Open data in Biodiversity Sciences: The 20 year journey of the Global Biodiversity Information Facility

Francisco Pando
GBIF Spain - CSIC

Centro Internacional de Encuentros Matemáticos de la Universidad de Cantabria, Castro Urdiales. 12 November 2022

Workshop: "Open Science, a landscape under construction with a horizon of possibilities"

CIEM
Centro Internacional de Encuentros Matemáticos

UC
UNIVERSIDAD DE CANTABRIA

Gbif.es

GOBIERNO DE ESPAÑA

CSIC
CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS

GBIF

Summary

- What GBIF is
- What GBIF does
- GBIF in Spain
- Data in GBIF
- Some thoughts

GBIF- Global Biodiversity Information Facility

--Infraestructura Mundial de Información sobre Biodiversidad--

- Independent intergovernmental organization
- Established at the instances of the OECD
- Supported by the country members
- Created from the science, for the science…and the society; to support biodiversity science, environmental research and evidence-based decision making
- Since 2001
- “Free and open access to biodiversity data”… from all to all

GBIF.es

www.gbif.es

GBIF Global Biodiversity Information Facility

www.gbif.org

41 VOTING PARTICIPANTS | 23 ASSOCIATE COUNTRY PARTICIPANTS | 43 OTHER ASSOCIATE PARTICIPANTS

2,249,497,561
Occurrence records

79,097
Datasets

1,926
Publishing institutions

7,918
Peer-reviewed papers using data

115.0 billion
Average records served per month (2022)

214,876
Avg monthly user sessions (2019)

https://datos.gbif.es/ https://www.gbif.es/ https://www.gbif.org/

GBIF in Spain

MINISTERIO DE CIENCIA Y TECNOLOGÍA

MINISTERIO DE EDUCACIÓN Y CIENCIA

GOBIERNO DE ESPAÑA | MINISTERIO DE ECONOMÍA Y COMPETITIVIDAD

GOBIERNO DE ESPAÑA | MINISTERIO DE CIENCIA E INNOVACIÓN

BOE núm. 166

13959 *RESOLUCIÓN de 12 de junio de 2002, d*
Estado de Política Científica y Tecnológ

CSIC
CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS

Consejo Superior de Investigaciones Científicas

Real Jardín Botánico

Museo Nacional de Ciencias Naturales

Gbif.es

Nodo Nacional de GBIF:
Unidad de coordinación + red de centros

TRAINNING

Alrededor de 150 talleres. Materiales disponibles en línea:

Talleres: http://www.gbif.es/formacion.php

Vídeos: http://www.gbif.es/videos/videos.php

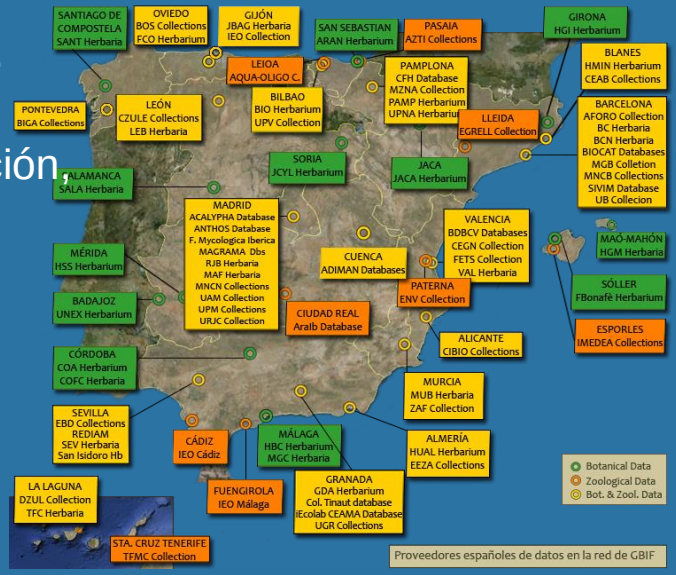
Temás: Georreferenciación - Toma de datos con dispositivos móviles
Modelos de nicho ecológico - Fichado de datos - Calidad de datos digitales -cibertaxonomía

