

Versioned Archive and Review of Biotic
Interactions and Taxon Names Found within
fonturbel-lab/pollination_catalogue
hash://md5/6f9ac7dbe67db53ae04d88355ecf2299

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

2025-07-14

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 fonturbel-lab/pollination_catalogue, has fingerprint hash://md5/6f9ac7dbe67db53ae04d88355ecf2299, is 1.91MiB in size and contains 2,619 interaction with 5 unique types of associations (e.g., pollinatedBy) between 356 primary taxa (e.g., *Alstroemeria ligta*) and 526 associated taxa (e.g., *Bombus dahlbomii*). This report includes detailed summaries of interaction data, a taxonomic review from multiple catalogs, and an archived version of the dataset from which the reviews are derived.

Contents

Introduction	2
Data Review and Archive	2
Methods	2
Results	4
Files	4
Archived Dataset	13

Biotic Interactions	13
Interaction Networks	18
Taxonomic Alignment	20
Additional Reviews	23
GloBI Review Badge	24
GloBI Index Badge	25
Discussion	25
Acknowledgements	26
Author contributions	26
References	26

Introduction

Data Review and Archive

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

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

Giselle Muschett & Francisco E. Fontúrbel. 2021. A comprehensive catalogue of plant – pollinator interactions for Chile
https://github.com/fonturbel-lab/pollination_catalogue/archive/8e987dcd5b67199fc54d46ca468f894ea382025-07-11T22:51:02.476Z hash://md5/6f9ac7dbe67db53ae04d88355ecf2299

For additional metadata related to this dataset, please visit https://github.com/fonturbel-lab/pollination_catalogue 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. 1.91MiB) under review
elton pull fonturbel-lab/pollination_catalogue

# generate review notes
elton review fonturbel-lab/pollination_catalogue\
> review.tsv

# export indexed interaction records
elton interactions fonturbel-lab/pollination_catalogue\
> interactions.tsv

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

or visually, in a process diagram.

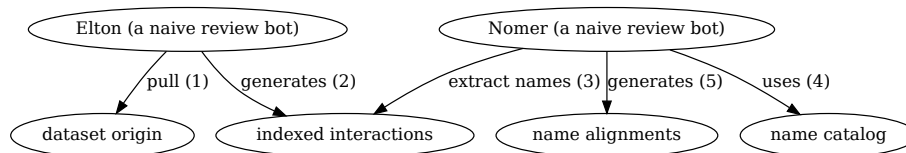


Figure 1: Review Process Overview

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

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

Results

In the following sections, the results of the review are summarized ². Then, links to the detailed review reports are provided.

Files

The following files are produced in this review:

filename	description
biblio.bib	list of bibliographic reference of this review
check-dataset.sh	data review workflow/process as expressed in a bash script
data.zip	a versioned archive of the data under review
HEAD	the digital signature of the data under review
index.docx	review in MS Word format
index.html	review in HTML format
index.md	review in Pandoc markdown format
index.pdf	review in PDF format
indexed-citations.csv.gz	list of distinct reference citations for reviewed species interaction claims in gzipped comma-separated values file format
indexed-citations.html.gz	list of distinct reference citations for reviewed species interactions claims in gzipped html file format
indexed-citations.tsv.gz	list of distinct reference citations for reviewed species interaction claims in gzipped tab-separated values format
indexed-interactions-col-family-col-family.svg	network diagram showing the taxon family to taxon family interaction claims in the dataset under review as interpreted by the Catalogue of Life via Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024)

²Disclaimer: The results in this review should be considered friendly, yet naive, notes from an unsophisticated robot. Please keep that in mind when considering the review results.

