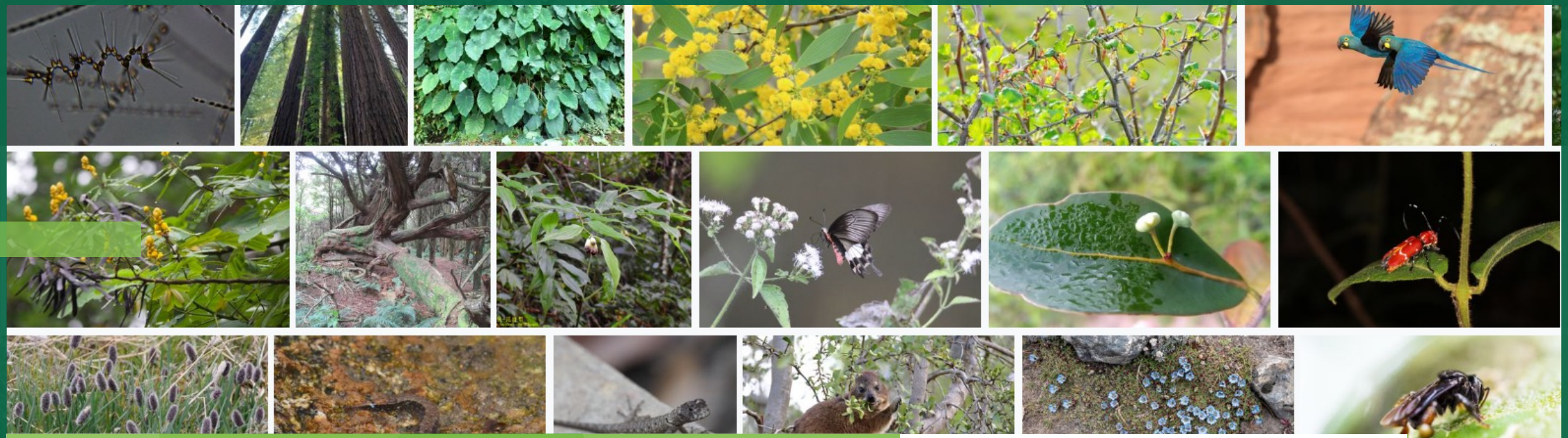


GBIF and data citations

Daniel Noesgaard | Science Communications Officer

GBIF Global Biodiversity Information Facility

Data Citation Community of Practice workshop, 8 June 2021

WHAT IS GBIF?

- International network and data infrastructure funded by the world's governments and aimed at providing anyone, anywhere, open access to data about all types of life on Earth
- 1.7 B records of organisms occurring in time and space, originating from 36,000 different datasets
- Data mediated by GBIF was used and cited in 1000+ scientific publications published in 2020, with citations increasing ~30% every year.

Photo: *Aspidimorpha sanctaecrucis* (Fabricius, 1792) observed in Indonesia by David Gil (CC BY-NC 4.0)

