Versioned Archive and Review of Biotic Interactions and Taxon Names Found within globalbioticinteractions/globalfungi hash://md5/4e29adcf334fd1f828dfb7d39173458f

by Nomer, Elton and Preston, three naive review bots
review@globalbioticinteractions.org
https://globalbioticinteractions.org/contribute
https://github.com/globalbioticinteractions/globalfungi/issues

2026-03-30

Abstract

Life on Earth is sustained by complex interactions between organisms and their environment. These biotic interactions can be captured in datasets and published digitally. We present a review and archiving process for such an openly accessible digital interactions dataset of known origin and discuss its outcome. The dataset under review, named globalbioticinteractions/globalfungi, has fingerprint hash://md5/4e29adcf334fd1f828dfb7d39173458f, is 966MiB in size and contains 1,698,482 interactions with 1 unique type of association (e.g., symbiontOf) between 2,510 primary taxa (e.g., Fagus sylvatica) and 16,053 associated taxa (e.g., Ascomycota sp.). This report includes detailed summaries of interaction data, a taxonomic review from multiple catalogs, and an archived version of the dataset from which the reviews are derived.

Contents

Introduction

Data Review and Archive

Data review and archiving can be a time-consuming process, especially when done manually. This review report aims to help facilitate both activities. It automates the archiving of datasets, including Darwin Core archives, and is a citable backup of a version of the dataset. Additionally, an automatic review of species interaction claims made in the dataset is generated and registered with Global Biotic Interactions (J. H. Poelen, Simons, and Mungall 2014).

This review includes summary statistics about, and observations about, the dataset under review :

> Větrovský, T., Morais, D., Kohout, P., Lepinay, C., Algora, C., Awokunle Hollá, S., Bahnmann, B.D., Bílohnědá, K., Brabcová, V., D'Alò, F., Human, Z.R., Jomura, M., Kolařík, M., Kvasničková, J., Lladó, S., López-Mondéjar, R., Martinović, T., Mašínová, T., Meszárošová, L., Michalčíková, L., Michalová, T., Mundra, S., Navrátilová, D., Odriozola, I., Piché-Choquette, S., Štursová, M., Švec, K., Tláskal, V., Urbanová, M., Vlk, L., Voříšková, J., Žifčáková, L., Baldrian, P., 2020. GlobalFungi, a global database of fungal occurrences from high-throughput-sequencing metabarcoding studies. Scientific Data 7, 228. https://doi.org/10.1038/s41597-020-0567-7
> https://github.com/globalbioticinteractions/globalfungi/archive/b089bae129a5228d5f847648434061d943
> 2026-03-28T01:40:58.139Z hash://md5/4e29adcf334fd1f828dfb7d39173458f

For additional metadata related to this dataset, please visit https://github.com/globalbioticinteractions/globalfungi and inspect associated metadata files including, but not limited to, *README.md*, *eml.xml*, and/or *globi.json*.

Methods

The review is performed through programmatic scripts that leverage tools like Preston (Elliott et al. 2025), Elton (Kuhn, Poelen, and Leinweber 2025), Nomer (Salim and Poelen 2025), globinizer (J. Poelen, Seltmann, and Mietchen 2024) combined with third-party tools like grep, mlr, tail and head.

Table 1: Tools used in this review process

tool name	version
preston	0.11.1
elton	0.16.7
nomer	0.6.2
globinizer	0.4.0
mlr	6.0.0
jq	1.6
yq	4.25.3
pandoc	3.1.6.1
duckdb	1.3.1
mapserver	7.6.4

The review process can be described in the form of the script below [1].

```
# get versioned copy of the dataset (size approx.  966MiB) under review
elton pull globalbioticinteractions/globalfungi

# generate review notes
elton review globalbioticinteractions/globalfungi\
 > review.tsv

# export indexed interaction records
elton interactions globalbioticinteractions/globalfungi\
 > interactions.tsv

# export names and align them with the Catalogue of Life using Nomer
elton names globalbioticinteractions/globalfungi\
 | nomer append col\
 > name-alignment.tsv
```

or visually, in a process diagram.

[1] Note that you have to first get the data (e.g., via elton pull globalbioticinteractions/globalfungi) before being able to generate reviews (e.g., elton review globalbioticinteractions/globalfungi), extract interaction claims (e.g., elton interactions globalbioticinteractions/globalfungi), or list taxonomic names (e.g., elton names globalbioticinteractions/globalfungi)

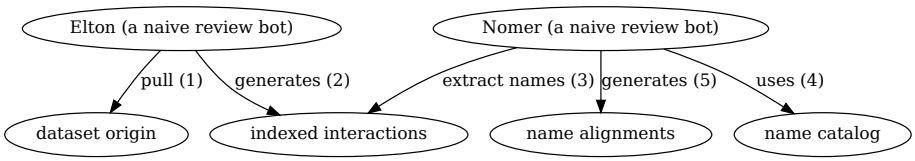

Figure 1: Review Process Overview

You can find a copy of the full review script at check-data.sh. See also GitHub and Codeberg.

Results

In the following sections, the results of the review are summarized [2]. Then, links to the detailed review reports are provided.

