Versioned Archive and Review of Biotic Interactions and Taxon Names Found within globalbioticinteractions/inaturalist hash://md5/f442a757bbbf9b85038c262728d6d46a

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

2025-06-03

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/inaturalist, has fingerprint hash://md5/f442a757bbbf9b85038c262728d6d46a, is 606MiB in size and contains 1,082,852 interaction with 21 unique types of associations (e.g., visitsFlowersOf) between 45,950 primary taxon (e.g., Apis mellifera) and 31,749 associated taxon (e.g., Solidago). This report includes detailed summaries of interaction data, a taxonomic review from multiple catalogs, and an archived version of the dataset from which the reviews are derived.

Contents

Introduction **2**
Data Review and Archive 2

Methods **2**

Results **4**
Files 4
Archived Dataset 11
Biotic Interactions 12

Interaction Networks 15
Taxonomic Alignment 15
Additional Reviews 23
GloBI Review Badge 24
GloBI Index Badge 24

Discussion **25**

Acknowledgements **25**

Author contributions **25**

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:

> http://iNaturalist.org is a place where you can record what you see in nature, meet other nature lovers, and learn about the natural world. https://www.inaturalist.org/observations/globi-observations-resource-relationships-dwca.zip 2025-05-31T00:18:49.785Z hash://md5/f442a757bbbf9b85038c262728d6d46a

For additional metadata related to this dataset, please visit https://github.com/globalbioticinteractions/inaturalist 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	0.15.12
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.  606MiB) under review
elton pull globalbioticinteractions/inaturalist

# generate review notes
elton review globalbioticinteractions/inaturalist\
 > review.tsv

# export indexed interaction records
elton interactions globalbioticinteractions/inaturalist\
 > interactions.tsv

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

or visually, in a process diagram.

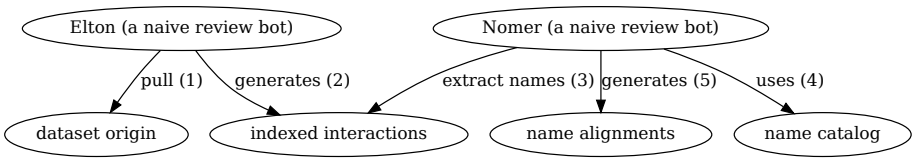

Figure 1: Review Process Overview

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

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)

[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

filename	description
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)
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.

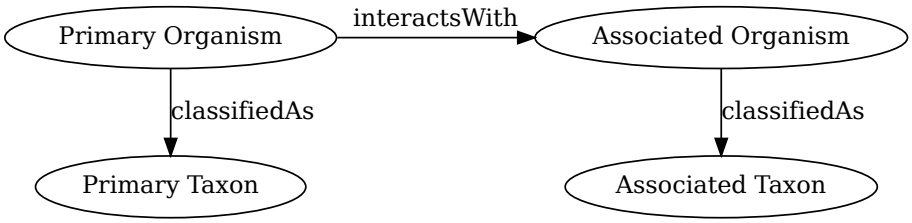

Figure 2: Biotic Interaction Data Model

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/inaturalist, has fingerprint hash://md5/f442a757bbbf9b85038c262728d6d46a, is 606MiB in size and contains 1,082,852 interaction with 21 unique types of associations (e.g., visitsFlowersOf) between 45,950 primary taxon (e.g., Apis mellifera) and 31,749 associated taxon (e.g., Solidago).

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
Leptosia nina	interactsWith	Cleome rutidosperma	https://www.inaturalist.org/observations/213158
Corvus corax	interactsWith	Aechmophorus	https://www.inaturalist.org/observations/213159
Graphium agamemnon	interactsWith	Monoon longifolium	https://www.inaturalist.org/observations/213161
Bombus impatiens	visitsFlowersOf	Rhododendron	https://www.inaturalist.org/observations/213162

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

interactionTypeName	count
visitsFlowersOf	380277
interactsWith	340476
eats	169324
visits	73737
hasHost	46726
preysOn	40753
preyedUponBy	9195
flowersVisitedBy	5433
parasiteOf	4721
eatenBy	4534
hasParasite	3871
pathogenOf	1173
hostOf	1031
hasPathogen	593
kills	393
parasitoidOf	271
symbiontOf	126
laysEggsOn	114
killedBy	84

