

Versioned Archive and Review of Biotic
Interactions and Taxon Names Found within
globalbioticinteractions/asu-asuhic
hash://md5/285ffd1b5d4b0c5777d3bfe92dbcd186

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

2025-06-28

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/asu-asuhic, has fingerprint hash://md5/285ffd1b5d4b0c5777d3bfe92dbcd186, is 20.4MiB in size and contains 24,931 interaction with 4 unique types of associations (e.g., interactsWith) between 578 primary taxa (e.g., *Oncerus floralis*) and 812 associated taxa (e.g., *Allionia incarnata* L. (trailing windmills) [USDA]). This report includes detailed summaries of interaction data, a taxonomic review from multiple catalogs, and an archived version of the dataset from which the reviews are derived.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	13

Biotic Interactions	13
Interaction Networks	18
Taxonomic Alignment	20
Additional Reviews	23
GloBI Review Badge	25
GloBI Index Badge	25
Discussion	25
Acknowledgements	26
Author contributions	26
References	26

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:

Arizona State University Hasbrouck Insect Collection <https://github.com/globalbioticinteractions/asu-asuhic/archive/4ed77cb9ca8e526269d4678692e2844c950022f8.zip>
2025-06-28T00:44:32.866Z hash://md5/285ffd1b5d4b0c5777d3bfe92dbcd186

For additional metadata related to this dataset, please visit <https://github.com/globalbioticinteractions/asu-asuhic> 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	0.15.13
nomer	0.5.17
globinizer	0.4.0
mlr	6.0.0
jq	1.6
yq	4.25.3
pandoc	3.1.6.1
duckdb	1.3.1

