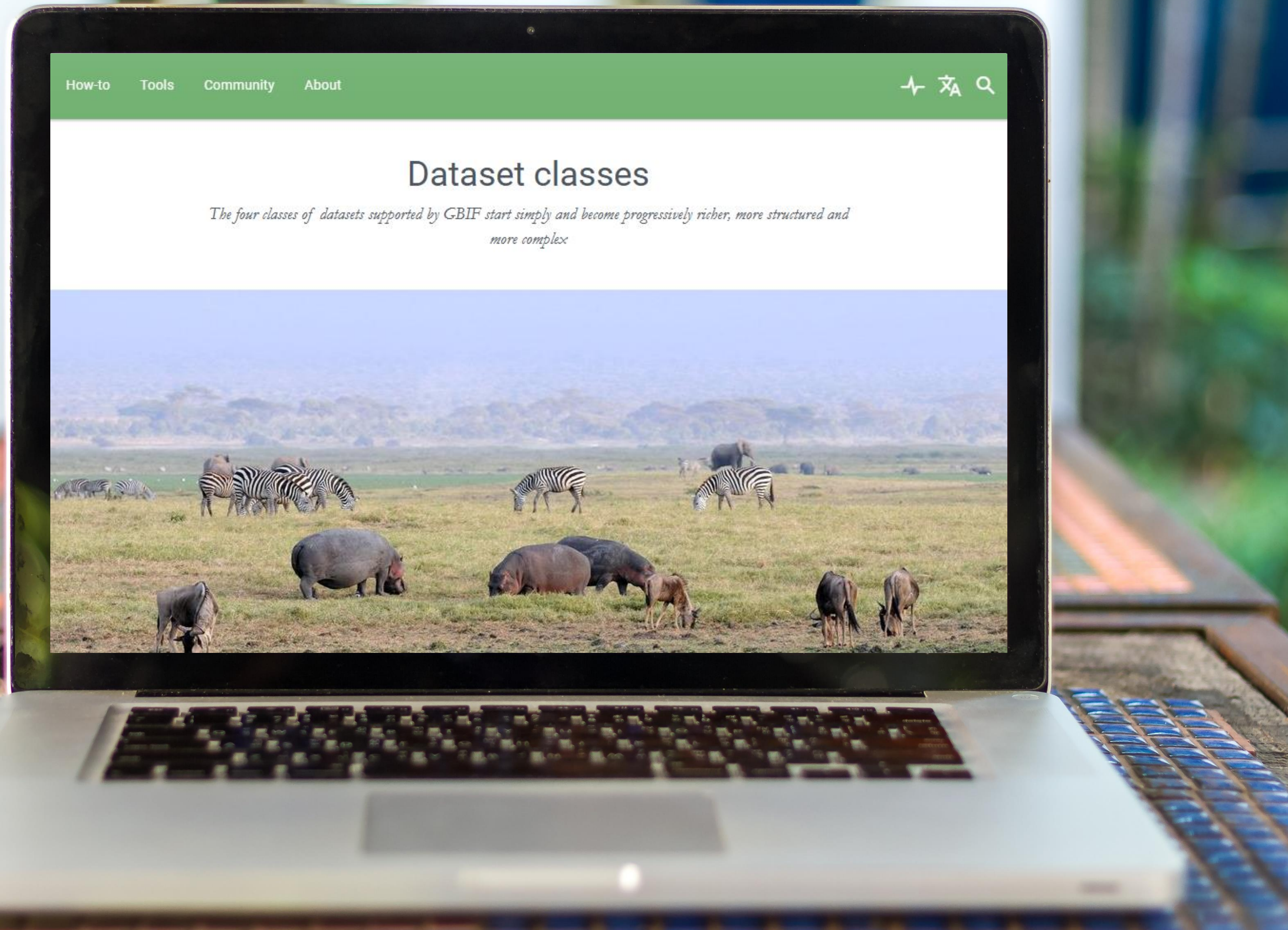
Meta data only

Checklist

Occurrence

Sampling-event

---







[posts](#) [community-forum](#) [gbif.org](#) [about](#)

## How to choose a dataset class on GBIF?

Marie Grosjean

2019-12-04 · GBIF · Publishing

If you are a (first time) publisher on GBIF and you are trying to decide which type of dataset would best fit your data, this blogpost is for you.

All the records shared on GBIF are organized into datasets. Each dataset is associated with some metadata describing its content (the classic "what, where, when, why, how"). The dataset's content

<https://data-blog.gbif.org/post/choose-dataset-type/>



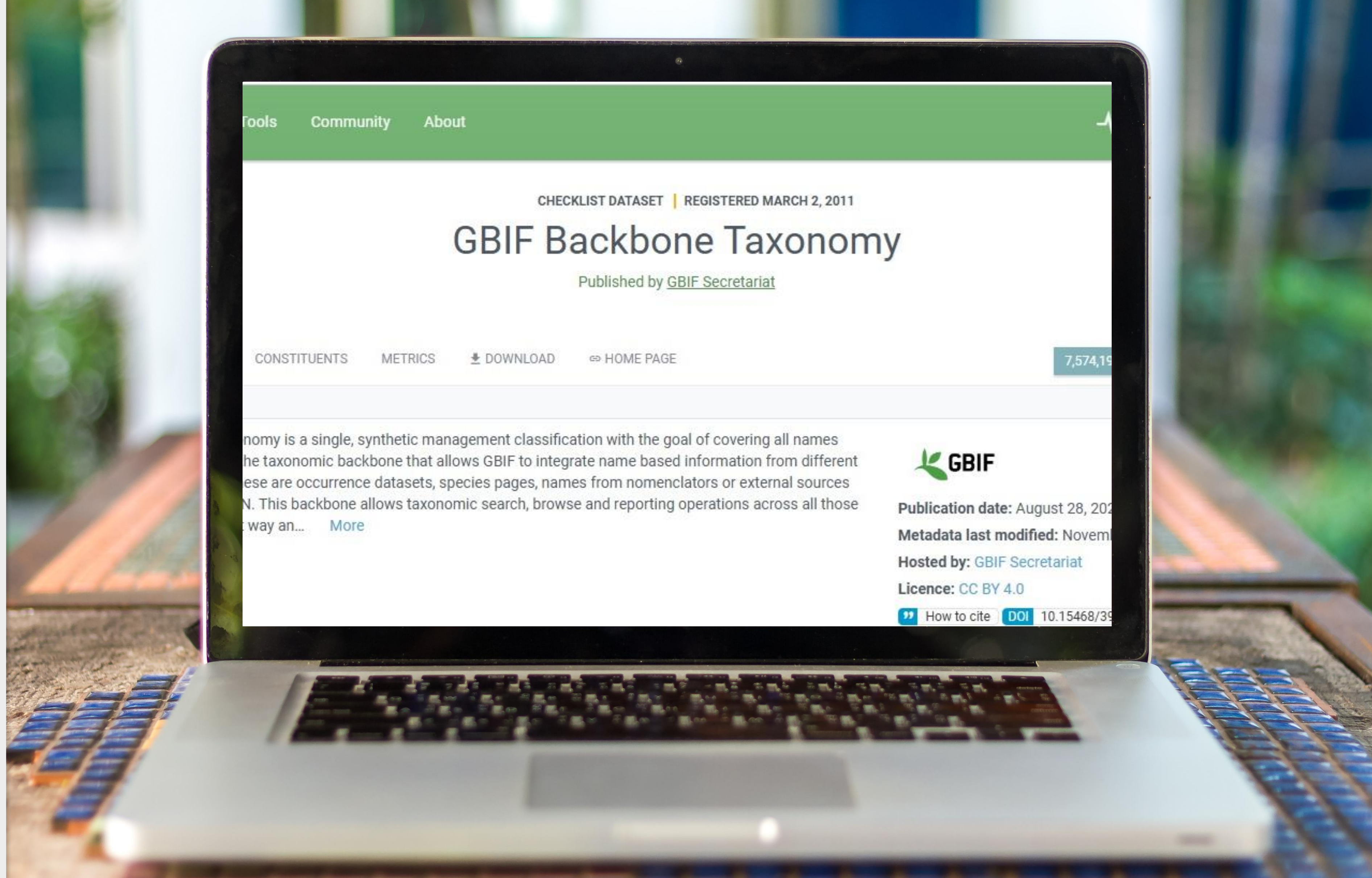
---

## Primary biodiversity data

### GBIF taxonomic backbone

The GBIF Backbone Taxonomy is a single, synthetic management classification with the goal of covering all names GBIF is dealing with. It's the taxonomic backbone that allows GBIF to integrate name based information from different resources, no matter if these are occurrence datasets, species pages, names from nomenclators or external sources like EOL, Genbank or IUCN.

---



[Description](#)[Data description](#)[GBIF registration](#)[Citation](#)

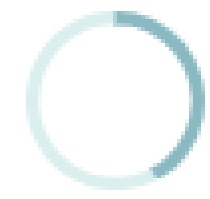
The following 105 sources have been used to assemble the GBIF backbone with number of names given in brackets:

- Catalogue of Life Checklist - 4766428 names
- International Barcode of Life project (iBOL) Barcode Index Numbers (BINs) - 635951 names
- UNITE - Unified system for the DNA based fungal species linked to the classification - 611208 names
- The Paleobiology Database - 212054 names
- World Register of Marine Species - 188857 names
- The Interim Register of Marine and Nonmarine Genera - 183894 names
- The World Checklist of Vascular Plants (WCVP) - 131891 names
- GBIF Backbone Taxonomy - 114350 names
- TAXREF - 109374 names
- The Leipzig catalogue of vascular plants - 75380 names
- ZooBank - 73549 names
- Integrated Taxonomic Information System (ITIS) - 68377 names
- [Plazi.org](#) taxonomic treatments database - 61346 names
- Genome Taxonomy Database r207 - 60545 names
- International Plant Names Index - 52329 names
- Fauna Europaea - 45077 names





4,114,717  
Accepted names

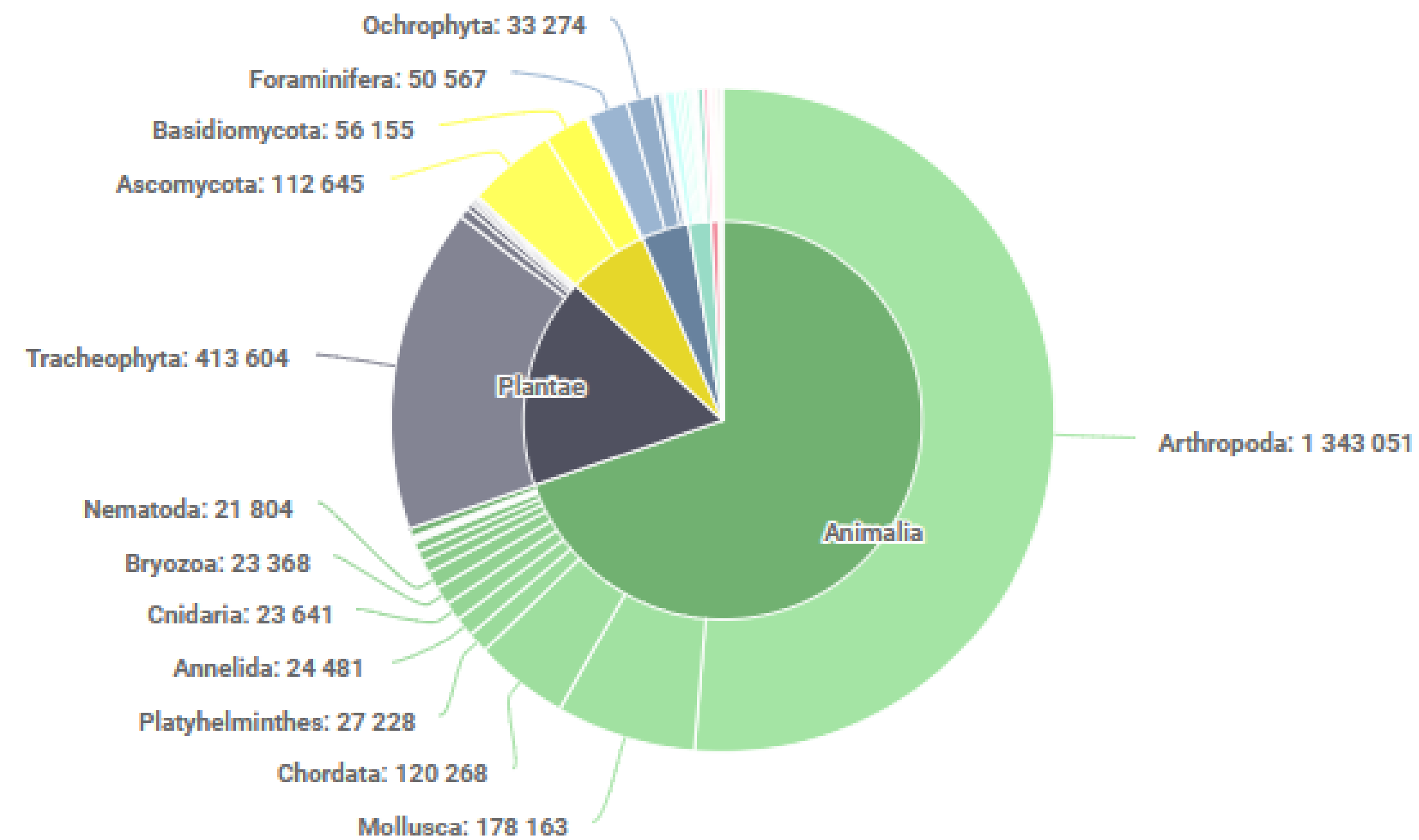


3,197,213  
Synonyms



59%  
Overlap with Catalogue of Life

### NUMBER OF ACCEPTED SPECIES BY HIGHER TAXON



---

## Principles of GBIF-mediated data

### Digital object identifiers

- datasets from the GBIF network
  - data downloads from GBIF.org
  - research articles and reports published by scientific journals, agencies and NGOs
  - materials deposited in a general-use repository
- 

## Citation guidelines

*These guidelines provide the most common examples of citation by GBIF users.*

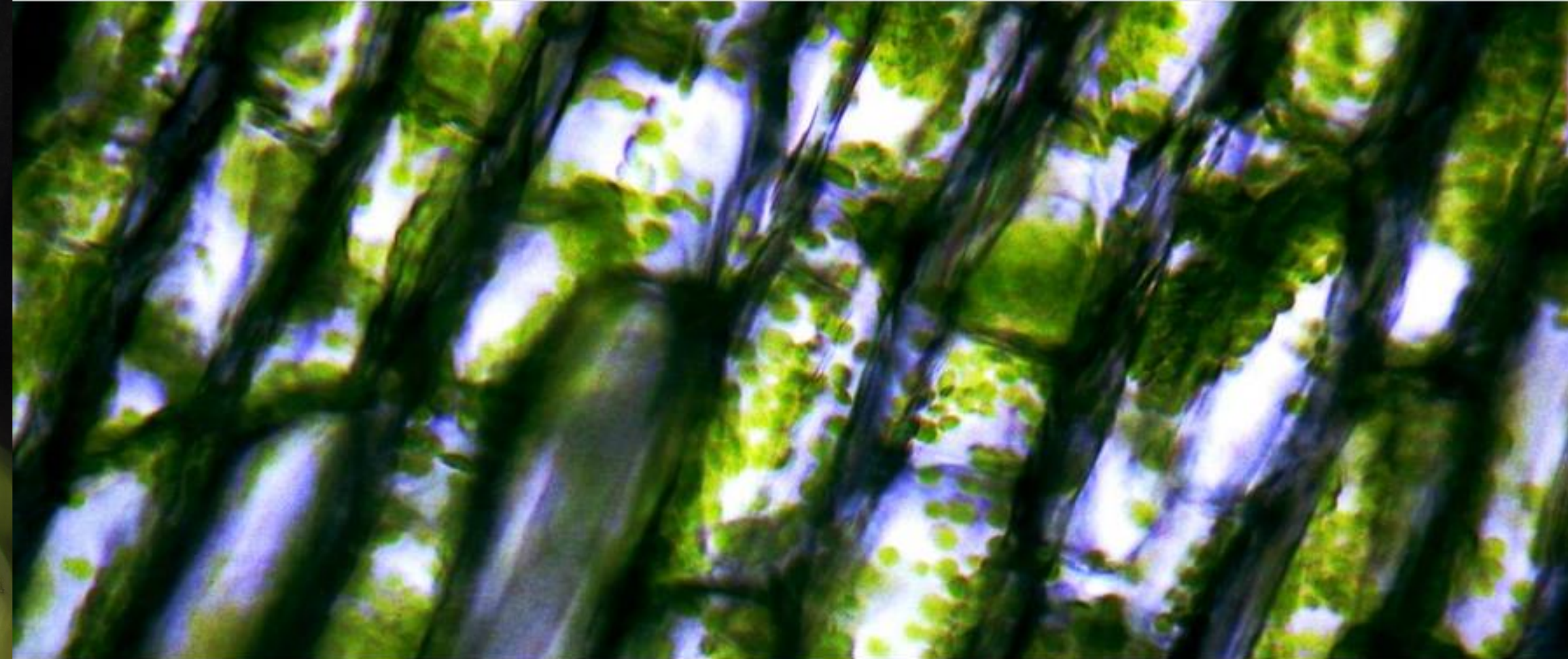


Sunfly (*Helophilus pendulus*) on Chicory (*Cichorium intybus*) by Donald Hobern. Photo licensed under CC BY 4.0.



## Data standards

*Shared data standards are the main enabler for bringing together the hundreds of millions of primary biodiversity records in the GBIF index.*



Macrophotograph of Elodea cells and chloroplast by Brandon Zierer. Licensed under CC BY-NC-SA 2.0.

