



**biodiversa+**

European Biodiversity Partnership

## Workshop on biodiversity monitoring data interoperability and harmonisation

1<sup>st</sup> of September – 9am to 12pm CEST

REC

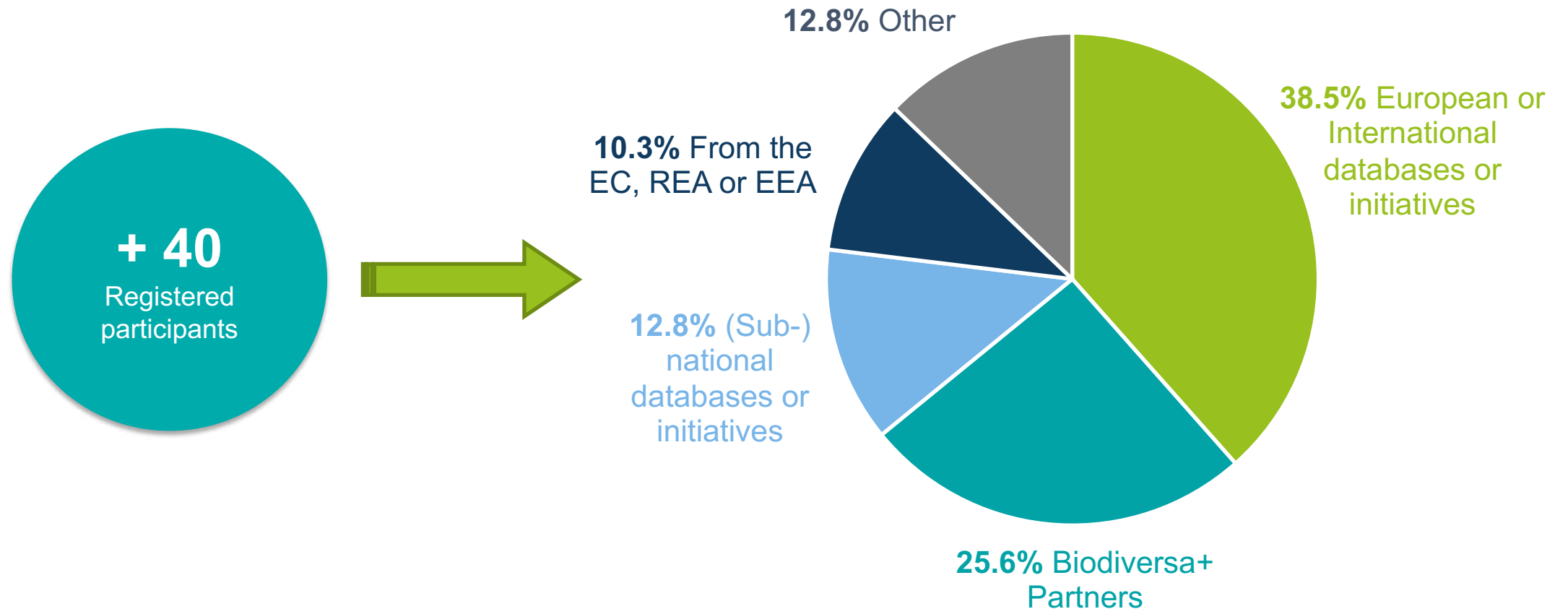
The plenary sessions of this meeting will be recorded and shared on the Biodiversa+ website and Youtube channel



# Welcome words and introduction of the meeting

By Alberto Basset, MUR

# Welcome to the workshop on biodiversity monitoring data interoperability and harmonisation



# Agenda of the workshop

- ✓ Introduction of the meeting and scene setting *by Alberto Basset, MUR and Hilde Eggermont, BelSPO, Biodiversa+ Chair and Coordinator*
- ✓ Data architecture and data flows in EuropaBON & EBVs, *by Dani Villero and Néstor Fernandez, EuropaBON*
- ✓ Presentation of data flows in GBIF and diversifying the GBIF model, *by Tim Hirsch, GBIF Secretariat*
- ✓ Presentation of Meta(data) structure and workflows in LifeWatch ERIC, *by Lucia Vaira, LifeWatch ERIC*
- ✓ DNA (meta)barcoding. BIOSCAN/iBOL Infrastructure developments, *by Rutger Vos, iBOL/BIOSCAN*
- ✓ Open discussion to identify barriers preventing data interoperability (i), factors encouraging data interoperability (ii) and possible support from Biodiversa+ (iii)
- ✓ Concluding words

# Aims of Biodiversa+ in this context

- ✓ Better understanding of data architecture & operating dataflow in European and international initiatives
- ✓ Exchange views with different actors about the level of harmonisation and interoperability the different databases
- ✓ Discuss how Biodiversa+ could help to harmonise databases and improve data interoperability, in order to support the overall objective of transnational biodiversity monitoring

## To be developed:

- ✓ In connection with the biodiversity monitoring priorities at the national, European, international level;
- ✓ In connection with the running activities in major international initiatives and projects;
- ✓ In connection with the running activities of the ESFRI landscape ERIs, mainly referring to the Biosphere domain.

# Harmonise databases and data interoperability

Considering that



Harmonisation



# Next steps after this workshop

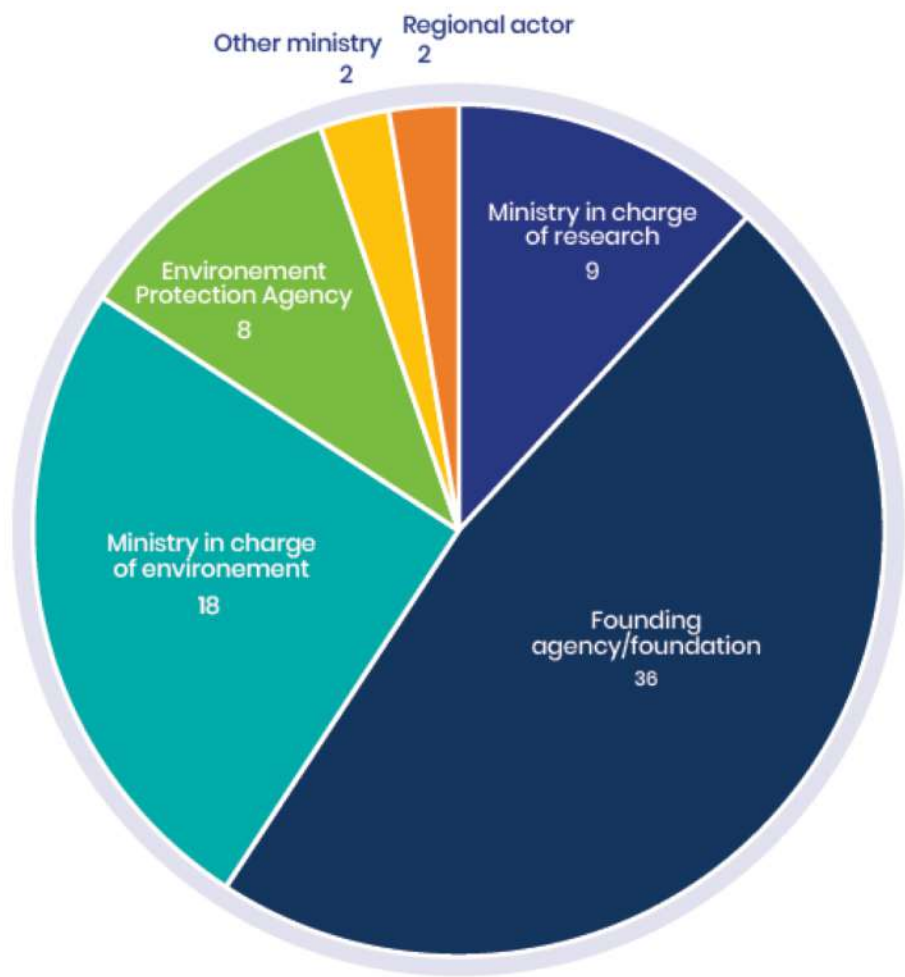
- ✓ Initial reflection of how to take up proposals from this WS in the Biodiversa+ workplan (possibly in collaboration with other key actors)
- ✓ Include the national perspective:
  - Showcases of (sub-)national data architecture plans and data workflows will be identified and mapped.
  - Possibly a follow-up workshop with presentations of (sub-)national showcases will be organised (in the coming months)
- ✓ A public report showcasing these national and global data architectures and workflows will be drawn. The objectives of this report will be to better understand the landscape (including its gaps), identify case studies enabling capacity building and knowledge transfer, identify good practices.
- Building on relevant EuropaBON activities

# Presentation of Biodiversa+

By Hilde Eggermont, BelSPO, Biodiversa+ Chair & Coordinator



# Biodiversa+ membership



# Biodiversa+ portfolio of activities and budget amplitude

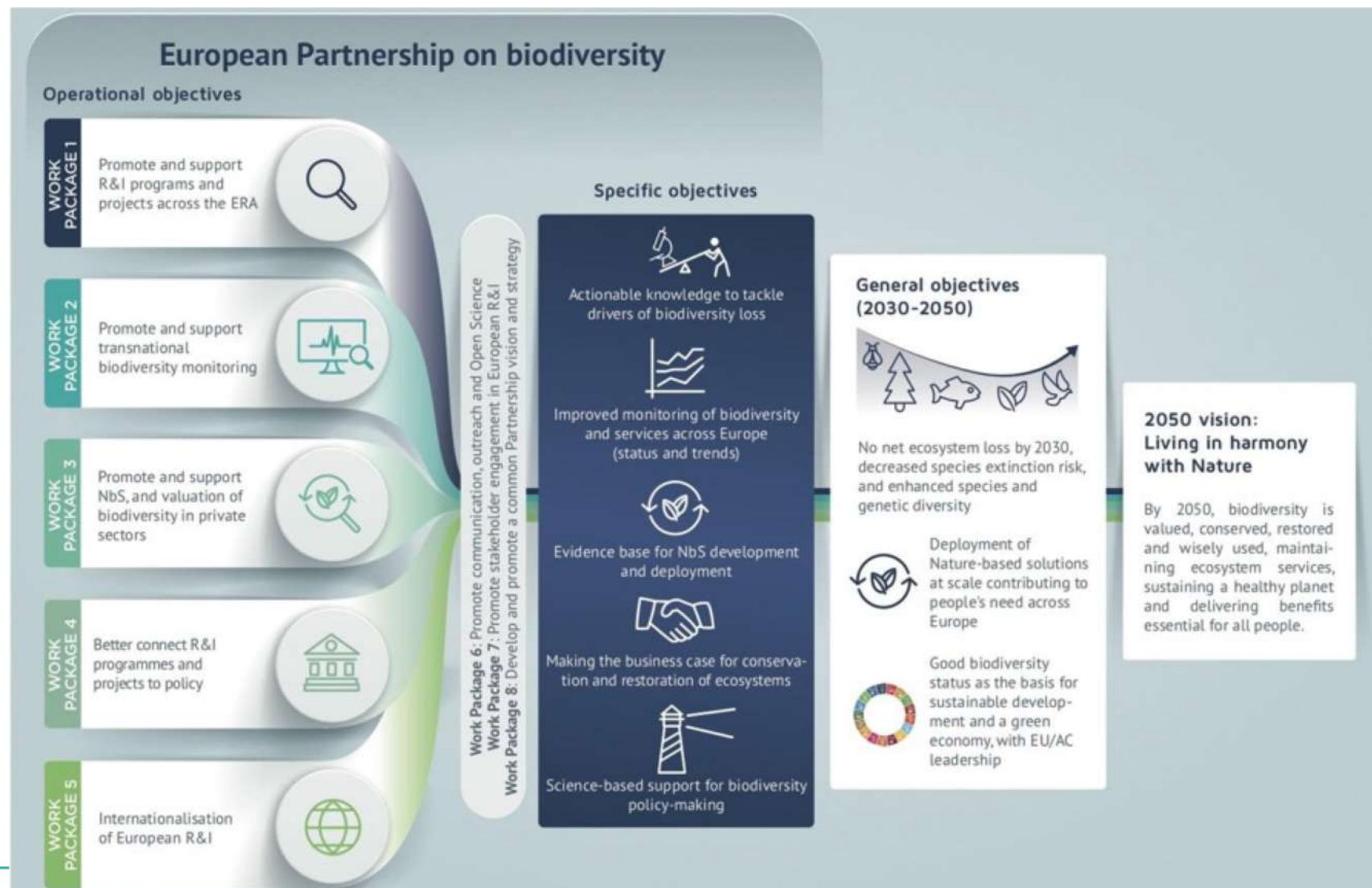
Budget of >800 Mio € over 7 yrs,  
combining in-cash and in-kind  
resources from its Partners and  
including 165 Mio € by the  
European Commission



# Aligned with the policy context

## EU Biodiversity Strategy 2030:

*“... making the bridge between science, policy & practice...”*



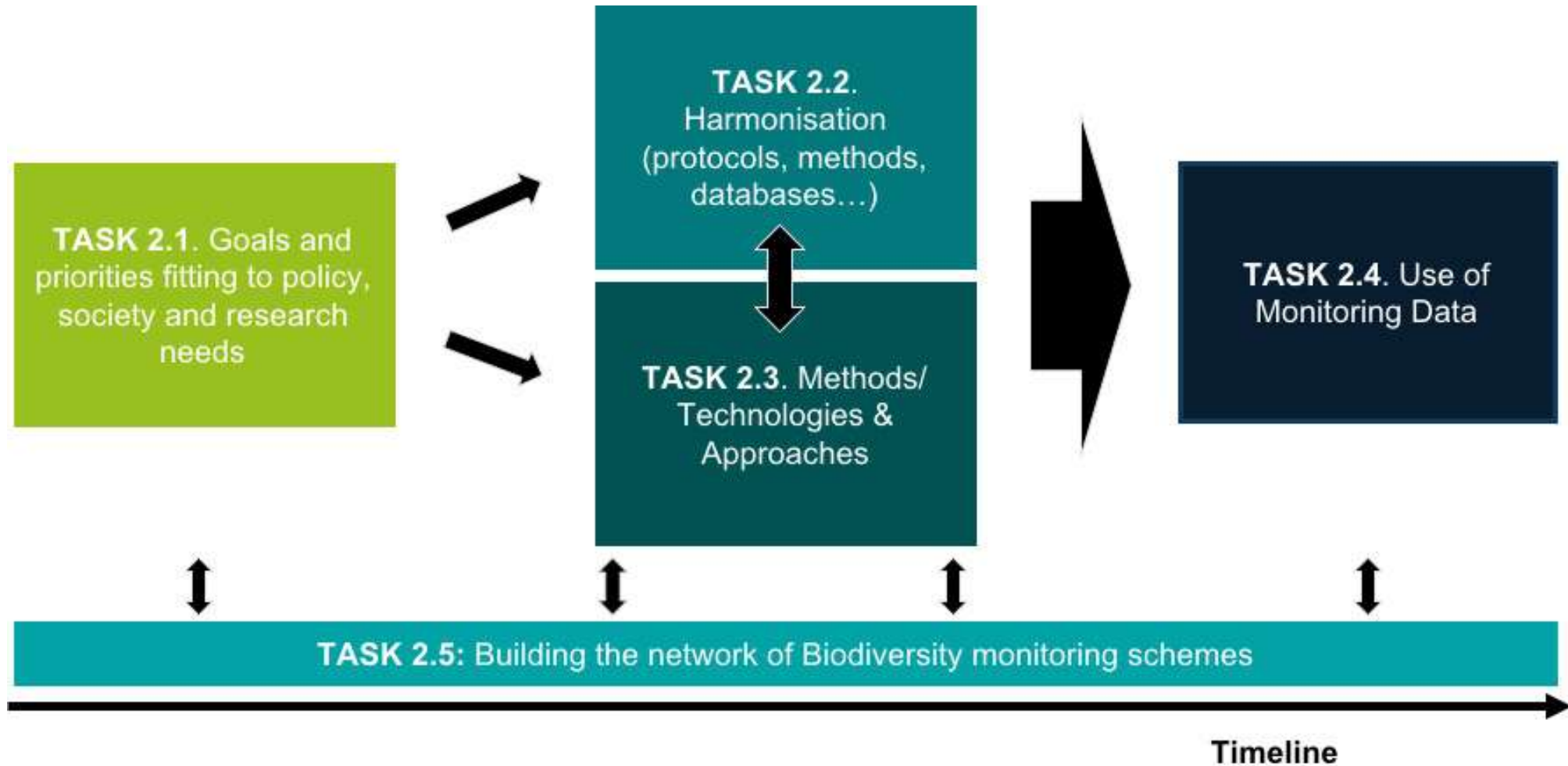


# Building on existing and new initiatives



... and many more!

# Biodiversity monitoring workstreams of activities



# Work on data architecture and data workflows in EuropaBON and EBVs

By Dani Villero and Nestor Fernandez, EuropaBON

# Mapping and making current European monitoring initiatives accessible via a web-based database.

Alejandra Morán-Ordóñez

David Martí Pino, Dani Villero, Sergi Herrando, Lluís Brotons



**CREAF**



EXCEL·L  
SEVERO  
OCHOA

**EBCC**  
European Bird Census Council



every bird counts



This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101007492.

**24/08/2022 ECCB Conference, Prague**





# Why is important to map monitoring efforts & workflows?

## Biodiversity monitoring capability across Europe

- what taxa are currently (well) monitored?
- what is the geographic coverage of biodiversity data?
- what is the nature of monitoring programs?
- is the current monitoring capability temporally sustainable?
- cost-efficiency of current monitoring schemes



# Why is important to map monitoring efforts & workflows?

## Biodiversity monitoring capability across Europe

- what taxa are currently (well) monitored? **Gaps**
- what is the geographic coverage of biodiversity data? **Gaps**
- what is the nature of monitoring programs? **Bottlenecks**
- is the current monitoring capability temporally sustainable? **Bottlenecks**
- cost-efficiency of current monitoring schemes **Bottlenecks**

Co-Design of EBVs, EESVs to track goals of (e.g.): EU restoration law, Biodiversity Strategy, etc.

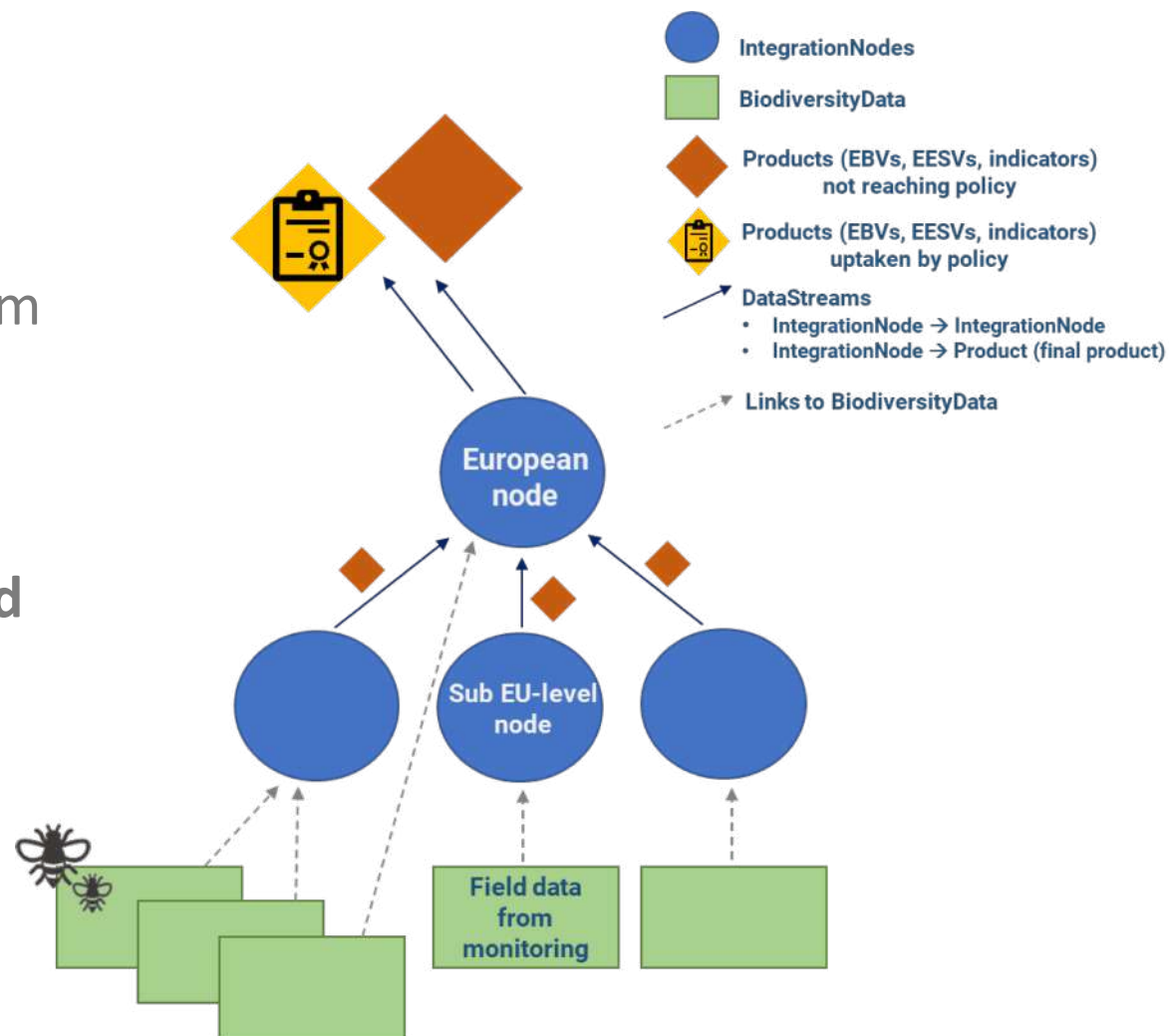


# EuropaBON: Inventory of monitoring initiatives

Initiatives/projects/programs – NETWORKS - that **integrate biodiversity data** from monitoring schemes and process them to **produce** aggregated biodiversity data or Essential Biodiversity Variables (**EBVs**), Essential Ecosystem Services Variables (**EESVs**) and/or other European policy-relevant **indicators**.

Priority on **monitoring efforts that are coordinated at a transnational level** (combining bottom up and top down approach).

