

A Review of Biotic Interactions and Taxon Names Found in `globalbioticinteractions/mcz` hash://md5/f0dd4a24369d5fd0104b440fb8d7465a

by Nomer, Elton and Preston, three naive review bots
review@globalbioticinteractions.org
<https://globalbioticinteractions.org/contribute>
<https://github.com/globalbioticinteractions/mcz/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/mcz`, has fingerprint hash://md5/f0dd4a24369d5fd0104b440fb8d7465a, is 3.26GiB in size and contains 9,957 interaction with 9 unique types of associations (e.g., `interactsWith`) between 1,990 primary taxon (e.g., *Gasterosteus aculeatus* Linnaeus, 1758) and 2,863 associated taxon (e.g., ground). This report includes detailed summaries of interaction data, a taxonomic review from multiple catalogs, and an archived version of the dataset from which the reviews are derived.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	22
Acknowledgements	23
Author contributions	23
References	23

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:

Museum of Comparative Zoology, Harvard University - Version
162.465 <http://digir.mcz.harvard.edu/ipt/archive.do?r=mczbase>
2025-04-05T01:17:56.929Z hash://md5/f0dd4a24369d5fd0104b440fb8d7465a

For additional metadata related to this dataset, please visit <https://github.com/globalbioticinteractions/mcz> 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

tool name	version
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. 3.26GiB) under review
elton pull globalbioticinteractions/mcz

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

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

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

or visually, in a process diagram.

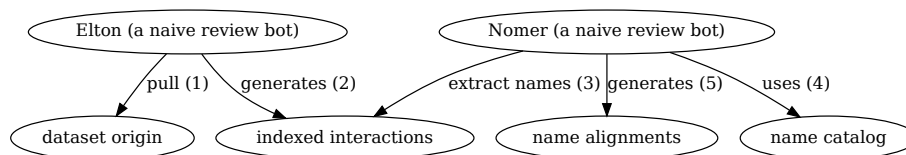


Figure 1: Review Process Overview

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

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

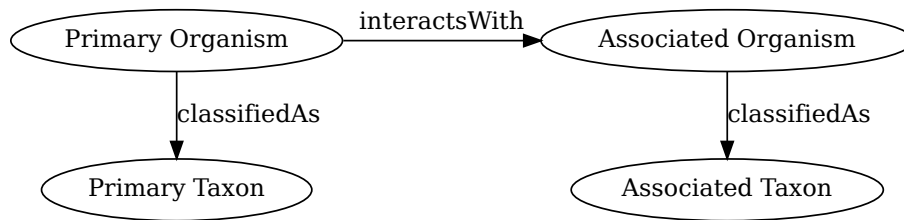


Figure 2: Biotic Interaction Data Model

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/mcz, has fingerprint hash://md5/f0dd4a24369d5fd0104b440fb8d7465a, is 3.26GiB in size and contains 9,957 interaction with 9 unique types of associations (e.g., interactsWith) between 1,990 primary taxon (e.g., *Gasterosteus aculeatus* Linnaeus, 1758) and 2,863 associated taxon (e.g., ground).

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
Lithobates catesbeianus (Shaw, 1802)	eats	MCZ:Herp:R- 200746	http://mczbase.mcz.harvard.edu/guid/MCZ:Herp:R-200746
Lithobates catesbeianus (Shaw, 1802)	eats	MCZ:Herp:R- 200745	http://mczbase.mcz.harvard.edu/guid/MCZ:Herp:R-200745
Pyxicephalus adspersus	interactsWith	duplicates of MCZ-A-20225 - 20228 are given new catalog numbers MCZ-A-154068 - 154070 02OCT2023 Tsuyoshi Takahashi	http://mczbase.mcz.harvard.edu/guid/MCZ:Herp:R-200746

sourceTaxonName	interactionTypeName	targetTaxonName	referenceCitation
Lissotriton vulgaris vulgaris (Linnaeus, 1758)	interactsWith	Ichthyosaura alpestris alpestris (Laurenti, 1768)	http://mczbase.mcz.harvard.edu/guid/MCZ:Herp 4724

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

interactionTypeName	count
interactsWith	5743
adjacentTo	2701
hasHost	764
parasiteOf	690
eats	31
hasParasite	10
visits	9
eatenBy	8
killedBy	1

