



BIODT
biodiversitydigitaltwin

School & Hackaton

March 3rd – March 7th
Officine Cantelmo, Lecce IT

DATA CITATION

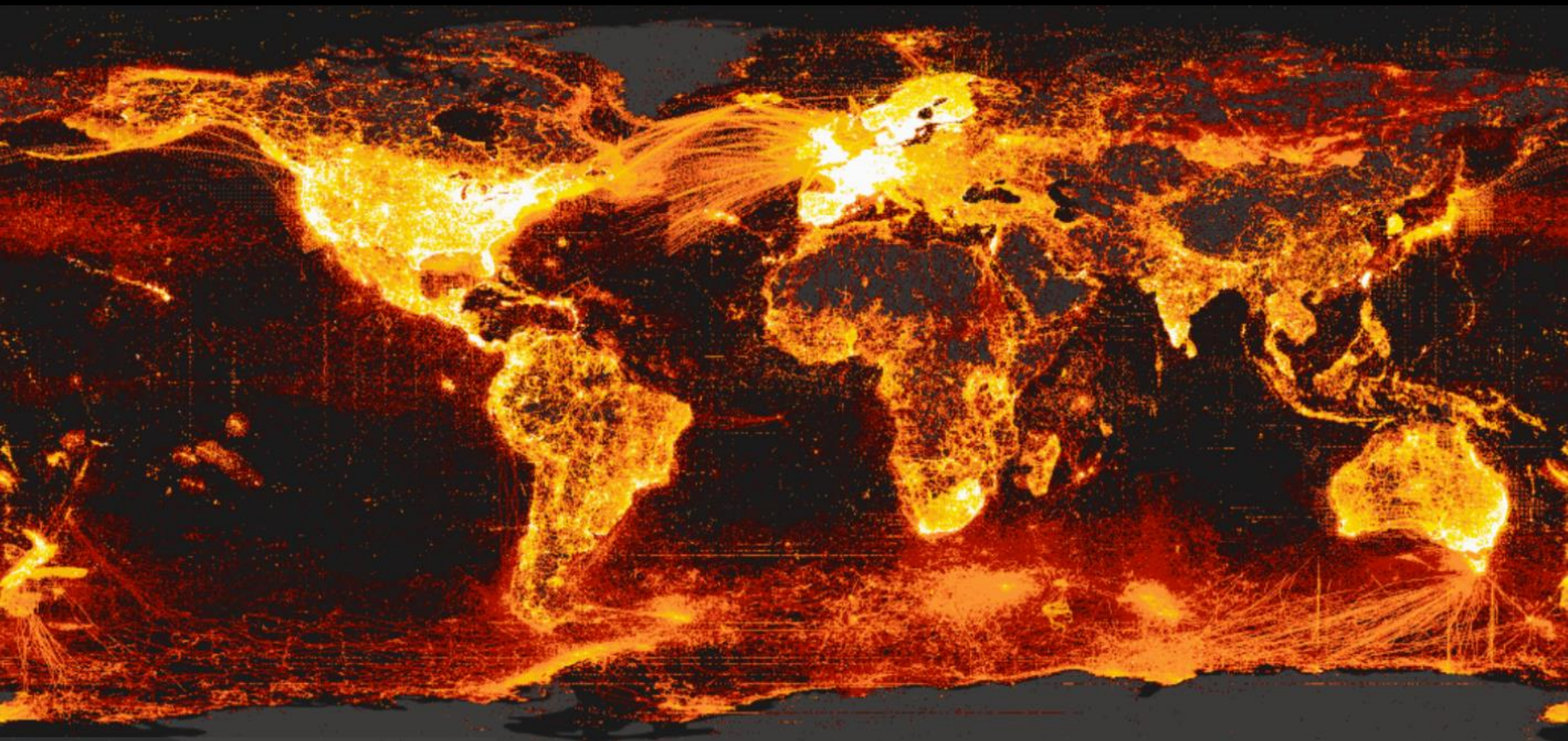
#CiteTheDOI

Dmitry Schigel – GBIF Secretariate

(supported by Dag Endresen, GBIF Norway, University of Oslo)



Funded by
the European Union



© OpenStreetMap contributors, © OpenMapTiles, GBIF.

Data Citation, slides by Scientific Officer Dmitry Schigel, GBIF Secretariate, March 2025

