Versioned Archive and Review of Biotic Interactions and Taxon Names Found within urn:lsid:checklistbank.org:dataset:1061 hash://md5/75592983b876ce66f69ccea9cbe01d3b

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

2026-06-04

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 urn:lsid:checklistbank.org:dataset:1061, has fingerprint hash://md5/75592983b876ce66f69ccea9cbe01d3b, is 18.5MiB in size and contains 180 interactions with 1 unique type of association (e.g., eats) between 63 primary taxa (e.g., Aphis (Aphis) nasturtii) and 161 associated taxa (e.g., Cornus sericea). This report includes detailed summaries of interaction data, a taxonomic review from multiple catalogs, and an archived version of the dataset from which the reviews are derived.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 4
Biotic Interactions 4
Interaction Networks 9
Geospatial Distribution 10

Taxonomic Alignment 11
Additional Reviews 14
GloBI Review Badge 15
GloBI Index Badge 15

Discussion **15**

Acknowledgements **16**

Author contributions **16**

Appendix A. Review Files **16**

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 :

> (**misc?**){ChecklistBankDataset1061, publisher = {Species File Group}, address = {Champaign, IL, United States of America}, version = {May 2026}, title = {Aphid Species File}, author = {{Favret}, {Colin}}, year = 2026, month = 5} https://api.checklistbank.org/dataset/1061/archive.zip 2026-06-03T00:37:07.811Z hash://md5/75592983b876ce66f69ccea9cbe01d3b

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	0.16.11
nomer	0.6.5
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. 18.5MiB) under review
elton pull urn:lsid:checklistbank.org:dataset:1061

# generate review notes
elton review urn:lsid:checklistbank.org:dataset:1061 \
 > review.tsv

# export indexed interaction records
elton interactions urn:lsid:checklistbank.org:dataset:1061 \
 > interactions.tsv

# export names and align them with the Catalogue of Life using Nomer
elton names urn:lsid:checklistbank.org:dataset:1061 \
 | 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 urn:lsid:checklistbank.org:dataset:1061) before being able to generate reviews (e.g., elton review urn:lsid:checklistbank.org:dataset:1061), extract interaction claims (e.g., elton interactions urn:lsid:checklistbank.org:dataset:1061), or list taxonomic names (e.g., elton names urn:lsid:checklistbank.org:dataset:1061)

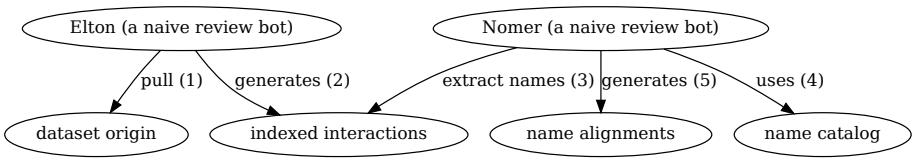


Figure 1: Review Process Overview

Results

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

Files

An extensive list of files produced as part of the review process can be found in Appendix A. Review Files.

Archived Dataset

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

Biotic Interactions

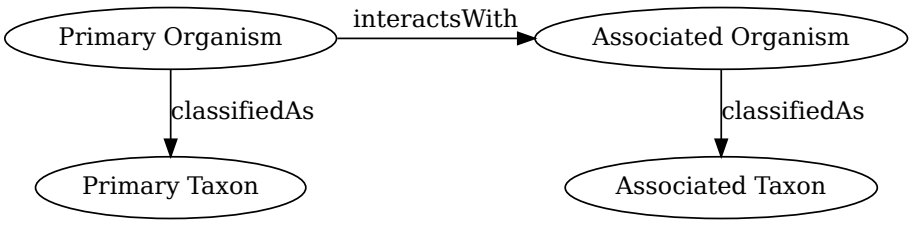


Figure 2: Biotic Interaction Data Model

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 urn:lsid:checklistbank.org:dataset:1061, has fingerprint hash://md5/75592983b876ce66f69ccea9cbe01d3b, is 18.5MiB in size

[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.

and contains 180 interactions with 1 unique type of association (e.g., eats) between 63 primary taxa (e.g., Aphis (Aphis) nasturtii) and 161 associated taxa (e.g., Cornus sericea).