filename	description
indexed-interactions-col-kingdom-col-kingdom.svg	network diagram showing the taxon kingdom to taxon kingdom interaction claims in the dataset under review as interpreted by the Catalogue of Life via Nomen Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024)
indexed-interactions.csv.gz	species interaction claims indexed from the dataset under review in gzipped comma-separated values format
indexed-interactions.html.gz	species interaction claims indexed from the dataset under review in gzipped html format
indexed-interactions.tsv.gz	species interaction claims indexed from the dataset under review in gzipped tab-separated values format
indexed-interactions.parquet	species interaction claims indexed from the dataset under review in Apache Parquet format
indexed-interactions-sample.csv	list of species interaction claims indexed from the dataset under review in gzipped comma-separated values format
indexed-interactions-sample.html	first 500 species interaction claims indexed from the dataset under review in html format
indexed-interactions-sample.tsv	first 500 species interaction claims indexed from the dataset under review in tab-separated values format
indexed-names.csv.gz	taxonomic names indexed from the dataset under review in gzipped comma-separated values format
indexed-names.html.gz	taxonomic names found in the dataset under review in gzipped html format
indexed-names.tsv.gz	taxonomic names found in the dataset under review in gzipped tab-separated values format
indexed-names.parquet	taxonomic names found in the dataset under review in Apache Parquet format

filename	description
indexed-names-resolved-col.csv.gz	taxonomic names found in the dataset under review aligned with the Catalogue of Life as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped comma-separated values format
indexed-names-resolved-col.html.gz	taxonomic names found in the dataset under review aligned with the Catalogue of Life as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped html format
indexed-names-resolved-col.tsv.gz	taxonomic names found in the dataset under review aligned with the Catalogue of Life as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped tab-separated values format
indexed-names-resolved-col.parquet	taxonomic names found in the dataset under review aligned with the Catalogue of Life as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in Apache Parquet format
indexed-names-resolved-discoverlife.csv.gz	taxonomic names found in the dataset under review aligned with Discover Life bee species checklist as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped comma-separated values format
indexed-names-resolved-discoverlife.html.gz	taxonomic names found in the dataset under review aligned with Discover Life bee species checklist as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped html format

filename	description
indexed-names-resolved-discoverlife.tsv.gz	taxonomic names found in the dataset under review aligned with Discover Life bee species checklist as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped tab-separated values format
indexed-names-resolved-discoverlife.parquet	taxonomic names found in the dataset under review aligned with Discover Life bee species checklist as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in Apache Parquet format
indexed-names-resolved-gbif.csv.gz	taxonomic names found in the dataset under review aligned with GBIF Backbone Taxonomy as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped comma-separated values format
indexed-names-resolved-gbif.html.gz	taxonomic names found in the dataset under review aligned with GBIF Backbone Taxonomy as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped html format
indexed-names-resolved-gbif.tsv.gz	taxonomic names found in the dataset under review aligned with GBIF Backbone Taxonomy as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped tab-separated values format
indexed-names-resolved-gbif.parquet	taxonomic names found in the dataset under review aligned with GBIF Backbone Taxonomy as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in Apache Parquet format

filename	description
indexed-names-resolved-itis.csv.gz	taxonomic names found in the dataset under review aligned with Integrated Taxonomic Information System (ITIS) as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped comma-separated values format
indexed-names-resolved-itis.html.gz	taxonomic names found in the dataset under review aligned with Integrated Taxonomic Information System (ITIS) as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped html format
indexed-names-resolved-itis.tsv.gz	taxonomic names found in the dataset under review aligned with Integrated Taxonomic Information System (ITIS) as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped tab-separated values format
indexed-names-resolved-itis.parquet	taxonomic names found in the dataset under review aligned with Integrated Taxonomic Information System (ITIS) as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in Apache Parquet format
indexed-names-resolved-mdd.csv.gz	taxonomic names found in the dataset under review aligned with the Mammal Diversity Database as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped comma-separated values format
indexed-names-resolved-mdd.html.gz	taxonomic names found in the dataset under review aligned with Mammal Diversity Database as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped html format

