A Review of Biotic Interactions and Taxon Names Found in globalbioticinteractions/psuc-ento hash://md5/bff87cf761af6ad12b60388b622c1288

by Nomer, Elton and Preston, three naive review bots
review@globalbioticinteractions.org
https://globalbioticinteractions.org/contribute
https://github.com/globalbioticinteractions/psuc-ento/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/psuc-ento, has fingerprint hash://md5/bff87cf761af6ad12b60388b622c1288, is 15.7MiB in size and contains 71,991 interaction with 7 unique types of associations (e.g., interactsWith) between 1,938 primary taxon (e.g., Augochlora pura) and 1,311 associated taxon (e.g., Prunus). This report includes detailed summaries of interaction data, a taxonomic review from multiple catalogs, and an archived version of the dataset from which the reviews are derived.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:

> Frost Entomological Museum, Pennsylvania State University https://github.com/globalbioticinteractions/psuc-ento/archive/30b1f96619a6e9f10da18b42fb93ff22cc4f72e2 2025-04-12T10:15:47.461Z hash://md5/bff87cf761af6ad12b60388b622c1288

For additional metadata related to this dataset, please visit https://github.com/globalbioticinteractions/psuc-ento 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

tool name	version
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. 15.7MiB) under review
elton pull globalbioticinteractions/psuc-ento

# generate review notes
elton review globalbioticinteractions/psuc-ento\
 > review.tsv

# export indexed interaction records
elton interactions globalbioticinteractions/psuc-ento\
 > interactions.tsv

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

or visually, in a process diagram.

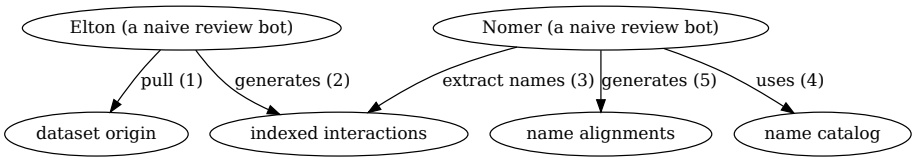

Figure 1: Review Process Overview

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

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

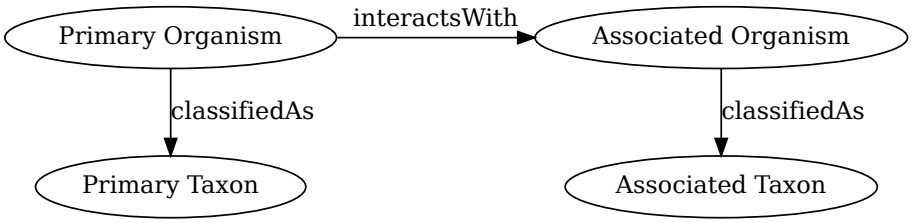

Figure 2: Biotic Interaction Data Model

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/psuc-ento, has fingerprint hash://md5/bff87cf761af6ad12b60388b622c1288, is 15.7MiB in size and contains 71,991 interaction with 7 unique types of associations (e.g., interactsWith) between 1,938 primary taxon (e.g., Augochlora pura) and 1,311 associated taxon (e.g., Prunus).

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
PSUC | ENTO | http://grbio.org/cool/29fv-ztxs | PSUC_FEM_7144 | db73d8c6-4525-4e5d-b50b-2d07eaa383b4	interactsWith	Prarie dog	https://scan-bugs.org:443/portal/collections/individual/index.
Hybophthirus notophallus	interactsWith	Cape Ant Bear or Aardvark	https://scan-bugs.org:443/portal/collections/individual/index.
PSUC | ENTO | http://grbio.org/cool/29fv-ztxs | PSUC_FEM_7082 | a61a4049-6127-40b6-9789-9220c144ac3d	interactsWith	Ovis dalli	https://scan-bugs.org:443/portal/collections/individual/index.

sourceTaxonName	interactionTypeName	targetTaxonName	referenceCitation
PSUC | ENTO | http://grbio.org/cool/29fv-ztxs | PSUC_FEM_7083 | 9d7d60df-b703-4e2f-9d2d-04f36990229b	interactsWith	Ovis dalli	https://scan-bugs.org:443/portal/collections/individual/index.

