

# A Review of Biotic Interactions and Taxon Names Found in globalbioticinteractions/min-umsp hash://md5/86be604c40899938f0544f43848f891e

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

2025-04-11

## 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/min-umsp, has fingerprint hash://md5/86be604c40899938f0544f43848f891e, is 1.66MiB in size and contains 7,681 interaction with 3 unique types of associations (e.g., interactsWith) between 230 primary taxa (e.g., *Orchopeas leucopus*) and 420 associated taxa (e.g., *Peromyscus maniculatus*). This report includes detailed summaries of interaction data, a taxonomic review from multiple catalogs, and an archived version of the dataset from which the reviews are derived.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

<b>Introduction</b>	<b>2</b>
Data Review and Archive . . . . .	2
<b>Methods</b>	<b>2</b>
<b>Results</b>	<b>4</b>
Files . . . . .	4
Archived Dataset . . . . .	11
Biotic Interactions . . . . .	11
Interaction Networks . . . . .	16

Taxonomic Alignment . . . . .	18
Additional Reviews . . . . .	21
GloBI Review Badge . . . . .	22
GloBI Index Badge . . . . .	22
<b>Discussion</b>	<b>22</b>
<b>Acknowledgements</b>	<b>23</b>
<b>Author contributions</b>	<b>23</b>
<b>References</b>	<b>23</b>

## 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:

UMSP / University of Minnesota / University of Minnesota  
Insect Collection <https://github.com/globalbioticinteractions/min-umsp/archive/3f1b9d32f947dcb80b9aaab50523e097f0e8776e.zip> 2025-04-05T01:29:44.551Z hash://md5/86be604c40899938f0544f43848f891e

