

Versioned Archive and Review of Biotic
Interactions and Taxon Names Found within
globalbioticinteractions/globalamfungi
hash://md5/ed0a063f879594b2a724016570963638

by Nomer, Elton and Preston, three naive review bots
review@globalbioticinteractions.org
<https://globalbioticinteractions.org/contribute>

2026-06-03

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/globalamfungi, has fingerprint hash://md5/ed0a063f879594b2a724016570963638, is 1.17GiB in size and contains 2,251,394 interactions with 1 unique type of association (e.g., hasArbuscularMycorrhizalHost) between 87 primary taxa (e.g., Funneliformis mosseae) and 579 associated taxa (e.g., Zea mays). This report includes detailed summaries of interaction data, a taxonomic review from multiple catalogs, and an archived version of the dataset from which the reviews are derived.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	2
Data Review and Archive	2
Methods	2
Results	4
Files	4
Archived Dataset	4
Biotic Interactions	4
Interaction Networks	9

Geospatial Distribution	9
Taxonomic Alignment	12
Additional Reviews	15
GloBI Review Badge	16
GloBI Index Badge	16
Discussion	17
Acknowledgements	17
Author contributions	18
Appendix A. Review Files	18
References	27

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 :

Geoffrey Zahn. (2025). gzahn/GlobalAMF_Database: Initial release (Version v1.0) [Computer software]. Zenodo.
<https://doi.org/10.5281/ZENODO.14812876> <https://github.com/globalbioticinteractions/globalamfungi>,
2026-06-02T15:12:00.626Z hash://md5/ed0a063f879594b2a724016570963638

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

tool name	version
elton	0.16.11
nomer	0.6.5
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. 1.17GiB) under review
elton pull globalbioticinteractions/globalamfungi

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

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

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

or visually, in a process diagram.

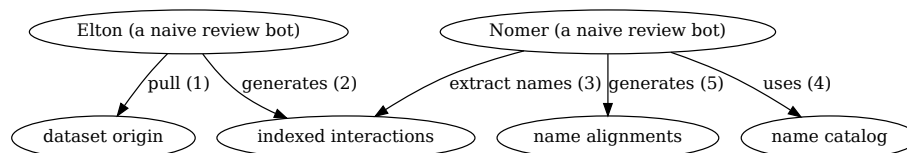


Figure 1: Review Process Overview

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

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

Results

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

Files

An extensive list of files produced as part of the review process can be found in Appendix A. Review Files.

Archived Dataset

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

Biotic Interactions

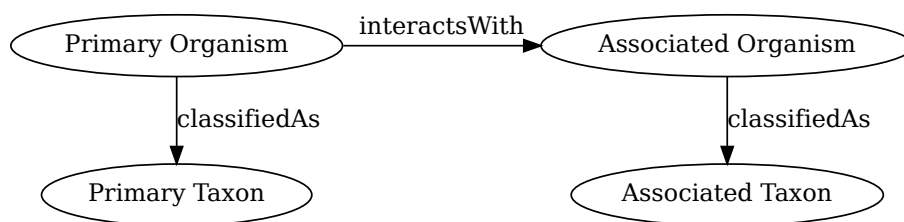


Figure 2: Biotic Interaction Data Model

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/globalamfungi`, has fingerprint hash://md5/ed0a063f879594b2a724016570963638, is 1.17GiB in size and contains 2,251,394 interactions with 1 unique type of association (e.g., `hasArbuscularMycorrhizalHost`) between 87 primary taxa (e.g., `Funnelformis mosseae`) and 579 associated taxa (e.g., `Zea mays`).

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.

²Disclaimer: 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.

