Versioned Archive and Review of Biotic Interactions and Taxon Names Found within globalbioticinteractions/refuted-biotic-interactions-by-eol hash://md5/5f958b14eb086b859bef150620376d2e

by Nomer, Elton and Preston, three naive review bots
review@globalbioticinteractions.org
https://globalbioticinteractions.org/contribute
https://github.com/globalbioticinteractions/refuted-biotic-interactions-by-eol/issues

2025-09-13

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/refuted-biotic-interactions-by-eol, has fingerprint hash://md5/5f958b14eb086b859bef150620376d2e, is 135MiB in size and contains 103,653 interactions with 13 unique types of associations (e.g., visitsFlowersOf) between 12,256 primary taxa (e.g., Apis mellifera) and 6,358 associated taxa (e.g., Achillea millefolium). This report includes detailed summaries of interaction data, a taxonomic review from multiple catalogs, and an archived version of the dataset from which the reviews are derived.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 13
Biotic Interactions 13
Interaction Networks 17
Taxonomic Alignment 18
Additional Reviews 25
GloBI Review Badge 25
GloBI Index Badge 26

Discussion **26**

Acknowledgements **27**

Author contributions **27**

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 :

> Katja Schulz. 2020. Collection of refuted species associations claims provided by Encyclopedia of Life. https://github.com/globalbioticinteractions/refuted-biotic-interactions-by-eol/archive/67dc570987c6897fca00b81247edbdb0b1861a36.zip 2025-09-06T02:52:52.830Z hash://md5/5f958b14eb086b859bef150620376d2e

For additional metadata related to this dataset, please visit https://github.com/globalbioticinteractions/refuted-biotic-interactions-by-eol 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	0.10.1
elton	0.15.13
nomer	0.5.17
globinizer	0.4.0
mlr	6.0.0
jq	1.6
yq	4.25.3
pandoc	3.1.6.1
duckdb	1.3.1