For additional metadata related to this dataset, please visit <https://github.com/globalbioticinteractions/min-umsp> 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, Seltsmann, 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 <sup>1</sup>.

```
# get versioned copy of the dataset (size approx. 1.66MiB) under review
elton pull globalbioticinteractions/min-umsp

# generate review notes
elton review globalbioticinteractions/min-umsp\
> review.tsv

# export indexed interaction records
elton interactions globalbioticinteractions/min-umsp\
> interactions.tsv

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

or visually, in a process diagram.

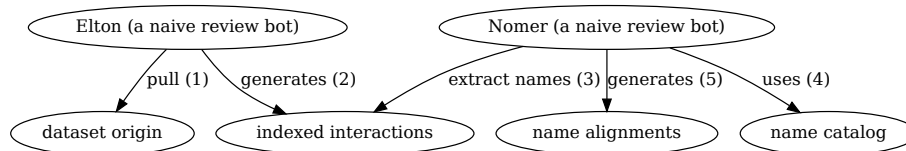


Figure 1: Review Process Overview

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

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

## Results

In the following sections, the results of the review are summarized <sup>2</sup>. 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)

---

<sup>2</sup>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 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

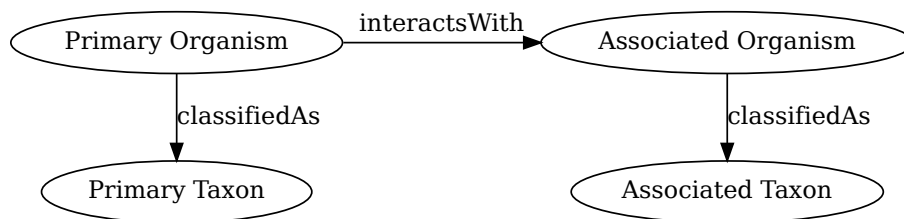


Figure 2: Biotic Interaction Data Model

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/min-umsp, has fingerprint hash://md5/86be604c40899938f0544f43848f891e, is 1.66MiB in size and contains 7,681 interaction with 3 unique types of associations (e.g., interactsWith) between 230 primary taxa (e.g., *Orchopeas leucopus*) and 420 associated taxa (e.g., *Peromyscus maniculatus*).

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](https://github.com/globalbioticinteractions/min-umsp/archive/3f1b9d32f947dcb80b9aaab50523e097f0e8776e.zip) 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
Anthrenus lepidus	interactsWith	Adenostoma	UMSP / University of Minnesota / University of Minnesota Insect Collection. Accessed at <a href="https://github.com/globalbioticinteractions/min-umsp/archive/3f1b9d32f947dcb80b9aaab50523e097f0e8776e.zip">https://github.com/globalbioticinteractions/min-umsp/archive/3f1b9d32f947dcb80b9aaab50523e097f0e8776e.zip</a> on 11 Apr 2025.

sourceTaxonName	interactionTypeName	targetTaxonName	referenceCitation
Anthrenus lepidus	interactsWith	Adenostoma	UMSP / University of Minnesota / University of Minnesota Insect Collection. Accessed at <a href="https://github.com/globalbiotici/interactions/min-umsp/archive/3f1b9d32f947dcb80b9aaab50523e097f0e8776e.zip">https://github.com/globalbiotici/interactions/min-umsp/archive/3f1b9d32f947dcb80b9aaab50523e097f0e8776e.zip</a> on 11 Apr 2025.
Dermestes caninus	interactsWith	Apocynum cannabinum var. pubescens	UMSP / University of Minnesota / University of Minnesota Insect Collection. Accessed at <a href="https://github.com/globalbiotici/interactions/min-umsp/archive/3f1b9d32f947dcb80b9aaab50523e097f0e8776e.zip">https://github.com/globalbiotici/interactions/min-umsp/archive/3f1b9d32f947dcb80b9aaab50523e097f0e8776e.zip</a> on 11 Apr 2025.

sourceTaxonName	interactionTypeName	targetTaxonName	referenceCitation
