Versioned Archive and Review of Biotic Interactions and Taxon Names Found within globalbioticinteractions/neon-biorepo hash://md5/ca6b3f7e2bf204e7d80b10b2e63a4a00

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

2026-01-28

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/neon-biorepo, has fingerprint hash://md5/ca6b3f7e2bf204e7d80b10b2e63a4a00, is 586MiB in size and contains 25,298 interactions with 3 unique types of associations (e.g., interactsWith) between 2,212 primary taxa (e.g., Acer rubrum) and 2,328 associated taxa (e.g., sharesOriginatingSample). This report includes detailed summaries of interaction data, a taxonomic review from multiple catalogs, and an archived version of the dataset from which the reviews are derived.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 **3**

Results **5**
Files 5
Archived Dataset 14

Biotic Interactions 15
Interaction Networks 18
Geospatial Distribution 18
Taxonomic Alignment 19
Additional Reviews 24
GloBI Review Badge 36
GloBI Index Badge 36

Discussion **36**

Acknowledgements **37**

Author contributions **37**

References **37**

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 :

> MSB-MAMM | Museum of Southwestern Biology at the University of New Mexico https://biorepo.neonscience.org/portal/content/dwca/MSB-MAMM_DwC-A.zip 2026-01-24T05:02:35.461Z NEON-APLC-HVCH | Aquatic Plant, Bryophyte, and Macroalgae Collection (Herbarium Vouchers [Clip Harvests]) https://biorepo.neonscience.org/portal/content/dwca/NEON-APLC-HVCH_DwC-A.zip 2026-01-24T05:02:35.461Z NEON-APLC-HVPC | Aquatic Plant, Bryophyte, and Macroalgae Collection (Herbarium Vouchers [Point Counts]) https://biorepo.neonscience.org/portal/content/dwca/NEON-APLC-HVPC_DwC-A.zip 2026-01-24T05:02:35.461Z NEON-APLC-HVSS | Aquatic Plant, Bryophyte, and Macroalgae Collection (Herbarium Vouchers [Standard Sampling]) https://biorepo.neonscience.org/portal/content/dwca/NEON-APLC-HVSS_DwC-A.zip 2026-01-24T05:02:35.461Z NEON-BMIC-CHI | Identified Benthic Macroinvertebrate Collection (Fluid-Preserved Chironomids) https://biorepo.neonscience.org/portal/content/dwca/NEON-BMIC-CHI_DwC-A.zip 2026-01-24T05:02:35.461Z NEON-BMIC-MSI | Identified Benthic Macroinvertebrate Collection (Slide-Mounted Chironomids and Oligochaetes) https://biorepo.neonscience.org/portal/content/dwca/NEON-

BMIC-MSI_DwC-A.zip 2026-01-24T05:02:35.461Z NEON-BMIC-TSI | Identified Benthic Macroinvertebrate Collection (Taxonomy Subsample) https://biorepo.neonscience.org/portal/content/dwca/NEON-BMIC-TSI_DwC-A.zip 2026-01-24T05:02:35.461Z NEON-DRHC | Herbaria Reference Samples (Retained at Domain Offices) https://biorepo.neonscience.org/portal/content/dwca/NEON-DRHC_DwC-A.zip 2026-01-24T05:02:35.461Z NEON-MAMC-DNA | Mammal Collection (DNA Extracts) https://biorepo.neonscience.org/portal/content/dwca/NEON-MAMC-DNA_DwC-A.zip 2026-01-24T05:02:35.461Z NEON-MAMC-EA | Mammal Collection (Ear Tissue) https://biorepo.neonscience.org/portal/content/dwca/NEON-MAMC-EA_DwC-A.zip 2026-01-24T05:02:35.461Z NEON-MAMC-FE | Mammal Collection (Fecal Samples) https://biorepo.neonscience.org/portal/content/dwca/NEON-MAMC-FE_DwC-A.zip 2026-01-24T05:02:35.461Z NEON-MAMC-HA | Mammal Collection (Hair Samples) https://biorepo.neonscience.org/portal/content/dwca/NEON-MAMC-HA_DwC-A.zip 2026-01-24T05:02:35.461Z NEON-MAMC-VSS | Mammal Collection (Vouchers [Standard Sampling]) https://biorepo.neonscience.org/portal/content/dwca/NEON-MAMC-VSS_DwC-A.zip 2026-01-24T05:02:35.461Z NEON-TICC-PE | Tick Collection (Pathogen Extracts) https://biorepo.neonscience.org/portal/content/dwca/NEON-TICC-PE_DwC-A.zip 2026-01-24T05:02:35.461Z NEON-TPLC-HV | Terrestrial Plant Collection (Herbarium Vouchers) https://biorepo.neonscience.org/portal/content/dwca/NEON-TPLC-HV_DwC-A.zip 2026-01-24T05:02:35.461Z NEON-TPLC-LT | Terrestrial Plant Collection (Leaf Tissue) https://biorepo.neonscience.org/portal/content/dwca/NEON-TPLC-LT_DwC-A.zip 2026-01-24T05:02:35.461Z hash://md5/ca6b3f7e2bf204e7d80b10b2e63a4a00

