Versioned Archive and Review of Biotic Interactions and Taxon Names Found within globalbioticinteractions/fishbase hash://md5/ace2d06b6d8f88a83b8dbd8e0df34a65

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

2025-09-01

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/fishbase, has fingerprint hash://md5/ace2d06b6d8f88a83b8dbd8e0df34a65, is 237MiB in size and contains 193,761 interactions with 2 unique types of associations (e.g., eats) between 8,813 primary taxa (e.g., Gadus morhua) and 17,175 associated taxa (e.g., benth. crust.). This report includes detailed summaries of interaction data, a taxonomic review from multiple catalogs, and an archived version of the dataset from which the reviews are derived.

Contents

Introduction

Data Review and Archive

Data review and archiving can be a time-consuming process, especially when done manually. This review report aims to help facilitate both activities. It automates the archiving of datasets, including Darwin Core archives, and is a citable backup of a version of the dataset. Additionally, an automatic review of species interaction claims made in the dataset is generated and registered with Global Biotic Interactions (J. H. Poelen, Simons, and Mungall 2014).

This review includes summary statistics about, and observations about, the dataset under review :

> Froese, R. and D. Pauly. Editors. 2018. FishBase. World Wide Web electronic publication. www.fishbase.org, version (10/2018). https://github.com/globalbioticinteractions/fishbase/archive/ffea863a41bd29f9677aa33d5e4733484d2208 2025-08-29T23:47:40.630Z hash://md5/ace2d06b6d8f88a83b8dbd8e0df34a65

For additional metadata related to this dataset, please visit https://github.com/globalbioticinteractions/fishbase 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.  237MiB) under review
elton pull globalbioticinteractions/fishbase

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

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

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

or visually, in a process diagram.

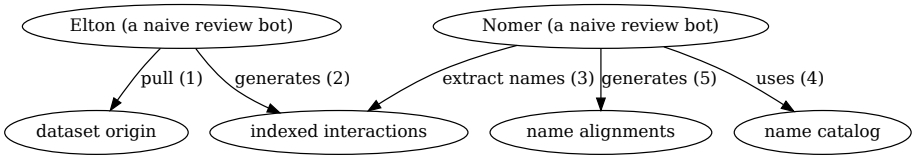

Figure 1: Review Process Overview

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

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.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/fishbase, has finger-

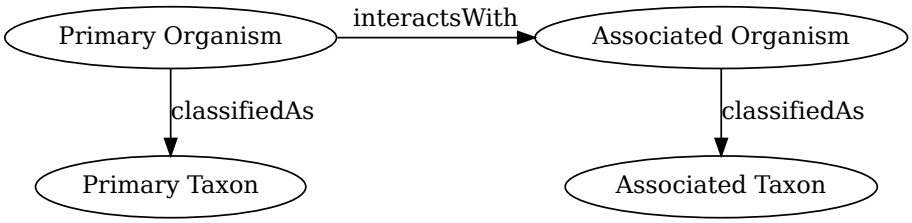

Figure 2: Biotic Interaction Data Model

print hash://md5/ace2d06b6d8f88a83b8dbd8e0df34a65, is 237MiB in size and contains 193,761 interactions with 2 unique types of associations (e.g., eats) between 8,813 primary taxa (e.g., Gadus morhua) and 17,175 associated taxa (e.g., benth. crust.).

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 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
Gadus morhua	eats	gastropods	Armstrong, M J. 1982. The predator-prey relationships of Irish Sea poor-cod (Trisopterus minutus L.), pouting (Trisopterus luscus L), and cod (Gadus morhua L.). J. Cons. Int. Explor. Mer. 40(2):135-152.

sourceTaxonName	interactionTypeName	targetTaxonName	referenceCitation
Gadus morhua	eats	Trisopterus minutus	Armstrong, M J. 1982. The predator-prey relationships of Irish Sea poor-cod (Trisopterus minutus L.), pouting (Trisopterus luscus L), and cod (Gadus morhua L.). J. Cons. Int. Explor. Mer. 40(2):135-152.
Gadus morhua	eats	Galatheidae	Armstrong, M J. 1982. The predator-prey relationships of Irish Sea poor-cod (Trisopterus minutus L.), pouting (Trisopterus luscus L), and cod (Gadus morhua L.). J. Cons. Int. Explor. Mer. 40(2):135-152.

sourceTaxonName	interactionTypeName	targetTaxonName	referenceCitation
Gadus morhua	eats	polychaetes	Armstrong, M J. 1982. The predator-prey relationships of Irish Sea poor-cod (Trisopterus minutus L.), pouting (Trisopterus luscus L), and cod (Gadus morhua L.). J. Cons. Int. Explor. Mer. 40(2):135-152.