---

## Principles of GBIF-mediated data

### Standards

- Darwin Core
  - Ecological Metadata Language (EML)
  - BioCase/ABCD: The Biological Collection Access Service
- 

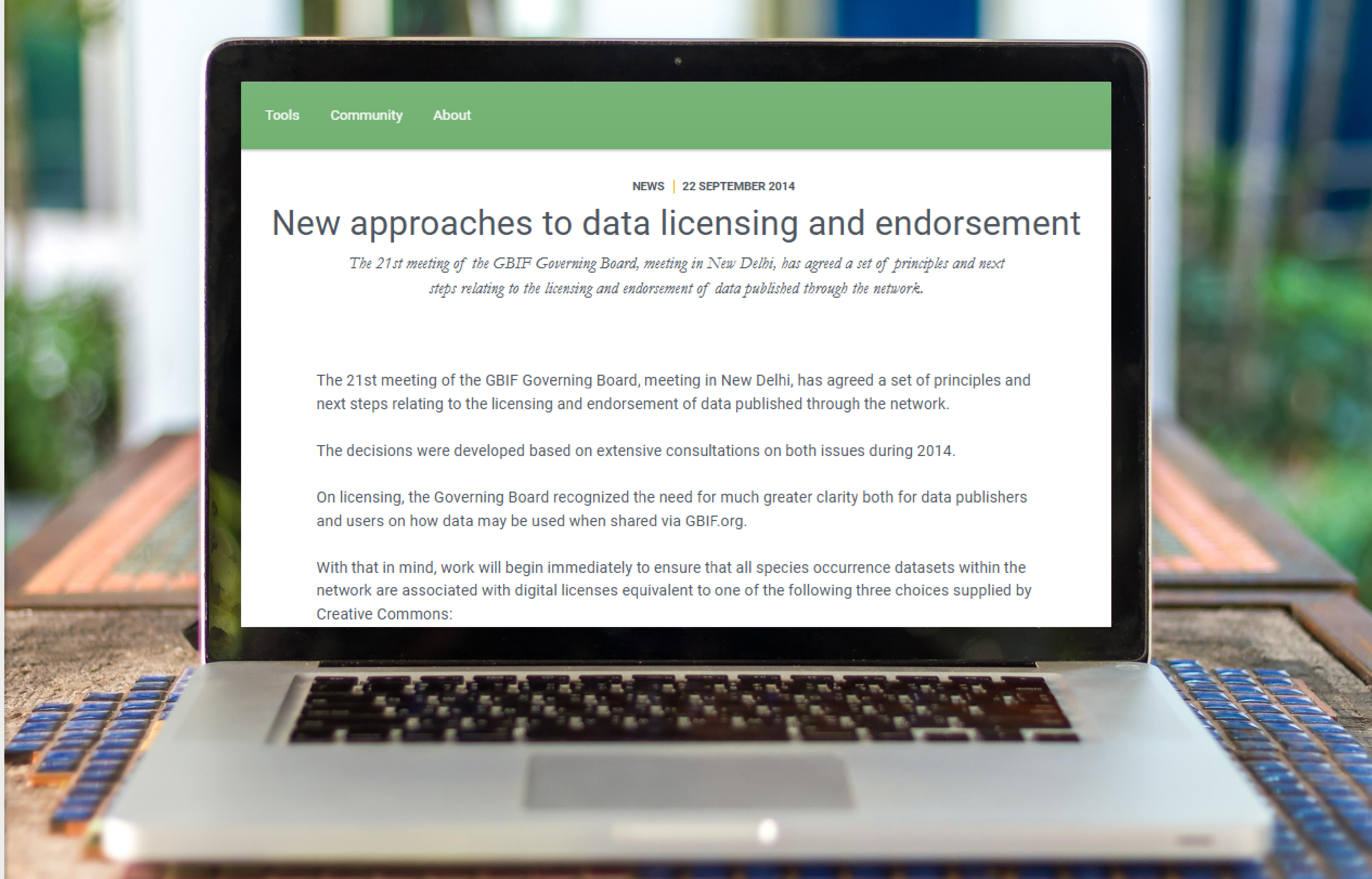


---

## Principles of GBIF-mediated data

### Open data

- [CCO](#) - no conditions for use
  - [CC-BY](#) - use with attribution
  - [CC-BY-NC](#) - non-commercial use with attribution
- 



---

## Principles of GBIF-mediated data

### FAIR data

- Findable
  - Accessible
  - Interoperable
  - Reusable
- 



### FAIR Principles

GO FAIR is committed to making data and services **findable, accessible, interoperable** and **reusable (FAIR)**.



**Findable:** Metadata and data should be easy to find for both humans and computers.



**Accessible:** The exact conditions under which the data is accessible should be provided in such a way that humans and machines can understand them.



**Interoperable:** The (meta)data should be based on standardized vocabularies, ontologies, thesauri etc. so that it integrates with existing applications or workflows.





Data found on GBIF.org are FAIR.

#### FINDABLE

GBIF has [requirements](#) for metadata and datasets. All datasets are identified by [Digital Object Identifiers](#) (DOIs).

#### ACCESSIBLE

The [GBIF Portal API](#) provides a machine readable interface (REST + JSON) and use the [Integrated Publishing Toolkit](#) (IPT) as trusted data repository.

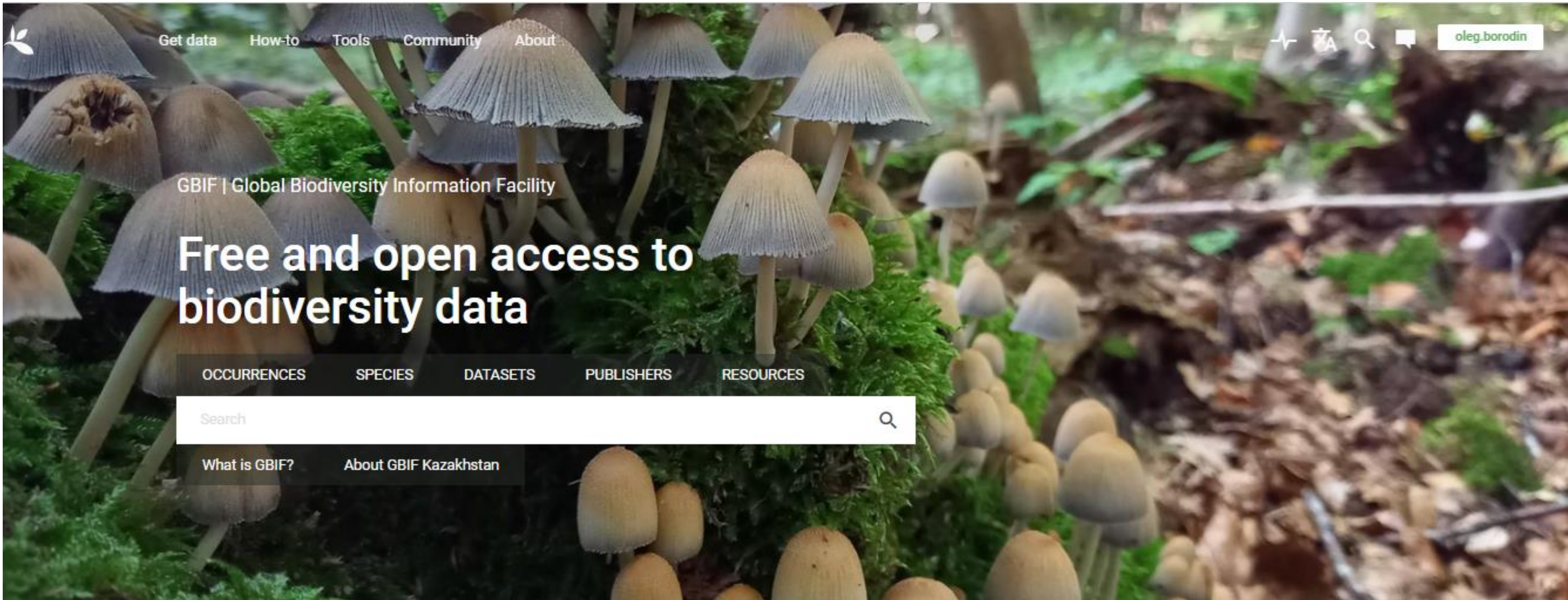
#### INTEROPERABLE

GBIF recommends using the [Ecological Metadata Language](#) (EML) for datasets and [Darwin Core](#) for occurrence data.

#### REUSABLE

GBIF require creative common data licenses ([CC0](#), [CC BY](#), or [CC BY-NC](#)). Provenance available from the GBIF portal.





[Get data](#)

[How-to](#)

[Tools](#)

[Community](#)

[About](#)



oleg.borodin

GBIF | Global Biodiversity Information Facility

# Free and open access to biodiversity data

[OCCURRENCES](#)

[SPECIES](#)

[DATASETS](#)

[PUBLISHERS](#)

[RESOURCES](#)

Search



[What is GBIF?](#)

[About GBIF Kazakhstan](#)



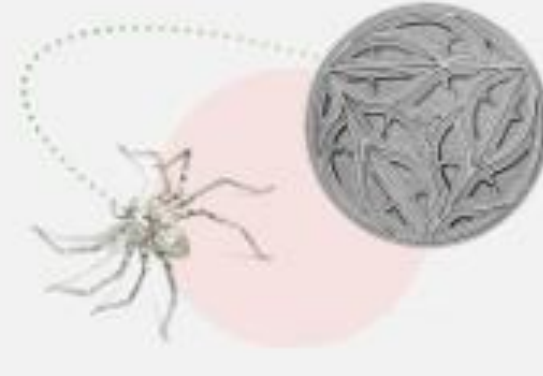
The image shows a screenshot of the GBIF website's navigation menu. The background features a forest floor with mushrooms. The top navigation bar includes links for 'Get data', 'How-to', 'Tools', 'Community', and 'About'. The 'Community' menu is open, displaying three columns of options: NETWORK, VOLUNTEERS, and ACTIVITIES. The 'Training and learning resources' option under the ACTIVITIES column is circled in red. Other visible text includes 'GBIF | Global Biodiversity Information', 'Free and open access to biodiversity data', 'OCCURRENCES', 'SPECIES', a search bar, and footer links for 'What is GBIF?' and 'About GBIF Kazakhstan'. A user profile 'oleg.borodin' is visible in the top right corner.

NETWORK	VOLUNTEERS	ACTIVITIES
Participant network	Mentors	Capacity enhancement
Nodes	Ambassadors	Programmes & projects
Publishers	Translators	<b>Training and learning resources</b>
Network contacts	Citizen scientists	Data Use Club
Community forum ↔		Living Atlases ↔
alliance for biodiversity knowledge ↔		





Introduction to GBIF



Biodiversity Data Mobilization



Introduction to using GBIF-mediated data



Accelerating biodiversity research through DNA barcodes, collection and observation data



Formal engagement and establishing Participant nodes





**Introduction to using GBIF-mediated data**

**Course details**

Course description

Acknowledgements

Citation

Navigating this course

Files for download

Software and accounts

Using GBIF-mediated data

Data processing and quality

Programmatic access to data

Course review

Course evaluation

Introduction to using GBIF-mediated data / Course details

# Introduction to using GBIF-mediated data



# Free and open access to biodiversity data

OCCURRENCES

SPECIES

DATASETS

PUBLISHERS

RESOURCES

Search



What is GBIF?

About GBIF Nepal



2,976,842,348

Occurrence records



106,482

Datasets



2,254

Publishing institutions



10,742

Peer-reviewed papers  
using data

## Data Access

Most common access points to GBIF-mediated data:

1. [www.gbif.org](http://www.gbif.org): a website driven by open-source data engineering that offers advanced searching, data analytics, visualisations, and downloads for registered users in 10 languages.
2. R ([www.r-project.org](http://www.r-project.org)): packages, including `rgbif` and `coordinatecleaner`, for data analysis, processing and visualisation.



# Free and open access to biodiversity data

OCCURRENCES

SPECIES

DATASETS

PUBLISHERS

RESOURCES

Search



What is GBIF?

About GBIF Nepal



2,976,842,348

Occurrence records



106,482

Datasets



2,254

Publishing institutions



10,742

Peer-reviewed papers  
using data

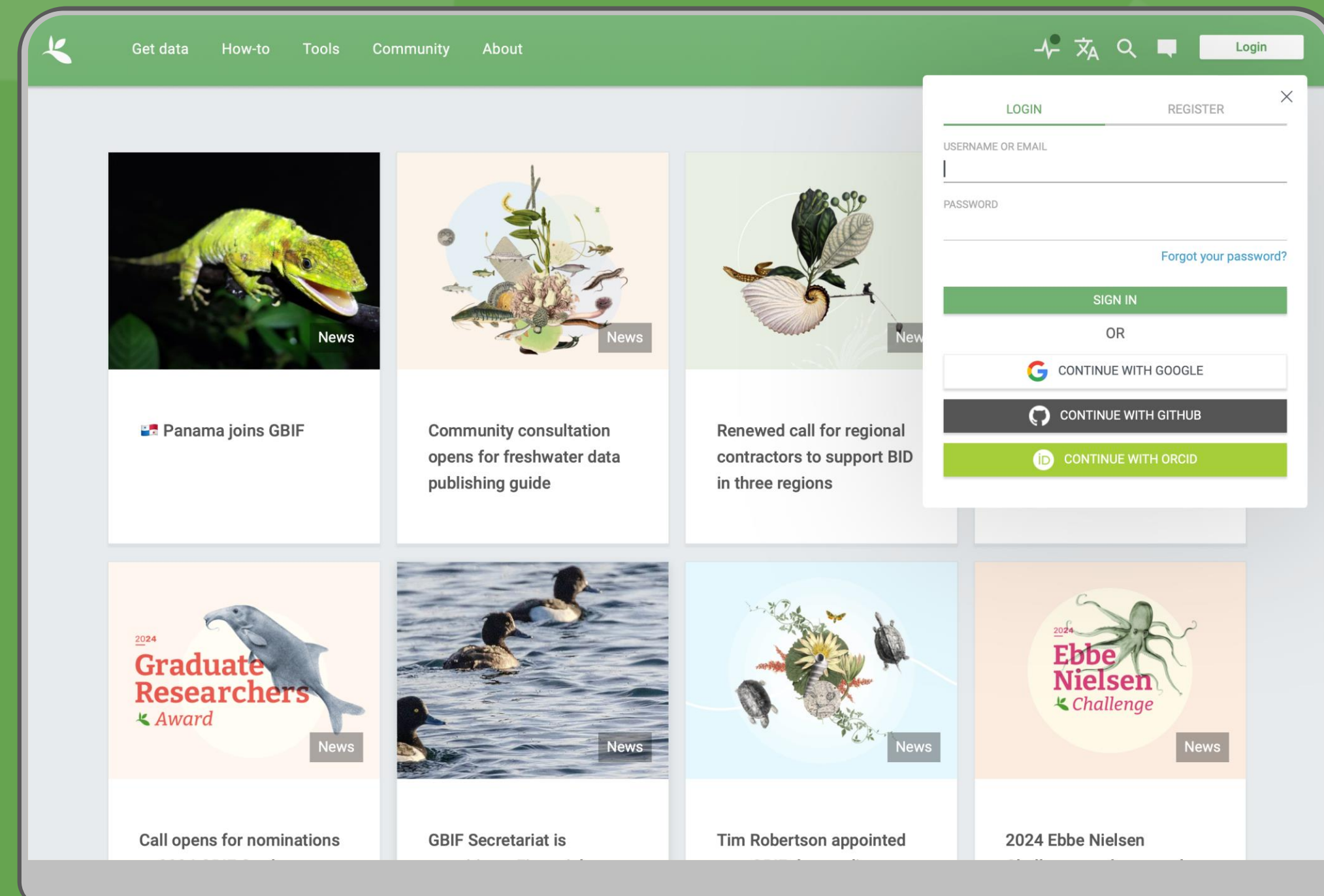
## GOLDEN RULES OF GBIF-MEDIATED DATA USE

1. Must have an account on [www.gbif.org](http://www.gbif.org)
2. Must agree to the Data User Agreement - <https://www.gbif.org/terms/data-user>
3. Document how you process your data
4. Correctly cite the data you use
5. Deposit used data in a public repository



# GOLDEN RULES OF GBIF-MEDIATED DATA USE

Must have an account  
on [GBIF.org](https://gbif.org)

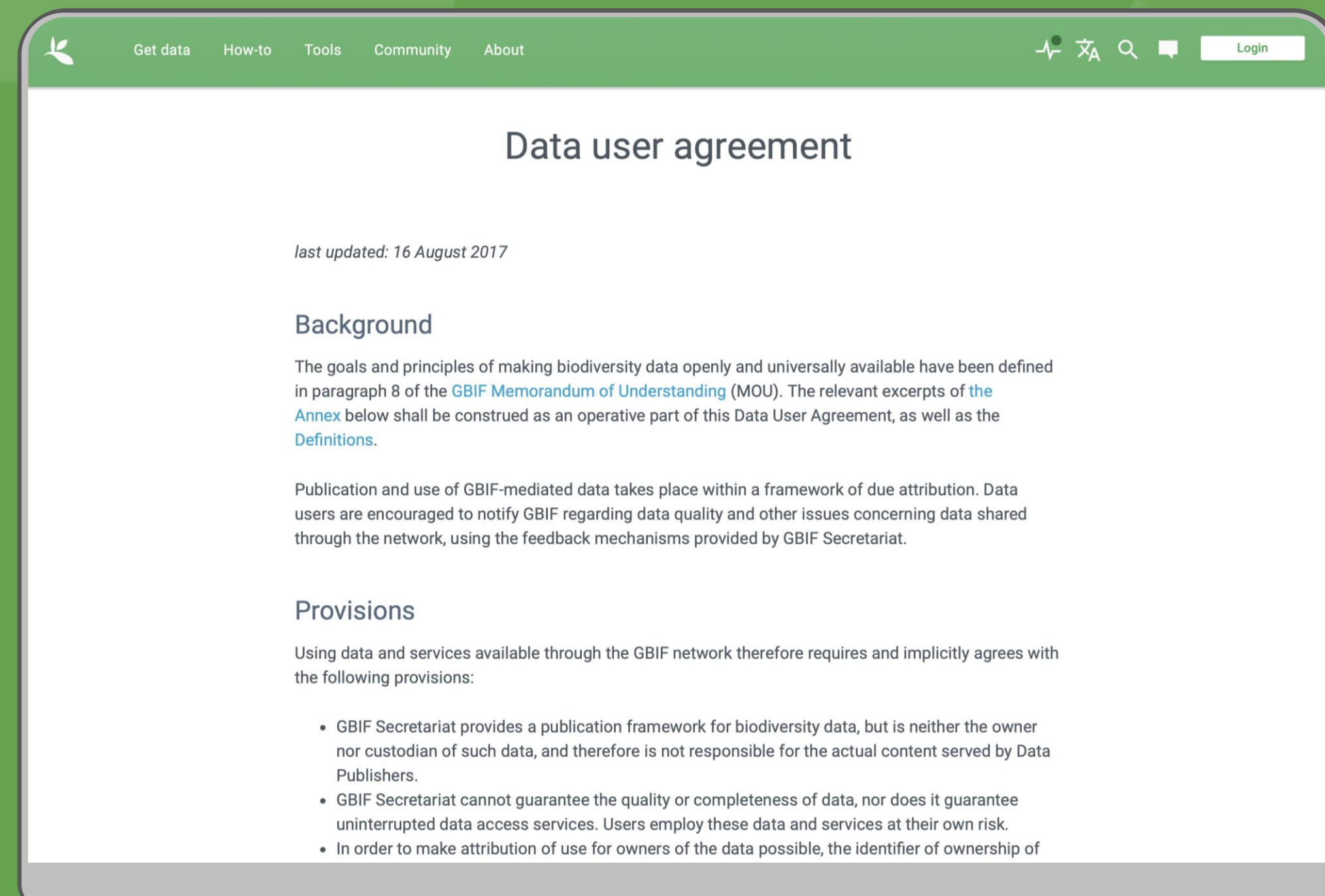




# GOLDEN RULES OF GBIF-MEDIATED DATA USE

## Must agree on GBIF the Data User Agreement

- Non-binding
- Sets out guiding principles of data use,  
including citation of data