The review process can be described in the form of the script below ¹.

```
# get versioned copy of the dataset (size approx. 20.4MiB) under review
elton pull globalbioticinteractions/asu-asuhic

# generate review notes
elton review globalbioticinteractions/asu-asuhic\
> review.tsv

# export indexed interaction records
elton interactions globalbioticinteractions/asu-asuhic\
> interactions.tsv

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

or visually, in a process diagram.

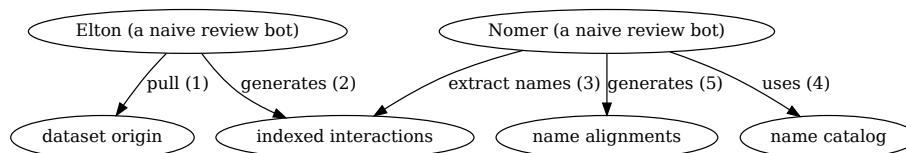


Figure 1: Review Process Overview

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

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

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 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 Nomen 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.parquet	species interaction claims indexed from the dataset under review in Apache Parquet 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.parquet	taxonomic names found in the dataset under review in Apache Parquet format

filename	description
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
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-col.parquet	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 Apache Parquet 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

filename	description
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-discoverlife.parquet	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 Apache Parquet 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
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-gbif.parquet	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 Apache Parquet format

filename	description
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-itis.parquet	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 Apache Parquet 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
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

filename	description
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-mdd.parquet	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 Apache Parquet 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-ncbi.parquet	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 Apache Parquet format

filename	description
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
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-pbdb.parquet	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 Apache Parquet 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

filename	description
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-tpt.parquet	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 Apache Parquet 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
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-wfo.parquet	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 Apache Parquet format

filename	description
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-resolved-worms.parquet	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 Apache Parquet 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
nanopub-sample.trig	first 500 species interaction claims as expressed in the nanopub format (Kuhn and Dumontier 2014)

filename	description
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 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/asu-asuhic, has fin-

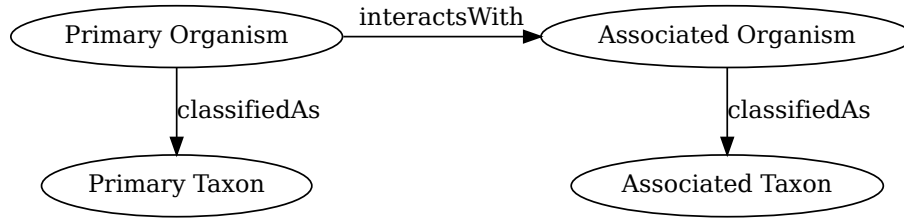


Figure 2: Biotic Interaction Data Model

gerprint hash://md5/285ffd1b5d4b0c5777d3bfe92dbcd186, is 20.4MiB in size and contains 24,931 interaction with 4 unique types of associations (e.g., interactsWith) between 578 primary taxa (e.g., *Oncerus floralis*) and 812 associated taxa (e.g., *Allionia incarnata* L. (trailing windmills) [USDA]).

An exhaustive list of indexed interaction claims can be found in gzipped csv, tsv and parquet archives. To facilitate discovery, a preview of claims available in the gzipped html page at [indexed-interactions.html.gz](https://indexing-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
Hirmoneura flavipes	adjacentTo	palm shaped leaves	https://ecdysis.org/collections/individual/index.
Neorhynchocephalus sackenii	hasHost	Ex. Morseiella flaviventris (Bruner)	https://ecdysis.org/collections/individual/index.
Neorhynchocephalus sackenii	hasHost	Ex. Morseiella flaviventris (Bruner)	https://ecdysis.org/collections/individual/index.
