Versioned Archive and Review of Biotic Interactions and Taxon Names Found within globalbioticinteractions/paz2021 hash://md5/d18421bca11d5b74b2716a86b76b6585

by Nomer, Elton and Preston, three naive review bots
review@globalbioticinteractions.org
https://globalbioticinteractions.org/contribute
https://github.com/globalbioticinteractions/paz2021/issues

2026-02-02

Abstract

Life on Earth is sustained by complex interactions between organisms and their environment. These biotic interactions can be captured in datasets and published digitally. We present a review and archiving process for such an openly accessible digital interactions dataset of known origin and discuss its outcome. The dataset under review, named globalbioticinteractions/paz2021, has fingerprint hash://md5/d18421bca11d5b74b2716a86b76b6585, is 104KiB in size and contains 239 interactions with 2 unique types of associations (e.g., eatenBy) between 188 primary taxa (e.g., Acrocomia sclerocarpa Mart.) and 29 associated taxa (e.g., Homo sapiens). This report includes detailed summaries of interaction data, a taxonomic review from multiple catalogs, and an archived version of the dataset from which the reviews are derived.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 14

Biotic Interactions 14
Interaction Networks 22
Geospatial Distribution 22
Taxonomic Alignment 23
Additional Reviews 26
GloBI Review Badge 26
GloBI Index Badge 27

Discussion **27**

Acknowledgements **28**

Author contributions **28**

References **28**

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 :

> Fabricia Sousa Paz, Carlos Eduardo Pinto, Rafael Melo de Brito, Vera Lucia Imperatriz-Fonseca, Tereza Cristina Giannini, Edible Fruit Plant Species in the Amazon Forest Rely Mostly on Bees and Beetles as Pollinators, Journal of Economic Entomology, Volume 114, Issue 2, April 2021, Pages 710–722, https://doi.org/10.1093/jee/toaa284 https://github.com/globalbioticinteractions/paz2021/archive/6c73c0294fc6d8fa1da417c7caaf11c2445b03 2026-01-31T04:45:58.194Z hash://md5/d18421bca11d5b74b2716a86b76b6585

For additional metadata related to this dataset, please visit https://github.com/globalbioticinteractions/paz2021 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	0.11.1
elton	0.16.4
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
mapserver	7.6.4

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

```
# get versioned copy of the dataset (size approx.  104KiB) under review
elton pull globalbioticinteractions/paz2021

# generate review notes
elton review globalbioticinteractions/paz2021\
 > review.tsv

# export indexed interaction records
elton interactions globalbioticinteractions/paz2021\
 > interactions.tsv

# export names and align them with the Catalogue of Life using Nomer
elton names globalbioticinteractions/paz2021\
 | nomer append col\
 > name-alignment.tsv
```

or visually, in a process diagram.

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

[1]Note that you have to first get the data (e.g., via elton pull globalbioticinteractions/paz2021) before being able to generate reviews (e.g., elton review globalbioticinteractions/paz2021), extract interaction claims (e.g., elton interactions globalbioticinteractions/paz2021), or list taxonomic names (e.g., elton names globalbioticinteractions/paz2021)

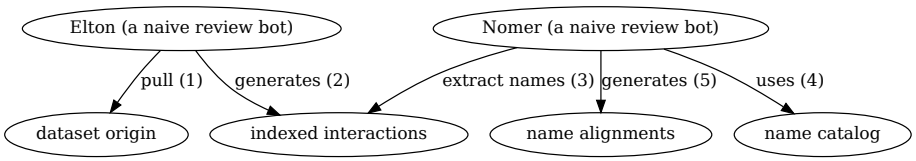


Figure 1: Review Process Overview

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

[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-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)
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.png	species interaction claims indexed from the dataset under review plotted on a map
indexed-interactions.gpkg	species interaction claims indexed from the dataset under review in GeoPackage format
indexed-interactions-h3.gpkg	geospatially clustered h3 species interaction claims indexed from the dataset under review in GeoPackage 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

filename	description
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
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

filename	description
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-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

filename	description
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
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

filename	description
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
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

filename	description
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
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

filename	description
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-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

filename	description
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
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

filename	description
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)
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

