# A Review of Biotic Interactions and Taxon Names Found in globalbioticinteractions/sdnhm-sdmc hash://md5/6cd67357f559b97c0e6916fcdbc27dbb

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

#### 2025-05-02

#### 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/sdnhm-sdmc, has fingerprint hash://md5/6cd67357f559b97c0e6916fcdbc27dbb, is 11.8MiB in size and contains 8,688 interaction with 4 unique types of associations (e.g., interactsWith) between 746 primary taxa (e.g., Gelechidae) and 380 associated taxa (e.g., Oak). This report includes detailed summaries of interaction data, a taxonomic review from multiple catalogs, and an archived version of the dataset from which the reviews are derived.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	<b>2</b>
Data Review and Archive	2
Methods	<b>2</b>
Results	4
Files	4
Archived Dataset	11
Biotic Interactions	11
Interaction Networks	15
Taxonomic Alignment	15

Additional Reviews	20
GloBI Review Badge	21
GloBI Index Badge	21
Discussion	<b>21</b>
Acknowledgements	<b>22</b>
Author contributions	<b>22</b>
References	<b>22</b>

## 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:

San Diego Natural History Museum https://github.com/globalbioticinteractions/sdnhm-sdmc/archive/7238d8b804f543250eb487b43144e1125fb3688a.zip 2025-04-26T03:59:03.065Z hash://md5/6cd67357f559b97c0e6916fcdbc27dbb

For additional metadata related to this dataset, please visit https://github.c om/globalbioticinteractions/sdnhm-sdmc 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.

Tabl	е 1	L:	Tool	$\mathbf{S}$	used	in	this	review	process
10001	~ -		- O O -	~			01110	1011011	p100000

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. 11.8MiB) under review
elton pull globalbioticinteractions/sdnhm-sdmc

```
# generate review notes
elton review globalbioticinteractions/sdnhm-sdmc\
    > review.tsv
# export indexed interaction records
elton interactions globalbioticinteractions/sdnhm-sdmc\
    > interactions.tsv
```

# export names and align them with the Catalogue of Life using Nomer elton names globalbioticinteractions/sdnhm-sdmc\

- | 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.

<sup>&</sup>lt;sup>1</sup>Note that you have to first get the data (e.g., via elton pull globalbioticinteractions/sdnhmsdmc) before being able to generate reviews (e.g., elton review globalbioticinteractions/sdnhmsdmc), extract interaction claims (e.g., elton interactions globalbioticinteractions/sdnhm-sdmc), or list taxonomic names (e.g., elton names globalbioticinteractions/sdnhm-sdmc)

## 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)

 $<sup>^{2}</sup>$ 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 gripped 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 gripped 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 gripped 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 gripped 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 gripped 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 gripped 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 tak experiented subset 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
indexed-names-sample.tsv	dataset under review in html format 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
prov na	processing worknow
provinq	expressed in rdf/nquads
review.csv.gz	review notes associated with the
1011011001182	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
review-sample html	first 500 review notes associated with
review-sample.nemi	the dataset under review in html
	format
review-sample.tsv	first 500 review notes associated with
Ĩ	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/sdnhm-sdmc, has fingerprint hash://md5/6cd67357f559b97c0e6916fcdbc27dbb, is 11.8MiB in size and contains 8,688 interaction with 4 unique types of associations (e.g., interactsWith) between 746 primary taxa (e.g., Gelechiidae) and 380 associated taxa (e.g., Oak).

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.

sourceTaxonName	interactionTypeN	ameargetTaxonName	referenceCitation
Poanes melane	interactsWith	Aira caespitosa	https://scan-
			bugs.org:443/portal/collections/individual/index
Tortricidae	interactsWith	Pine	https://scan-
			bugs.org:443/portal/collections/individual/index
Tortricidae	hasHost	diseased pine tree	https://scan-
			bugs.org:443/portal/collections/individual/index
Argyrotaenia	interactsWith	Christmas cactus	https://scan-
citrana			bugs.org:443/portal/collections/individual/index

 Table 3: Sample of Indexed Interaction Claims

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

interactionTypeName	count
interactsWith	8159
adjacentTo	440
hasHost	88

interactionTypeName	count
eats	1

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

sourceTaxonName	count
Gelechiidae	292
Orthosia erythrolita	192
Glaucina epiphysaria	169
Malacosoma californicum andreasum	162
Plataea personaria	145
Gnorimoschema	136
Apodemia mormo	132
Hymenoptera	122
Stamnodes annellata	118
Zenophleps lignicolorata	109
Ethmia discostrigella	104
Miodera stigmata	102
Amydria	96
Eucosma pulveratana	92
Notarctia proxima	80
Stamnodes coenonymphata	79
Grammia ornata	78
Speranza marcescaria	77
Pherne subpunctata	67