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

sourceTaxonName	count
Gasterosteus aculeatus Linnaeus, 1758	1032
Perdita sp. Smith, 1853	536
Xylocopa darwini Cockerell, 1926	482
Camallanus Railliet & Henry, 1915	338
Subdoluseps bowringii (Günther, 1864) (Günther, 1864)	311
Symbion americanus Obst, Funch & Kristensen, 2006	208
Camponotus sp. Mayr, 1861	197
Pheidole sp. Westwood, 1839	184
Cephalochlamys Blanchard, 1908	169
Eutropis multifasciata multifasciata (Kuhl, 1820)	114
Andrena hirticincta Provancher, 1888	102
Hepialidae Stephens, 1829	94
Bdelloura candida (Girard, 1850)	65
Andrena sp. Fabricius 1775	63
Pontoniinae	60
Formicidae Latreille, 1809	57
Pheidole Westwood, 1839	54
Hylaeus modestus Say, 1837	54

sourceTaxonName	count
Globicephala melaena (Traill, 1809)	49

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

targetTaxonName	count
ground	437
Mites	228
Ceanothus	176
Stylopized	172
Heterotheca	141
ex rotten log	131
Solidago	119
Ophiocordyceps	94
ex dead twig	89
under	88
low vegetation	86
Mentzelia	83
Haplopappus	80
tree	60
gilled mushrooms	56
vegetation	49
Chamaesaracha	46
trout	44
Trout	44

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

sourceTaxonName	interactionType	targetTaxonName	count
Xylocopa darwini Cockerell, 1926	interactsWith	Mites	167
Perdita sp. Smith, 1853	interactsWith	Heterotheca	136
Hepialidae Stephens, 1829	interactsWith	Ophiocordyceps	94
Andrena hirticincta Provancher, 1888	interactsWith	Stylopized	90

sourceTaxonName	interactionTypeNam	targetTaxonName	count
Perdita sp. Smith, 1853	interactsWith	Mentzelia	83
Perdita sp. Smith, 1853	interactsWith	Haplopappus	77
Perdita sp. Smith, 1853	interactsWith	Chamaesaracha	46
Camponotus sp. Mayr, 1861	adjacentTo	ground	45
Gasterosteus aculeatus Linnaeus, 1758	interactsWith	trout	44
Gasterosteus aculeatus Linnaeus, 1758	interactsWith	Trout	44
Globicephala melaena (Traill, 1809)	interactsWith	Globicephala melaena (Traill, 1809)	40
Lapethus newtoni Slipinski	adjacentTo	gilled mushrooms	40
Pseudopanurgus pauper (Cresson, 1878)	interactsWith	Ceanothus	40
Gasterosteus aculeatus Linnaeus, 1758	interactsWith	sculpin	39
Trigona ceophloe Schwarz	hasHost	Ex stomach of Ceoploeus lineatus	38
Perdita sp. Smith, 1853	interactsWith	Solidago	38
Xylocopa darwini Cockerell, 1926	interactsWith	mites	38
Crematogaster laeviuscula Mayr, 1870	adjacentTo	cotton wood tree	37
Gasterosteus aculeatus Linnaeus, 1758	interactsWith	Salmonids	35

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.

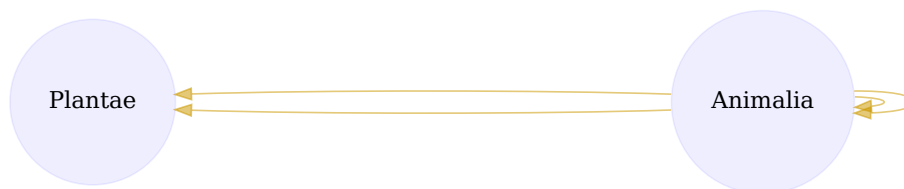


Figure 3: Interactions on taxonomic kingdom rank as interpreted by the Catalogue of Life download svg



Figure 4: Interactions on the taxonomic family 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.