GBIF

DOWNLOADING DATA FROM GBIF

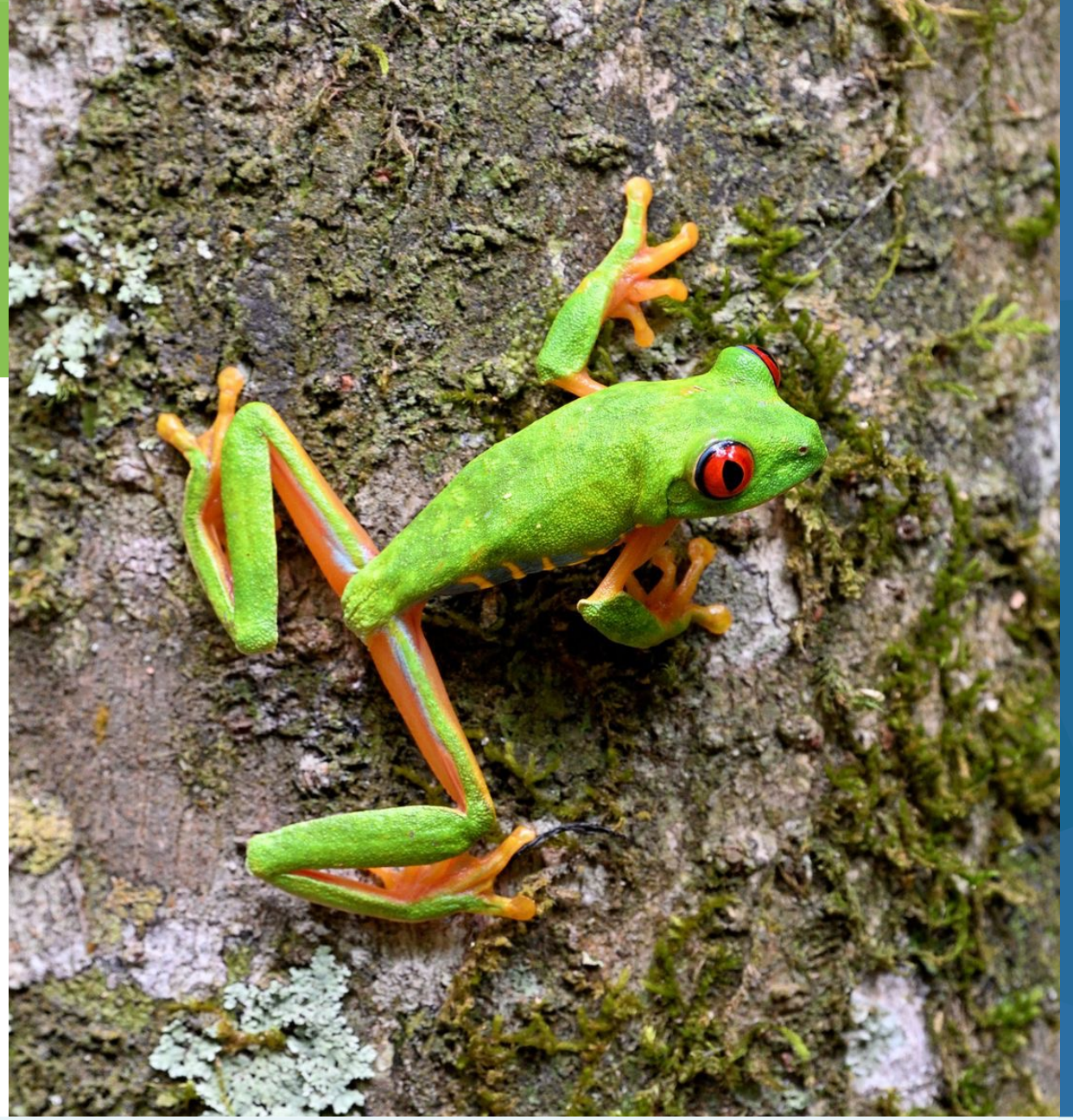


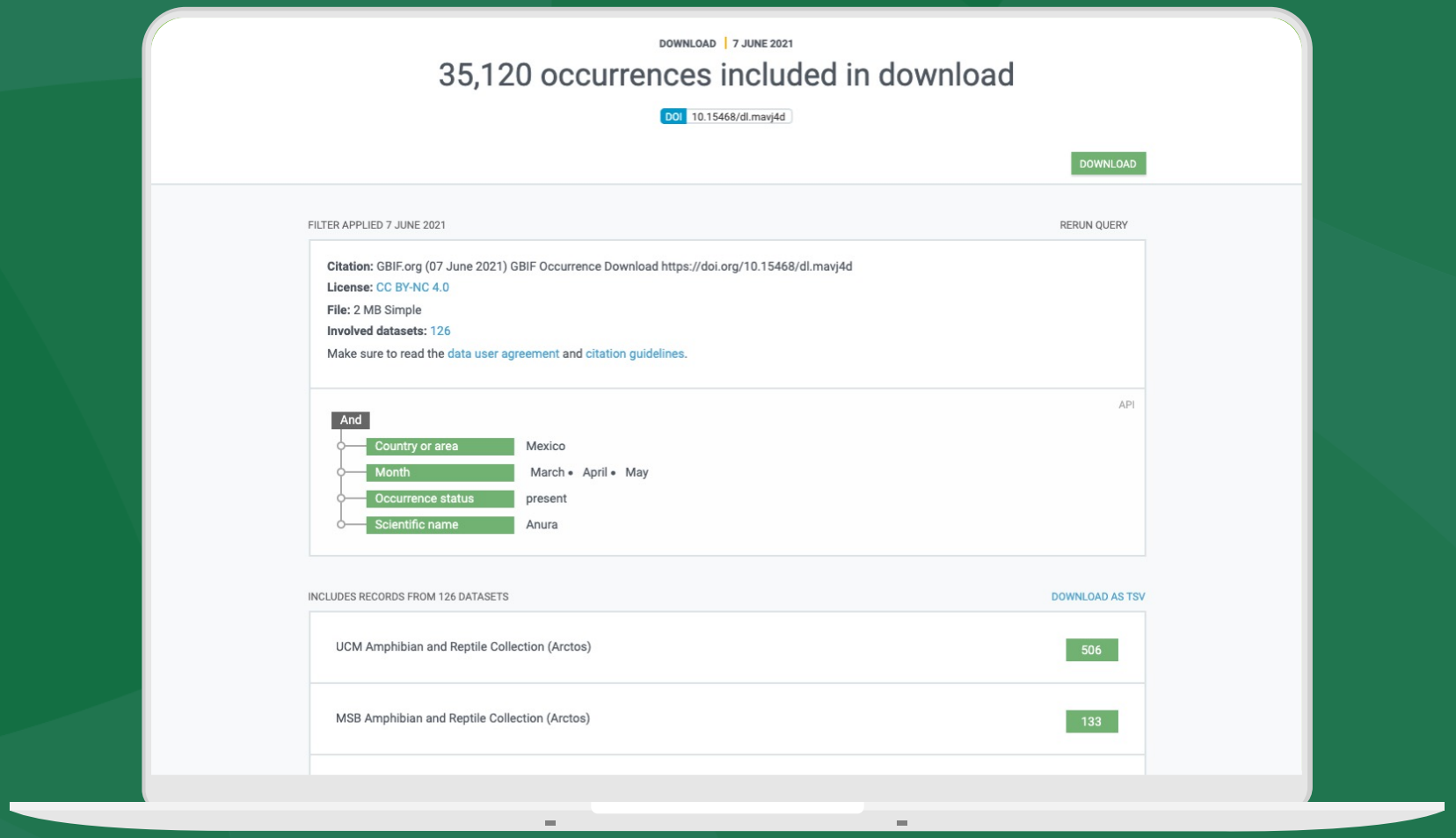
Photo: *Agalychnis callidryas* (Cope, 1862) observed in Mexico by Luis Guillermo (CC BY-NC 4.0)

DATA DOWNLOADS

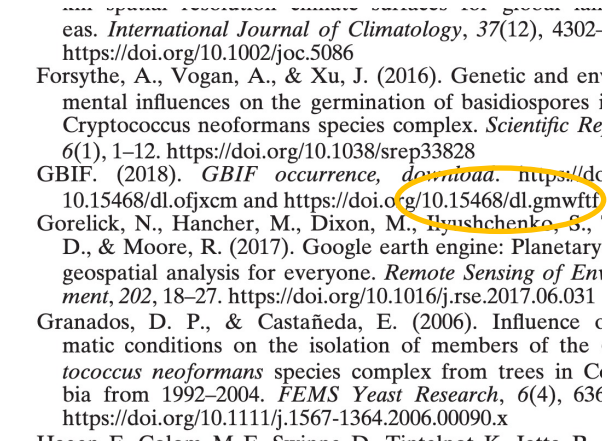
Number of records
Date
Filters
Contributing datasets