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

interactionTypeName	count
eats	187460
preysOn	6301

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

sourceTaxonName	count
Gadus morhua	2294
Melanogrammus aeglefinus	2066
Merlangius merlangus	1754
	1522
Scomber scombrus	855
Pollachius virens	777
Eutrigla gurnardus	750
Gadus macrocephalus	750
Amblyraja radiata	722
Salmo trutta	699
Raja clavata	669
Clupea harengus	657

sourceTaxonName	count
Hippoglossoides platessoides	599
Merluccius merluccius	580
Squalus acanthias	570
Gadus chalcogrammus	556
Boreogadus saida	488
Morone saxatilis	482
Trisopterus minutus	482

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

targetTaxonName	count
benth. crust.	23079
finfish	16346
unidentified	13511
plank. crust.	10807
insects	7982
mollusks	6594
worms	5507
other plants	4746
phytoplankton	3198
other plank. invertebrates	2911
detritus	2539
cephalopods	2527
other benth. invertebrates	2504
echinoderms	1760
bony fish	1347
fish (early stages)	1087
polychaetes	1005
benth. crustaceans	819
amphipods	801

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

sourceTaxonName	interactionTypeName	targetTaxonName	count
Gadus morhua	eats	finfish	319
Gadus morhua	eats	benth. crust.	319
Melanogrammus aeglefinus	eats	benth. crust.	283
Merlangius merlangus	eats	benth. crust.	242

sourceTaxonName	interactionTypeName	targetTaxonName	count
Salmo trutta	eats	insects	217
Clupea harengus	eats	plank. crust.	213
Melanogrammus aeglefinus	eats	mollusks	204
Sprattus sprattus	eats	plank. crust.	184
Raja clavata	eats	benth. crust.	167
Hiodon alosoides	eats	insects	161
Micromesistius poutassou	eats	plank. crust.	157
Merlangius merlangus	eats	finfish	154
Gadus chalcogrammus	eats	plank. crust.	152
Melanogrammus aeglefinus	eats	worms	147
Trisopterus minutus	eats	benth. crust.	142
Gadus morhua	eats	worms	138
Eutrigla gurnardus	eats	benth. crust.	137
Melanogrammus aeglefinus	eats	echinoderms	135
Boreogadus saida	eats	plank. crust.	133

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 alignments can help review name usage or aid in selecting of a suitable taxonomic name resource.

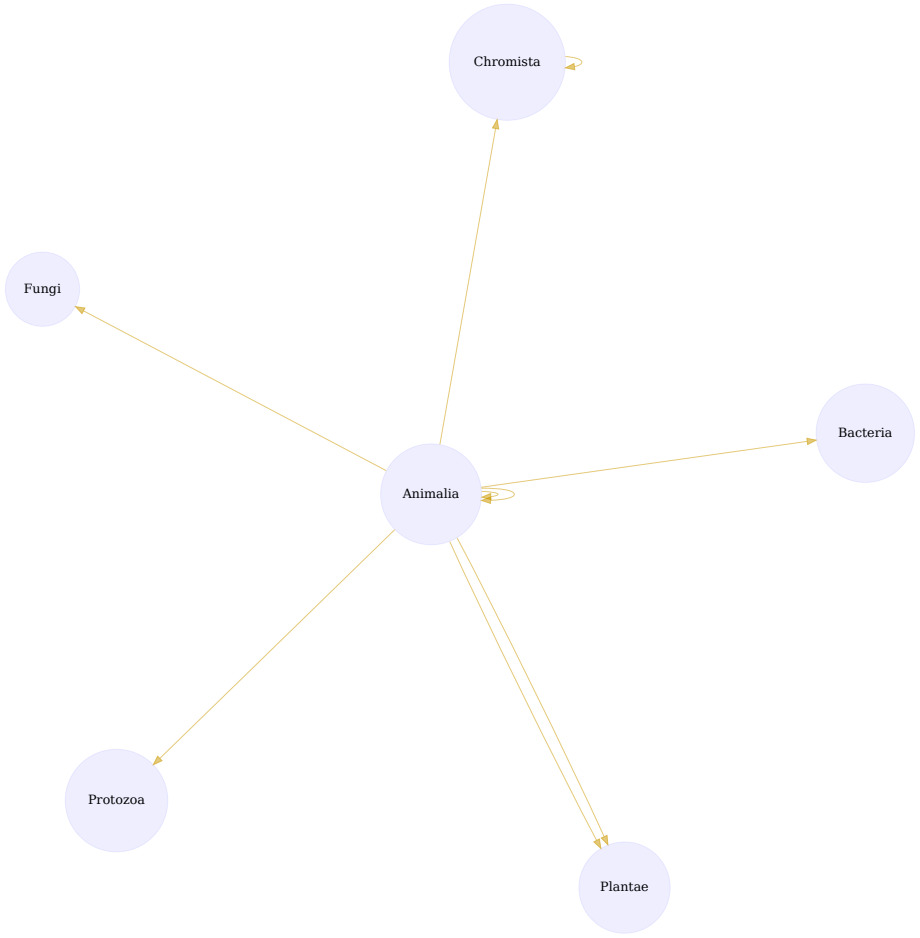

Figure 3: Interactions on taxonomic kingdom rank as interpreted by the Catalogue of Life download svg

