

# Publicación de datos en GBIF y portal del publicador (ipt.gbif.es): Introducción

Katia Cezón  
GBIF.ES, Unidad de  
Coordinación - CSIC

Jornada GBIF.ES 2023:  
Real Jardín Botánico – CSIC.  
Madrid, 18 de abril de 2023

Gbif.es



GBIF

# Contenidos

- ¿En qué consiste la publicación de datos de biodiversidad?
- ¿Por qué publicar datos de biodiversidad?
- ¿Qué datos se pueden publicar?
- ¿Cómo publicar?



# Biodiversidad

## Datos primarios de biodiversidad

A los textos digitales o registros de datos multimedia que detallan hechos acerca del **qué, dónde, quién, cómo** y **cuándo** de la aparición y el registro correspondiente (GBIF, 2011).



A los datos crudos (sin procesar) que los investigadores generan como parte de su investigación.

**QUÉ** Datos taxonómicos y nomenclaturales

**DÓNDE** Datos espaciales

**QUIÉN** Datos colecta

**CUÁNDO** Datos temporales

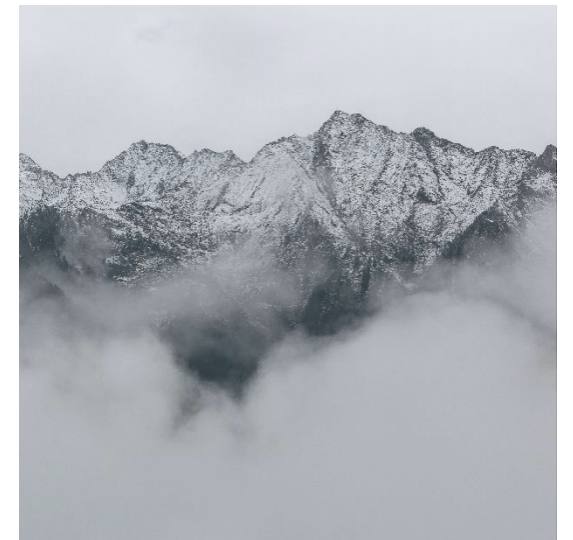
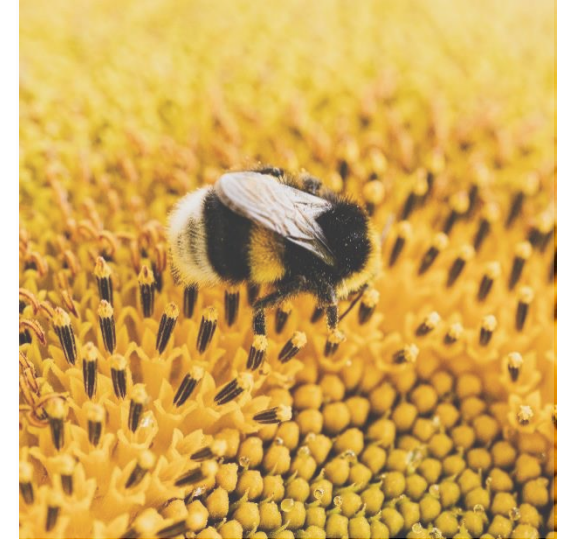
**CÓMO** Datos descriptivos



# Biodiversidad

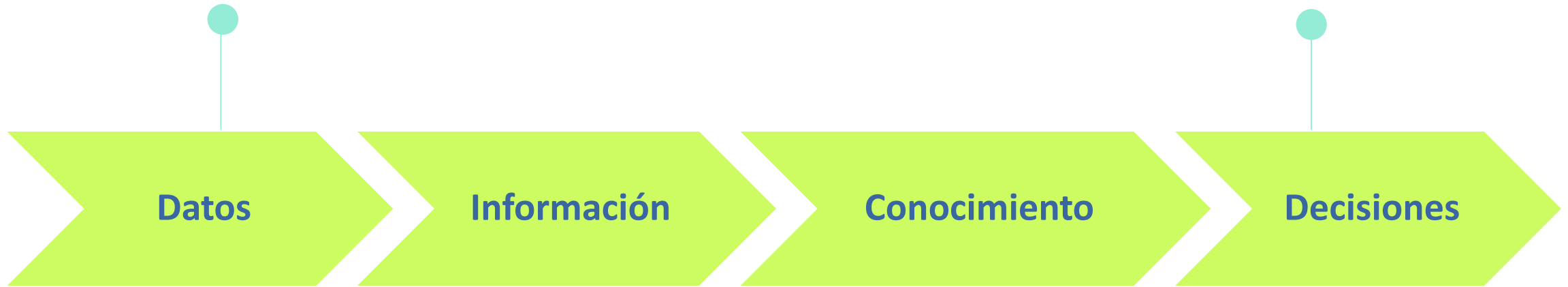
## Uso de los datos de biodiversidad

- Agricultura, montes y pesca
- Perspectivas basadas en productos biológicos
- Salud y seguridad públicas
- Medicina forense
- Ecoturismo
- Planificación de infraestructuras humanas
- Etc.