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

sourceTaxonName	count
Apis mellifera	38999
Bombus impatiens	30432
Danaus plexippus	19791
Bombus griseocollis	12086
Bombus bimaculatus	7527
Xylocopa virginica	7205
Pieris rapae	6942
Bombus vosnesenskii	6639
Pyrobombus	6152
Bombus melanopygus	6016
Vanessa cardui	5997
Halictus ligatus	5565
Hylephila phyleus	5103
Eristalis tenax	4597
Papilio glaucus	4496

sourceTaxonName	count
Bombus ternarius	4453
Toxomerus marginatus	4360
Sphex ichneumoneus	4208
Epargyreus clarus	4141

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

targetTaxonName	count
Solidago	28558
Asclepias syriaca	12128
Quercus lobata	10234
Pycnanthemum muticum	8844
Quercus agrifolia	8396
Echinacea purpurea	7982
Buddleja davidii	7756
Symphyotrichum	7499
Monarda fistulosa	7233
Taraxacum	7126
Bidens alba	6470
Daucus carota	6239
Lantana	6180
Quercus	6026
Trifolium pratense	5964
Asclepias incarnata	5801
Actinopterygii	5368
Quercus alba	5227
Salix	5141

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

sourceTaxonName	interactionTypeName	targetTaxonName	count
Bombus impatiens	visitsFlowersOf	Solidago	2135
Danaus plexippus	interactsWith	Asclepias syriaca	1235
Megacyllene robiniae	visitsFlowersOf	Solidago	1223
Murgantia histrionica	interactsWith	Cleomella arborea	1194

sourceTaxonName	interactionTypeName	targetTaxonName	count
Bombus griseocollis	visitsFlowersOf	Echinacea purpurea	1149
Disholcaspis quercusmamma	interactsWith	Quercus macrocarpa	1027
Bombus impatiens	visitsFlowersOf	Symphyotrichum	1006
Polistes fuscatus	visitsFlowersOf	Solidago	1001
Eristalis tenax	visitsFlowersOf	Chrysanthemum	925
Apis mellifera	visitsFlowersOf	Solidago	872
Feron kingi	interactsWith	Quercus lobata	841
Andricus quercus-californicus	hasHost	Quercus lobata	836
Andricus quercus-californicus	interactsWith	Quercus lobata	836
Danaus plexippus	visitsFlowersOf	Buddleja davidii	834
Feron kingi	hasHost	Quercus lobata	793
Cynips douglasii	interactsWith	Quercus lobata	769
Apis mellifera	visitsFlowersOf	Taraxacum	752
Apis mellifera	visitsFlowersOf	Trifolium repens	738
Eurosta solidaginis	interactsWith	Solidago	737

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.

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

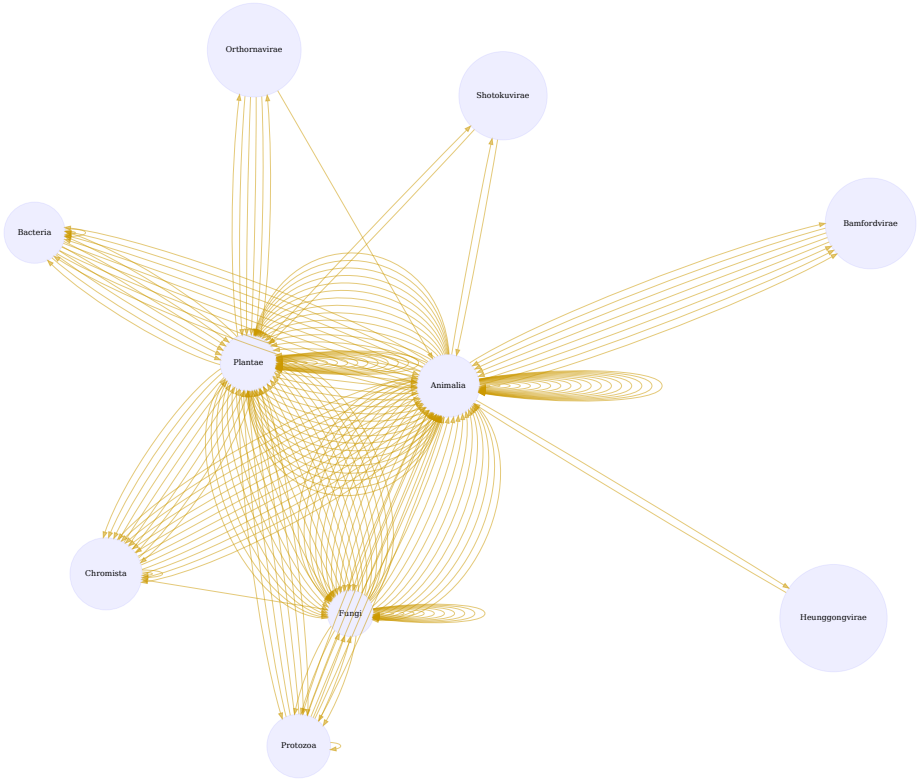

Figure 3: Interactions on taxonomic kingdom rank as interpreted by the Catalogue of Life download svg