Table 2: Sample of Indexed Interaction Claims

sourceTaxonName	interactionType	targetTaxonName	referenceCitation
Acaulospora alpina	hasArbuscularMycorrhiza	ArizadHost elatius	Řezáčová, V., Řezáč, M., Gryndler, M., Hršelová, H., Gryndlerová, H. and Michalová, T., 2021. Plant invasion alters community structure and decreases diversity of arbuscular mycorrhizal fungal communities.
Acaulospora alpina	hasArbuscularMycorrhiza	ShiidaHost canadensis	Řezáčová, V., Řezáč, M., Gryndler, M., Hršelová, H., Gryndlerová, H. and Michalová, T., 2021. Plant invasion alters community structure and decreases diversity of arbuscular mycorrhizal fungal communities.

sourceTaxonName	interactionTypeName	targetTaxonName	referenceCitation
Acaulospora alpina	hasArbuscularMycorrhizalHost	ArizalHost elatius	Řezáčová, V., Řezáč, M., Gryndler, M., Hršelová, H., Gryndlerová, H. and Michalová, T., 2021. Plant invasion alters community structure and decreases diversity of arbuscular mycorrhizal fungal communities.
Acaulospora brasiliensis	hasArbuscularMycorrhizalHost	ArizalHost sphaerocephalus	Řezáčová, V., Řezáč, M., Gryndler, M., Hršelová, H., Gryndlerová, H. and Michalová, T., 2021. Plant invasion alters community structure and decreases diversity of arbuscular mycorrhizal fungal communities.

Table 3: Most Frequently Mentioned Interaction Types (up to 20 most frequent)

interactionTypeName	count
hasArbuscularMycorrhizalHost	2251394

Table 4: Most Frequently Mentioned Primary Taxa (up to 20 most frequent)

sourceTaxonName	count
Funneliformis mosseae	496324
Entrophospora lamellosa	269038
Rhizophagus irregularis	238278
Rhizophagus intraradices	183074
Diversispora aurantia	170805
Funneliformis caledonius	141749
Racocetra castanea	109962
Acaulospora koskei	86844
Gigaspora decipiens	84694
Rhizophagus manihotis	51874
Sacculospora felinonii	37065
Rhizophagus clarus	36694
Rhizophagus prolifer	32710
Septoglomus viscosum	29022
Paraglomus occultum	27979
Microkamsiensia perpusilla	24234
Sclerocystis sinuosa	23700
Diversispora epigaea	22758
Cetraspora gilmorei	18749

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

targetTaxonName	count
Zea mays	216462
Triticum aestivum	189145
Glycine max	139103
Theobroma cacao	108959
Bromus tectorum	103585
Daucus carota	97607
Tanacetum vulgare	92401
Medicago sativa	85975
Pancreaticum maritimum	73891
Artemisia tridentata	69111
Arrhenatherum elatius	66261
Capsicum annuum	58374
Eragrostis tef	56530
Solidago canadensis	55929
Hordeum vulgare	38819

targetTaxonName	count
Citrus sinensis	37595
Achillea millefolium	37371
Plantago lanceolata	35204
Asclepias speciosa	33977

Table 6: Most Frequent Interactions between Primary and Associate Taxa (up to 20 most frequent)

sourceTaxonName	interactionType	targetTaxonName	count
Funneliformis mosseae	hasArbuscularMycorrhiza	ZizaniaHost	112139
Funneliformis mosseae	hasArbuscularMycorrhiza	BrizaHost	93116
Entrophospora lamellosa	hasArbuscularMycorrhiza	ThizidHost aestivum	90669
Funneliformis mosseae	hasArbuscularMycorrhiza	GlycineHost Host	89271
Funneliformis caledonius	hasArbuscularMycorrhiza	ThizidHost aestivum	82983
Rhizophagus irregularis	hasArbuscularMycorrhiza	ArizidHost tridentata	57418
Diversispora aurantia	hasArbuscularMycorrhiza	PhaseolHost Host	56661
Funneliformis mosseae	hasArbuscularMycorrhiza	MedicagoHost Host	54856
Racocetra castanea	hasArbuscularMycorrhiza	ThizidHost cacao	54743
Diversispora aurantia	hasArbuscularMycorrhiza	ThizidHost vulgare	50379
Gigaspora decipiens	hasArbuscularMycorrhiza	ThizidHost cacao	40371
Funneliformis mosseae	hasArbuscularMycorrhiza	ClaytonHost annuum	39572
Sacculospora felinonii	hasArbuscularMycorrhiza	PhaseolHost maritimum	37062
Rhizophagus intraradices	hasArbuscularMycorrhiza	HypochaHost hirta	30970
Funneliformis mosseae	hasArbuscularMycorrhiza	CitrusHost Host	29180
Diversispora aurantia	hasArbuscularMycorrhiza	ShizidHost canadensis	26535

