

A Review of Biotic Interactions and Taxon Names Found in `globalbioticinteractions/tamuic-ent` hash://md5/584fe9abb15456ace49af72da712111a

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

2025-04-13

Abstract

Life on Earth is sustained by complex interactions between organisms and their environment. These biotic interactions can be captured in datasets and published digitally. We present a review and archiving process for such an openly accessible digital interactions dataset of known origin and discuss its outcome. The dataset under review, named `globalbioticinteractions/tamuic-ent`, has fingerprint hash://md5/584fe9abb15456ace49af72da712111a, is 574MiB in size and contains 12,520 interaction with 4 unique types of associations (e.g., `interactsWith`) between 1,215 primary taxon (e.g., `Anthonomus`) and 1,857 associated taxon (e.g., `post oak savanna w/pasture`). This report includes detailed summaries of interaction data, a taxonomic review from multiple catalogs, and an archived version of the dataset from which the reviews are derived.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	15
Additional Reviews	20
GloBI Review Badge	21
GloBI Index Badge	21
Discussion	22
Acknowledgements	22
Author contributions	22
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:

Texas A&M University Insect Collection <https://github.com/globalbioticinteractions/tamuic-ent/archive/f261a8c192021408da67c39626a4aac56e3bac41.zip> 2025-04-12T10:51:07.262Z hash://md5/584fe9abb15456ace49af72da712111a

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

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. 574MiB) under review
elton pull globalbioticinteractions/tamuic-ent

# generate review notes
elton review globalbioticinteractions/tamuic-ent\
> review.tsv

# export indexed interaction records
elton interactions globalbioticinteractions/tamuic-ent\
> interactions.tsv

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

or visually, in a process diagram.

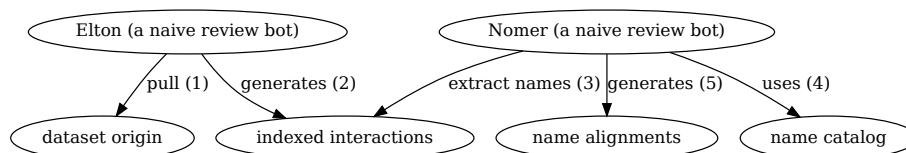


Figure 1: Review Process Overview

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

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

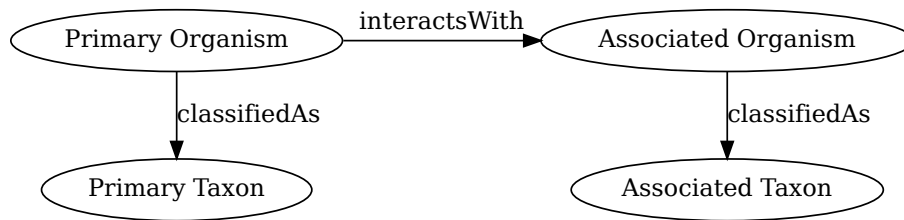


Figure 2: Biotic Interaction Data Model

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/tamuic-ent, has fingerprint hash://md5/584fe9abb15456ace49af72da712111a, is 574MiB in size and contains 12,520 interaction with 4 unique types of associations (e.g., interactsWith) between 1,215 primary taxon (e.g., Anthonomus) and 1,857 associated taxon (e.g., post oak savanna w/pasture).

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
Gnaphosa sericata	interactsWith	sandy open prairie	(L. Koch),
Drassyllus creolus	interactsWith	live oak woodland	Chamberlin & Gertsch,
Anthonomus suturalis	interactsWith	Phylloxera gall on pecan	LeConte,
Anthonomus suturalis	interactsWith	Phylloxera gall on pecan	LeConte,