# Biodiversidad

Los datos como base para el conocimiento

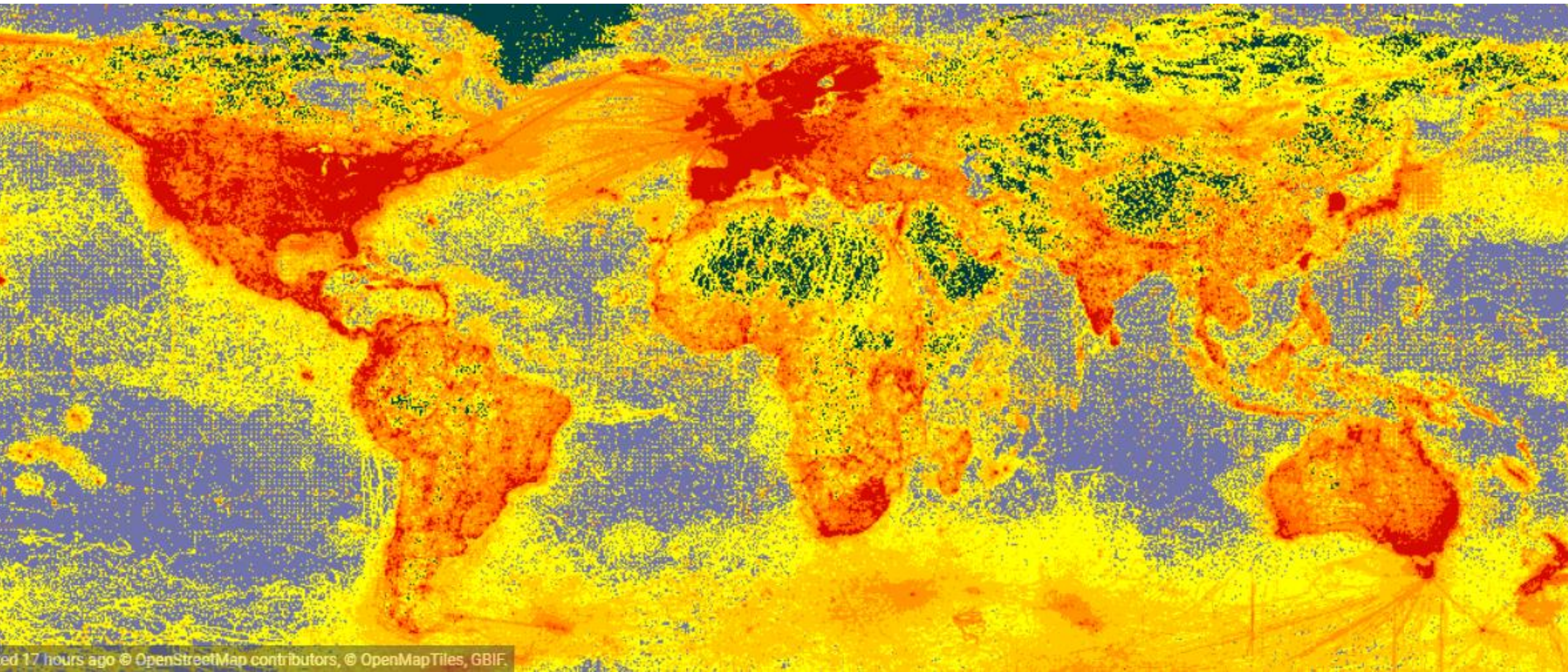


Necesitamos un universo de datos lo  
suficientemente representativo y de calidad