For additional metadata related to this dataset, please visit https://github.com/globalbioticinteractions/neon-biorepo 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	0.16.1
nomer	0.5.17
globinizer	0.4.0
mlr	6.0.0
jq	1.6

tool name	version
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.  586MiB) under review
elton pull globalbioticinteractions/neon-biorepo

# generate review notes
elton review globalbioticinteractions/neon-biorepo\
 > review.tsv

# export indexed interaction records
elton interactions globalbioticinteractions/neon-biorepo\
 > interactions.tsv

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

or visually, in a process diagram.

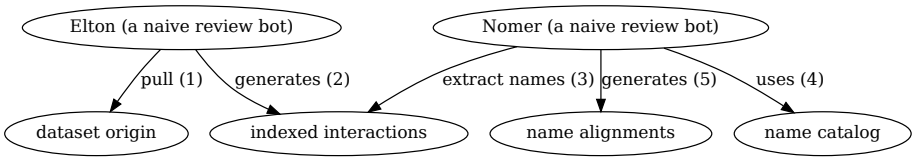


Figure 1: Review Process Overview

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/neon-biorepo) before being able to generate reviews (e.g., elton review globalbioticinteractions/neon-biorepo), extract interaction claims (e.g., elton interactions globalbioticinteractions/neon-biorepo), or list taxonomic names (e.g., elton names globalbioticinteractions/neon-biorepo)

Results

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

Files

The following files are produced in this review:

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

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

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

filename	description
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
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

filename	description
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
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

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

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

filename	description
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
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

filename	description
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
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

filename	description
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
indexed-names-sample.html	first 500 taxonomic names found in the dataset under review in html format
indexed-names-sample.tsv	first 500 taxonomic names found in the dataset under review in tab-separated values format

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

Archived Dataset

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

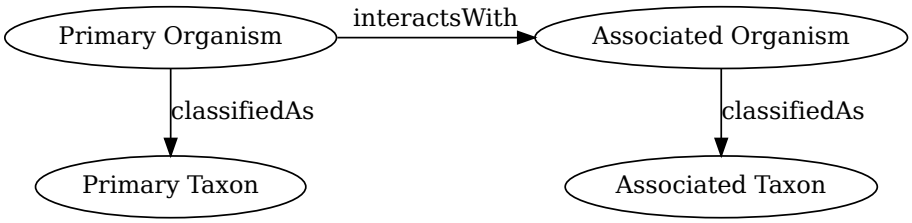


Figure 2: Biotic Interaction Data Model

Biotic Interactions

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

The dataset under review, named globalbioticinteractions/neon-biorepo, has fingerprint hash://md5/ca6b3f7e2bf204e7d80b10b2e63a4a00, is 586MiB in size and contains 25,298 interactions with 3 unique types of associations (e.g., interactsWith) between 2,212 primary taxa (e.g., Acer rubrum) and 2,328 associated taxa (e.g., sharesOriginatingSample).