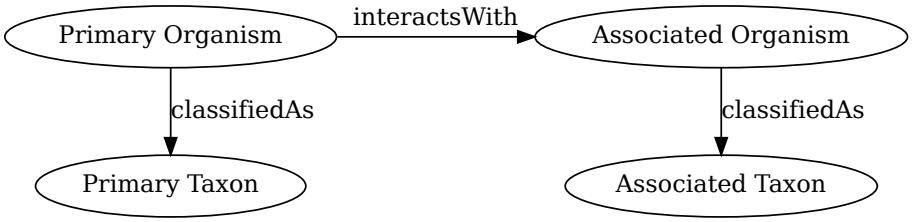


Figure 2: Biotic Interaction Data Model

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/paz2021, has fingerprint hash://md5/d18421bca11d5b74b2716a86b76b6585, is 104KiB in size and contains 239 interactions with 2 unique types of associations (e.g., eatenBy) between 188 primary taxa (e.g., Acrocomia sclerocarpa Mart.) and 29 associated taxa (e.g., Homo sapiens).

An exhaustive list of indexed interaction claims can be found in gzipped csv, tsv, geopackage 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
Acrocomia sclerocarpa Mart.	pollinatedBy	no species quoted	LORENZI, G. M. A. C.; NEGRELLE, R. R. B. 2006. Acrocomia aculeata (Jacq.) Lodd. ex Mart.: aspectos ecológicos. Visão Acadêmica 7. 10.5380/acd.v7i1.9021.
Allantoma lineata (Mart. & Berg) Miers	pollinatedBy	no species quoted	MORI, S. A. 1987. Biologia da polinização em Lecythi-daceae. Acta Bot. Bras. 1:121-124.
Anacardium giganteum Hanc. Ex Engl.	pollinatedBy	Apis mellifera	TAKEHANA, C.L.I. et al. 2013. Biologia floral e visitantes florais de Anacardium giganteum W. Hancock ex Engl. (Anacardiaceae) no município de Bragança, Pará. Revista de Ciências Agrárias 56:202-211.

sourceTaxonName	interactionTypeName	targetTaxonName	referenceCitation
Anacardium occidentale L.	pollinatedBy	Centris, Megachile	FREITAS, B.M.; CRUZ, D.O. 2010. As abelhas na polinização de culturas agricolas. In. Embrapa. II Semana dos Polinizadores: Palestras e Resumos. Embrapa, Petrolina, p. 140-156.

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

interactionTypeName	count
eatenBy	188
pollinatedBy	51

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

sourceTaxonName	count
Acrocomia sclerocarpa Mart.	2
Allantoma lineata (Mart. & Berg) Miers	2
Anacardium giganteum Hanc. Ex Engl.	2
Anacardium occidentale L.	2
Annona densicoma Mart.	2
Annona montana Macf.	2
Annona squamosa L.	2
Astrocaryum vulgare Mart.	2
Averrhoa carambola L.	2
Bactris maraja Mart.	2
Bertholletia excelsa Bonpland	2
Bixa orellana L.	2
Byrsonima crassifolia (L.) Rich.	2
Carica papaya L.	2
Caryocar brasiliense Camb.	2

sourceTaxonName	count
Couepia longipendula Pilger	2
Cucumis melo L.	2
Eugenia brasiliensis Lam.	2
Eugenia stipitata McVaugh	2

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

targetTaxonName	count
Homo sapiens	188
no species quoted	20
Apis mellifera	4
Colopterus (Nitidulidae), Curculionidae, Chrysomelidae	2
Centris, Megachile	1
Trigona sp	1
Bombus brevivillus, B. transversalis, C. americana, C. carrikeri, C. denudans, C. ferruginea, C. similis, Epicharis affinis, Ep. conica, Ep. flava, Ep. rustica, Ep. umbraculata, Ep. zonata, Eufriesea flaviventris, Euf. purpurata, Eulaema cingulata, Eul. meriana, Eul. mocsaryi, Eul. nigrita, Xylocopa frontalis	1
Bombus transversalis, Epicharis affinis, Epicharis flava, Epicharis rustica, Eulaema cingulata, Eulaema meriana, Melipona fasciculata, Melipona melanoventer, Xylocopa aurulenta, Xylocopa frontalis	1