# GBIF

<https://www.gbif.org/> marzo 2023





# Biodiversidad

Una preocupación global

2.294.227.595

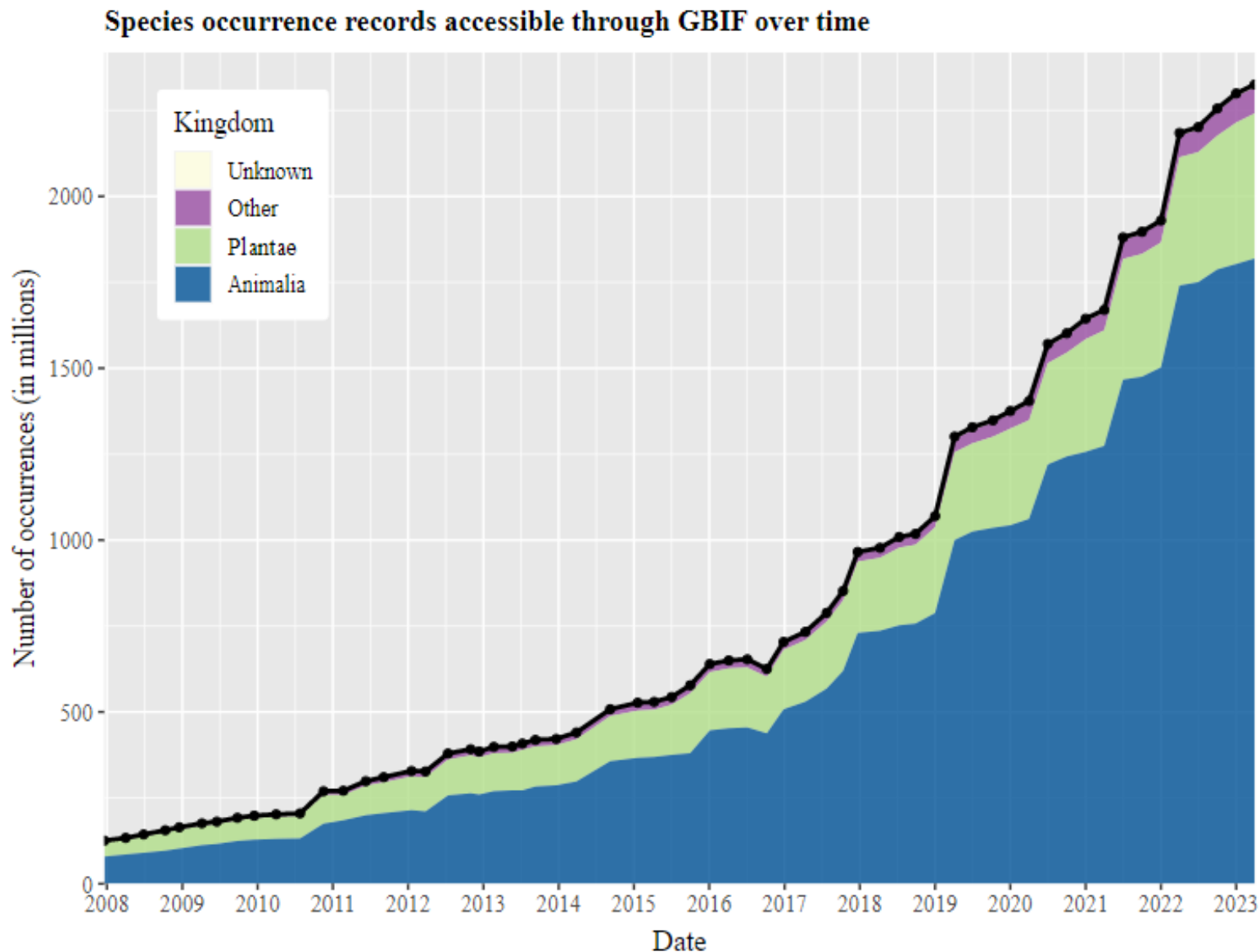
Registros

83.927

Conjuntos de datos

2003

Instituciones



<https://demo.gbif.org/analytics/global> (04/2023)

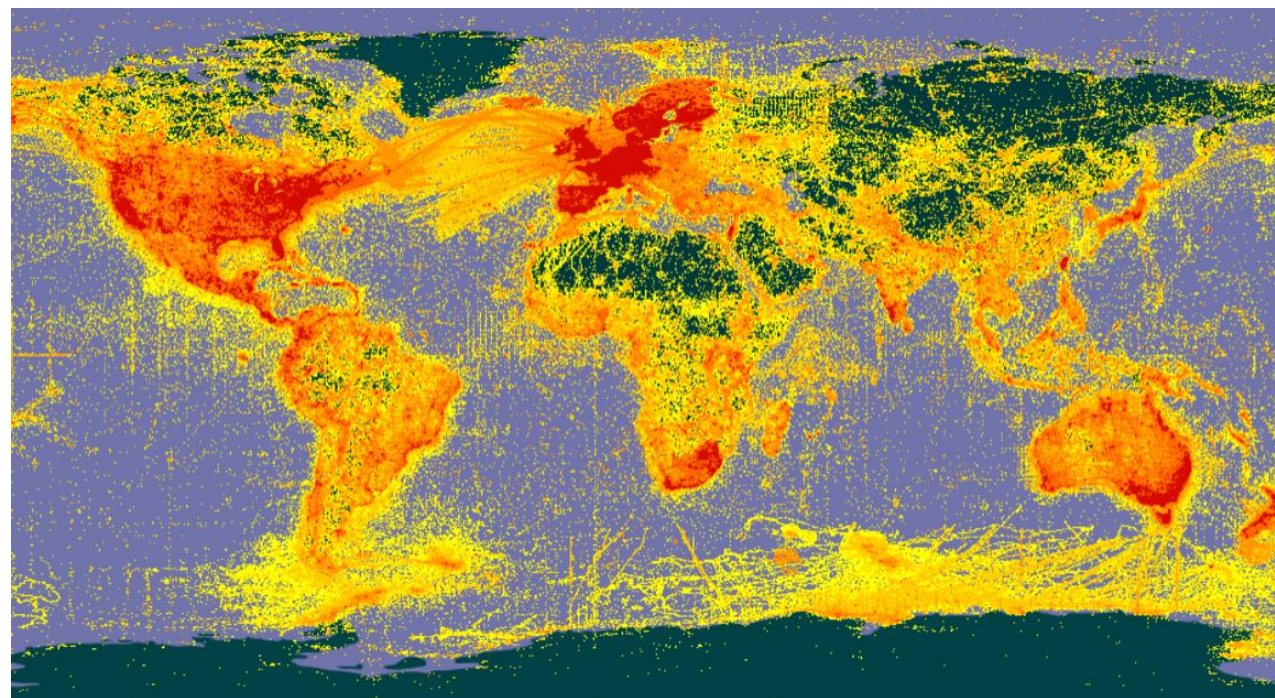
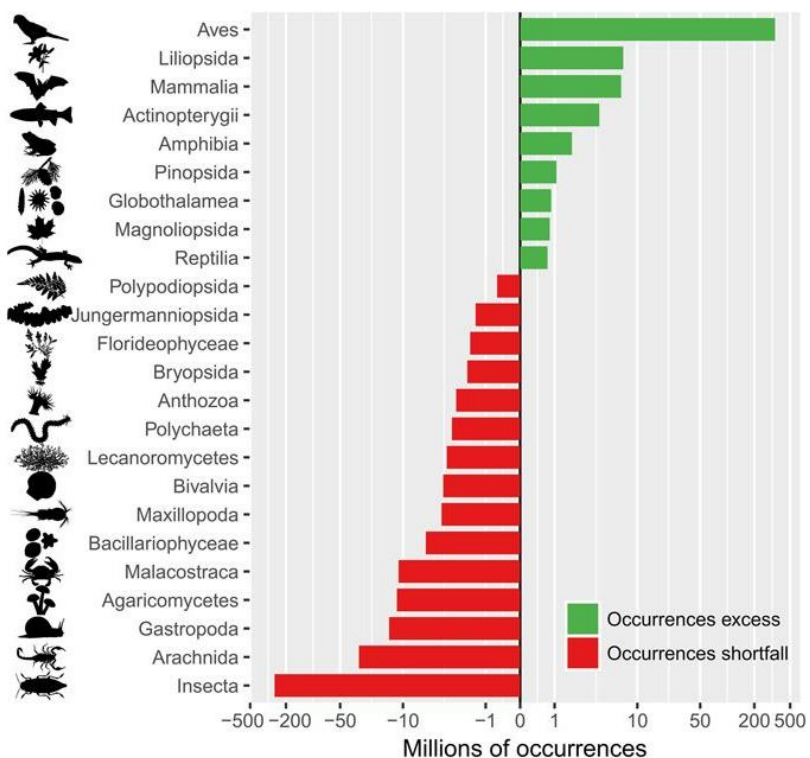
The background of the slide features a dense pattern of green fern fronds. The fronds are bright green and have a feathery, segmented appearance. They are set against a solid blue background that has a subtle gradient, being slightly darker at the top and bottom edges. The ferns are oriented in various directions, creating a sense of depth and texture.