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

```
# get versioned copy of the dataset (size approx.  135MiB) under review
elton pull globalbioticinteractions/refuted-biotic-interactions-by-eol

# generate review notes
elton review globalbioticinteractions/refuted-biotic-interactions-by-eol\
 > review.tsv

# export indexed interaction records
elton interactions globalbioticinteractions/refuted-biotic-interactions-by-eol\
 > interactions.tsv

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

or visually, in a process diagram.

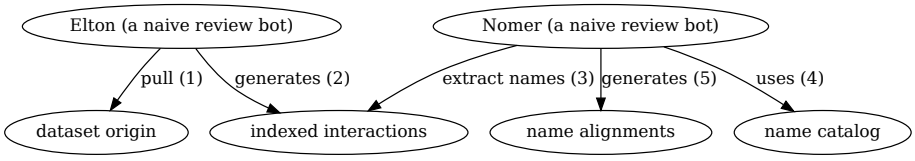


Figure 1: Review Process Overview

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

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
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)

[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-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-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

filename	description
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
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

filename	description
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
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

filename	description
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
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

filename	description
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
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

filename	description
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
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

filename	description
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
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

filename	description
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
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)

filename	description
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

Archived Dataset

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

Biotic Interactions

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/refuted-biotic-

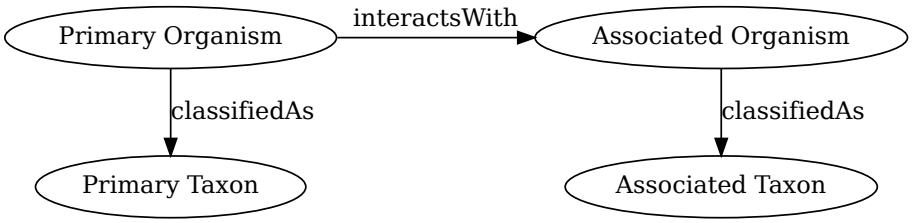


Figure 2: Biotic Interaction Data Model

interactions-by-eol, has fingerprint hash://md5/5f958b14eb086b859bef150620376d2e, is 135MiB in size and contains 103,653 interactions with 13 unique types of associations (e.g., visitsFlowersOf) between 12,256 primary taxa (e.g., Apis mellifera) and 6,358 associated taxa (e.g., Achillea millefolium).

An exhaustive list of indexed interaction claims can be found in gzipped csv, tsv 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
Diglossa humeralis	visitsFlowersOf	Tristerix longebracteatus	Records of organisms other than plants having flower visitors are probably errors
Diglossa brunneiventris	visitsFlowersOf	Tristerix longebracteatus	Records of organisms other than plants having flower visitors are probably errors
Ectemnius trifasciatus	visitsFlowersOf	Zizia aurea	Records of organisms other than plants having flower visitors are probably errors

sourceTaxonName	interactionTypeName	targetTaxonName	referenceCitation
Ectemnius trifasciatus	visitsFlowersOf	Pastinaca sativa	Records of organisms other than plants having flower visitors are probably errors

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

interactionTypeName	count
visitsFlowersOf	77356
parasiteOf	10914
flowersVisitedBy	5728
eats	4504
hasHost	2947
pathogenOf	1061
visits	554
preysOn	525
endoparasiteOf	29
ectoparasiteOf	24
parasitoidOf	9
kleptoparasiteOf	1
laysEggsOn	1

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

sourceTaxonName	count
Apis mellifera	2062
Digenea	1274