filename	description
indexed-names-resolved-mdd.tsv.gz	taxonomic names found in the dataset under review aligned with Mammal Diversity Database as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped tab-separated values format
indexed-names-resolved-mdd.parquet	taxonomic names found in the dataset under review aligned with Mammal Diversity Database as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in Apache Parquet format
indexed-names-resolved-ncbi.csv.gz	taxonomic names found in the dataset under review aligned with the NCBI Taxonomy as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped comma-separated values format
indexed-names-resolved-ncbi.html.gz	taxonomic names found in the dataset under review aligned with the NCBI Taxonomy as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped html format
indexed-names-resolved-ncbi.tsv.gz	taxonomic names found in the dataset under review aligned with the NCBI Taxonomy as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped tab-separated values format
indexed-names-resolved-ncbi.parquet	taxonomic names found in the dataset under review aligned with the NCBI Taxonomy as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in Apache Parquet format

filename	description
indexed-names-resolved-pbdb.csv.gz	taxonomic names found in the dataset under review aligned with the Paleobiology Database as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped comma-separated values format
indexed-names-resolved-pbdb.html.gz	taxonomic names found in the dataset under review aligned with Paleobiology Database as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped html format
indexed-names-resolved-pbdb.tsv.gz	taxonomic names found in the dataset under review aligned with Paleobiology Database as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped tab-separated values format
indexed-names-resolved-pbdb.parquet	taxonomic names found in the dataset under review aligned with Paleobiology Database as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in Apache Parquet format
indexed-names-resolved-tpt.csv.gz	taxonomic names found in the dataset under review aligned with the Terrestrial Parasite Tracker (TPT) Taxonomic Resource as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped comma-separated values format
indexed-names-resolved-tpt.html.gz	taxonomic names found in the dataset under review aligned with the Terrestrial Parasite Tracker (TPT) Taxonomic Resource as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped html format

filename	description
indexed-names-resolved-tpt.tsv.gz	taxonomic names found in the dataset under review aligned with the Terrestrial Parasite Tracker (TPT) Taxonomic Resource as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped tab-separated values format
indexed-names-resolved-tpt.parquet	taxonomic names found in the dataset under review aligned with the Terrestrial Parasite Tracker (TPT) Taxonomic Resource as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in Apache Parquet format
indexed-names-resolved-wfo.csv.gz	taxonomic names found in the dataset under review aligned with the World of Flora Online as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped comma-separated values format
indexed-names-resolved-wfo.html.gz	taxonomic names found in the dataset under review aligned with the World of Flora Online as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped html format
indexed-names-resolved-wfo.tsv.gz	taxonomic names found in the dataset under review aligned with the World of Flora Online as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped tab-separated values format
indexed-names-resolved-wfo.parquet	taxonomic names found in the dataset under review aligned with the World of Flora Online as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in Apache Parquet format

filename	description
indexed-names-resolved-worms.csv.gz	taxonomic names found in the dataset under review aligned with the World Register of Marine Species (WoRMS) as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped comma-separated values format
indexed-names-resolved-worms.html.gz	taxonomic names found in the dataset under review aligned with the World Register of Marine Species (WoRMS) as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped html format
indexed-names-resolved-worms.tsv.gz	taxonomic names found in the dataset under review aligned with the World Register of Marine Species (WoRMS) as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped tab-separated values format
indexed-names-resolved-worms.parquet	taxonomic names found in the dataset under review aligned with the World Register of Marine Species (WoRMS) as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in Apache Parquet format
indexed-names-sample.csv	first 500 taxonomic names found in the dataset under review in comma-separated values format
indexed-names-sample.html	first 500 taxonomic names found in the dataset under review in html format
indexed-names-sample.tsv	first 500 taxonomic names found in the dataset under review in tab-separated values format
interaction.svg	diagram summarizing the data model used to index species interaction claims
nanopub-sample.trig	first 500 species interaction claims as expressed in the nanopub format (Kuhn and Dumontier 2014)