Dermestes caninus	adjacentTo	dead mouse	UMSP / University of Minnesota / University of Minnesota Insect Collection. Accessed at <a href="https://github.com/globalbiotici/interactions/min-umsp/archive/3f1b9d32f947dcb80b9aaab50523e097f0e8776e.zip">https://github.com/globalbiotici/interactions/min-umsp/archive/3f1b9d32f947dcb80b9aaab50523e097f0e8776e.zip</a> on 11 Apr 2025.

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

interactionTypeName	count
interactsWith	7534
hasHost	76
adjacentTo	71

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

sourceTaxonName	count
Orchopeas leucopus	1063
Ctenophthalmus pseudagyrtis	494
Megabothris quirini	465
Orchopeas caedens	343
Monopsyllus vison	275
Cediopsylla simplex	273
Aetheca wagneri	229
Oropsylla bruneri	223
Foxella ignota ignota	218
Orchopeas howardi	214
Epitedia wenmanni	204
Buprestidae	188

sourceTaxonName	count
Ceratophyllus styx riparius	187
Myodopsylla insignis	143
Ctenocephalides felis	126
Trogoderma versicolor	117
Euchoplopsyllus glacialis affinis	112
Xenopsylla cheopis	112
Peromyscopsylla hesperomys	104

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

targetTaxonName	count
Peromyscus maniculatus	1151
Clethrionomys gapperi	868
Tamiasciurus hudsonicus	564
Peromyscus leucopus	286
Sylvilagus floridanus	149
Blarina brevicauda	149
Geomys bursarius	146
Microtus pennsylvanicus	136
Myotis lucifugus	135
Bank Swallow	129
Peromyscus maniculatus gracilis	110
Rattus norvegicus	107
Phalacrocorax auritus	97
Quercus rubra	92
Sciurus carolinensis	80
Vulpes fulva	76
Citellus tridecemlineatus	76
Larix occidentalis	69
Pinus banksiana	67

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

sourceTaxonName	interactionType	targetTaxonName	count
Orchopeas leucopus	interactsWith	Peromyscus maniculatus	742
Megabothris quirini	interactsWith	Clethrionomys gapperi	390

sourceTaxonName	interactionTypeNam	targetTaxonName	count
Ctenophthalmus pseudagyrtes	interactsWith	Clethrionomys gapperi	246
Orchopeas caedens	interactsWith	Tamiasciurus hudsonicus	241
Monopsyllus vison	interactsWith	Tamiasciurus hudsonicus	221
Aetheca wagneri	interactsWith	Peromyscus maniculatus	152
Foxella ignota ignota	interactsWith	Geomys bursarius	146
Orchopeas leucopus	interactsWith	Peromyscus leucopus	138
Myodopsylla insignis	interactsWith	Myotis lucifugus	135
Ceratophyllus styx riparius	interactsWith	Bank Swallow	124
Cediopsylla simplex	interactsWith	Sylvilagus floridanus	107
Peromyscopsylla hesperomys	interactsWith	Peromyscus maniculatus	103
Xenopsylla cheopis	interactsWith	Rattus norvegicus	101
Ceratophyllus niger	interactsWith	Phalacrocorax auritus	97
Buprestidae	interactsWith	Quercus rubra	79
Orchopeas howardi	interactsWith	Sciurus carolinensis	74
Orchopeas leucopus	interactsWith	Microtus pennsylvanicus	71
Ctenocephalides felis	interactsWith	Cat	66
Phaenops drummondi	interactsWith	Larix occidentalis	65

## 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](#)





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

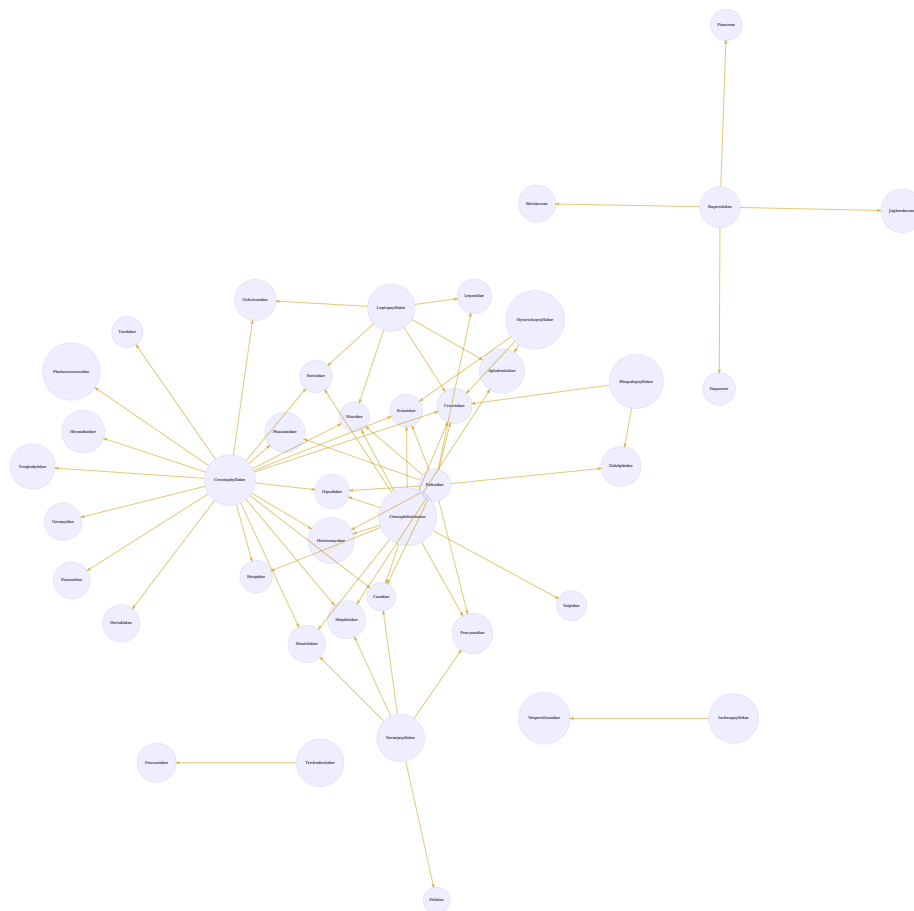


Figure 4: Interactions on the taxonomic family rank as interpreted by the Catalogue of Life. download svg

.gz. A tab-separated file can be found at [indexed-interactions.tsv.gz](#)