Bombus pascuorum	794
Episyrphus balteatus	715
Bombus terrestris	693
Bombus lapidarius	653
Eristalis tenax	642
Sphaerophoria scripta	569
Saxifraga	537
Apoidea	471
Gloiosiphonia	427

sourceTaxonName	count
Bombus impatiens	407
Hymenolepis	405
Bombus	404
Diptera	402
Syritta pipiens	386
Lepidoptera	378
Bombus lucorum	370
Lasioglossum	356

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

targetTaxonName	count
Achillea millefolium	1306
Brassica napus	1145
Monarda	1133
Daucus carota	918
Trifolium pratense	880
Cirsium arvense	849
Rhinolophus bat coronavirus HKU2	846
Leucanthemum vulgare	832
Trifolium repens	806
Helianthus annuus	796
Asteraceae	781
Heracleum sphondylium	763
Swine acute diarrhea syndrome related coronavirus	667
Lotus corniculatus	649
Knautia arvensis	591
Lyssavirus rabies	589
Ranunculus repens	572
Centaurea jacea	555
Rosaceae	541

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

sourceTaxonName	interactionTypeName	targetTaxonName	count
Zephyranthes	hasHost	Bos taurus	247
Apis mellifera	flowersVisitedBy	Apis mellifera	173

sourceTaxonName	interactionTypeName	targetTaxonName	count
Bombus bimaculatus	flowersVisitedBy	Monarda	143
Conostegia xalapensis	parasiteOf	BOLD:AAA0099	140
Bombus impatiens	flowersVisitedBy	Dahlia	127
Hymenolepis	hasHost	Anas platyrhynchos	120
Trichogramma pretiosum	parasiteOf	Swine acute diarrhea syndrome related coronavirus	115
Hemaris thysbe	flowersVisitedBy	Monarda	108
Eukaryota	visitsFlowersOf	Eukaryota	105
Trichogramma	parasiteOf	Swine acute diarrhea syndrome related coronavirus	104
Conostegia xalapensis	parasiteOf	BOLD:AAA0222	103
Bombus robustus	flowersVisitedBy	Bombus robustus	84
Baccharoides	hasHost	Gallus gallus	67
Hemaris diffinis	flowersVisitedBy	Monarda	67
Bombus auricomus	flowersVisitedBy	Monarda	65
Cestoda	parasiteOf	Astrovirus sp.	65
Bombus griseocollis	flowersVisitedBy	Monarda	61
Bourreria costaricensis	parasiteOf	BOLD:AAA6593	61
Megachile sculpturalis	flowersVisitedBy	Latirus	60

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

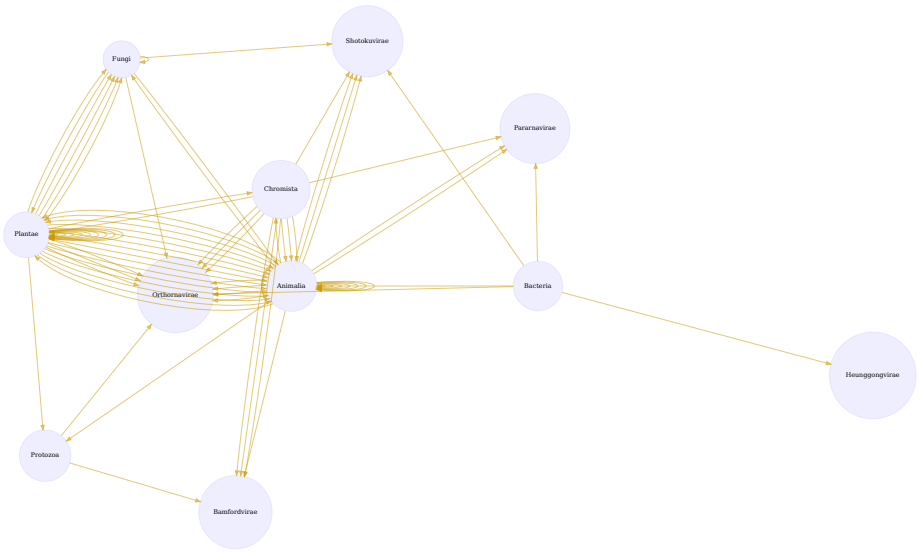


Figure 3: Interactions on taxonomic kingdom rank as interpreted by the Catalogue of Life download svg

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
Ischnopsyllus	HAS_ACCEPTED_NAME	col	Ischnopsyllus
Carduiceps zonarius	HAS_ACCEPTED_NAME	col	Carduiceps zonarius
Chelopistes bicolor	HAS_ACCEPTED_NAME	col	Chelopistes bicolor
Colinicola similis	HAS_ACCEPTED_NAME	col	Colinicola similis

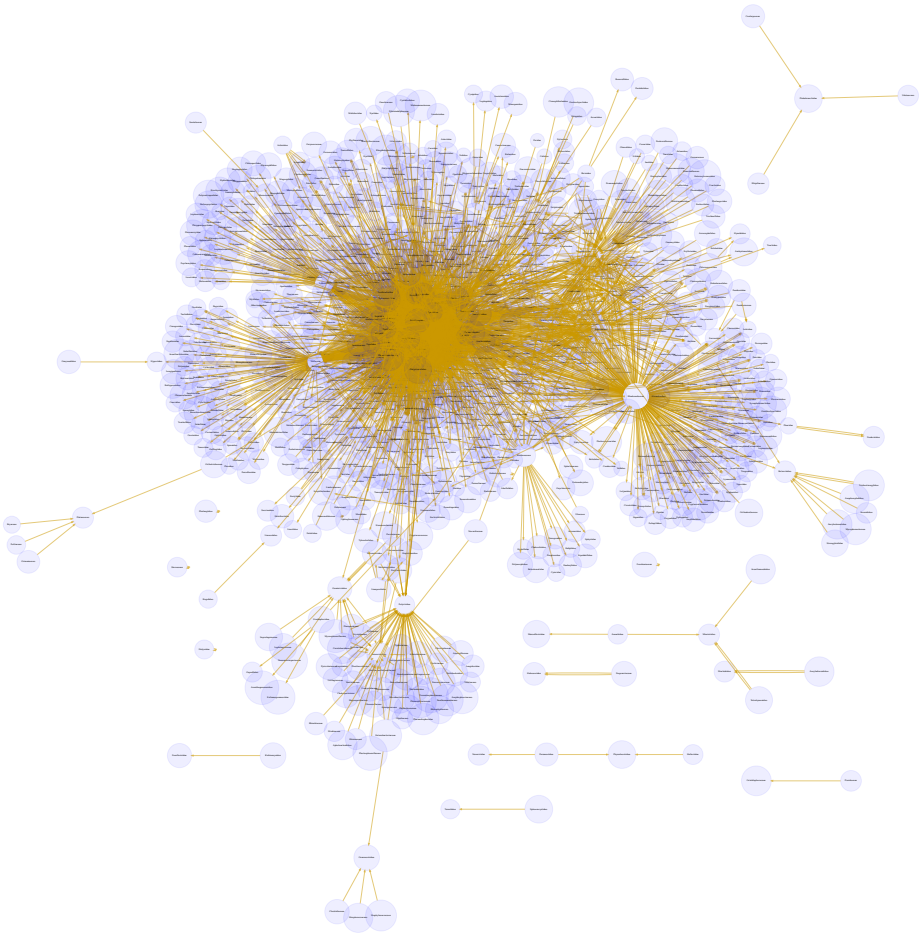


Figure 4: Interactions on the taxonomic family rank as interpreted by the Catalogue of Life. download svg

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	2139
col	class	60
col	family	410
col	genus	2225
col	gigaclass	1
col	infraclass	1
col	infraorder	1
col	infraphylum	1
col	infraspecific name	1
col	kingdom	4
col	order	138
col	other	2
col	parvorder	1
col	phylum	31
col	species	12040
col	subclass	3
col	subfamily	28
col	subgenus	50
col	suborder	3
col	subphylum	1
col	subspecies	270
col	subterclass	1
col	subtribe	2
col	superfamily	13
col	superorder	2
col	tribe	16
col	unranked	1
col	variety	22
discoverlife	NA	15168
discoverlife	species	2095
gbif	NA	930
gbif	class	64
gbif	family	445
gbif	form	10
gbif	genus	2518
gbif	kingdom	5
gbif	order	138
gbif	phylum	35
gbif	species	12861

resolvedCatalogName	resolvedRank	count
gbif	subspecies	430
gbif	variety	79
itis	NA	7311
itis	class	44
itis	division	6
itis	family	380
itis	genus	1681
itis	infrakingdom	2
itis	infraorder	2
itis	infraphylum	1
itis	kingdom	5
itis	order	130
itis	phylum	16
itis	species	7420
itis	subclass	11
itis	subfamily	35
itis	subgenus	7
itis	suborder	16
itis	subphylum	3
itis	subspecies	175
