A Review of Biotic Interactions and Taxon Names Found in globalbioticinteractions/byu-byuc hash://md5/fe46d23dd90804b7735503c9254496f9

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

2025-04-10

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/byu-byuc, has fingerprint hash://md5/fe46d23dd90804b7735503c9254496f9, is 18.0MiB in size and contains 37,912 interaction with 7 unique types of associations (e.g., hasHost) between 1,424 primary taxon (e.g., Aetheca wagneri) and 2,371 associated taxon (e.g., Peromyscus maniculatus). This report includes detailed summaries of interaction data, a taxonomic review from multiple catalogs, and an archived version of the dataset from which the reviews are derived.

Contents

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:

> Brigham Young University Arthropod Museum https://github.com/globalbioticinteractions/byu-byuc/archive/4a609ac6a9a03425e2720b6cdebca6438488f029.zip 2025-04-04T23:30:07.302Z hash://md5/fe46d23dd90804b7735503c9254496f9

For additional metadata related to this dataset, please visit https://github.com/globalbioticinteractions/byu-byuc 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. 18.0MiB) under review
elton pull globalbioticinteractions/byu-byuc

# generate review notes
elton review globalbioticinteractions/byu-byuc\
 > review.tsv

# export indexed interaction records
elton interactions globalbioticinteractions/byu-byuc\
 > interactions.tsv

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

or visually, in a process diagram.

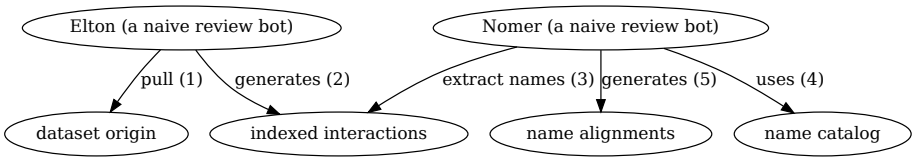

Figure 1: Review Process Overview

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

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

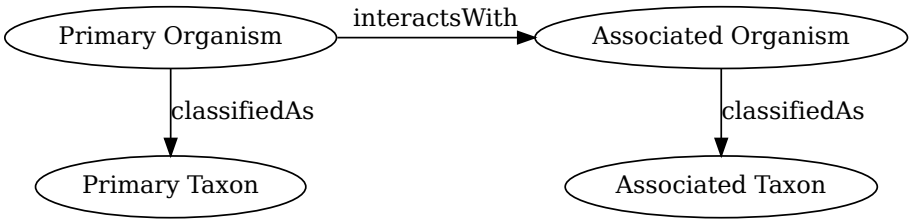

Figure 2: Biotic Interaction Data Model

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/byu-byuc, has fingerprint hash://md5/fe46d23dd90804b7735503c9254496f9, is 18.0MiB in size and contains 37,912 interaction with 7 unique types of associations (e.g., hasHost) between 1,424 primary taxon (e.g., Aetheca wagneri) and 2,371 associated taxon (e.g., Peromyscus maniculatus).

An exhaustive list of indexed interaction claims can be found in gzipped csv and tsv archives. To facilitate discovery, a preview of claims available in the gzipped html page at indexed-interactions.html.gz 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
Axinopalpus biplagiatus	hasHost	Ex nest of Neotoma	https://scan-bugs.org:443/portal/collections/individual/index.
Axinopalpus biplagiatus	hasHost	Ex nest of Neotoma	https://scan-bugs.org:443/portal/collections/individual/index.
Axinopalpus biplagiatus	hasHost	Ex nest of Neotoma	https://scan-bugs.org:443/portal/collections/individual/index.
Amara farcta	interactsWith	Salsola pestifer	https://scan-bugs.org:443/portal/collections/individual/index.

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

interactionTypeName	count
hasHost	24965
interactsWith	12649
adjacentTo	251

interactionTypeName	count
hostOf	36
hasParasite	9
eats	1
visits	1

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

sourceTaxonName	count
Aetheca wagneri	4727
Meringis parkeri	2951
Thrassis bacchi	1364
Spilopsyllus inaequalis	1010
Orchopeas sexdentatus	933
Hoplopsyllus anomalus	911
Monopsyllus wagneri wagneri	851
Eumolpianus eumolpi	840
Thrassis pandorae	730
Opisodasys keeni	641
Atyphloceras multidentatus multidentatus	636
Oropsylla idahoensis	632
Foxella ignota	629
Hoplopleura hesperomydis	604
Thrassis aridis	586
Thrassis francisi	468
Atyphloceras multidentatus	453
Oropsylla montanus	443
Anomiopsyllus amphibolus	429