# Problemas



# El problema de los datos

Sesgos taxonómicos, espaciales y temporales

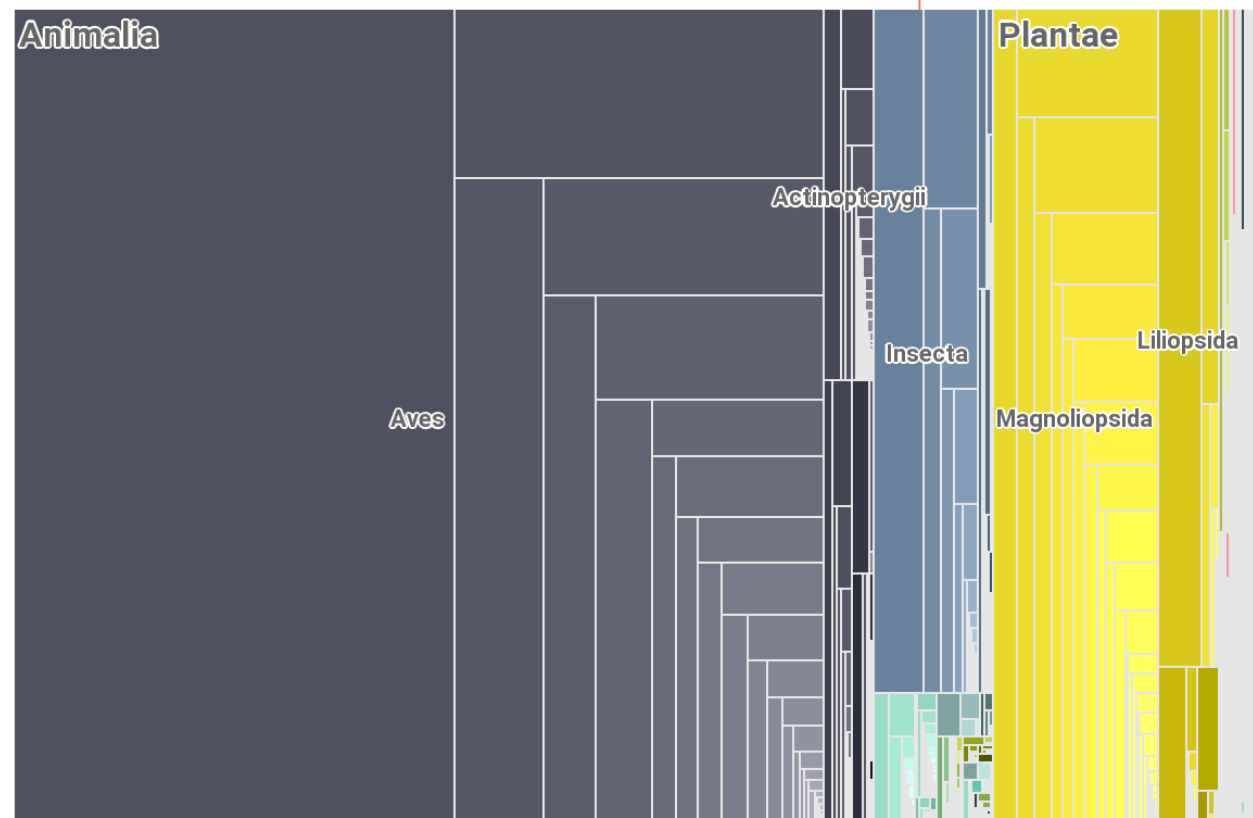
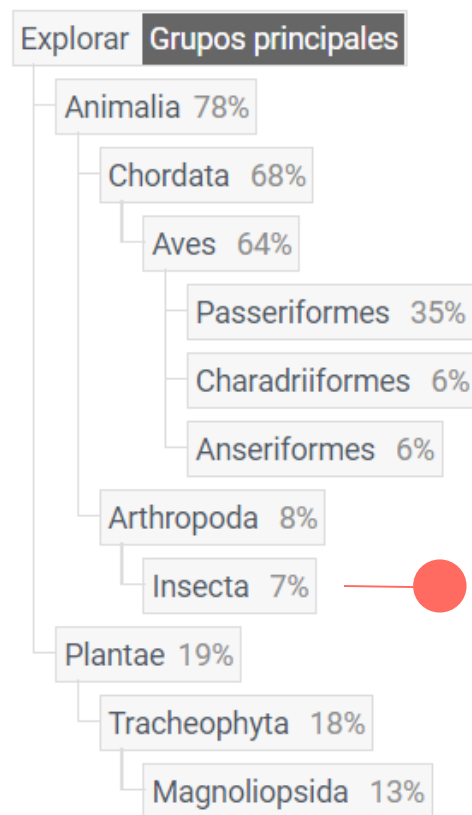


*Taxonomic bias in biodiversity data and societal preferences*  
doi:10.1038/s41598-017-09084-6