filename	description
nanopub.trig.gz	species interaction claims as expressed in the nanopub format (Kuhn and Dumontier 2014)
process.svg	diagram summarizing the data review processing workflow
prov.nq	origin of the dataset under review as expressed in rdf/nquads
review.csv.gz	review notes associated with the dataset under review in gzipped comma-separated values format
review.html.gz	review notes associated with the dataset under review in gzipped html format
review.tsv.gz	review notes associated with the dataset under review in gzipped tab-separated values format
review-sample.csv	first 500 review notes associated with the dataset under review in comma-separated values format
review-sample.html	first 500 review notes associated with the dataset under review in html format
review-sample.tsv	first 500 review notes associated with the dataset under review in tab-separated values format
review.svg	a review badge generated as part of the dataset review process
zenodo.json	metadata of this review expressed in Zenodo record metadata

Archived Dataset

Note that *data.zip* file in this archive contains the complete, unmodified archived dataset under review.

Biotic Interactions

In this review, biotic interactions (or biotic associations) are modeled as a primary (aka subject, source) organism interacting with an associate (aka object, target) organism. The dataset under review classified the primary/associate organisms with specific taxa. The primary and associate organisms The kind of interaction is documented as an interaction type.

The dataset under review, named fonturbel-lab/pollination_catalogue, has fin-

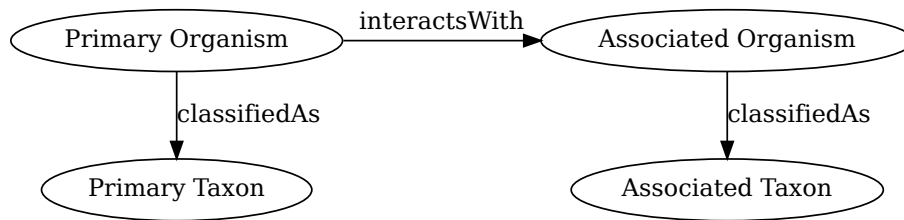


Figure 2: Biotic Interaction Data Model

gerprint hash://md5/6f9ac7dbe67db53ae04d88355ecf2299, is 1.91MiB in size and contains 2,619 interaction with 5 unique types of associations (e.g., pollinatedBy) between 356 primary taxa (e.g., *Alstroemeria ligtu*) and 526 associated taxa (e.g., *Bombus dahlbomii*).

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
Acrisione	pollinatedBy	Insecta	Aizen & Escurra. 1998. High incidence of plant-animal mutualisms in the woody flora of the temperate forest of southern South America: biogeographical origin and present ecological significance. <i>Ecologia Austral</i> , 8, 217-236

sourceTaxonName	interactionType	targetTaxonName	referenceCitation
Aextoxicon punctatum	pollinatedBy	Insecta	Aizen & Escurra. 1998. High incidence of plant-animal mutualisms in the woody flora of the temperate forest of southern South America: biogeographical origin and present ecological significance. Ecologia Austral, 8, 217-236
Aextoxicon punctatum	pollinatedBy	Insecta	Aizen & Escurra. 1998. High incidence of plant-animal mutualisms in the woody flora of the temperate forest of southern South America: biogeographical origin and present ecological significance. Ecologia Austral, 8, 217-236

sourceTaxonName	interactionTypeName	targetTaxonName	referenceCitation
Amomyrtus	pollinatedBy	Insecta	Aizen & Escurra. 1998. High incidence of plant-animal mutualisms in the woody flora of the temperate forest of southern South America: biogeographical origin and present ecological significance. Ecologia Austral, 8, 217-236

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

interactionTypeName	count
pollinatedBy	2245
flowersVisitedBy	362
mutualistOf	6
commensalistOf	5
eatenBy	1

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

sourceTaxonName	count
Alstroemeria ligtu	288
Magnoliopsida	287
Eucryphia cordifolia	169
Mimulus luteus	72
Viola portalesia	68
Embothrium coccineum	65
Echinopsis chiloensis	48
Tepualia stipularis	45
Amomyrtus meli	35
Schizanthus hookeri	34

sourceTaxonName	count
Escallonia myrtoidea	32
Myrceugenia ovata	31
Senecio bustillosianus	31
Chaetanthera lycopodioides	31
Chuquiraga oppositifolia	26
Stachys albicaulis	26
Luma apiculata	25
Phacelia secunda	24
Escallonia pulverulenta	23