Talleres: presenciales / en línea;
nacionales / internacionales




DATA INFRAESTRUCTURA

Alojamiento de datos e imágenes
Plataforma de publicación recursos
Asesoramiento



PORTALS

Portal de comunicación

http://www.gbif.es

Portal de datos nacional

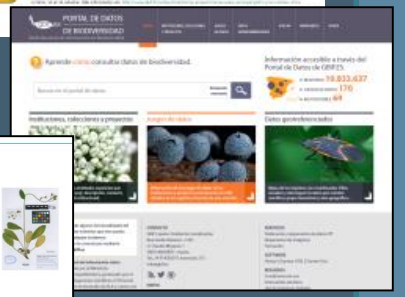
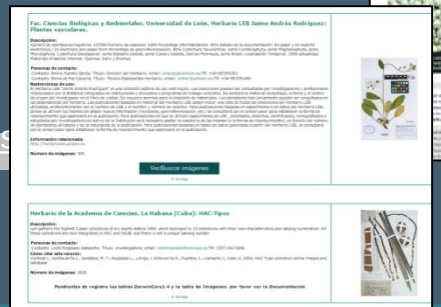
http://datos.gbif.es

Plataforma IPT para publicación de datos

http://www.gbif.es:8080/ipt/

Portal de imágenes

http://www.gbif.es/Imagenes

SOFTWARE

Calidad de datos: DARWIN TEST
Herramienta de validación

http://www.gbif.es/darwin_test/

Gestión de colecciones:

ELYSIA y HZL

Herramientas de gestión de colecciones:

http://www.gbif.es/zoorbar/

http://www.gbif.es/herbar/

http://www.gbif.es/h

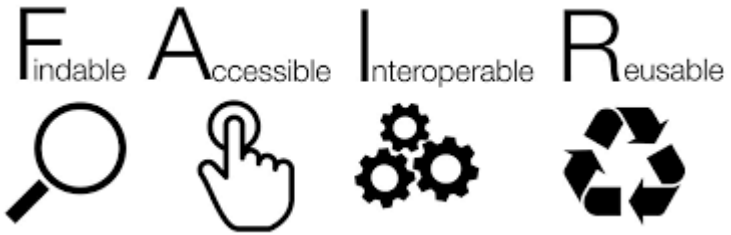
GEOLocate

Elysia Gestor de colecciones biológicas

HZL Herbar/Zoorbar Ligero

GBIF is an innovator and a reference

- A science initiative conceived as a research infrastructure since 2001
- Virtual, distributed
- Unified access, but data providers are credited and in control
- Data and metadata are standardized and fully indexed
- Pioneer and promoter of data papers
- DOIs for data sets and downloads since 2015
- Data under the FAIR* principles (since before they were invented)