Files

The following files are produced in this review:

filename	description
biblio.bib	list of bibliographic reference of this review
check-dataset.sh	data review workflow/process as expressed in a bash script
data.zip	a versioned archive of the data under review
HEAD	the digital signature of the data under review
index.docx	review in MS Word format
index.html	review in HTML format
index.md	review in Pandoc markdown format
index.pdf	review in PDF format
indexed-citations.csv.gz	list of distinct reference citations for reviewed species interaction claims in gzipped comma-separated values file format
indexed-citations.html.gz	list of distinct reference citations for reviewed species interactions claims in gzipped html file format

[2]Disclaimer: The results in this review should be considered friendly, yet naive, notes from an unsophisticated robot. Please keep that in mind when considering the review results.

filename	description
indexed-citations.tsv.gz	list of distinct reference citations for reviewed species interaction claims in gzipped tab-separated values format
indexed-interactions-col-family-col-family.svg	network diagram showing the taxon family to taxon family interaction claims in the dataset under review as interpreted by the Catalogue of Life via Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024)
indexed-interactions-col-kingdom-col-kingdom.svg	network diagram showing the taxon kingdom to taxon kingom interaction claims in the dataset under review as interpreted by the Catalogue of Life via Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024)
indexed-interactions.csv.gz	species interaction claims indexed from the dataset under review in gzipped comma-separated values format
indexed-interactions.html.gz	species interaction claims indexed from the dataset under review in gzipped html format
indexed-interactions.tsv.gz	species interaction claims indexed from the dataset under review in gzipped tab-separated values format
indexed-interactions.parquet	species interaction claims indexed from the dataset under review in Apache Parquet format
indexed-interactions.png	species interaction claims indexed from the dataset under review plotted on a map
indexed-interactions.map	mapserver configuration to plot species interaction claims indexed from the dataset under review on a map
indexed-interactions.gpkg	species interaction claims indexed from the dataset under review in GeoPackage format
indexed-interactions-h3.gpkg	geospatially clustered h3 species interaction claims indexed from the dataset under review in GeoPackage format

filename	description
indexed-interactions-sample.csv	list of species interaction claims indexed from the dataset under review in gzipped comma-separated values format
indexed-interactions-sample.html	first 500 species interaction claims indexed from the dataset under review in html format
indexed-interactions-sample.tsv	first 500 species interaction claims indexed from the dataset under review in tab-separated values format
indexed-names.csv.gz	taxonomic names indexed from the dataset under review in gzipped comma-separated values format
indexed-names.html.gz	taxonomic names found in the dataset under review in gzipped html format
indexed-names.tsv.gz	taxonomic names found in the dataset under review in gzipped tab-separated values format
indexed-names.parquet	taxonomic names found in the dataset under review in Apache Parquet format
indexed-names-resolved-col.csv.gz	taxonomic names found in the dataset under review aligned with the Catalogue of Life as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped comma-separated values format
indexed-names-resolved-col.html.gz	taxonomic names found in the dataset under review aligned with the Catalogue of Life as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped html format
indexed-names-resolved-col.tsv.gz	taxonomic names found in the dataset under review aligned with the Catalogue of Life as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped tab-separated values format

filename	description
indexed-names-resolved-col.parquet	taxonomic names found in the dataset under review aligned with the Catalogue of Life as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in Apache Parquet format
indexed-names-resolved-discoverlife.csv.gz	taxonomic names found in the dataset under review aligned with Discover Life bee species checklist as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped comma-separated values format
indexed-names-resolved-discoverlife.html.gz	taxonomic names found in the dataset under review aligned with Discover Life bee species checklist as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped html format
indexed-names-resolved-discoverlife.tsv.gz	taxonomic names found in the dataset under review aligned with Discover Life bee species checklist as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped tab-separated values format
indexed-names-resolved-discoverlife.parquet	taxonomic names found in the dataset under review aligned with Discover Life bee species checklist as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in Apache Parquet format
indexed-names-resolved-gbif.csv.gz	taxonomic names found in the dataset under review aligned with GBIF Backbone Taxonomy as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped comma-separated values format

filename	description
indexed-names-resolved-gbif.html.gz	taxonomic names found in the dataset under review aligned with GBIF Backbone Taxonomy as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped html format
indexed-names-resolved-gbif.tsv.gz	taxonomic names found in the dataset under review aligned with GBIF Backbone Taxonomy as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped tab-separated values format
indexed-names-resolved-gbif.parquet	taxonomic names found in the dataset under review aligned with GBIF Backbone Taxonomy as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in Apache Parquet format
indexed-names-resolved-itis.csv.gz	taxonomic names found in the dataset under review aligned with Integrated Taxonomic Information System (ITIS) as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped comma-separated values format
indexed-names-resolved-itis.html.gz	taxonomic names found in the dataset under review aligned with Integrated Taxonomic Information System (ITIS) as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped html format
indexed-names-resolved-itis.tsv.gz	taxonomic names found in the dataset under review aligned with Integrated Taxonomic Information System (ITIS) as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped tab-separated values format