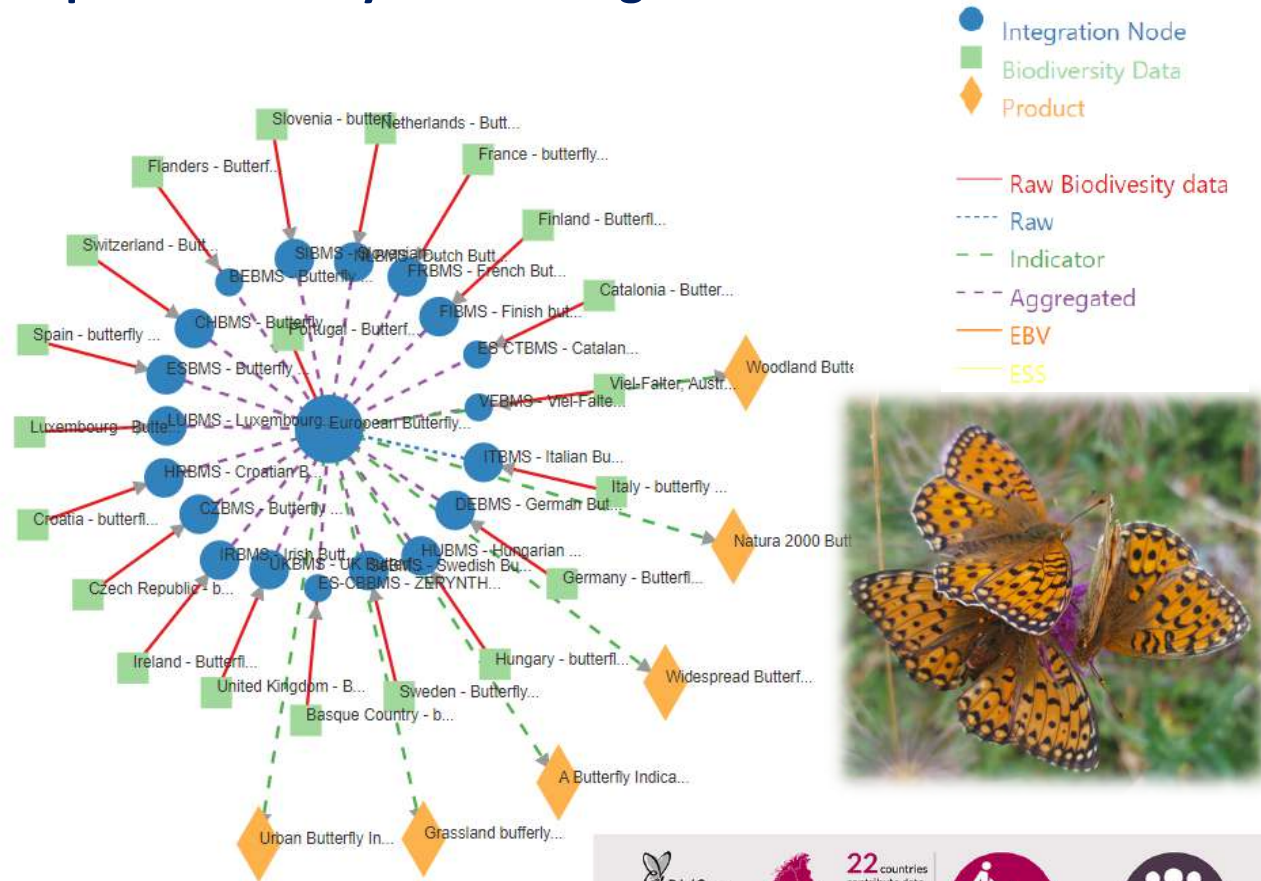
Focus on **data workflows** not only on data collection.



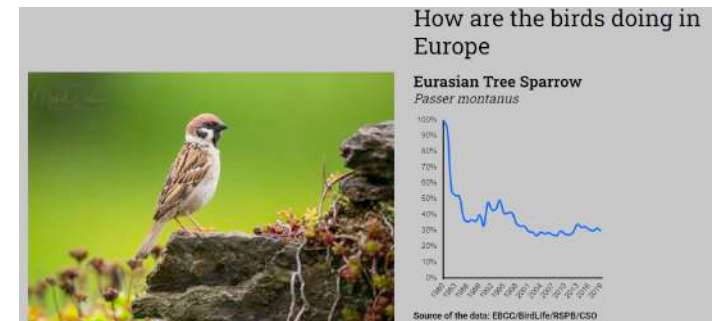
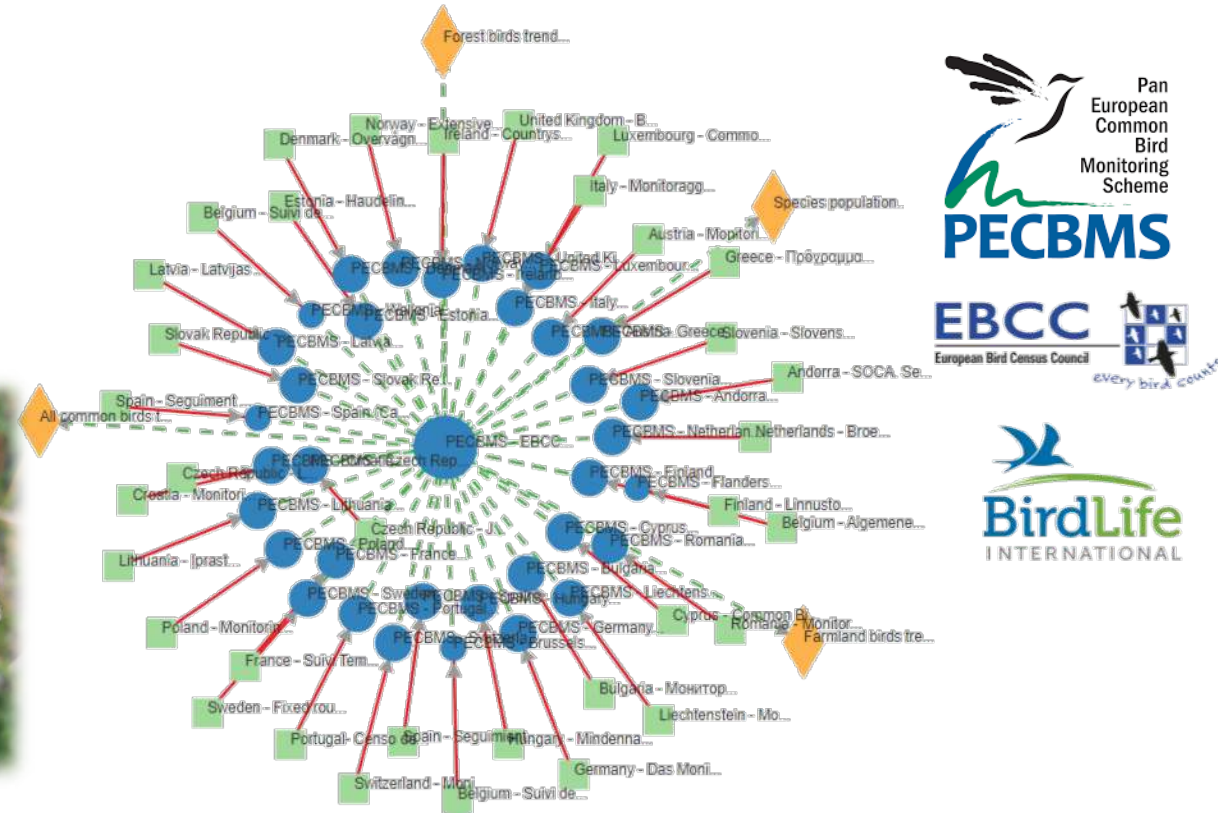


# Integration initiatives: examples

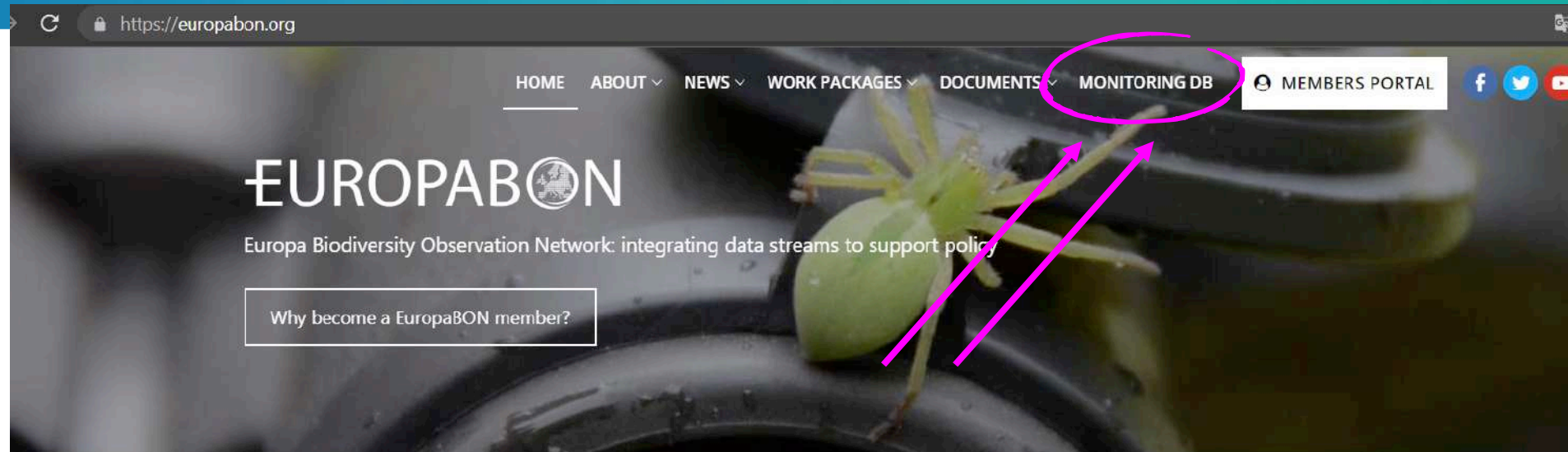
## European Butterfly Monitoring Scheme



## PanEuropean Common Bird Monitoring Scheme



# Web-based database (open): access via europabon.org



Latest news



**New EuropaBON-supported paper investigates the updates in the global mining land use**  
*August 5, 2022*

A recently published paper, studying the current updates in global mining land use is now available in the Scientific Data Journal. The study has been conducted by a team of leading international experts, amongst whom is EuropaBON's member Ian McCallum (International Institute for Applied Systems Analysis, IIASA). The growing demand for minerals has pushed mining activities...



**The June issue of EuropaBON's newsletter has just been published!**  
*June 8, 2022*

Tweets by @EuropaBon\_H2020

**EuropaBON**  
@EuropaBon\_H2020

Would you like to join our ever-growing and ambitious EuropaBON team? Go check out the open @idiv student positions and become a part of the #biodiversity monitoring experience that EuropaBON is! Find more info in the post below!  
<https://twitter.com/idiv/status/1554730319663726592>

Aug 4, 2022

EuropaBON Retweeted

**REST-COAST**  
@RESTCOAST\_H2020

REST-COAST is attending the ECSA59 symposium "Using the best scientific



# Web-based database (open)

Website open: anyone can query content (permissions required for data input)

## EUROPABON

[Admin panel](#)[Alejandra Moran](#)[Log out](#)

### Main tables

- Home
- Integration nodes**
- BiodiversityData
- All Data streams
  - IntegrationNode → IntegrationNode
  - Products
- Users in projects
- Institutions
- Thesaurus

### Search and visualize

- Visualizations
- Search
- Geo. coverage

### Stats

- Global stats

### Metadata

### Schematic diagram of the dataflow across an integration initiative

**IntegrationNodes**

Institutions/projects/platforms integrating/processing biodiversity data to generate Essential Biodiversity Variables (EBVs), Ecosystem Services Indicators (ESS) or any other indicators with potential relevance to environmental policy, particularly at the European level. One single integration initiative can be composed of many nodes operating at different scales (e.g. subnational, national, European-level) in a coordinated manner. One single institution can act as an integration node of different integration initiatives.

[Integration node list](#)

**BiodiversityData**

Biodiversity monitoring initiatives responsible for the collection of biodiversity information (data from field observations). The focus of WP3.1 is not to map all monitoring initiatives in Europe, but to ensure that those monitoring initiatives providing data for current biodiversity workflows are well represented in the database.

[BiodiversityData list](#)

**Datastreams**

These represent data flows and connect IntegrationNodes across different scales. These contain two types of information: information related to the data flowing between nodes (dataset) and whether this has been integrated ("products": EBVs, ESS, indicators) or not (raw and aggregated data) (dataprocess).

If the data has been integrated/processed, it also contains information about the integration method. A data stream that doesn't connect with another integration node represents a final Product (EBVs, ESS, Indicator) that can eventually be uptaken by policy or not (this is represented in Figure 1 with diamonds of different colours)

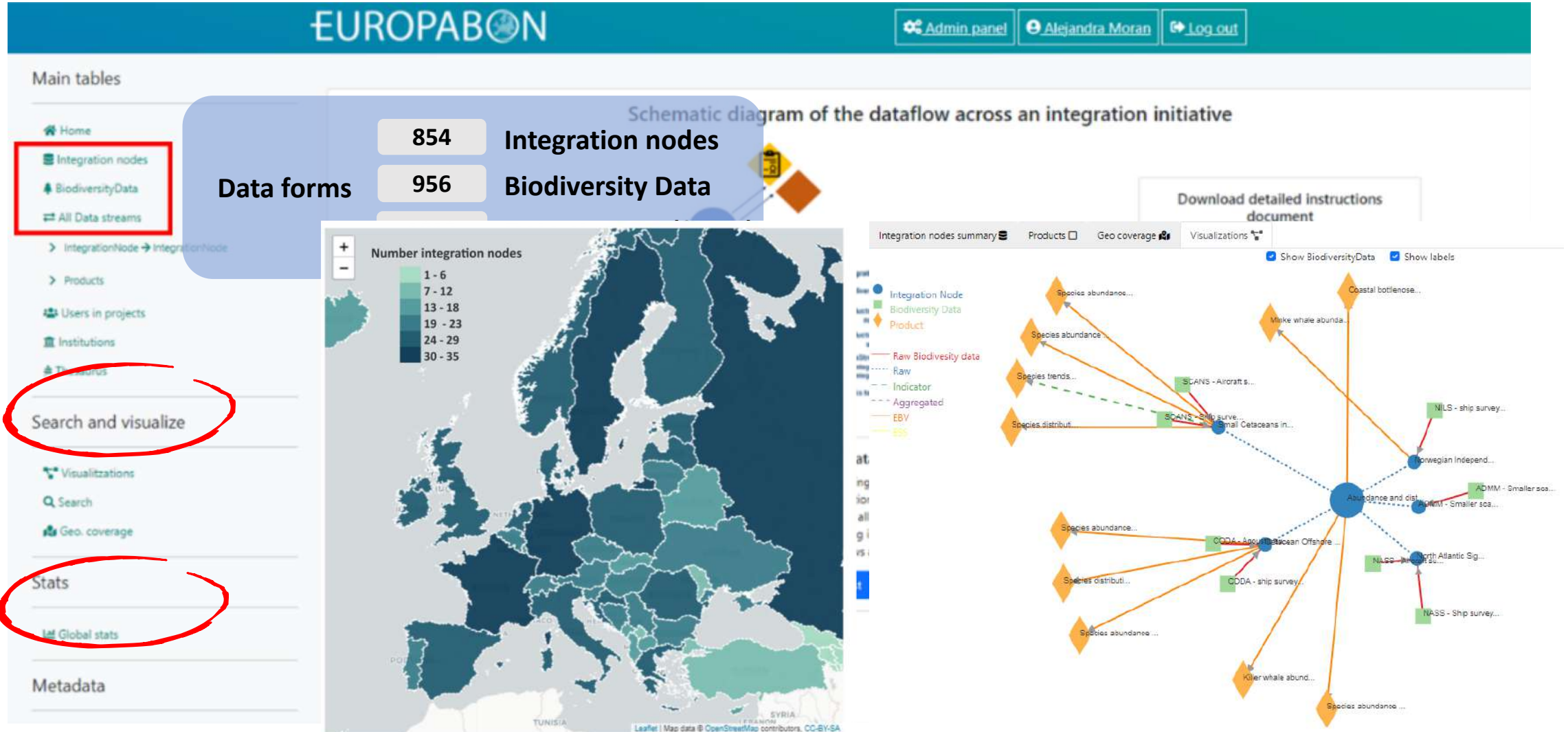
[DataStreams list](#)

[Download detailed instructions document](#)

[Instructions\\_WP3\\_task3.1](#)

# Web-based database (open)

Website open: anyone can query content (permissions required for data input)

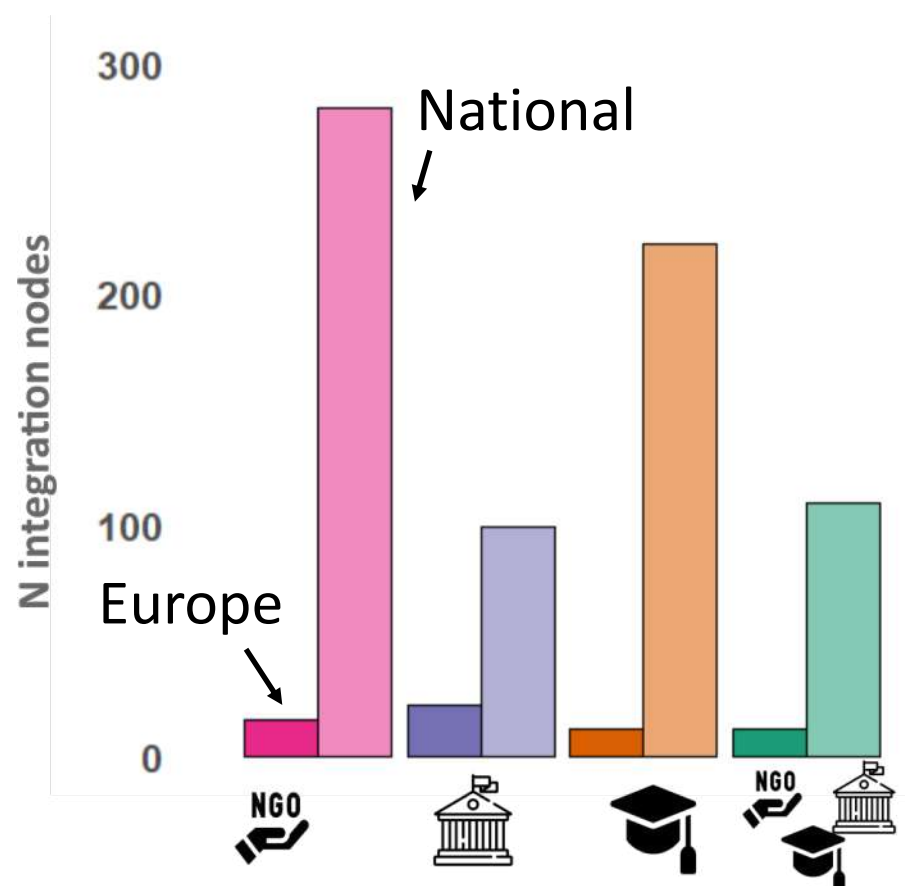





# Results: integration nodes

## Who is integrating the data?

NGOs process and integrate most biodiversity data across Europe (38% of National integration nodes), followed by research institutions (35%) and governmental institutions (19,8%)



 **NGOs** (e.g.): IUCN, Birdlife International, Wetlands International, European Mammal Foundation, Ornithological Societies

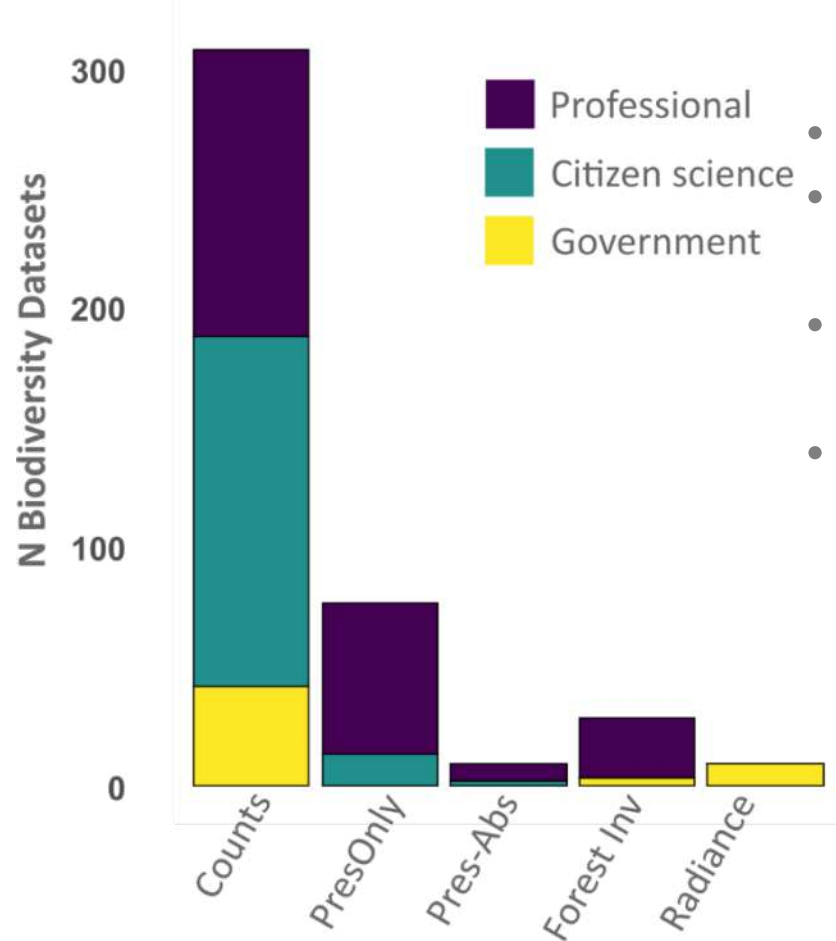
 **Supra-/Governmental** (e.g.): Joint Research Center (EC), Croatian Agency for Environment and Nature, Swedish Species Information Centre (SSIC), Centre Suisse de Cartographie de la Faune (CSCF)

 **Research bodies** (e.g.): Research Institute for Nature and Forest (INBO, BE), Institute of Zoology (Academy of Sciences of Moldova), Sapienza University of Rome (IT), University of Aarhus (DK), Museu de Ciències Naturals de Granollers (ES)

*Number of integration nodes coordinated by NGOs, Governmental or supra-governmental institutions, Research bodies (including Universities) or by a combination of these (EuropaBON database numbers by end November 2021). In each group, the left column corresponds to European nodes and the right column to National nodes.*

# Results: biodiversity data

## Origins and data types of biodiversity monitoring data

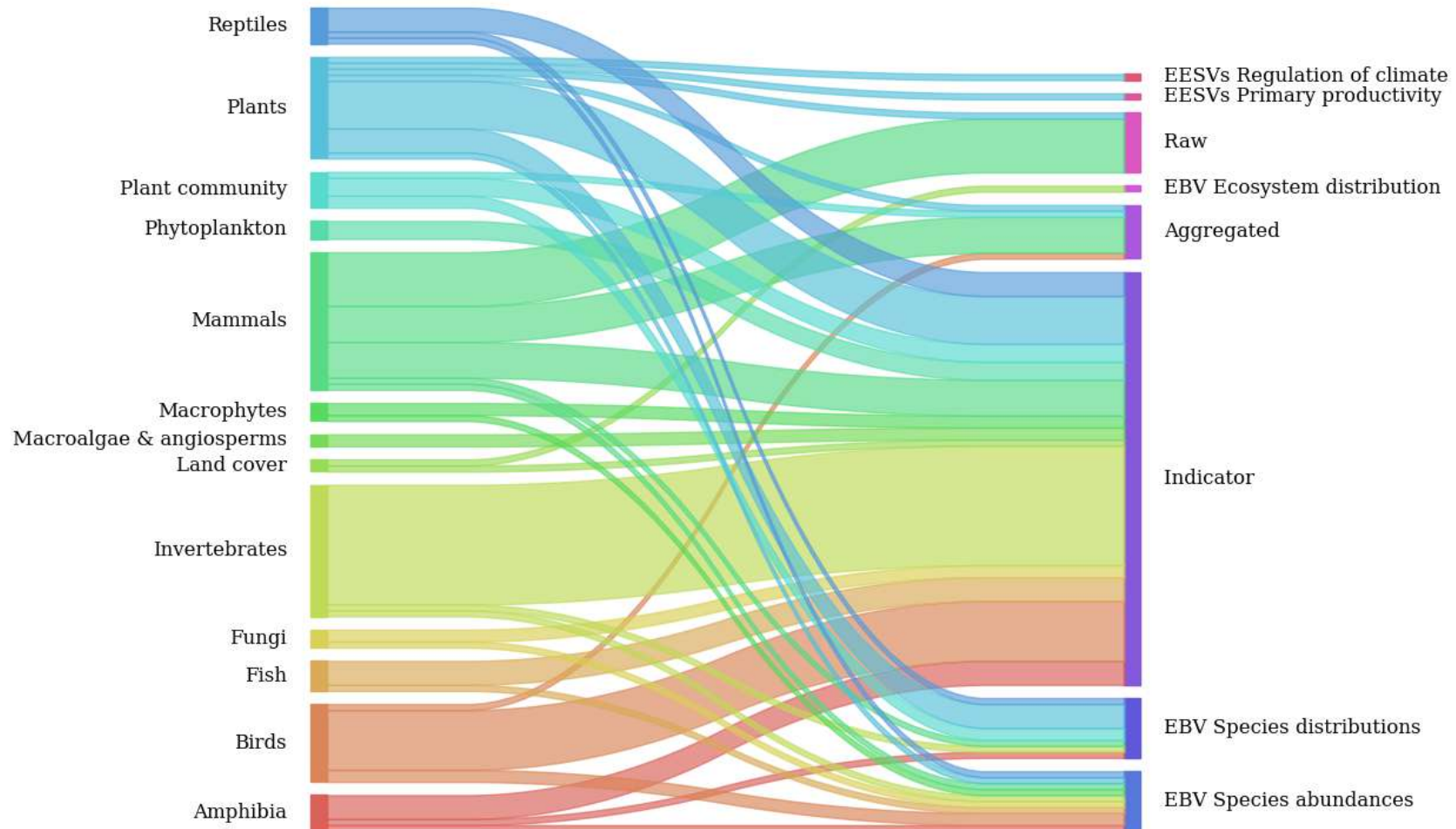


- **62 % corresponded to systematic monitoring programs**
- **Count data** for abundance estimation were the most common type of data collected (**80% of data**)
- The purpose of most biodiversity and environmental data sources gathered is **to advance ecological knowledge of species**
- **Only** a small fraction of all biodiversity and environmental data sources collected are **open access (10.6%)**