Data citation and GBIF powered literature

Dmitry Schigel | Scientific officer

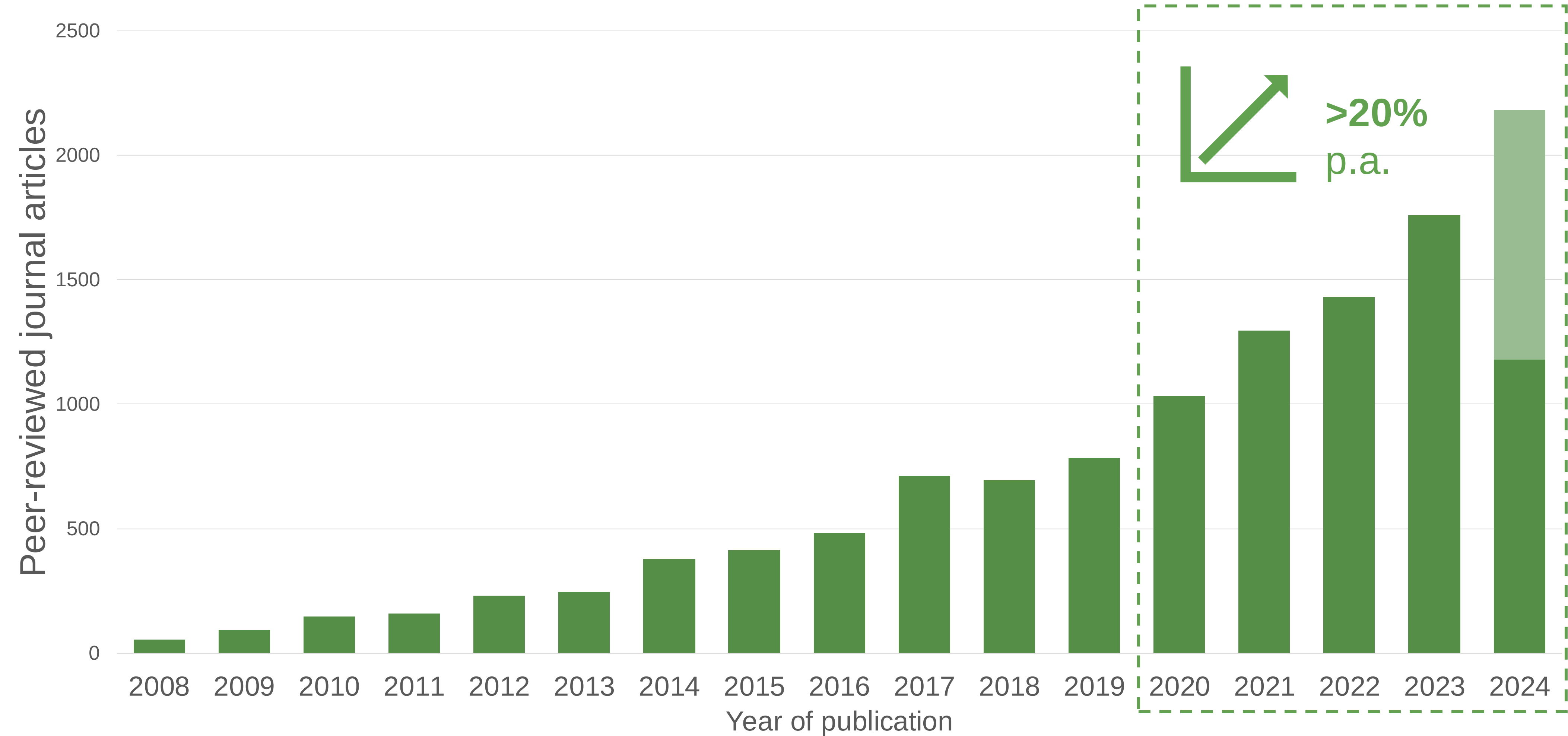


Why track citations?

