

A Review of Biotic Interactions and Taxon Names Found in globalbioticinteractions/osal-ar hash://md5/13a7912d20fa969dca345fe0028b14f1

by Nomer, Elton and Preston, three naive review bots
review@globalbioticinteractions.org
<https://globalbioticinteractions.org/contribute>
<https://github.com/globalbioticinteractions/osal-ar/issues>

2025-04-12

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/osal-ar, has fingerprint hash://md5/13a7912d20fa969dca345fe0028b14f1, is 460MiB in size and contains 42,487 interaction with 5 unique types of associations (e.g., interactsWith) between 3,000 primary taxon (e.g., Hannemania) and 2,988 associated taxon (e.g., Bos taurus). This report includes detailed summaries of interaction data, a taxonomic review from multiple catalogs, and an archived version of the dataset from which the reviews are derived.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	11
Biotic Interactions	11
Interaction Networks	15
Taxonomic Alignment	17

Additional Reviews	21
GloBI Review Badge	22
GloBI Index Badge	22
Discussion	23
Acknowledgements	23
Author contributions	24
References	24

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:

Klompen H, Johnson N (2018). Ohio State Acarology Laboratory (OSAL), Ohio State University. Museum of Biological Diversity, The Ohio State University. Accessed via https://mbd-db.osu.edu/hol/taxon_name/05fbf4bb-f8e1-404e-a27c-759d345aa4d0 on 2023-11-06 hash://sha256/fb23140e60f4889de35ae174b2570cf294012bff4f2c8c419c292af51c
<https://github.com/globalbioticinteractions/osal-ar/archive/eaa0ff7fa87f0217b2d1313eec7f2c9ddb0ad232.zip>
2025-04-12T10:06:47.585Z hash://md5/13a7912d20fa969dca345fe0028b14f1

For additional metadata related to this dataset, please visit <https://github.com/globalbioticinteractions/osal-ar> 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	0.15.9
nomer	0.5.13
globinizer	0.4.0
mlr	6.0.0
jq	1.6
yq	4.25.3
pandoc	3.1.6.1