Elements to make a difference

Unique and persistent identifiers that provide traceability and metrics

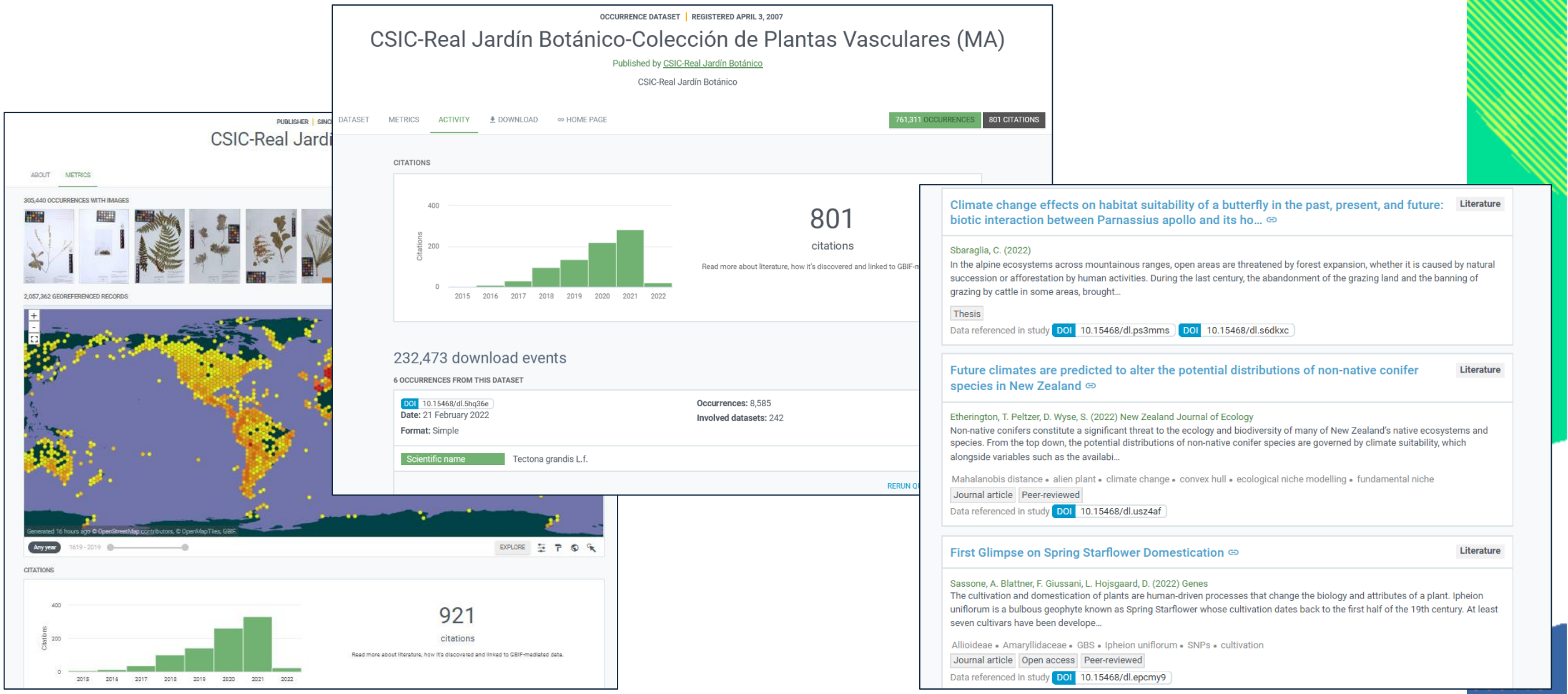
Standardized data and metadata

Standardized use licenses to facilitate data reutilization

Deep linking to any content and result instead of taking users to a generic page (such as the home page)

https://registros.gbif.es/occurrences/search?q=Tetraclinix+articulata#tab_mapView

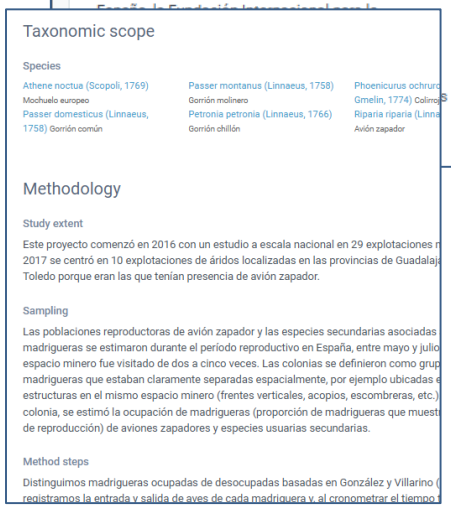
DOIs (identifiers) ⇒ GBIF: credit, impact relevance



https://www.gbif.org/resource/search?contentType=literature&gbifDatasetKey=834c9918-f762-11e1-a439-00145eb45e9a

Data and Metadata Standardization

Metadata (EML)