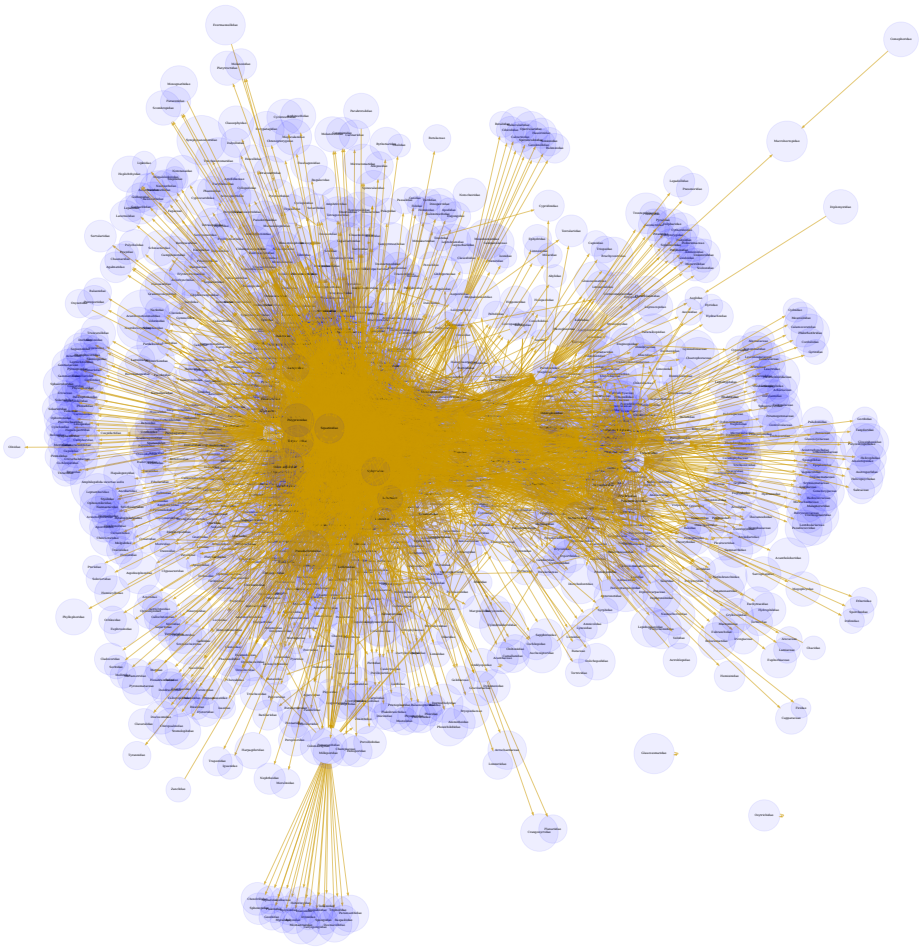

Figure 4: Interactions on the taxonomic family rank as interpreted by the Catalogue of Life. download svg

Table 8: Sample of Name Alignments

providedName	relationName	resolvedCatalogName	resolvedName
Organic material	NONE	col	Organic material
Aaptos aaptos	HAS_ACCEPTED_NAME	col	Aaptos aaptos
Ablabesmyia	HAS_ACCEPTED_NAME	col	Ablabesmyia
Ablennes hians	HAS_ACCEPTED_NAME	col	Ablennes hians

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	1980
col	class	49
col	family	504
col	form	1
col	genus	2093
col	gigaclass	1
col	infraorder	10
col	infraphylum	2
col	infraspecific name	35
col	kingdom	5
col	megaclass	1
col	order	91
col	other	2
col	parvorder	1
col	parvphylum	2
col	phylum	29
col	section	1
col	species	12599
col	subclass	10
col	subfamily	4
col	subgenus	53
col	suborder	22
col	subphylum	5
col	subspecies	60
col	superfamily	18
col	superorder	2
col	tribe	4
discoverlife	NA	17465
gbif	NA	1309