filename	description
indexed-names-resolved-itis.parquet	taxonomic names found in the dataset under review aligned with Integrated Taxonomic Information System (ITIS) as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in Apache Parquet format
indexed-names-resolved-mdd.csv.gz	taxonomic names found in the dataset under review aligned with the Mammal Diversity Database as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped comma-separated values format
indexed-names-resolved-mdd.html.gz	taxonomic names found in the dataset under review aligned with Mammal Diversity Database as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped html format
indexed-names-resolved-mdd.tsv.gz	taxonomic names found in the dataset under review aligned with Mammal Diversity Database as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped tab-separated values format
indexed-names-resolved-mdd.parquet	taxonomic names found in the dataset under review aligned with Mammal Diversity Database as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in Apache Parquet format
indexed-names-resolved-ncbi.csv.gz	taxonomic names found in the dataset under review aligned with the NCBI Taxonomy as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped comma-separated values format

filename	description
indexed-names-resolved-ncbi.html.gz	taxonomic names found in the dataset under review aligned with the NCBI Taxonomy as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped html format
indexed-names-resolved-ncbi.tsv.gz	taxonomic names found in the dataset under review aligned with the NCBI Taxonomy as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped tab-separated values format
indexed-names-resolved-ncbi.parquet	taxonomic names found in the dataset under review aligned with the NCBI Taxonomy as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in Apache Parquet format
indexed-names-resolved-pbdb.csv.gz	taxonomic names found in the dataset under review aligned with the Paleobiology Database as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped comma-separated values format
indexed-names-resolved-pbdb.html.gz	taxonomic names found in the dataset under review aligned with Paleobiology Database as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped html format
indexed-names-resolved-pbdb.tsv.gz	taxonomic names found in the dataset under review aligned with Paleobiology Database as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped tab-separated values format

filename	description
indexed-names-resolved-pbdb.parquet	taxonomic names found in the dataset under review aligned with Paleobiology Database as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in Apache Parquet format
indexed-names-resolved-tpt.csv.gz	taxonomic names found in the dataset under review aligned with the Terrestrial Parasite Tracker (TPT) Taxonomic Resource as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped comma-separated values format
indexed-names-resolved-tpt.html.gz	taxonomic names found in the dataset under review aligned with the Terrestrial Parasite Tracker (TPT) Taxonomic Resource as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped html format
indexed-names-resolved-tpt.tsv.gz	taxonomic names found in the dataset under review aligned with the Terrestrial Parasite Tracker (TPT) Taxonomic Resource as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped tab-separated values format
indexed-names-resolved-tpt.parquet	taxonomic names found in the dataset under review aligned with the Terrestrial Parasite Tracker (TPT) Taxonomic Resource as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in Apache Parquet format

filename	description
indexed-names-resolved-wfo.csv.gz	taxonomic names found in the dataset under review aligned with the World of Flora Online as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped comma-separated values format
indexed-names-resolved-wfo.html.gz	taxonomic names found in the dataset under review aligned with the World of Flora Online as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped html format
indexed-names-resolved-wfo.tsv.gz	taxonomic names found in the dataset under review aligned with the World of Flora Online as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped tab-separated values format
indexed-names-resolved-wfo.parquet	taxonomic names found in the dataset under review aligned with the World of Flora Online as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in Apache Parquet format
indexed-names-resolved-worms.csv.gz	taxonomic names found in the dataset under review aligned with the World Register of Marine Species (WoRMS) as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped comma-separated values format
indexed-names-resolved-worms.html.gz	taxonomic names found in the dataset under review aligned with the World Register of Marine Species (WoRMS) as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped html format

filename	description
indexed-names-resolved-worms.tsv.gz	taxonomic names found in the dataset under review aligned with the World Register of Marine Species (WoRMS) as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped tab-separated values format
indexed-names-resolved-worms.parquet	taxonomic names found in the dataset under review aligned with the World Register of Marine Species (WoRMS) as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in Apache Parquet format
indexed-names-sample.csv	first 500 taxonomic names found in the dataset under review in comma-separated values format
indexed-names-sample.html	first 500 taxonomic names found in the dataset under review in html format
indexed-names-sample.tsv	first 500 taxonomic names found in the dataset under review in tab-separated values format
interaction.svg	diagram summarizing the data model used to index species interaction claims
nanopub-sample.trig	first 500 species interaction claims as expressed in the nanopub format (Kuhn and Dumontier 2014)
nanopub.trig.gz	species interaction claims as expressed in the nanopub format (Kuhn and Dumontier 2014)
process.svg	diagram summarizing the data review processing workflow
prov.nq	origin of the dataset under review as expressed in rdf/nquads
review.csv.gz	review notes associated with the dataset under review in gzipped comma-separated values format
review.html.gz	review notes associated with the dataset under review in gzipped html format

filename	description
review.tsv.gz	review notes associated with the dataset under review in gzipped tab-separated values format
review-sample.csv	first 500 review notes associated with the dataset under review in comma-separated values format
review-sample.html	first 500 review notes associated with the dataset under review in html format
review-sample.tsv	first 500 review notes associated with the dataset under review in tab-separated values format
review.svg	a review badge generated as part of the dataset review process
zenodo.json	metadata of this review expressed in Zenodo record metadata

Archived Dataset

Note that *data.zip* file in this archive contains the complete, unmodified archived dataset under review.