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

targetTaxonName	count
Peromyscus maniculatus	7445
Dipodomys ordii	2147
Ammospermophilus leucurus	1255
Microtus	1172
Neotoma lepida	1019
Spermophilus armatus	865
Eutamias minimus	644

targetTaxonName	count
Spermophilus townsendii	620
Lepus californicus	560
Onychomys leucogaster	532
Citellus leucurus	479
Peromyscus Maniculatus	430
Neotoma cinerea	427
Citellus variegatus	374
Lepus Californicus	333
Citellus armatus	320
Dipodomys microps	309
Sylvilagus	301
Abrothrix longipilis	270

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

sourceTaxonName	interactionTypeName	targetTaxonName	count
Aetheca wagneri	hasHost	Peromyscus maniculatus	3455
Meringis parkeri	hasHost	Dipodomys ordii	1745
Monopsyllus wagneri wagneri	interactsWith	Peromyscus maniculatus	742
Atyphloceras multidentatus multidentatus	hasHost	Microtus	636
Hoplopleura hesperomydis	hasHost	Peromyscus maniculatus	570
Thrassis bacchi	interactsWith	Ammospermophilus leucurus	536
Aetheca wagneri	interactsWith	Peromyscus maniculatus	455
Thrassis pandorae	hasHost	Spermophilus armatus	410
Thrassis bacchi	hasHost	Ammospermophilus leucurus	382
Opisodasys keeni	interactsWith	Peromyscus maniculatus	329
Atyphloceras multidentatus	hasHost	Microtus	325
Aetheca wagneri	hasHost	Peromyscus Maniculatus	318

sourceTaxonName	interactionTypeName	targetTaxonName	count
Anomiopsyllus amphibolus	interactsWith	Neotoma lepida	292
Thrassis francisi	interactsWith	Spermophilus townsendii	282
Meringis parkeri	interactsWith	Dipodomys ordii	266
Oropsylla montanus	hasHost	Citellus variegatus	262
Oropsylla idahoensis	hasHost	Spermophilus armatus	235
Hoplopsyllus anomalus	hasHost	Ammospermophilus leucurus	235
Opisocrostis t. tuberculatus	hasHost	Spermophilus townsendii	219

Interaction Networks

The figures below provide a graph view on the dataset under review. The first shows a summary network on the kingdom level, and the second shows how interactions on the family level. It is important to note that both network graphs were first aligned taxonomically using the Catalogue of Life. Please refer to the original (or verbatim) taxonomic names for a more original view on the interaction data.

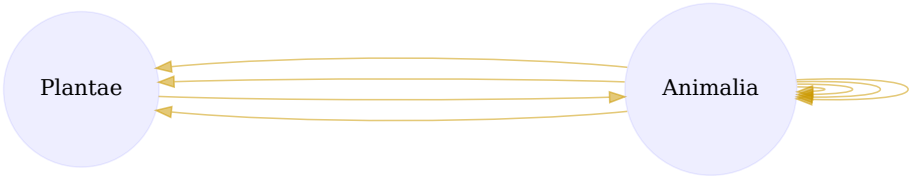

Figure 3: Interactions on taxonomic kingdom rank as interpreted by the Catalogue of Life download svg

You can download the indexed dataset under review at indexed-interactions.csv .gz. A tab-separated file can be found at indexed-interactions.tsv.gz