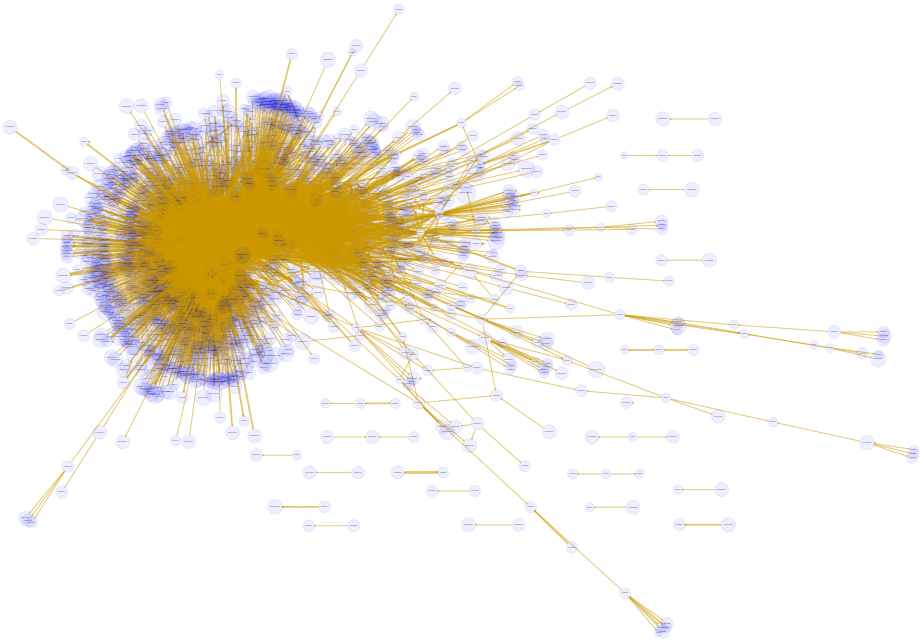

Figure 4: Interactions on the taxonomic family rank as interpreted by the Catalogue of Life. download svg

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
Degeeriella fulva	HAS_ACCEPTED_NAME	col	Degeeriella fulva
Siphonolejeunea raharahane-hemiae	HAS_ACCEPTED_NAME	col	Siphonolejeunea raharahane-hemiae
Neuroterus christi	NONE	col	Neuroterus christi
Melikaiella flora	NONE	col	Melikaiella flora

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	11130
col	class	47
col	epifamily	1
col	family	736
col	form	2
col	genus	4708
col	gigaclass	1
col	infraclass	2
col	infraorder	10
col	infraspecific name	26
col	kingdom	5
col	order	175
col	phylum	20
col	section	3
col	species	46031
col	subclass	7
col	subfamily	222
col	subgenus	110
col	suborder	19
col	subphylum	2
col	subspecies	2374
col	subterclass	3
col	subtribe	15
col	superfamily	64
col	superorder	7
col	supertribe	1
col	tribe	94
col	unranked	1
col	variety	158
discoverlife	NA	63205
discoverlife	species	1959
gbif	NA	4563
gbif	class	49
gbif	family	767
gbif	form	65
gbif	genus	5052
gbif	kingdom	6
gbif	order	174
gbif	phylum	19

resolvedCatalogName	resolvedRank	count
gbif	species	51064
gbif	subspecies	3804
gbif	variety	792
itis	NA	32048
itis	class	47
itis	division	8
itis	family	727
itis	genus	3686
itis	infraclass	1
itis	infrakingdom	1
itis	infraorder	17
itis	kingdom	5
itis	order	172
itis	phylum	11
itis	species	26281
itis	subclass	17
itis	subdivision	3
itis	subfamily	198
itis	subgenus	9
itis	subkingdom	1
itis	suborder	50
itis	subphylum	6
itis	subspecies	1677
itis	subtribe	9
itis	superclass	1
itis	superfamily	53
itis	superorder	6
itis	tribe	80
itis	variety	86
mdd	NA	65164
ncbi	NA	17195
ncbi	biotype	1
ncbi	clade	23
ncbi	class	42
ncbi	cohort	1
ncbi	family	730
ncbi	forma	2
ncbi	forma specialis	1
ncbi	genus	4686
ncbi	infraclass	1
ncbi	infraorder	23
ncbi	kingdom	2
ncbi	order	176
ncbi	parvorder	2