Biotic Interactions

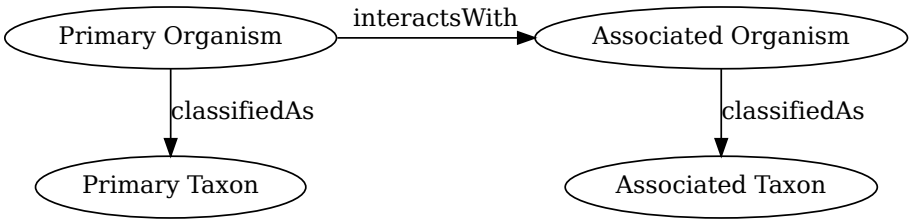

Figure 2: Biotic Interaction Data Model

In this review, biotic interactions (or biotic associations) are modeled as a primary (aka subject, source) organism interacting with an associate (aka object, target) organism. The dataset under review classified the primary/associate organisms with specific taxa. The primary and associate organisms The kind of interaction is documented as an interaction type.

The dataset under review, named globalbioticinteractions/globalfungi, has fingerprint hash://md5/4e29adcf334fd1f828dfb7d39173458f, is 966MiB in size and contains 1,698,482 interactions with 1 unique type of association (e.g., sym-

biontOf) between 2,510 primary taxa (e.g., Fagus sylvatica) and 16,053 associated taxa (e.g., Ascomycota sp.).

An exhaustive list of indexed interaction claims can be found in gzipped csv, tsv, geopackage and parquet archives. To facilitate discovery, a preview of claims available in the gzipped html page at indexed-interactions.html.gz are shown below.

The exhaustive list was used to create the following data summaries below.

Table 3: Sample of Indexed Interaction Claims

sourceTaxonName	interactionTypeName	targetTaxonName	referenceCitation
Dryas octopetala	symbiontOf	Leotiomycetes sp.	Botnen, S., Vik, U., Carlsen, T., Eidesen, P.B., Davey, M.L. and Kauserud, H., 2014. Low host specificity of root-associated fungi at an Arctic site.. Molecular ecology. doi:10.1111/mec.12646
Dryas octopetala	symbiontOf	Hyphodontiella multiseptata	Botnen, S., Vik, U., Carlsen, T., Eidesen, P.B., Davey, M.L. and Kauserud, H., 2014. Low host specificity of root-associated fungi at an Arctic site.. Molecular ecology. doi:10.1111/mec.12646

sourceTaxonName	interactionTypeName	targetTaxonName	referenceCitation
Glycine max	symbiontOf	Fusarium tonkinense	Zhang, B., Zhang, J., Liu, Y., Shi, P. and Wei, G., 2018. Co-occurrence patterns of soybean rhizosphere microbiome at a continental scale.. Soil Biology and Biochemistry. doi:10.1016/j.soilbio.2017.12.011
Glycine max	symbiontOf	Stachybotryaceae sp.	Zhang, B., Zhang, J., Liu, Y., Shi, P. and Wei, G., 2018. Co-occurrence patterns of soybean rhizosphere microbiome at a continental scale.. Soil Biology and Biochemistry. doi:10.1016/j.soilbio.2017.12.011

Table 4: Most Frequently Mentioned Interaction Types (up to 20 most frequent)

interactionTypeName	count
symbiontOf	1698482

Table 5: Most Frequently Mentioned Primary Taxa (up to 20 most frequent)

sourceTaxonName	count
Fagus sylvatica	105794
Glycine max	49183
Picea abies	41301

sourceTaxonName	count
Triticum aestivum	40771
Festuca hallii	34438
Hylocomium splendens	28325
Panicum virgatum	27125
Populus deltoides	26532
Malus domestica	25687
Populus trichocarpa x deltoides	22409
Vitis vinifera	20540
Vitis vinifera L.	20052
Fraxinus excelsior	19362
Schizachyrium scoparium	18771
Thymus zygis	18505
Fuscospora cliffortioides	17378
Bouteloua gracilis	16809
Plantago lanceolata	15157
Holcus lanatus	15123

Table 6: Most Frequently Mentioned Associate Taxa (up to 20 most frequent)

targetTaxonName	count
Ascomycota sp.	19668
Fungi sp.	16697
Cladosporium herbarum	13592
Helotiales sp.	12814
Pleosporales sp.	11168
Dothideomycetes sp.	9143
Cladophialophora sp.	9023
Sordariomycetes sp.	9009
Agaricomycetes sp.	8080
Hypocreales sp.	7626
Sordariales sp.	7071
Eukaryota sp.	7007
Chaetothyriales sp.	6649
Leotiomycetes sp.	6478
Mortierella sp.	6220
Aureobasidium pullulans	6150
Agaricales sp.	6140
Alternaria eichhorniae	5956
Penicillium sp.	5615

Table 7: Most Frequent Interactions between Primary and Associate Taxa (up to 20 most frequent)

sourceTaxonName	interactionTypeName	targetTaxonName	count
Fagus sylvatica	symbiontOf	Ascomycota sp.	763
Fagus sylvatica	symbiontOf	Helotiales sp.	706
Fagus sylvatica	symbiontOf	Fungi sp.	689
Fagus sylvatica	symbiontOf	Cladosporium herbarum	626
Fuscospora cliffortioides	symbiontOf	Helotiales sp.	610
Fagus sylvatica	symbiontOf	Hyaloscyphaceae sp.	605
Fuscospora cliffortioides	symbiontOf	Oidiodendron sp.	600
Fagus sylvatica	symbiontOf	Cladophialophora sp.	595
Fagus sylvatica	symbiontOf	Leotiomycetes sp.	579
Festuca hallii	symbiontOf	Fungi sp.	572
Fagus sylvatica	symbiontOf	Dothideomycetes sp.	570
Festuca hallii	symbiontOf	Mortierella sp.	570
Fagus sylvatica	symbiontOf	Pleosporales sp.	559
Festuca hallii	symbiontOf	Ascomycota sp.	547
Fuscospora cliffortioides	symbiontOf	Penicillium sp.	543
Fagus sylvatica	symbiontOf	Hypocreaceae sp.	539
Fagus sylvatica	symbiontOf	Sordariomycetes sp.	538
Fagus sylvatica	symbiontOf	Agaricomycetes sp.	529
Fuscospora cliffortioides	symbiontOf	Oidiodendron pilicola	520