resolvedCatalogName	resolvedRank	count
gbif	class	50
gbif	family	545
gbif	form	8
gbif	genus	2277
gbif	kingdom	5
gbif	order	84
gbif	phylum	30
gbif	species	13181
gbif	subspecies	57
gbif	variety	19
itis	NA	3115
itis	class	54
itis	division	9
itis	family	514
itis	form	1
itis	genus	2027
itis	infraclass	1
itis	infrakingdom	1
itis	infraorder	13
itis	infraphylum	3
itis	kingdom	5
itis	order	105
itis	phylum	26
itis	species	11444
itis	subclass	25
itis	subfamily	10
itis	suborder	37
itis	subphylum	11
itis	subspecies	52
itis	superclass	4
itis	superdivision	1
itis	superfamily	14
itis	superorder	6
itis	tribe	4
itis	variety	3
mdd	NA	17464
ncbi	NA	3964
ncbi	clade	18
ncbi	class	53
ncbi	cohort	1
ncbi	family	502
ncbi	genus	2016
ncbi	infraclass	1
ncbi	infraorder	10

resolvedCatalogName	resolvedRank	count
ncbi	kingdom	2
ncbi	order	92
ncbi	parvorder	1
ncbi	phylum	32
ncbi	series	1
ncbi	species	10643
ncbi	subclass	18
ncbi	subfamily	7
ncbi	subgenus	17
ncbi	suborder	22
ncbi	subphylum	5
ncbi	subspecies	62
ncbi	superclass	2
ncbi	superfamily	17
ncbi	superkingdom	1
ncbi	superorder	6
ncbi	tribe	4
ncbi	varietas	2
pbdb	NA	14663
pbdb	class	60
pbdb	family	423
pbdb	genus	1078
pbdb	informal	1
pbdb	infraclass	2
pbdb	infraorder	8
pbdb	kingdom	4
pbdb	order	101
pbdb	phylum	36
pbdb	species	995
pbdb	subclass	19
pbdb	subfamily	11
pbdb	subgenus	2
pbdb	subkingdom	3
pbdb	suborder	37
pbdb	subphylum	7
pbdb	subspecies	5
pbdb	superclass	5
pbdb	superfamily	17
pbdb	superorder	7
pbdb	superphylum	2
pbdb	tribe	4
pbdb	unranked clade	28
tpt	NA	17418
tpt	family	2

resolvedCatalogName	resolvedRank	count
tpt	genus	7
tpt	order	2
tpt	species	35
wfo	NA	17248
wfo	family	9
wfo	genus	134
wfo	order	1
wfo	phylum	1
wfo	species	71
worms	NA	1949
worms	class	59
worms	family	484
worms	forma	3
worms	genus	2040
worms	gigaclass	1
worms	infraclass	3
worms	infraorder	10
worms	infraphylum	3
worms	kingdom	6
worms	megaclass	1
worms	order	101
worms	parvphylum	2
worms	phylum	31
worms	phylum (division)	5
worms	section	1
worms	species	12676
worms	subclass	17
worms	subfamily	6
worms	subgenus	2
worms	suborder	32
worms	subphylum	11
worms	subspecies	20
worms	superclass	1
worms	superfamily	15
worms	superorder	5
worms	tribe	4
worms	variety	7

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	NONE	3302
col	HAS_ACCEPTED_NAME	21922
col	SYNONYM_OF	2943
discoverlife	NONE	24671
gbif	NONE	2715
gbif	HAS_ACCEPTED_NAME	23438
gbif	SYNONYM_OF	4023
itis	NONE	4255
itis	HAS_ACCEPTED_NAME	19320
itis	SYNONYM_OF	1343
mdd	NONE	24406
mdd	HAS_ACCEPTED_NAME	33
mdd	SYNONYM_OF	1
ncbi	NONE	5582
ncbi	SAME_AS	18260
ncbi	SYNONYM_OF	1062
ncbi	COMMON_NAME_OF	10
pbdb	NONE	17564
pbdb	HAS_ACCEPTED_NAME	7439
pbdb	SYNONYM_OF	1748
tpt	NONE	24189
tpt	HAS_ACCEPTED_NAME	248
tpt	SYNONYM_OF	3
wfo	NONE	24006
wfo	SYNONYM_OF	140
wfo	HAS_ACCEPTED_NAME	275
wfo	HAS_UNCHECKED_NAME	44
worms	NONE	3097
worms	HAS_ACCEPTED_NAME	19605
worms	SYNONYM_OF	2623

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
2025-09-01T19:42:44Z	note	target taxon name missing
2025-09-01T19:42:45Z	note	target taxon name missing
2025-09-01T19:42:50Z	summary	https://github.com/globalbioticinteractions/fishbase/ar
2025-09-01T19:42:50Z	summary	193761 interaction(s)

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
target taxon name missing	2

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

GloBI Review Badge

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

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

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

GloBI Index Badge

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

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

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

Discussion

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

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

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

[5] At time of writing (2025-09-01) the version of the GloBI dataset index was available at https://globalbioticinteractions.org/datasets

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

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

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

Acknowledgements

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

Author contributions

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

References

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

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

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

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

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

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

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

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

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

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

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