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
Sarmentypnum exannulatum	interactsWith	algae sp.	https://biorepo.neonscience.org/portal/collectio
Sarmentypnum exannulatum	interactsWith	algae sp.	https://biorepo.neonscience.org/portal/collectio
Sarmentypnum exannulatum	interactsWith	algae sp.	https://biorepo.neonscience.org/portal/collectio
Sarmentypnum exannulatum	interactsWith	algae sp.	https://biorepo.neonscience.org/portal/collectio

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

interactionTypeName	count
interactsWith	25258
killedBy	20
adjacentTo	20

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

sourceTaxonName	count
Acer rubrum	242
Picea mariana	130
Abies balsamea	129
Elaeagnus commutata	123
Bouteloua gracilis	116
Aralia nudicaulis	108
Artemisia tridentata	104
Vaccinium vitis-idaea	104
Poa pratensis	99
Liriodendron tulipifera	90
Bouteloua eriopoda	89
Acer pensylvanicum	89
Larrea tridentata	88
Artemisia absinthium	85
Celtis laevigata	84
Bromus tectorum	80
Erodium botrys	80
Rhododendron groenlandicum	78
Populus tremuloides	77

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

targetTaxonName	count
sharesOriginatingSample	1315
Acer rubrum	429
Bromus inermis	220
Poa pratensis	215
Pinus palustris	202

targetTaxonName	count
Populus tremuloides	201
Liquidambar styraciflua	171
Vaccinium vitis-idaea	167
Liriodendron tulipifera	164
Tsuga canadensis	161
Betula glandulosa/nana	157
Maianthemum canadense	154
Fagus grandifolia	145
Vaccinium uliginosum	139
Artemisia ludoviciana	138
Picea mariana	133
Abies balsamea	132
Acer saccharum	132
Pinus strobus	131

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

sourceTaxonName	interactionTypeName	targetTaxonName	count
Eukiefferiella discoloripes	interactsWith	sharesOriginatingSample	72
Rheotanytarsus exiguus	interactsWith	sharesOriginatingSample	48
Thienemanniella	interactsWith	sharesOriginatingSample	47
Synorthocladius	interactsWith	sharesOriginatingSample	41
Cricotopus	interactsWith	sharesOriginatingSample	39
Corynoneura	interactsWith	sharesOriginatingSample	37
Cricotopus/Orthocladius	interactsWith	sharesOriginatingSample	35
Thienemannimyia	interactsWith	sharesOriginatingSample	29
Polypedilum illinoense	interactsWith	sharesOriginatingSample	26
Tanytarsus	interactsWith	sharesOriginatingSample	24
Simulium	interactsWith	sharesOriginatingSample	24
Picea mariana	interactsWith	Vaccinium vitis-idaea	21
Cricotopus bicinctus	interactsWith	sharesOriginatingSample	20
Rheotanytarsus pellucidus	interactsWith	sharesOriginatingSample	20
Stenochironomus	interactsWith	sharesOriginatingSample	18
Rheocricotopus	interactsWith	sharesOriginatingSample	18
Poa pratensis	interactsWith	Bromus inermis	18

sourceTaxonName	interactionTypeName	targetTaxonName	count
Parakiefferiella	interactsWith	sharesOriginatingSample	17
Nais	interactsWith	sharesOriginatingSample	17

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.