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

targetTaxonName	count
Oak	2100
Sycamore	1885
Sage	1327
Willow	853
Manzanita	516
Coleogyne ramosissima	153
Hemizonia conjugens	103
Bebbia	86
Eriogonum fasciculatum	69
Bebbia flowers	63

targetTaxonName	count
Pogogyne nudiuscula	45
Pine	44
Penstemon	42
Yucca	40
Baccharis sarothroides	36
Alfalfa	33
Foeniculum vulgare	26
Lasthenia californica	24
Eriogonum	21

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

sourceTaxonName	interactionTypeNar	m¢argetTaxonName	count
Gelechiidae	interactsWith	Sycamore	95
Gelechiidae	interactsWith	Sage	83
Gelechiidae	interactsWith	Oak	75
Malacosoma	interactsWith	Coleogyne	75
californicum		ramosissima	
andreasum			
Malacosoma	hasHost	Coleogyne	75
californicum		ramosissima	
andreasum			
Apodemia mormo	interactsWith	Bebbia	66
Orthosia	interactsWith	Sycamore	58
erythrolita			
Stamnodes	interactsWith	Oak	57
annellata			
Stamnodes	interactsWith	Sycamore	56
annellata			
Apodemia mormo	adjacentTo	Bebbia flowers	55
Glaucina	interactsWith	Sage	51
epiphysaria			
Ethmia	interactsWith	Sycamore	50
discostrigella			
Ethmia	interactsWith	Oak	50
discostrigella			
Glaucina	interactsWith	Sycamore	49
epiphysaria			
Plataea	interactsWith	Sage	49
personaria			
Gnorimoschema	interactsWith	Willow	46

sourceTaxonName	$interaction Type Nam {\tt e} arget Taxon Name$		count
Orthosia	interactsWith	Oak	44
erythrolita			
Gnorimoschema	interactsWith	Sage	43
Miodera stigmata	interactsWith	Sage	43

#### **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.

Table	8:	Sample	of	Name	Alignm	ents
-------	----	--------	----	------	--------	------

providedName	relationName	resolvedCatalogNam	resolvedName
Anobiidae	HAS_ACCEPTED_	₫ø₩E	Anobiidae

providedName	relationName	resolvedCatal	ogNamæesolvedName
Anticarsia gemmatalis	HAS_ACCEPTED_	_ðøÅME	Anticarsia gemmatalis
Anticarsia gemmatalis	SYNONYM_OF	$\operatorname{col}$	Anticarsia acutilinea
Anticarsia gemmatalis	SYNONYM_OF	col	Anticarsia anisospila

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.

${\it resolved} Catalog Name$	$\operatorname{resolvedRank}$	$\operatorname{count}$
col	NA	325
col	family	44
col	genus	107
col	infraspecific name	1
$\operatorname{col}$	kingdom	1
$\operatorname{col}$	order	4
$\operatorname{col}$	species	557
col	subfamily	8
col	subgenus	1
col	subspecies	19
col	subtribe	1
$\operatorname{col}$	superfamily	1
col	tribe	2
col	variety	3
discoverlife	NA	1055
discoverlife	species	12
gbif	NA	232
gbif	family	46
gbif	genus	118
gbif	kingdom	1
gbif	order	4
gbif	species	625
gbif	subspecies	49
gbif	variety	7
itis	NA	375
itis	family	44
itis	genus	87
itis	kingdom	1
itis	order	4

resolvedCatalogName	$\operatorname{resolvedRank}$	count
itis	phylum	1
itis	species	485
itis	subfamily	9
itis	suborder	1
itis	subspecies	53
itis	superfamily	1
itis	superorder	1
itis	tribe	1
itis	variety	6
mdd	NA	1067
ncbi	NA	375
ncbi	cohort	1
ncbi	family	44
ncbi	genus	102
ncbi	order	4
ncbi	species	501
ncbi	subfamily	9
ncbi	subgenus	4
ncbi	suborder	1
ncbi	subspecies	26
ncbi	superfamily	1
ncbi	tribe	2
pbdb	NA	977
pbdb	family	40
pbdb	genus	30
pbdb	infraclass	1
pbdb	infraorder	1
pbdb	kingdom	1
pbdb	order	4
pbdb	species	3
pbdb	subfamily	9
pbdb	suborder	1
pbdb	superfamily	1
tpt	NA	1067
wfo	NA	899
wfo	genus	66
wfo	species	98
wfo	subspecies	4
wfo	variety	1
worms	NA	914
worms	family	31
worms	genus	45
worms	kingdom	1
worms	order	4

resolvedCatalogName	$\operatorname{resolvedRank}$	count
worms	species	71
worms	suborder	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).