Interaction Networks

The figures below provide a graph view on the dataset under review. The first shows a summary network on the kingdom level, and the second shows how interactions on the family level. It is important to note that both network graphs were first aligned taxonomically using the Catalogue of Life. Please refer to the original (or verbatim) taxonomic names for a more original view on the interaction data.

You can download the indexed dataset under review at indexed-interactions.csv.gz. A tab-separated file can be found at indexed-interactions.tsv.gz

Geospatial Distribution

If geospatial information was extracted from the dataset under review, the map below will show their distribution. These maps were generated using MapServer (McKenna et al. 2025) tools configured via map configuration indexed-interactions.map :

```
MAP
  SIZE 1600 800
  EXTENT -180 -90 180 90
```

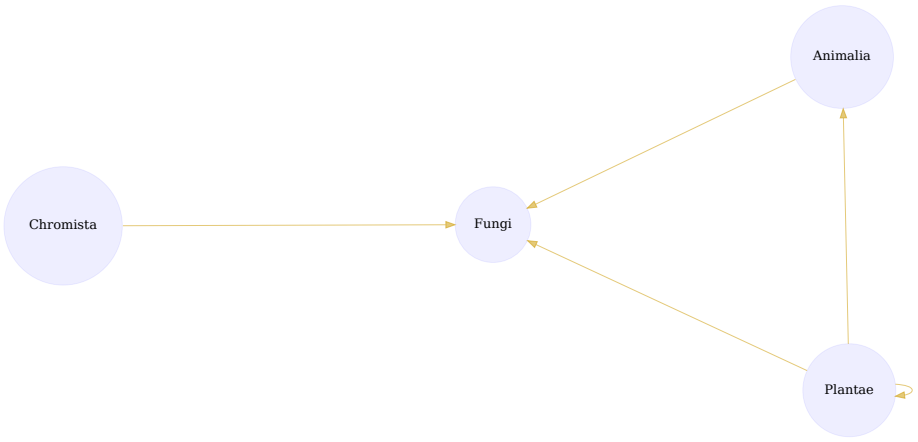

Figure 3: Interactions on taxonomic kingdom rank as interpreted by the Catalogue of Life download svg

```
  PROJECTION
    "init=epsg:4326"
  END
  LAYER # MODIS WMS map from NASA
    NAME         "modis_nasa"
    TYPE         RASTER
    OFFSITE      0 0 0
    STATUS       ON
    CONNECTIONTYPE WMS
    CONNECTION "https://gibs.earthdata.nasa.gov/wms/epsg4326/best/wms.cgi?"

    METADATA
      "wms_srs" "EPSG:4326"
      "wms_name" "OSM_Land_Water_Map"
      "wms_server_version" "1.1.1"
      "wms_format" "image/jpeg"
    END
    CLASS
      STYLE
        COLOR        232 232 232
        OUTLINECOLOR 32 32 32
      END
    END
  END
  LAYER
    NAME "indexed-interactions"
    TYPE POLYGON
```