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

interactionTypeName	count
interactsWith	12104
hasHost	209
adjacentTo	169

interactionType	Name	count
eats		38

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

sourceTaxonName	count
Anthonomus	560
Anthonomus hunteri	260
Tricorynus	240
Dieunomia heteropoda	230
Amblyomma cajennense	222
Anthonomus tectus	216
Anthonomus squamans	207
Anthonomopsis mixta	180
Anthonomus testaceosquamosus	179
Phantyna segregata	172
Falconina gracilis	169
Anthonomus rileyi	143
Talanites exlineae	138
Dermacentor variabilis	137
Amblyomma inornatum	135
Anthonomus rufipennis	127
Elaver excepta	118
Anthonomus sphaeralciae	116
Drassyllus	110

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

targetTaxonName	count
post oak savanna w/pasture	612
pecan	359
dense coastal brush	317
Byrsonima crassifolia (L.) H. B. K.	305
palm forest	280
cotton	224
Cottontail	169
Javelina	165
Cotton	162
re-vegetated site	155

targetTaxonName	count
Hampea trilobata	153
ebony-guayacan association	149
managed	136
Raccoon	131
loblolly pine	130
Horse	125
peanuts	121
spices	120
Tandoori blend	120

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

sourceTaxonName	interactionType	targetTaxonName	count
Anthonomus	interactsWith	Byrsonima crassifolia (L.) H. B. K.	305
Anthonomus hunteri	interactsWith	Hampea trilobata	150
Tricorynus	interactsWith	spices	120
Tricorynus	interactsWith	Tandoori blend	120
Anthonomus sphaeralciae	interactsWith	Sphaeralcea emoryi	107
Anthonomus hunteri	interactsWith	Hamp. trilobata	105
Anthonomus tectus	interactsWith	Heterotheca	96
Haemaphysalis leporispalustris	interactsWith	Cottontail	94
Anthonomopsis mixta	interactsWith	Prunus sp.	88
Anthonomopsis mixta	interactsWith	Wild prune	86
Amblyomma cajennense	interactsWith	Javelina	85
Falconina gracilis	interactsWith	post oak savanna w/pasture	79
Anthonomus rileyi	interactsWith	Juniperus ashei	77
Dieunomia heteropoda	interactsWith	Verbesina encelioides	72

sourceTaxonName	interactionTypeNam	targetTaxonName	count
Tabanus stenocephalus	interactsWith	Horse	68
Mazax pax	interactsWith	palm forest	67
Anthonomus squamans	interactsWith	Malvastrum aurantiacum	67
Anthonomus squamans	interactsWith	Sphaeralcea lindheimeri	65
Dieunomia heteropoda	interactsWith	sunflower	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.



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.

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.