- 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

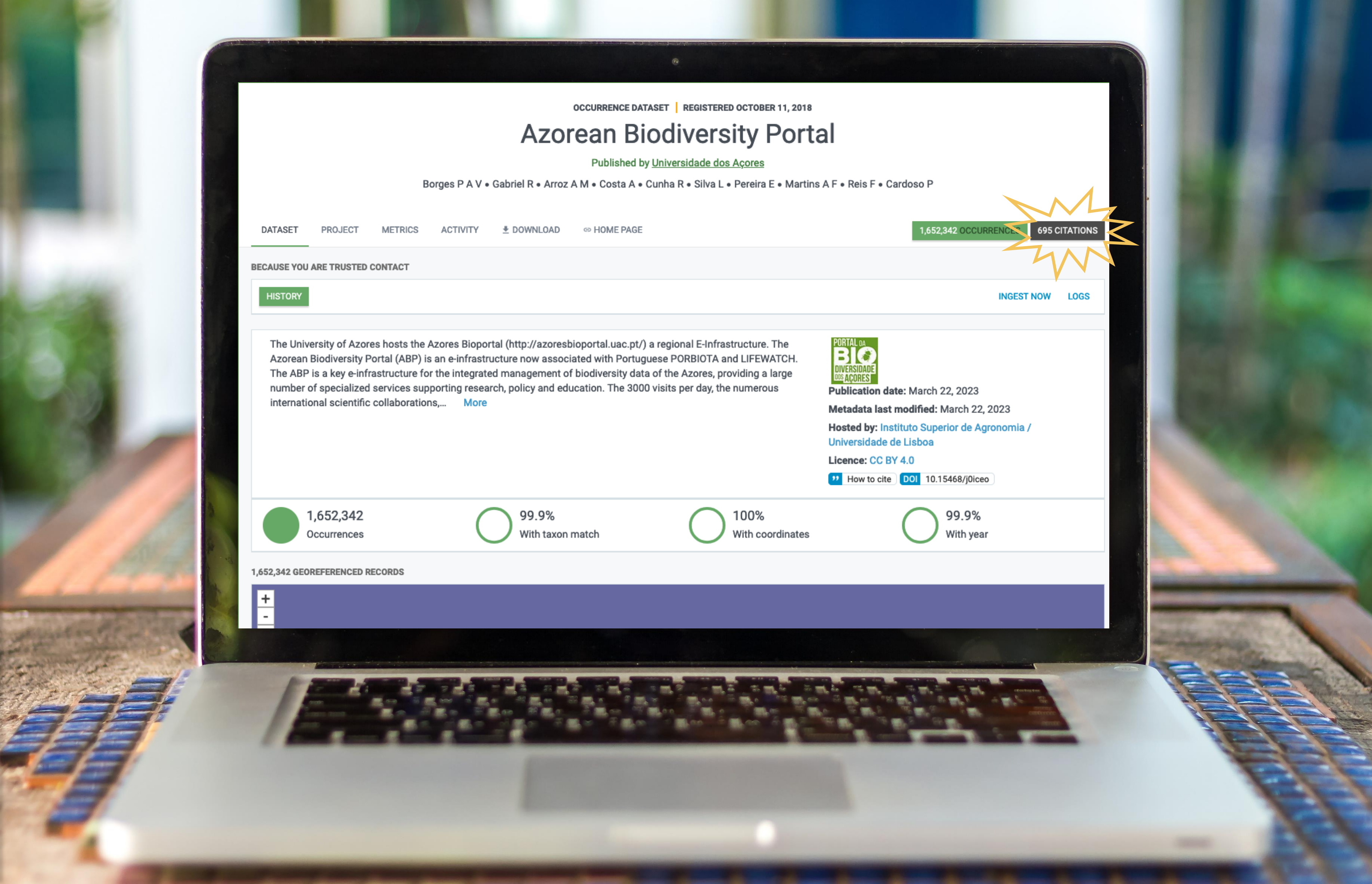


Use of GBIF-mediated data in research



Data citations: All about the DOIs

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

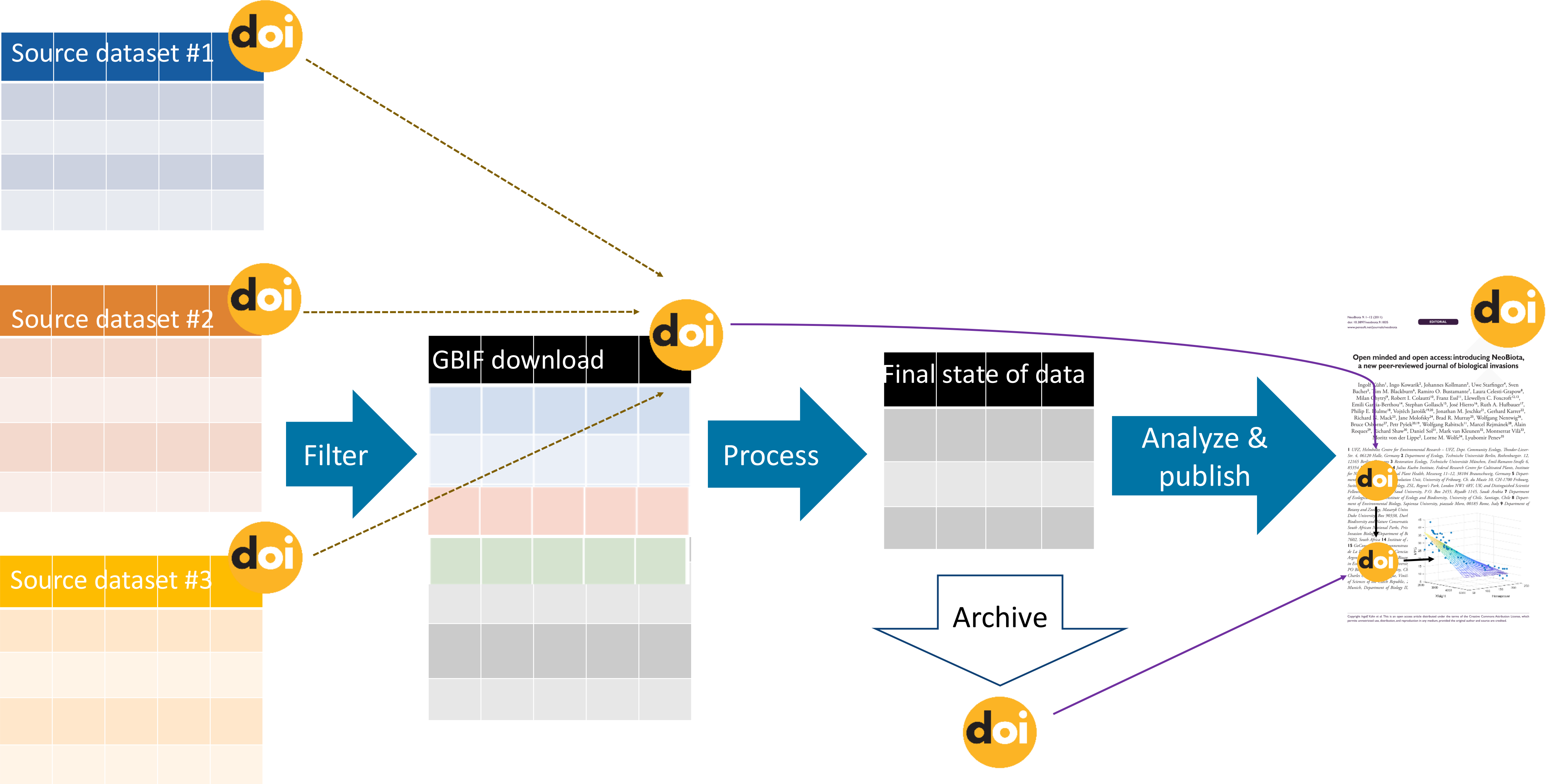


Dataset DOIs

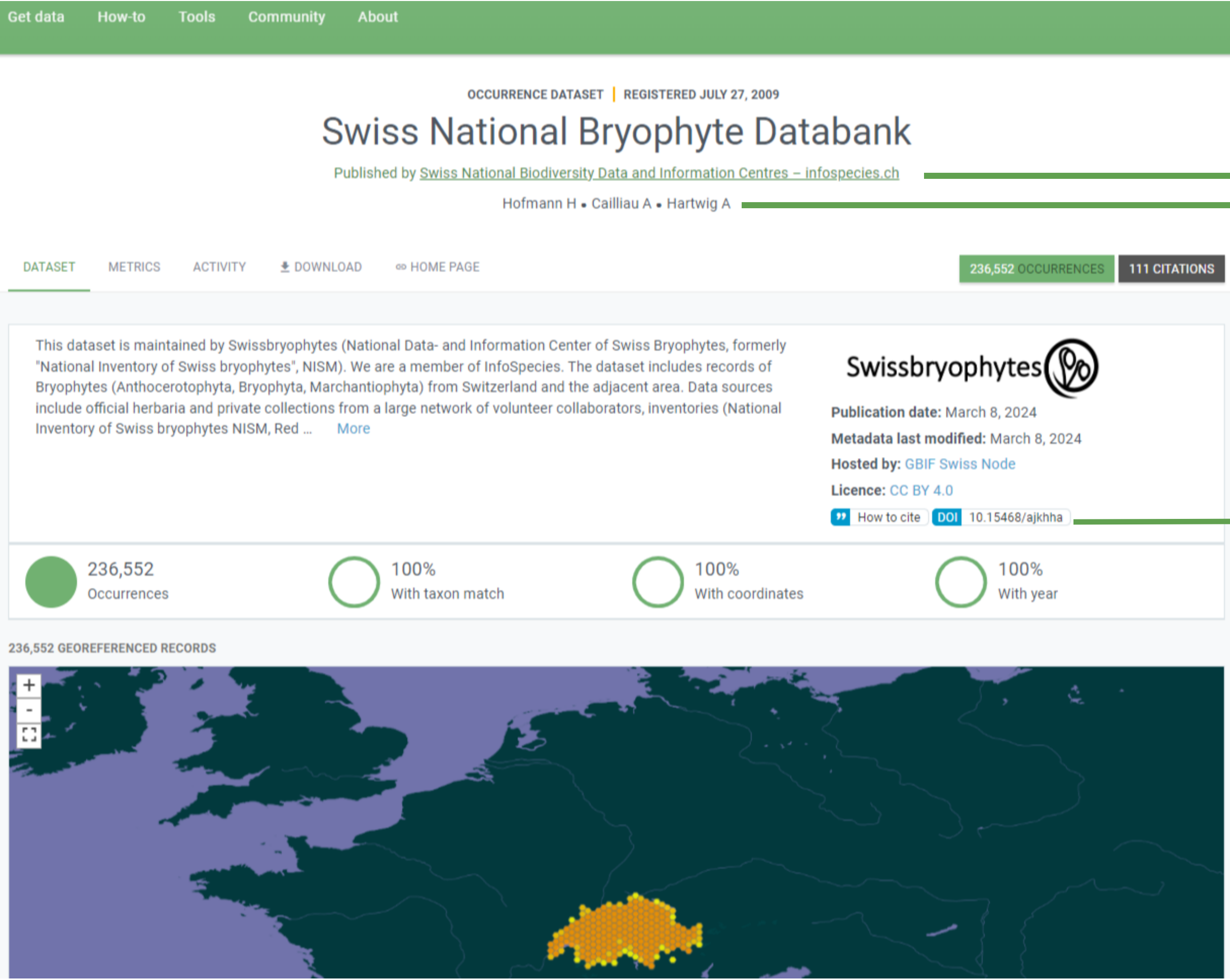
Download DOI

Archive DOI

Bibliographic DOI



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

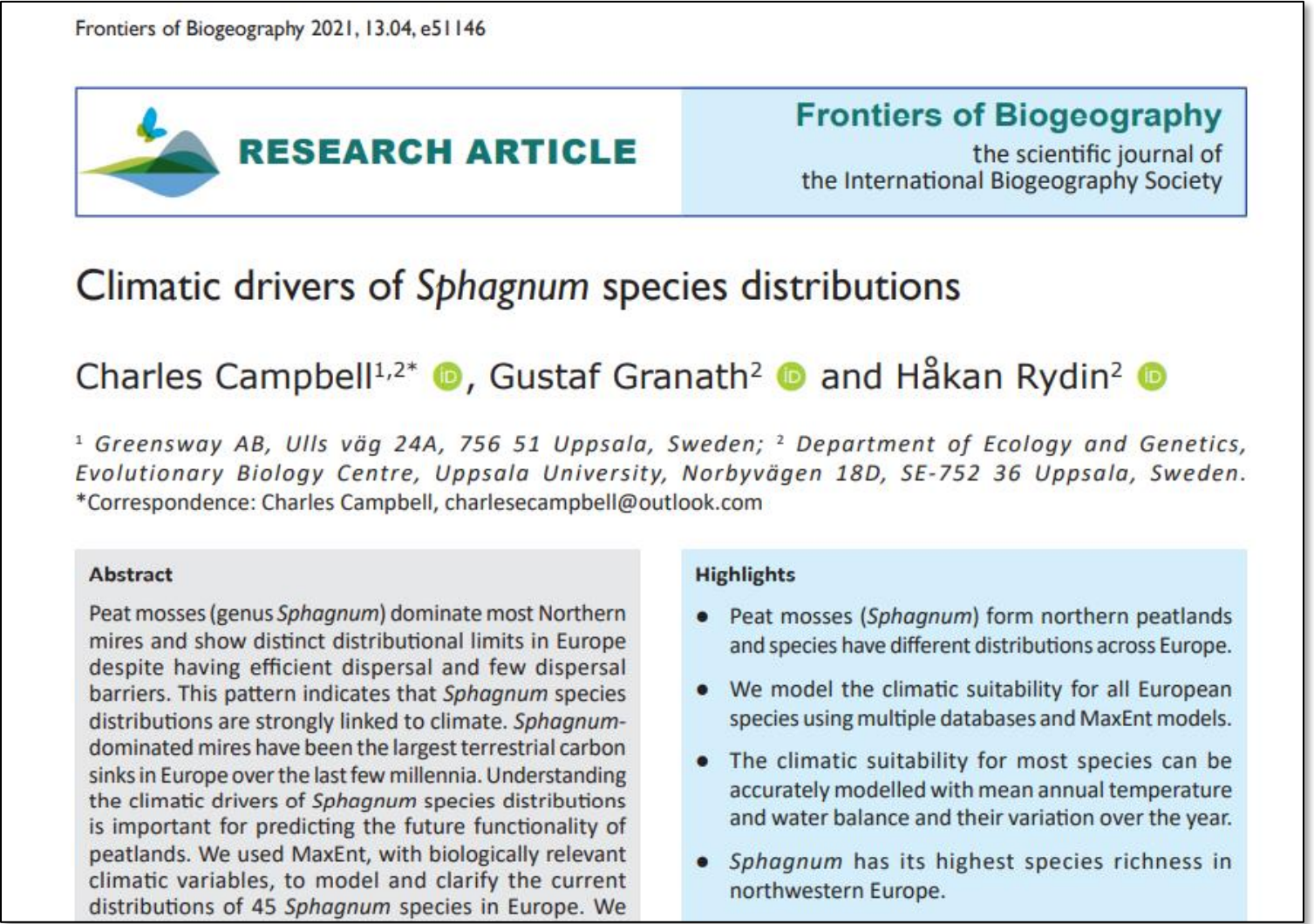


1 Affiliation