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

targetTaxonName	count
Bombus dahlbomii	143
Sephanoides sephaniodes	105
Apis mellifera	91
Insecta	70
Bombus terrestris	62
Centris nigerrima	60
Scaeva melanostoma	58
Bombus ruderatus	56
Diptera	42
Megachile semirufa	40
Sarcophagidae	36
Megachile	35
Alloscirtetica gayi	34
Manuelia gayi	29
Scirtidae	27
Sephanoides fernandensis	24
Lasia corvina	24
Centris cineraria	23
Villa	22

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

sourceTaxonName	interactionTypeName	targetTaxonName	count
Alstroemeria ligtu	pollinatedBy	Centris nigerrima	29
Alstroemeria ligtu	pollinatedBy	Alloscirtetica gayi	26

sourceTaxonName	interactionTypeName	targetTaxonName	count
Alstroemeria ligtu	pollinatedBy	Manuelia gayi	24
Magnoliopsida	flowersVisitedBy	Bombus dahlbomii	20
Alstroemeria ligtu	pollinatedBy	Lasia corvina	17
Alstroemeria ligtu	pollinatedBy	Ruizantheda mutabilis	14
Alstroemeria ligtu	pollinatedBy	Ruizantheda proxima	14
Alstroemeria ligtu	pollinatedBy	Lasia aenea	13
Magnoliopsida	flowersVisitedBy	Bombus ruderatus	12
Alstroemeria ligtu	pollinatedBy	Svastrides melanura	12
Embothrium coccineum	pollinatedBy	Sephanoides sephaniodes	11
Alstroemeria ligtu	pollinatedBy	Epiclopus gayi	11
Eucryphia cordifolia	pollinatedBy	Sarcophagidae	11
Tristerix corymbosus	pollinatedBy	Sephanoides sephaniodes	10
Alstroemeria ligtu	pollinatedBy	Megachile cinerea	10
Alstroemeria ligtu	pollinatedBy	Tabanidae	10
Eucryphia cordifolia	pollinatedBy	Calliphoridae	10
Mimulus luteus	pollinatedBy	Bombus dahlbomii	10
Mimulus luteus	pollinatedBy	Centris nigerrima	9

Interaction Networks

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

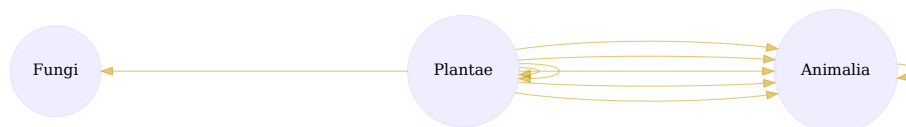


Figure 3: Interactions on taxonomic kingdom rank as interpreted by the Catalogue of Life download svg