GBIF | Global Biodiversity Information Facility

# Free and open access to biodiversity data

OCCURRENCES | SPECIES | DATASETS | PUBLISHERS | RESOURCES

What is GBIF? | About GBIF Kazakhstan

Glistening inkcap - *Coprinellus micaceus* (Bull.) Vilgalys, Hopple & Jacq-Johnson - observed by Ella Buck in Farum Nørreskov, Denmark. Photo via Danish Mycological Society (CC BY-NC 4.0)



3 024 408 823

Occurrence records



109 882

Datasets



2 315

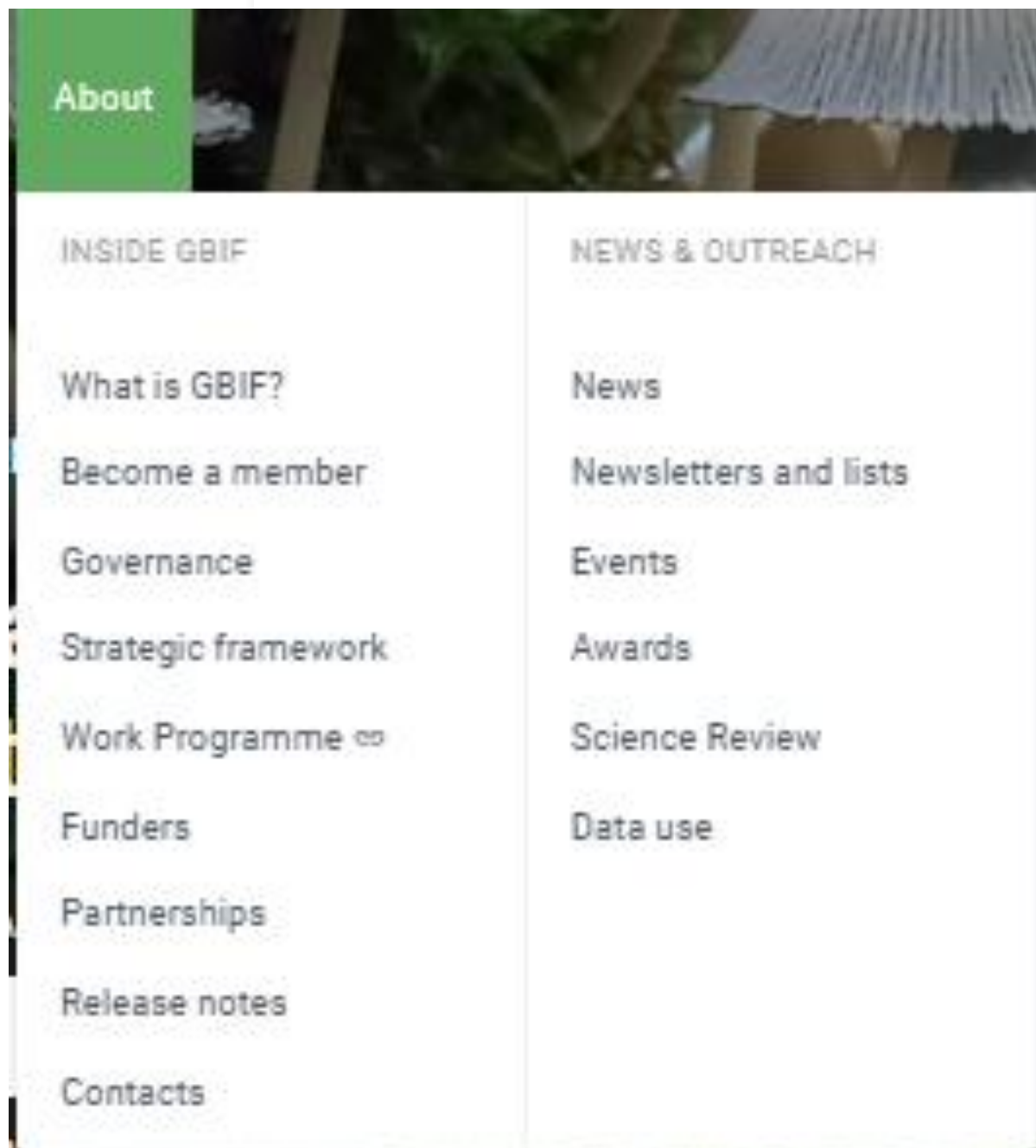
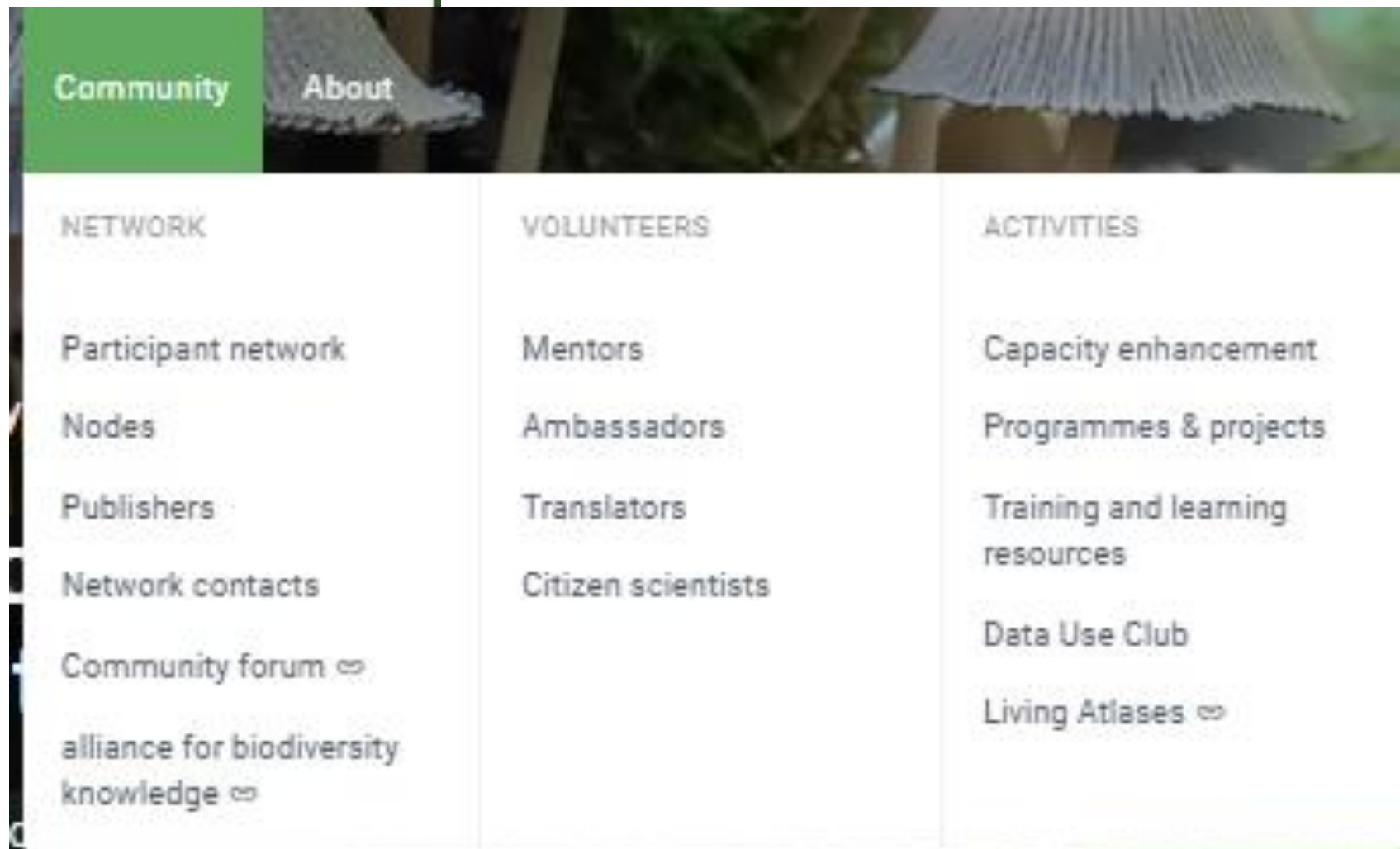
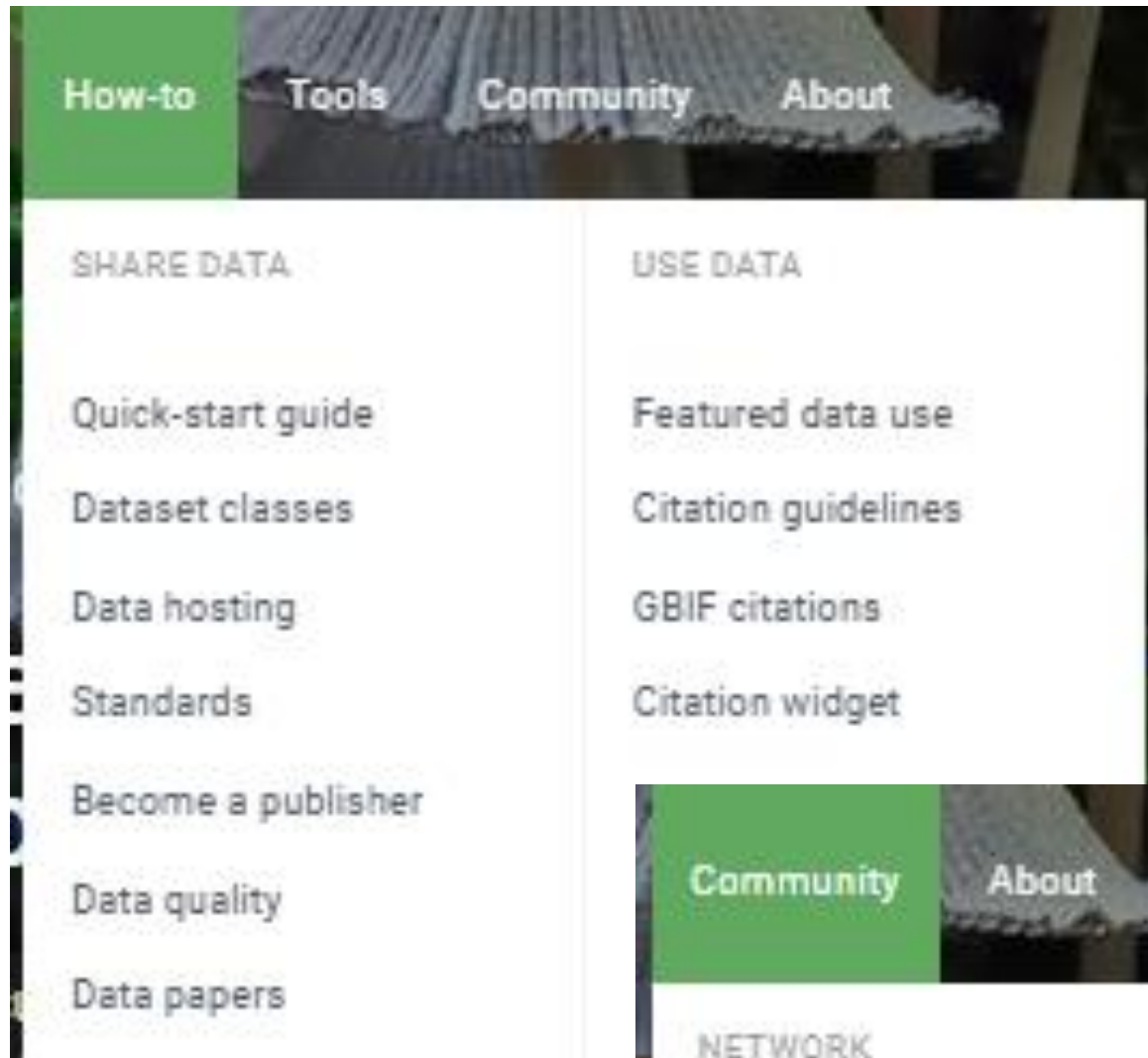
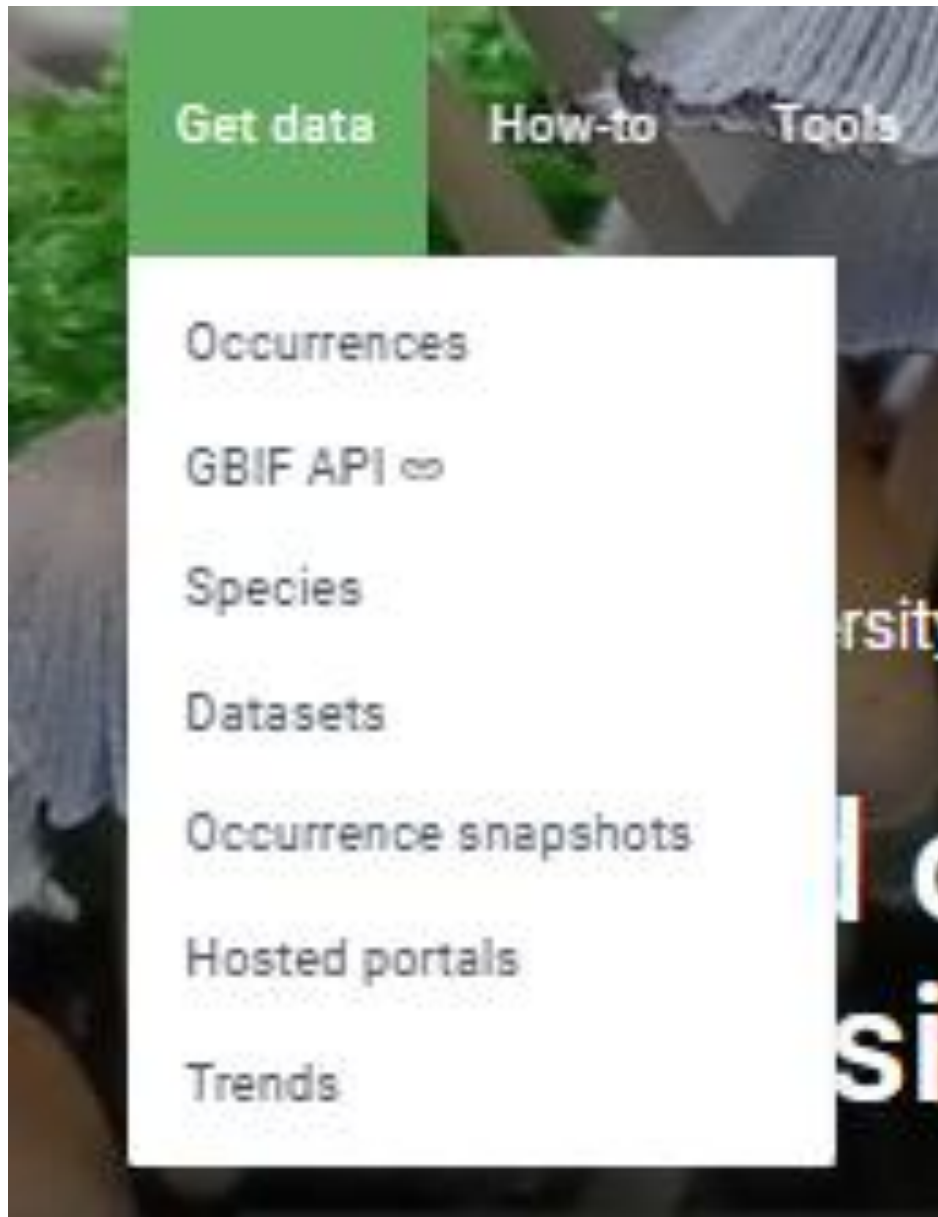
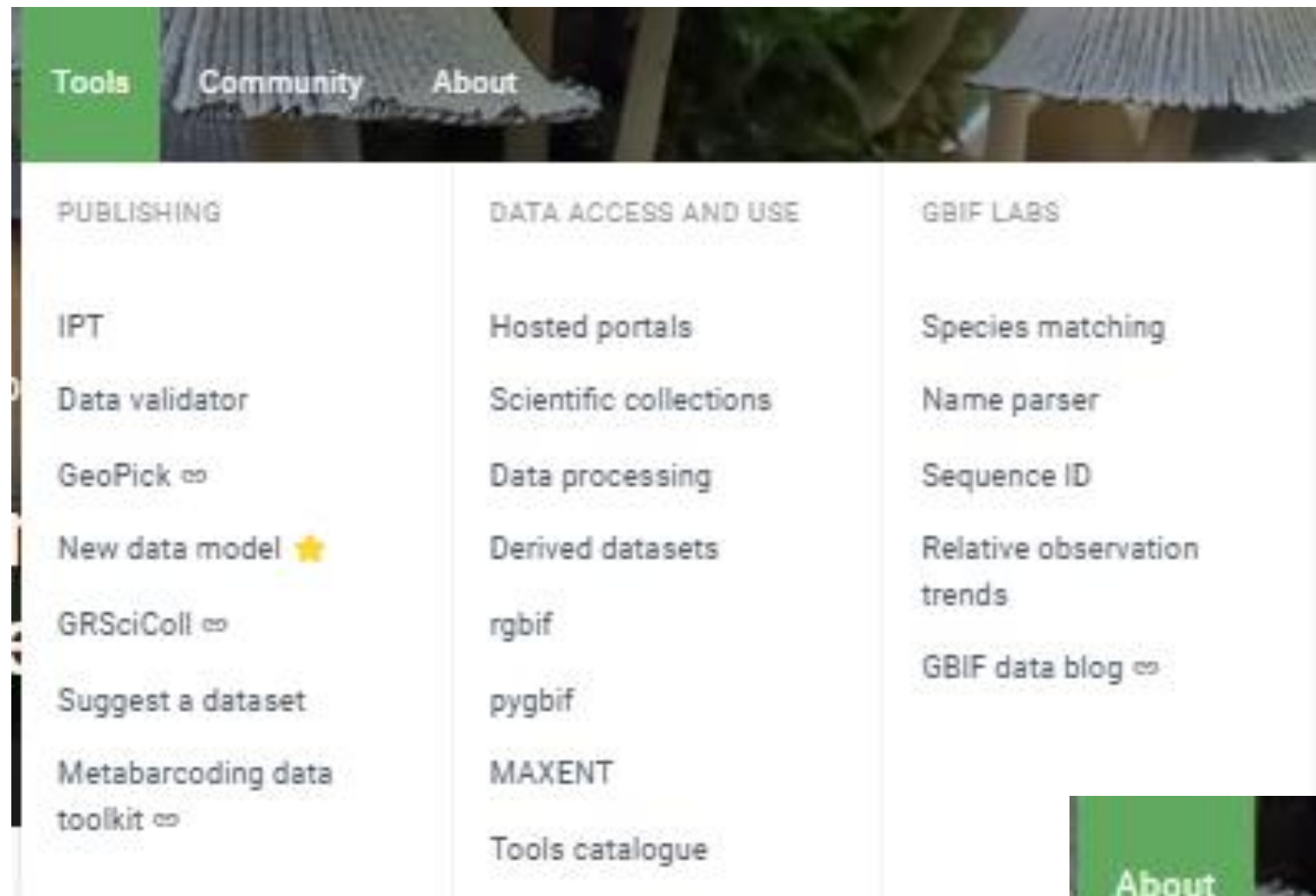
Publishing institutions



11 603

Peer-reviewed papers using data





[Get data](#)[How-to](#)[Tools](#)[Community](#)[About](#)oleg.borodin

< Occurrences

SEARCH OCCURRENCES | 3,068,936,550 RESULTS

Search all fields



Simple filters

All filters

Occurrence status 1

Licence

Scientific name

Basis of record

Year

Month

Location

Administrative areas (gadm.org)

Country or area

Continent

Dataset

Publisher

TABLE

GALLERY

MAP

TAXONOMY

METRICS

DOWNLOAD

Scientific name	Country or area	Coordinates	Event date	Occurrence status	Basis of record
<i>Mareca strepera</i> (Linnaeus, 1758)	France	48.9N, 2.8E	2024 Jan 07	Present	Human observation
<i>Ondatra zibethicus</i> (Linnaeus, 1766)	Netherlands (Kingdom of...)	51.5N, 6.1E	2024 Jan 18	Present	Human observation
<i>Sitta europaea</i> Linnaeus, 1758	Denmark	55.5N, 11.9E	2024 Jan 28	Present	Human observation
<i>Prunella modularis</i> (Linnaeus, 1758)	Germany	49.2N, 7.2E	2024 Jan 11	Present	Human observation
<i>Callidemum</i> Blanchard, 1853	Australia	35.3S, 149.1E	2024 Jan 05	Present	Human observation
<i>Pyrrhula pyrrhula</i> (Linnaeus, 1758)	Russian Federation	54.9N, 73.5E	2024 Jan 03	Present	Human observation
Oecophoridae	Australia	35.3S, 149.1E	2024 Jan 30	Present	Human observation
<i>Cyclamen hederifolium</i> Aiton	United Kingdom of Great ...	50.9N, 0.2W	2024 Jan 13	Present	Human observation
<i>Aegithalos caudatus</i> (Linnaeus, 1758)	Russian Federation	55.4N, 38.4E	2024 Jan 21	Present	Human observation
<i>Pyrrhula pyrrhula</i> (Linnaeus, 1758)	Russian Federation	55.4N, 38.5E	2024 Jan 21	Present	Human observation
<i>Corvus cornix</i> Linnaeus, 1758	Russian Federation	55.4N, 38.5E	2024 Jan 21	Present	Human observation





Occurrences

SEARCH OCCURRENCES | 3,068,936,550 RESULTS

Search all fields

Simple filters All filters

Occurrence status

Licence

Scientific name

Cicadellidae

**Cicadellidae** Family

Animalia > Arthropoda > Insecta > Hemiptera

Simplicillium **cicadellidae** W.H.Chen, C.Li...