2 Authorship

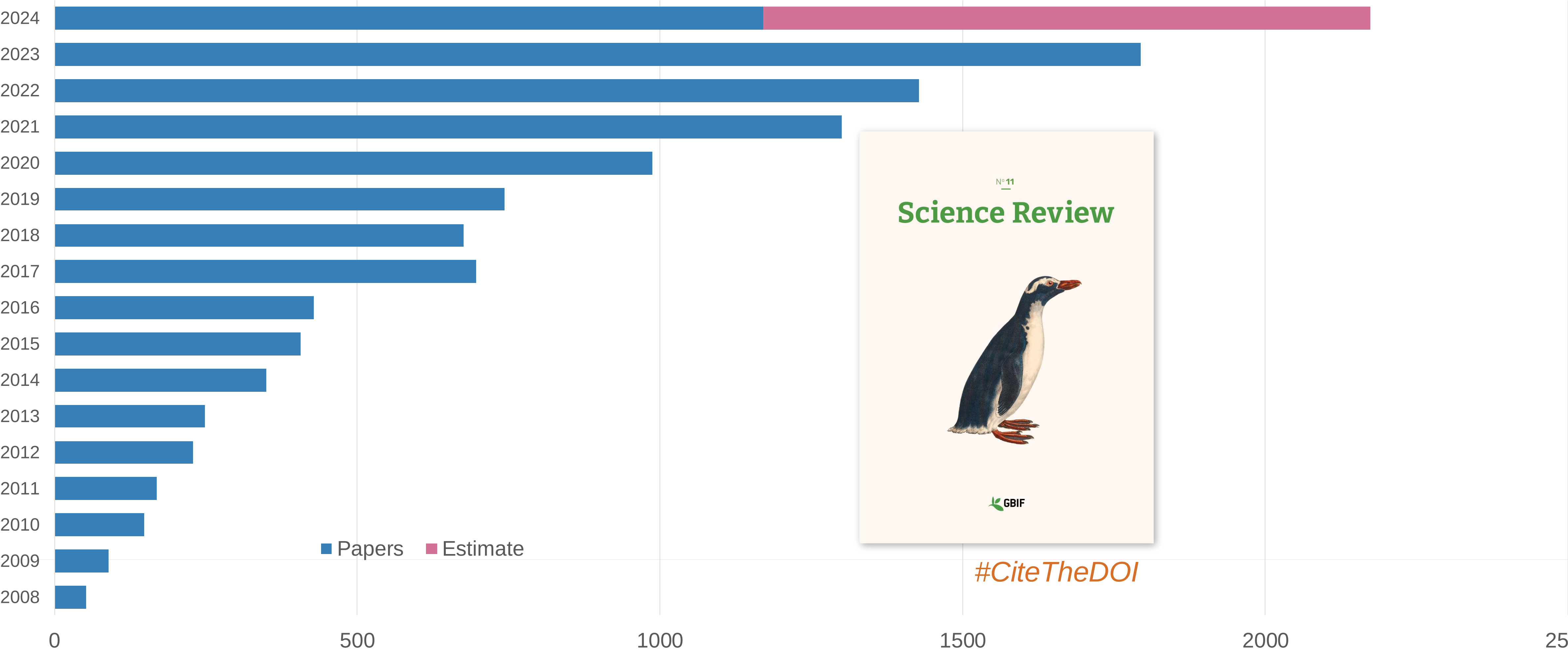
3 Data citations

4 DOI



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



#CiteTheDOI



Data citations

Ideal world

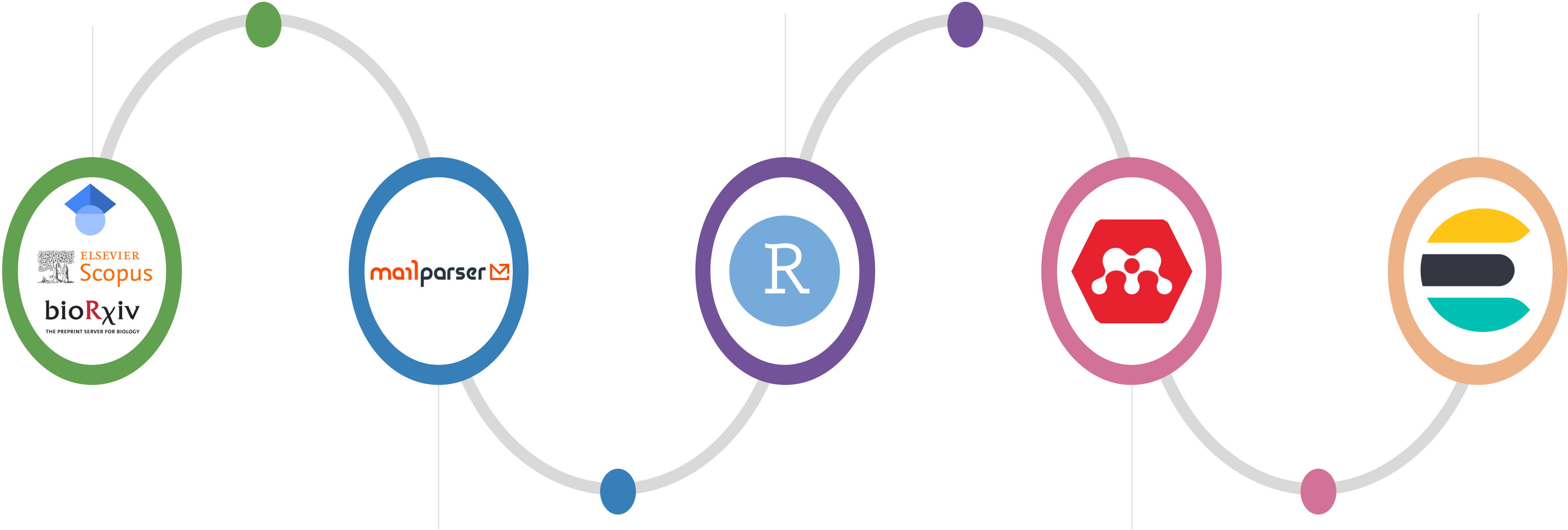
- All citations include a DOI, uniquely and specifically identifying the data used in a publication
- All data citations are included in the reference list of an article
- All data citations are machine-readable links between DOIs included in the DOI metadata, easily interpreted