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

```
# get versioned copy of the dataset (size approx. 460MiB) under review
elton pull globalbioticinteractions/osal-ar

# generate review notes
elton review globalbioticinteractions/osal-ar\
> review.tsv

# export indexed interaction records
elton interactions globalbioticinteractions/osal-ar\
> interactions.tsv

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

or visually, in a process diagram.

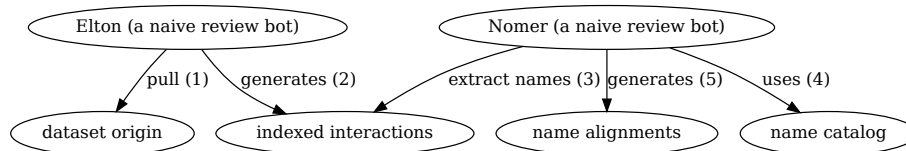


Figure 1: Review Process Overview

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/osal-ar`) before being able to generate reviews (e.g., `elton review globalbioticinteractions/osal-ar`), extract interaction claims (e.g., `elton interactions globalbioticinteractions/osal-ar`), or list taxonomic names (e.g., `elton names globalbioticinteractions/osal-ar`)

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 Preston (Elliott et al. 2025) 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)

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.

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 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-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-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

filename	description
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-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-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-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-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-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-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

filename	description
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-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-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

filename	description
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-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-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

filename	description
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

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

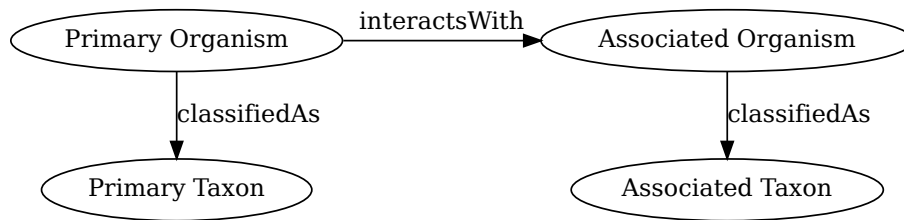


Figure 2: Biotic Interaction Data Model

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/osal-ar, has fingerprint hash://md5/13a7912d20fa969dca345fe0028b14f1, is 460MiB in size and contains 42,487 interaction with 5 unique types of associations (e.g., interactsWith) between 3,000 primary taxon (e.g., Hannemania) and 2,988 associated taxon (e.g., Bos taurus).

An exhaustive list of indexed interaction claims can be found in gzipped csv and tsv 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	interactionTypeNam	targetTaxonName	referenceCitation
Cheiroseius aciculatus	interactsWith	Bambusa	https://mbd-db.osu.edu/hol/collecting_units/0eae5d09-bf7b-3aaa-e053-0100007f2cc9
Cheiroseius aciculatus	interactsWith	Eucalyptus	https://mbd-db.osu.edu/hol/collecting_units/0eae5d09-bf7b-3aaa-e053-0100007f2cc9
Cheiroseius browningi	interactsWith	Pinus	https://mbd-db.osu.edu/hol/collecting_units/0eae5d09-bf7c-3aaa-e053-0100007f2cc9
Cheiroseius browningi	interactsWith	Pinus	https://mbd-db.osu.edu/hol/collecting_units/0eae5d09-bf80-3aaa-e053-0100007f2cc9

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

interactionTypeName	count