Neorhynchocephalus sackenii	hasHost	Ex. Morseiella flaviventris (Bruner)	https://ecdysis.org/collections/individual/index.

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

interactionTypeName	count
interactsWith	14755
adjacentTo	9901
hasHost	263

interactionTypeName	count
eats	12

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

sourceTaxonName	count
Oncerus floralis	1182
Acmaeodera pubiventris lanata	745
Acmaeodera gibbula	689
Paracotalpa ursina	675
Acmaeodera lanata	673
Apidae	646
Acmaeodera diffusa	615
Acmaeodera bowditchi	609
Acmaeodera pulchella	589
Acmaeodera acanthicola	583
Acmaeodera decipiens	560
Acmaeodera cazieri	544
Acmaeodera flavopicta	521
Gyascutus planicosta cribriceps	509
Acmaeodera yuccavora	466
Rhipiphorus vierecki	434
Hippomelas sphenicus	422
Acmaeodera parkeri	408
Hippomelas planicosta	404

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

targetTaxonName	count
Allionia incarnata L. (trailing windmills) [USDA]	1284
Chaenactis fremontii A. Gary (pincushion flower) [USDA]	1134
Prosopis juliflora (Sw.) DC. (mesquite) [USDA]	878
Sphaeralcea laxa Wooton & Standl. (caliche globemallow) [USDA]	757
Asclepias erosa Torr. (desert milkweed) [USDA]	669

targetTaxonName	count
Larrea tridentata (DC.) Coville (creosote bush) [USDA]	622
Poliomintha incana (Torr.) A. Gray (frosted mint) [USDA]	591
Cercidum microphyllum (Torr.) Rose & I.M. Johnst. (yellow paloverde) [USDA]	536
Poliomintha incana (Torr.) Gray (frosted mint) [USDA]	464
Tamarix pentandra Pall. (salt cedar) [USDA]	445
Ephedra trifurca Torr. ex S. Watson (longleaf jointfir) [USDA]	429
a 15 watt 12 volt battery black light	402
Baileya multiradiata Harv. & A. Gray ex A. Gray (desert marigold) [USDA]	380
dead Hilaria rigida (Thurb.) Benth	379
Dalea spinosa A. Gray (smoketree) [USDA]	378
grass stems	367
Acacia constricta Benth. (whitethorn acacia) [USDA]	360
propylene glycol pit fall cup trap	346
yellow Compositae sp.	318

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

sourceTaxonName	interactionType	targetTaxonName	count
Oncerus floralis	interactsWith	Chaenactis fremontii A. Gary (pincushion flower) [USDA]	1134
Acmaeodera diffusa	adjacentTo	Sphaeralcea laxa Wooton & Standl. (caliche globemallow) [USDA]	559

sourceTaxonName	interactionType	targetTaxonName	count
Acmaeodera acanthicola	interactsWith	Cercidum microphyllum (Torr.) Rose & I.M. Johnst. (yellow paloverde) [USDA]	517
Acmaeodera yuccavora	adjacentTo	Allionia incarnata L. (trailing windmills) [USDA]	462
Dinacoma caseyi	interactsWith	a 15 watt 12 volt battery black light	394
Hippomelas planicosta	interactsWith	Larrea tridentata (DC.) Coville (creosote bush) [USDA]	382
Acmaeodera parkeri	adjacentTo	Allionia incarnata L. (trailing windmills) [USDA]	381
Acmaeodera pubiventris lanata	adjacentTo	dead Hilaria rigida (Thurb.) Benth	379
Acmaeodera barri	interactsWith	Dalea spinosa A. Gray (smoketree) [USDA]	376
Heteropogon johnsoni	adjacentTo	grass stems	359
Hippomelas sphenicus	adjacentTo	Prosopis juliflora (Sw.) DC. (mesquite) [USDA]	353
Paracotalpa ursina	interactsWith	propylene glycol pit fall cup trap	346
Gyascutus planicosta cribriceps	interactsWith	Tamarix pentandra Pall. (salt cedar) [USDA]	345

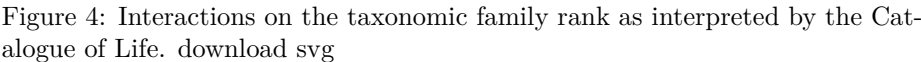
sourceTaxonName	interactionType	targetTaxonName	count
Acmaeodera cazieri	adjacentTo	Allionia incarnata L. (trailing windmills) [USDA]	335
Acmaeodera pubiventris lanata	interactsWith	Poliomintha incana (Torr). Gray (frosted mint) [USDA]	332
Macrosiagon fernaldum	interactsWith	Asclepias erosa Torr. (desert milkweed) [USDA]	329
Acmaeodera decipiens	adjacentTo	yellow Compositae sp.	252
Apidae	interactsWith	Poliomintha incana (Torr.) A. Gray (frosted mint) [USDA]	239
Paracotalpa ursina	interactsWith	propylene glycol pit fall trap set 12-December, 2019	239

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, pldb, 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
Aciurina	HAS_ACCEPTED_NAME	col	Aciurina
mexicana			mexicana
Aciurina trixa	HAS_ACCEPTED_NAME	col	Aciurina trixa
Archilestris	HAS_ACCEPTED_NAME	col	Archilestris
magnificus			magnificus
Blepharepium	HAS_ACCEPTED_NAME	col	Blepharepium
secabile			secabile

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	541
col	family	17
col	genus	118
col	infraspecific name	2
col	kingdom	1
col	section	1
col	species	463
col	subfamily	1
col	subgenus	2
col	subspecies	10
col	superfamily	2
col	tribe	1

resolvedCatalogName	resolvedRank	count