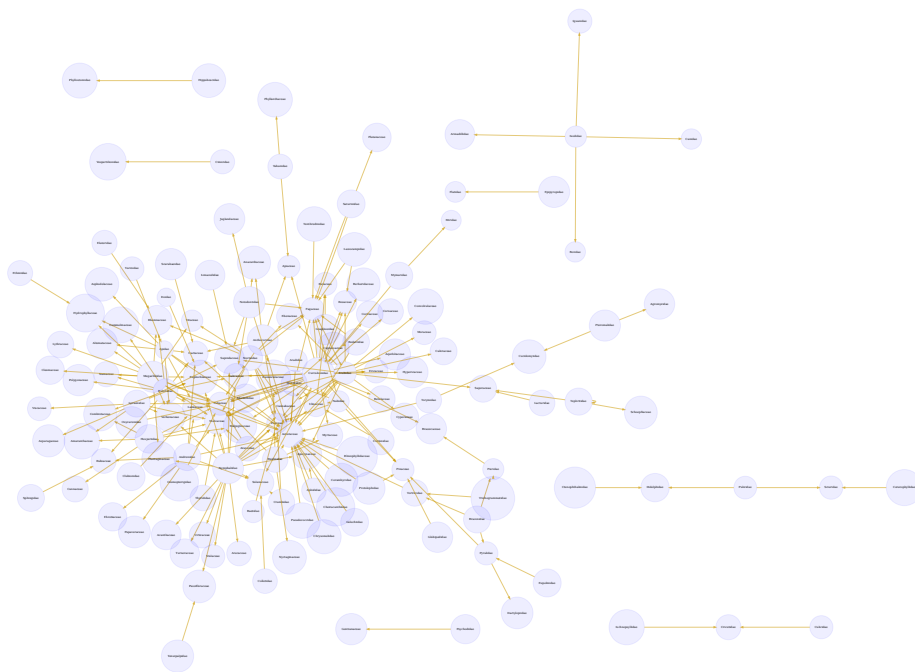


Figure 4: Interactions on the taxonomic family rank as interpreted by the Catalogue of Life. [download svg](#)

Table 8: Sample of Name Alignments

providedName	relationName	resolvedCatalogName	resolvedName
Year old girl	NONE	col	Year old girl
Ft high tree	NONE	col	Ft high tree
Abies fraseri	HAS_ACCEPTED_NAME	col	Abies fraseri
Abies	HAS_ACCEPTED_NAME	col	Abies

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	1261
col	family	8
col	genus	284
col	kingdom	2
col	order	1
col	species	1243
col	subfamily	1
col	subgenus	10
col	subspecies	45
col	superfamily	1
col	variety	17
discoverlife	NA	2744
discoverlife	species	95
gbif	NA	1119
gbif	family	8
gbif	form	1
gbif	genus	306
gbif	kingdom	2
gbif	order	1
gbif	species	1331
gbif	subspecies	67
gbif	variety	28
itis	NA	1461
itis	family	8
itis	genus	247
itis	kingdom	2
itis	order	2
itis	phylum	1
itis	species	1049

resolvedCatalogName	resolvedRank	count
itis	subfamily	1
itis	subspecies	54
itis	variety	15
mdd	NA	2838
ncbi	NA	1557
ncbi	family	7
ncbi	genus	277
ncbi	kingdom	1
ncbi	order	1
ncbi	species	957
ncbi	subfamily	1
ncbi	subgenus	8
ncbi	subspecies	37
ncbi	varietas	2
pbdb	NA	2671
pbdb	family	8
pbdb	genus	123
pbdb	infraorder	1
pbdb	kingdom	2
pbdb	order	2
pbdb	species	30
pbdb	subfamily	1
pbdb	suborder	2
pbdb	unranked clade	2
tpt	NA	2787
tpt	genus	2
tpt	species	49
wfo	NA	2313
wfo	family	6
wfo	genus	107
wfo	phylum	1
wfo	species	399
wfo	subspecies	9
wfo	variety	13
worms	NA	2484
worms	family	7
worms	genus	126
worms	kingdom	2
worms	order	1
worms	species	212
worms	suborder	1
worms	subspecies	3
worms	variety	2

Table 10: Name relationship types per catalog. Name relationship type “NONE” means that a name was not recognized by the associated catalog. “SAME_AS” indicates either a “HAS_ACCEPTED_NAME” or “SYNONYM_OF” name relationship type. We recognize that “SYNONYM_OF” encompasses many types of nomenclatural synonymies (ICZN 1999) (e.g., junior synonym, senior synonyms).

resolvedCatalogName	relationName	count
col	NONE	1404
col	HAS_ACCEPTED_NAME	1535
col	SYNONYM_OF	478
discoverlife	NONE	2975
discoverlife	HAS_ACCEPTED_NAME	95
discoverlife	SYNONYM_OF	13
discoverlife	HOMONYM_OF	1
gbif	NONE	1256
gbif	HAS_ACCEPTED_NAME	1804
gbif	SYNONYM_OF	628
itis	NONE	1604
itis	HAS_ACCEPTED_NAME	1341
itis	SYNONYM_OF	156
mdd	NONE	3056
mdd	HAS_ACCEPTED_NAME	6
ncbi	NONE	1718
ncbi	SAME_AS	1287
ncbi	SYNONYM_OF	81
ncbi	COMMON_NAME_OF	7
pbdb	NONE	2861
pbdb	HAS_ACCEPTED_NAME	201
pbdb	SYNONYM_OF	19
tpt	NONE	3011
tpt	HAS_ACCEPTED_NAME	62
tpt	SYNONYM_OF	7
wfo	NONE	2465
wfo	HAS_ACCEPTED_NAME	494
wfo	HAS_UNCHECKED_NAME	35
wfo	SYNONYM_OF	191
worms	NONE	2661
worms	HAS_ACCEPTED_NAME	392
worms	SYNONYM_OF	53