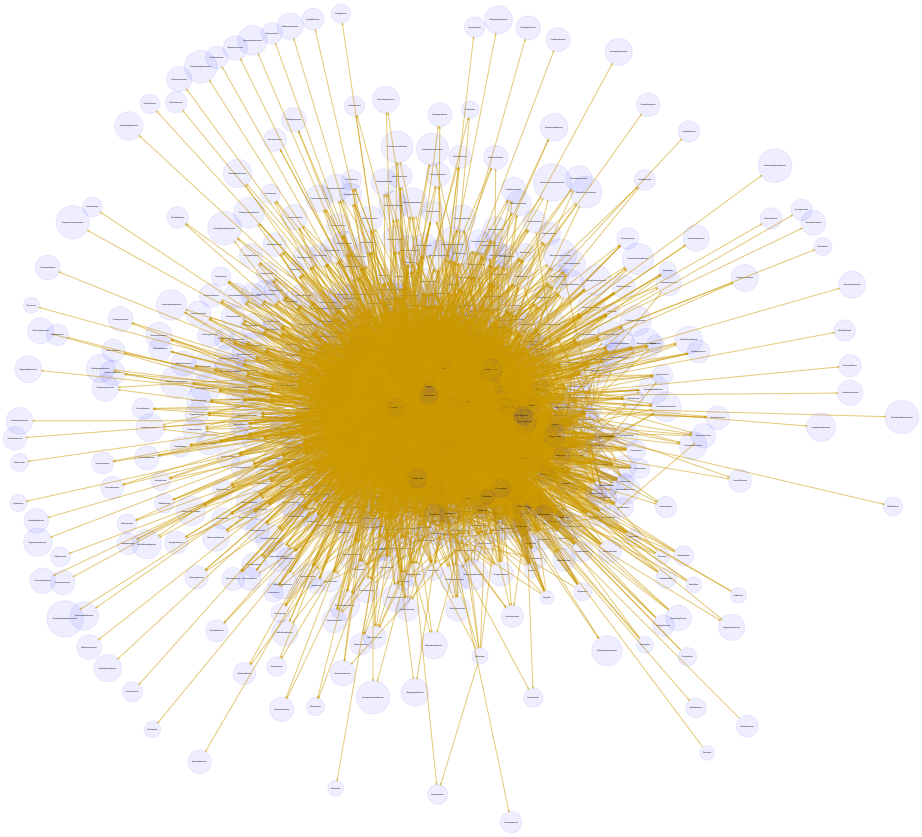

Figure 4: Interactions on the taxonomic family rank as interpreted by the Catalogue of Life. download svg

```
    STATUS ON
    CONNECTIONTYPE OGR
    CONNECTION "indexed-interactions-h3.gpkg"
    DATA "indexed-interactions-h3"
    CLASS
      STYLE
        COLORRANGE 253.0 231.0 37.0 32.0 164.0 134.0
        DATARANGE 0.47712125471966244 4.747404038684875
        RANGEITEM "log_number_of_records"
        OUTLINECOLOR 0 0 0
      END
    END
  END
END
```

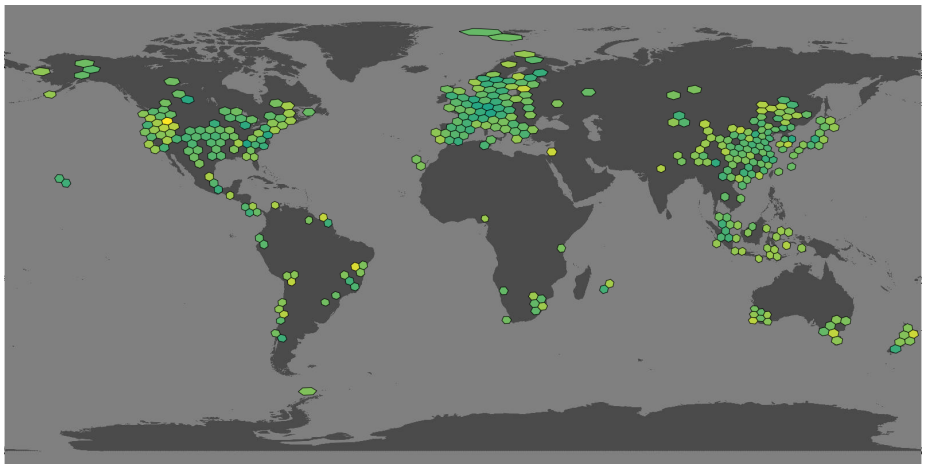

Figure 5: Hexagonal grid cells indicate that interactions claims are available for selected geospatial area: light yellow means relatively fewer claims, dark green relatively more claims.

Associated data can be found in the geopackage files at indexed-interactions .gpkg for point data and indexed-interactions-h3.gpkg for data clustered in geospatial h3 hexagonals.

Learn more about the structure of this download at GloBI website, by opening a GitHub issue, or by sending an email.

Another way to discover the dataset under review is by searching for it on the GloBI website.

Taxonomic Alignment

As part of the review, all names are aligned against various name catalogs (e.g., col, ncbi, discoverlife, gbif, itis, wfo, mdd, tpt, pbdb, and worms). These alignments can help review name usage or aid in selecting of a suitable taxonomic name resource.

Table 8: Sample of Name Alignments

providedName	relationName	resolvedCatalogName	resolvedName
Aaosphaeria arxii	HAS_ACCEPTED_NAME	col	Aaosphaeria arxii
Abies alba	HAS_ACCEPTED_NAME	col	Abies alba
Abies amabilis	HAS_ACCEPTED_NAME	col	Abies amabilis
Abies balsamea	HAS_ACCEPTED_NAME	col	Abies balsamea

Table 9: Distribution of Taxonomic Ranks of Aligned Names by Catalog. Names that were not aligned with a catalog are counted as NAs. So, the total number of unaligned names for a catalog will be listed in their NA row.

resolvedCatalogName	resolvedRank	count
col	NA	1975
col	class	28
col	family	400
col	genus	1685
col	kingdom	1
col	order	132
col	phylum	9
col	section	1
col	species	14268
col	subfamily	3
col	subgenus	2
col	subspecies	27
col	subtribe	1
col	superorder	1
col	variety	6
discoverlife	NA	18503
gbif	NA	2135
gbif	class	25
gbif	family	424
gbif	form	7
gbif	genus	1722
gbif	kingdom	1
gbif	order	129