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
Ft circled	NONE	col	Ft circled
Ft	NONE	col	Ft
Pm	NONE	col	Pm
Acmaeodera pulchella	HAS_ACCEPTED_NAME	col	Acmaeodera pulchella

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	279
col	class	1
col	family	1
col	genus	39
col	species	267
col	subgenus	6
col	subspecies	16
discoverlife	NA	603
gbif	NA	222
gbif	class	1
gbif	family	1
gbif	genus	47
gbif	species	293
gbif	subspecies	39
gbif	variety	1

resolvedCatalogName	resolvedRank	count
itis	NA	382
itis	class	1
itis	family	1
itis	genus	29
itis	species	181
itis	subspecies	8
mdd	NA	602
ncbi	NA	339
ncbi	class	1
ncbi	family	1
ncbi	genus	36
ncbi	species	213
ncbi	subgenus	1
ncbi	subspecies	13
pbdb	NA	463
pbdb	class	1
pbdb	family	1
pbdb	genus	23
pbdb	species	113
pbdb	subspecies	1
tpt	NA	378
tpt	genus	15
tpt	species	209
wfo	NA	589
wfo	genus	3
wfo	species	10
worms	NA	541
worms	class	1
worms	family	1
worms	genus	12
worms	species	47

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	316

resolvedCatalogName	relationName	count
col	HAS_ACCEPTED_NAME	300
col	SYNONYM_OF	57
discoverlife	NONE	650
gbif	NONE	258
gbif	HAS_ACCEPTED_NAME	347
gbif	SYNONYM_OF	74
itis	NONE	419
itis	HAS_ACCEPTED_NAME	198
itis	SYNONYM_OF	34
mdd	NONE	556
mdd	HAS_ACCEPTED_NAME	83
mdd	SYNONYM_OF	9
ncbi	NONE	379
ncbi	SAME_AS	238
ncbi	SYNONYM_OF	32
ncbi	COMMON_NAME_OF	2
pbdb	NONE	500
pbdb	HAS_ACCEPTED_NAME	120
pbdb	SYNONYM_OF	30
tpt	NONE	418
tpt	HAS_ACCEPTED_NAME	271
tpt	SYNONYM_OF	44
wfo	NONE	635
wfo	HAS_ACCEPTED_NAME	11
wfo	SYNONYM_OF	3
worms	NONE	581
worms	HAS_ACCEPTED_NAME	61
worms	SYNONYM_OF	7

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
2025-04-11T14:03:37Z	note	invalid date string [06-06-1928]
2025-04-11T14:03:37Z	note	invalid date string [06-06-1928]
2025-04-11T14:03:37Z	note	invalid date string [06-06-1928]
2025-04-11T14:03:37Z	note	invalid date string [06-06-1928]

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
date [0007-01-01T00:00:00Z] occurred in the first century AD	602
invalid date string [05-05-1961]	470
date [0008-01-01T00:00:00Z] occurred in the first century AD	436
invalid date string [11-08-1963]	196

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 <sup>3</sup>

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.



Figure 6: Picture of a GloBI Index Badge <sup>4</sup>

If you'd like to keep track of reviews or index status of the dataset under review, please visit GloBI's dataset index <sup>5</sup> 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

---

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

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

<sup>5</sup>At time of writing (2025-04-11) the version of the GloBI dataset index was available at <https://globalbioticinteractions.org/datasets>

indication of the open-ness<sup>6</sup> 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 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.

---

<sup>6</sup>According 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.”

- 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 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>.