<https://www.gbif.org/> Marzo 2023

# El problema de los datos

Sesgos taxonómicos, espaciales y temporales



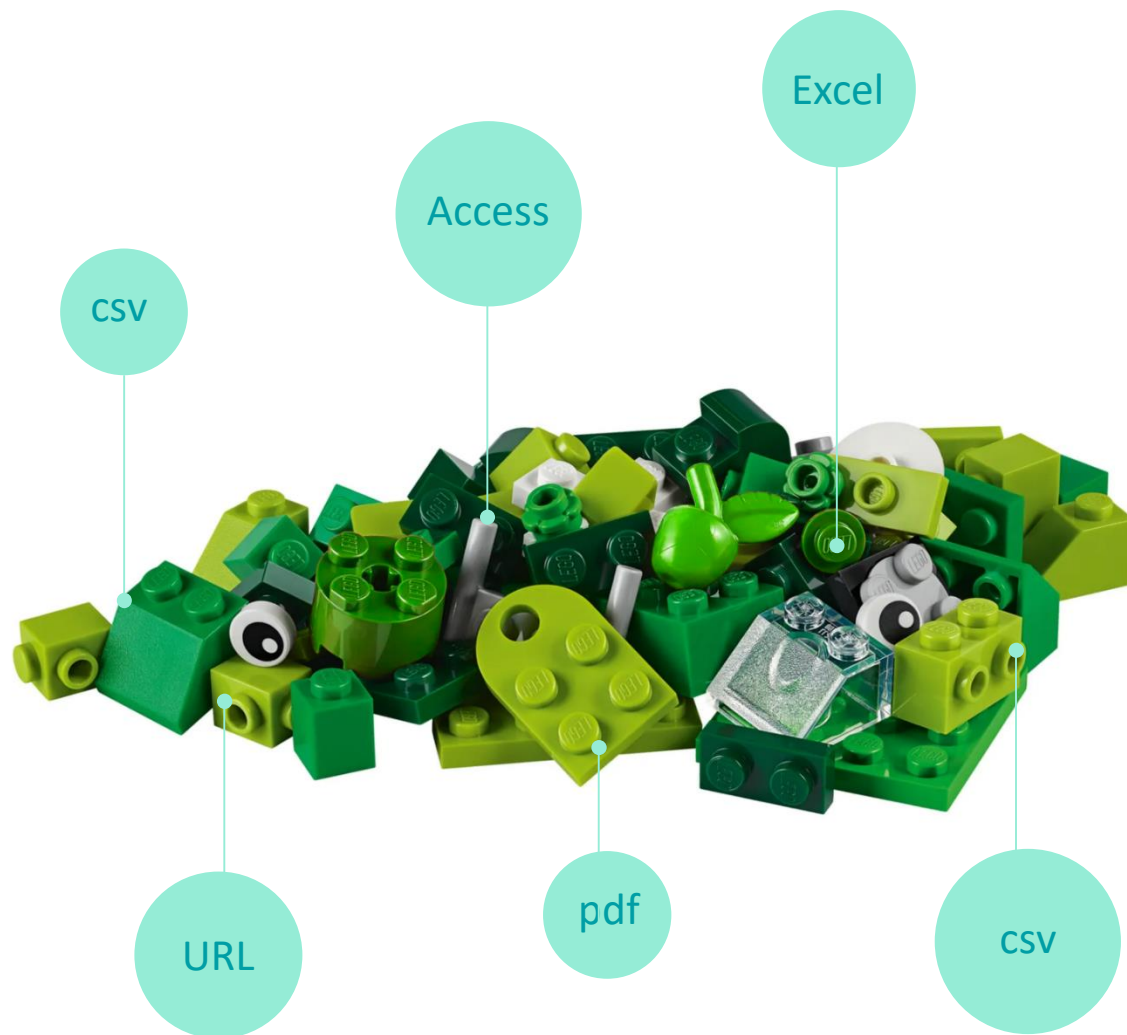
Distribución taxonómica de los registros de presencia en GBIF.org

[https://www.gbif.org/occurrence/taxonomy?occurrence\\_status=present](https://www.gbif.org/occurrence/taxonomy?occurrence_status=present) sept 2021