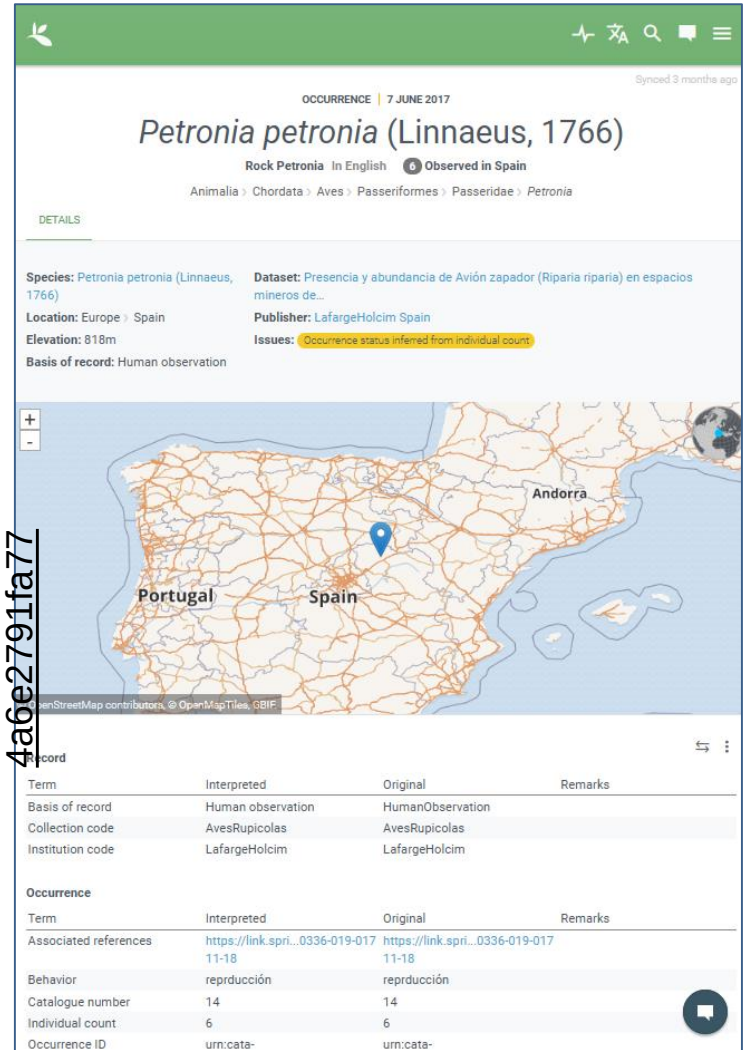


https://doi.org/10.15470/wffiti

Datasets/ records (Darwin Core)