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
2025-04-13T19:47:25Z	note	target taxon name missing
2025-04-13T19:47:25Z	note	target taxon name missing
2025-04-13T19:47:25Z	note	target taxon name missing
2025-04-13T19:47:25Z	note	target taxon name missing

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

Table 13: Most frequently occurring review notes, if any.

reviewComment	count
target taxon name missing	1093026
found unsupported interaction type with name: [larval food plant]	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.

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 (2025-04-13) the version of the GloBI dataset index was available at <https://globalbioticinteractions.org/datasets>

Discussion

This review and archive provides a means of creating citable versions of datasets that change frequently. This may be useful for dataset managers, including natural history collection data managers, as a backup archive of a shared Darwin Core archive. It also serves as a means of creating a trackable citation for the dataset in an automated way, while also including some information about the contents of the dataset.

This review aims to provide a perspective on the dataset to aid in understanding of species interaction claims discovered. However, it is important to note that this review does *not* assess the quality of the dataset. Instead, it serves as an indication of the open-ness⁶ and FAIRness (Wilkinson et al. 2016; Trekels et al. 2023) of the dataset: to perform this review, the data was likely openly available, **F**indable, **A**ccessible, **I**nteroperable and **R**eusable. The current Open-FAIR assessment is qualitative, and a more quantitative approach can be implemented with specified measurement units.

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

Acknowledgements

We thank the many humans that created us and those who created and maintained the data, software and other intellectual resources that were used for producing this review. In addition, we are grateful for the natural resources providing the basis for these human and bot activities. Also, thanks 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.

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.”

References

- Elliott, Michael, Jorrit Poelen, Icaro Alzuru, Emilio Berti, and partha04patel. 2025. “Bio-Guoda/Preston: 0.10.5.” Zenodo. <https://doi.org/10.5281/zenodo.14662206>.
- ICZN. 1999. “International Code of Zoological Nomenclature.” The International Trust for Zoological Nomenclature, London, UK. <https://www.iczn.org/the-code/the-code-online/>.
- Kuhn, Tobias, and Michel Dumontier. 2014. “Trusty URIs: Verifiable, Immutable, and Permanent Digital Artifacts for Linked Data.” In *The Semantic Web: Trends and Challenges*, edited by Valentina Presutti, Claudia d’Amato, Fabien Gandon, Mathieu d’Aquin, Steffen Staab, and Anna Tordai, 395–410. Cham: Springer International Publishing.
- Kuhn, Tobias, Jorrit Poelen, and Katrin Leinweber. 2025. “Globalbioticinteractions/Elton: 0.15.1.” Zenodo. <https://doi.org/10.5281/zenodo.14927734>.
- Poelen, Jorrit H. (ed.). 2024. “Nomer Corpus of Taxonomic Resources Hash://Sha256/ B60c0d25a16ae77b24305782017b1a270b79b5d1746f832650 F2027ba536e276 Hash://Md5/17f1363a277ee0e4ecaf1b91c665e47e.” Zenodo. <https://doi.org/10.5281/zenodo.12695629>.
- Poelen, Jorrit H., James D. Simons, and Chris J. Mungall. 2014. “Global Biotic Interactions: An Open Infrastructure to Share and Analyze Species-Interaction Datasets.” *Ecological Informatics* 24 (November): 148–59. <https://doi.org/10.1016/j.ecoinf.2014.08.005>.
- Poelen, Jorrit, Katja Seltmann, and Daniel 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>.