and used for quantifying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

Real world

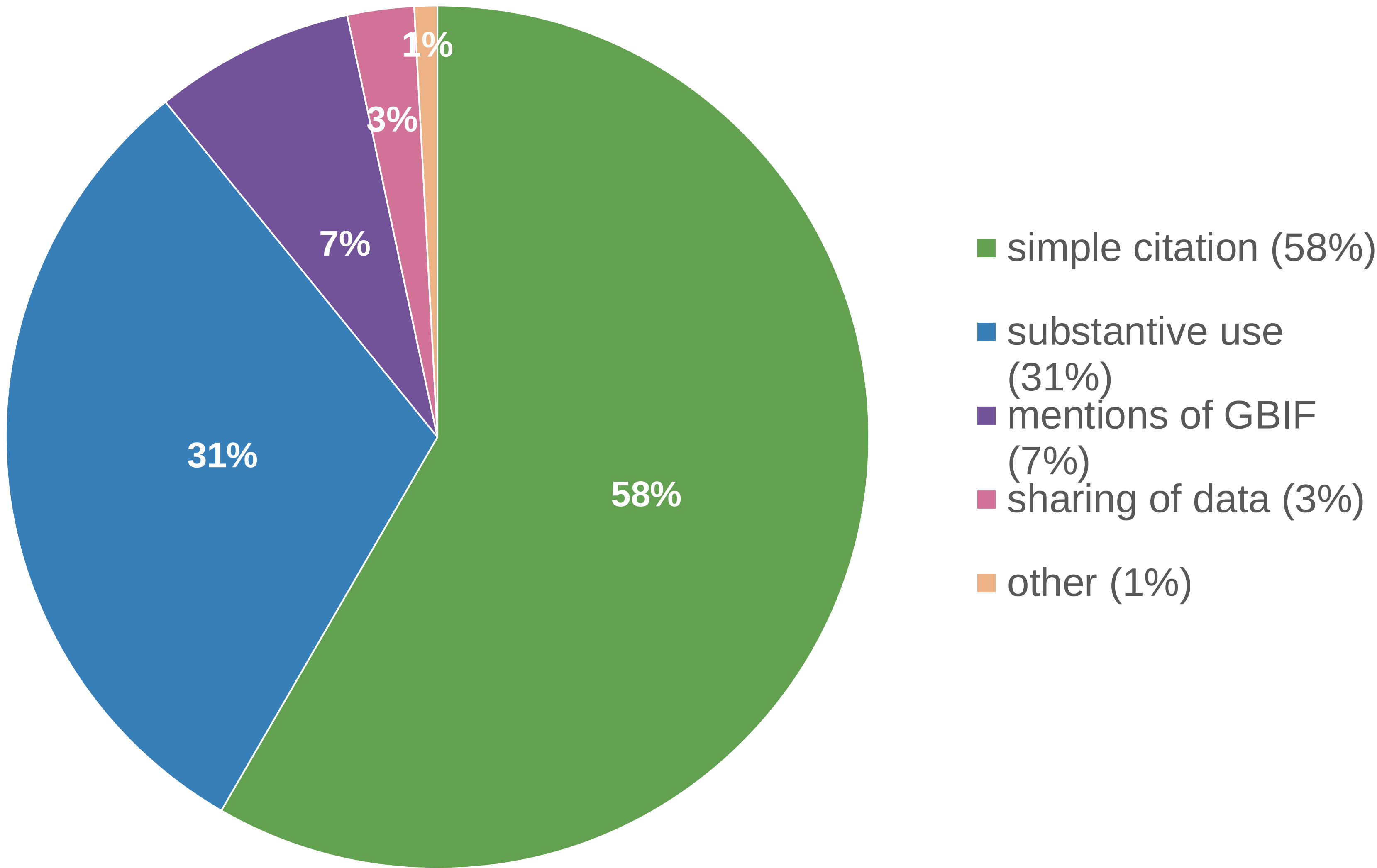
- Far from all citations include a DOI, and some tend to cite more data than is actually used
- Some citations are included in the body text of papers, some in data availability statements, and some are even hidden in supplements
- Even when done right, not all citations become available as machine-readable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



