A Review of Biotic Interactions and Taxon Names Found in globalbioticinteractions/ummz-ummzi hash://md5/00d704285e620cf6ad08775401da42b7

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

2025-04-22

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/ummz-ummzi, has fingerprint hash://md5/00d704285e620cf6ad08775401da42b7, is 36.1MiB in size and contains 54,789 interaction with 2 unique types of associations (e.g., interactsWith) between 1,890 primary taxon (e.g., Proctophyllodes) and 4,219 associated taxon (e.g., Mammalia Chiroptera Vespertilionidae Myotis velifer). This report includes detailed summaries of interaction data, a taxonomic review from multiple catalogs, and an archived version of the dataset from which the reviews are derived.

Contents

Introduction

Data Review and Archive

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

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

> University of Michigan Museum of Zoology, Division of Insects https://github.com/globalbioticinteractions/ummz-ummzi/archive/0f20ed99288f6668fb894b69b98c01658c08debe.zip 2025-04-19T03:52:58.590Z hash://md5/00d704285e620cf6ad08775401da42b7

For additional metadata related to this dataset, please visit https://github.com/globalbioticinteractions/ummz-ummzi 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. 36.1MiB) under review
elton pull globalbioticinteractions/ummz-ummzi

# generate review notes
elton review globalbioticinteractions/ummz-ummzi\
 > review.tsv

# export indexed interaction records
elton interactions globalbioticinteractions/ummz-ummzi\
 > interactions.tsv

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

or visually, in a process diagram.

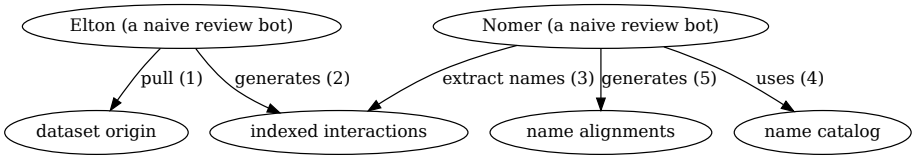

Figure 1: Review Process Overview

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

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

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)

[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-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

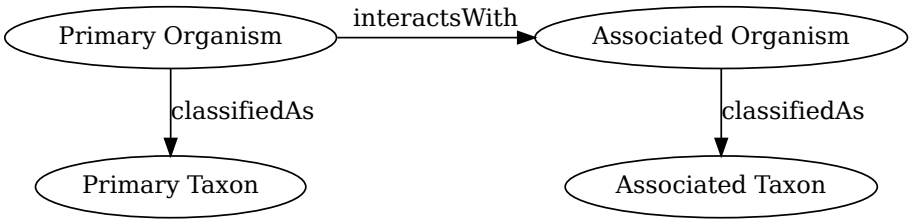

Figure 2: Biotic Interaction Data Model

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/ummz-ummzi, has fingerprint hash://md5/00d704285e620cf6ad08775401da42b7, is 36.1MiB in size and contains 54,789 interaction with 2 unique types of associations (e.g., interactsWith) between 1,890 primary taxon (e.g., Proctophyllodes) and 4,219 associated taxon (e.g., Mammalia Chiroptera Vespertilionidae Myotis velifer).

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	interactionTypeName	targetTaxonName	referenceCitation
Agathis pumila	interactsWith	Coleophoridae Coleophora laricella	53814451-56ee-4222-a6a0-ea61b4789e2c
Agathis pumila	interactsWith	Coleophoridae Coleophora laricella	afab9dee-1270-4793-8e89-16b1ef20c24e
Agathis pumila	interactsWith	Coleophoridae Coleophora laricella	caed85f4-b30c-45b9-ae01-4bb8765a878d
Agathis pumila	interactsWith	Coleophoridae Coleophora laricella	a3cbebba-acae-4b91-9487-9a7d74042b9b

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

interactionTypeName	count
interactsWith	54788
adjacentTo	1

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

sourceTaxonName	count
Proctophyllodes	3206
Chiroptoglyphus americanus	1641
Macronyssus crosbyi	1553
Proterothrix	1170
Proctophyllodes polyxenus	925
Macronyssus unidens	882
Acari	671
Pterodectes	625
Freyana anatina	597
Trouessartia trouessarti	552
Amerodectes	498
Alloptellus coniventris	489
Freyana largifolia	486
Prolistrophorus bidentatus	478
Proctophyllodes spini	471
Eustathia cultrifera	457
Freyana nyrocae	441
Trouessartia lonchurae	411
Pterolichus obtusus	392