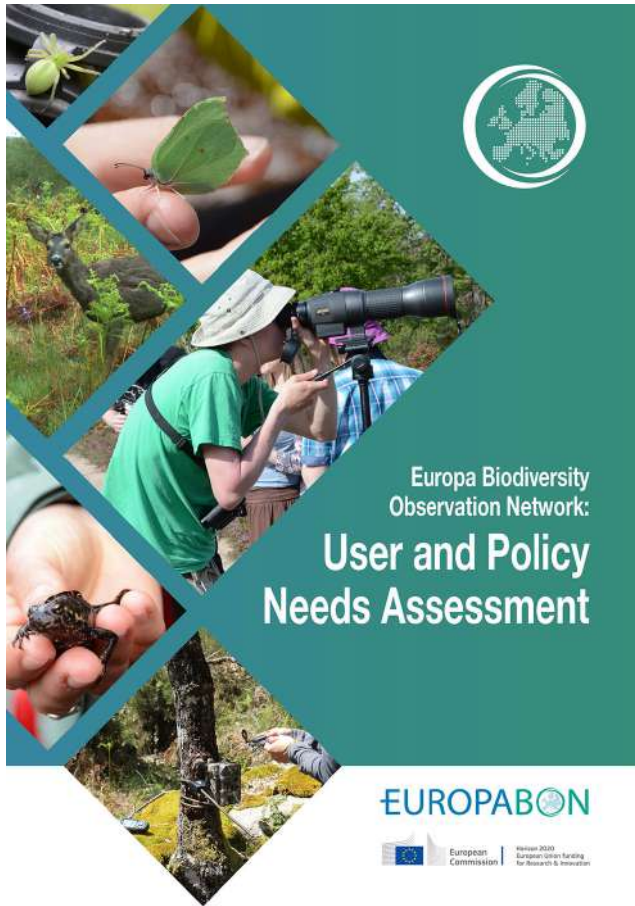
*Number of Biodiversity Datasets categorized by their data origin (citizen science, governmental authorities, professional monitoring) and data type (counts, presence-only, Presence-Absence, Forest Inventory data and spectral radiance - remote sensing)*

# Data flows: European level

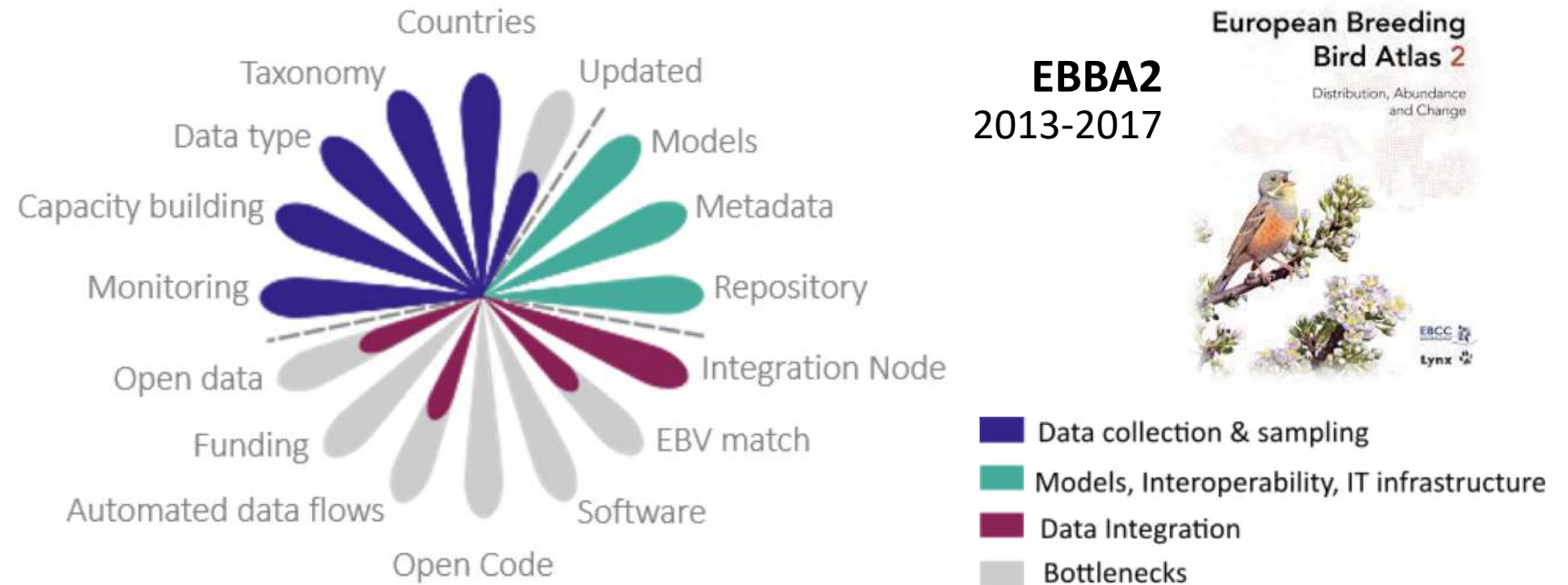
Only around 22 % of data flows have been documented to generate EBVs and EESVs: species distributions 55%, species abundance 17%, measures of taxonomic/phylogenetic diversity 7% and EESVs <1%



# Next developments: identification of gaps & bottlenecks



## EBV: species distributions all terrestrial birds

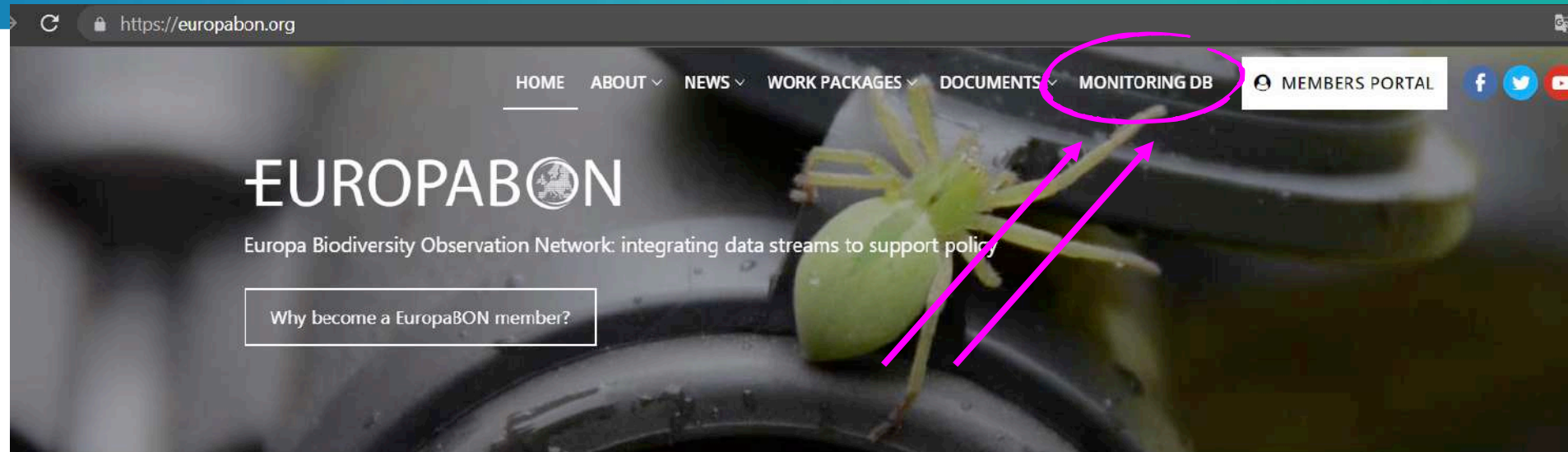


Thanks for your attention!

<https://monitoring.europabon.org/>  
[d.villero@creaf.uab.cat](mailto:d.villero@creaf.uab.cat)



# Web-based database (open): access via europabon.org



### Latest news



#### New EuropaBON-supported paper investigates the updates in the global mining land use

August 5, 2022

A recently published paper, studying the current updates in global mining land use is now available in the Scientific Data Journal. The study has been conducted by a team of leading international experts, amongst whom is EuropaBON's member Ian McCallum (International Institute for Applied Systems Analysis, IIASA). The growing demand for minerals has pushed mining activities...



#### The June issue of EuropaBON's newsletter has just been published!

June 4, 2022

### Tweets by @EuropaBon\_H2020

**EuropaBON** @EuropaBon\_H2020  
Would you like to join our ever-growing and ambitious EuropaBON team? Go check out the open @idiv student positions and become a part of the #biodiversity monitoring experience that EuropaBON is! Find more info in the post below! <https://twitter.com/idiv/status/1554730319663726592>

Aug 4, 2022

EuropaBON Retweeted

**REST-COAST** @RESTCOAST\_H2020  
REST-COAST is attending the ECSA59 symposium "Using the best scientific



**iDiv**

German Centre for Integrative Biodiversity Research (iDiv)  
Halle-Jena-Leipzig

# Mainstreaming Essential Biodiversity Variables:

A Data Portal and a standard for spatiotemporal biodiversity data

*Néstor Fernández, Christian Langer, Luise Quoss, Miguel Fernández, José Valdez, Henrique M. Pereira*

[portal.geobon.org](https://portal.geobon.org)



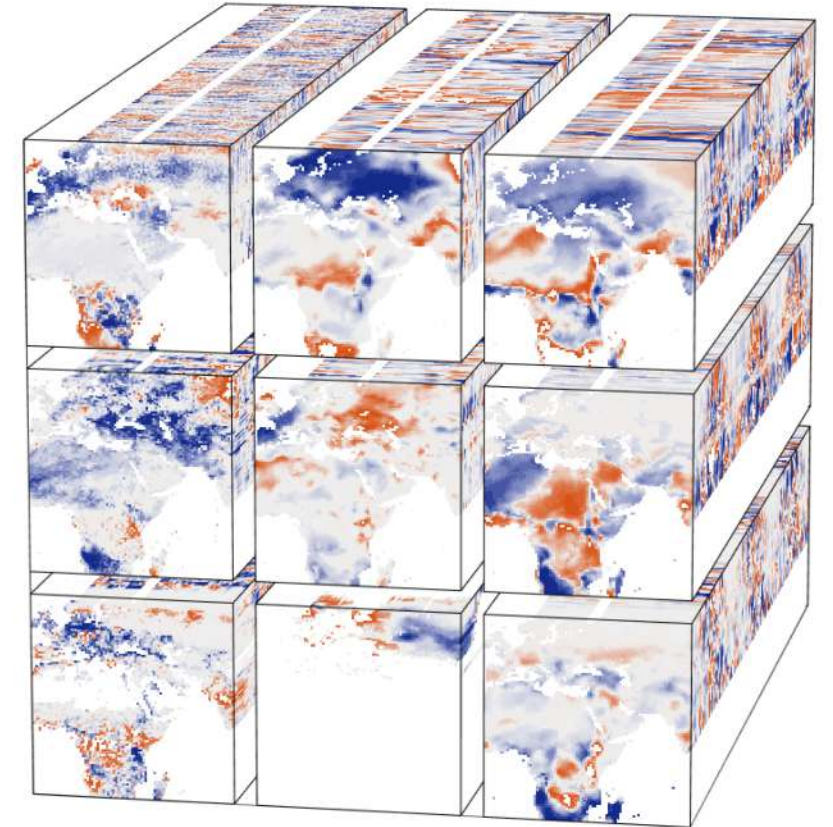
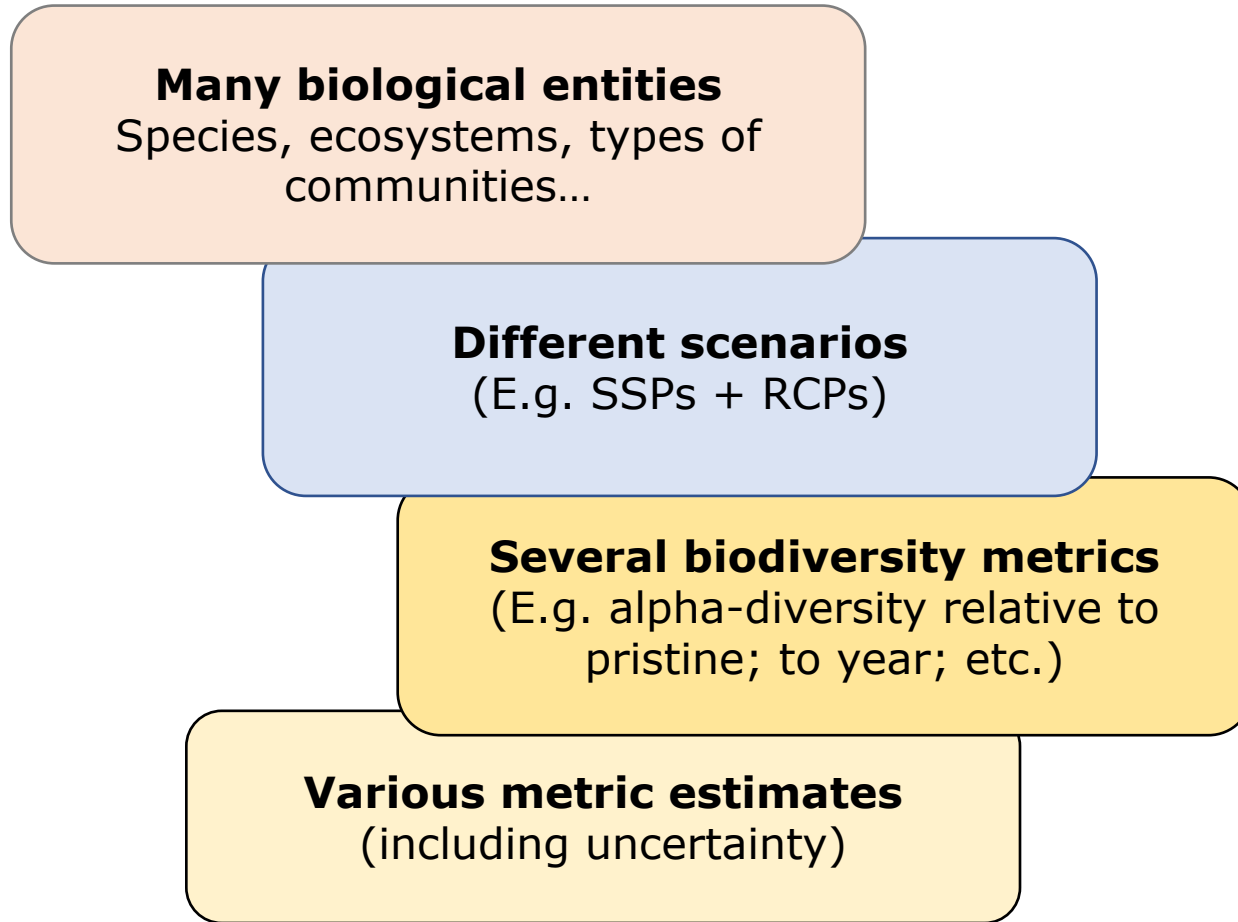
Group on Earth Observations  
Biodiversity Observation Network



iDiv is a research centre of the  
**DFG** Deutsche  
Forschungsgemeinschaft

# Mainstreaming Essential Biodiversity Variables

Biodiversity datasets are multidimensional but no standard exists for their organization

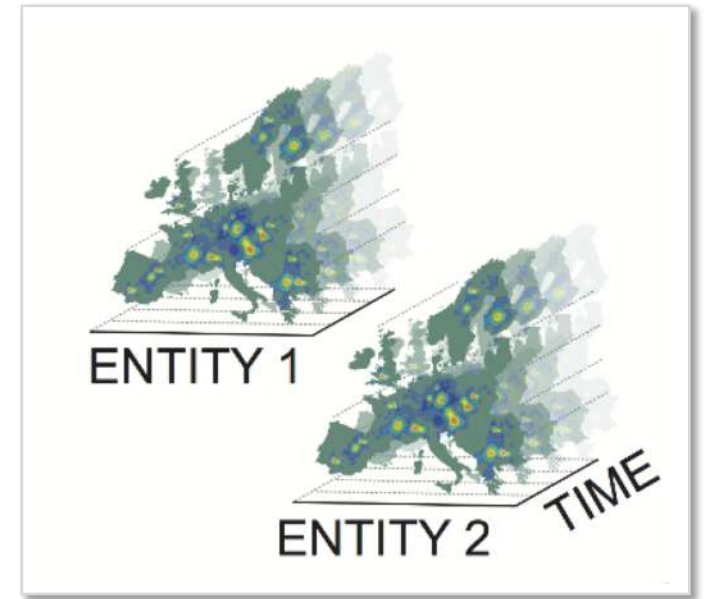
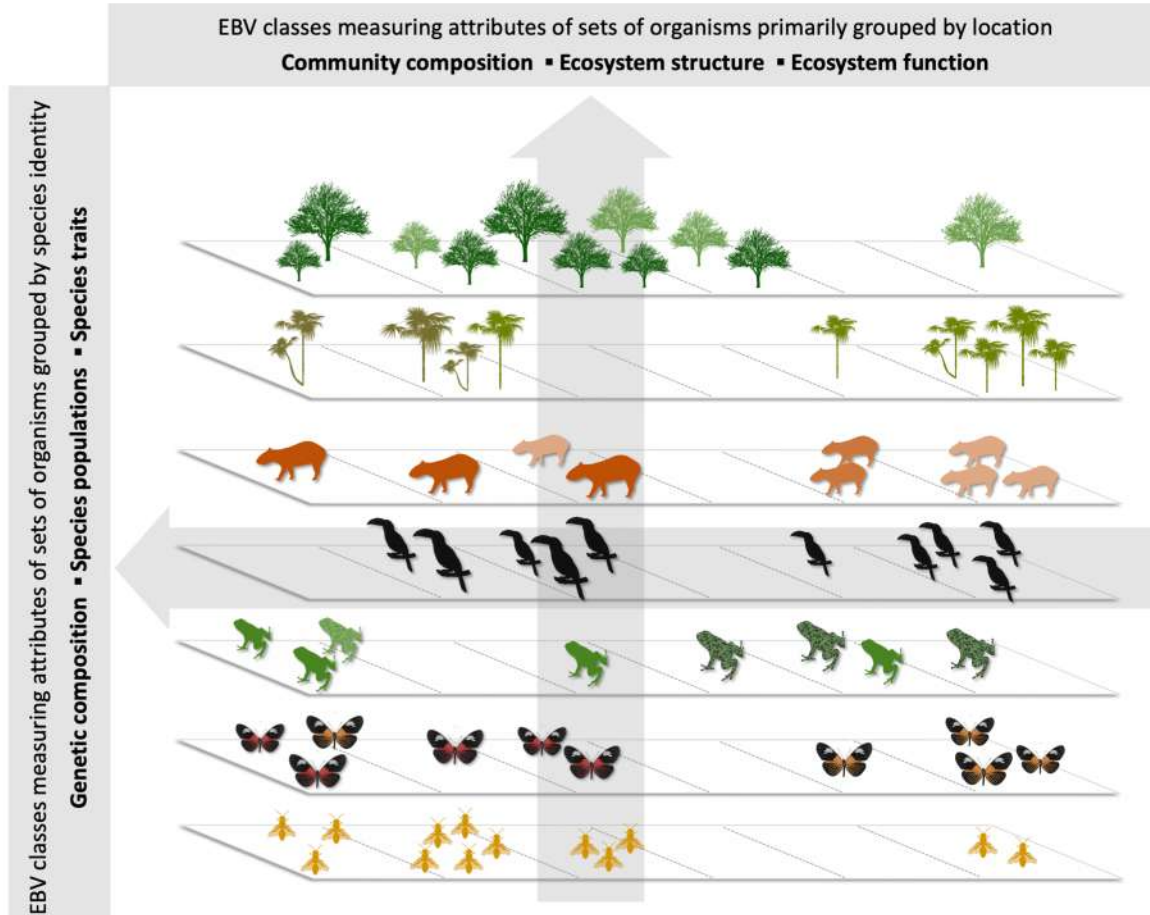


*Picture produced by the Earth System Data Lab*  
[www.earthsystemdatalab.net](http://www.earthsystemdatalab.net)



# Mainstreaming Essential Biodiversity Variables

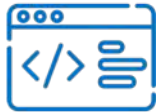
Comparable data outputs across EBV classes



# Mainstreaming Essential Biodiversity Variables

Comparable data outputs across EBV classes

How should we organize such disparity of data?



**Consistent metadata**



**Interoperable**



**Easy access**



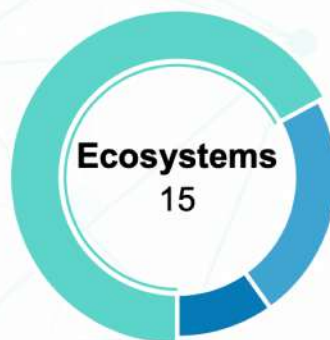
portal.geobon.org

# EBV Data Portal

The EBV Data Portal includes a variety of EBV raster datasets.  
You can import these datasets into the map with a single click. You can also upload your own EBV dataset for sharing with others.