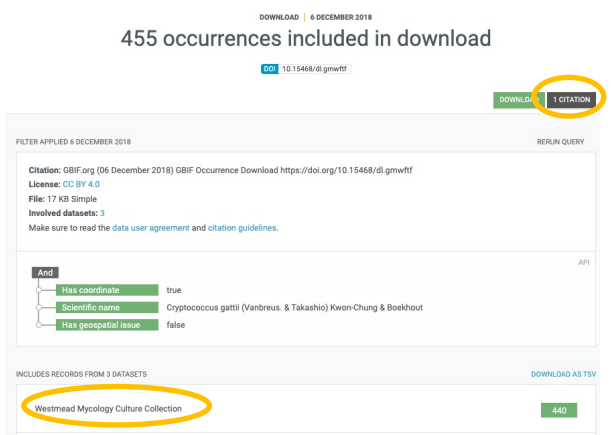
Download
Rerun

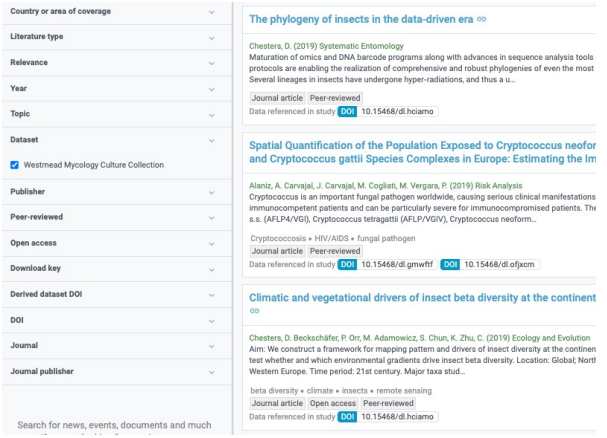
Recommended citation
DOI



LINKING USE TO DATA


eas. *International Journal of Climatology*, *37*(12), 4302-
https://doi.org/10.1002/joc.5086
Forsythe, A., Vogan, A., & Xu, J. (2016). Genetic and en
mental influences on the germination of basidiospores i
Cryptococcus neoformans species complex. *Scientific Re*
6(1), 1–12. https://doi.org/10.1038/srep33828
GBIF. (2018). *GBIF occurrence, download*. https://do
10.15468/dl.ofjxcm and https://doi.org/10.15468/dl.gmwftf
Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S.,
D., & Moore, R. (2017). Google earth engine: Planetary
geospatial analysis for everyone. *Remote Sensing of En*
ment, *202*, 18–27. https://doi.org/10.1016/j.rse.2017.06.031
Granados, D. P., & Castañeda, E. (2006). Influence
matic conditions on the isolation of members of the
tococcus neoformans species complex from trees in C
bia from 1992–2004. *FEMS Yeast Research*, *6*(4), 63
https://doi.org/10.1111/j.1567-1364.2006.00090.x





Paper is published citing a GBIF download DOI

The citation is logged against the download, shown on the landing page as 1 citation

Unpacking the citations to the individual datasets and publishers in GBIF

Data publishers can assess how much their data is being used and by which publications

USE OF GBIF DATA

- 1,700 papers logged citing specific data downloads
- In total, ~12,000 specific downloads cited
- More than one in three papers now cite GBIF using DOIs