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

targetTaxonName	count
Mammalia Chiroptera Vespertilionidae Myotis velifer	3929
Aves Passeriformes Acrocephalidae Acrocephalus arundinaceus	800
Aves Passeriformes Icteridae Quiscalus quiscula	438
Aves Suliformes Fregatidae Fregata ariel ariel	437
Aves Passeriformes Turdidae Turdus migratorius	363
Aves Galliformes Phasianidae Gallus gallus domesticus	359

targetTaxonName	count
Aves Passeriformes Fringillidae Spinus tristis	345
Aves Passeriformes Cardinalidae Pheucticus ludovicianus	309
Aves Passeriformes Parulidae Seiurus aurocapilla	297
Salix	271
Aves Apodiformes Apodidae Chaetura pelagica	267
Aves Apodiformes Apodidae Apus apus apus	267
Aves Passeriformes Fringillidae Haemorhous mexicanus	263
Aves Passeriformes Bombycillidae Bombycilla cedrorum	253
Mammalia Rodentia Cricetidae Akodon aerosus	244
Aves Strigiformes Strigidae Bubo virginianus	223
Apodiformes Apodidae Collocalia C. esculenta	220
Mammalia Didelphimorphia Didelphidae Metachirus nudicaudatus	215
Charadrius vociferus	207

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

sourceTaxonName	interactionTypeName	targetTaxonName	count
Macronyssus crosbyi	interactsWith	Mammalia Chiroptera Vespertilionidae Myotis velifer	1544
Chiroptoglyphus americanus	interactsWith	Mammalia Chiroptera Vespertilionidae Myotis velifer	1530
Macronyssus unidens	interactsWith	Mammalia Chiroptera Vespertilionidae Myotis velifer	855
Trouessartia trouessarti	interactsWith	Aves Passeriformes Acrocephalidae Acrocephalus arundinaceus	441
Trouessartia trouessartii	interactsWith	Aves Passeriformes Acrocephalidae Acrocephalus arundinaceus	348

sourceTaxonName	interactionTypeName	targetTaxonName	count
Pterolichus obtusus	interactsWith	Aves Galliformes Phasianidae Gallus gallus domesticus	346
Proctophyllodes	interactsWith	Aves Passeriformes Turdidae Turdus migratorius	336
Proctophyllodes spini	interactsWith	Aves Passeriformes Fringillidae Spinus tristis	308
Prolistrophorus bidentatus	interactsWith	Mammalia Rodentia Cricetidae Akodon aerosus	244
Alloptellus coniventris	interactsWith	Aves Suliformes Fregatidae Fregata ariel ariel	243
Proctophyllodes ampelidis	interactsWith	Aves Passeriformes Bombycillidae Bombycilla cedrorum	239
Eustathia cultrifera	interactsWith	Aves Apodiformes Apodidae Apus apus apus	228
Echineustathia tricapitosetosa	interactsWith	Aves Apodiformes Apodidae Chaetura pelagica	208
Proctophyllodes	interactsWith	Aves Passeriformes Fringillidae Haemorhous mexicanus	204
Prolistrophorus bidentatus	interactsWith	Mammalia Rodentia Cricetidae Akodon torques	170

sourceTaxonName	interactionTypeName	targetTaxonName	count
Metachiroecius brasiliensis	interactsWith	Mammalia Didelphimorphia Didelphidae Metachirus nudicaudatus	170
Proctophyllodes polyxenus	interactsWith	Aves Passeriformes Passerellidae Junco hyemalis	162
Plicatalloptes fregatae	interactsWith	Aves Suliformes Fregatidae Fregata ariel ariel	160
Kramerella	interactsWith	Aves Strigiformes Strigidae Bubo virginianus	153

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.

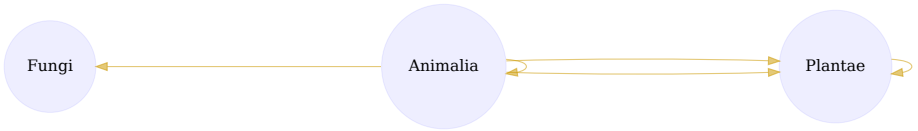

Figure 3: Interactions on taxonomic kingdom rank as interpreted by the Catalogue of Life download svg

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

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

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

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

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
Acanthonyssus	HAS_ACCEPTED_NAME	col	Acanthonyssus
Acanthonyssus dentipes	HAS_ACCEPTED_NAME	col	Acanthonyssus dentipes
Acanthonyssus proechimys	HAS_ACCEPTED_NAME	col	Acanthonyssus proechimys
Acari	NONE	col	Acari