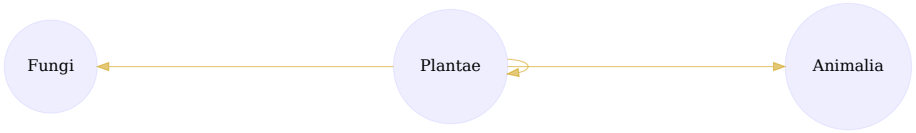


Figure 3: Interactions on taxonomic kingdom rank as interpreted by the Catalogue of Life download svg

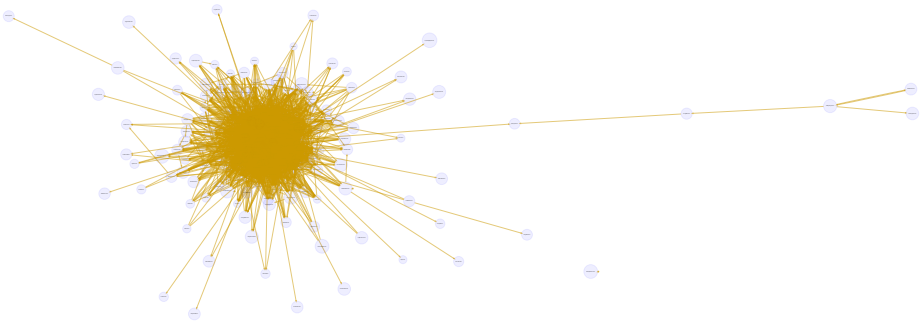


Figure 4: Interactions on the taxonomic family rank as interpreted by the Catalogue of Life. download svg

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

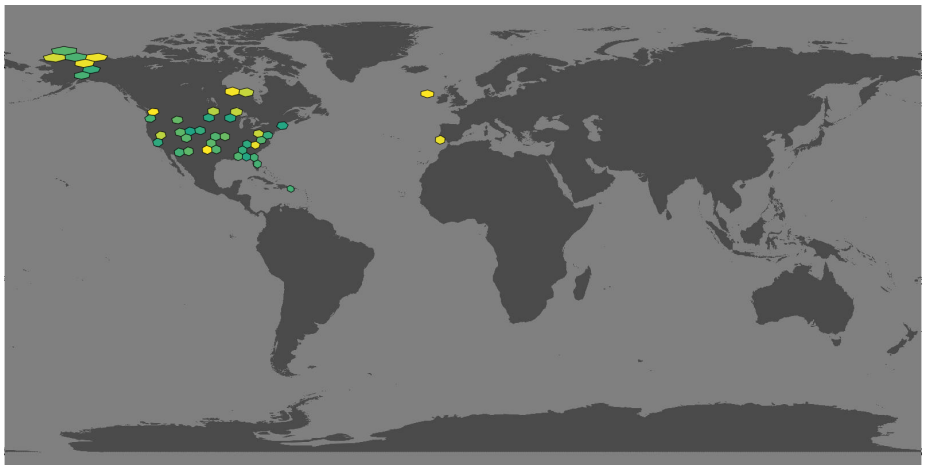


Figure 5: Hexagonal grid cells indicate that interactions claims are available for selected geospatial area: light yellow means relatively fewer claims, dark green relatively more claims.

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 8: Sample of Name Alignments

providedName	relationName	resolvedCatalogName	resolvedName
Caulanthus lasiophyllus	SYNONYM_OF	col	Streptanthus lasiophyllus
Chaetogaster diaphanus	HAS_ACCEPTED_NAME	col	Chaetogaster diaphanus
Corydalidae	HAS_ACCEPTED_NAME	col	Corydalidae
Chimarra	HAS_ACCEPTED_NAME	col	Chimarra

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	314
col	class	5
col	family	43
col	genus	293
col	kingdom	1
col	order	5
col	other	2
col	phylum	2
col	species	2547
col	subclass	1
col	subfamily	3
col	subgenus	9
col	suborder	3
col	subspecies	80
col	superfamily	2
col	tribe	3
col	variety	37
discoverlife	NA	3256
gbif	NA	165
gbif	class	6
gbif	family	43
gbif	form	2
gbif	genus	302
gbif	kingdom	1
gbif	order	5
gbif	phylum	2
gbif	species	2644
gbif	subspecies	132
gbif	variety	117
itis	NA	285
itis	class	5
itis	division	1
itis	family	43
itis	genus	292
itis	kingdom	1
itis	order	5
itis	phylum	2
itis	species	2504
itis	subclass	1