resolved Catalog Name	relationName	$\operatorname{count}$
col	HAS_ACCEPTED_NAME	715
col	SYNONYM_OF	100
col	NONE	370
discoverlife	NONE	1114
discoverlife	HAS_ACCEPTED_NAME	11
discoverlife	SYNONYM_OF	7
gbif	HAS_ACCEPTED_NAME	804
gbif	SYNONYM_OF	194
gbif	NONE	275
itis	HAS_ACCEPTED_NAME	673
itis	NONE	420
itis	SYNONYM_OF	35
mdd	NONE	1126
ncbi	SAME_AS	669
ncbi	NONE	420
ncbi	SYNONYM_OF	44
pbdb	SYNONYM_OF	4
pbdb	NONE	1029
pbdb	HAS_ACCEPTED_NAME	96
tpt	NONE	1126
wfo	NONE	944
wfo	HAS_ACCEPTED_NAME	164
wfo	SYNONYM_OF	31
wfo	HAS_UNCHECKED_NAME	3
worms	HAS_ACCEPTED_NAME	151
worms	NONE	968
worms	SYNONYM_OF	16



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

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)

Table 11: List of Available Name Alignment Reports

associated names alignments report	in
gzipped html, csv, and tsv)	

#### **Additional Reviews**

worms

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
2025-05-02T16:28:18Z	note	found unsupported interaction type with
2025-05-02T16:28:19Z	note	name: [attacked by] source taxon name missing: using
		institution- Code/collectionCode/collectionId/catalogNumber/occur
2025-05-02T16:28:19Z	note	as placeholder source taxon name missing: using
		institution- Code/collectionCode/collectionId/catalogNumber/occur
2025-05-02T16:28:19Z	note	as placeholder 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	46	
institution-		
Code/collectionCode/collectionId/catalogNumber/occurrenceId		
as placeholder		

20

reviewComment	count
found unsupported interaction type with name: [attacked by]	1
found unsupported interaction type with name: [15]	1

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

#### **GloBI Review Badge**

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

review 🎎

Figure 5: Picture of a GloBI Review Badge <sup>3</sup>

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.

GloBI 🗸

Figure 6: Picture of a GloBI Index Badge <sup>4</sup>

If you'd like to keep track of reviews or index status of the dataset under review, please visit GloBI's dataset index <sup>5</sup> 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

 $<sup>^{3}\</sup>mathrm{Up}\text{-to-date}$  status of the GloBI Review Badge can be retrieved from the GloBI Review Depot

<sup>&</sup>lt;sup>4</sup>Up-to-date status of the GloBI Index Badge can be retrieved from GloBI's API

 $<sup>{}^{5}</sup>$ At time of writing (2025-05-02) 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<sup>6</sup> 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/zeno do.14662206.

 $<sup>^{6}\</sup>rm According$  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/thecode/the-code-online/.
- Kuhn, Tobias, and Michel Dumontier. 2014. "Trusty URIs: Verifiable, Immutable, and Permanent Digital Artifacts for Linked Data." In *The Semantic Web: Trends and Challenges*, edited by Valentina Presutti, Claudia d'Amato, Fabien Gandon, Mathieu d'Aquin, Steffen Staab, and Anna Tordai, 395–410. Cham: Springer International Publishing.
- Kuhn, Tobias, Jorrit Poelen, and Katrin Leinweber. 2025. "Globalbioticinteractions/Elton: 0.15.1." Zenodo. https://doi.org/10.5281/zenodo.14927734.
- Poelen, Jorrit H. (ed.). 2024. "Nomer Corpus of Taxonomic Resources Hash://Sha256/ B60c0d25a16ae77b24305782017b1a270b79b5d1746f832650 F2027ba536e276 Hash://Md5/17f1363a277ee0e4ecaf1b91c665e47e." Zenodo. https://doi.org/10.5281/zenodo.12695629.
- Poelen, Jorrit H., James D. Simons, and Chris J. Mungall. 2014. "Global Biotic Interactions: An Open Infrastructure to Share and Analyze Species-Interaction Datasets." *Ecological Informatics* 24 (November): 148–59. https: //doi.org/10.1016/j.ecoinf.2014.08.005.
- Poelen, Jorrit, Katja Seltmann, and Daniel Mietchen. 2024. "Globalbioticinteractions/Globinizer: 0.4.0." Zenodo. https://doi.org/10.5281/zenodo.10647565.
- Salim, José Augusto, and Jorrit Poelen. 2025. "Globalbioticinteractions/Nomer: 0.5.15." Zenodo. https://doi.org/10.5281/zenodo.14893840.
- Trekels, Maarten, Debora Pignatari Drucker, José Augusto Salim, Jeff Ollerton, Jorrit Poelen, Filipi Miranda Soares, Max Rünzel, Muo Kasina, Quentin Groom, and Mariano Devoto. 2023. "WorldFAIR Project (D10.1) Agriculturerelated 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.