itis	superclass	1
itis	superdivision	1
itis	superfamily	13
itis	superorder	5
itis	tribe	10
itis	variety	9
mdd	NA	17263
ncbi	NA	4641
ncbi	clade	9
ncbi	class	65
ncbi	cohort	1
ncbi	family	419
ncbi	genus	2185
ncbi	infraorder	3
ncbi	kingdom	2
ncbi	order	156
ncbi	phylum	35
ncbi	section	1
ncbi	series	1
ncbi	species	9510
ncbi	species group	1
ncbi	subclass	6
ncbi	subfamily	46
ncbi	subgenus	78

resolvedCatalogName	resolvedRank	count
ncbi	suborder	9
ncbi	subphylum	1
ncbi	subspecies	115
ncbi	superclass	1
ncbi	superfamily	14
ncbi	superkingdom	4
ncbi	superorder	5
ncbi	tribe	11
ncbi	varietas	5
pbdb	NA	14864
pbdb	class	38
pbdb	family	351
pbdb	genus	988
pbdb	infraclass	3
pbdb	infraorder	4
pbdb	kingdom	3
pbdb	order	116
pbdb	phylum	21
pbdb	species	785
pbdb	subclass	11
pbdb	subfamily	41
pbdb	subgenus	1
pbdb	subkingdom	1
pbdb	suborder	16
pbdb	subphylum	1
pbdb	subspecies	4
pbdb	superclass	2
pbdb	superfamily	14
pbdb	superorder	5
pbdb	tribe	10
pbdb	unranked clade	32
tpt	NA	16149
tpt	family	37
tpt	genus	166
tpt	order	30
tpt	species	881
tpt	subspecificepithet	1
wfo	NA	13371
wfo	family	68
wfo	genus	546
wfo	order	27
wfo	section	1
wfo	species	3242
wfo	subfamily	1

resolvedCatalogName	resolvedRank	count
wfo	subgenus	1
wfo	subspecies	28
wfo	variety	13
worms	NA	13937
worms	class	48
worms	family	337
worms	genus	1031
worms	gigaclass	1
worms	infraclass	2
worms	infrakingdom	1
worms	infraorder	1
worms	infraphylum	1
worms	kingdom	6
worms	order	130
worms	phylum	19
worms	phylum (division)	6
worms	species	1707
worms	subclass	7
worms	subfamily	6
worms	suborder	8
worms	subphylum	4
worms	subspecies	16
worms	subterclass	1
worms	superclass	1
worms	superfamily	3
worms	superorder	4
worms	variety	2

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	17249
col	NONE	2257
col	SYNONYM_OF	2506
discoverlife	NONE	16966
discoverlife	HAS_ACCEPTED_NAME	2313

resolvedCatalogName	relationName	count
discoverlife	SYNONYM_OF	544
discoverlife	HOMONYM_OF	127
gbif	HAS_ACCEPTED_NAME	20376
gbif	SYNONYM_OF	4265
gbif	NONE	993
itis	NONE	7668
itis	HAS_ACCEPTED_NAME	11525
itis	SYNONYM_OF	594
mdd	NONE	19049
mdd	HAS_ACCEPTED_NAME	307
mdd	SYNONYM_OF	2
ncbi	SAME_AS	14417
ncbi	NONE	4766
ncbi	SYNONYM_OF	517
pbdb	NONE	16187
pbdb	HAS_ACCEPTED_NAME	3188
pbdb	SYNONYM_OF	259
tpt	HAS_ACCEPTED_NAME	1270
tpt	SYNONYM_OF	40
tpt	NONE	18105
wfo	NONE	15080
wfo	SYNONYM_OF	607
wfo	HAS_ACCEPTED_NAME	4189
wfo	HAS_UNCHECKED_NAME	397
worms	NONE	15270
worms	HAS_ACCEPTED_NAME	4197
worms	SYNONYM_OF	358

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)
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)

catalog name	alignment results
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
2025-09-13T00:20:25Z	summary	https://github.com/globalbioticinteractions/refuted-biotic-interactions-by-eol/archive/67dc570987c6897fca00b81247edbdb0b1861a
2025-09-13T00:20:25Z	summary	103653 interaction(s)
2025-09-13T00:20:25Z	summary	0 note(s)
2025-09-13T00:20:25Z	summary	103653 info(s)

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

: Most frequently occurring review notes, if any.

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.

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



Figure 5: 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 6: 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-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.

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

[5] At time of writing (2025-09-13) 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."

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.

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.

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.