## Portal statistics

Entity Type



Spatial Domain



Environmental Domain



25 EBV datasets in total

Show All ▾ datasets

Search datasets:

portal.geobon.org

# EBV Data Portal

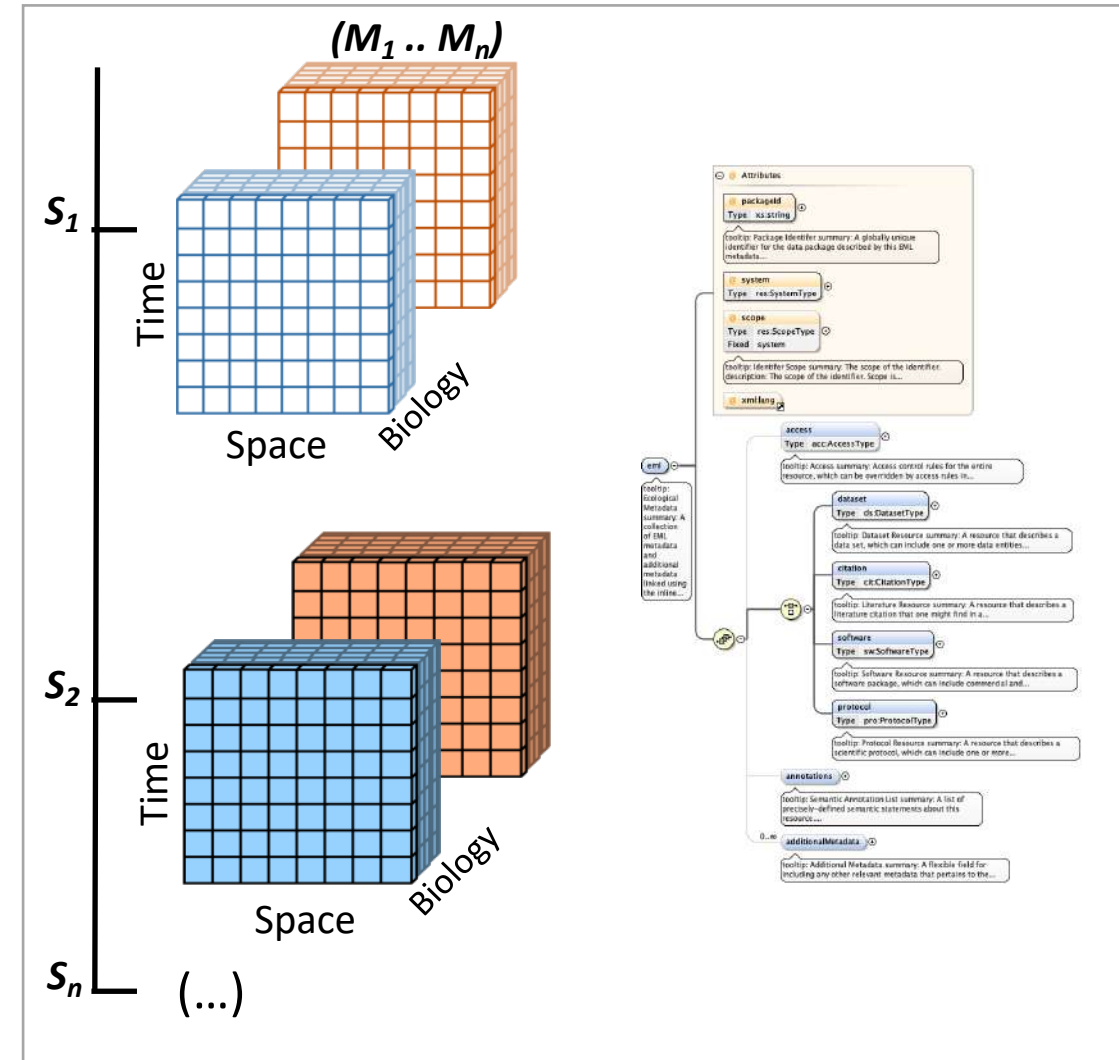
The EBV Data Portal includes a variety of EBV raster datasets.  
You can import these datasets into the map with a single click. You can also upload your own EBV dataset for sharing with others.

## Datasets:

- ❖ Consistently structured across **thematic, spatial and temporal dimensions**
- ❖ **Consistently documented** in a way that maximizes usability
- ❖ **Traceable** (both resources and production pipelines)

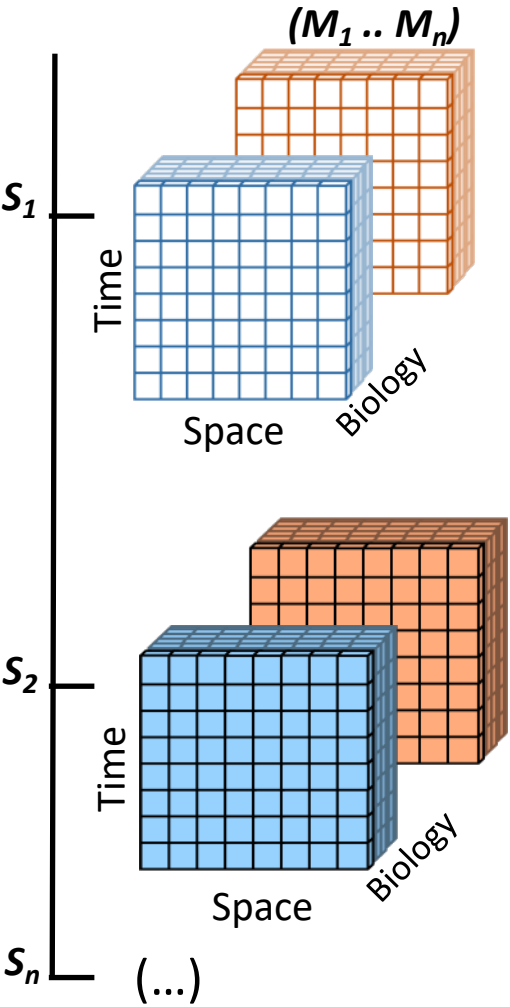
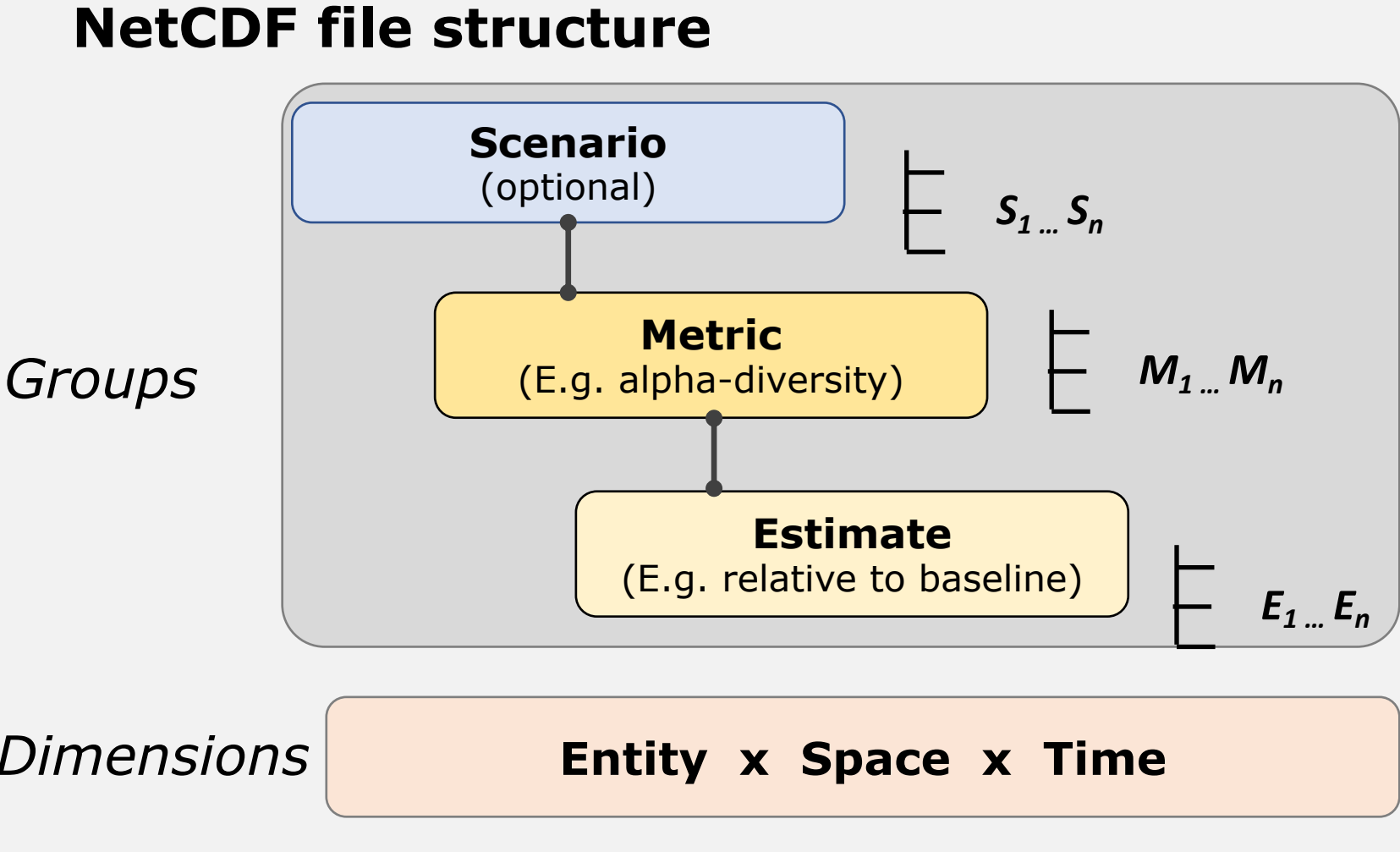
# Components of the EBV Cube Standard

- **NetCDF** specification with a **unified hierarchical structure** for organizing EBV data
- A **minimum information** specification using **ACDD** terms (with translation into **EML**)
  - *Self-described*
  - *Compliant with FAIR and GEOSS-DMP*





# Components of the EBV Cube Standard



# Components of the EBV Cube Standard

## Embedded metadata

### ACDD 1.3

*Attribute Convention for Data Discovery*

- 1 Attribution
- 2 Publication
- 3 History and processing level
- 4 Geospatial and temporal attributes
- 5 Data, processing and models

### EBV terms

- 6 EBV class, name, entities
- 7 Scenario and metric descriptions

```
..
"source": "sfsfsf",
▼ "coverage_content_type": [
  "referenceInformation"
],
"processing_level": "N/A",
"project": "N/A",
"project_url": "N/A",
▼ "creator": {
  "creator_name": "Abel Ramoelo",
  "creator_email": "N/A",
  "creator_institution": "Aarhus University",
  "creator_country": "Denmark"
},
"contributor_name": "N/A",
"license": "https://creativecommons.org/licenses/by-nc-sa/4.
▼ "publisher": {
  "publisher_name": "Christian Langer",
  "publisher_email": "christian.langer@idiv.de",
  "publisher_institution": "German Centre for Integrative E
  "publisher_country": "Germany"
},
▼ "ebv": {
  "ebv_class": "Species populations",
  "ebv_name": "Species distributions"
},
▼ "ebv_entity": {
  "ebv_entity_type": "Communities",
  "ebv_entity_scope": "sfsfsfsf",
  "ebv_entity_classification_name": "N/A",
  "ebv_entity_classification_url": "N/A"
},
▼ "ebv_metric": {
  ▼ "ebv_metric_1": {
    ":standard_name": "fsfsf",
    ":long_name": "sfsfsf",
    ":units": "sfsfsf"
  }
}
```



# Components of the EBV Cube Standard

## *Interoperable information standards*



### **Attribute Convention for Data Discovery**

❖ Standard for earth sciences data



*Ecological  
Metadata Language*

❖ LTER, NEON, DataOne...

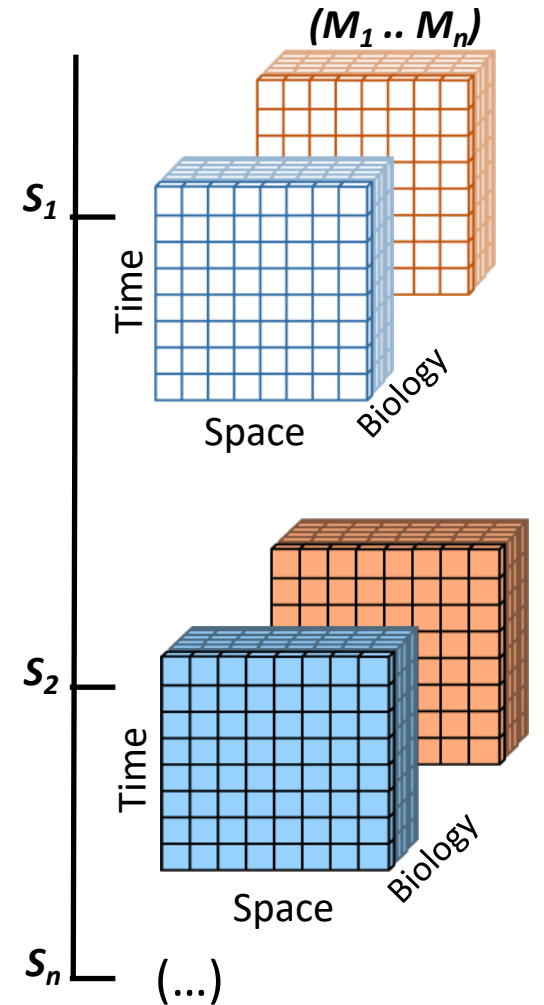


.json

❖ Web / server applications

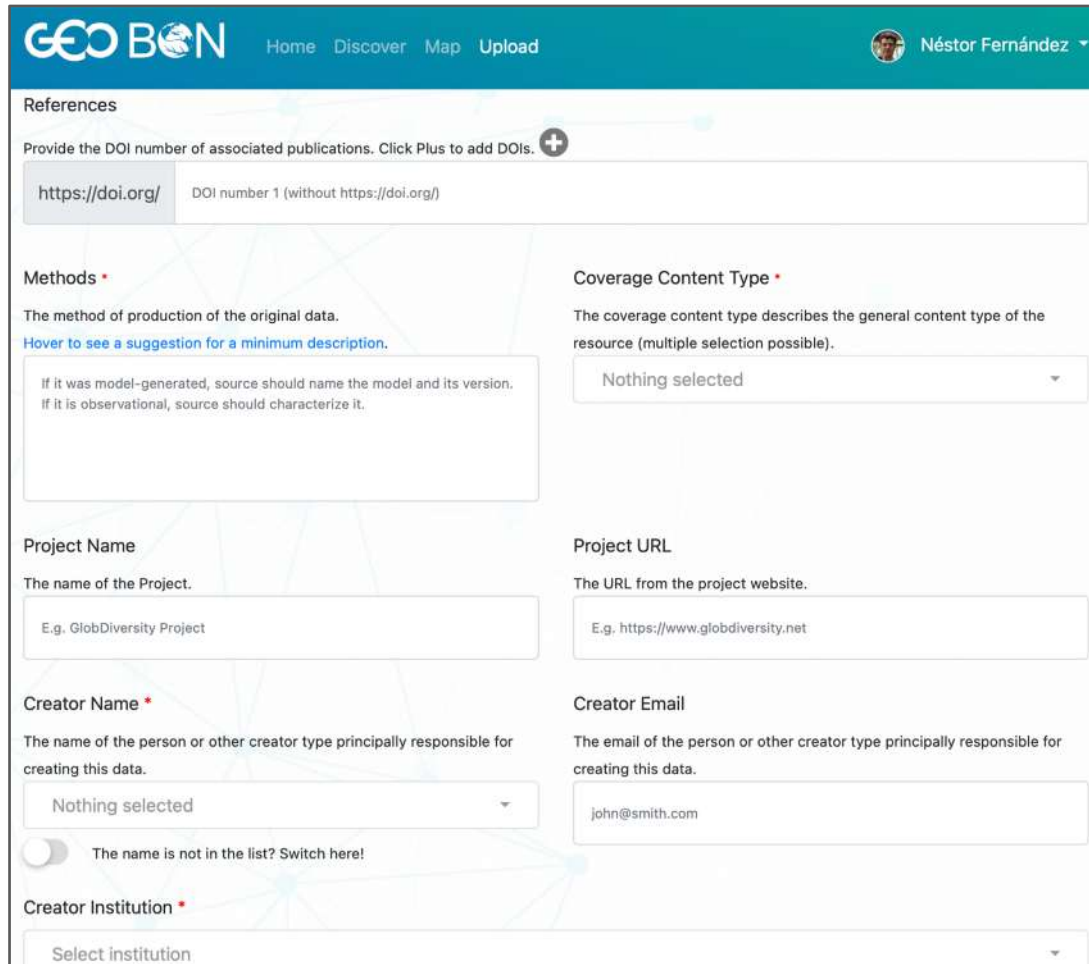
*Implementation of  
multiple taxonomies*

❖ GBIF backbone  
taxonomy



# EBV Cube tools

## EBV Portal Upload tool



The screenshot shows the EBV Portal Upload tool interface. At the top, there is a header with the GEO BON logo and navigation links: Home, Discover, Map, and Upload. A user profile for Néstor Fernández is visible in the top right corner. The main section is titled "References" and contains a form for uploading data. The form includes a text input for the DOI number, a "Methods" section with a description and a "Coverage Content Type" dropdown menu. Below these are fields for "Project Name", "Project URL", "Creator Name", "Creator Email", and "Creator Institution". A toggle switch is also present for "The name is not in the list? Switch here!".

**References**

Provide the DOI number of associated publications. Click Plus to add DOIs. +

DOI number 1 (without https://doi.org/)

**Methods \***

The method of production of the original data.

[Hover to see a suggestion for a minimum description.](#)

If it was model-generated, source should name the model and its version.  
If it is observational, source should characterize it.

**Coverage Content Type \***

The coverage content type describes the general content type of the resource (multiple selection possible).

**Project Name**

The name of the Project.

**Project URL**

The URL from the project website.

**Creator Name \***

The name of the person or other creator type principally responsible for creating this data.

☐ The name is not in the list? Switch here!

**Creator Email**

The email of the person or other creator type principally responsible for creating this data.

**Creator Institution \***



ebvcube

- ❖ Create NetCDF-EBV cubes
- ❖ Access the data
- ❖ Basic descriptions
- ❖ Visualization



# EBV Data Portal

The EBV Data Portal includes a variety of EBV raster datasets. You can import these datasets into the map with a single click. You can also upload your own EBV dataset for sharing with others.

<https://portal.geobon.org>

# Q&A



# Presentation of data architecture and data workflows in GBIF

By Tim Hirsch, Deputy Director, GBIF Secretariat

# Data flows in GBIF and diversifying the GBIF data model

Tim Hirsch  
Deputy Director, GBIF Secretariat



**GBIF**

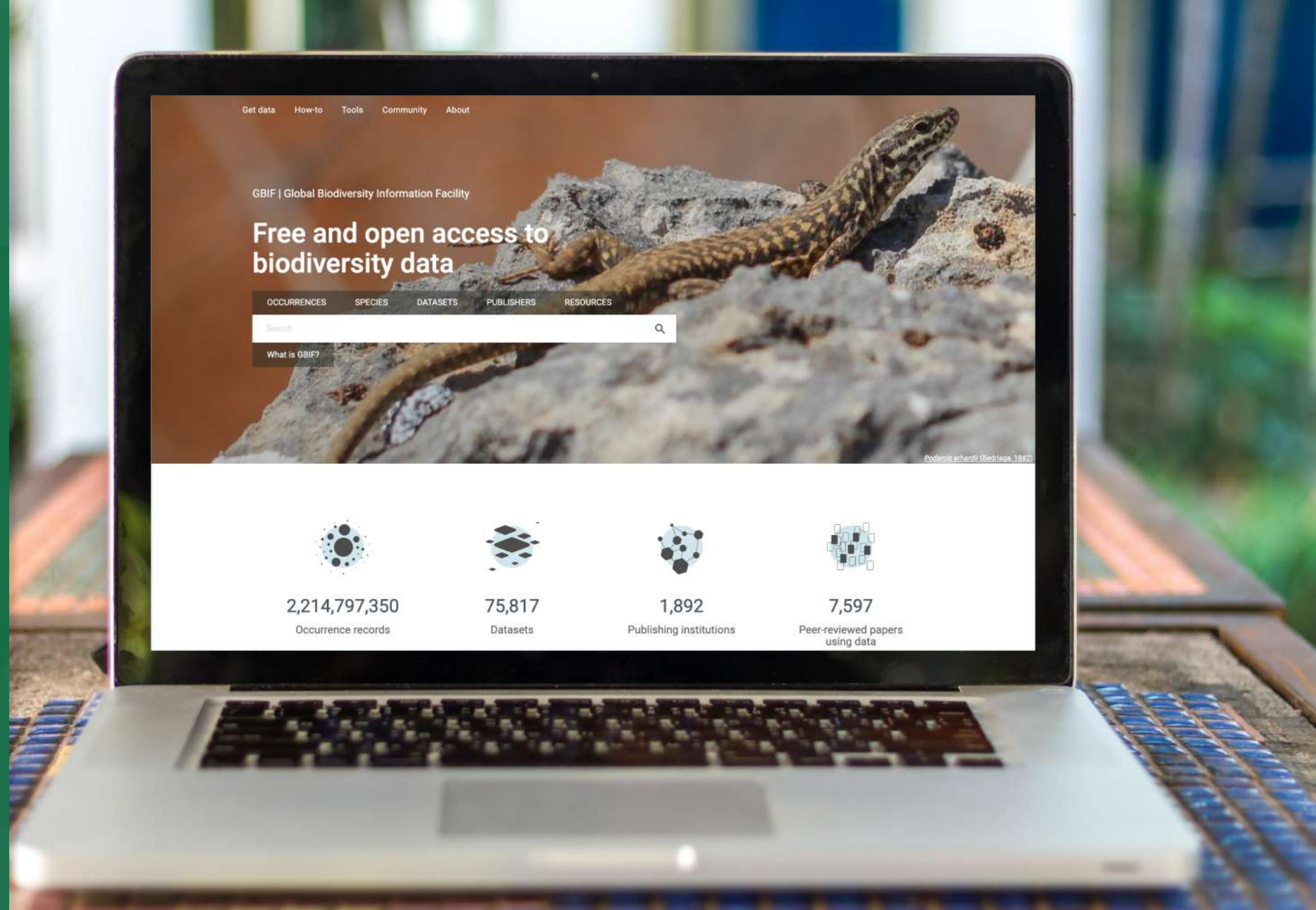
Global Biodiversity  
Information Facility