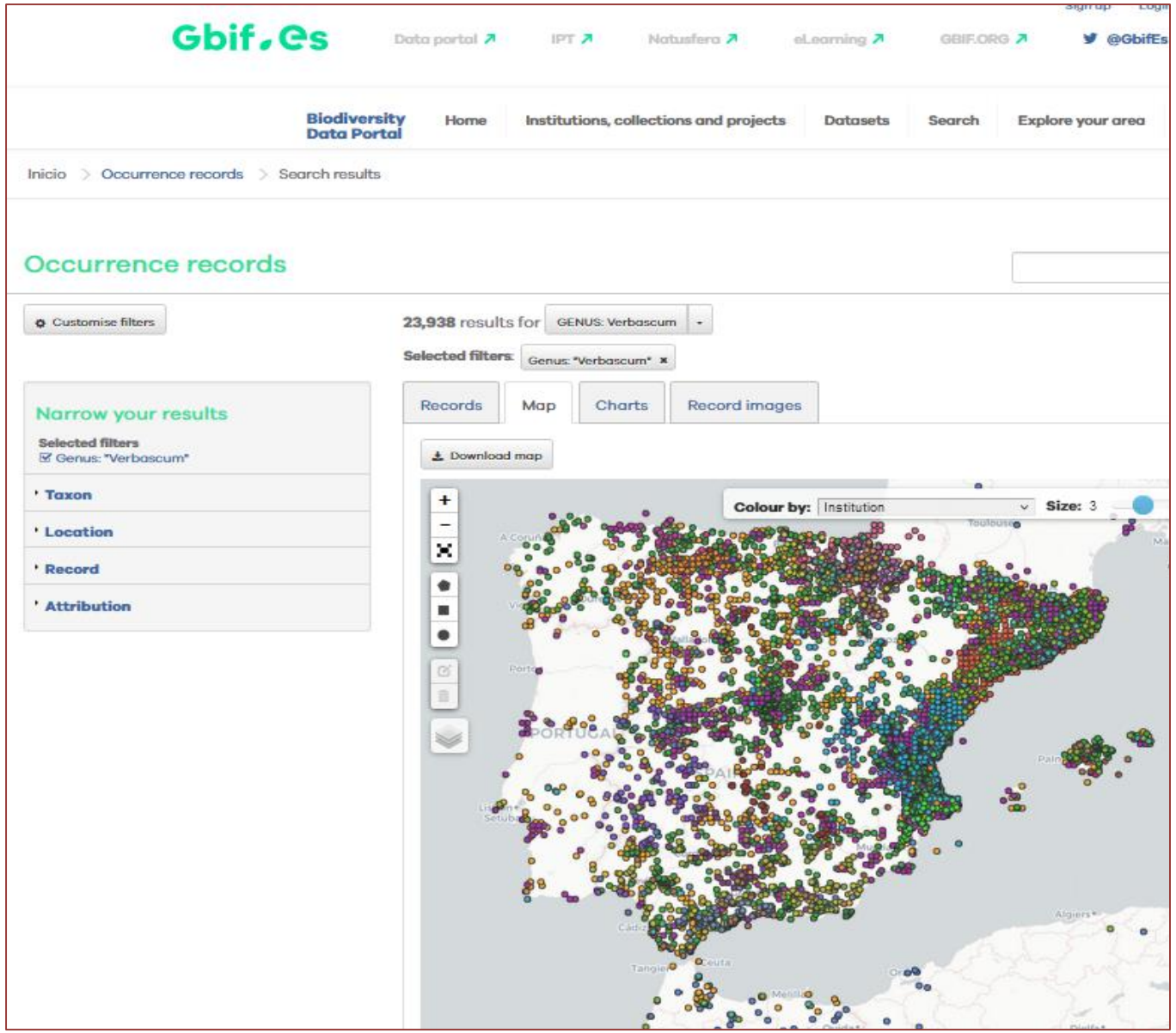
https://www.gbif.org/dataset/34b5f4ce-bbbe-43b8-a82e-4a6e2791fa77

https://www.gbif.org/occurrence/search?dataset_key=34b5f4ce-bbbe-43b8-a82e-4a6e2791fa77