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

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

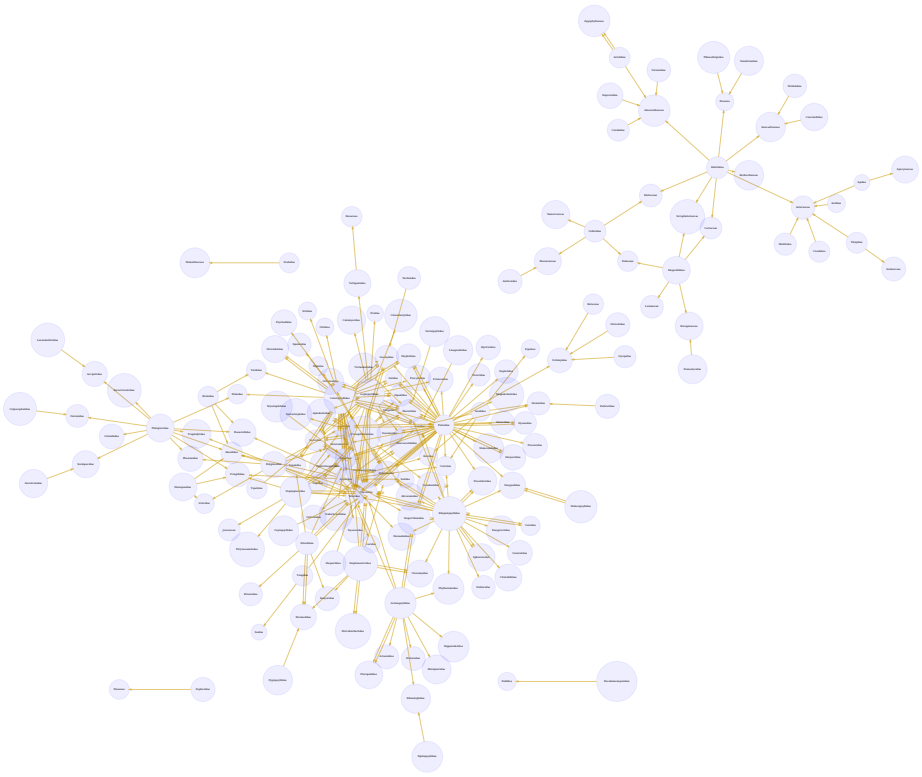

Figure 4: Interactions on the taxonomic family rank as interpreted by the Catalogue of Life. download svg

Taxonomic Alignment

As part of the review, all names are aligned against various name catalogs (e.g., col, ncbi, discoverlife, gbif, itis, wfo, mdd, tpt, pbdb, and worms). These alignments can help review name usage or aid in selecting of a suitable taxonomic name resource.

Table 8: Sample of Name Alignments

providedName	relationName	resolvedCatalogName	resolvedName
Baculomeris schmidti	HAS_ACCEPTED_NAME	col	Baculomeris schmidti
Barreropsylla excelsa	HAS_ACCEPTED_NAME	col	Barreropsylla excelsa
Caenopsylla mira	HAS_ACCEPTED_NAME	col	Caenopsylla mira
Callopsylla caspia	HAS_ACCEPTED_NAME	col	Callopsylla caspia

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	1744
col	family	17
col	genus	218
col	kingdom	1
col	nanorder	1
col	order	2
col	parvorder	1
col	phylum	2
col	species	1194
col	subfamily	3
col	subgenus	29
col	subspecies	102
col	subtribe	1
col	variety	3
discoverlife	NA	3262
discoverlife	species	19
gbif	NA	1660
gbif	family	17
gbif	genus	241
gbif	kingdom	2
gbif	order	2

resolvedCatalogName	resolvedRank	count
gbif	phylum	2
gbif	species	1244
gbif	subspecies	121
gbif	variety	3
itis	NA	2332
itis	family	17
itis	genus	164
itis	kingdom	1
itis	order	3
itis	phylum	2
itis	species	703
itis	subfamily	3
itis	subgenus	1
itis	suborder	2
itis	subspecies	54
itis	variety	1
mdd	NA	3280
ncbi	NA	2140
ncbi	family	17
ncbi	genus	207
ncbi	order	3
ncbi	phylum	2
ncbi	species	862
ncbi	subfamily	3
ncbi	subgenus	8
ncbi	suborder	1
ncbi	subspecies	46
pbdb	NA	2817
pbdb	family	13
pbdb	genus	133
pbdb	kingdom	1
pbdb	order	3
pbdb	phylum	2
pbdb	species	309
pbdb	subfamily	2
pbdb	suborder	1
pbdb	unranked clade	1
tpt	NA	2152
tpt	family	7
tpt	genus	109
tpt	order	1
tpt	species	1011
wfo	NA	3195
wfo	family	1

resolvedCatalogName	resolvedRank	count
wfo	genus	30
wfo	species	52
wfo	subspecies	2
wfo	variety	2
worms	NA	3069
worms	family	12
worms	genus	85
worms	kingdom	1
worms	order	3
worms	phylum	2
worms	species	106
worms	suborder	2
worms	subspecies	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	1822
col	SYNONYM_OF	285
col	NONE	1915
discoverlife	NONE	3808
discoverlife	SYNONYM_OF	6
discoverlife	HAS_ACCEPTED_NAME	18
gbif	HAS_ACCEPTED_NAME	1897
gbif	SYNONYM_OF	434
gbif	NONE	1830
itis	NONE	2565
itis	HAS_ACCEPTED_NAME	1096
itis	SYNONYM_OF	177
mdd	NONE	3318
mdd	HAS_ACCEPTED_NAME	472
mdd	SYNONYM_OF	29
ncbi	SAME_AS	1334
ncbi	NONE	2365
ncbi	SYNONYM_OF	141

resolvedCatalogName	relationName	count
ncbi	COMMON_NAME_OF	4
pbdb	NONE	3072
pbdb	HAS_ACCEPTED_NAME	671
pbdb	SYNONYM_OF	84
tpt	HAS_ACCEPTED_NAME	1577
tpt	SYNONYM_OF	230
tpt	NONE	2383
wfo	NONE	3727
wfo	SYNONYM_OF	25
wfo	HAS_ACCEPTED_NAME	76
wfo	HAS_UNCHECKED_NAME	5
worms	NONE	3475
worms	HAS_ACCEPTED_NAME	307
worms	SYNONYM_OF	57