# El problema de los datos

Accesibilidad



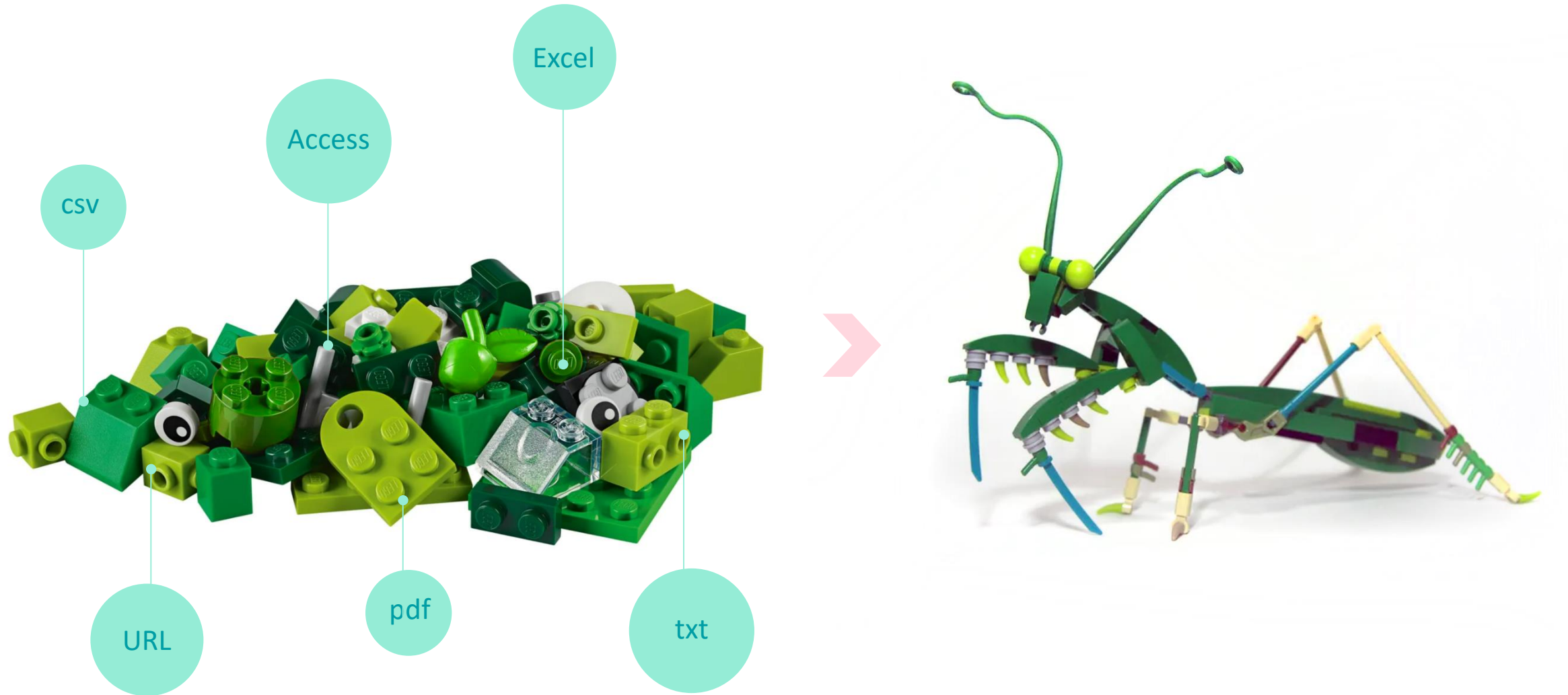
Repositorios de **datos aislados**

Datos en **diversos formatos**

Ausencias de **estándares y protocolos**

# El problema de los datos

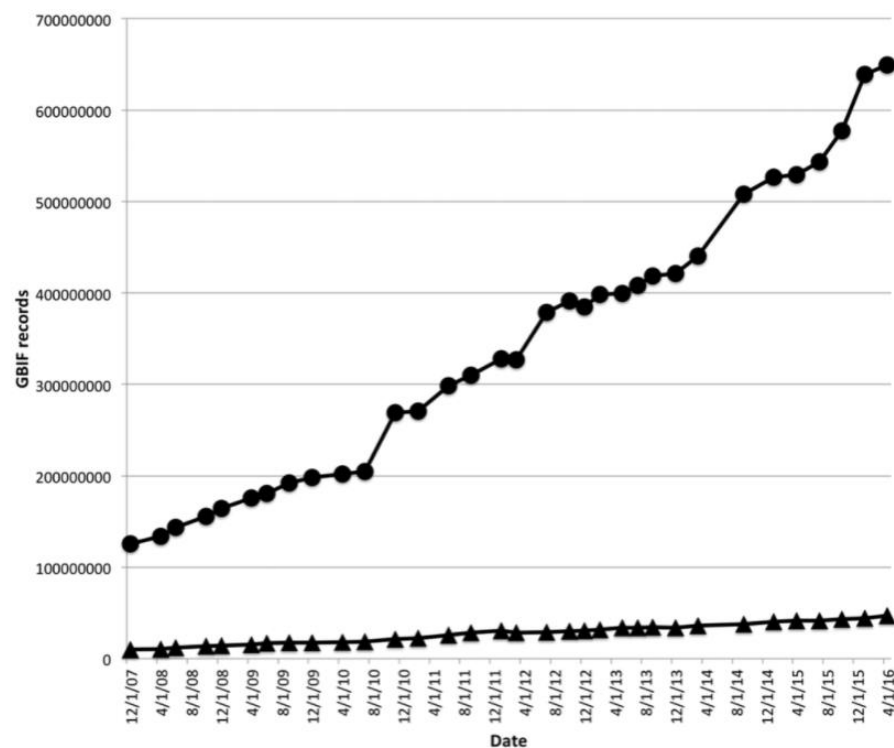
Accesibilidad





# El problema de los datos

## Accesibilidad



▲ Número de registros de insectos en GBIF.org


● Número de especímenes en colecciones

Datos  
accesibles

Datos no  
accesibles

# El problema de los datos

Accesibilidad



La información  
debe estar en  
formato **digital**,  
**accesible**,  
**localizable e**  
**integrado**