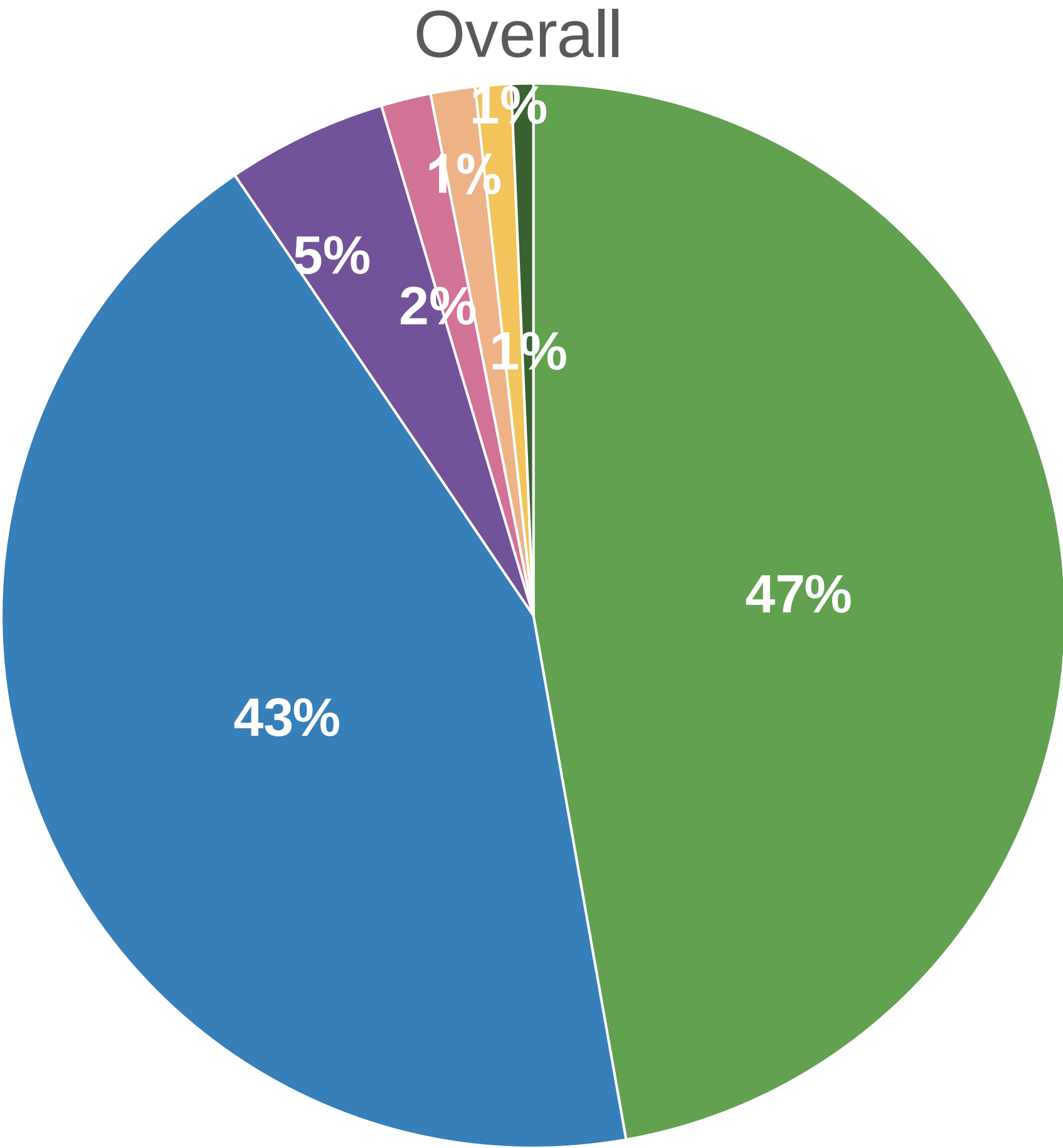
Literature tracking in GBIF



Why are papers mentioning GBIF?