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

interactionTypeName	count
interactsWith	70737
hasHost	645
adjacentTo	576
eats	25
hostOf	5
visits	2
parasiteOf	1

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

sourceTaxonName	count
Augochlora pura	12396
Ceratina calcarata	7022
Eucera hamata	6564
Bombus impatiens	3922
Apis mellifera	3382
Lasioglossum pilosum	3071
Melissodes bimaculatus	2829
Bombus vagans	2569
Agapostemon virescens	2566
Peponapis pruinosa	1402
Ceratina mikmaqi	1029
Halictus ligatus	1010
Bombus bimaculatus	891
Augochlorella aurata	820
Chauliognathus pensylvanicus	672
Lasioglossum tegulare	653

sourceTaxonName	count
Bombus perplexus	596
Andrena miserabilis	553
Bombus griseocollis	527

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

targetTaxonName	count
Prunus	31289
Malus	23658
Agastache Blue Fortune	2233
Rudbeckia Herbstonne	993
Agastache Black Adder	729
Rudbeckia Triloba	603
Homo sapiens	551
Rudbeckia Goldsturm	496
Rubus	489
Nepeta Walker's Low	437
Rudbeckia Fulgida	379
Agastache Foeniculum	334
Monarda	290
Quercus alba	242
Canis	241
-77.702398	190
Callorhinus ursinus	178
Pycnanthemum muticum	158
Nepeta Faassenii	156

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

sourceTaxonName	interactionTypeName	targetTaxonName	count
Augochlora pura	interactsWith	Malus	7157
Eucera hamata	interactsWith	Prunus	6279
Augochlora pura	interactsWith	Prunus	4327
Ceratina calcarata	interactsWith	Malus	3463
Ceratina calcarata	interactsWith	Prunus	3335
Lasioglossum pilosum	interactsWith	Prunus	2642
Agapostemon virescens	interactsWith	Prunus	2174
Melissodes bimaculatus	interactsWith	Prunus	2100

sourceTaxonName	interactionTypeName	targetTaxonName	count
Bombus vagans	interactsWith	Malus	1714
Apis mellifera	interactsWith	Prunus	1612
Bombus impatiens	interactsWith	Malus	1419
Bombus impatiens	interactsWith	Prunus	1244
Apis mellifera	interactsWith	Malus	1132
Peponapis pruinosa	interactsWith	Prunus	852
Melissodes bimaculatus	interactsWith	Malus	698
Bombus vagans	interactsWith	Prunus	664
Ceratina mikmaqi	interactsWith	Malus	652
Peponapis pruinosa	interactsWith	Malus	537
Lasioglossum tegulare	interactsWith	Malus	481

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.

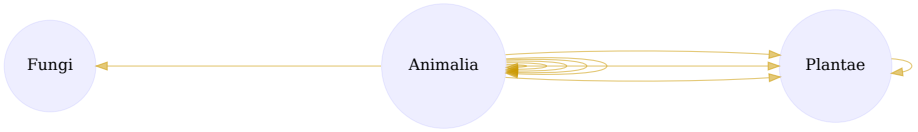

Figure 3: Interactions on taxonomic kingdom rank as interpreted by the Catalogue of Life download svg

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.