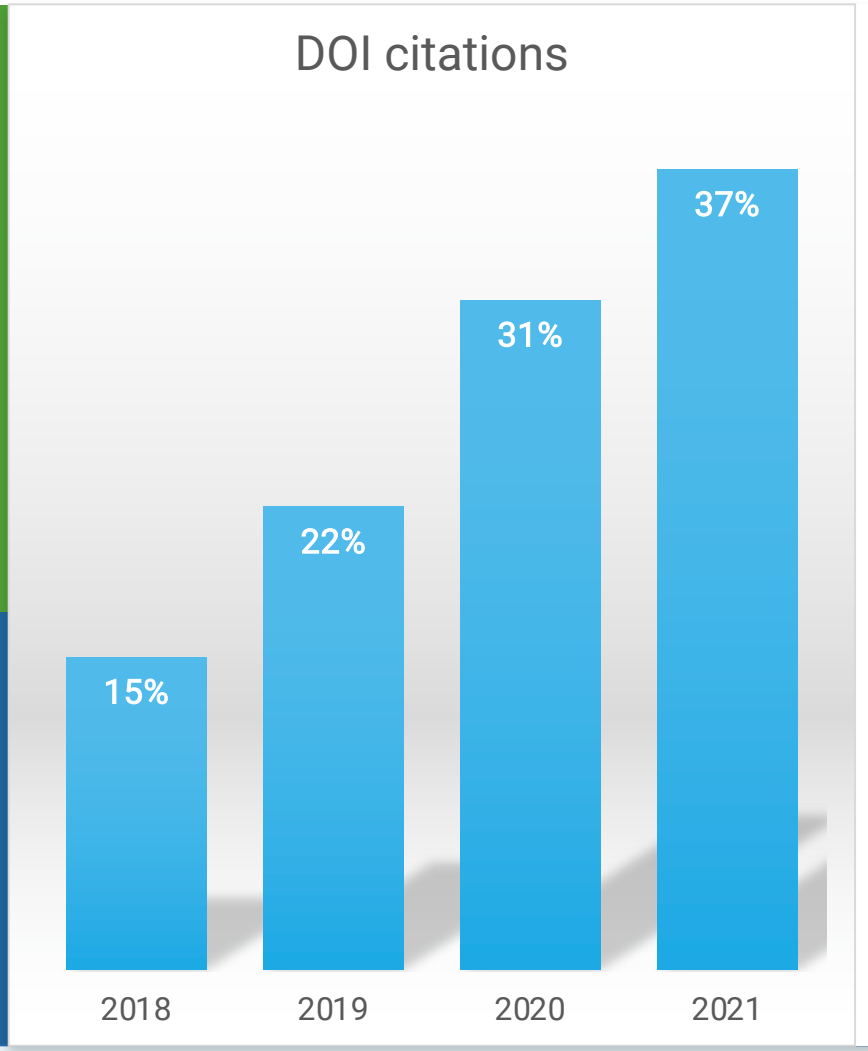


Photo: *Disphyma australe* (Sol.) J.M.Black observed in New Zealand by Christopher Stephens (CC BY-SA 4.0)

CHALLENGES

When there’s no (suitable) DOI to cite

Photo: *Gorilla beringei* subsp. *beringei* observed in Rwanda by Laura Keene (CC BY-NC 4.0)

NO SUITABLE DOI

- downloads that have been filtered/reduced significantly, or
- data accessed in a cloud computing environment, or
- data obtained by any means for which no DOI was assigned (e.g. third-party tools accessing the GBIF search API)

Photo: *Holacanthella paucispinosa* (J.T.Salmon, 1941) observed in New Zealand by damienbr (CC BY-NC 4.0)