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	1321
col	class	6
col	family	26
col	genus	171
col	kingdom	1
col	order	29
col	parvorder	1
col	phylum	2
col	species	1231
col	subgenus	11
col	subspecies	38
col	variety	5
discoverlife	NA	2510
discoverlife	species	308
gbif	NA	937
gbif	class	6
gbif	family	28
gbif	genus	280
gbif	kingdom	1
gbif	order	29
gbif	phylum	2
gbif	species	1501
gbif	subspecies	56
gbif	variety	12
itis	NA	1461
itis	class	7
itis	division	1
itis	family	24
itis	genus	153
itis	kingdom	1
itis	order	30
itis	phylum	1
itis	species	1111
itis	subclass	1
itis	subgenus	1
itis	subspecies	22
itis	variety	6
mdd	NA	2817
ncbi	NA	1475

resolvedCatalogName	resolvedRank	count
ncbi	clade	1
ncbi	class	6
ncbi	family	24
ncbi	genus	198
ncbi	kingdom	1
ncbi	order	29
ncbi	phylum	2
ncbi	species	1066
ncbi	subclass	1
ncbi	subfamily	2
ncbi	subgenus	3
ncbi	subspecies	11
ncbi	varietas	2
pbdb	NA	2454
pbdb	class	7
pbdb	family	15
pbdb	genus	60
pbdb	kingdom	1
pbdb	order	30
pbdb	phylum	1
pbdb	species	248
pbdb	subclass	1
pbdb	superorder	1
pbdb	unranked clade	6
tpt	NA	1261
tpt	family	10
tpt	genus	24
tpt	order	24
tpt	species	1498
wfo	NA	2529
wfo	family	2
wfo	genus	87
wfo	order	2
wfo	species	191
wfo	subspecies	14
wfo	variety	4
worms	NA	2433
worms	class	4
worms	family	16
worms	genus	78
worms	kingdom	1
worms	order	26
worms	phylum	1
worms	phylum (division)	1

resolvedCatalogName	resolvedRank	count
worms	species	254
worms	subclass	2
worms	subspecies	3
worms	variety	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	HAS_ACCEPTED_NAME	4710
col	NONE	1338
col	SYNONYM_OF	236
discoverlife	NONE	5807
discoverlife	HAS_ACCEPTED_NAME	301
discoverlife	SYNONYM_OF	99
discoverlife	HOMONYM_OF	23
gbif	HAS_ACCEPTED_NAME	5392
gbif	NONE	946
gbif	SYNONYM_OF	440
itis	HAS_ACCEPTED_NAME	4550
itis	SYNONYM_OF	111
itis	NONE	1483
mdd	NONE	5967
mdd	HAS_ACCEPTED_NAME	142
mdd	SYNONYM_OF	6
ncbi	NONE	1504
ncbi	SAME_AS	4562
ncbi	SYNONYM_OF	63
pbdb	NONE	2490
pbdb	SYNONYM_OF	10109
pbdb	HAS_ACCEPTED_NAME	3616
tpt	NONE	4161
tpt	HAS_ACCEPTED_NAME	1852
tpt	SYNONYM_OF	406
wfo	NONE	5806
wfo	HAS_ACCEPTED_NAME	275
wfo	HAS_UNCHECKED_NAME	84

resolvedCatalogName	relationName	count
wfo	SYNONYM_OF	83
worms	NONE	2485
worms	HAS_ACCEPTED_NAME	3627
worms	SYNONYM_OF	59

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
2025-04-22T17:23:51Z	note	source taxon name missing: using institution-Code/collectionCode/collectionId/catalogNumber/occur as placeholder
2025-04-22T17:23:51Z	note	source taxon name missing: using institution-Code/collectionCode/collectionId/catalogNumber/occur as placeholder
2025-04-22T17:23:57Z	note	issue handling date range [1962-01-01/1961-01-01]: The end instant must be greater than the start instant
2025-04-22T17:23:57Z	note	issue handling date range [1962-01-01/1961-01-01]: The end instant must be greater than the start instant

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
issue handling date range [1971-10/1971-03]: The end instant must be greater than the start instant	140
issue handling date range [1972-10/1972-03]: The end instant must be greater than the start instant	12
issue handling date range [1920-12-01/1920-08-16]: The end instant must be greater than the start instant	6

reviewComment	count
issue handling date range [1944-06-09/1944-06-03]: The end instant must be greater than the start instant	4

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 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

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

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

[5] At time of writing (2025-04-22) the version of the GloBI dataset index was available at https://globalbioticinteractions.org/datasets

natural history collection data managers, as a backup archive of a shared Darwin Core archive. It also serves as a means of creating a trackable citation for the dataset in an automated way, while also including some information about the contents of the dataset.

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

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

Acknowledgements

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

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.

[6]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.”

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.