Fungi > Ascomycota > Sordariomycetes >

Hypocreales > Cordycipitaceae > Simplicillium

Bacteria 23,367,230

incertae sedis 22,341,189

Chromista 17,304,679

Protozoa 1,584,702

Viruses 914,859

Archaea 422,323

Basis of record

TABLE GALLERY MAP TAXONOMY METRICS DOWNLOAD

	Scientific name	Country or area	Coordinates	Event date	Occurrence status	Basis of record
	<i>Mareca strepera</i> (Linnaeus, 1758)	France	48.9N, 2.8E	2024 Jan 07	Present	Human observation
	<i>Ondatra zibethicus</i> (Linnaeus, 1766)	Netherlands (Kingdom of...)	51.5N, 6.1E	2024 Jan 18	Present	Human observation
	<i>Sitta europaea</i> Linnaeus, 1758	Denmark	55.5N, 11.9E	2024 Jan 28	Present	Human observation
	<i>Prunella modularis</i> (Linnaeus, 1758)	Germany	49.2N, 7.2E	2024 Jan 11	Present	Human observation
	<i>Callidemum</i> Blanchard, 1853	Australia	35.3S, 149.1E	2024 Jan 05	Present	Human observation
	<i>Pyrrhula pyrrhula</i> (Linnaeus, 1758)	Russian Federation	54.9N, 73.5E	2024 Jan 03	Present	Human observation
	Oecophoridae	Australia	35.3S, 149.1E	2024 Jan 30	Present	Human observation
	<i>Cyclamen hederifolium</i> Aiton	United Kingdom of Great ...	50.9N, 0.2W	2024 Jan 13	Present	Human observation
	<i>Aegithalos caudatus</i> (Linnaeus, 1758)	Russian Federation	55.4N, 38.4E	2024 Jan 21	Present	Human observation
	<i>Pyrrhula pyrrhula</i> (Linnaeus, 1758)	Russian Federation	55.4N, 38.5E	2024 Jan 21	Present	Human observation
	<i>Corvus cornix</i> Linnaeus, 1758	Russian Federation	55.4N, 38.5E	2024 Jan 21	Present	Human observation
	<i>Pica pica</i> (Linnaeus, 1758)	Russian Federation	55.4N, 38.5E	2024 Jan 21	Present	Human observation



Occurrences  2

SEARCH OCCURRENCES | 852,507 RESULTS

Search all fields

Simple filters | All filters

Occurrence status 1

Licence

Scientific name

Cicadellidae

Basis of record

Year

Between 1993 - 2024

[CLEAR](#) [ADD](#)

Month

Location

Administrative areas (gadm.org)


Country or area

[TABLE](#)
[GALLERY](#)
[MAP](#)
[TAXONOMY](#)
[METRICS](#)
[DOWNLOAD](#)

Scientific name	Country or area	Coordinates	Event date	Occurrence status	Basis of record
<i>Eurymeloides adspersa</i> (Signoret, 1850)	Australia	35.3S, 149.1E	2024 Jan 05	Present	Human observation
<i>Orosius orientalis</i> (Matsumura, 1914)	Australia	35.3S, 149.1E	2024 Jan 30	Present	Human observation
<i>Acericerus ribauti</i> Nickel & Remane, 2002	Germany	50.8N, 12.9E	2024 Jan 28	Present	Human observation
<i>Katipo signoreti</i> Evans, 1934	Australia	35.3S, 149.2E	2024 Jan 16	Present	Human observation
<i>Brunotartessus fulvus</i> (Walker, 1851)	Australia	35.3S, 149.1E	2024 Jan 16	Present	Human observation
<i>Acericerus heydenii</i> (Kirschbaum, 1868)	United Kingdom of Great ...	55.1N, 1.7W	2024 Jan 12	Present	Human observation
<i>Empoasca</i> Walsh, 1862	United Kingdom of Great ...	54.0N, 1.1W	2024 Jan 24	Present	Human observation
<i>Zygina</i> Fieber, 1866	United Kingdom of Great ...	54.3N, 0.6W	2024 Jan 31	Present	Human observation
<i>Reuplemmeles hobartensis</i> (Evans, 1938)	Australia	35.3S, 149.1E	2024 Jan 28	Present	Human observation
<i>Eurymeloides</i> Ashmead, 1889	Australia	35.2S, 149.1E	2024 Jan 19	Present	Human observation
<i>Tremulicerus vitreus</i> (Fabricius, 1803)	United Kingdom of Great ...	51.5N, 0.2W	2024 Jan 05	Present	Preserved specimen





**Occurrences**  3

Cicadellidae

Basis of record ▾

Year ▲

Between ▾ ×  
1993 2024

CLEAR ADD

Month ▾

Location ▾

Administrative areas (gadm.org) ▾

Country or area ▲

Kazakhstan 8

Search

United States of America 138 252

Canada 133 601

Costa Rica 128 961

SEARCH OCCURRENCES | 8 RESULTS

TABLE GALLERY MAP TAXONOMY METRICS DOWNLOAD

Scientific name	Country or area	Coordinates	Event date	Occurrence status	Basis of record
<i>Eremophlepsius sexnotatus</i> (Kusnezov, 1929)	Kazakhstan		2022 Jun 24	Present	Material citation
<i>Pseudophlepsius binotatus</i> (Signoret, 1880)	Kazakhstan		2022 Jun 24	Present	Material citation
<i>Pseudophlepsius binotatus</i> (Signoret, 1880)	Kazakhstan		2019 Jun 27	Present	Material citation
<i>Taurotettix modestus</i> (Mitjaev, 1971)	Kazakhstan		2019 Jun 25	Present	Material citation
<i>Taurotettix beckeri</i> (Fieber, 1885)	Kazakhstan		2019 Jun 11	Present	Material citation
<i>Fieberiella septentrionalis</i> Wagner, 1963	Kazakhstan	43.3N, 77.2E	2019 Sep 25	Present	Human observation
<i>Cicadella viridis</i> (Linnaeus, 1758)	Kazakhstan	49.9N, 82.7E	2016 Jul 15	Present	Human observation
<i>Taurotettix modestus</i> (Mitjaev, 1971)	Kazakhstan		2004 Jul 02	Present	Material citation

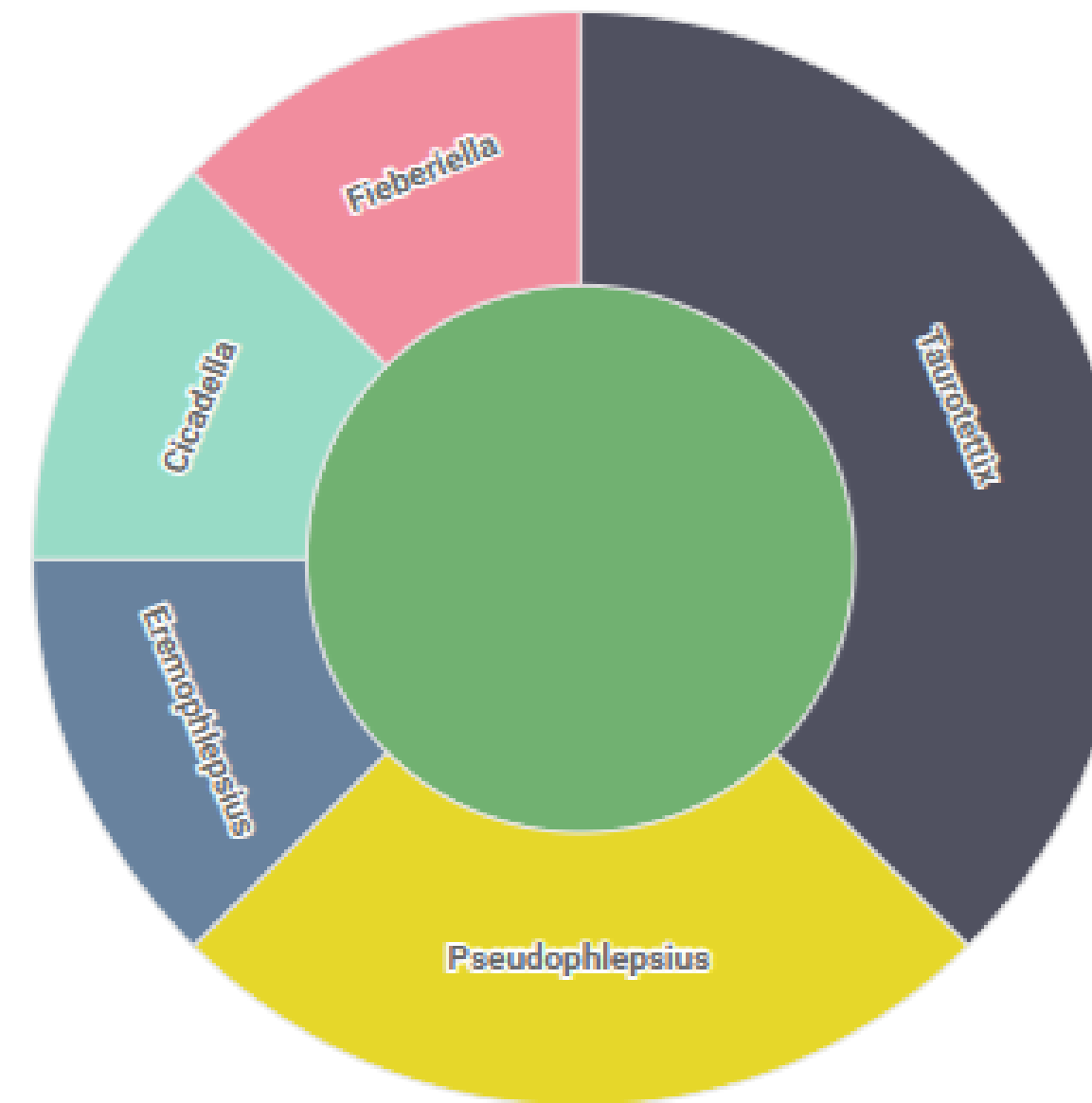


TAXONOMIC DISTRIBUTION OF OCCURRENCES

Explore Major groups

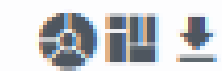
Animalia	8
Chordata	250,792
Arthropoda	67,736
Mollusca	139
Annelida	111
Platyhelminthes	99
Nematoda	98
Echinodermata	75
Porifera	9
Bryozoa	8
Unknown phylum	-319,059

TAXONOMIC DISTRIBUTION OF OCCURRENCES





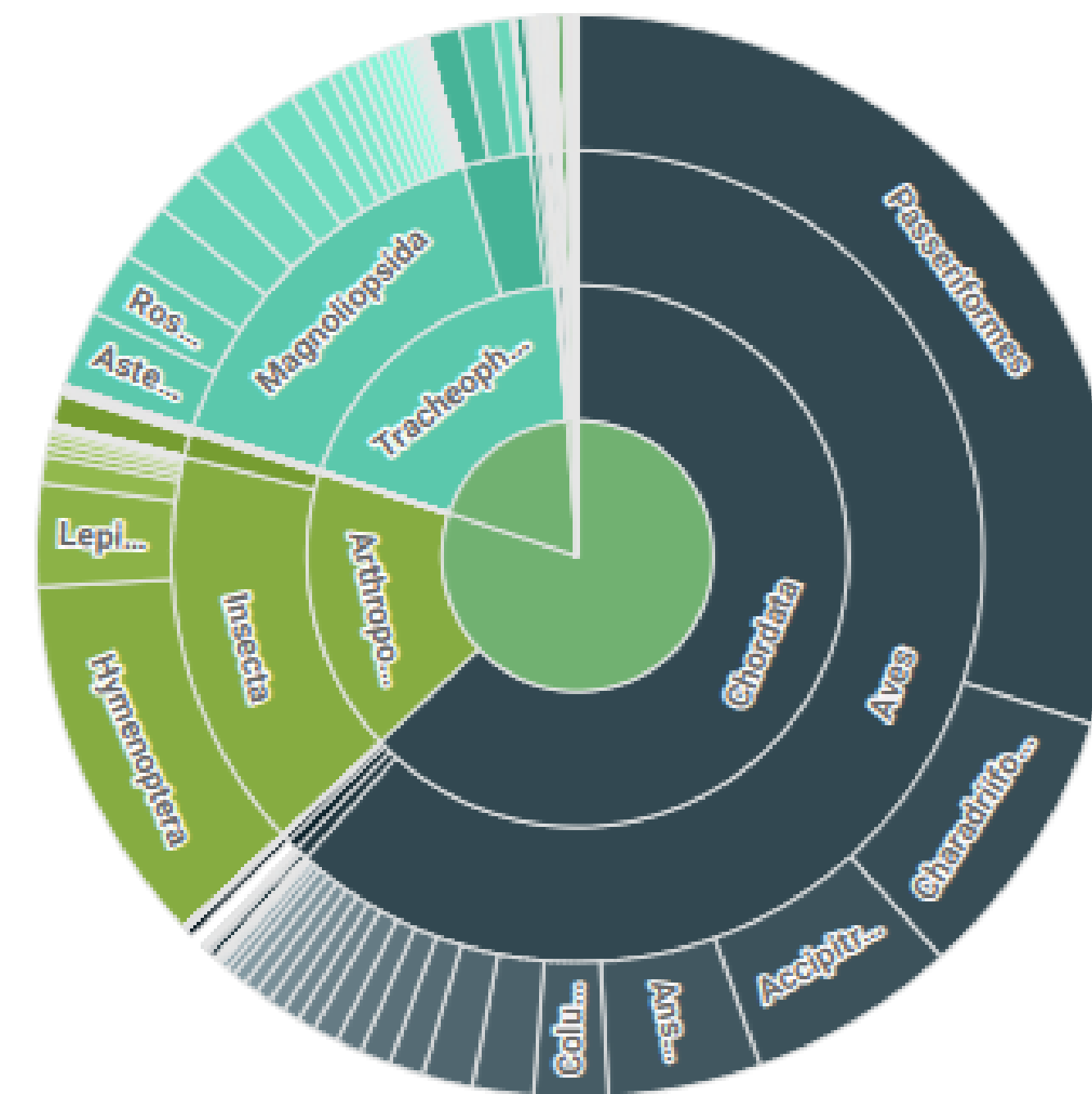
TAXONOMIC DISTRIBUTION OF OCCURRENCES

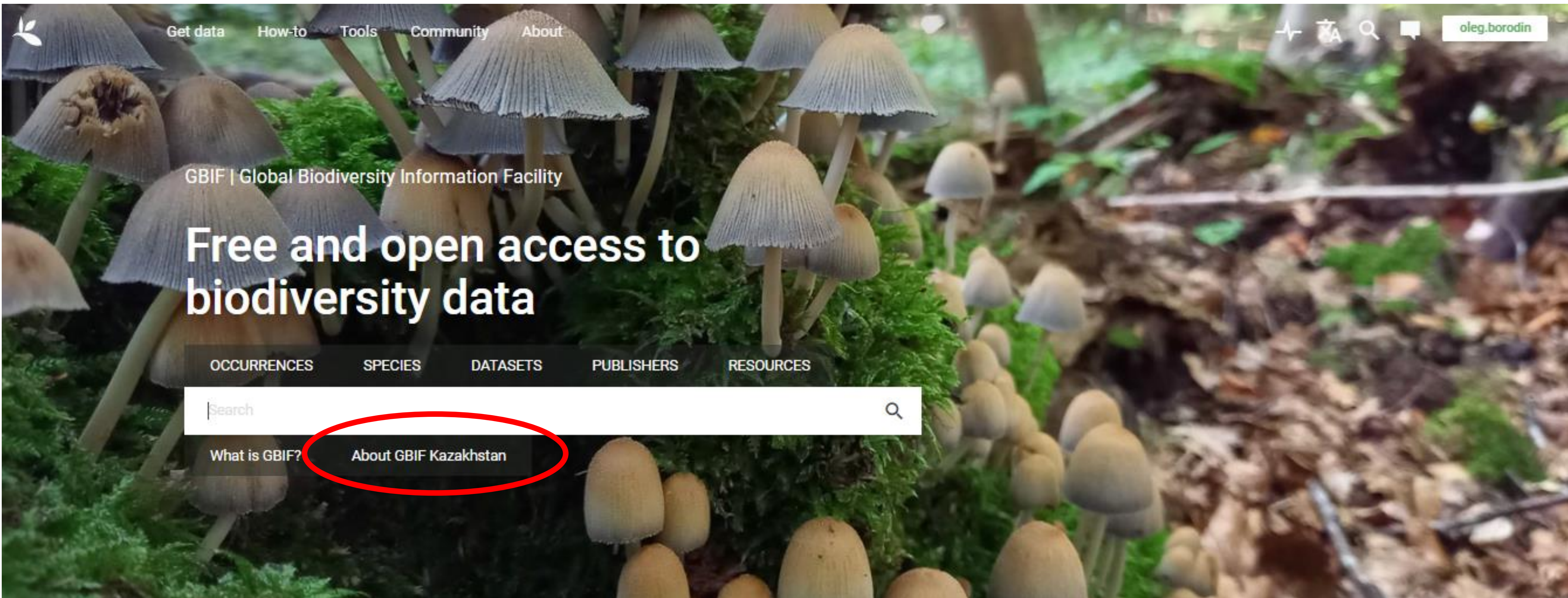


Explore Major groups

Animalia	357,189
Plantae	175,610
Fungi	3,678
Protozoa	1,781
incertae sedis	1,088
Viruses	719
Bacteria	406
Chromista	394
Archaea	7