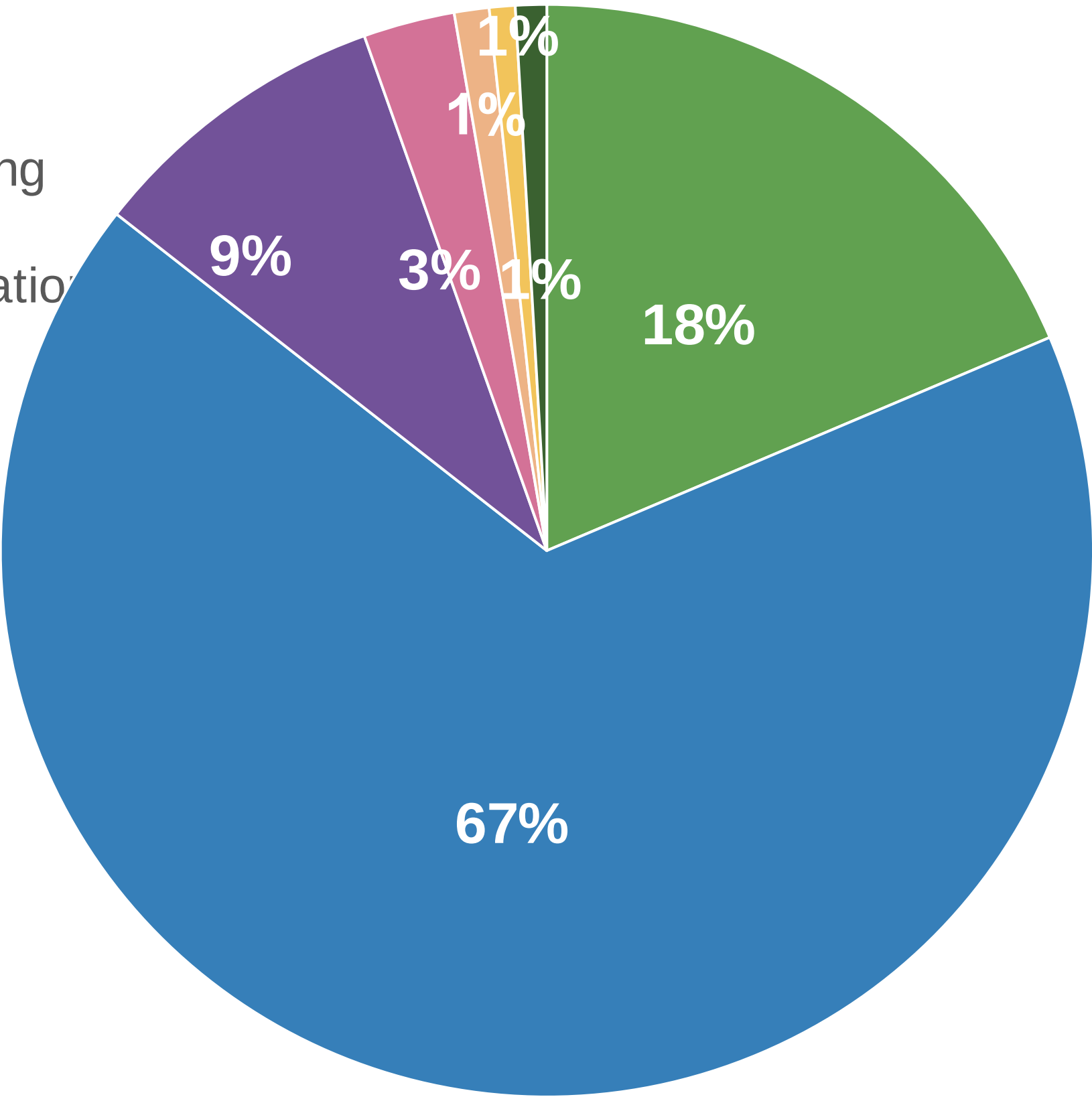


Types of literature tracked



- Reports
- Journal articles
- Preprints and working papers
- Theses and dissertations
- Books