https://www.gbif.org/occurrence/2465765583

How it works, how it is used

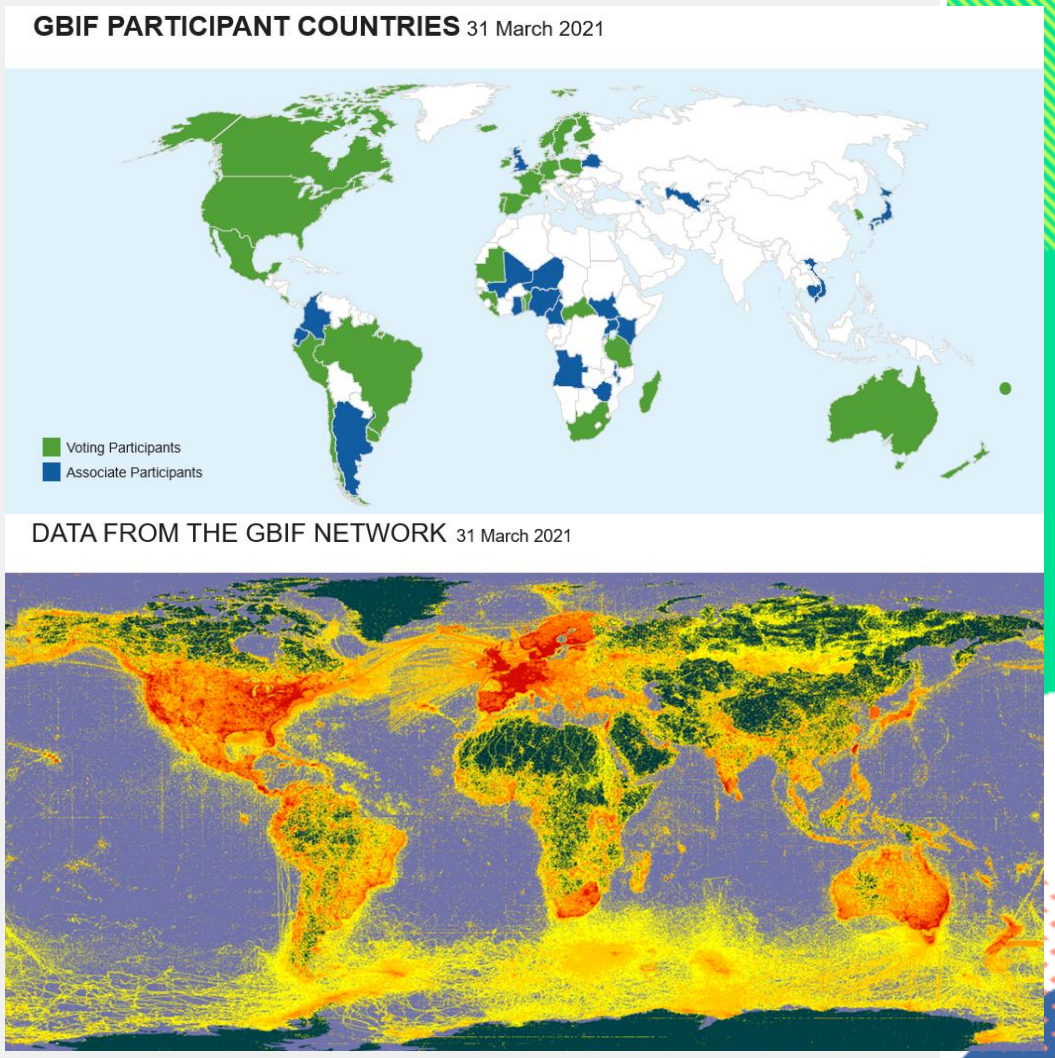


https://registros.gbif.es/occurrences/search?taxa=verbascum&submit=Submit%20Query&lang=en#tab_mapView

GBIF is FAIR data

Data: biological species occurrences, species check-lists, sample based datasets

- Findable?
  - Indexed: data & metadata
  - Standardized: data & metadata
- Interoperable?
  - Inside network: full
  - Outside network: limited
  - SPARQL endpoint: no
- Accessible?
  - Cached: yes
  - Downdable: yes
  - Visualizations: yes
- Reusable?
  - Downloadable: yes
  - Licenses schema: yes
  - DOIs for datasets & downloads
  - UUIDs for records: uneven implementation