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
2025-04-10T08:49:30Z	note	source taxon name missing: using institution-Code/collectionCode/collectionId/catalogNumber/occur as placeholder
2025-04-10T08:49:30Z	note	source taxon name missing: using institution-Code/collectionCode/collectionId/catalogNumber/occur as placeholder
2025-04-10T08:49:30Z	note	source taxon name missing: using institution-Code/collectionCode/collectionId/catalogNumber/occur as placeholder
2025-04-10T08:49:30Z	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	273
found unsupported interaction type with name: [Peromyscus maniculatus sonoriensis (found]	68

reviewComment	count
found unsupported interaction type with name: [hasH]	25
found unsupported interaction type with name: [lfp]	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.

Figure 5: Picture of a GloBI Review Badge [3]

Note that if the badge is green, no review notes were generated. If the badge is yellow, the review bots may need some help with interpreting the species interaction data.

GloBI Index Badge

If the dataset under review has been registered with GloBI, and has been succesfully indexed by GloBI, the GloBI Index Status Badge will turn green. This means that the dataset under review was indexed by GloBI and is available through GloBI services and derived data products.

Figure 6: Picture of a GloBI Index Badge [4]

If you'd like to keep track of reviews or index status of the dataset under review, please visit GloBI's dataset index [5] for badge examples.

Discussion

This review and archive provides a means of creating citable versions of datasets that change frequently. This may be useful for dataset managers, including

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

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

[5] At time of writing (2025-04-10) the version of the GloBI dataset index was available at https://globalbioticinteractions.org/datasets

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

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

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

Acknowledgements

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

Author contributions

Nomer was responsible for name alignments. Elton carried out dataset extraction, and generated the review notes. Preston tracked, versioned, and packaged, the dataset under review.

References

Elliott, Michael, Jorrit Poelen, Icaro Alzuru, Emilio Berti, and partha04patel. 2025. "Bio-Guoda/Preston: 0.10.5." Zenodo. https://doi.org/10.5281/zenodo.14662206.

[6]According to http://opendefinition.org/: "Open data is data that can be freely used, re-used and redistributed by anyone - subject only, at most, to the requirement to attribute and sharealike."

ICZN. 1999. “International Code of Zoological Nomenclature.” The International Trust for Zoological Nomenclature, London, UK. https://www.iczn.org/the-code/the-code-online/.

Kuhn, Tobias, and Michel Dumontier. 2014. “Trusty URIs: Verifiable, Immutable, and Permanent Digital Artifacts for Linked Data.” In *The Semantic Web: Trends and Challenges*, edited by Valentina Presutti, Claudia d’Amato, Fabien Gandon, Mathieu d’Aquin, Steffen Staab, and Anna Tordai, 395–410. Cham: Springer International Publishing.

Kuhn, Tobias, Jorrit Poelen, and Katrin Leinweber. 2025. “Globalbioticinteractions/Elton: 0.15.1.” Zenodo. https://doi.org/10.5281/zenodo.14927734.

Poelen, Jorrit H. (ed.). 2024. “Nomer Corpus of Taxonomic Resources Hash://Sha256/ B60c0d25a16ae77b24305782017b1a270b79b5d1746f832650F2027ba536e276 Hash://Md5/17f1363a277ee0e4ecaf1b91c665e47e.” Zenodo. https://doi.org/10.5281/zenodo.12695629.

Poelen, Jorrit H., James D. Simons, and Chris J. Mungall. 2014. “Global Biotic Interactions: An Open Infrastructure to Share and Analyze Species-Interaction Datasets.” *Ecological Informatics* 24 (November): 148–59. https://doi.org/10.1016/j.ecoinf.2014.08.005.

Poelen, Jorrit, Katja Seltmann, and Daniel Mietchen. 2024. “Globalbioticinteractions/Globinizer: 0.4.0.” Zenodo. https://doi.org/10.5281/zenodo.10647565.

Salim, José Augusto, and Jorrit Poelen. 2025. “Globalbioticinteractions/Nomer: 0.5.15.” Zenodo. https://doi.org/10.5281/zenodo.14893840.

Trekels, Maarten, Debora Pignatari Drucker, José Augusto Salim, Jeff Ollerton, Jorrit Poelen, Filipi Miranda Soares, Max Rünzel, Muo Kasina, Quentin Groom, and Mariano Devoto. 2023. “WorldFAIR Project (D10.1) Agriculture-related pollinator data standards use cases report.” Zenodo. https://doi.org/10.5281/zenodo.8176978.

Wilkinson, Mark D., Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, et al. 2016. “The FAIR Guiding Principles for Scientific Data Management and Stewardship.” *Scientific Data* 3 (1). https://doi.org/10.1038/sdata.2016.18.