TAXONOMIC DISTRIBUTION OF OCCURRENCES





## DATA ABOUT KAZAKHSTAN

540,872

Occurrences

793

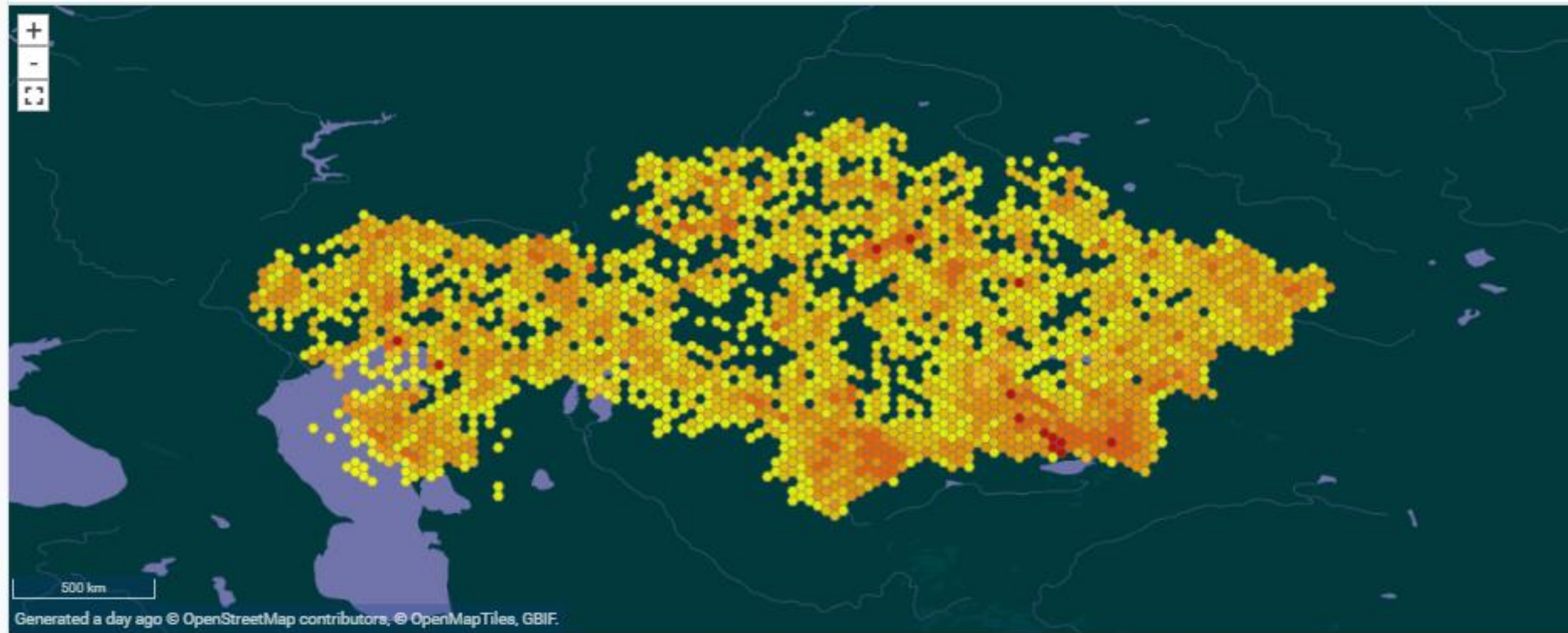
Datasets

44

Countries and areas contribute  
data

305

Publishers



Generated a day ago © OpenStreetMap contributors, © OpenMapTiles, GBIF.

Any year

1765 - 2024

EXPLORE AREA



DATA FROM KAZAKHSTAN

221,986

Published occurrences

20

Published datasets

15

Countries and areas covered by data from Kazakhstan

10

Publishers from Kazakhstan




Any year 1842 - 2024

EXPLORE AREA



[Get data](#)[How-to](#)[Tools](#)[Community](#)[About](#)oleg.borodin

**Occurrences**  3

Cicadellidae

Basis of record ▾

Year ▲

Between ▾ ×

1993 2024

[CLEAR](#) [ADD](#)

Month ▾

Location ▾

Administrative areas (gadm.org) ▾

Country or area ▲

Kazakhstan 8

Search

United States of America 138 252

Canada 133 601

Costa Rica 128 961

SEARCH OCCURRENCES | 8 RESULTS

[TABLE](#)
[GALLERY](#)
[MAP](#)
[TAXONOMY](#)
[METRICS](#)
[DOWNLOAD](#)

Scientific name	Country or area	Coordinates	Event date	Occurrence status	Basis of record
<i>Eremophlepsius sexnotatus</i> (Kusnezov, 1929)	Kazakhstan		2022 Jun 24	Present	Material citation
<i>Pseudophlepsius binotatus</i> (Signoret, 1880)	Kazakhstan		2022 Jun 24	Present	Material citation
<i>Pseudophlepsius binotatus</i> (Signoret, 1880)	Kazakhstan		2019 Jun 27	Present	Material citation
<i>Taurotettix modestus</i> (Mitjaev, 1971)	Kazakhstan		2019 Jun 25	Present	Material citation
<i>Taurotettix beckeri</i> (Fieber, 1885)	Kazakhstan		2019 Jun 11	Present	Material citation
<i>Fieberiella septentrionalis</i> Wagner, 1963	Kazakhstan	43.3N, 77.2E	2019 Sep 25	Present	Human observation
<i>Cicadella viridis</i> (Linnaeus, 1758)	Kazakhstan	49.9N, 82.7E	2016 Jul 15	Present	Human observation
<i>Taurotettix modestus</i> (Mitjaev, 1971)	Kazakhstan		2004 Jul 02	Present	Material citation



	Raw data	Interpreted data	Multimedia	Coordinates	Format	Estimated data size
<a href="#">↓ SIMPLE</a>	X	✓	X	✓ (if available)	Tab-delimited CSV (for use in Excel, etc.) <a href="#">?</a>	<b>4 KB</b> (952 Bytes zipped for download)
<a href="#">↓ DARWIN CORE ARCHIVE</a>	✓	✓	✓ (links)	✓ (if available)	Tab-delimited CSV (for use in Excel, etc.) <a href="#">?</a>	<b>13 KB</b> (3 KB zipped for download)
<a href="#">↓ SPECIES LIST</a>	X	✓	X	X	Tab-delimited CSV (for use in Excel, etc.) <a href="#">?</a>	

#### DOWNLOAD REPORT

**Total:** 8

**Licence:** CC BY-NC 4.0

**Year range:** 2004–2022

**With year:** 100 %

**With coordinates:** 25 %

**With taxon match:** 100 %

#### Known issues

A part of the GBIF processing is to flag occurrences that have suspicious fields

6 Occurrence status inferred from individual count   2 Continent derived from coordinates   2 Taxon match fuzzy



	Raw data	Interpreted data	Multimedia	Coordinates	Format	Estimated data size
<a href="#">↓ SIMPLE</a>	X	✓	X	✓ (if available)	Tab-delimited CSV (for use in Excel, etc.) ⓘ	<b>291 MB</b> (64 MB zipped for download)
<a href="#">↓ DARWIN CORE ARCHIVE</a>	✓	✓	✓ (links)	✓ (if available)	Tab-delimited CSV (for use in Excel, etc.) ⓘ	<b>891 MB</b> (197 MB zipped for download)
<a href="#">↓ SPECIES LIST</a>	X	✓	X	X	Tab-delimited CSV (for use in Excel, etc.) ⓘ	

#### DOWNLOAD REPORT

**Total:** 540 872

**Licence:** CC BY-NC 4.0

**Year range:** 1765–2024

**With year:** 100 %

**With coordinates:** 76 %

**With taxon match:** 99.8 %

#### Known issues

A part of the GBIF processing is to flag occurrences that have suspicious fields

400 295

Continent derived from coordinates

22 064

Occurrence status inferred from individual count

15 877

Country derived from coordinates





## Occurrences 3

Search all fields [Simple filters](#) [All filters](#)Occurrence status Licence Scientific name  CicadellidaeBasis of record Year  Between start of 1993 and end of 2024Month Location Administrative areas (gadm.org) [Raw data](#)[Interpreted data](#)[Multimedia](#)[Coordinates](#)[Format](#)

Estimated data size

### Free of cost – not free of responsibilities

While data from GBIF.org is free and open, please remember that by downloading this data, you are agreeing:

- to abide by the [GBIF user agreement](#)
- and, if you use the data, to [cite it appropriately](#)

Please make sure your citation includes the unique **DOI** (shown on the page once it refreshes). The use of properly formatted data citations ensures scientific transparency and reproducibility and enables proper attribution of credit to the data providers.

If you are analysing the data you will download, please consider referencing this citation in your Materials and methods section.

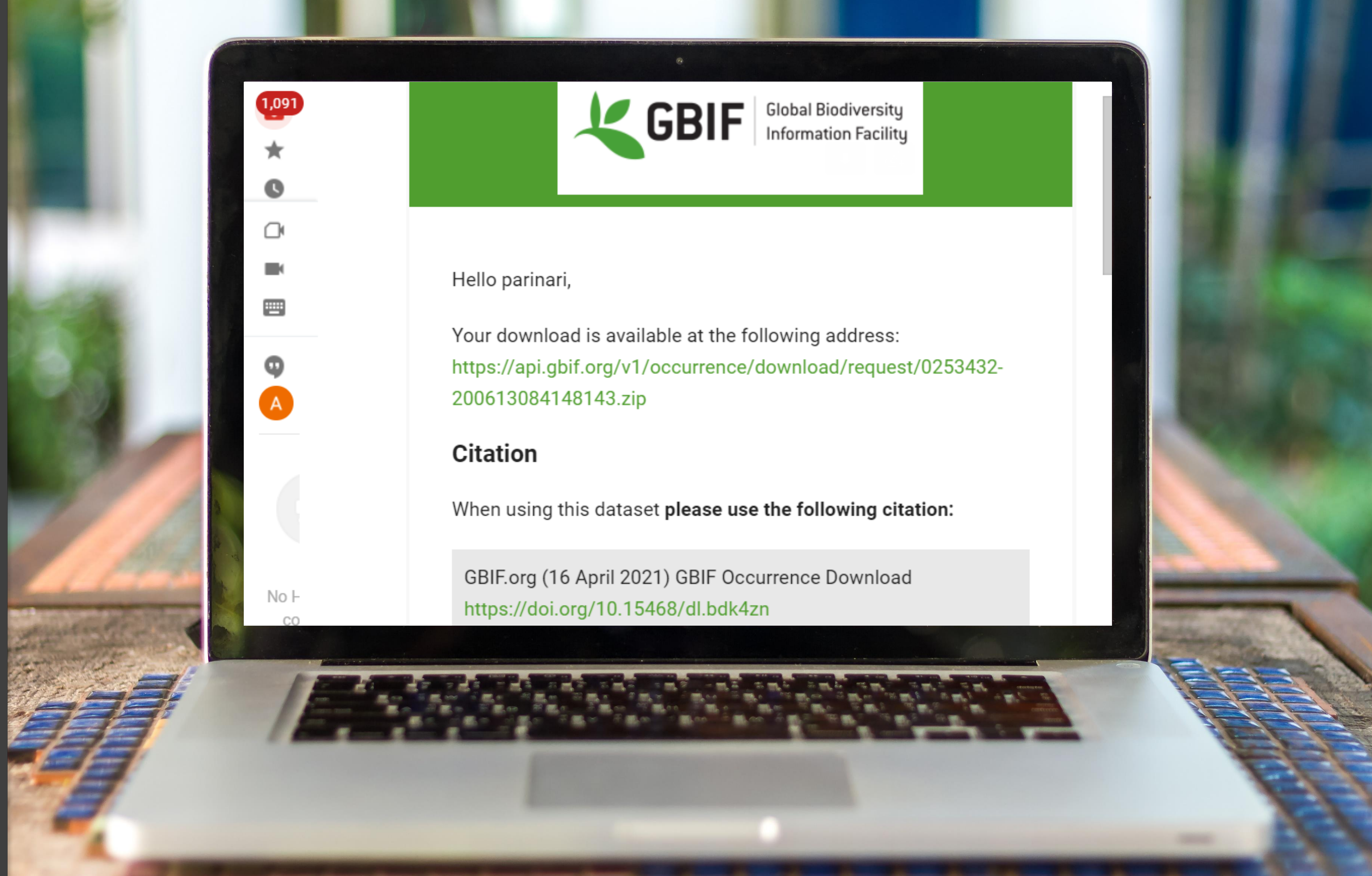
**New to GBIF?** Check out the [Data Use Club](#).