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

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

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

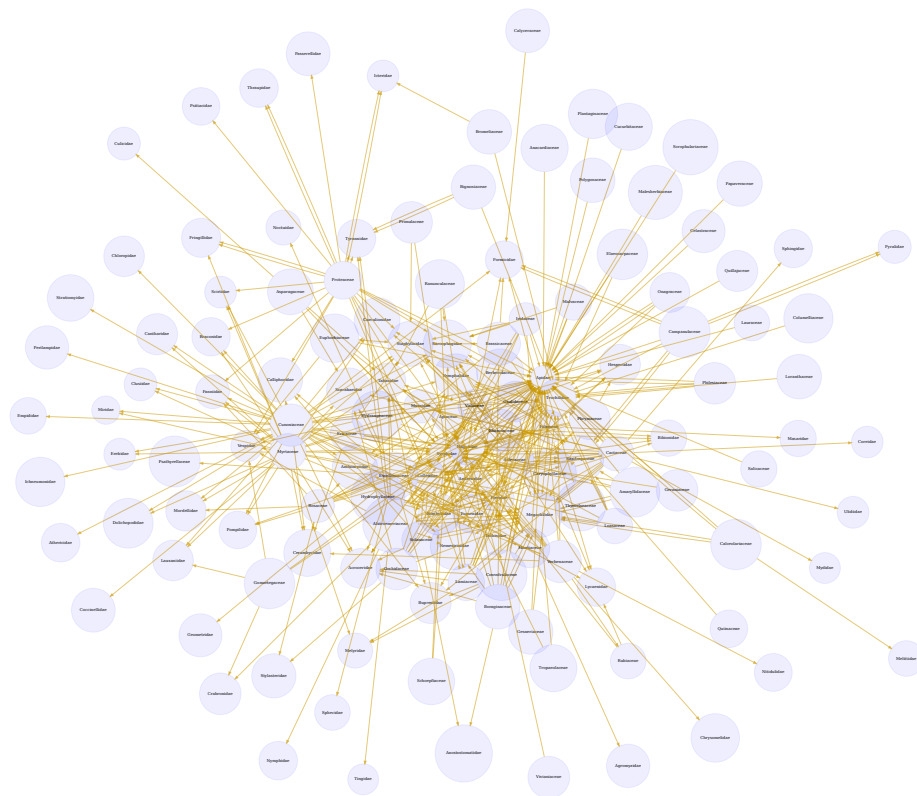


Figure 4: Interactions on the taxonomic family rank as interpreted by the Catalogue of Life. [download svg](#)

Taxonomic Alignment

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

Table 8: Sample of Name Alignments

providedName	relationName	resolvedCatalogName	resolvedName
Acacia caven	SYNONYM_OF	col	Vachellia caven
Acacia dealbata	HAS_ACCEPTED_NAME	col	Acacia dealbata
Acacia dealbata	SYNONYM_OF	col	Acacia leucolobia
Acaena pinnatifida	HAS_ACCEPTED_NAME	col	Acaena pinnatifida

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	107
col	class	2
col	family	45
col	genus	141
col	order	4
col	species	573
col	subfamily	4
col	subgenus	1
col	subspecies	4
col	superfamily	1
col	tribe	1
col	variety	4
discoverlife	NA	734
discoverlife	species	145
gbif	NA	67
gbif	class	2
gbif	family	49
gbif	genus	167
gbif	order	4
gbif	species	591
gbif	subspecies	8
gbif	variety	2
itis	NA	467

resolvedCatalogName	resolvedRank	count
itis	class	2
itis	family	44
itis	genus	82
itis	order	4
itis	species	273
itis	subfamily	5
itis	superfamily	1
itis	variety	1
mdd	NA	879
ncbi	NA	238
ncbi	class	2
ncbi	family	45
ncbi	genus	138
ncbi	order	4
ncbi	species	443
ncbi	subfamily	5
ncbi	subgenus	5
ncbi	superfamily	1
ncbi	varietas	1
pbdb	NA	766
pbdb	class	2
pbdb	family	48
pbdb	genus	49
pbdb	order	4
pbdb	species	7
pbdb	subfamily	3
pbdb	superfamily	1
tpt	NA	857
tpt	family	1
tpt	species	21
wfo	NA	528
wfo	family	1
wfo	genus	71
wfo	species	275
wfo	subspecies	2
wfo	variety	2
worms	NA	726
worms	class	1
worms	family	35
worms	genus	44
worms	order	4
worms	species	66
worms	subclass	1
worms	subfamily	1

resolvedCatalogName	resolvedRank	count
worms	subspecies	1

Table 10: Name relationship types per catalog. Name relationship type “NONE” means that a name was not recognized by the associated catalog. “SAME_AS” indicates either a “HAS_ACCEPTED_NAME” or “SYNONYM_OF” name relationship type. We recognize that “SYNONYM_OF” encompasses many types of nomenclatural synonymies (ICZN 1999) (e.g., junior synonym, senior synonyms).