resolvedCatalogName	resolvedRank	count
gbif	phylum	7
gbif	species	14046
gbif	subspecies	50
gbif	variety	31
itis	NA	16881
itis	class	21
itis	division	6
itis	family	277
itis	genus	326
itis	infrakingdom	2
itis	infraorder	1
itis	kingdom	1
itis	order	98
itis	species	882
itis	subdivision	2
itis	subkingdom	1
itis	subspecies	3
itis	variety	3
mdd	NA	18503
ncbi	NA	2468
ncbi	clade	1
ncbi	class	30
ncbi	family	390
ncbi	genus	1707
ncbi	infraorder	1
ncbi	kingdom	3
ncbi	order	129
ncbi	phylum	10
ncbi	species	13739
ncbi	subfamily	7
ncbi	subgenus	1
ncbi	subphylum	2
ncbi	subspecies	5
ncbi	superkingdom	1
ncbi	superorder	1
ncbi	tribe	7
ncbi	varietas	4
pbdb	NA	18156
pbdb	class	5
pbdb	family	53
pbdb	genus	187
pbdb	kingdom	3
pbdb	order	13
pbdb	phylum	3

resolvedCatalogName	resolvedRank	count
pbdb	species	75
pbdb	subfamily	1
pbdb	subkingdom	1
pbdb	subphylum	1
pbdb	unranked clade	11
tpt	NA	18502
tpt	genus	1
wfo	NA	17237
wfo	family	43
wfo	genus	196
wfo	order	4
wfo	section	1
wfo	species	1009
wfo	subfamily	4
wfo	subspecies	12
wfo	tribe	2
wfo	variety	5
worms	NA	17095
worms	class	19
worms	family	225
worms	genus	506
worms	infrakingdom	1
worms	kingdom	2
worms	order	82
worms	phylum (division)	5
worms	species	562
worms	subfamily	1
worms	subkingdom	1
worms	subphylum (subdivision)	1
worms	subspecies	2
worms	superorder	1
worms	variety	3

Table 10: Name relationship types per catalog. Name relationship type "NONE" means that a name was not recognized by the associated catalog. "SAME_AS" indicates either a "HAS_ACCEPTED_NAME" or "SYNONYM_OF" name relationship type. We recognize that "SYNONYM_OF" encompasses many types of nomenclatural synonymies (ICZN 1999) (e.g., junior synonym, senior synonyms).

resolvedCatalogName	relationName	count
col	HAS_ACCEPTED_NAME	16001
col	SYNONYM_OF	1024
col	NONE	1976
discoverlife	NONE	18562
gbif	HAS_ACCEPTED_NAME	16620
gbif	SYNONYM_OF	2173
gbif	NONE	2136
itis	NONE	16884
itis	HAS_ACCEPTED_NAME	1632
itis	SYNONYM_OF	95
mdd	NONE	18562
ncbi	SAME_AS	15584
ncbi	NONE	2469
ncbi	SYNONYM_OF	566
pbdb	NONE	18184
pbdb	HAS_ACCEPTED_NAME	385
pbdb	SYNONYM_OF	22
tpt	NONE	18561
tpt	HAS_ACCEPTED_NAME	1
wfo	NONE	17240
wfo	HAS_ACCEPTED_NAME	1238
wfo	SYNONYM_OF	211
wfo	HAS_UNCHECKED_NAME	115
worms	NONE	17123
worms	HAS_ACCEPTED_NAME	1452
worms	SYNONYM_OF	80

Table 11: List of Available Name Alignment Reports

catalog name	alignment results
col	associated names alignments report in gzipped html, csv, and tsv)
ncbi	associated names alignments report in gzipped html, csv, and tsv)

catalog name	alignment results
discoverlife	associated names alignments report in gzipped html, csv, and tsv)
gbif	associated names alignments report in gzipped html, csv, and tsv)
itis	associated names alignments report in gzipped html, csv, and tsv)
wfo	associated names alignments report in gzipped html, csv, and tsv)
mdd	associated names alignments report in gzipped html, csv, and tsv)
tpt	associated names alignments report in gzipped html, csv, and tsv)
pbdb	associated names alignments report in gzipped html, csv, and tsv)
worms	associated names alignments report in gzipped html, csv, and tsv)

Additional Reviews

Elton, Nomer, and other tools may have difficulties interpreting existing species interaction datasets. Or, they may misbehave, or otherwise show unexpected behavior. As part of the review process, detailed review notes are kept that document possibly misbehaving, or confused, review bots. An sample of review notes associated with this review can be found below.

Table 12: First few lines in the review notes.

reviewDate	reviewCommentType	reviewComment
2026-03-30T05:20:23Z	note	target taxon name missing
2026-03-30T05:20:23Z	note	target taxon name missing
2026-03-30T05:20:23Z	note	target taxon name missing
2026-03-30T05:20:23Z	note	target taxon name missing

In addition, you can find the most frequently occurring notes in the table below.

Table 13: Most frequently occurring review notes, if any.

reviewComment	count
source taxon name missing: using institution-Code/collectionCode/collectionId/catalogNumber/occurrenceId as placeholder	5985205
found unsupported interaction type with name: [deadwood]	133903
target taxon name missing	71957
found unsupported interaction type with name: [lichen]	9100

For additional information on review notes, please have a look at the first 500 Review Notes in html format or the download full gzipped csv or tsv archives.

GloBI Review Badge

As part of the review, a review badge is generated. This review badge can be included in webpages to indicate the review status of the dataset under review.

Figure 6: Picture of a GloBI Review Badge [3]

Note that if the badge is green, no review notes were generated. If the badge is yellow, the review bots may need some help with interpreting the species interaction data.

GloBI Index Badge

If the dataset under review has been registered with GloBI, and has been succesfully indexed by GloBI, the GloBI Index Status Badge will turn green. This means that the dataset under review was indexed by GloBI and is available through GloBI services and derived data products.