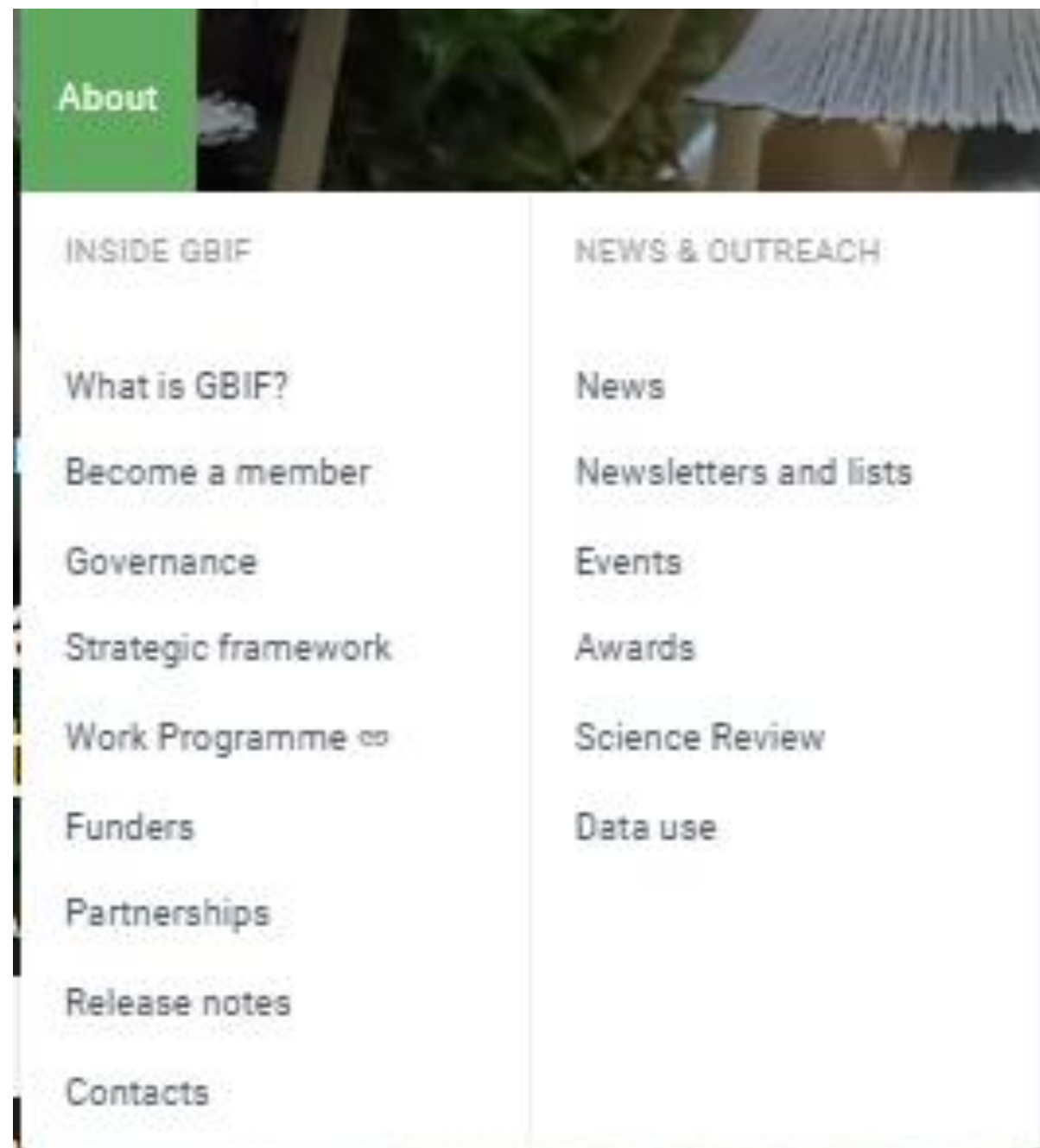
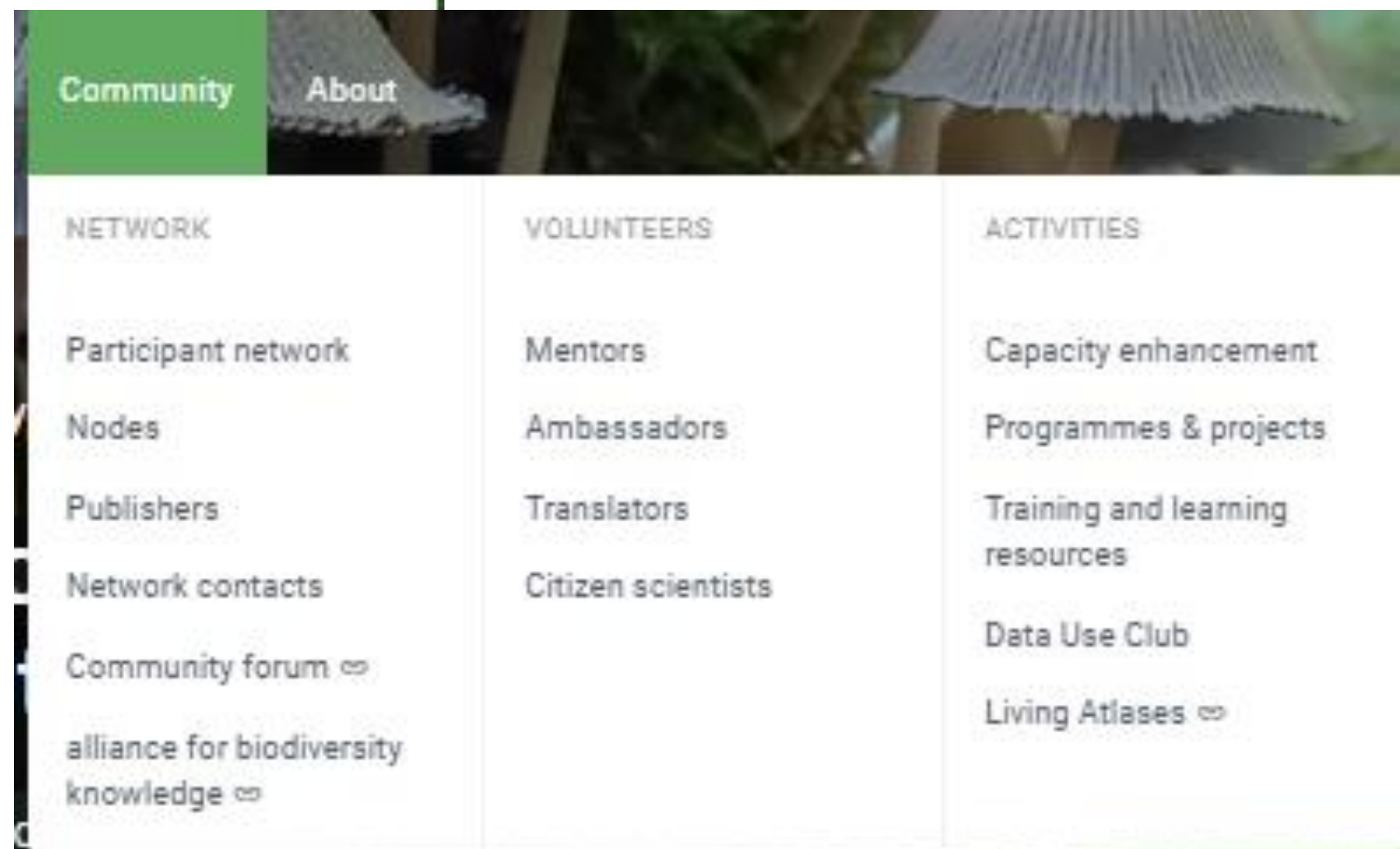
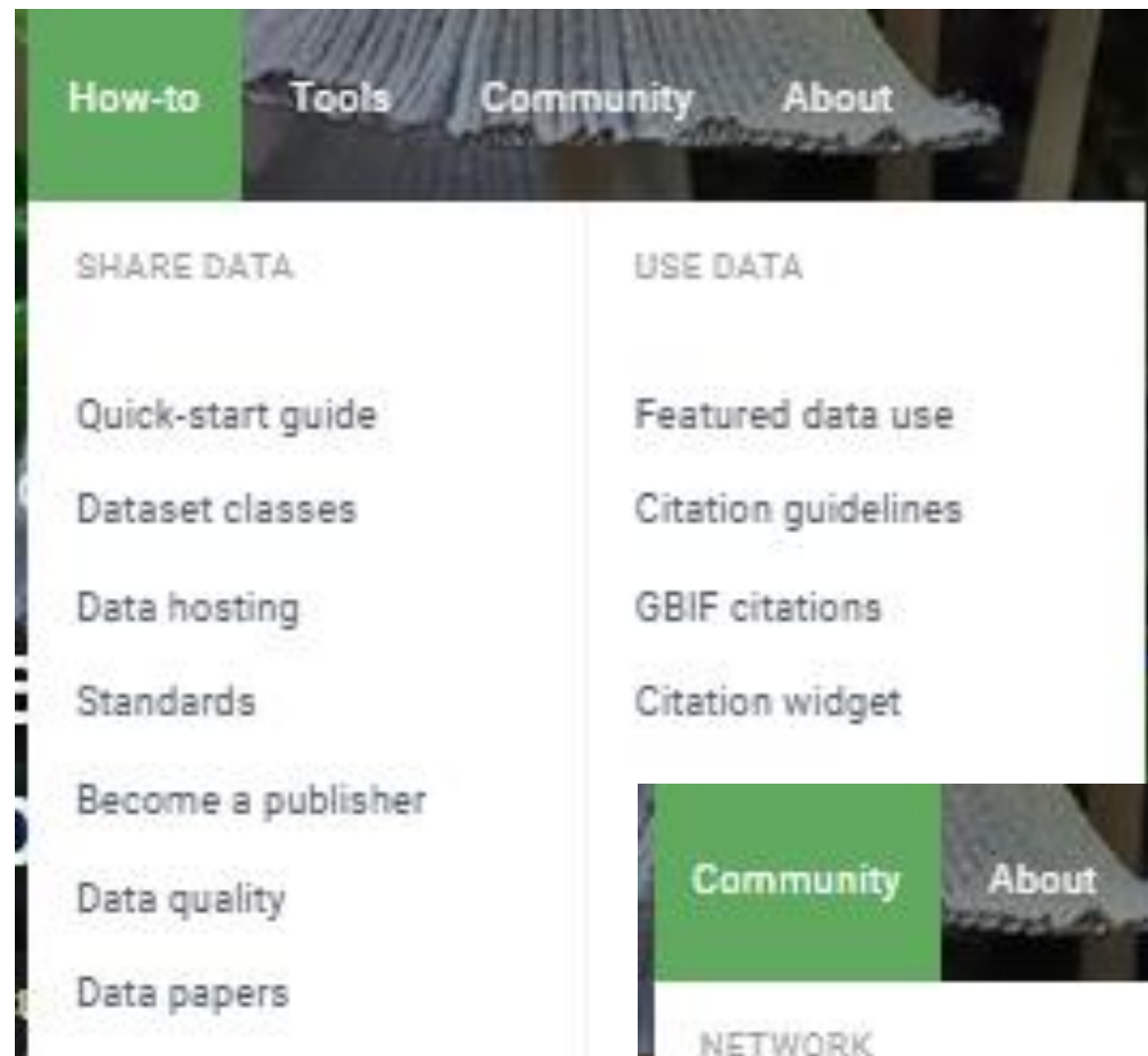
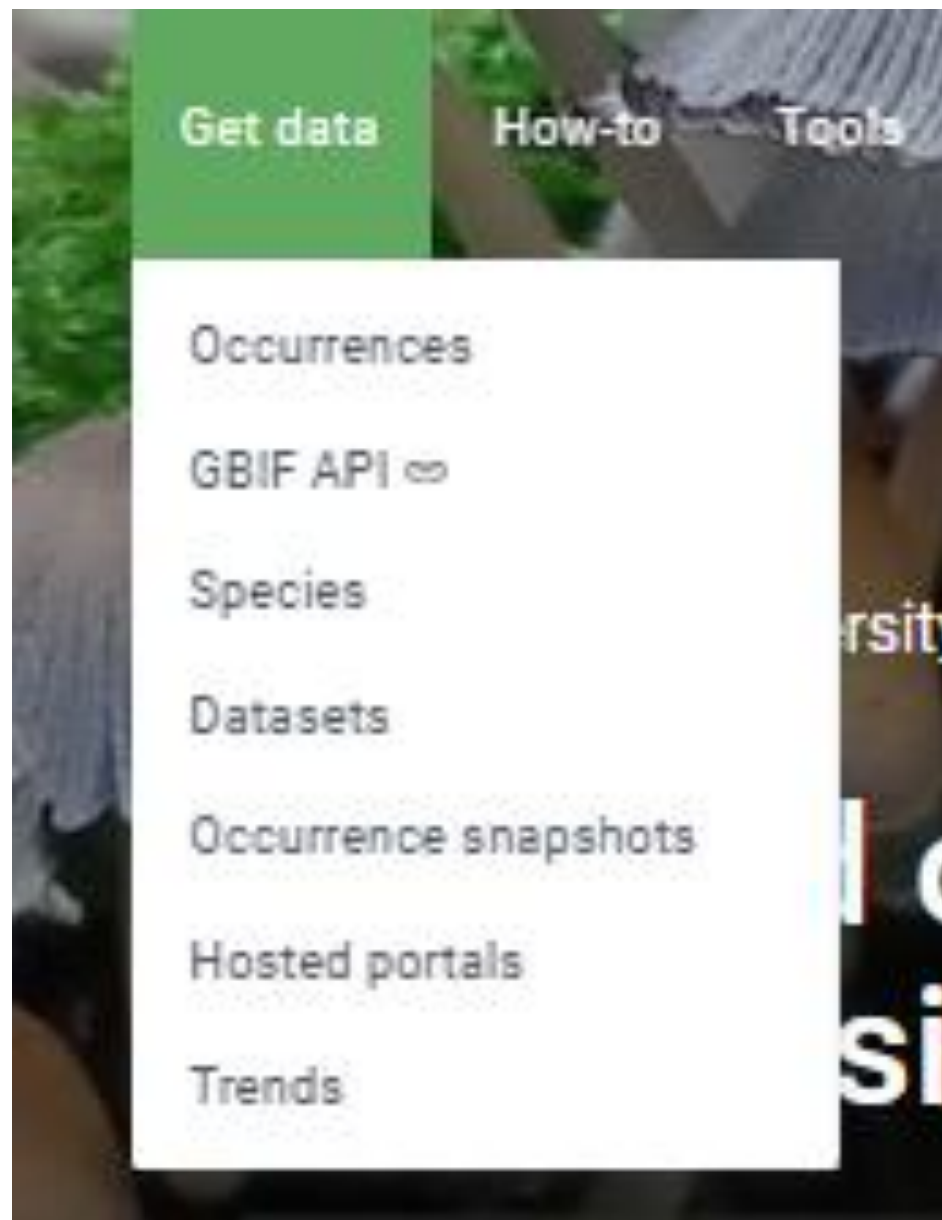
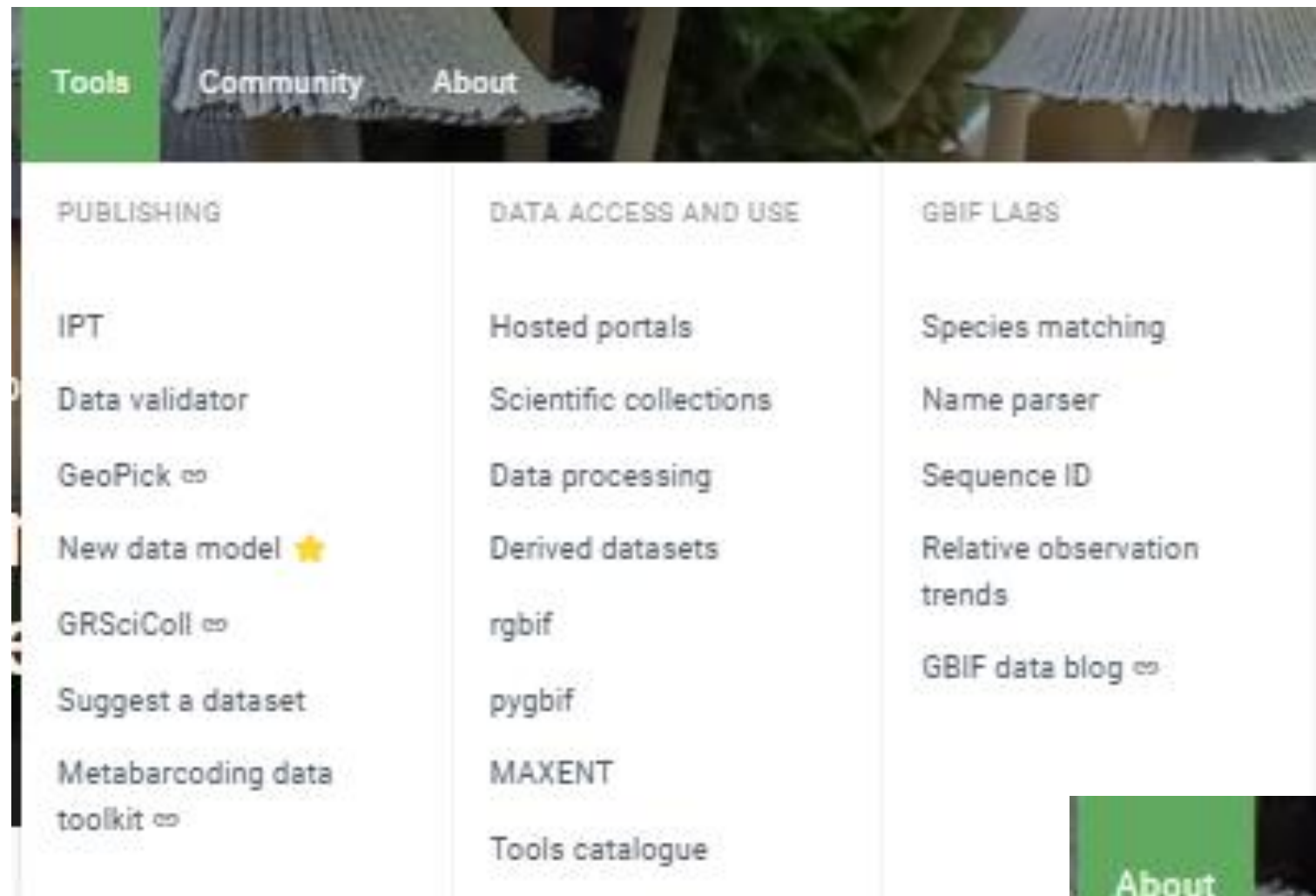
[Go back and add more filters](#)[UNDERSTOOD](#)**Known issues**



# Post-Download

Citations are provided at the time of download and in the confirmation email.







Resources

SEARCH RESOURCES | 12 RESULTS

Search 🔍

ALL NEWS DATA USE **EVENTS** PROJECTS PROGRAMMES TOOLS DOCUMENTS LITERATURE

Upcoming Past

Search for news, events, documents and much more. If you are looking for species occurrences then try our occurrence search.

[OCCURRENCE SEARCH](#)

+ SUGGEST EVENT SUBSCRIBE TO CALENDAR RSS

**Training workshop - Working with biodiversity data** Event

The event is aimed at the current and aspiring coordinators of biodiversity data activities at the national and institutional levels, community managers in need of data and GBIF skills, and at the adv...

📍 Kazakhstan - Almaty  
📅 November 20, 2024 - November 22, 2024

20 Nov  
2024

ADD TO CALENDAR [VIEW ON GBIF](#)

**Africa Regional Nodes Meeting | Engagement meeting for BID Africa** Event

Meeting goals The Africa nodes meeting aims to: Strengthen the GBIF Africa community Provide a forum for GBIF nodes in Africa and the GBIF Secretariat to share latest updates and best practices Showc...

📍 Zimbabwe - Harare - This event is by invitation only  
📅 November 27, 2024 - November 29, 2024  
👤 Attended by GBIF Secretariat

27 Nov  
2024

ADD TO CALENDAR

**International Conference on Research Infrastructures** Event

The International Conference on Research Infrastructures is the major professional forum for the global research infrastructure community. ICRI 2024 will bring together policymakers, research institut...

📍 Australia - Brisbane  
📅 December 3, 2024 - December 5, 2024

3 Dec  
2024



Resources

SEARCH RESOURCES | 449 RESULTS

ALL NEWS DATA USE EVENTS PROJECTS PROGRAMMES TOOLS DOCUMENTS LITERATURE

RSS

### Republic of Benin becomes a GBIF Voting Participant

News

West African nation's exemplary effort in building capacity and national communities of practitioners has made it a key centre of research- and policy-relevant data mobilization ... Building on its inc...

Published November 14, 2024



### GBIF appoints four contractors to boost network capacity for BID programme

News

Experts from Africa and Latin America to extend support of regional communities of practice for the EU-funded Biodiversity Information for Development ... Contractors from Uganda, Madagascar, Colombia ...

Published November 7, 2024



### Liam Lysaght re-elected as chair of the GBIF Governing Board

News

André Heughebaert of Belgium and Soledad Ceccarelli of Argentina voted into new governance positions amid other re-elections and reports on current work across the GBIF network ... GBIF's voting partic...

Published October 15, 2024



### ChatIPT system wins the 2024 Ebbe Nielsen Challenge

News

Assistant developed by Norwegian engineer Rukaya Johaadien helps transform spreadsheets into standardized GBIF-ready datasets; Planetary Knowledge Base and CoreTech Assistant place second and third in...



ch for news, events, documents and much  
If you are looking for species occurrences  
then try our occurrence search.