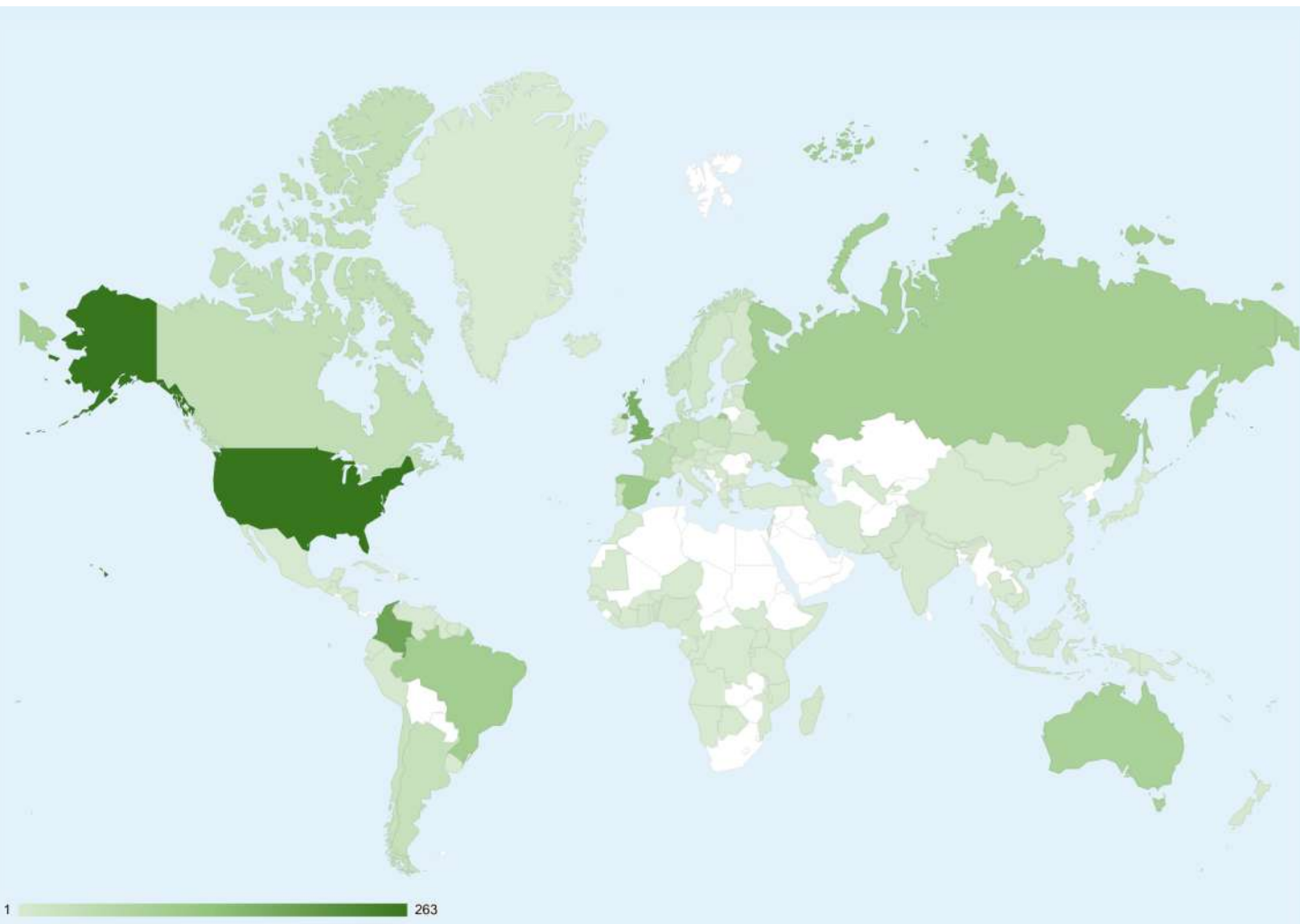
Open data

[www.gbif.org](http://www.gbif.org)

[api.gbif.org](http://api.gbif.org)



# GBIF Network of Data publishing institutions 30 June 2022

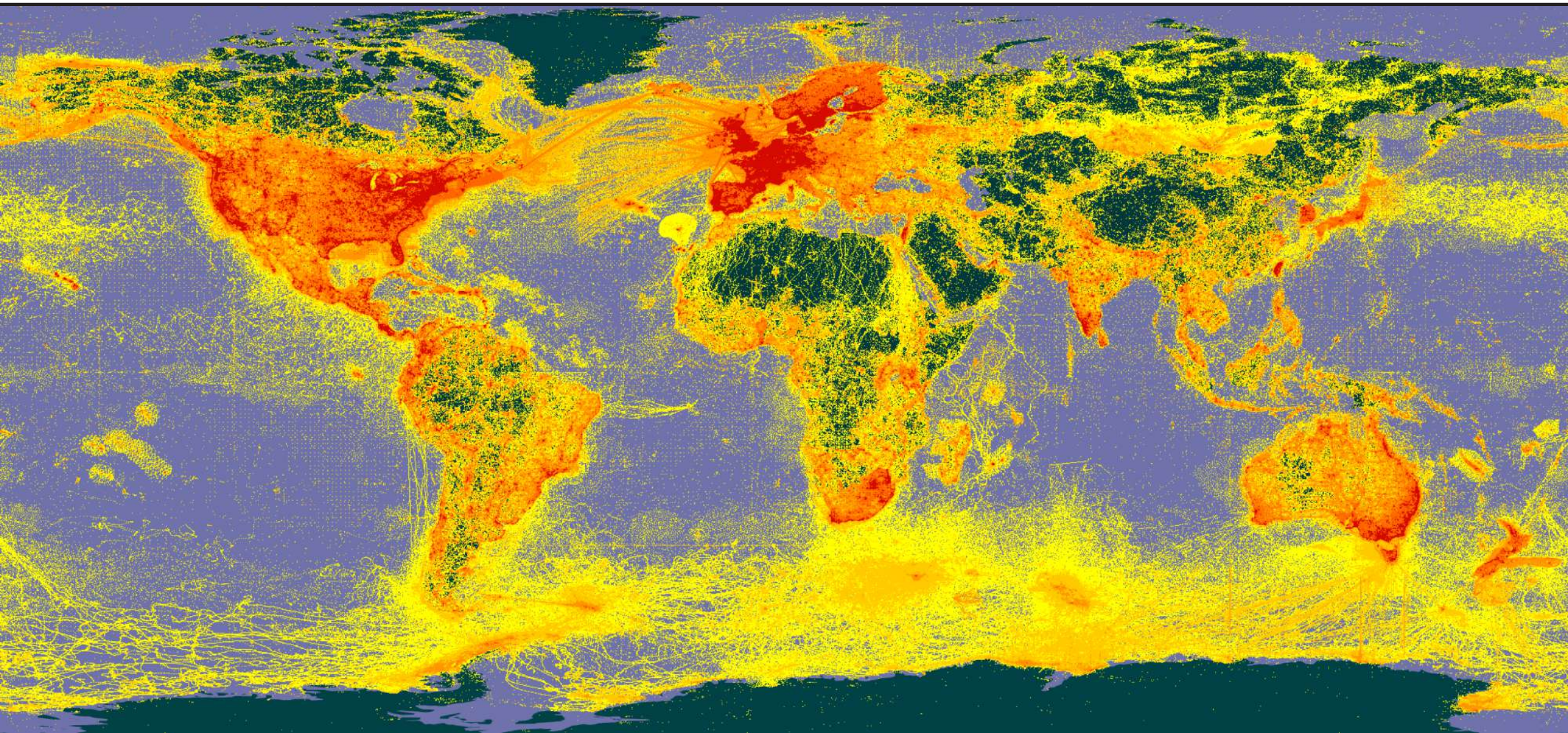


**132** countries/territories  
with institutions  
sharing data  
through GBIF

Top 10 countries: number of data publishers		
1	United States	263
2	Colombia	183
3	United Kingdom	165
4	Spain	111
5	Brazil	100
6	Russian Federation	95
7	Australia	87
8	France	56
9	Canada	44
10	Netherlands	41



Data From the GBIF Network 30 June 2022





# Species occurrence records with MultiMedia evidence

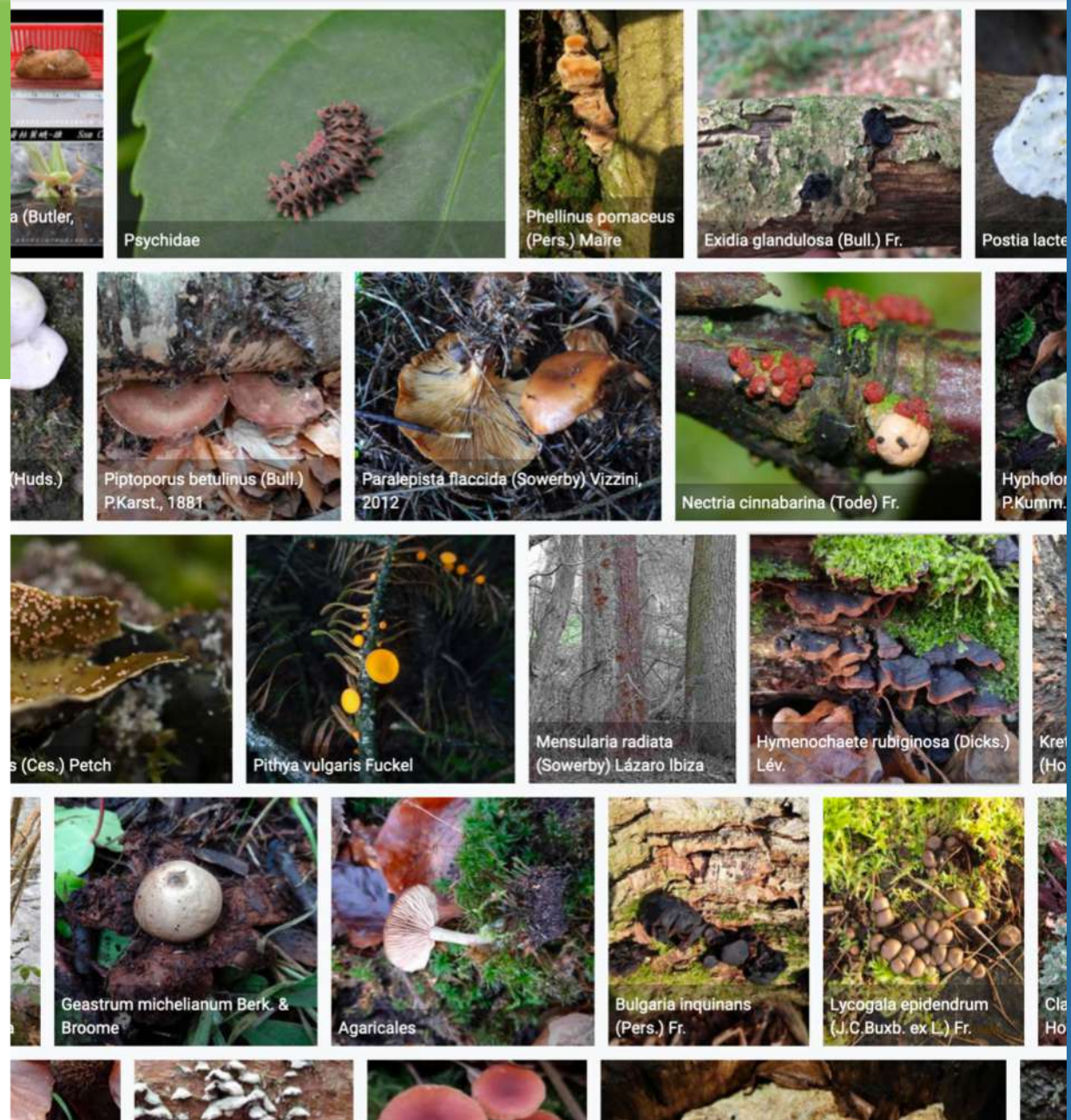
30 June 2022

## 107 million records with taxonomically identified images

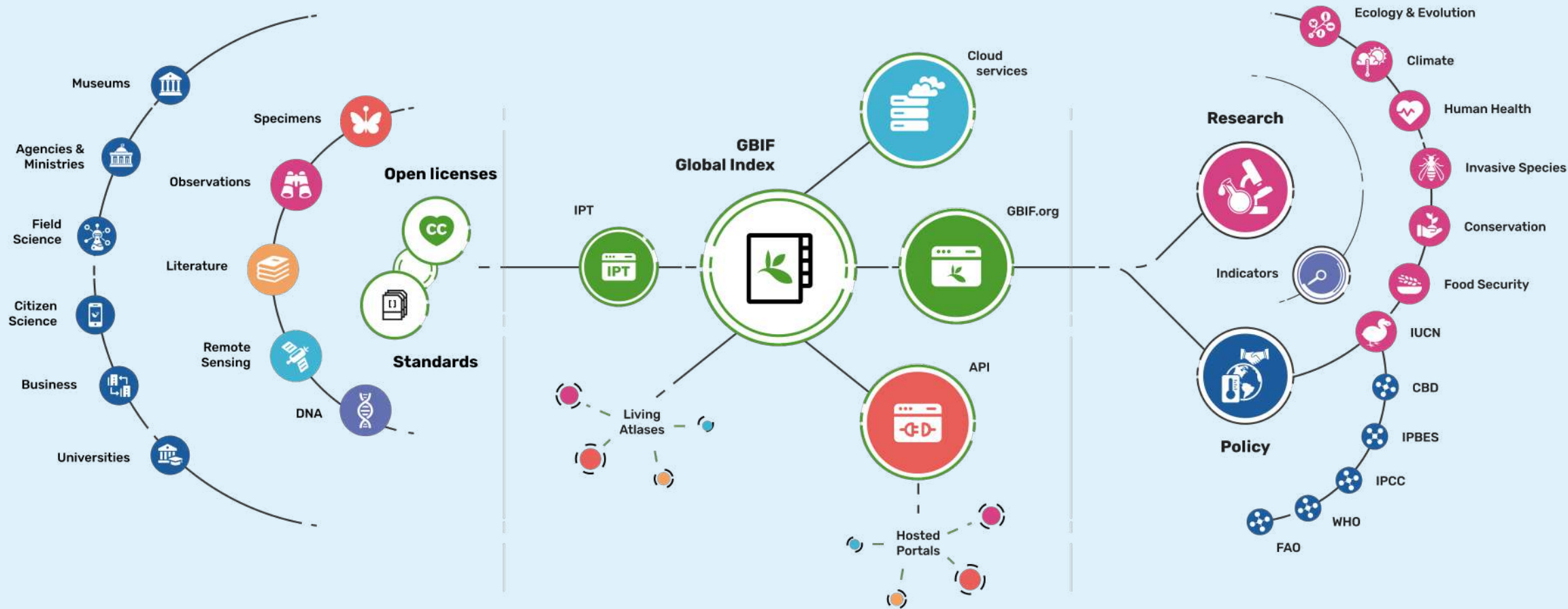
- 61.4 million human observations
- 41.2 million specimens
- 2.3 million material samples
- 1.4 million fossil specimens