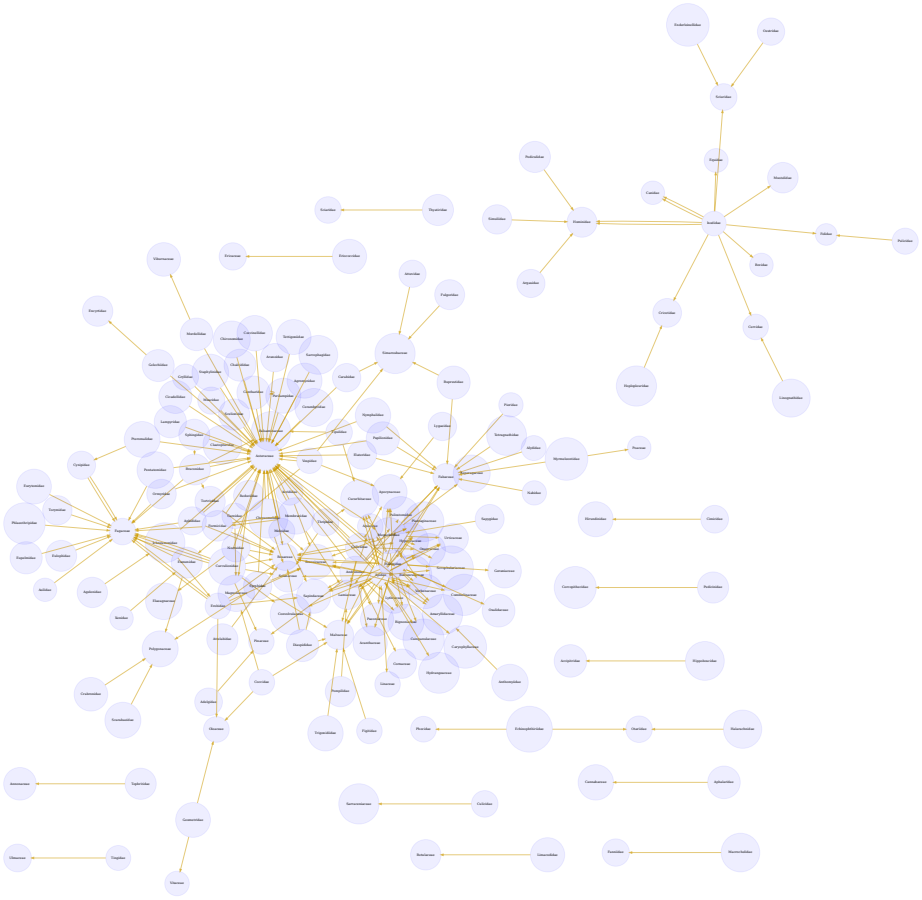

Figure 4: Interactions on the taxonomic family rank as interpreted by the Catalogue of Life. download svg

Table 8: Sample of Name Alignments

providedName	relationName	resolvedCatalogName	resolvedName
Baryscapus	HAS_ACCEPTED_NAME	col	Baryscapus
Bracon	HAS_ACCEPTED_NAME	col	Bracon
Celastrina	HAS_ACCEPTED_NAME	col	Celastrina
Chrysis	NONE	col	Chrysis

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	1520
col	class	2
col	family	137
col	genus	234
col	kingdom	2
col	nanorder	1
col	order	21
col	parvorder	1
col	phylum	1
col	species	974
col	subfamily	3
col	subgenus	10
col	suborder	1
col	subspecies	16
col	subterclass	1
col	superfamily	5
col	tribe	2
col	variety	5
discoverlife	NA	2649
discoverlife	species	255
gbif	NA	1313
gbif	class	2
gbif	family	143
gbif	form	1
gbif	genus	262
gbif	kingdom	3
gbif	order	21
gbif	phylum	1
gbif	species	1158