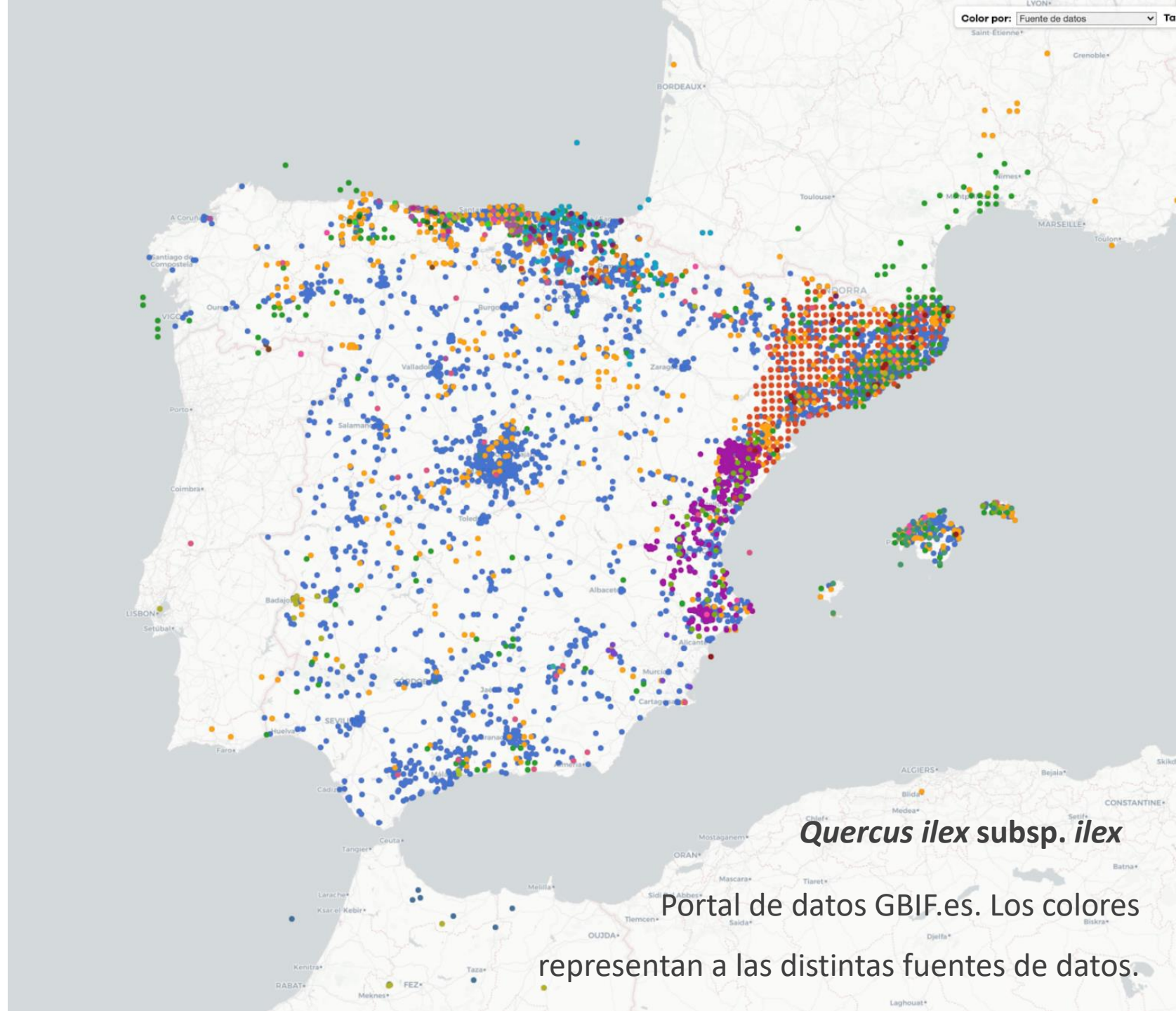


# GBIF

## Estándares

Garantizará el uso de un **lenguaje común**, para ofrecer los datos de una manera **estructurada** y más homogénea, que lo que hará es **facilitar el acceso, disponibilidad e intercambio** de datos primarios sobre biodiversidad a nivel global.

<https://registros.gbif.es/occurrences/>



# GBIF

## Estándares

### Estándar Darwin Core

El estándar Darwin Core (DwC) ofrece marco común, marco estable, sencillo y flexible para recopilar datos de biodiversidad procedentes de diferentes fuentes.

Desarrollado por **Biodiversity Information Standards** (TDWG)  
<https://dwc.tdwg.org/terms/>



**Registros presencia**  
**Eventos**  
**Listas**

Núcleo

Simple: 169 términos.