Figure 7: Picture of a GloBI Index Badge [4]

[3] Up-to-date status of the GloBI Review Badge can be retrieved from the GloBI Review Depot

[4] Up-to-date status of the GloBI Index Badge can be retrieved from GloBI's API

If you'd like to keep track of reviews or index status of the dataset under review, please visit GloBI's dataset index [5] for badge examples.

Discussion

This review and archive provides a means of creating citable versions of datasets that change frequently. This may be useful for dataset managers, including natural history collection data managers, as a backup archive of a shared Darwin Core archive. It also serves as a means of creating a trackable citation for the dataset in an automated way, while also including some information about the contents of the dataset.

This review aims to provide a perspective on the dataset to aid in understanding of species interaction claims discovered. However, it is important to note that this review does *not* assess the quality of the dataset. Instead, it serves as an indication of the open-ness[6] and FAIRness (Wilkinson et al. 2016; Trekels et al. 2023) of the dataset: to perform this review, the data was likely openly available, **F**indable, **A**ccessible, **I**nteroperable and **R**eusable. The current Open-FAIR assessment is qualitative, and a more quantitative approach can be implemented with specified measurement units.

This report also showcases the reuse of machine-actionable (meta)data, something highly recommended by the FAIR Data Principles (Wilkinson et al. 2016). Making (meta)data machine-actionable enables more precise procesing by computers, enabling even naive review bots like Nomer and Elton to interpret the data effectively. This capability is crucial for not just automating the generation of reports, but also for facilitating seamless data exchanges, promoting interoperability.

Acknowledgements

We thank the many humans that created us and those who created and maintained the data, software and other intellectual resources that were used for producing this review. In addition, we are grateful for the natural resources providing the basis for these human and bot activities. Also, thanks to https://github.com/zygoballus for helping improve the layout of the review tables.

[5] At time of writing (2026-03-30) the version of the GloBI dataset index was available at https://globalbioticinteractions.org/datasets

[6] According to http://opendefinition.org/: "Open data is data that can be freely used, reused and redistributed by anyone - subject only, at most, to the requirement to attribute and sharealike."

Author contributions

Nomer was responsible for name alignments. Elton carried out dataset extraction, and generated the review notes. Preston tracked, versioned, and packaged, the dataset under review.

References

Elliott, Michael, Jorrit Poelen, Icaro Alzuru, Emilio Berti, and partha04patel. 2025. "Bio-Guoda/Preston: 0.10.5." Zenodo. https://doi.org/10.5281/zenodo.14662206.

ICZN. 1999. "International Code of Zoological Nomenclature." The International Trust for Zoological Nomenclature, London, UK. https://www.iczn.org/the-code/the-code-online/.

Kuhn, Tobias, and Michel Dumontier. 2014. "Trusty URIs: Verifiable, Immutable, and Permanent Digital Artifacts for Linked Data." In *The Semantic Web: Trends and Challenges*, edited by Valentina Presutti, Claudia d'Amato, Fabien Gandon, Mathieu d'Aquin, Steffen Staab, and Anna Tordai, 395–410. Cham: Springer International Publishing.

Kuhn, Tobias, Jorrit Poelen, and Katrin Leinweber. 2025. "Globalbioticinteractions/Elton: 0.15.1." Zenodo. https://doi.org/10.5281/zenodo.14927734.

McKenna, Jeff, Steve Lime, Thomas Bonfort, Jérome Boué, Howard Butler, Seth Girvin, Tom Kralidis, et al. 2025. "MapServer." Zenodo. https://doi.org/10.5281/zenodo.17807263.

Poelen, Jorrit H. (ed.). 2024. "Nomer Corpus of Taxonomic Resources Hash://Sha256/ B60c0d25a16ae77b24305782017b1a270b79b5d1746f832650F2027ba536e276 Hash://Md5/17f1363a277ee0e4ecaf1b91c665e47e." Zenodo. https://doi.org/10.5281/zenodo.12695629.

Poelen, Jorrit H., James D. Simons, and Chris J. Mungall. 2014. "Global Biotic Interactions: An Open Infrastructure to Share and Analyze Species-Interaction Datasets." *Ecological Informatics* 24 (November): 148–59. https://doi.org/10.1016/j.ecoinf.2014.08.005.

Poelen, Jorrit, Katja Seltmann, and Daniel Mietchen. 2024. "Globalbioticinteractions/Globinizer: 0.4.0." Zenodo. https://doi.org/10.5281/zenodo.10647565.

Salim, José Augusto, and Jorrit Poelen. 2025. "Globalbioticinteractions/Nomer: 0.5.15." Zenodo. https://doi.org/10.5281/zenodo.14893840.

Trekels, Maarten, Debora Pignatari Drucker, José Augusto Salim, Jeff Ollerton, Jorrit Poelen, Filipi Miranda Soares, Max Rünzel, Muo Kasina, Quentin Groom, and Mariano Devoto. 2023. "WorldFAIR Project (D10.1) Agriculture-related pollinator data standards use cases report." Zenodo. https://doi.org/10.5281/zenodo.8176978.

Wilkinson, Mark D., Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, et al. 2016. "The FAIR Guiding Principles for Scientific Data Management and Stewardship." *Scientific Data* 3 (1). https://doi.org/10.1038/sdata.2016.18.