resolvedCatalogName	resolvedRank	count
ncbi	phylum	17
ncbi	section	1
ncbi	series	1
ncbi	species	40386
ncbi	species group	2
ncbi	subclass	14
ncbi	subfamily	294
ncbi	subgenus	123
ncbi	suborder	34
ncbi	subphylum	6
ncbi	subspecies	1163
ncbi	subtribe	40
ncbi	superclass	2
ncbi	superfamily	69
ncbi	superkingdom	2
ncbi	superorder	10
ncbi	tribe	171
ncbi	varietas	29
pbdb	NA	59274
pbdb	class	40
pbdb	family	676
pbdb	genus	1779
pbdb	infraclass	4
pbdb	infraorder	26
pbdb	kingdom	4
pbdb	order	177
pbdb	phylum	17
pbdb	species	2682
pbdb	subclass	16
pbdb	subfamily	229
pbdb	subgenus	1
pbdb	suborder	39
pbdb	subphylum	4
pbdb	subspecies	34
pbdb	subtribe	9
pbdb	superclass	3
pbdb	superfamily	79
pbdb	superorder	7
pbdb	superphylum	1
pbdb	tribe	92
pbdb	unranked clade	60
tpt	NA	61141
tpt	family	84
tpt	genus	211

resolvedCatalogName	resolvedRank	count
tpt	order	26
tpt	species	3700
tpt	specificepithet	1
tpt	subspecificepithet	3
wfo	NA	48074
wfo	class	3
wfo	family	162
wfo	form	6
wfo	genus	2066
wfo	order	47
wfo	section	2
wfo	species	14626
wfo	subfamily	15
wfo	subsection	1
wfo	subspecies	191
wfo	subtribe	21
wfo	supertribe	1
wfo	tribe	23
wfo	variety	104
worms	NA	54794
worms	class	42
worms	family	584
worms	genus	1996
worms	gigaclass	1
worms	infraclass	4
worms	infraorder	9
worms	infraphylum	1
worms	kingdom	6
worms	order	166
worms	phylum	11
worms	phylum (division)	8
worms	section	2
worms	species	7268
worms	subclass	16
worms	subfamily	46
worms	subgenus	1
worms	suborder	22
worms	subphylum	4
worms	subphylum (subdivision)	2
worms	subspecies	113
worms	subterclass	3
worms	subtribe	29
worms	superclass	1
worms	superfamily	33

resolvedCatalogName	resolvedRank	count
worms	superorder	10
worms	tribe	22
worms	variety	29

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	53191
col	NONE	11291
col	SYNONYM_OF	9131
discoverlife	NONE	63873
discoverlife	HAS_ACCEPTED_NAME	1918
discoverlife	SYNONYM_OF	402
discoverlife	HOMONYM_OF	83
gbif	HAS_ACCEPTED_NAME	63225
gbif	SYNONYM_OF	16965
gbif	NONE	4702
itis	NONE	32292
itis	HAS_ACCEPTED_NAME	32575
itis	SYNONYM_OF	1677
mdd	NONE	64673
mdd	HAS_ACCEPTED_NAME	1166
mdd	SYNONYM_OF	11
ncbi	SAME_AS	47371
ncbi	NONE	17361
ncbi	SYNONYM_OF	1438
pbdb	NONE	59794
pbdb	HAS_ACCEPTED_NAME	5969
pbdb	SYNONYM_OF	429
tpt	HAS_ACCEPTED_NAME	4197
tpt	NONE	61776
tpt	SYNONYM_OF	216
wfo	NONE	48533
wfo	HAS_UNCHECKED_NAME	1418
wfo	HAS_ACCEPTED_NAME	16554
wfo	SYNONYM_OF	2359

resolvedCatalogName	relationName	count
worms	NONE	55293
worms	HAS_ACCEPTED_NAME	10808
worms	SYNONYM_OF	929

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
2025-06-03T03:12:15Z	summary	https://github.com/globalbioticinteractions/inaturalist/
2025-06-03T03:12:15Z	summary	1082852 interaction(s)

reviewDate	reviewCommentType	reviewComment
2025-06-03T03:12:15Z	summary	0 note(s)
2025-06-03T03:12:15Z	summary	8 info(s)

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

: Most frequently occurring review notes, if any.

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.

[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 (2025-06-03) the version of the GloBI dataset index was available at https://globalbioticinteractions.org/datasets

Discussion

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

This review aims to provide a perspective on the dataset to aid in understanding of species interaction claims discovered. However, it is important to note that this review does *not* assess the quality of the dataset. Instead, it serves as an indication of the open-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.

[6] According 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."

References

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

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

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

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

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

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

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

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

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

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