col	variety	1
discoverlife	NA	1101
discoverlife	species	55
gbif	NA	484
gbif	family	18
gbif	genus	127
gbif	kingdom	1
gbif	species	507
gbif	subspecies	21
gbif	variety	3
itis	NA	558
itis	family	17
itis	genus	112
itis	kingdom	1
itis	species	441
itis	subfamily	1
itis	subspecies	16
itis	superfamily	2
itis	tribe	1
itis	variety	8
mdd	NA	1156
ncbi	NA	749
ncbi	family	16
ncbi	genus	118
ncbi	species	265
ncbi	subgenus	10
ncbi	subspecies	3
ncbi	superfamily	2
ncbi	tribe	1
ncbi	varietas	1
pbdb	NA	1068
pbdb	family	18
pbdb	genus	62
pbdb	kingdom	1
pbdb	species	4
pbdb	superfamily	2
pbdb	tribe	1
tpt	NA	1154
tpt	genus	2
wfo	NA	1007
wfo	family	1
wfo	genus	73
wfo	phylum	1
wfo	species	69

resolvedCatalogName	resolvedRank	count
wfo	subspecies	6
wfo	variety	1
worms	NA	1066
worms	family	9
worms	genus	54
worms	kingdom	1
worms	species	22
worms	subspecies	1
worms	superfamily	1
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	HAS_ACCEPTED_NAME	657
col	SYNONYM_OF	115
col	NONE	676
discoverlife	NONE	1337
discoverlife	SYNONYM_OF	18
discoverlife	HAS_ACCEPTED_NAME	50
gbif	HAS_ACCEPTED_NAME	751
gbif	SYNONYM_OF	173
gbif	NONE	619
itis	HAS_ACCEPTED_NAME	630
itis	SYNONYM_OF	84
itis	NONE	694
mdd	NONE	1392
ncbi	NONE	891
ncbi	SAME_AS	493
ncbi	SYNONYM_OF	28
pbdb	NONE	1255
pbdb	HAS_ACCEPTED_NAME	138
pbdb	SYNONYM_OF	3
tpt	NONE	1390
tpt	HAS_ACCEPTED_NAME	2
wfo	NONE	1154

resolvedCatalogName	relationName	count
wfo	HAS_ACCEPTED_NAME	192
wfo	HAS_UNCHECKED_NAME	12
wfo	SYNONYM_OF	60
worms	NONE	1248
worms	HAS_ACCEPTED_NAME	132
worms	SYNONYM_OF	22

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
2025-06-28T05:15:44Z	note	source taxon name missing: using institution- Code/collectionCode/collectionId/catalogNumber/occurrenceId as placeholder
2025-06-28T05:15:44Z	note	source taxon name missing: using institution- Code/collectionCode/collectionId/catalogNumber/occurrenceId as placeholder
2025-06-28T05:15:45Z	note	found unsupported interaction type with name: [7]
2025-06-28T05:15:45Z	note	source taxon name missing: using institution- Code/collectionCode/collectionId/catalogNumber/occurrenceId 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
found unsupported interaction type with name: [attacks]	54
source taxon name missing: using institution- Code/collectionCode/collectionId/catalogNumber/occurrenceId as placeholder	27
found unsupported interaction type with name: [Cupressaceae]	3
found unsupported interaction type with name: [Pinaceae]	3

For additional information on review notes, please have a look at the first 500 Review Notes in html format or the download full gzipped csv or tsv archives.

GloBI Review Badge

As part of the review, a review badge is generated. This review badge can be included in webpages to indicate the review status of the dataset under review.



Figure 5: Picture of a GloBI Review Badge ³

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

GloBI Index Badge

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



Figure 6: Picture of a GloBI Index Badge ⁴

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

Discussion

This review and archive provides a means of creating citable versions of datasets that change frequently. This may be useful for dataset managers, including natural history collection data managers, as a backup archive of a shared Darwin 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

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 (2025-06-28) the version of the GloBI dataset index was available at <https://globalbioticinteractions.org/datasets>

indication of the open-ness⁶ and FAIRness (Wilkinson et al. 2016; Trekels et al. 2023) of the dataset: to perform this review, the data was likely openly available, **F**indable, **A**ccessible, **I**nteroperable and **R**eusable. The current Open-FAIR assessment is qualitative, and a more quantitative approach can be implemented with specified measurement units.

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

Acknowledgements

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

Author contributions

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

References

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

⁶According to <http://opendefinition.org/>: “Open data is data that can be freely used, reused and redistributed by anyone - subject only, at most, to the requirement to attribute and sharealike.”

- 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>.