**914,887 audio files**

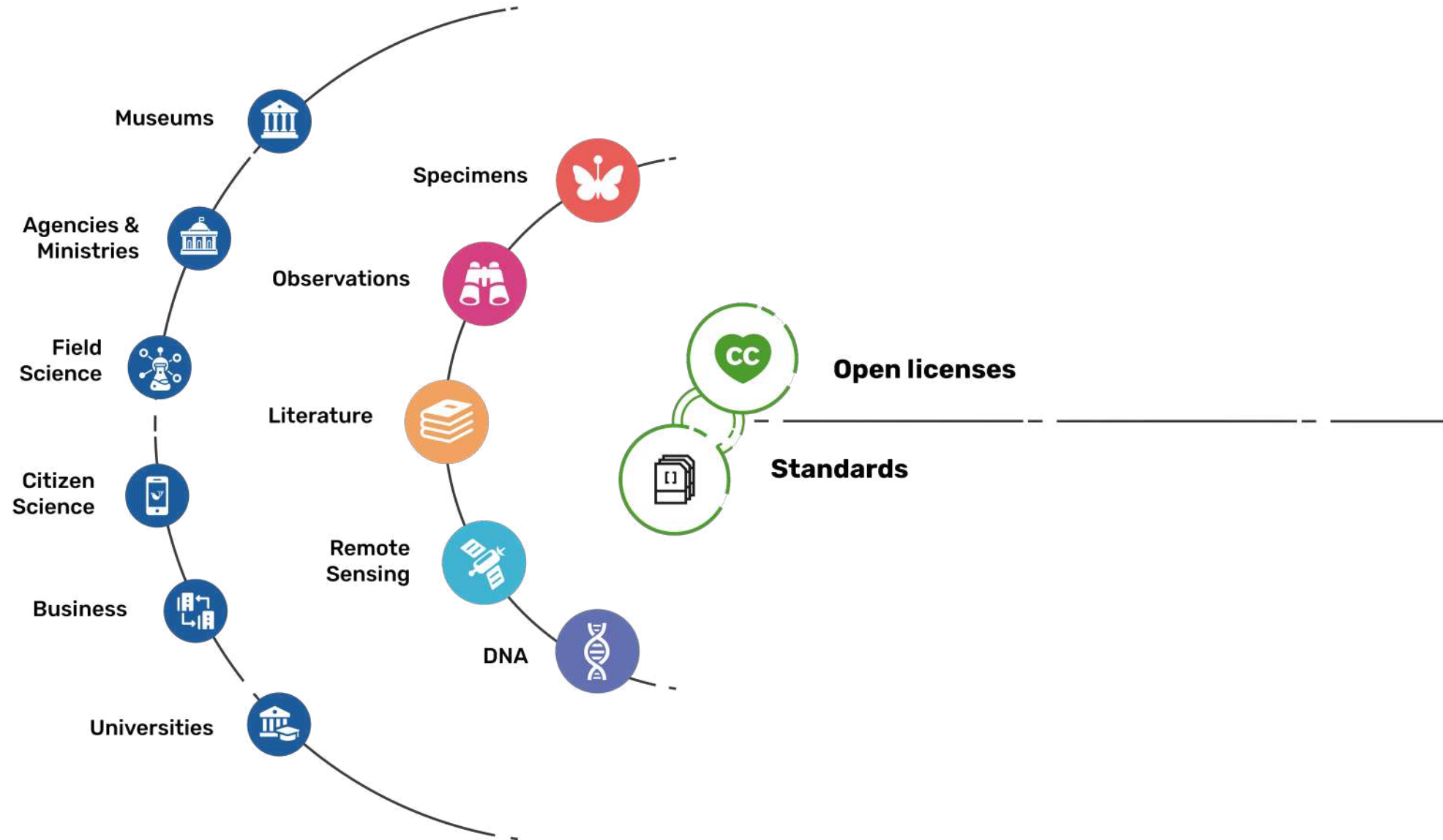
**3,526 videos**



# Providing biodiversity evidence for research and policy

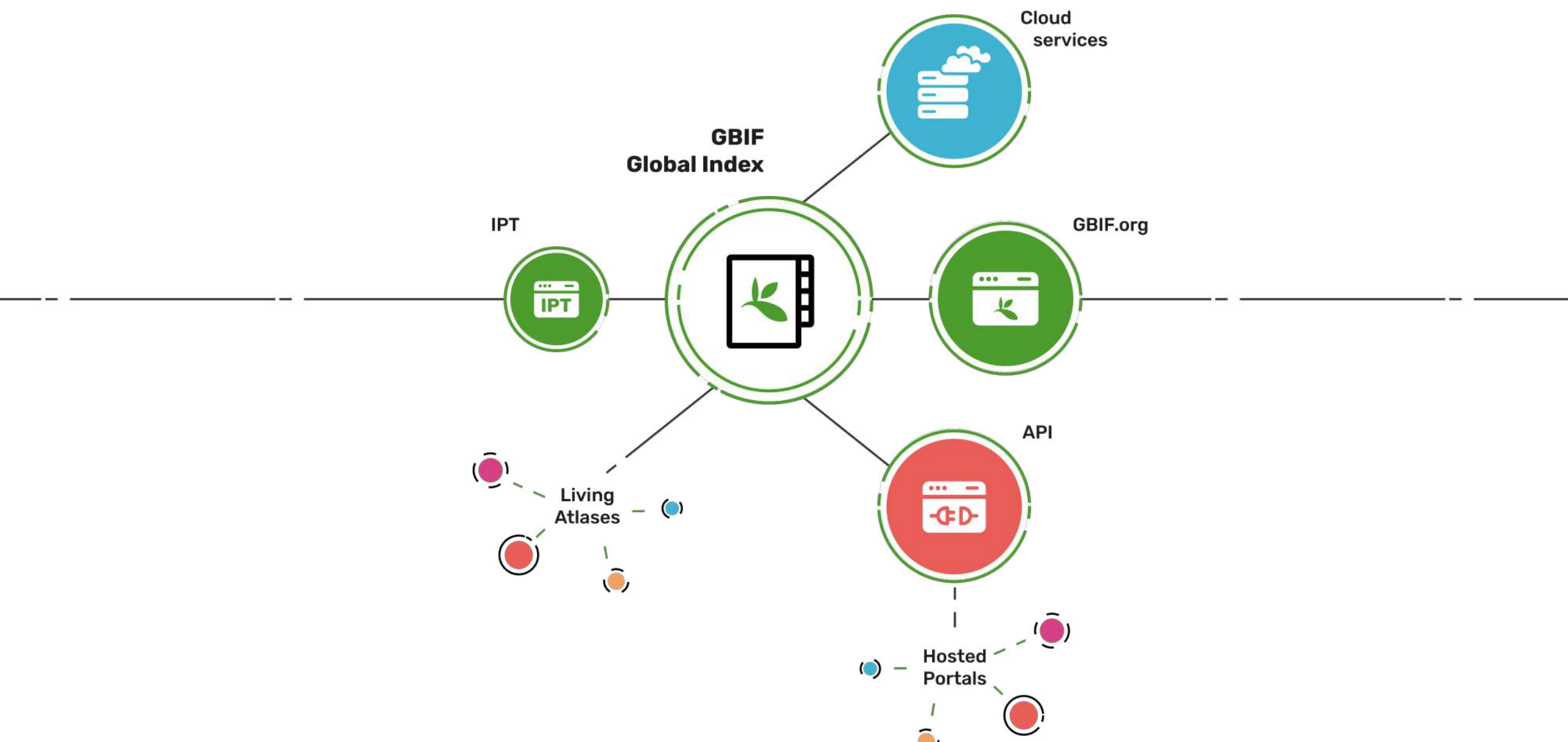


# SOURCES OF BIODIVERSITY EVIDENCE

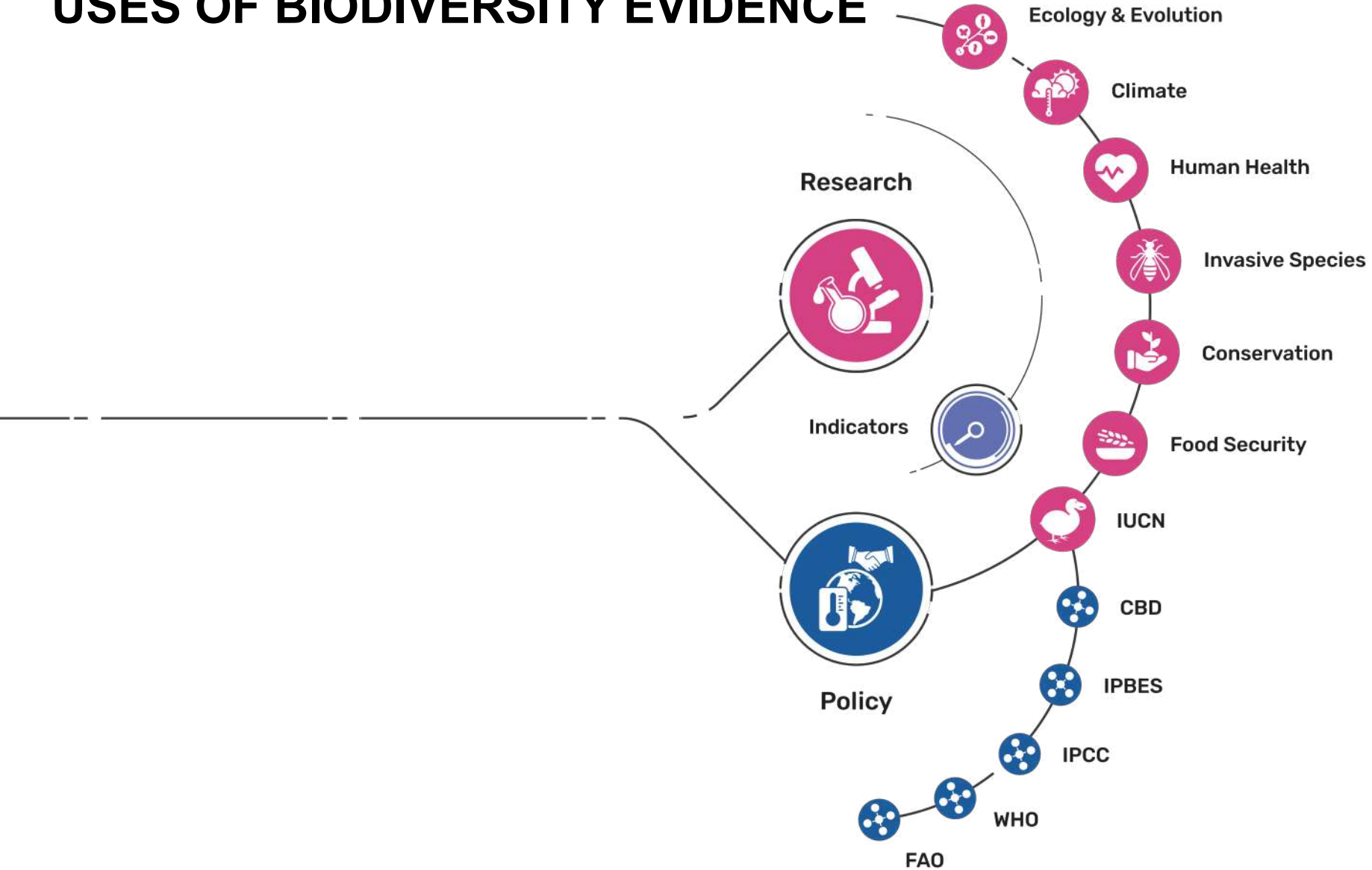




# ACCESS TO BIODIVERSITY EVIDENCE

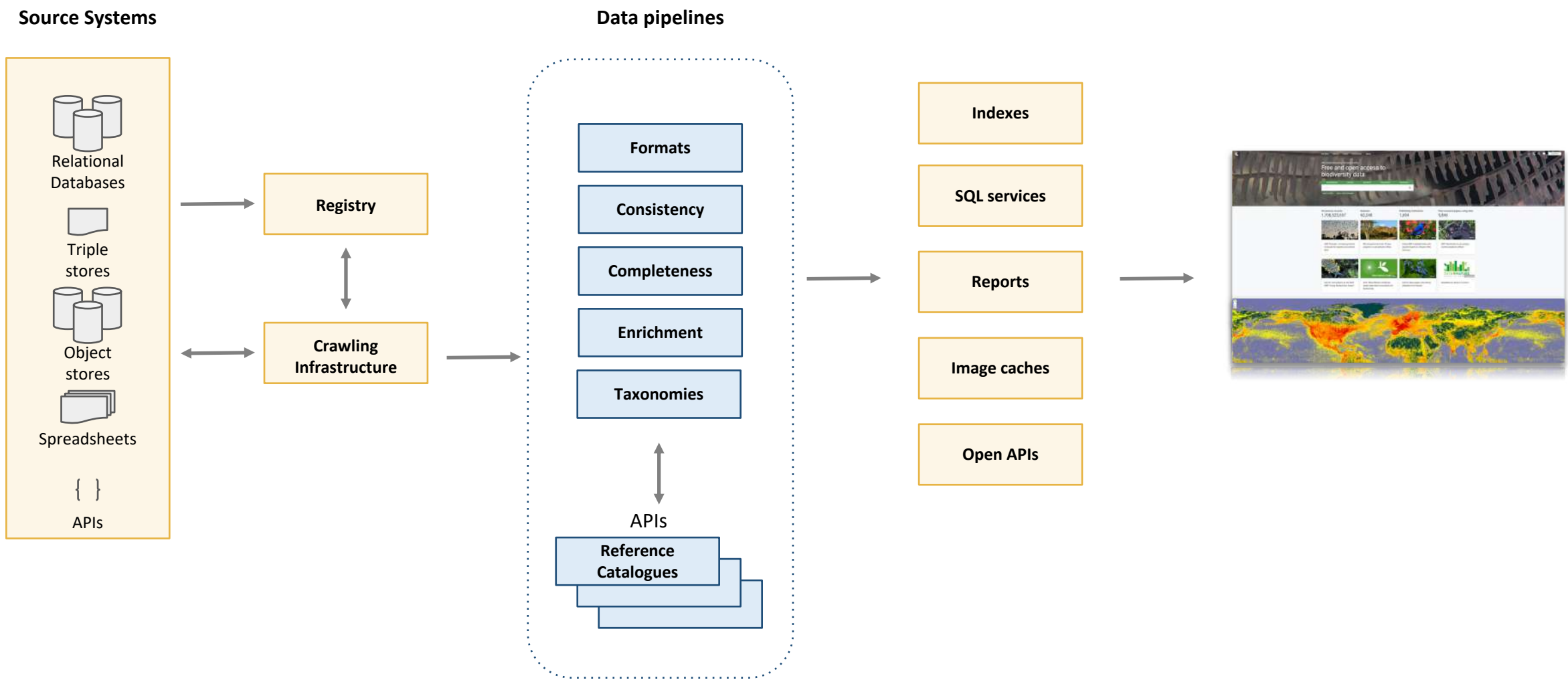


# USES OF BIODIVERSITY EVIDENCE



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

# Data warehouse



## Data Catalog

The Planetary Computer Data Catalog includes petabytes of environmental monitoring data, in consistent, analysis-ready formats. All of the datasets below can be accessed via Azure Blob Storage, and can be used by developers whether you're working within or outside of our Planetary Computer Hub.

## Additional datasets

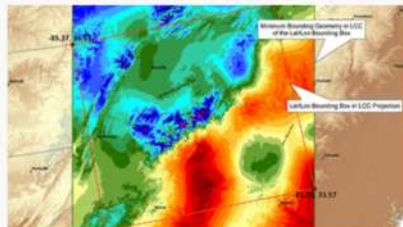
The following datasets are available on Azure, for use within or outside of the Planetary Computer Hub.



### GBIF

Species occurrences shared through the Global Biodiversity Information Facility

[Get GBIF occurrence data >](#)



### Daymet

Gridded temperature data across North America

[Get Daymet temperature data >](#)



### Global Hydro Estimator

Global precipitation estimates

[Get global precipitation data >](#)



### GOES-16

Weather imagery of the Americas

[Get GOES-16 weather data >](#)

# Global Biodiversity Information Facility (GBIF) Species Occurrences

[biodiversity](#) [bioinformatics](#) [conservation](#) [earth observation](#) [life sciences](#)

## Description

The Global Biodiversity Information Facility (GBIF) is an international network and data infrastructure funded by the world's governments providing global data that document the occurrence of species. GBIF currently integrates datasets documenting over 1.6 billion species occurrences, growing daily. The GBIF occurrence dataset combines data from a wide array of sources including specimen-related data from natural history museums, observations from citizen science networks and environment recording schemes. While these data are constantly changing at GBIF.org, periodic snapshots are taken and made available on AWS.

## Update Frequency

Snapshots of GBIF are taken on a monthly basis

## License

This dataset is available under a [CC-BY](#) license and with the GBIF [terms of use](#). Please refer to the GBIF [citation guidelines](#) when using this dataset.

## Documentation

Documentation can be found [here](#). You can learn more about GBIF [here](#).

## Managed By

The Global Biodiversity Information Facility (GBIF)

See all datasets managed by GBIF.

## Contact

[helpdesk@gbif.org](mailto:helpdesk@gbif.org)

## Resources on AWS

### Description

GBIF species occurrence data in Parquet format (af-south-1 region).

### Resource type

S3 Bucket

### Amazon Resource Name (ARN)

`arn:aws:s3:::gbif-open-data-af-south-1`

### AWS Region

`af-south-1`

### AWS CLI Access (No AWS account required)

`aws s3 ls s3://gbif-open-data-af-south-1/ --region af-south-1 --no-sign-request`

### Explore

[Browse bucket](#)

### Description

GBIF species occurrence data in Parquet format (ap-southeast-2 region).



## Hosted portals building on subsets of data shared through GBIF



GBIF.us



Biodiversidad.co



SANBI-GBIF



Natural History Museum  
Rotterdam



Pacific Biodiversity  
Information Facility



Legume Data Portal



Virtual Herbarium Germany




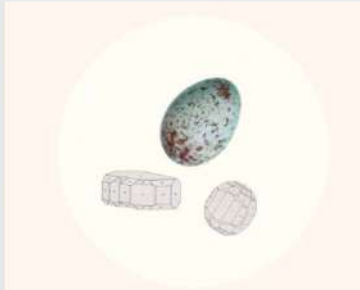
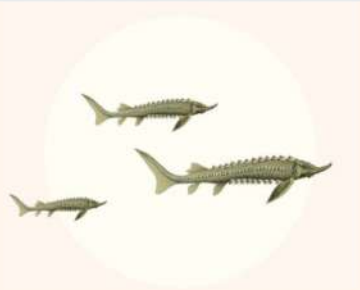
Living Norway Ecological  
Data Network









# Diversifying the GBIF Data Model








Explore the current ideas for a unified common model capable of supporting expanded data-publishing capabilities and potential directions for evolving the Darwin Core standard



## Case studies

		
► Community webinar recording 12 April 2022	► Intro presentation: Diversifying the GBIF data model	Slides: Diversifying the data model

 Introduction	 eDNA barcoding	 Camera trap	 Tissue samples
 Automatic moth trap	 Global malaise programme	 iNaturalist	 Specimens with media

 Environmental and community measurements	 Taxonomic treatments in literature	 Malaise trapping for reference barcode collection	 Taxonomic checklist
 Ecological survey data exchange specification	 Biotic interactions	 Glossary	

# THANK YOU

[thirsch@gbif.org](mailto:thirsch@gbif.org)



[www.gbif.org](http://www.gbif.org)



# Q&A



# Presentation of Meta(data) structure and workflows in LifeWatch ERIC

By Lucia Vaira, LifeWatch ERIC

# Meta(data) structure and workflows in



Biodiversa+ WS biodiversity monitoring databases

**Lucia Vaira** | LifeWatch ERIC Service Centre, Lecce – Italy | [lucia.vaira@lifewatch.eu](mailto:lucia.vaira@lifewatch.eu)

# Outline

- LifeWatch ERIC
  - Organisation and governance
- The (meta)data interoperability need
  - The FAIR principles
  - Alignment with the EOSC Interoperability Framework
- Meta(data) structure and workflows in LifeWatch ERIC
  - The LifeWatch ERIC Metadata Catalogue
  - The LifeWatch ERIC EcoPortal
  - The interoperability in LifeWatch ERIC

# LifeWatch ERIC

**LifeWatch ERIC** is a European Infrastructure Consortium that offers e-Science research facilities to scientists investigating Biodiversity organization and Ecosystem functions and services.

LifeWatch ERIC's mission: to be the leading worldwide provider of content and services for the biodiversity research community, creating new opportunities for large-scale scientific development, and enabling accelerated data capture and modelling thanks to the use of innovative technologies.



LifeWatch ERIC is a distributed research e-infrastructure consortium consisting of 8 EU Member States.

Its structure mirrors its nature, with central components (Common Facilities) located in 3 Member States (Spain, Italy and the Netherlands), and National Nodes in all 8 countries.

LifeWatch ERIC's current members are: Belgium, Bulgaria, Greece, Italy, the Netherlands, Portugal, Slovenia and Spain.

Slovakia participates as an Observer.



# LifeWatch ERIC

With LifeWatch ERIC users can:

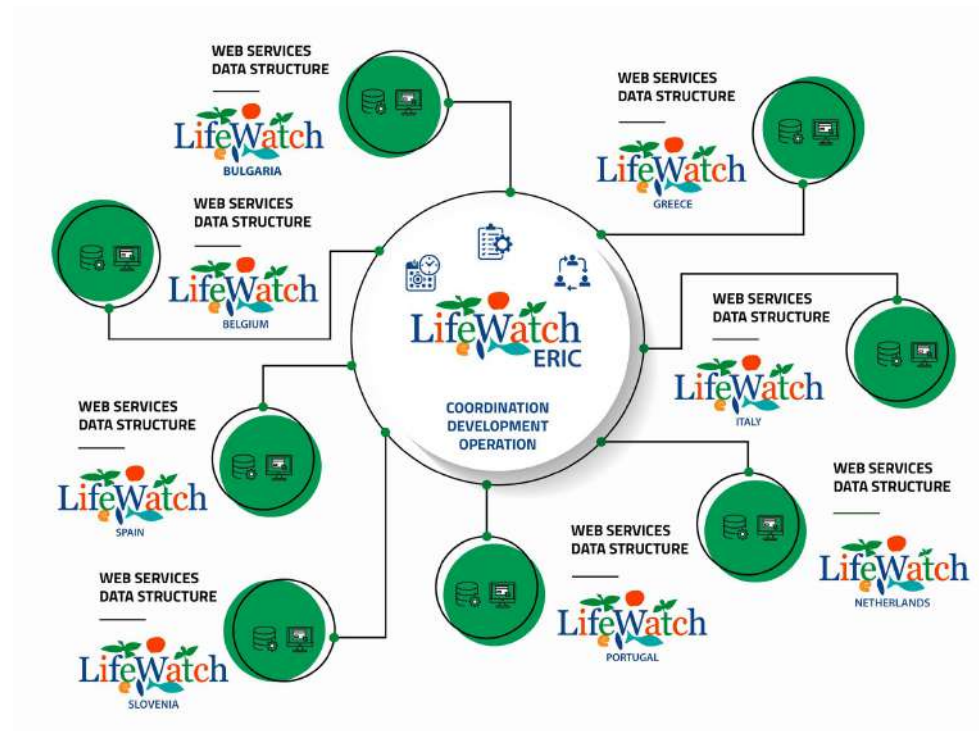
- find, access, work with and reuse FAIR data collected by science at a global level, of different typologies and scales. Advanced search functions, thesauri and ontologies are available on our catalogues to further combine data and generate new services;
- process and analyse data in our Virtual Research Environments (VREs);
- get trained in the use of our services, as well as on key scientific issues through Master's and PhD courses, summer schools, webinar programmes and educational initiatives;
- receive support through our Helpdesk and its experts specialised in different domains.



# LifeWatch ERIC

Over the last 5 years the Common Facilities and Distributed Centres have gone from being isolated elements to fully integrated into LifeWatch ERIC.

Rules and policies are continuously being developed and updated to support this integration process.



# The (meta)data interoperability need

Considering the FAIR principles, interoperability is considered in relation to the fact that “research data usually need to be integrated with other data; in addition, the data need to interoperate with applications or workflows for analysis, storage, and processing”.

The following principles are proposed:

- I1. (Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
- I2. (Meta)data use vocabularies that follow FAIR principles.
- I3. (Meta)data include qualified references to other (meta)data.

LifeWatch ERIC: achieving interoperability is essential to allow the federation of **all the LifeWatch ERIC assets**  
→ provide added value for the final users.

# Alignment with the EOSC Interoperability Framework

EOSC Interoperability Framework\*: 4 main points of view and a set of recommendations has been proposed:

- **Technical:** the “ability of different information technology systems and software applications to communicate and exchange data”, *EOSC IF*.
- **Semantic:** the “ability of computer systems to transmit data with unambiguous, shared meaning”, *FAIRsFAIR deliverable D2.1*
- **Organisational:** the “way in which organisations align their business processes, responsibilities and expectations to achieve commonly agreed and mutually beneficial goals”, *EIF*
- **Legal:** “it deals, in particular, how data should be re-used”, *EOSC IF*.

\* European Commission, Directorate-General for Research and Innovation, Corcho, O., Eriksson, M., Kurowski, K., et al., *EOSC interoperability framework : report from the EOSC Executive Board Working Groups FAIR and Architecture*, Publications Office, 2021, <https://data.europa.eu/doi/10.2777/620649>



# Alignment with the EOSC Interoperability Framework

## Technical interoperability

1. The definition of a common security and privacy framework (including Authorisation and Authentication Infrastructure) is needed to ensure secure and trustworthy data exchange.
2. Easy-to-understand Service-Level Agreements for all resource providers.
3. Easy access to data sources available in several formats to facilitate overcoming their heterogeneity and allow integrating data across communities.
4. Coarse-grained and fine-grained dataset (and other research objects) search tools need to be made available.
5. Multiple service providers for different types of PIDs exist. As a result, sometimes the identifiers are not resolvable. A clear PID policy needs to be defined.

# Alignment with the EOSC Interoperability Framework

## Semantic interoperability

1. Clear and precise, publicly-available definitions for all concepts, metadata and data schemas should be available in the communities.
2. Semantic artefacts should be available preferably with open licenses and must have sufficient associated documentation.
3. There is a lack of common reference repositories of semantic artefacts.
4. A minimum metadata model should be adopted to ease discovery over existing federated research data and metadata.
5. There should be extensibility options to allow for disciplinary metadata allowing users to add annotations.
6. There should be clear protocols and building blocks for the federation/harvesting of semantic artefacts catalogues.

# Alignment with the EOSC Interoperability Framework

## Organisational interoperability

1. A clearly-defined governance structure that includes the governance framework that will deal with interoperability across organisations and disciplines.
2. A clear description of the “terms and conditions” and “acceptable use policies” that will rule the provisioned services.
3. The rules of participation recommendations should be completed with aspects related to interoperability.
4. Usage recommendations of standardised data formats and/or vocabularies, with their corresponding metadata.
5. A clear management of permanent organisation names and functions needs to be provided (sustainability).

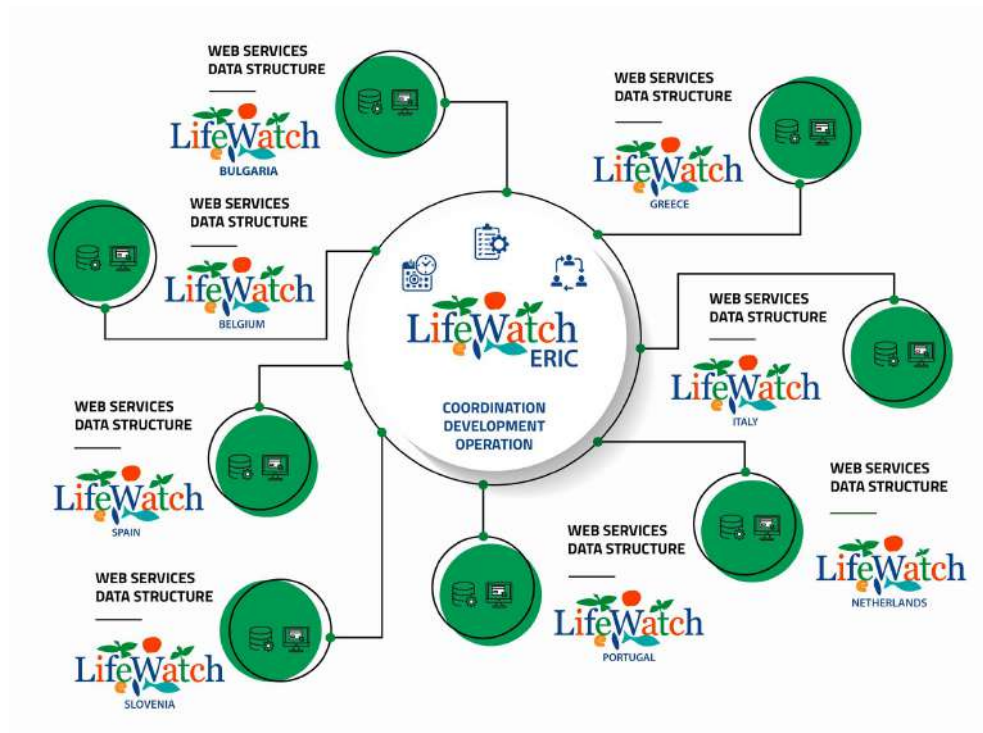
# Alignment with the EOSC Interoperability Framework

## Legal interoperability

1. Information on who holds the rights to the data, including any embedded data.
2. Metadata to be available without restrictions on (re-)use to facilitate the FAIRness of the data it describes.
3. Licensing requirements should be both human and machine readable and allow data providers and users to understand license compatibility.
4. Clear status of historic copyrightable datasets and metadata with no license or unclear licenses arrangements (orphan data).
5. Need to track the license evolution over time for datasets.
6. Need of harmonised policy and guidance to deal with patents or trade secrets disclosure.
7. Need to ensure adequate protection of personal data and general compliance with the GDPR and EU Member States' domestic law and guidelines.
8. Harmonised terms of use across repositories.



# Meta(data) structure and workflows in LifeWatch ERIC



LifeWatch ERIC: central node for the coordination, development and operation.

Main goal: to increase collaboration within and among national nodes but also among external organisations to:

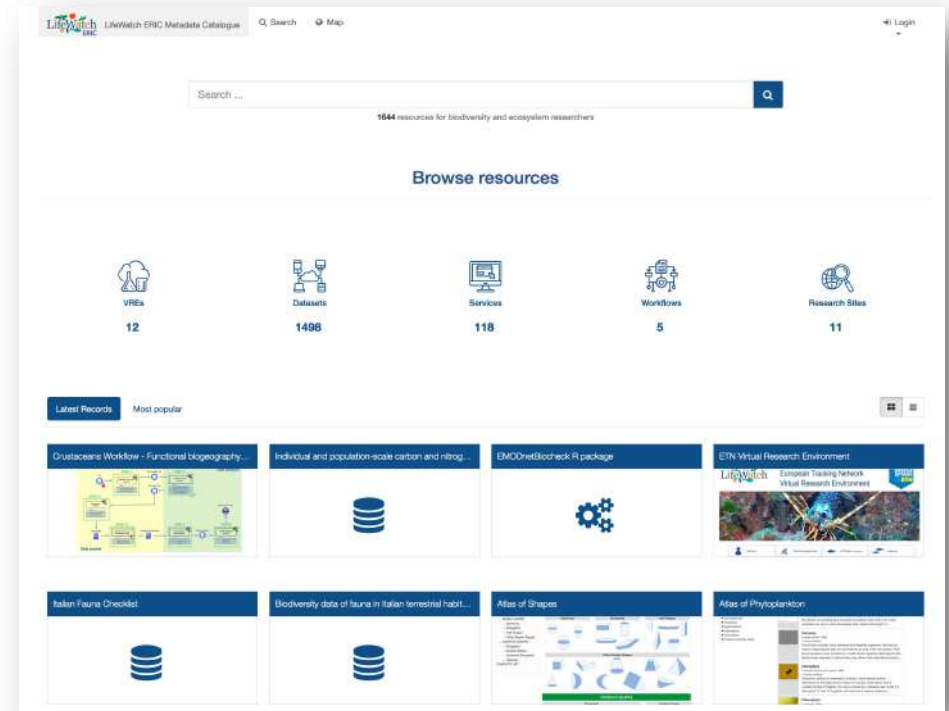
- enhance findability;
- reduce duplication;
- improve information consistency and quality.

A standard-based information management system based on GeoNetwork 3.10.

It enables access to several resources from a variety of providers through descriptive metadata, enhancing and promoting the information exchange and sharing among organisations and research infrastructures.

It allows to manage descriptive metadata related to **datasets** (EML 2.2.0 standard), **research sites**, **services**, **Virtual Research Environments**, and **workflows** (ISO 19139 standard). Metadata attributes can be optional/mandatory and can require single/multiple values.

It allows (upon validation and verification) the creation of Digital Object Identifiers (DOIs) for resources that do not have it, by exploiting the GeoNetwork – DataCite connection.