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 2: Sample of Indexed Interaction Claims

sourceTaxonName	interactionTypeName	targetTaxonName	referenceCitation
Aphis (Aphis) spiraephila	eats	Spiraea alba	Lagos, D. (2007) Species of the genus Aphis in the Midwestern States of United States of America.
Aphis (Aphis) thaspii	eats	Thaspium trifoliatum (L.) Gray var. aureum Britt.	Lagos, D. (2007) Species of the genus Aphis in the Midwestern States of United States of America.

sourceTaxonName	interactionTypeName	targetTaxonName	referenceCitation
Aphis (Aphis) ulmariae	eats	Filipendula ulmaria (L.)Maxim.	(**misc?**){ChecklistBankDataset1061, publisher = {Species File Group}, address = {Champaign, IL, United States of America}, version = {May 2026}, title = {Aphid Species File}, author = {{Favret}, {Colin}}, year = 2026, month = 5}. Accessed at https://api.checklistbank.org/dataset/1061/archive.zip on 04 Jun 2026.
Aphis (Aphis) viburni	eats	Viburnum sp.	(**misc?**){ChecklistBankDataset1061, publisher = {Species File Group}, address = {Champaign, IL, United States of America}, version = {May 2026}, title = {Aphid Species File}, author = {{Favret}, {Colin}}, year = 2026, month = 5}. Accessed at https://api.checklistbank.org/dataset/1061/archive.zip on 04 Jun 2026.

Table 3: Most Frequently Mentioned Interaction Types (up to 20 most frequent)

interactionTypeName	count
eats	180

Table 4: Most Frequently Mentioned Primary Taxa (up to 20 most frequent)

sourceTaxonName	count
Aphis (Aphis) nasturtii	62
Aphis (Aphis) fabae	8
Aphis (Aphis) asclepiadis	6
Aphis (Aphis) craccivora	6
Aphis (Aphis) spiraecola	6
Aphis (Aphis) gossypii	5
Anoecia (Anoecia) corni	4
Aphis (Aphis) caliginosa	3
Aphis (Aphis) nerii	3
Aphis (Aphis) polygonata	3
Aphis (Aphis) rumicis	3
Aphis (Aphis) vernoniae	3
Aphis (Aphis) thaspii	2
Aphis (Aphis) coreopsidis	2
Aphis (Aphis) cornifoliae	2
Aphis (Aphis) decepta	2
Aphis (Aphis) diospyri	2
Aphis (Aphis) feminea	2
Aphis (Aphis) glycines	2

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

targetTaxonName	count
Cornus sericea	4
Rhamnus cathartica L.	3
Asclepias syriaca L.	2
Cornus racemosa	2
Glycines max L.	2
Pastinaca sativa L.	2
Hedera helix L.	2

targetTaxonName	count
Taraxacum officinale (Weber)	2
Oenothera biennis L.	2
Malus pumila	2
Rubus spp.	2
Rumex crispus L.	2
Asclepias spp.	2
Vicia spp.	2
Alisma spp	2
Ranunculus spp	2
Spiraea alba	1
Thaspium trifoliatum (L.) Gray var. aureum Britt.	1
Filipendula ulmaria (L.)Maxim.	1