[OCCURRENCE SEARCH](#)



Occurrences 1 SEARCH OCCURRENCES | 47,003 WITH IMAGES

SEARCH ALL FIELDS

Simple filters All filters

Occurrence status 1

Presence

Scientific name

Cercopidae

Search

Explore Major groups

- Animalia 120,584
- Chordata 2,137,565,042
- Arthropoda 276,200,078
- Mollusca 21,701,046
- Annelida 6,462,415
- Cnidaria 4,718,470
- Echinodermata 2,979,127
- Porifera 1,677,637
- Brachiopoda 941,068
- Bryozoa 805,524
- Nematoda 709,820
- Platyhelminthes 565,676

TABLE GALLERY MAP TAXONOMY METRICS DOWNLOAD

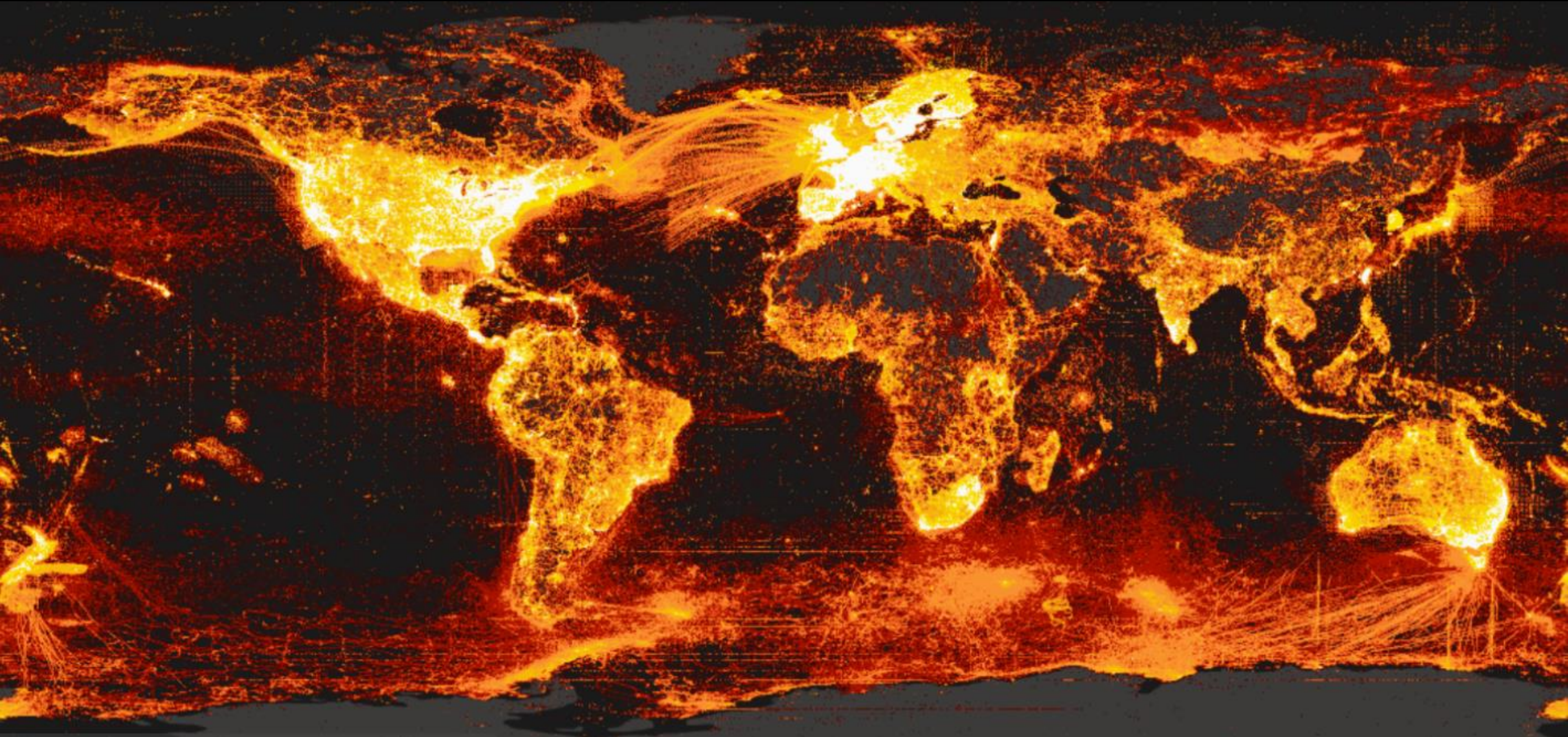
*Locris rubida* (Stål, 1855) *Notozulia enterriana* (Berg, 1879) *Locris arithmetica* (Walker, 1851) *Locris areata* (Walker, 1851) *Mahanarva rubripennis* (Schmidt, 1922) *Prosapia simulans* (Walker, 1858)

*Locris transversa* (Thunberg, 1822) *Maxantonia stabilis* Naest, 1979 *Sphenorhina rubra* (Linnaeus, 1758) *Anyllis leiala* Kirkaldy, 1906 *Mahanarva costaricensis* (Distant, 1879) *Sphenorhina rubra* (Linnaeus, 1758)

*Caloscarta capitata* (Stål, 1865) *Locris arithmetica* (Walker, 1851) *Sphenorhina limbata* (Lallemand, 1927) *Prosapia bicincta* (Say, 1830) *Callitettix versicolor* (Fabricius, 1794)

Thank you for your attention!!!





# Data citation and GBIF powered literature

Dmitry Schigel | Scientific officer



*Biodiversity data in montane and arid Eurasia  
Almaty, Kazakhstan*

*18-19 November 2024*



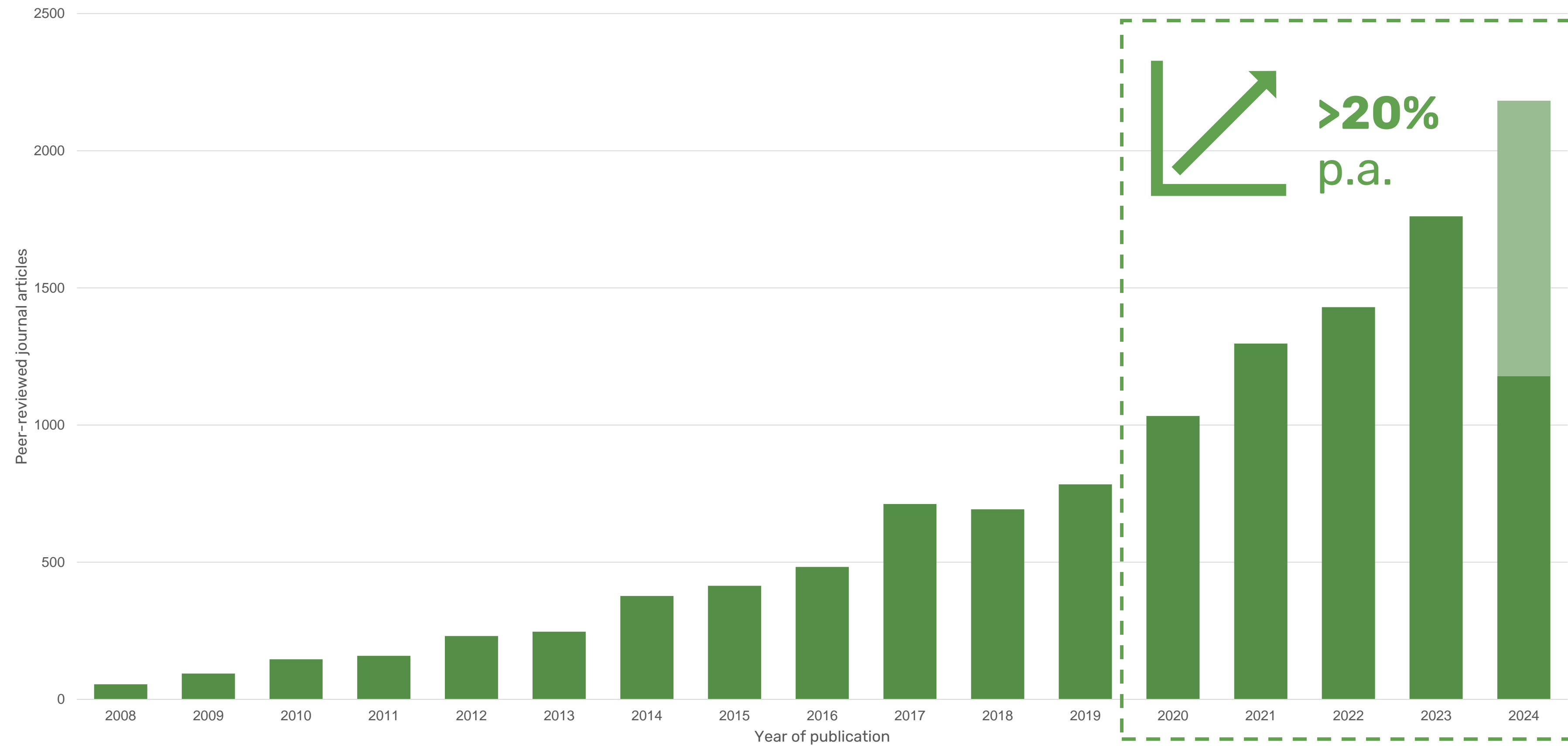
# Why track citations?

- Knowledge of where, when, how and by whom data is used
- Evidence of the impact
- Credit to data publishers
- Started in 2010
- Optimized, streamlined and automated since 2017






# Use of GBIF-mediated data in research



## Data citations: All about the DOIs



- **DOIs** for datasets
- **DOIs** for downloads
- **DOIs** for derived datasets
  
- Paper → (download)  dataset(s)
- Dataset citations

OCCURRENCE DATASET | REGISTERED OCTOBER 11, 2018

# Azorean Biodiversity Portal



Published by [Universidade dos Açores](#)

Borges P A V • Gabriel R • Arroz A M • Costa A • Cunha R • Silva L • Pereira E • Martins A F • Reis F • Cardoso P



DATASET PROJECT METRICS ACTIVITY  DOWNLOAD  HOME PAGE

1,652,342 OCCURRENCES 695 CITATIONS

BECAUSE YOU ARE TRUSTED CONTACT



HISTORY  INGEST NOW  LOGS

The University of Azores hosts the Azores Bioportal (<http://azoresbioportal.uac.pt/>) a regional E-Infrastructure. The Azorean Biodiversity Portal (ABP) is an e-infrastructure now associated with Portuguese PORBIOTA and LIFEWATCH. The ABP is a key e-infrastructure for the integrated management of biodiversity data of the Azores, providing a large number of specialized services supporting research, policy and education. The 3000 visits per day, the numerous international scientific collaborations,... [More](#)

 **Publication date:** March 22, 2023  
**Metadata last modified:** March 22, 2023  
**Hosted by:** [Instituto Superior de Agronomia / Universidade de Lisboa](#)  
**Licence:** [CC BY 4.0](#)  
 How to cite [DOI 10.15468/jdiceo](#)

1,652,342 Occurrences 99.9% With taxon match 100% With coordinates 99.9% With year

1,652,342 GEOREFERENCED RECORDS



# Data in 108,709 datasets: attribution, credit and affiliation

Get data | How-to | Tools | Community | About

OCCURRENCE DATASET | REGISTERED JULY 27, 2009

## Swiss National Bryophyte Databank

Published by [Swiss National Biodiversity Data and Information Centres – infospecies.ch](https://www.infospecies.ch)

Hofmann H • Cailliau A • Hartwig A

DATASET | METRICS | ACTIVITY | DOWNLOAD | HOME PAGE

236,552 OCCURRENCES | 111 CITATIONS

This dataset is maintained by Swissbryophytes (National Data- and Information Center of Swiss Bryophytes, formerly "National Inventory of Swiss bryophytes", NISM). We are a member of InfoSpecies. The dataset includes records of Bryophytes (Anthocerotophyta, Bryophyta, Marchantiophyta) from Switzerland and the adjacent area. Data sources include official herbaria and private collections from a large network of volunteer collaborators, inventories (National Inventory of Swiss bryophytes NISM, Red ... [More](#))

**Swissbryophytes**

Publication date: March 8, 2024  
Metadata last modified: March 8, 2024  
Hosted by: GBIF Swiss Node  
Licence: CC BY 4.0  
How to cite | DOI: 10.15468/ajkhha

236,552 Occurrences | 100% With taxon match | 100% With coordinates | 100% With year

236,552 GEOREFERENCED RECORDS

1 Affiliation

2 Authorship

3 Data citations

4 DOI

Frontiers of Biogeography 2021, 13.04, e51146

**RESEARCH ARTICLE** | Frontiers of Biogeography  
the scientific journal of the International Biogeography Society

### Climatic drivers of *Sphagnum* species distributions

Charles Campbell<sup>1,2\*</sup>, Gustaf Granath<sup>2</sup> and Håkan Rydin<sup>2</sup>

<sup>1</sup> Greensway AB, Ulls väg 24A, 756 51 Uppsala, Sweden; <sup>2</sup> Department of Ecology and Genetics, Evolutionary Biology Centre, Uppsala University, Norbyvägen 18D, SE-752 36 Uppsala, Sweden.  
\*Correspondence: Charles Campbell, charlescampbell@outlook.com

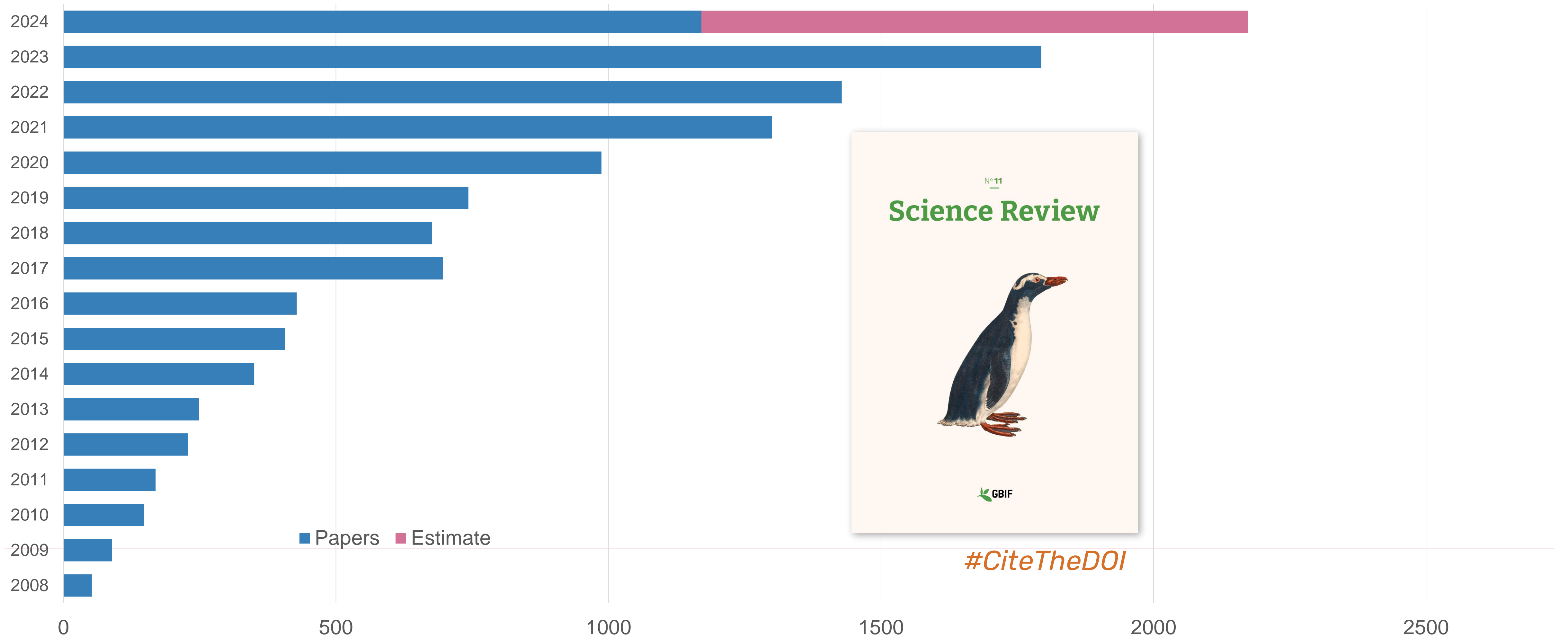
**Abstract**  
Peat mosses (genus *Sphagnum*) dominate most Northern mires and show distinct distributional limits in Europe despite having efficient dispersal and few dispersal barriers. This pattern indicates that *Sphagnum* species distributions are strongly linked to climate. *Sphagnum*-dominated mires have been the largest terrestrial carbon sinks in Europe over the last few millennia. Understanding the climatic drivers of *Sphagnum* species distributions is important for predicting the future functionality of peatlands. We used MaxEnt, with biologically relevant climatic variables, to model and clarify the current distributions of 45 *Sphagnum* species in Europe. We

**Highlights**

- Peat mosses (*Sphagnum*) form northern peatlands and species have different distributions across Europe.
- We model the climatic suitability for all European species using multiple databases and MaxEnt models.
- The climatic suitability for most species can be accurately modelled with mean annual temperature and water balance and their variation over the year.
- *Sphagnum* has its highest species richness in northwestern Europe.

Hofmann H, Kiebacher T, Moser T, Meier M (2021). Swiss National Bryophyte Databank. Swiss National Biodiversity Data and Information Centres – infospecies.ch. Occurrence dataset <https://doi.org/10.15468/ajkhha> accessed via GBIF.org on 2022-04-28.

# Peer-reviewed publications using GBIF-mediated data



# Data citations

## Ideal world

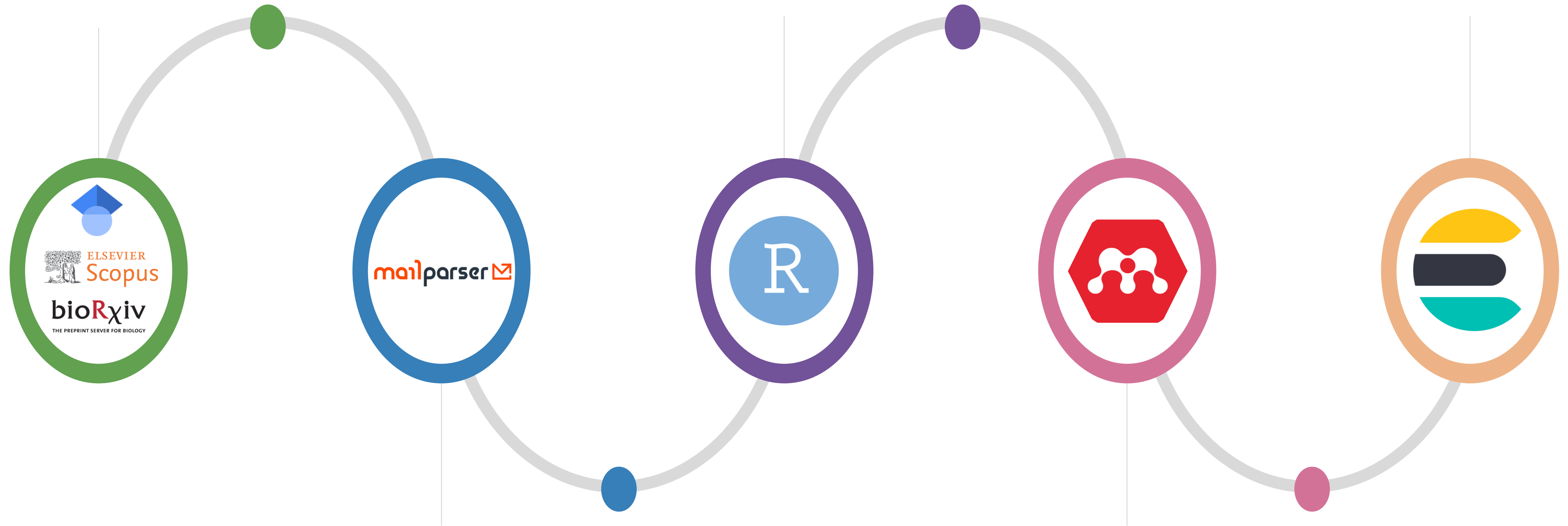
- All citations include a DOI, uniquely and specifically identifying the data used in a publication
- All data citations are included in the reference list of an article
- All data citations are machine-readable links between DOIs included in the DOI metadata, easily interpreted and used for quantifying citations
- 100% automatic 🤖