<https://metadatatatalogue.lifewatch.eu>



**EML2.2.0 standard - 77 metadata fields**



**ISO 19139 standard - 34 metadata fields**



**ISO 19139 standard - 16 metadata fields**



**ISO 19139 standard - 24 metadata fields**




**ISO 19139 standard - 24 metadata fields**

## Crustaceans Workflow - Functional biogeography of invaders: the case of two widely distributed omnivorous crustaceans

Invaders are to date acknowledged as significant environmental and economic threats, yet the identification of key ecological traits determining species invasiveness has remained elusive. One unappreciated source of variation concerns dietary flexibility and the ability to shift trophic position within invaded food webs. Trophic plasticity may greatly influence invasion success as it regulates resource availability as a functional constraint for introduced individuals; in addition, the impact of an invader might increase with its dietary plasticity since a greater proportion of the resident assemblage of species and multiple trophic levels are affected.

The validation case will focus on two invasive crustaceans widely distributed in marine and freshwater European waters, i.e., the Atlantic blue crab *Callinectes sapidus* and the Louisiana crayfish *Procambarus clarkii*. Key questions that will be addressed are:


- a) does the trophic position of *C. sapidus* vary between invaded Mediterranean and native west Atlantic habitats, testifying the occurrence of significant post-invasion shifts in dietary habits?
- b) does the trophic position of *P. clarkii* vary in invaded habitats, and which are the main ecological factors involved?
- c) do bioclimatic drivers influence broad-scale spatial patterns of variation in the trophic position of *C. sapidus* and *P. clarkii*?



**workflow**

### About this resource

Status	Under development / Pre operational
Version	1.0



**Info page**  
<http://www.lifewatch.eu/ji-4-callinectes-sapidus>

Open link

### Central Services

• Service 1

Name	Crustaceans WoRMS Taxonomic Checker
Description	A service that aims at verifying the species names (performing the taxonomy check) by using the input file SIA.csv. It represents the Step 1.
Reference	<a href="https://metadatas2014lifewatch.eu/en/catalog/catalog_search#/metadatas2014lifewatch-36ad-4e7-4c70-ef09a25381e">https://metadatas2014lifewatch.eu/en/catalog/catalog_search#/metadatas2014lifewatch-36ad-4e7-4c70-ef09a25381e</a>

• Service 2

Name	Trophic Position Analyser
Description	A service that aims at analyzing the trophic position of the "invader" by using a Bayesian model to estimate the NA values of the trophic position that corresponds to the column 13 ("TP" of the "Tax validated_SIA.csv" file by means of a loop that allows extracting data of single populations. It represents the Step 2.
Reference	<a href="https://metadatas2014lifewatch.eu/en/catalog/catalog_search#/metadatas2014lifewatch-5912-48ef-8244-bd93d177b41">https://metadatas2014lifewatch.eu/en/catalog/catalog_search#/metadatas2014lifewatch-5912-48ef-8244-bd93d177b41</a>


• Service 3

⌚ Temporal extent


**Creation date**  
2019-07-04

**Publication date**  
2021-02-25

🔍 Source

 LifeWatch ERIC

🔗 Provided by

 LifeWatch ERIC






📅 Updated:

a month ago


📅 Updated:

a month ago


📢 Share on social sites

Under development / Pre operational



**workflow**



# Trophic Position Analyzer

## service

A service that aims at estimating the trophic position of an organism by using a Bayesian model to estimate the  $\delta^{15}\text{N}$  values of the trophic position and converges to the means 13 "TP" of the "Tax\_validated\_SIA.csv" file by means of a loop that allows extracting data of single populations.

represents the Step 2 of the Crustaceans Workflow within the Internal Joint Initiative.

### About this resource

<b>Keywords</b>	<ul style="list-style-type: none"> <li>functional trait <a href="#">🔗</a></li> <li>alien invasive species <a href="#">🔗</a></li> <li>native species <a href="#">🔗</a></li> <li>trophic position <a href="#">🔗</a></li> <li>Crustaceans <a href="#">🔗</a></li> <li>LiJ <a href="#">🔗</a></li> </ul>
<b>Status</b>	Under development / Pre operational
<b>Version</b>	1.0

### Contain Operations

• Operation 1

<b>Name</b>	Tax_validated_SIA.csv
<b>Website</b>	<a href="http://metacatcatalogue.lifewatch.eu/en/catalog/search?lifeWatchData/46d073f0-ac-4d50-433f-a713-d5d553caw415">http://metacatcatalogue.lifewatch.eu/en/catalog/search?lifeWatchData/46d073f0-ac-4d50-433f-a713-d5d553caw415</a>
<b>Description</b>	It is an intermediate file of the Crustaceans workflow that contains the validated SIA file by identifying the native and invasive species.
<b>Function</b>	Input file

• Operation 2

<b>Name</b>	Exp_troph_SIA.csv
<b>Website</b>	<a href="http://metacatcatalogue.lifewatch.eu/en/catalog/search?lifeWatchData/7977d018-d419-4038-a38d-c156b0beb191">http://metacatcatalogue.lifewatch.eu/en/catalog/search?lifeWatchData/7977d018-d419-4038-a38d-c156b0beb191</a>
<b>Description</b>	It is an intermediate file of the Crustaceans workflow that contains the trophic position of the invader identified in Step 1 (Crustaceans WoRMS Taxonomic Checker) of the Workflow.
<b>Function</b>	Output file

### Service Classification Information

<b>Topic Categories</b>	
<b>Service Categories</b>	Service Category <span>data analysis</span>
<b>Service TRL</b>	TRL 4 - Technology validated in lab

### Other Information


Service helpdesk <https://www.lifewatch.eu/help-desk>

**Temporal extent**


Creation date  
2019-07-04

Publication date  
2021-02-25

**Source**



**Provided by**



**Updated:**

a month ago

**Updated:**

a month ago

**Share on social sites**

[🐦](#)
[f](#)
[in](#)
[📧](#)
[🔗](#)

[Under development / Pre operational](#)

[illegible]

**Link to access data**



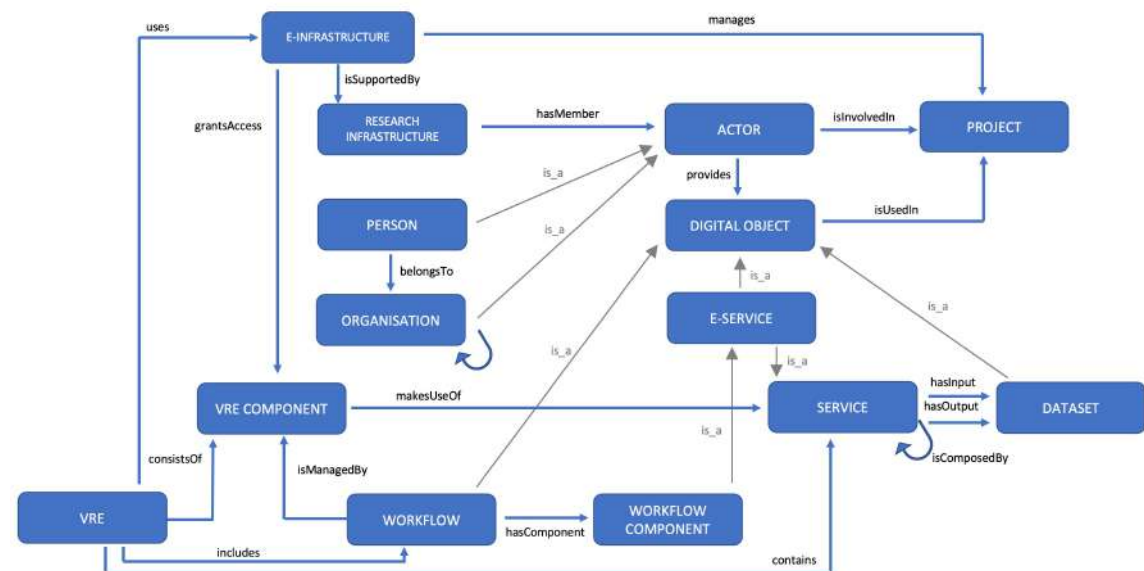
# The LifeWatch ERIC Upper Ontology

The usage of the same terminology concepts is highly relevant for the interoperability in the LifeWatch ERIC context.

The resources hosted in the Metadata Catalogue exploit the work already done within the Working Group D of the LifeWatch ERIC Internal Joint Initiative where **LUPO, the LifeWatch ERIC UPper Ontology\*** has been defined so that the key concepts (dataset, service, Virtual Research Environment, workflow, etc.) are well defined and agreed within the scientific community.

An upper ontology (top-level ontology, upper model, or foundation ontology) is an ontology which consists of very general terms (such as "object", "property", "relation") that are common across all the Research Infrastructure domains.

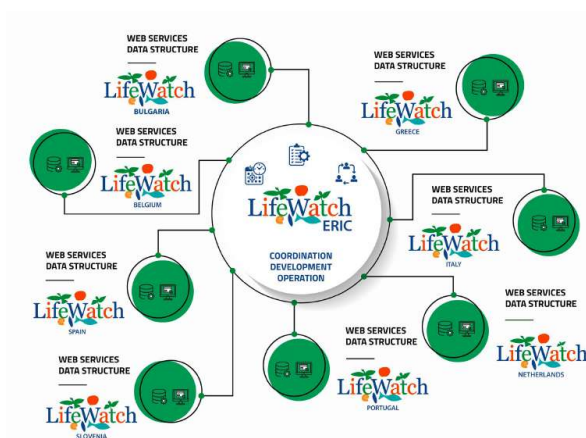
Main goal: to support broad semantic interoperability among a large number of domain-specific ontologies by providing a common starting point for the formulation of definitions.



\* LUPO Ontology: <http://ecoportal.lifewatch.eu/ontologies/LUPO>

# LifeWatch ERIC Metadata Catalogue

## Harvesting workflows



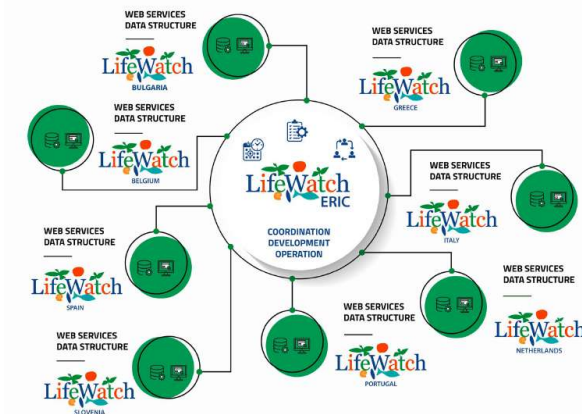
Harvesting procedures to  
import/export datasets via UI or API  
→ humans + machine2machine

OAI/PMH\*: our most used  
protocol due to the possibility to  
specify the exact endpoint, the  
prefixes, etc.

Used to harvest datasets metadata  
records from 4 LifeWatch ERIC  
national nodes

\*OAI-PMH - Open Archives Initiative Protocol for Metadata Harvesting

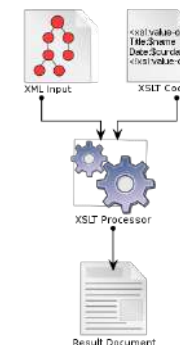
# LifeWatch ERIC Metadata Catalogue Mapping workflows



Mapping procedures to convert formats, map into LifeWatch ERIC metadata profile, align standards to the same version, etc.

XSLT\*: our most used method due to the possibility to specify the desired output according to the specific input.

Used to map harvested datasets metadata records from 4 LifeWatch ERIC national nodes



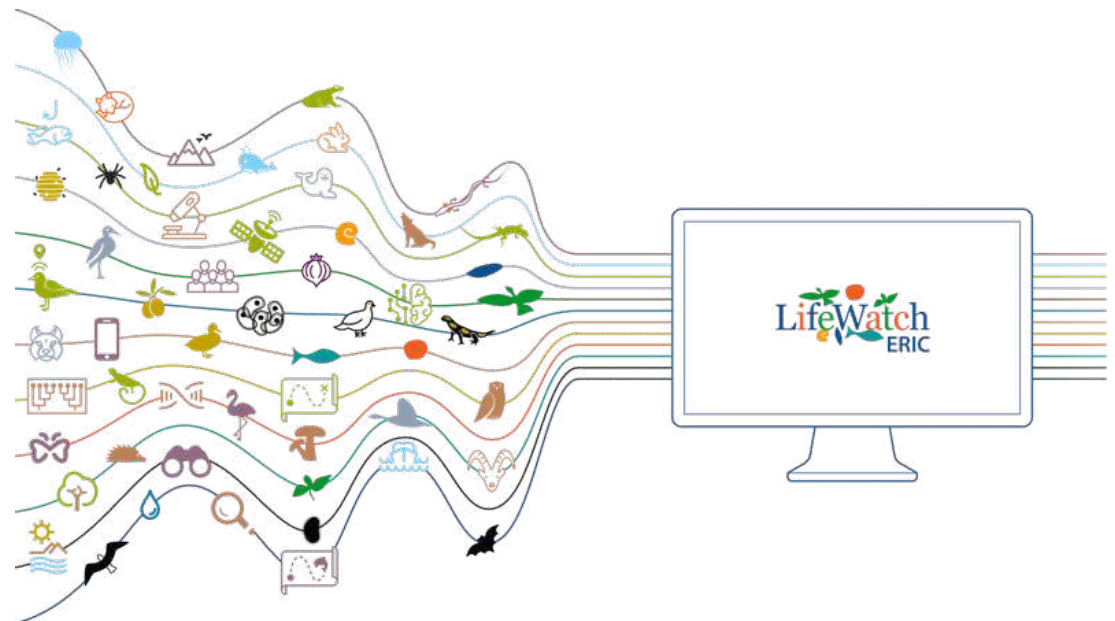
\*XSLT: eXtensible Stylesheet Language Transformations

# LifeWatch ERIC Metadata Catalogue

## Next steps

New version of the Catalogue (by the end of 2023)

- Semantic search
- Semantic annotation
- Map features enhancement
- Metrics and statistics
- Versioning mechanism
- Editing tool improvement
- Validation mechanism improvement
- Editorial workflow improvement
- ...





# LifeWatch ERIC EcoPortal

The use of semantic resources (vocabularies & ontologies), giving well defined meaning and understanding, ensure the exchange of information among machines and people.

**Controlled vocabularies (CVs)** play an important role in metadata standards → define the meaning of metadata elements and the allowed values.

Help to find relevant data, or provide information on how to interpret data and reuse it:

- the use of CVs helps to improve the data interoperability by facilitating the data interpretation and harmonization;
- the use of CVs improve the discovery, linking, understanding and reuse of research data within research communities;
- CVs can be used for the (meta)data annotation to comply with the FAIR principles.

## **Main issues:**

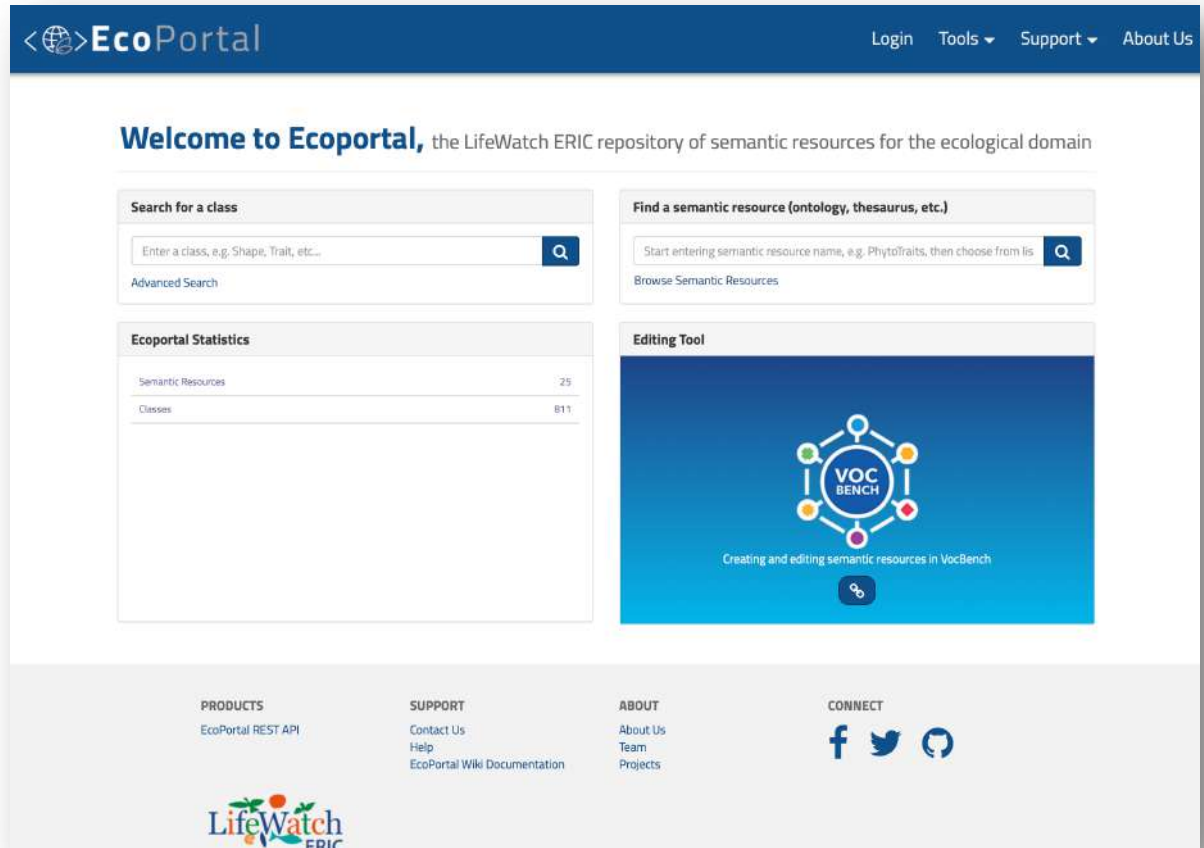
- Lack of semantic artefacts across communities (or with no open licenses and adequate documentation).
- Lack of common reference repositories of semantic artefacts like ontology catalogues .

# LifeWatch ERIC EcoPortal

The LifeWatch ERIC repository of semantic resources (ontologies, vocabularies, thesauri, etc.) for the ecological domain.

A web-based application for supporting the community in the:

- identification and selection of semantic resources;
- creation, management, mapping and alignment of semantic resources in the ecological domain.



The screenshot shows the EcoPortal interface. At the top is a dark blue header with the 'EcoPortal' logo and navigation links: 'Login', 'Tools', 'Support', and 'About Us'. Below the header, a welcome message reads: 'Welcome to Ecoportal, the LifeWatch ERIC repository of semantic resources for the ecological domain'. The main content area is divided into four sections: 1. 'Search for a class' with a text input field and a search button. 2. 'Find a semantic resource (ontology, thesaurus, etc.)' with a text input field and a search button. 3. 'Ecoportal Statistics' showing a table with two rows: 'Semantic Resources' with a value of '25' and 'Classes' with a value of '811'. 4. 'Editing Tool' featuring a blue box with a 'VOC BENCH' logo and the text 'Creating and editing semantic resources in VocBench'. The footer contains four columns: 'PRODUCTS' (EcoPortal REST API), 'SUPPORT' (Contact Us, Help, EcoPortal Wiki Documentation), 'ABOUT' (About Us, Team, Projects), and 'CONNECT' (Facebook, Twitter, GitHub icons). The LifeWatch ERIC logo is centered at the bottom of the footer.

<http://ecoportal.lifewatch.eu>

# The interoperability in LifeWatch ERIC status & challenges

Almost aligned with the EOSC Interoperability Framework.

**Big challenge:** overcome data heterogeneity and allow integrating data across communities.

- Research data available in multiple general-purpose formats (CSV, database dumps, JSON, XML, etc.) or community-based models (Darwin Core, NetCDF, etc.) → hard to align when reusing datasets across communities.
- Semantic interoperability issues also appear (lack of agreement in attributes or column headers, absence of headers or adequate documentation, etc.).
- There should be extensibility options to allow for disciplinary metadata that is typical for some research communities, allowing users/researchers to add annotations according to the established practices of their communities → LifeWatch ERIC Semantic Platform – *work in progress*
- But the possibility to integrate data across communities depends not only on the (meta)data structure and/or format but also on other aspects like the heterogeneity in the adopted tools and methods for collecting and processing data.

---

# Thank you



**Lucia Vaira** | LifeWatch ERIC Service Centre, Lecce – Italy | [lucia.vaira@lifewatch.eu](mailto:lucia.vaira@lifewatch.eu)

---



# Q&A

# Presentation of data architecture and data workflows in BIOSCANEurope

By Rutger Vos, iBOL/BIOSCAN

# DNA (meta)barcoding for biodiversity monitoring

BIOSCAN/iBOL Infrastructural developments

*Sujeewan Ratnasingham & Rutger Vos*

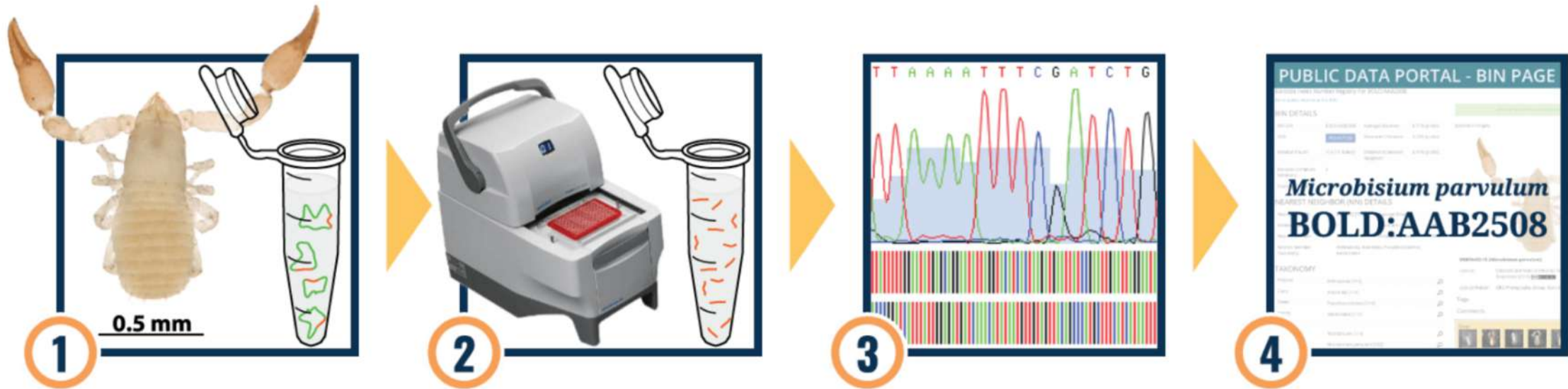
**BIOSCAN  
EUROPE**



international  
BARCODE  
OF LIFE



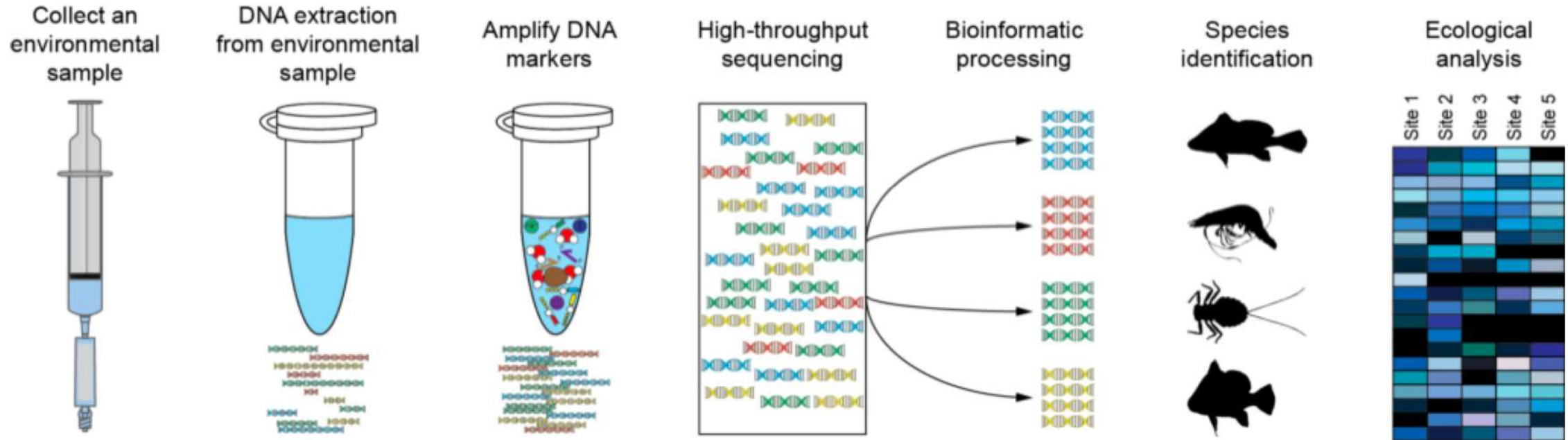
# DNA barcoding



- DNA extraction, amplification, sequencing and processing of (a) specific marker gene(s) from specimens from known species and those from novel species
- Broad data model including specimen records, images, sequences, and provenance
- Results in annotated reference data in which taxonomically-broad patterns of sequence diversity and specific identifications can be investigated
- In existence since Sanger sequencing, scaled up using high-throughput sequencing methods



# Metabarcoding



- DNA extraction, amplification and high-throughput sequencing of mixtures, e.g. soil, water, air, but also gut contents, commercial products, etc.
- Requires annotated barcode reference databases to perform species identification.
- Results in lists of taxonomically identified entities (OTUs, ASVs), which can then be analyzed for alpha/beta/gamma indices of biodiversity.