Table 6: Most Frequent Interactions between Primary and Associate Taxa (up to 20 most frequent)

sourceTaxonName	interactionTypeName	targetTaxonName	count
Protaphis knowltoni	eats	Taraxacum officinale (Weber)	2
Aphis (Aphis) nasturtii	eats	Alisma spp	2
Aphis (Aphis) nasturtii	eats	Ranunculus spp	2
Aphis (Aphis) spiraephila	eats	Spiraea alba	1
Aphis (Aphis) thaspii	eats	Thaspium trifoliatum (L.) Gray var. aureum Britt.	1
Aphis (Aphis) ulmariae	eats	Filipendula ulmaria (L.)Maxim.	1
Aphis (Aphis) viburni	eats	Viburnum sp.	1
Aphis (Aphis) viburniphila	eats	Viburnum opulus	1
Aphis (Aphis) asclepiadis	eats	Asclepias syriaca L.	1
Aphis (Aphis) asclepiadis	eats	Asclepias hirtella(Pennell) Woodson	1

sourceTaxonName	interactionTypeName	targetTaxonName	count
Aphis (Aphis) asclepiadis	eats	Asclepias sullivantii Engelm.	1
Aphis (Aphis) asclepiadis	eats	Apocynum cannabinum L.	1
Aphis (Aphis) caliginosa	eats	Cornus racemosa	1
Aphis (Aphis) cephalanthi	eats	Cephalanthus occidentalis	1
Aphis (Aphis) chetansapa	eats	Prunus americana	1
Aphis (Aphis) coreopsidis	eats	Bidens aristosa (Michx.) Britt.	1
Aphis (Aphis) coreopsidis	eats	Eupatorium rugosum Houtt.	1
Aphis (Aphis) cornifoliae	eats	Cornus sericea	1
Aphis (Aphis) craccivora	eats	Astragalus sp.	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

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

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

Geospatial Distribution

If geospatial information was extracted from the dataset under review, the map below will show their distribution. These maps were generated using MapServer (McKenna et al. 2025) tools configured via map configuration indexed-interact ions.map :

```
MAP
  SIZE 1600 800
  EXTENT -180 -90 180 90
  PROJECTION
    "init=epsg:4326"
  END
  LAYER # MODIS WMS map from NASA
    NAME         "modis_nasa"
    TYPE         RASTER
    OFFSITE      0 0 0
    STATUS       ON
    CONNECTIONTYPE WMS
    CONNECTION "https://gibs.earthdata.nasa.gov/wms/epsg4326/best/wms.cgi?"

    METADATA
      "wms_srs" "EPSG:4326"
      "wms_name" "OSM_Land_Water_Map"
      "wms_server_version" "1.1.1"
      "wms_format" "image/jpeg"
    END
    CLASS
      STYLE
        COLOR        232 232 232
        OUTLINECOLOR 32 32 32
      END
    END
  END
  LAYER
    NAME "indexed-interactions"
    TYPE POLYGON
    STATUS ON
```

```
    CONNECTIONTYPE OGR
    CONNECTION "indexed-interactions-h3.gpkg"
    DATA "indexed-interactions-h3"
    CLASS
      STYLE
        COLORRANGE 253.0 231.0 37.0 32.0 164.0 134.0
        DATARANGE NULL NULL
        RANGEITEM "log_number_of_records"
        OUTLINECOLOR 0 0 0
      END
    END
  END
END
```

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.

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.

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 7: Sample of Name Alignments

providedName	relationName	resolvedCatalogName	resolvedName
Anoecia	HAS_ACCEPTED_NAME	col	Anoecia
Anoecia	SYNONYM_OF	col	Anoeciinae
Protaphis middletonii	HAS_ACCEPTED_NAME	col	Protaphis middletonii

providedName	relationName	resolvedCatalogName	resolvedName
Protaphis knowltoni	HAS_ACCEPTED_NAME	NA	Protaphis knowltoni

Table 8: 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	15
col	genus	76
col	species	75
col	subfamily	1
col	subspecies	4
discoverlife	NA	166
gbif	NA	12
gbif	genus	76
gbif	species	75
gbif	subspecies	10
gbif	variety	1
itis	NA	21
itis	genus	75
itis	species	66
itis	subspecies	3
itis	variety	1
mdd	NA	166
ncbi	NA	19
ncbi	genus	76
ncbi	species	71
ncbi	subgenus	1
pbdb	NA	124
pbdb	genus	37
pbdb	species	5
tpt	NA	166
wfo	NA	24
wfo	genus	73
wfo	species	69
wfo	subspecies	1
worms	NA	87
worms	genus	59
worms	species	20