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
Triungulin	NONE	col	Triungulin
Pungitius	HAS_ACCEPTED_NAME	col	Pungitius
Salmonids	NONE	col	Salmonids
High plants	NONE	col	High plants

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	1581
col	class	5
col	family	65
col	genus	377
col	infraspecific name	1
col	kingdom	1
col	order	6
col	phylum	2
col	section	1
col	species	1494
col	subfamily	5
col	subgenus	6
col	suborder	1
col	subspecies	53
col	tribe	1
col	variety	7
discoverlife	NA	3239
discoverlife	species	351
gbif	NA	1315
gbif	class	5
gbif	family	69
gbif	genus	418
gbif	kingdom	2

resolvedCatalogName	resolvedRank	count
gbif	order	6
gbif	phylum	2
gbif	species	1717
gbif	subspecies	65
gbif	variety	12
itis	NA	1792
itis	class	5
itis	family	64
itis	genus	352
itis	kingdom	1
itis	order	6
itis	phylum	2
itis	species	1301
itis	subclass	1
itis	subfamily	4
itis	suborder	1
itis	subspecies	50
itis	tribe	1
itis	variety	12
mdd	NA	3589
ncbi	NA	1920
ncbi	class	6
ncbi	family	62
ncbi	genus	369
ncbi	infraorder	1
ncbi	order	6
ncbi	phylum	2
ncbi	species	1200
ncbi	subfamily	4
ncbi	subgenus	10
ncbi	subspecies	17
ncbi	tribe	1
ncbi	varietas	2
pdb	NA	3191
pdb	class	7
pdb	family	55
pdb	genus	214
pdb	informal	1
pdb	kingdom	1
pdb	order	6
pdb	phylum	2
pdb	species	110
pdb	subclass	1
pdb	subfamily	2

resolvedCatalogName	resolvedRank	count
pbdb	suborder	1
pbdb	tribe	1
pbdb	unranked clade	1
tpt	NA	3532
tpt	genus	3
tpt	species	54
wfo	NA	3250
wfo	family	4
wfo	genus	130
wfo	phylum	1
wfo	species	194
wfo	subspecies	5
wfo	variety	8
worms	NA	2872
worms	class	5
worms	family	54
worms	genus	239
worms	kingdom	1
worms	order	6
worms	phylum	2
worms	species	403
worms	subclass	1
worms	subfamily	2
worms	subgenus	1
worms	suborder	1
worms	subspecies	4
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	NONE	2665
col	HAS_ACCEPTED_NAME	1965
col	SYNONYM_OF	406
discoverlife	NONE	4445
discoverlife	HAS_ACCEPTED_NAME	331

resolvedCatalogName	relationName	count
discoverlife	SYNONYM_OF	114
discoverlife	HOMONYM_OF	24
gbif	NONE	2390
gbif	HAS_ACCEPTED_NAME	2456
gbif	SYNONYM_OF	609
itis	NONE	2880
itis	HAS_ACCEPTED_NAME	1792
itis	SYNONYM_OF	152
mdd	NONE	4758
mdd	HAS_ACCEPTED_NAME	15
ncbi	NONE	3029
ncbi	SAME_AS	1717
ncbi	SYNONYM_OF	76
ncbi	COMMON_NAME_OF	3
pbdb	NONE	4309
pbdb	HAS_ACCEPTED_NAME	462
pbdb	SYNONYM_OF	40
tpt	NONE	4710
tpt	HAS_ACCEPTED_NAME	65
tpt	SYNONYM_OF	1
wfo	NONE	4400
wfo	SYNONYM_OF	100
wfo	HAS_ACCEPTED_NAME	312
wfo	HAS_UNCHECKED_NAME	55
worms	NONE	3995
worms	HAS_ACCEPTED_NAME	745
worms	SYNONYM_OF	130

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
2025-04-11T12:48:38Z	note	found unresolved reference [MCZ:Cryo:14052]
2025-04-11T12:48:38Z	note	found unresolved reference [MCZ:Cryo:14068]
2025-04-11T12:48:38Z	note	found unresolved reference [MCZ:Cryo:14075]
2025-04-11T12:48:38Z	note	found unresolved reference [MCZ:Cryo:14076]

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
found unresolved reference [MCZ:Cryo:14052]	1
found unresolved reference [MCZ:Cryo:14068]	1

reviewComment	count
found unresolved reference [MCZ:Cryo:14075]	1
found unresolved reference [MCZ:Cryo:14076]	1

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.



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

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

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

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

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 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>.

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.”

- 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 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>.