targetTaxonName	count
Apis melífera, Augochloropsis crassigena, Centris bicolor, Centris byrsonimae, Centris caxienses, Centris decolorata, Centris flavifrons, Centris fuscata, Centris longimana, Centris rhodoprocta, Centris spilopoda, Centris sponsa, Centris tarsata, Centris trigonoides, Dicranthidium arenarium, Epicharis analis, Epicharis bicolor, Epicharis flava, Epicharis umbraculata, Lophopedia pygmaea, Nannotrigona punctata, Paratetrapedia leucostoma, Paratetrapedia testacea, Plebeia fallax, Scaptotrigona postica, Tetragona beebei, Tetrapedia diversipes, Trigona fulviventris, Trigona fuscipennis, Trigona pallens, Trigonisca extrema, Trigonisca pediculana, Tropidopedia punctifrons, Xanthopedia globulosa	1
Apis mellifera, Eulaema mocsaryi, Ptilotrigona lurida	1
197 species quoted: 51 Coleoptera, 36 Diptera, 104 Hymenoptera (48 Apidae), 1 Hemiptera, 1 Neuroptera, 1 Lepidoptera	1
Sphingidae, Hesperiidae, Nymphalidae, Apidae (Euglossini)	1
Centris americana, Centris flavifrons, Centris longimana, Centris rhodoprocta, Centris terminata, Melipona paraensis, Tetragona clavipes	1
Palpada vinetorum, Belvosia bicincta, Musca domestica (Diptera), Apis mellifera (Apidae)	1
Ornidia obesa, Eristalis sp (Syrphidae)	1
Cyclocephala distincta, Pachymerus sp, Epitragini sp, Belopeus carmelitus	1
Glossophaga soricina, Anoura caudifer, Phyllostomus discolor	1

targetTaxonName	count
Apis mellifera, Exomalopsis auropilosa, Melipona compressipes fasciculata, Melipona fasciculata, M. seminigra, M. flavolineata, Nannotrigona punctata, Partamona sp., Partamona cupira, Trigona branneri, Trigona pallens, Trigona recursa	1
Cyclocephala marginalis, Cyclocephala sarpedon, others	1

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

sourceTaxonName	interactionTypeName	targetTaxonName	count
Acrocomia sclerocarpa Mart.	pollinatedBy	no species quoted	1
Allantoma lineata (Mart. & Berg) Miers	pollinatedBy	no species quoted	1
Anacardium giganteum Hanc. Ex Engl.	pollinatedBy	Apis mellifera	1
Anacardium occidentale L.	pollinatedBy	Centris, Megachile	1
Annona densicoma Mart.	pollinatedBy	Colopterus (Nitidulidae), Curculionidae, Chrysomelidae	1
Annona montana Macf.	pollinatedBy	Colopterus (Nitidulidae), Curculionidae, Chrysomelidae	1
Annona squamosa L.	pollinatedBy	no species quoted	1
Astrocaryum vulgare Mart.	pollinatedBy	Apis mellifera	1
Averrhoa carambola L.	pollinatedBy	Trigona sp	1
Bactris maraja Mart.	pollinatedBy	no species quoted	1

sourceTaxonName	interactionTypeName	targetTaxonName	count
Bertholletia excelsa Bonpland	pollinatedBy	Bombus brevivillus, B. transversalis, C. americana, C. carrikeri, C. denudans, C. ferruginea, C. similis, Epicharis affinis, Ep. conica, Ep. flava, Ep. rustica, Ep. umbraculata, Ep. zonata, Eufriesea flaviventris, Euf. purpurata, Eulaema cingulata, Eul. meriana, Eul. mocsaryi, Eul. nigrita, Xylocopa frontalis	1
Bixa orellana L.	pollinatedBy	Bombus transversalis, Epicharis affinis, Epicharis flava, Epicharis rustica, Eulaema cingulata, Eulaema meriana, Melipona fasciculata, Melipona melanoventer, Xylocopa aurulenta, Xylocopa frontalis	1

sourceTaxonName	interactionTypeName	targetTaxonName	count
Byrsonima crassifolia (L.) Rich.	pollinatedBy	Apis melífera, Augochloropsis crassigena, Centris bicolor, Centris byrsonimae, Centris caxienses, Centris decolorata, Centris flavifrons, Centris fuscata, Centris longimana, Centris rhodoprocta, Centris spilopoda, Centris sponsa, Centris tarsata, Centris trigonoides, Dicranthidium arenarium, Epicharis analis, Epicharis bicolor, Epicharis flava, Epicharis umbraculata, Lophopedia pygmaea, Nannotrigona punctata, Paratetrapedia leucostoma, Paratetrapedia testacea, Plebeia fallax, Scaptotrigona postica, Tetragona beebei, Tetrapedia diversipes, Trigona fulviventris, Trigona fuscipennis, Trigona pallens, Trigonisca extrema, Trigonisca pediculana, Tropidopedia punctifrons, Xanthopedia	1