interactsWith	21617
parasiteOf	19229
eats	1598
preysOn	42
coRoostsWith	1

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

sourceTaxonName	count
Hannemania	1568
Euschoengastia peromysci	856
Euschoengastia brevipes	787
Uropodina	695
Trombiculidae	674
Harpirhynchidae	631
Uropodidae	618
Euschoengastia setosa	612
Harpirhynchus	598
Amblyomma testudinarium	567
Hyalomma anatolicum anatolicum	547
Boophilus microplus	505
Tarsonemus krantzi	390
Hyalomma	340
Dermacentor variabilis	337
Parasitidae	323
Rhipicephalus sanguineus	313
Insectolaelaps quadrisetus	311
Hyalomma dromedarii	292

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

targetTaxonName	count
Bos taurus	1233
Peromyscus	1009
Peromyscus leucopus	968

targetTaxonName	count
Pinus	853
Lithobates pipiens	852
Quercus	817
Pinus taeda	734
Myodes	726
Myodes gapperi	707
Sus scrofa	697
Bubalus bubalis	610
Dendroctonus frontalis	428
Equus caballus	423
Mus musculus	418
Capra hircus	323
Ovis aries	308
Bambusa	304
Canis lupus	283
Ips typographus	280

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

sourceTaxonName	interactionTypeName	targetTaxonName	count
Hannemania	parasiteOf	Lithobates pipiens	705
Euschoengastia brevipes	parasiteOf	Myodes	551
Euschoengastia setosa	parasiteOf	Peromyscus leucopus	365
Trombiculidae	parasiteOf	Myodes gapperi	350
Amblyomma testudinarium	parasiteOf	Sus scrofa	344
Boophilus microplus	parasiteOf	Bos taurus	286
Myobia murismusculi	parasiteOf	Mus musculus	251
Insectolaelaps quadrisetus	interactsWith	Ips typographus	245
Euschoengastia setosa	interactsWith	Peromyscus	241
Euschoengastia peromysci	parasiteOf	Myodes gapperi	239
Euschoengastia peromysci	interactsWith	Peromyscus	198
Tarsonemus krantzi	interactsWith	Dendroctonus frontalis	195
Tarsonemus krantzi	interactsWith	Pinus taeda	195
Galendromus occidentalis	interactsWith	Vitis vinifera	183
Hyalomma anatolicum anatolicum	parasiteOf	Bubalus bubalis	174
Euschoengastia peromysci	parasiteOf	Peromyscus leucopus	167
Hyalomma anatolicum anatolicum	parasiteOf	Bos taurus	166
Neotrombicula japonica	parasiteOf	Apodemus agrarius	161
Bdellorhynchus	interactsWith	Anas clypeata	158

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.

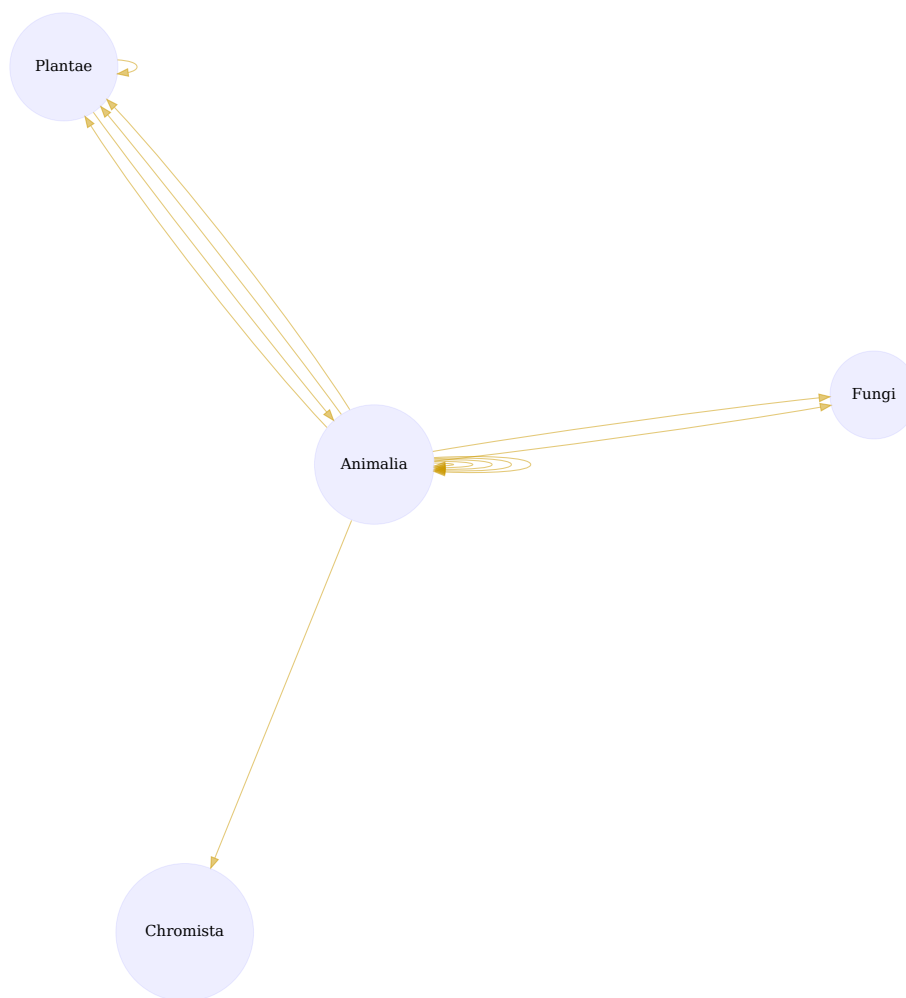


Figure 3: Interactions on taxonomic kingdom rank as interpreted by the Catalogue of Life download svg

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](#)

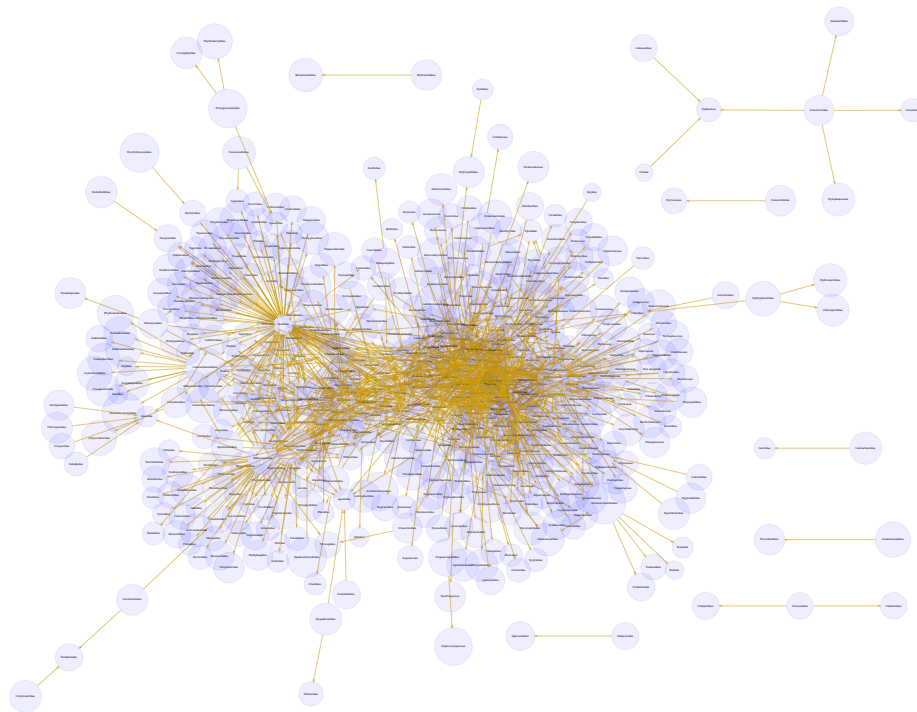


Figure 4: Interactions on the taxonomic family rank as interpreted by the Catalogue of Life. [download svg](#)

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
Dysdercus splendidus	NONE	col	Dysdercus splendidus
Melanerpes formicivorus	HAS_ACCEPTED_NAME	col	Melanerpes formicivorus
Laelaps stupkai	NONE	col	Laelaps stupkai
Psorergates musculus	HAS_ACCEPTED_NAME	col	Psorergates musculus

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	1776
col	class	9
col	family	246
col	genus	735
col	infraorder	2
col	kingdom	1
col	order	29
col	phylum	1
col	species	3109
col	subclass	1
col	subfamily	1
col	subgenus	52
col	suborder	2
col	subspecies	54

resolvedCatalogName	resolvedRank	count
col	subtribe	1
col	superfamily	2
col	superorder	1
col	tribe	1
col	variety	10
discoverlife	NA	5936
discoverlife	species	9
gbif	NA	959
gbif	class	10
gbif	family	249
gbif	genus	1034
gbif	kingdom	1
gbif	order	29
gbif	phylum	1
gbif	species	3633