Table 9: Name relationship types per catalog. Name relationship type "NONE" means that a name was not recognized by the associated catalog. "SAME_AS" indicates either a "HAS_ACCEPTED_NAME" or "SYNONYM_OF" name relationship type. We recognize that "SYNONYM_OF" encompasses many types of nomenclatural synonymies (ICZN 1999) (e.g., junior synonym, senior synonyms).

resolvedCatalogName	relationName	count
col	HAS_ACCEPTED_NAME	274
col	SYNONYM_OF	340
col	NONE	15
discoverlife	NONE	224
gbif	HAS_ACCEPTED_NAME	256
gbif	SYNONYM_OF	250
gbif	NONE	12
itis	NONE	22
itis	HAS_ACCEPTED_NAME	193
itis	SYNONYM_OF	17
mdd	NONE	224
ncbi	SAME_AS	261
ncbi	NONE	19
ncbi	SYNONYM_OF	5
pbdb	NONE	125
pbdb	HAS_ACCEPTED_NAME	102
pbdb	SYNONYM_OF	1
tpt	NONE	224
wfo	NONE	81
wfo	HAS_ACCEPTED_NAME	140
wfo	HAS_UNCHECKED_NAME	16
wfo	SYNONYM_OF	19
worms	NONE	89
worms	HAS_ACCEPTED_NAME	139
worms	SYNONYM_OF	19

Table 10: 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)

catalog name	alignment results
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 11: First few lines in the review notes.

reviewDate	reviewCommentType	reviewComment
2026-06-04T00:02:22Z	note	source taxon name missing
2026-06-04T00:02:22Z	note	source taxon name missing
2026-06-04T00:02:22Z	note	source taxon name missing
2026-06-04T00:02:22Z	note	source taxon name missing

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

Table 12: Most frequently occurring review notes, if any.

reviewComment	count
source taxon name missing	36942

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

[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 (2026-06-04) the version of the GloBI dataset index was available at https://globalbioticinteractions.org/datasets

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.

Appendix A. Review 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

[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."

filename	description
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)
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

filename	description
indexed-interactions.map	mapserver configuration to plot species interaction claims indexed from the dataset under review 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
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

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

filename	description
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
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

filename	description
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
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

filename	description
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
indexed-names-resolved-pbdb.csv.gz	taxonomic names found in the dataset under review aligned with the Paleobiology Database as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped comma-separated values format

filename	description
indexed-names-resolved-pbdb.html.gz	taxonomic names found in the dataset under review aligned with Paleobiology Database as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped html format
indexed-names-resolved-pbdb.tsv.gz	taxonomic names found in the dataset under review aligned with Paleobiology Database as accessed through the Nomer Corpus of Taxonomic Resources (J. H. (ed.). Poelen 2024) in gzipped tab-separated values format
indexed-names-resolved-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
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

filename	description
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
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

filename	description
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)
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

filename	description
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

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.

McKenna, Jeff, Steve Lime, Thomas Bonfort, Jérome Boué, Howard Butler, Seth Girvin, Tom Kralidis, et al. 2025. "MapServer." Zenodo. https://doi.org/10.5281/zenodo.17807263.

Poelen, Jorrit H. (ed.). 2024. "Nomer Corpus of Taxonomic Resources Hash://Sha256/ B60c0d25a16ae77b24305782017b1a270b79b5d1746f832650

F2027ba536e276 Hash://Md5/17f1363a277ee0e4ecaf1b91c665e47e." Zenodo. https://doi.org/10.5281/zenodo.12695629.

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

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

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

Trekels, Maarten, Debora Pignatari Drucker, José Augusto Salim, Jeff Ollerton, Jorrit Poelen, Filipi Miranda Soares, Max Rünzel, Muo Kasina, Quentin Groom, and Mariano Devoto. 2023. "WorldFAIR Project (D10.1) 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.