## Real world

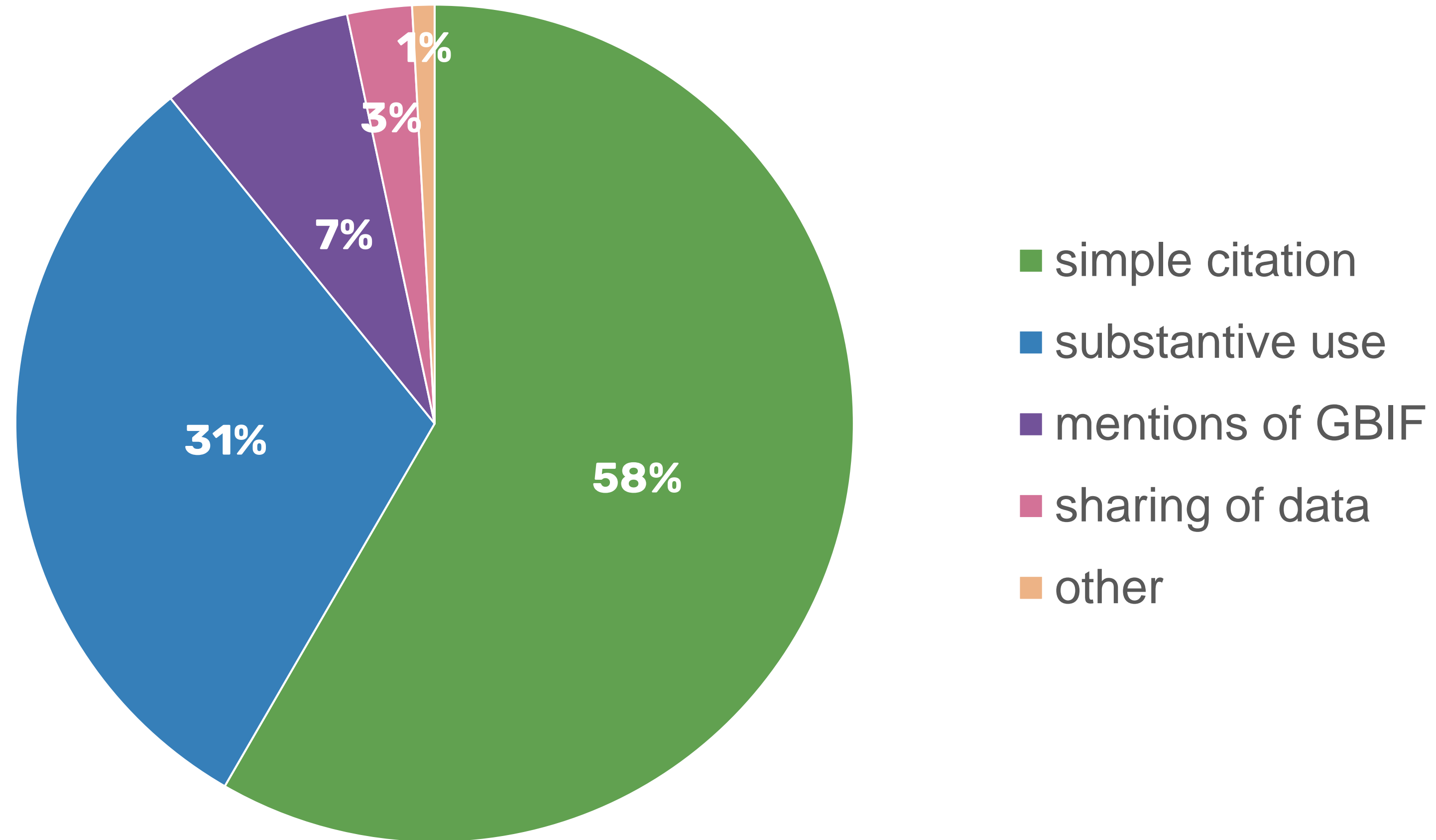
- Far from all citations include a DOI, and some tend to cite more data than is actually used
- Some citations are included in the body text of papers, some in data availability statements, and some are even hidden in supplements
- Even when done right, not all citations become available as machine-readable links between papers and dataset
- Automatic? 😬



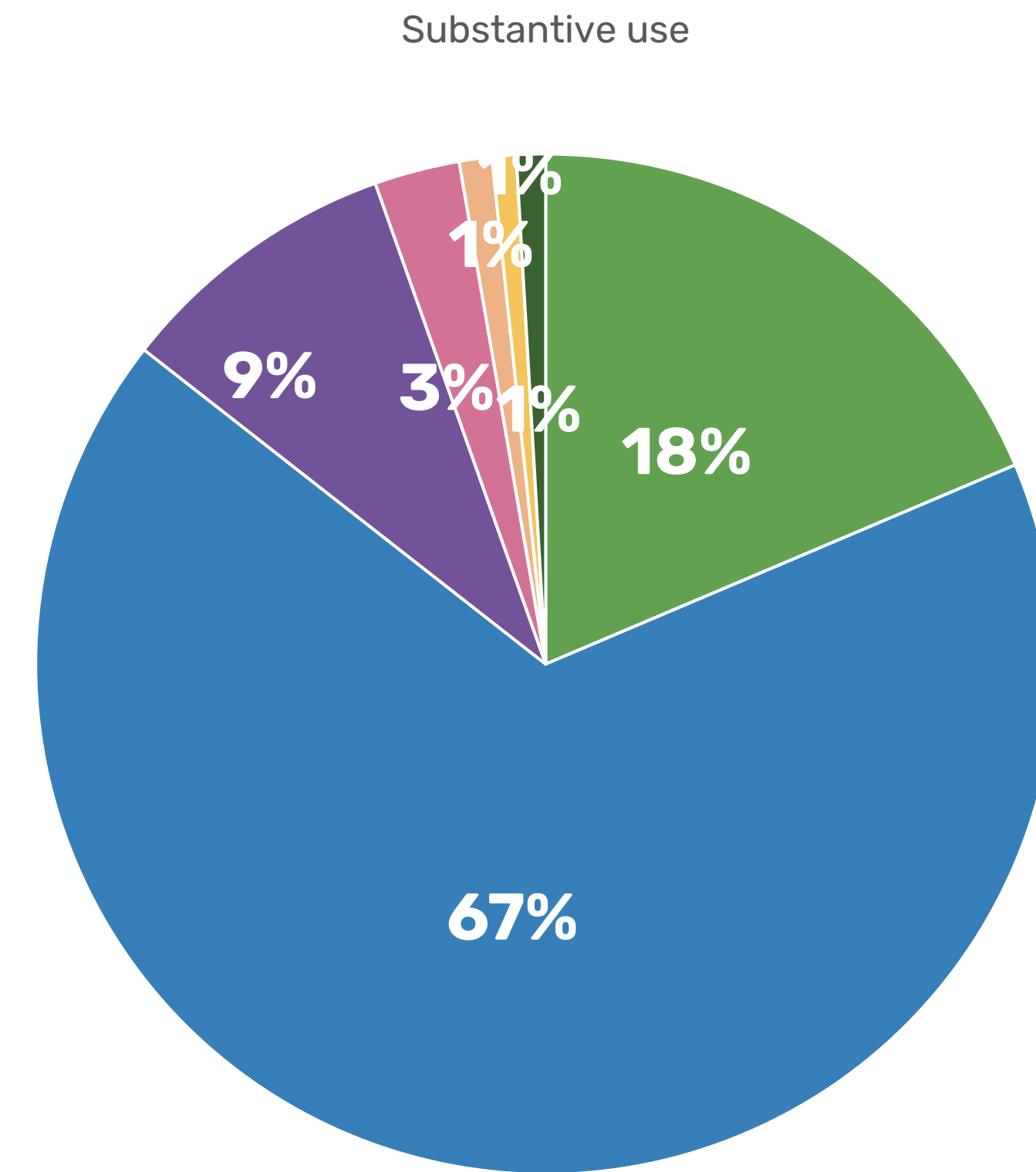
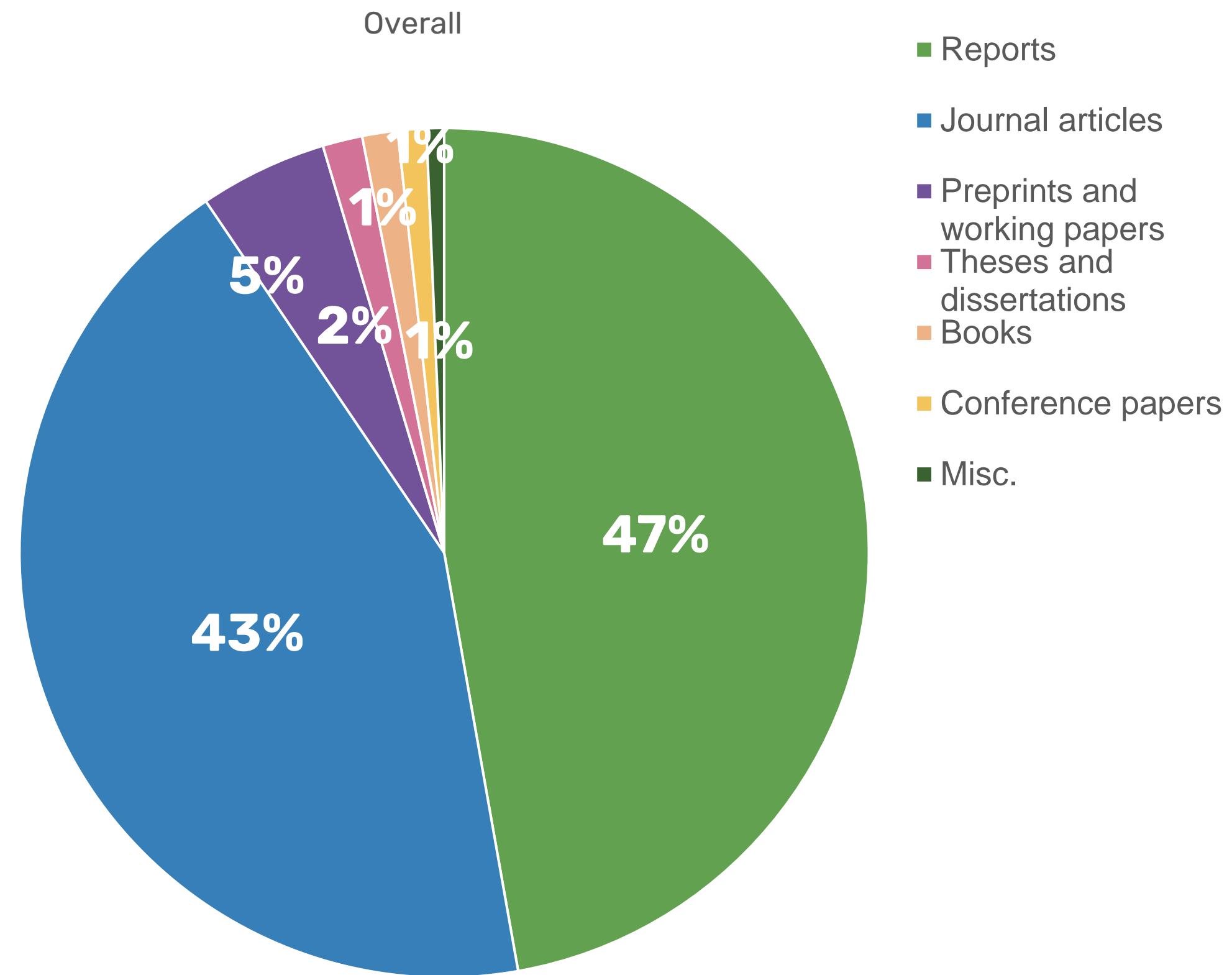
# Literature tracking in GBIF



# Why are papers mentioning GBIF?



# Types of literature tracked





Nº 11

# Science Review



 GBIF





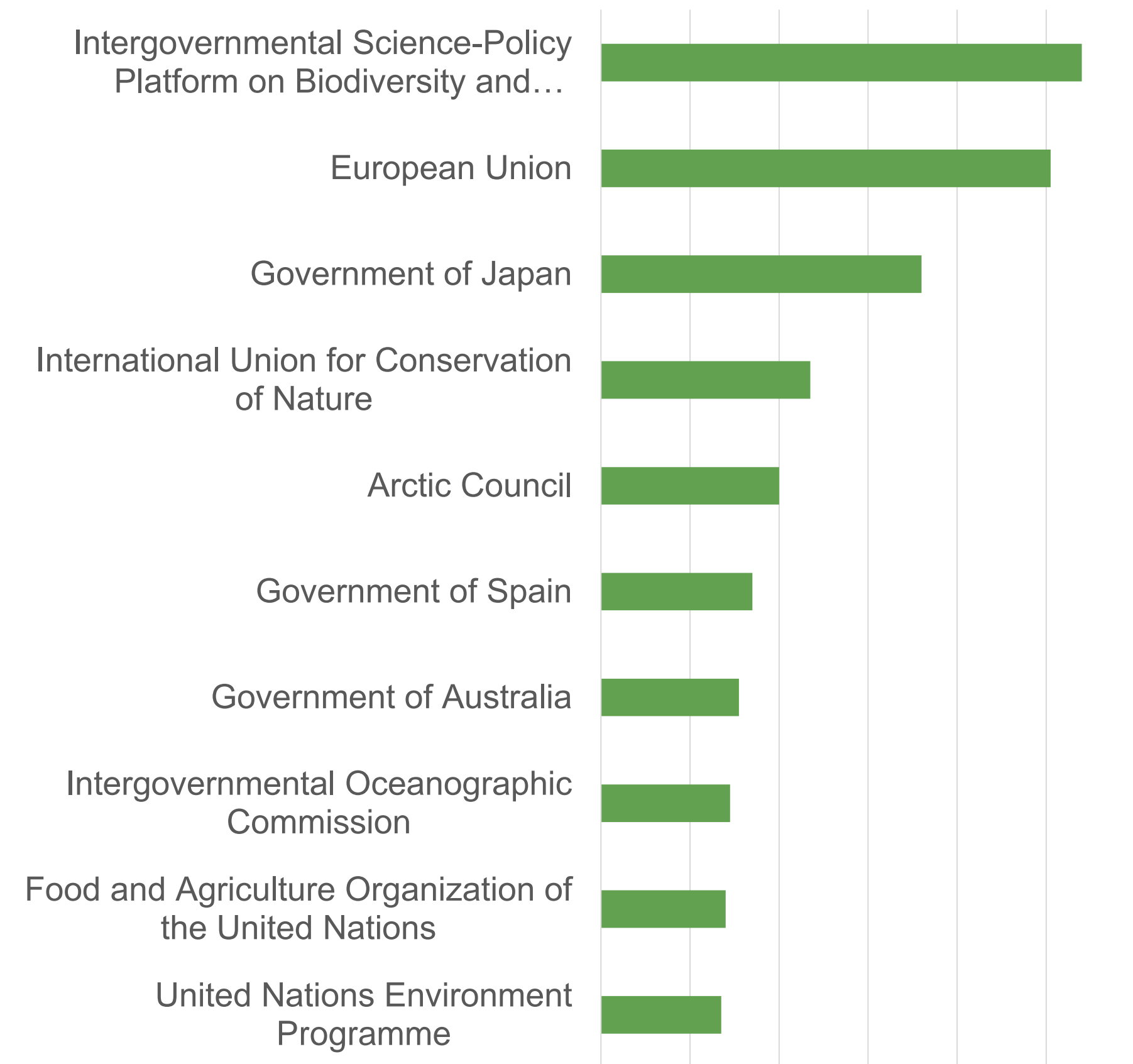
# GBIF and policy

More than just research articles

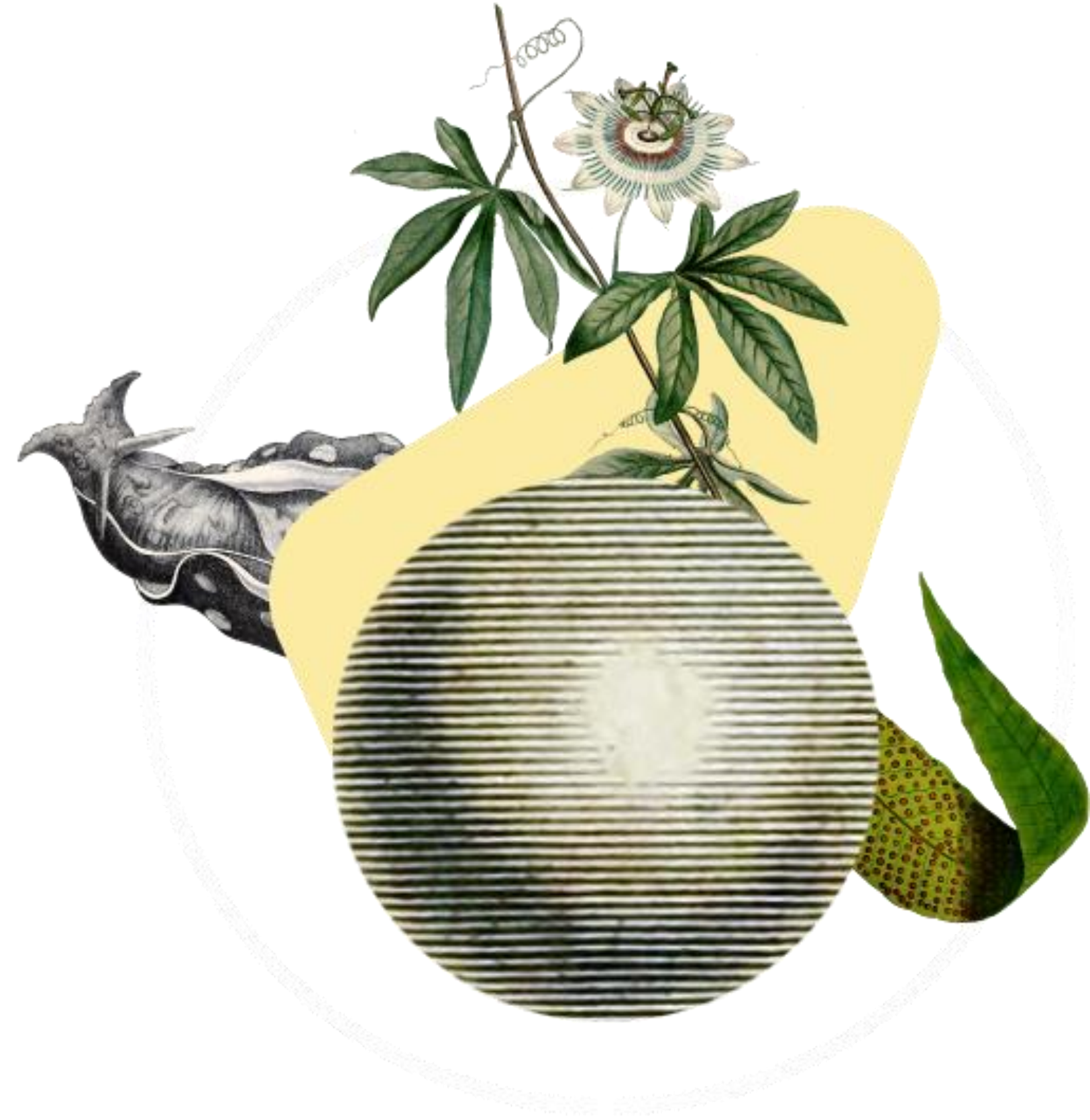
# Overton

The world's largest searchable index of policy documents, guidelines, think tank publications and working papers

- ~1,400 GBIF-relevant documents identified, published by 350 bodies including more than 100 national, regional and municipal governments
- Other top contributors include IPBES, IUCN, Arctic Council, IOC-UNESCO and FAO



## ~10 years of tracking data use and citations



- Manuscript in progress
- Process of tracking literature
- Compiling all the findings from the programme including new deep dives into download and citation practices, taxonomic and geographic focus and more
- Stay tuned!



Dmitry Schigel  
dschigel@gbif.org

@dschigel