resolvedCatalogName	resolvedRank	count
gbif	subspecies	26
gbif	variety	11
itis	NA	1493
itis	class	2
itis	division	1
itis	family	137
itis	genus	207
itis	kingdom	2
itis	order	19
itis	species	1012
itis	subclass	1
itis	subfamily	5
itis	suborder	8
itis	subspecies	10
itis	superfamily	5
itis	superorder	2
itis	tribe	1
itis	variety	3
mdd	NA	2903
ncbi	NA	1576
ncbi	clade	1
ncbi	class	2
ncbi	cohort	1
ncbi	family	137
ncbi	genus	231
ncbi	infraorder	1
ncbi	kingdom	1
ncbi	order	21
ncbi	species	914
ncbi	subclass	1
ncbi	subfamily	4
ncbi	subgenus	9
ncbi	suborder	3
ncbi	subspecies	3
ncbi	subtribe	1
ncbi	superfamily	5
ncbi	superorder	1
ncbi	tribe	1
pbdb	NA	2546
pbdb	class	2
pbdb	family	134
pbdb	genus	129
pbdb	infraclass	1
pbdb	infraorder	2

resolvedCatalogName	resolvedRank	count
pbdb	kingdom	2
pbdb	order	23
pbdb	phylum	1
pbdb	species	53
pbdb	subfamily	4
pbdb	suborder	3
pbdb	superfamily	5
pbdb	tribe	2
pbdb	unranked clade	3
tpt	NA	2780
tpt	family	8
tpt	genus	15
tpt	order	4
tpt	species	96
wfo	NA	2597
wfo	family	5
wfo	genus	100
wfo	order	1
wfo	species	197
wfo	subspecies	4
wfo	subtribe	1
wfo	variety	2
worms	NA	2498
worms	family	109
worms	genus	113
worms	kingdom	2
worms	order	19
worms	phylum (division)	1
worms	species	154
worms	subclass	2
worms	suborder	3
worms	subspecies	1
worms	subterclass	1
worms	subtribe	1
worms	superfamily	1
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	1420
col	NONE	1764
col	SYNONYM_OF	354
discoverlife	NONE	3007
discoverlife	HAS_ACCEPTED_NAME	240
discoverlife	SYNONYM_OF	76
discoverlife	HOMONYM_OF	19
gbif	HAS_ACCEPTED_NAME	1741
gbif	SYNONYM_OF	752
gbif	NONE	1546
itis	SYNONYM_OF	142
itis	HAS_ACCEPTED_NAME	1431
itis	NONE	1737
mdd	NONE	3221
mdd	HAS_ACCEPTED_NAME	33
mdd	SYNONYM_OF	1
ncbi	SAME_AS	1380
ncbi	SYNONYM_OF	76
ncbi	NONE	1824
ncbi	COMMON_NAME_OF	3
pbdb	NONE	2861
pbdb	HAS_ACCEPTED_NAME	388
pbdb	SYNONYM_OF	39
tpt	NONE	3122
tpt	HAS_ACCEPTED_NAME	128
tpt	SYNONYM_OF	37
wfo	NONE	2883
wfo	HAS_ACCEPTED_NAME	356
wfo	SYNONYM_OF	73
wfo	HAS_UNCHECKED_NAME	45
worms	NONE	2790
worms	SYNONYM_OF	53
worms	HAS_ACCEPTED_NAME	481

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
2025-04-13T08:06:04Z	note	source taxon name missing: using institution-Code/collectionCode/collectionId/catalogNumber/occur as placeholder

reviewDate	reviewCommentType	reviewComment
2025-04-13T08:06:04Z	note	source taxon name missing: using institution-Code/collectionCode/collectionId/catalogNumber/occur as placeholder
2025-04-13T08:06:04Z	note	source taxon name missing: using institution-Code/collectionCode/collectionId/catalogNumber/occur as placeholder
2025-04-13T08:06:04Z	note	source taxon name missing: using institution-Code/collectionCode/collectionId/catalogNumber/occur as placeholder

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
source taxon name missing: using institution-Code/collectionCode/collectionId/catalogNumber/occurrenceId as placeholder	569
found unsupported interaction type with name: [parasitoid of]	78
found unsupported interaction type with name: [inquiline of]	78
target taxon name missing	12

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.

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

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 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.

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

[5] At time of writing (2025-04-13) the version of the GloBI dataset index was available at https://globalbioticinteractions.org/datasets

[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."

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.

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.