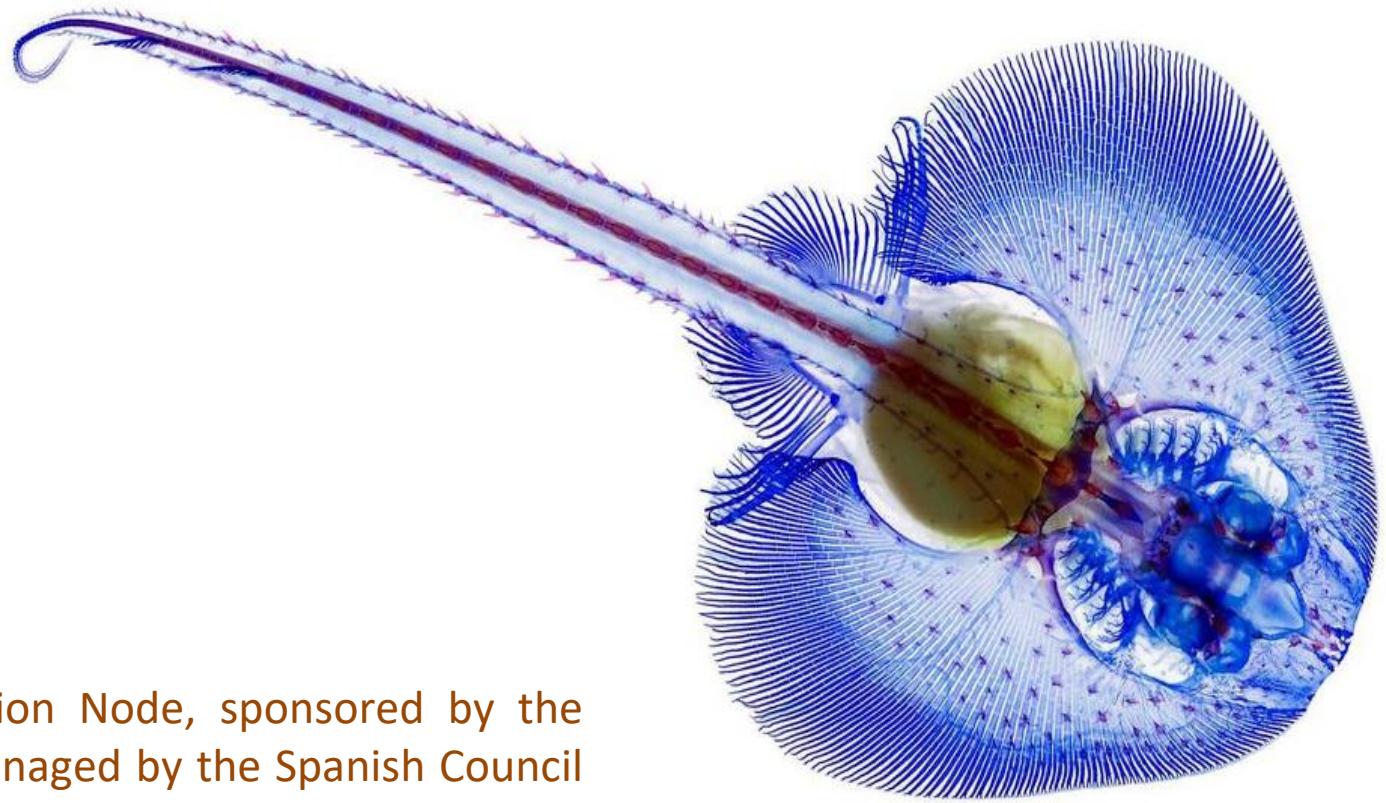
Some thoughts, maybe some lessons

- There are not “GBIF data”; there are data from 79.000+ datasets ; each of these with is own characteristics, methodologies, scopes, etc.… and authors
- This is a long-run game. Training, support, dissemination, mainstreaming,… need to be done as permanent operations, not as campaigns or projects
- Go for the people, not for the data
- Biodiversity goes so much beyond science. That is a challenge, it is also beautiful

Francisco Pando

GBIF.ES, Unidad de Coordinación
CSIC
Joaquín Costa, 22
280o2 Madrid, Spain

pando@gbif.es



https://www.smithsonianmag.com/science-nature/true-inner-beauty-fishes-180949391/

GBIF-ES is the National Biodiversity Information Node, sponsored by the Spanish Ministry of Science and Innovation, managed by the Spanish Council for Scientific Research (CSIC). https://www.gbif.es/ , https://datos.gbif.es, https://elearning.gbif.es

MINISTERIO DE CIENCIA E INNOVACIÓN

CSIC

Gbif.es