sourceTaxonName	interactionTypeName	targetTaxonName	count
Carica papaya L.	pollinatedBy	no species quoted	1
Caryocar brasiliense Camb.	pollinatedBy	no species quoted	1
Couepia longipendula Pilger	pollinatedBy	no species quoted	1
Cucumis melo L.	pollinatedBy	Apis mellifera	1
Eugenia brasiliensis Lam.	pollinatedBy	no species quoted	1
Eugenia stipitata McVaugh	pollinatedBy	Apis mellifera, Eulaema mocsaryi, Ptilotrigona lurida	1

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

Figure 4: Interactions on the taxonomic family 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

Geospatial Distribution

If geospatial information was extracted from the dataset under review, the map below will show their distribution.

Associated data can be found in the geopackage files at indexed-interactions.gpkg for point data and indexed-interactions-h3.gpkg for data clustered in geospatial h3 hexagonals.

Hexagonal grid cells indicate that interactions claims are available for selected geospatial area: light yellow means relatively fewer claims, dark green relatively more claims.

Figure 5: Hexagonal grid cells indicate that interactions claims are available for selected geospatial area: light yellow means relatively fewer claims, dark green relatively more claims.

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

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

Taxonomic Alignment

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

Table 8: Sample of Name Alignments

providedName	relationName	resolvedCatalogName	resolvedName
Quoted	NONE	col	Quoted
Acrocomia sclerocarpa	SYNONYM_OF	col	Acrocomia aculeata
Agonandra brasiliensis	HAS_ACCEPTED_NAME	col	Agonandra brasiliensis
Aiphanes aculeata	SYNONYM_OF	col	Aiphanes horrida

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	30
col	family	3
col	genus	7
col	species	176
discoverlife	NA	215
discoverlife	species	1

resolvedCatalogName	resolvedRank	count
gbif	NA	27
gbif	family	3
gbif	form	1
gbif	genus	7
gbif	species	178
gbif	variety	1
itis	NA	147
itis	family	3
itis	genus	7
itis	species	59
mdd	NA	216
ncbi	NA	74
ncbi	family	3
ncbi	genus	7
ncbi	species	132
pbdb	NA	205
pbdb	family	3
pbdb	genus	4
pbdb	species	4
tpt	NA	214
tpt	family	1
tpt	species	1
wfo	NA	41
wfo	genus	3
wfo	species	172
worms	NA	198
worms	family	2
worms	genus	4
worms	species	12

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	31
col	SYNONYM_OF	58
col	HAS_ACCEPTED_NAME	157

resolvedCatalogName	relationName	count
discoverlife	NONE	216
discoverlife	HAS_ACCEPTED_NAME	1
gbif	NONE	28
gbif	SYNONYM_OF	99
gbif	HAS_ACCEPTED_NAME	184
itis	NONE	148
itis	SYNONYM_OF	5
itis	HAS_ACCEPTED_NAME	64
mdd	NONE	216
mdd	HAS_ACCEPTED_NAME	1
ncbi	NONE	75
ncbi	SAME_AS	133
ncbi	SYNONYM_OF	9
pbdb	NONE	206
pbdb	HAS_ACCEPTED_NAME	11
tpt	NONE	215
tpt	HAS_ACCEPTED_NAME	2
wfo	NONE	42
wfo	HAS_ACCEPTED_NAME	148
wfo	SYNONYM_OF	45
wfo	HAS_UNCHECKED_NAME	5
worms	NONE	199
worms	HAS_ACCEPTED_NAME	16
worms	SYNONYM_OF	3

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)

catalog name	alignment results
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
2026-02-02T12:20:24Z	note	target taxon name missing
2026-02-02T12:20:24Z	note	target taxon name missing
2026-02-02T12:20:24Z	note	target taxon name missing
2026-02-02T12:20:24Z	note	target taxon name missing

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	137
found [14] column definitions, but only [1] values: assuming undefined values are empty.	2
source 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 6: 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 7: 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-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

[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 (2026-02-02) the version of the GloBI dataset index was available at https://globalbioticinteractions.org/datasets

[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."

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.

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.