sourceTaxonName	interactionType	targetTaxonName	count
Rhizophagus intraradices	hasArbuscularMycorrhiza	Asclepias speciosa	26512
Acaulospora koskei	hasArbuscularMycorrhiza	Arizadendron elatius	25552
Rhizophagus clarus	hasArbuscularMycorrhiza	Zizania	25008

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.

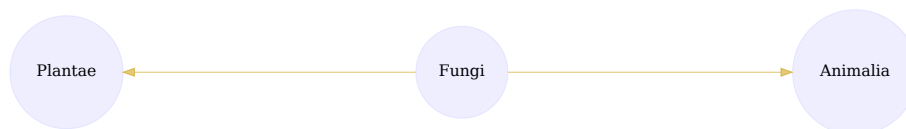


Figure 3: Interactions on taxonomic kingdom rank as interpreted by the Catalogue of Life download svg

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

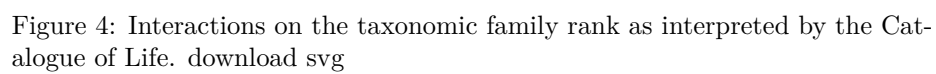
Geospatial Distribution

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
  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
LAYER
  NAME "indexed-interactions"
  TYPE POLYGON
  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 0.3010299956639812 5.611386572725673
      RANGEITEM "log_number_of_records"
      OUTLINECOLOR 0 0 0
    END
  END
END
END

```

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.

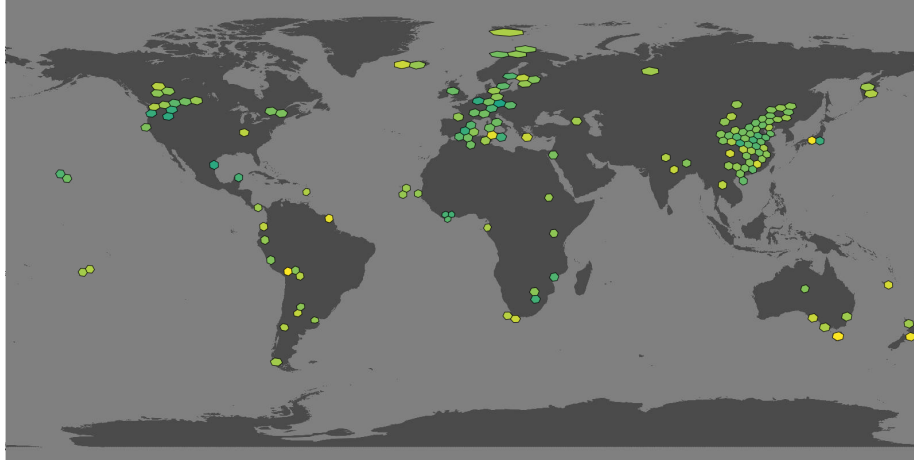


Figure 5: Hexagonal grid cells indicate that interactions claims are available for selected geospatial area: light yellow means relatively fewer claims, dark green relatively more claims.

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 7: Sample of Name Alignments

providedName	relationName	resolvedCatalogName	resolvedName
Abies lasiocarpa	HAS_ACCEPTED_NAME	col	Abies lasiocarpa
Abies lasiocarpa	SYNONYM_OF	col	Abies grandis
Acacia koa	HAS_ACCEPTED_NAME	col	Acacia koa
Acacia nigrescens	SYNONYM_OF	col	Senegalia nigrescens

Table 8: 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	45
col	family	1
col	genus	29

resolvedCatalogName	resolvedRank	count
col	species	580
col	subspecies	23
col	variety	5
discoverlife	NA	663
gbif	NA	37
gbif	family	1
gbif	genus	29
gbif	species	587
gbif	subspecies	26
gbif	variety	10
itis	NA	284
itis	family	1
itis	genus	29
itis	species	346
itis	subspecies	1
itis	variety	6
mdd	NA	663
ncbi	NA	58
ncbi	family	1
ncbi	genus	29
ncbi	species	570
ncbi	subfamily	1
ncbi	subspecies	3
ncbi	varietas	1
pbdb	NA	636
pbdb	family	1
pbdb	genus	13
pbdb	species	12
pbdb	unranked clade	1
tpt	NA	663
wfo	NA	121
wfo	family	1
wfo	genus	29
wfo	species	509
wfo	subspecies	8
wfo	variety	1
worms	NA	509
worms	family	1
worms	genus	15
worms	species	137
worms	subspecies	1
worms	variety	2

Table 9: 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	559
col	SYNONYM_OF	257
col	NONE	45
discoverlife	NONE	670
gbif	HAS_ACCEPTED_NAME	736
gbif	SYNONYM_OF	362
gbif	NONE	37
itis	HAS_ACCEPTED_NAME	338
itis	NONE	288
itis	SYNONYM_OF	70
mdd	NONE	670
ncbi	SAME_AS	566
ncbi	SYNONYM_OF	50
ncbi	NONE	58
pbdb	NONE	641
pbdb	HAS_ACCEPTED_NAME	28
pbdb	SYNONYM_OF	1
tpt	NONE	670
wfo	SYNONYM_OF	135
wfo	HAS_ACCEPTED_NAME	494
wfo	NONE	125
wfo	HAS_UNCHECKED_NAME	44
worms	NONE	513
worms	HAS_ACCEPTED_NAME	167
worms	SYNONYM_OF	45

Table 10: 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)
tppt	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 11: First few lines in the review notes.

reviewDate	reviewCommentType	reviewComment
2026-06-03T10:11:56Z	note	found [6] columns, but only [5] columns are defined: ignoring remaining undefined columns.
2026-06-03T10:11:56Z	note	found [6] columns, but only [5] columns are defined: ignoring remaining undefined columns.