# Trends

## Producers

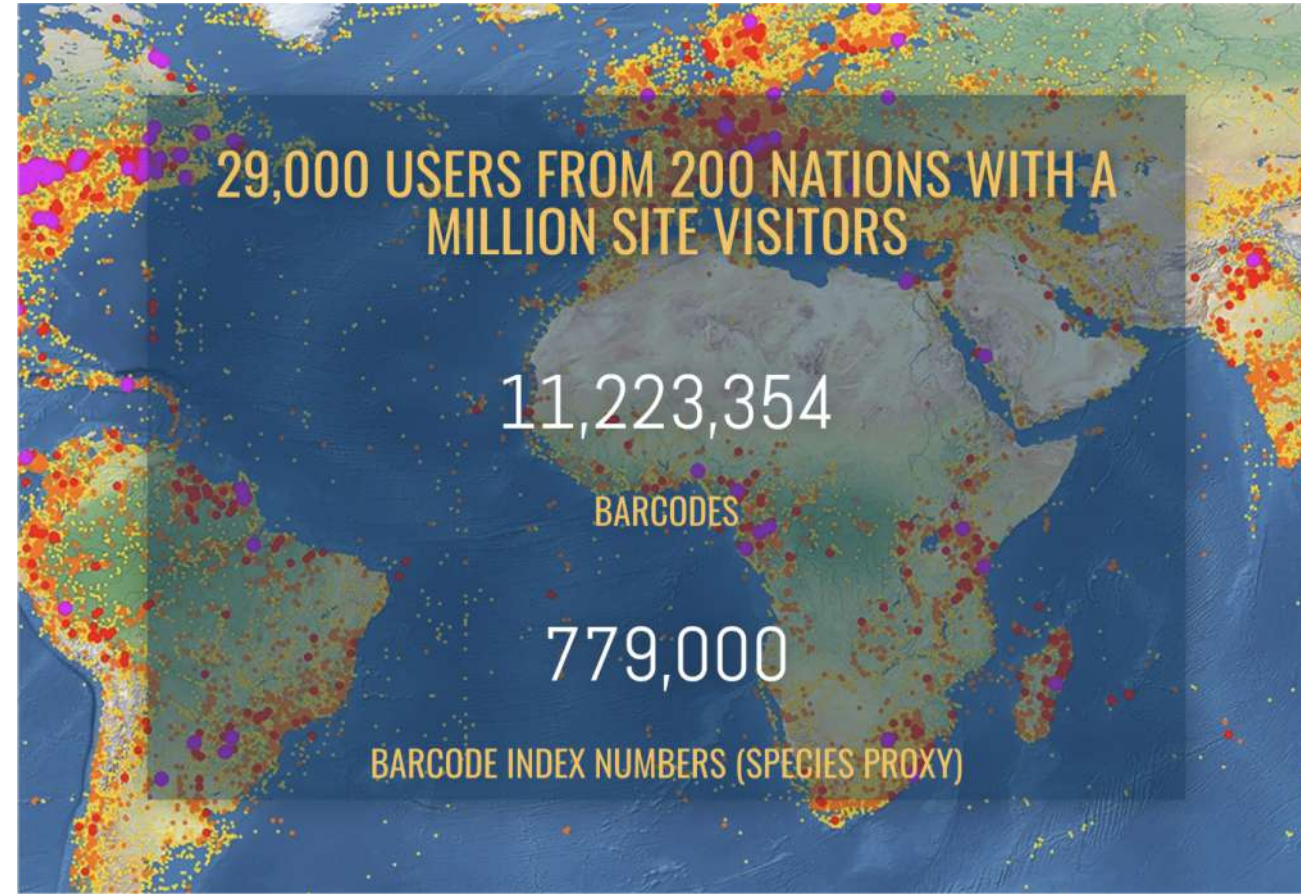
- High-throughput sequencing (HTS) allows for broader taxonomic coverage at low cost compared to Sanger
- Using HTS for barcoding results in the production of multiple barcodes per specimen
- Most barcodes are produced through institutional work processes rather than *ad hoc*, small-scale studies by individual researchers.

## Consumers

- Large-scale, integrative projects (e.g. UNITE, ARISE) seek to harvest well-annotated sequences at scale
- Law enforcement
- Rising demand from corporations
- FAIR data considerations gain importance

# iBOL

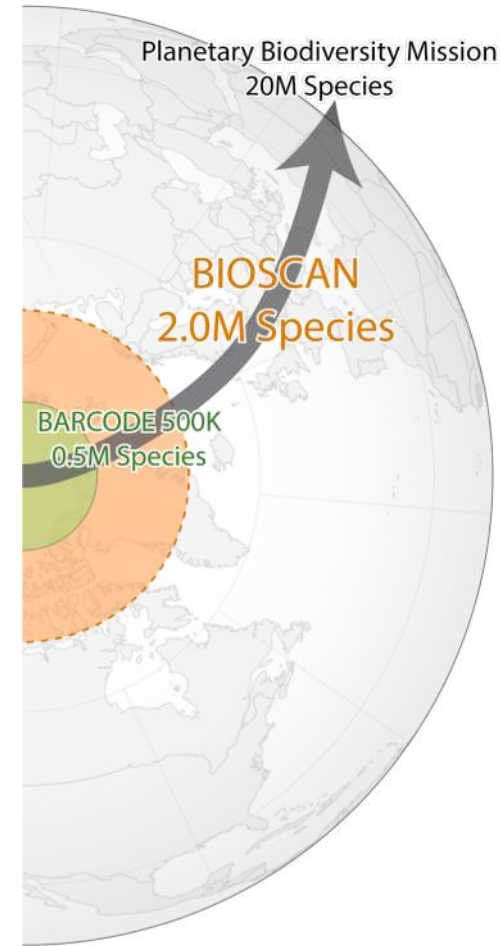
- Research alliance of nations
- Established in 2008
- Builds:
  - barcode reference libraries
  - data and process standards
  - steering and capacity building
  - international collaboration
  - **informatics platforms**



# BioScan

- Launched in 2019
- Focus on building baselines for biomonitoring
- Establishes the European node of iBOL
- Complementary with genome skimming (Biodiversity Genomics Europe)

**BIOSCAN**  
**EUROPE**





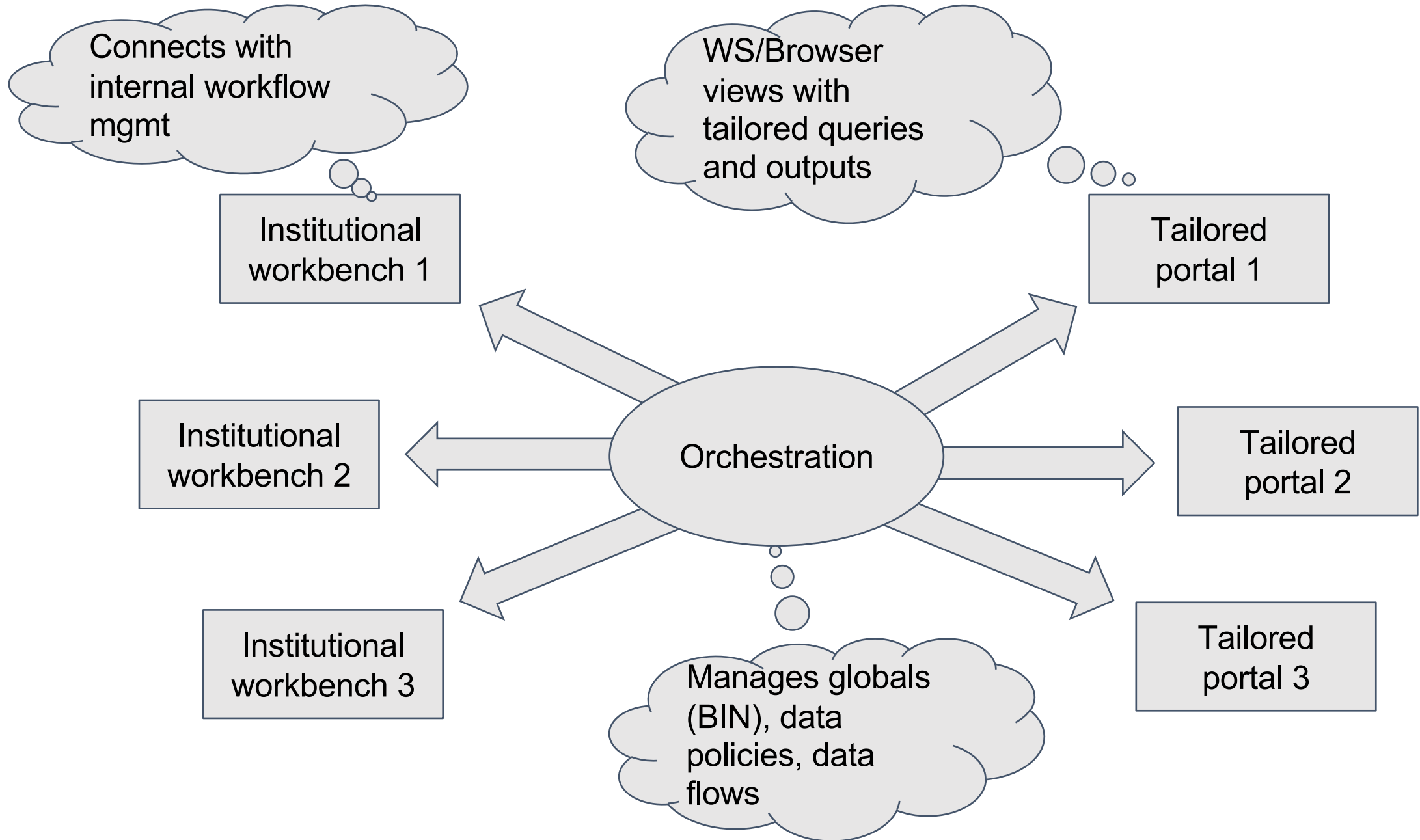
# Towards a decentralized infrastructure for DNA barcodes

## **Back-end/producer considerations**

- Higher data volumes make manual submissions to repositories cumbersome
- More of the preprocessing workflow is automated, e.g. in LIMS and workflow mgmt tools
- Need for tailored, automated submission infrastructure

## **Front-end/consumer considerations**

- Larger-scale initiatives in data reuse, integration and curation
- Data collection moving from browser-based/manual to APIs
- Need for curated collections and data packages
- Need for tailored queries, with provenance and versioning



# Implementation

## Planning

### **Q4-2022..Q3-2023**


- Initial clone of BoLD deployed
- Initiation of OS code dev
- Development of governance
- Technical roadmap

### **Q4-2023-Q2-2026**

- Roadmap implementation

## Participants

- **CBG** - data producer
- **Naturalis** - data producer
- **EBI** - consumer
- **Unite** - consumer
- **Elixir** - community steward
- **RBGE** - community steward

A long-exposure photograph of a waterfall in a lush forest. Sunlight rays stream through the trees, creating a magical atmosphere. The text "Thank you for your attention" is centered in white.

Thank you for your attention



# Q&A

# Open discussion session - How Biodiversa+ can help achieve effective data interoperability

# Objectives of this session

By Hilde Eggermont, BelSPO, Biodiversa+ Chair & Coordinator

# Objectives

- Identify barriers preventing data interoperability
- Factors encouraging data interoperability
- Possible support from Biodiversa+



# Support from Biodiversa+

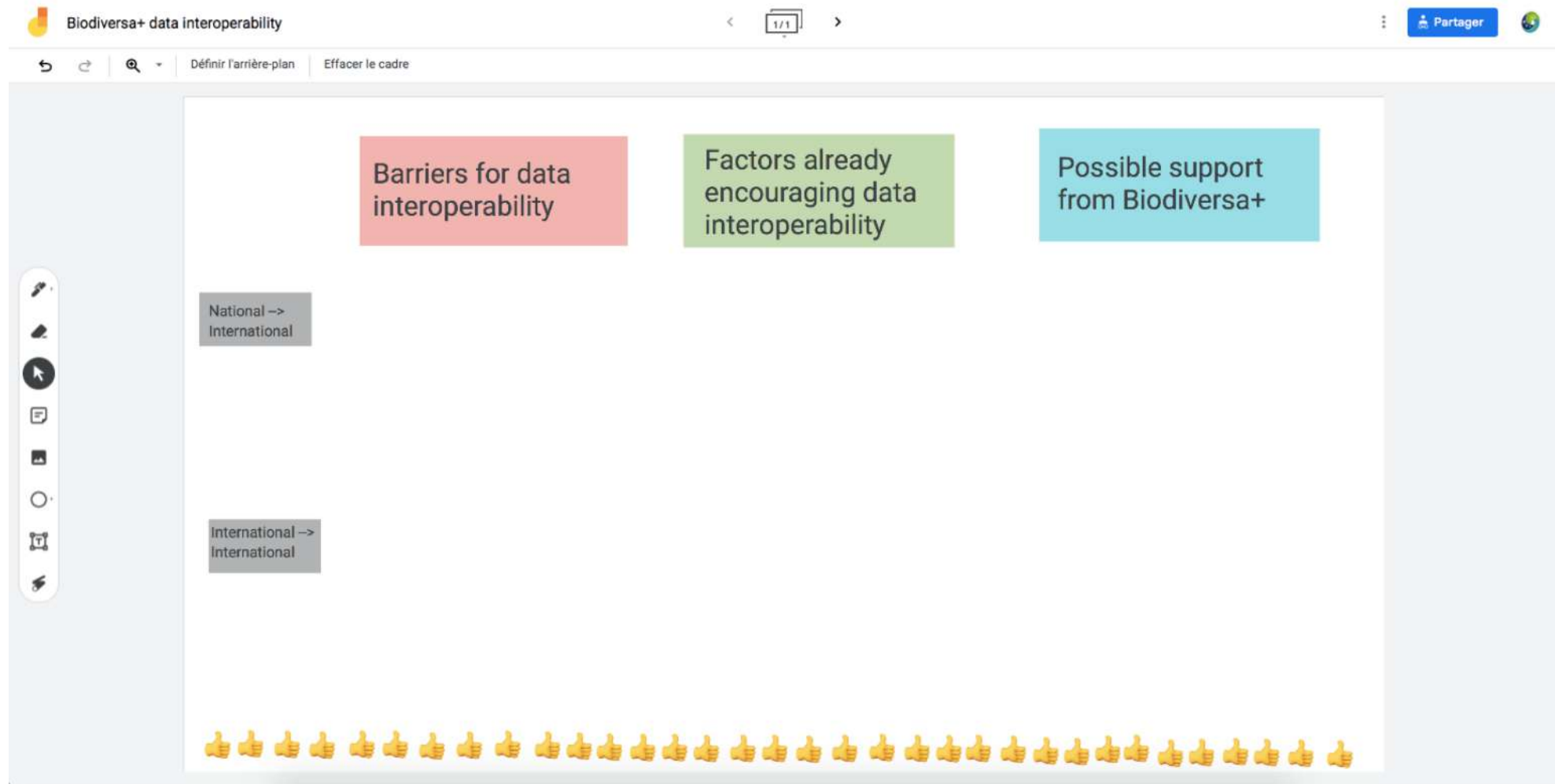
- ✓ Enriching the EuropaBON work in relation to description of data flows (from raw observation, to indicators, to multiple uses);
- ✓ Mapping of (sub)national data architectures => identify best practices;
- ✓ Promoting best practices through capacity building activities;
- ✓ Promoting open data and standards for exchanging information, in line with the EU digitization strategy;
- ✓ Promoting use of European and global research infrastructures; ...

# How to use Jamboard?

By Cécile Mandon, FRB, Biodiversa+ Officer



# 1. Click on the link we will share in the chat and join the board!



# Add post-its

The screenshot shows the Biodiversa+ data interoperability Jamboard interface. The title bar includes the Biodiversa+ logo, the text "Biodiversa+ data interoperability", navigation icons, and a "Partager" button. The main workspace contains several sticky notes: a red one titled "Barriers for data interoperability", a green one titled "Factors already encouraging data interoperability", and a light blue one titled "Support in Biodiversa+". A sidebar on the left contains various tool icons, with the "Add Post-it" icon (a notepad) highlighted by a red box and a red arrow pointing to it from the instruction "1. Click here to add a post-it". A yellow sticky note titled "Pense-bête" is in the center, containing the text "Test" and "2. Writes here what you want to say". A red box surrounds this note, with a red arrow pointing to it from the sidebar icon. A teal box with the instruction "3. Select a colour for your post-it" has a teal arrow pointing to the color selection icons (yellow, green, blue, pink, orange) at the top right of the yellow note. A blue box with the instruction "4. Save you post-it and publish it on Jamboard" has a blue arrow pointing to the "Enregistrer" button at the bottom right of the yellow note. The bottom of the workspace features a row of thumbs-up emojis.

1. Click here to add a post-it

3. Select a colour for your post-it

2. Writes here what you want to say

4. Save you post-it and publish it on Jamboard



# Move your post-it

1. Click here
2. Select your post-it by clicking on it
3. Move your post-it on the board

The screenshot shows a web-based collaborative workspace titled "Biodiversa+ data interoperability". The interface includes a top navigation bar with a logo, a title, and a "Partager" (Share) button. Below the navigation bar is a toolbar with icons for undo, redo, search, and zoom. The main workspace area contains three large colored boxes: a red box labeled "Barriers for data interoperability", a green box labeled "Factors already encouraging data interoperability", and a blue box labeled "Possible support from Biodiversa+". To the left of these boxes is a vertical toolbar with icons for erasing, moving, and creating post-its. A red square highlights the move icon (a hand cursor). Below the main workspace area is a row of yellow thumbs-up icons. The workspace also features a sidebar on the left with a search icon and a "Définir l'arrière-plan" (Set background) button, and a bottom bar with an "Effacer le cadre" (Erase frame) button.



# interoperability issues to be tackled

Use of new technologies to tackle not only interoperability but also their use and impact

Semantic artefact development requires people who know not only one Institution's data and metadata architectures, but multiple ones. Exchange of experts for short period of times across organisations

Best practices in semantic artefact development/management/governance for establishing guidelines.

Proliferation of metadata standards and paradigms. Hard to make the "correct" choice.

Specific funding and support for standards development are needed (TDWG, RI etc)

Can propose best practices for leveraging funds for the modest but non-negligible costs of standardising and sharing data, otherwise this effort is left out of budgets and plans

Adherence to the relevant standards is required in funding programs, including guidance and a list of standards (for information purposes, the list may not be comprehensive)

## Possible support from Biodiversa+

Much existing data is collected in templates that lose the raw data, e.g. for EU reporting.

There is no clear guidance/mandate from funding agencies to adhere to specific standards that facilitate interoperability

Define clear tasks and goals for the different organisations and align these (exactly - who is doing what)

There will continue to be a proliferation of metadata standards. The issue is how to translate the information between standards

facilitate integration across sectors (agri, forest, nature, water etc.) to create co-benefits and improve cost effectiveness also in line with global and EU policy targets

harmonise data and methods across the heterogeneous research landscape and link with national, regional (EU) and global institutions

Help documenting the evolving landscape incl. key standards (decision tree?), and continue promoting interoperability and integration

Biodiversa+ could coordinate the interaction between all these organisations/initiatives to build something big that works

Always needed: a one stop shop for collecting the relevant information needed in terms of monitoring initiatives and FAIR data (e.g. standards, protocols)

Machine observations have large data storage needs and IT infrastructure is not keeping up

Creation / application of standards (e.g. metadata); provision of a central register for those standards (one stop shop)

Although we have provenance tools, such as ORCID, ROR, DOIs etc they are still not being consistently used by the community

Proliferation of ontologies which are not easy to cross-walk. Requires initial human investigation and sense-making to build machine-readable repos.

Mandating data management plan and the use of certain data management strategies within Biodiversa+ project

Define "playbooks" for data management at different stages of the cycle (collect, analyze, publish, etc.)

data interoperability and semantic meaning (mentioned by Life Watch): foster this to allow interdisciplinary use

actor together with the EU science service in make (BioAgora) and the KCBD to see how these infrastructures and biodiv monitoring (Monitoring centre on Biodiv on European level?) could be a pillar under the

From the management perspective it is important to see what dataflows are/can be channeled into official reporting mechanisms

Integration of machine learning results, with human interpreted results has no standards

Machine learning processes are not always transparently documented (e.g. training sets)

Data collected by different agencies within government apply different standards - so biodiversity data interoperability can be seen as part of the process of mainstreaming.

A lack of harmonization on what to monitor

common guidelines to facilitate interoperability between monitoring data from the management and research infrastructures

## Establish common best practices and guidelines

Awareness/ increase the understanding of existing tools / standards (uptake at the relevant levels)

Metadata review and validation protocols to ensure the quality and consistence of information workflows

Concepts such as interoperability, ontologies, semantics etc are still unfamiliar to many data collectors

facilitate implementation of domain specific standards and standards which can be common across sectors

Too little attention to capacity building in the poorer areas of Europe

capitalise on increasing profile of / evolving techniques for 'explainable AI' with biodiversity- and provenance-specific funding calls

Identification of the minimum metadata sets of information that can be translated in all the major existing metadata schemas

provide knowledge on existing methods and standards for harmonisation and data interoperability to key actors and stakeholders (technical knowledge)

Biodiv+ can provide expertise in data management and interoperability



# Conclusions of the workshop

By Alberto Basset, MUR and Hilde Eggermont, BelSPO, Biodiversa+ Chair and Coordinator



**biodiversa+**  
European Biodiversity Partnership

EUROPEAN PARTNERSHIP



Co-funded by  
the European Union

**Thank you!**



[www.biodiversa.org](http://www.biodiversa.org)



[contact@biodiversa.org](mailto:contact@biodiversa.org)



BiodiversaPlus