resolvedCatalogName	relationName	count
col	SYNONYM_OF	169
col	HAS_ACCEPTED_NAME	891
col	NONE	128
discoverlife	NONE	888
discoverlife	HAS_ACCEPTED_NAME	171
discoverlife	SYNONYM_OF	42
discoverlife	HOMONYM_OF	3
gbif	SYNONYM_OF	272
gbif	HAS_ACCEPTED_NAME	1038
gbif	NONE	79
itis	SYNONYM_OF	25
itis	HAS_ACCEPTED_NAME	510
itis	NONE	553
mdd	NONE	1078
mdd	HAS_ACCEPTED_NAME	1
ncbi	SYNONYM_OF	40
ncbi	SAME_AS	780
ncbi	NONE	273
pbdb	NONE	927
pbdb	HAS_ACCEPTED_NAME	151
pbdb	SYNONYM_OF	4
tpt	NONE	1048
tpt	HAS_ACCEPTED_NAME	31
wfo	SYNONYM_OF	63
wfo	HAS_ACCEPTED_NAME	357
wfo	HAS_UNCHECKED_NAME	30
wfo	NONE	696
worms	NONE	885
worms	HAS_ACCEPTED_NAME	198
worms	SYNONYM_OF	24

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
2025-07-14T07:09:43Z	note	issue handling date range [1999-09-15/1990-12-06]: The end instant must be greater than the start instant

reviewDate	reviewCommentType	reviewComment
2025-07-14T07:09:43Z	note	issue handling date range [1999-09-15/1990-12-06]: The end instant must be greater than the start instant
2025-07-14T07:09:43Z	note	issue handling date range [1999-09-15/1990-12-06]: The end instant must be greater than the start instant
2025-07-14T07:09:43Z	note	issue handling date range [1999-09-15/1990-12-06]: The end instant must be greater than the start instant

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
issue handling date range [1999-09-15/1990-12-06]: The end instant must be greater than the start instant	8

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

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

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

GloBI Index Badge

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



Figure 6: Picture of a GloBI Index Badge ⁴

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

Discussion

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

This review aims to provide a perspective on the dataset to aid in understanding of species interaction claims discovered. However, it is important to note that this review does *not* assess the quality of the dataset. Instead, it serves as an indication of the open-ness⁶ and FAIRness (Wilkinson et al. 2016; Trekels et al. 2023) of the dataset: to perform this review, the data was likely openly available, **F**indable, **A**ccessible, **I**nteroperable and **R**eusable. The current Open-FAIR assessment is qualitative, and a more quantitative approach can be implemented with specified measurement units.

This report also showcases the reuse of machine-actionable (meta)data, something highly recommended by the FAIR Data Principles (Wilkinson et al. 2016). Making (meta)data machine-actionable enables more precise processing by computers, enabling even naive review bots like Nomer and Elton to interpret the

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

⁵At time of writing (2025-07-14) the version of the GloBI dataset index was available at <https://globalbioticinteractions.org/datasets>

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

data effectively. This capability is crucial for not just automating the generation of reports, but also for facilitating seamless data exchanges, promoting interoperability.

Acknowledgements

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

Author contributions

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

References

- Elliott, Michael, Jorrit Poelen, Icaro Alzuru, Emilio Berti, and partha04patel. 2025. “Bio-Guoda/Preston: 0.10.5.” Zenodo. <https://doi.org/10.5281/zenodo.14662206>.
- ICZN. 1999. “International Code of Zoological Nomenclature.” The International Trust for Zoological Nomenclature, London, UK. <https://www.iczn.org/the-code/the-code-online/>.
- Kuhn, Tobias, and Michel Dumontier. 2014. “Trusty URIs: Verifiable, Immutable, and Permanent Digital Artifacts for Linked Data.” In *The Semantic Web: Trends and Challenges*, edited by Valentina Presutti, Claudia d’Amato, Fabien Gandon, Mathieu d’Aquin, Steffen Staab, and Anna Tordai, 395–410. Cham: Springer International Publishing.
- 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.10647>

565.

- 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>.