2026-06-03T10:11:56Z	note	found [6] columns, but only [5] columns are defined: ignoring remaining undefined columns.

reviewDate	reviewCommentType	reviewComment
2026-06-03T10:11:56Z	note	found [6] columns, but only [5] columns are defined: ignoring remaining undefined columns.

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

Table 12: Most frequently occurring review notes, if any.

reviewComment	count
found malformed doi [Applied Soil Ecology]	517899
target taxon name missing	235234
found malformed doi [Plant and Soil]	234475
found malformed doi [Symbiosis]	208411

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 successfully 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.

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

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



Figure 7: Picture of a GloBI Index Badge ⁴

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-ness⁶ 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 processing 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

⁵At time of writing (2026-06-03) the version of the GloBI dataset index was available at <https://globalbioticinteractions.org/datasets>

⁶According 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.”

to <https://github.com/zygoballus> for helping improve the layout of the review tables.

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.

Appendix A. Review 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
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)

filename	description
indexed-interactions-col-kingdom-col-kingdom.svg	network diagram showing the taxon kingdom to taxon kingdom interaction claims in the dataset under review as interpreted by the Catalogue of Life via Nomen 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
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

filename	description
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
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

filename	description
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
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

filename	description
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
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

filename	description
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
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

filename	description
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
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

filename	description
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
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

filename	description
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
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

filename	description
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
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

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>

- rg/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, Jérôme 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/ B60c0d25a16ae77b24305782017b1a270b79b5d1746f832650 F2027ba536e276 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>.