Substantive use



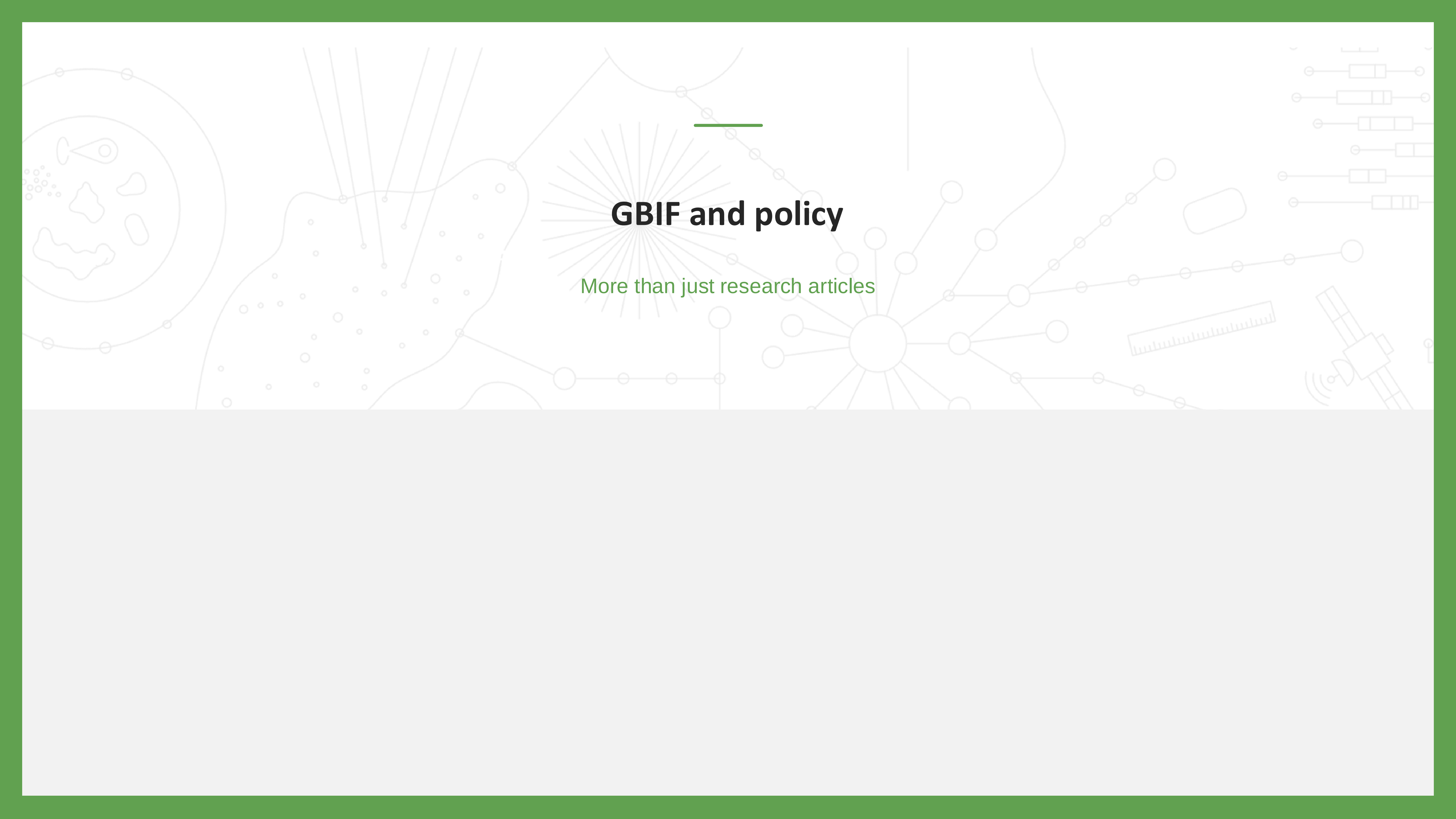
Nº 11

Science Review



 GBIF





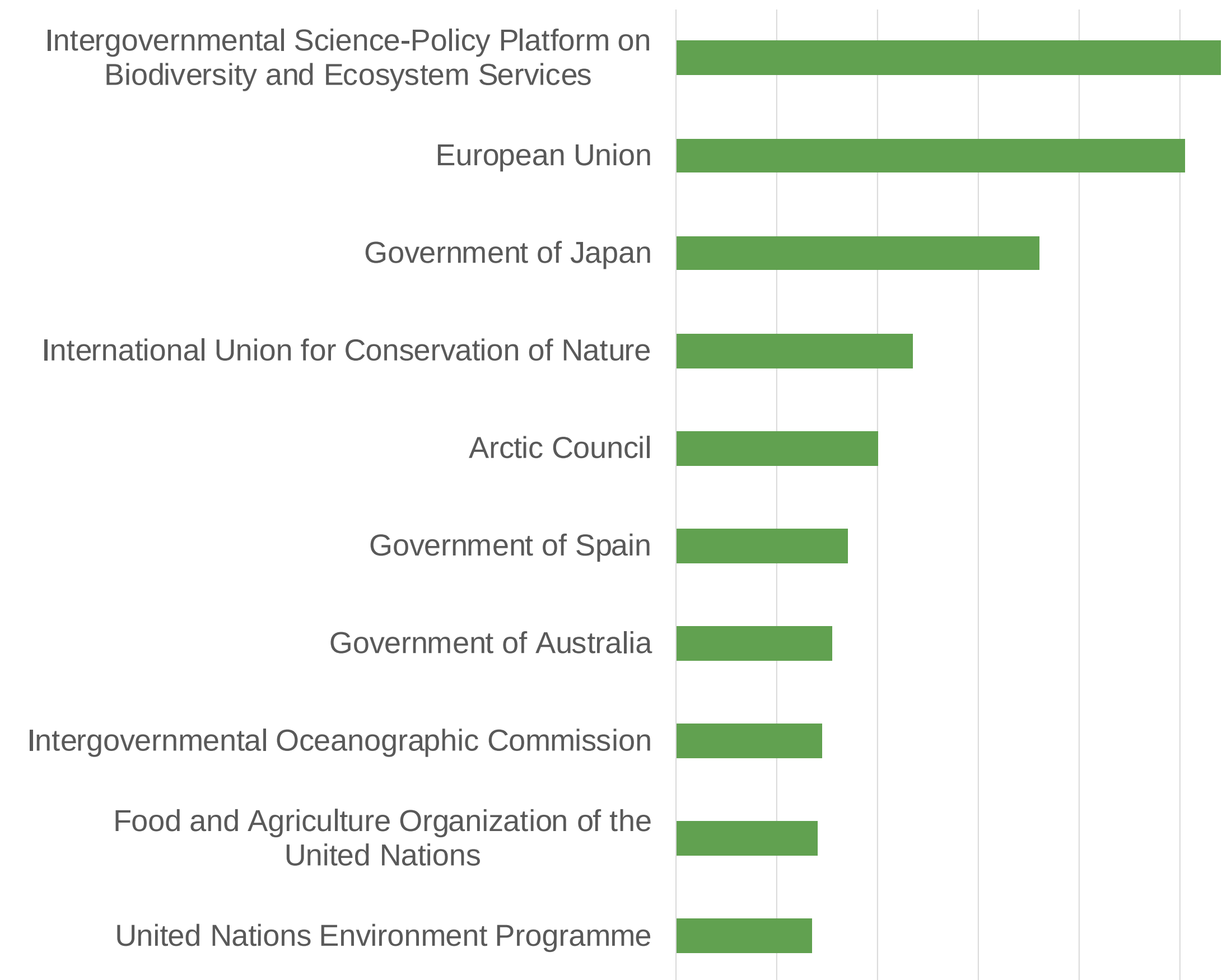
GBIF and policy

More than just research articles

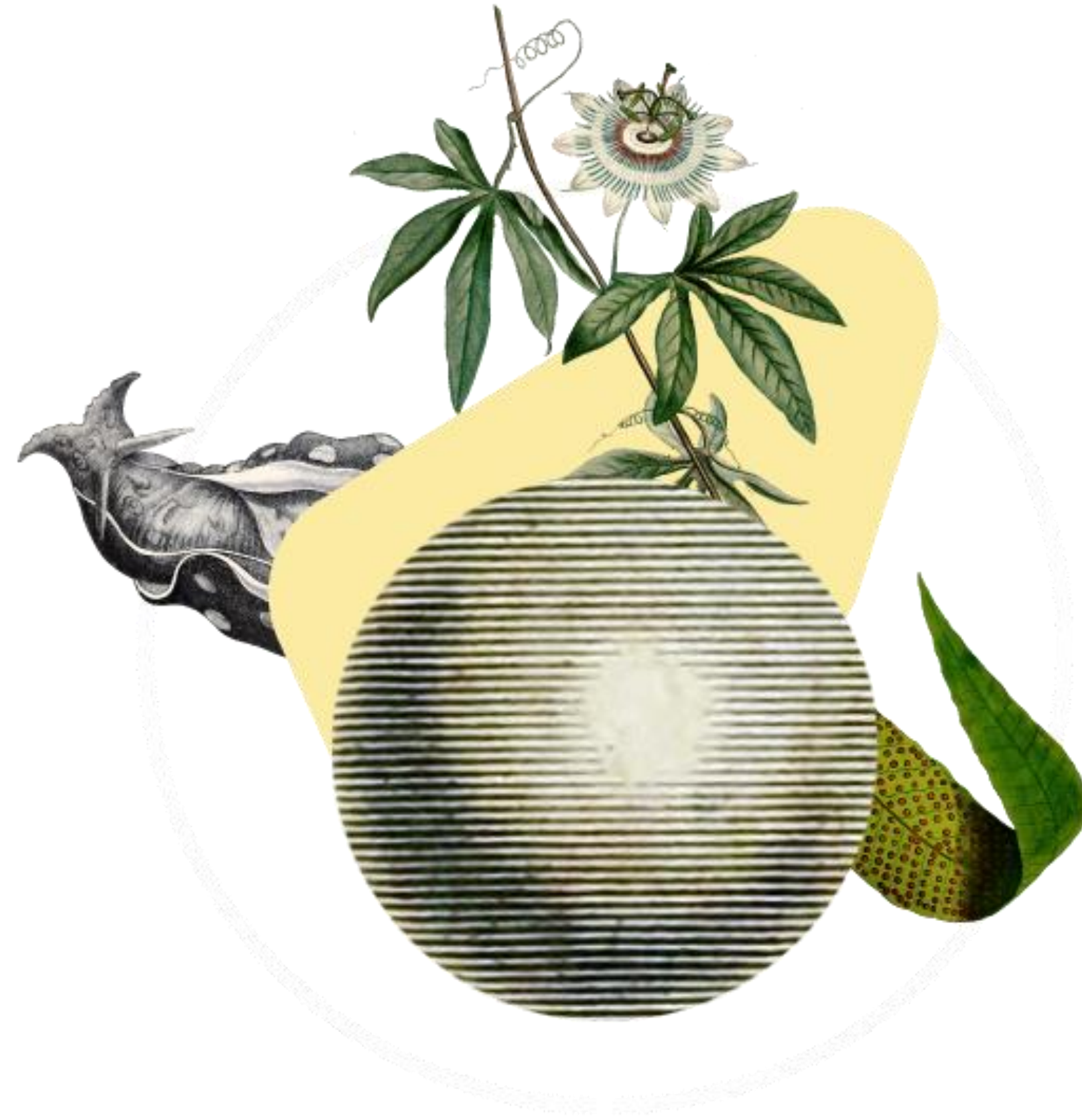
Overton

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