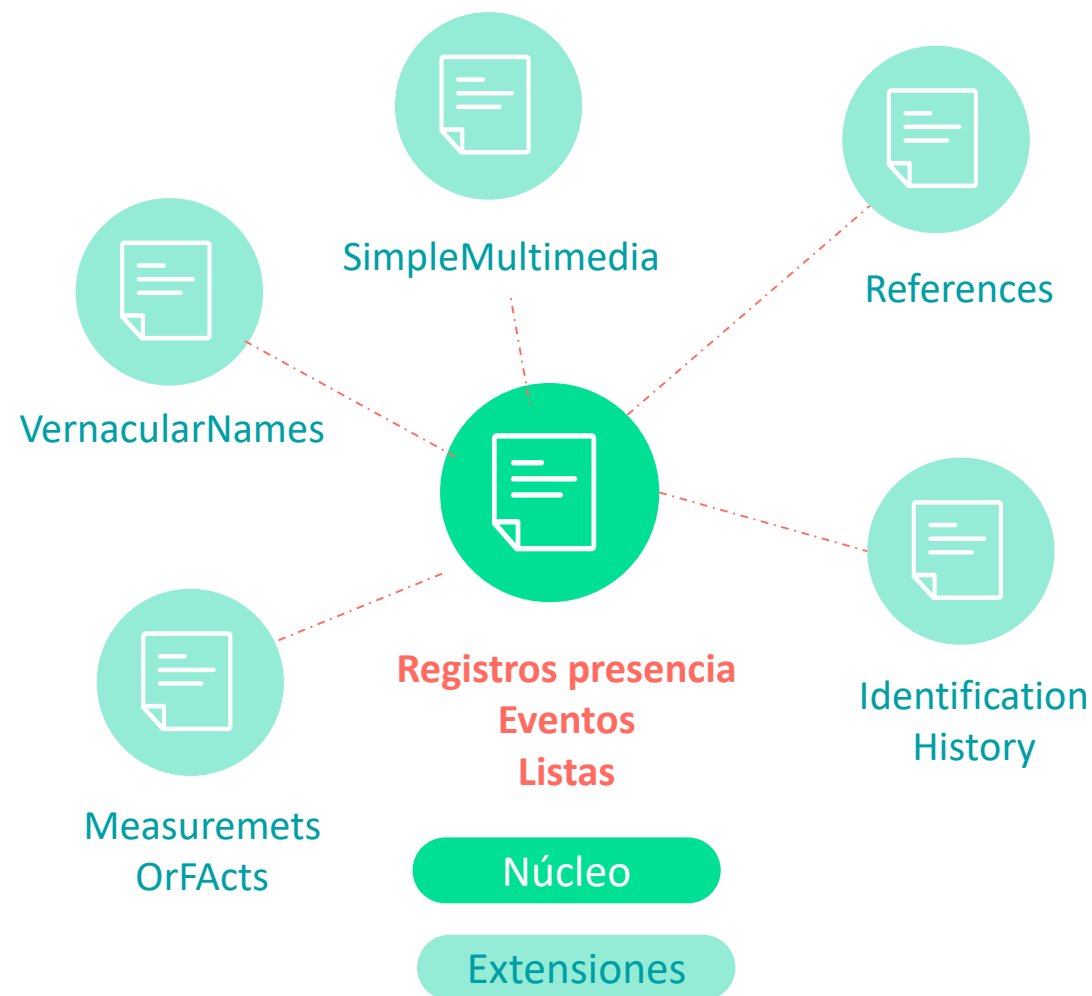
# GBIF

## Estándares

### Estándar Darwin Core

El estándar Darwin Core (DwC) ofrece marco común, marco estable, sencillo y flexible para recopilar datos de biodiversidad procedentes de diferentes fuentes.

Desarrollado por **Biodiversity Information Standards** (TDWG)  
<https://dwc.tdwg.org/terms/>



Simple: 169 términos. Extendido: 23 extensiones



**GBIF**

Estándares



Metadatos

Desarrollado por Biodiversity Information Standards (TDWG)

## Ecological Metadata Language

Lenguaje de Metadatos Ecológicos, o EML, es un estándar de metadatos que recoge información sobre juegos de datos ecológicos en una serie de documentos XML modulares y extensibles.

**quién, qué, cómo, cuándo y dónde**



# Fuentes de datos de biodiversidad



# Fuentes de datos



Especímenes preservados en colecciones



Especímenes preservados en herbarios



Libretas de campo, Tesis, literatura, informes



Ciencia ciudadana



Imágenes de satélite



Capturas de audio, vídeos, fototrampeo



Checklist



EIA, inventarios, etc.



## ¿Qué datos se pueden publicar?



### Datos de presencia

Evidencia sobre la presencia de un taxón en un lugar y una fecha determinados.



### Listas de especies

Catálogo o lista con nombres de taxones siguiendo criterios taxonómicos, geográficos o temáticos.



### Eventos de muestreo

Indican métodos, eventos, y la abundancia relativa de las especies registradas en un muestreo.

# ¿Qué datos se pueden publicar?

## DATOS DE PRESENCIA (OCCURRENCES)



Registros de **especímenes conservados** en herbarios y colecciones biológicas, u **observaciones** de individuos registradas en el campo.

- Pliegos de herbarios, bancos de germoplasma, semillas, etc., especímenes conservados en alcohol, formol, montados, etc.
- Fósiles, restos, muestras de tejidos, ADN, etc.
- Inventarios biológicos, etc.

# ¿Qué datos se pueden publicar?

## EVENTOS DE MUESTREO (SAMPLING EVENT)



Datos generados a partir de **estudios de monitoreo** como parcelas de vegetación, cámaras trampa, censos de aves, muestreos en aguas, etc. Indican claramente los **métodos, eventos y abundancia** relativa de las especies registradas.



# ¿Qué datos se pueden publicar?

## LISTAS DE ESPECIES (CHECKLIST)



Consiste en **listas de nombres científicos** o taxones que pueden estar restringidas a un grupo taxonómico, región geográfica, temática concreta o a una combinación de las tres.

- Checklist
- Listas de referencia taxonómica
- Listas rojas


# ¿Por qué publicar datos de biodiversidad?

- **Contribuir al conocimiento global** sobre la biodiversidad y, por tanto, a soluciones que promuevan su conservación y uso sostenible.


- Dar visibilidad y acceso a los datos aumenta las **posibilidades de su uso** y reutilización para la investigación y toma de decisiones.




# ¿Por qué publicar datos de biodiversidad?




Obtener **visibilidad y reconocimiento** a tú trabajo, obteniendo mayor reconocimiento a nivel institucional y personal.




Reconocimiento académico al publicarlos como un **artículo de datos (*data papers*)**.



Mayores posibilidades para **rastrear el uso de los datos** a través de DOI.



**Retroalimentación** a los datos por parte de la comunidad global que ayuda a mejorar su calidad.



Nuevas **oportunidades de colaboración** entre propietarios de datos e investigadores



Katia Cezón

GBIF.ES, Unidad de Coordinación  
CSIC  
Joaquín Costa, 22  
28002 Madrid, Spain

[katia@gbif.es](mailto:katia@gbif.es)



<https://creativecommons.org/licenses/by/4.0/deed.es>

GBIF-ES es el Nodo Nacional de Información en Biodiversidad patrocinado por el Ministerio Español de Ciencia e Innovación, gestionado por el Consejo Superior de Investigaciones Científicas (CSIC). <https://www.gbif.es/> <https://datos.gbif.es>, <https://elearning.gbif.es>



MINISTERIO  
DE CIENCIA  
E INNOVACIÓN



CSIC

Gbif.Es