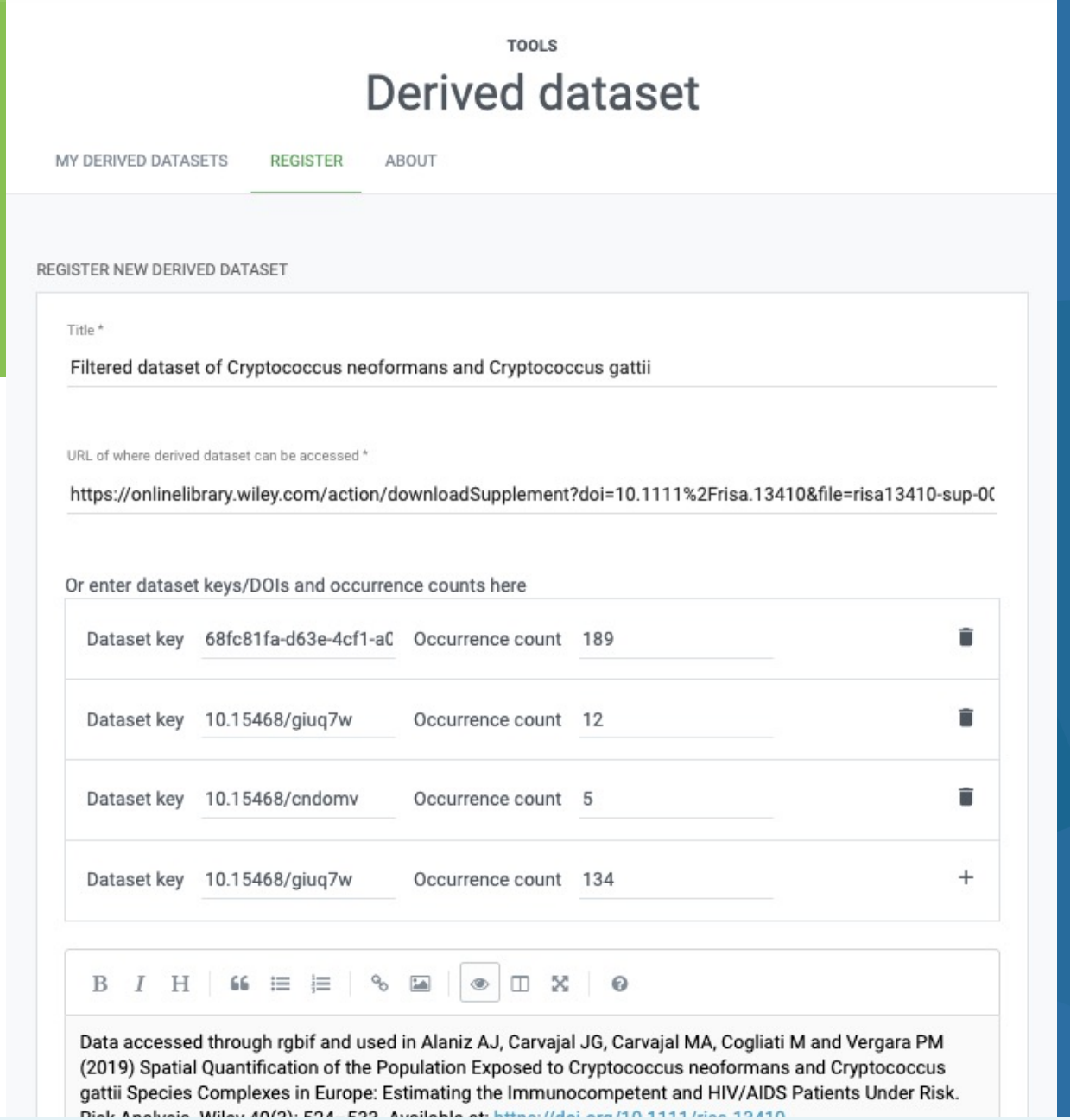
THE SOLUTION

A new service: derived datasets

To mimick what we do with downloads but based on data that has already been extracted for which no suitable DOI exists

User must extract list of contributing datasets with counts of records and provide a URL of where the dataset can be accessed

In return they get a DOI to cite which can be used to credit data publishers



https://www.gbif.org/derived-dataset/register

#CITETHEDOI

Daniel Noesgaard

dnoesgaard@gbif.org

GBIF