gbif	subspecies	63
gbif	variety	16
itis	NA	2188
itis	class	9
itis	division	1
itis	family	247
itis	genus	664
itis	infraorder	1
itis	kingdom	1
itis	order	32
itis	species	2757
itis	subclass	2
itis	subfamily	1
itis	suborder	2
itis	subphylum	1
itis	subspecies	36
itis	superfamily	3
itis	superorder	1
itis	variety	3
mdd	NA	5945
ncbi	NA	2306
ncbi	clade	4
ncbi	class	8
ncbi	family	240
ncbi	genus	833
ncbi	infraorder	1
ncbi	kingdom	1
ncbi	order	30
ncbi	species	2498

resolvedCatalogName	resolvedRank	count
ncbi	subclass	2
ncbi	subfamily	2
ncbi	subgenus	17
ncbi	suborder	2
ncbi	subphylum	1
ncbi	subspecies	15
ncbi	superfamily	2
ncbi	superorder	1
ncbi	varietas	2
pdbb	NA	4600
pdbb	class	9
pdbb	family	156
pdbb	genus	405
pdbb	infraorder	3
pdbb	kingdom	1
pdbb	order	34
pdbb	phylum	1
pdbb	species	726
pdbb	subclass	2
pdbb	subfamily	3
pdbb	suborder	2
pdbb	subspecies	3
pdbb	superclass	1
pdbb	superfamily	1
pdbb	superorder	1
pdbb	tribe	1
pdbb	unranked clade	8
tpt	NA	3283
tpt	family	44
tpt	genus	151
tpt	order	6
tpt	species	2458
tpt	specific epithet	2
tpt	subspecific epithet	5
wfo	NA	5166
wfo	class	1
wfo	family	29
wfo	genus	278
wfo	order	5
wfo	species	457
wfo	subspecies	9
wfo	variety	7
worms	NA	4793
worms	class	7

resolvedCatalogName	resolvedRank	count
worms	family	160
worms	genus	369
worms	infraorder	1
worms	kingdom	1
worms	order	30
worms	phylum (division)	1
worms	species	575
worms	subclass	3
worms	subgenus	1
worms	suborder	2
worms	subspecies	3
worms	superfamily	2
worms	superorder	1

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	NONE	1798
col	HAS_ACCEPTED_NAME	3997
col	SYNONYM_OF	919
discoverlife	NONE	5975
discoverlife	HAS_ACCEPTED_NAME	9
discoverlife	SYNONYM_OF	2
gbif	NONE	970
gbif	HAS_ACCEPTED_NAME	5079
gbif	SYNONYM_OF	1310
itis	NONE	2215
itis	HAS_ACCEPTED_NAME	3495
itis	SYNONYM_OF	357
mdd	NONE	5406
mdd	HAS_ACCEPTED_NAME	565
mdd	SYNONYM_OF	13
ncbi	NONE	2319
ncbi	SAME_AS	3461
ncbi	SYNONYM_OF	244
pdbb	NONE	4635

resolvedCatalogName	relationName	count
pbdb	HAS_ACCEPTED_NAME	1323
pbdb	SYNONYM_OF	92
tpt	NONE	3313
tpt	HAS_ACCEPTED_NAME	2715
tpt	SYNONYM_OF	778
wfo	NONE	5205
wfo	HAS_ACCEPTED_NAME	708
wfo	SYNONYM_OF	146
wfo	HAS_UNCHECKED_NAME	75
worms	NONE	4826
worms	HAS_ACCEPTED_NAME	1122
worms	SYNONYM_OF	94

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)
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
2025-04-12T16:17:16Z	note	target taxon name missing
2025-04-12T16:17:16Z	note	invalid date string [Mar-1926]
2025-04-12T16:17:16Z	note	invalid date string [Mar-1926]
2025-04-12T16:17:16Z	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
target taxon name missing	19777
invalid date string [id-68]	152
invalid date string [DEC-1987]	98
invalid date string [OCT-1926]	74

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 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 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



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-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

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

⁵At time of writing (2025-04-12) 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, re-used 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.

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/ 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 Mitchen. 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>.