resolvedCatalogName	resolvedRank	count
itis	subfamily	3
itis	suborder	3
itis	subspecies	75
itis	superfamily	1
itis	tribe	3
itis	variety	41
mdd	NA	3256
ncbi	NA	464
ncbi	clade	2
ncbi	class	5
ncbi	family	43
ncbi	genus	290
ncbi	order	4
ncbi	phylum	1
ncbi	species	2419
ncbi	subclass	2
ncbi	subfamily	3
ncbi	subgenus	5
ncbi	suborder	3
ncbi	subspecies	9
ncbi	superfamily	2
ncbi	tribe	3
ncbi	varietas	6
pbdb	NA	2946
pbdb	class	7
pbdb	family	36
pbdb	genus	147
pbdb	informal	1
pbdb	infraorder	1
pbdb	kingdom	1
pbdb	order	5
pbdb	phylum	2
pbdb	species	103
pbdb	subclass	1
pbdb	subfamily	3
pbdb	suborder	1
pbdb	superfamily	2
pbdb	tribe	3
pbdb	unranked clade	1
tpt	NA	3238
tpt	genus	1
tpt	species	17
wfo	NA	549
wfo	class	2

resolvedCatalogName	resolvedRank	count
wfo	family	9
wfo	genus	180
wfo	order	1
wfo	species	2488
wfo	subspecies	36
wfo	variety	22
worms	NA	2458
worms	class	4
worms	family	31
worms	genus	156
worms	kingdom	1
worms	order	4
worms	phylum	1
worms	phylum (division)	1
worms	species	586
worms	subclass	3
worms	subfamily	3
worms	suborder	2
worms	subspecies	7
worms	superfamily	2
worms	tribe	2
worms	variety	5

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	1260
col	HAS_ACCEPTED_NAME	4077
col	NONE	434
discoverlife	NONE	4825
gbif	SYNONYM_OF	2584
gbif	HAS_ACCEPTED_NAME	5305
gbif	NONE	224
itis	HAS_ACCEPTED_NAME	4151
itis	SYNONYM_OF	463
itis	NONE	381

resolvedCatalogName	relationName	count
mdd	NONE	4808
mdd	HAS_ACCEPTED_NAME	17
ncbi	SAME_AS	3992
ncbi	NONE	607
ncbi	SYNONYM_OF	248
pbdb	NONE	4394
pbdb	HAS_ACCEPTED_NAME	434
pbdb	SYNONYM_OF	22
tpt	NONE	4802
tpt	HAS_ACCEPTED_NAME	27
tpt	SYNONYM_OF	1
wfo	SYNONYM_OF	845
wfo	NONE	730
wfo	HAS_ACCEPTED_NAME	3850
wfo	HAS_UNCHECKED_NAME	360
worms	NONE	3662
worms	HAS_ACCEPTED_NAME	1266
worms	SYNONYM_OF	197

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.

reviewDate	reviewCommentType	reviewComment

Table 12: First few lines in the review notes.

reviewDate	reviewCommentType	reviewComment
2026-01-28T20:41:30Z	note	[] Caused by: org.eol.globi.data.StudyImporterException: failed to read archive [https://biorepo.neonscience.org/portal/collections/mis at org.eol.globi.data.DatasetImporterForDwCA.importStu at org.eol.globi.util.DatasetImportUtil.importDataset(Dat at org.eol.globi.util.DatasetImportUtil.importDataset(Dat at org.eol.globi.util.DatasetImportUtil.indexDatasets(Data at org.eol.globi.util.DatasetImportUtil.indexUnresolvedDe at org.eol.globi.util.DatasetImportUtil.resolveAndImportD at org.eol.globi.data.DatasetImporterForRSS.importStudy at org.eol.globi.util.DatasetImportUtil.importDataset(Dat at org.globalbioticinteractions.elton.cmd.CmdReview.revie at org.globalbioticinteractions.elton.cmd.CmdReview.revie at org.globalbioticinteractions.elton.cmd.CmdReview.doRu at org.globalbioticinteractions.elton.cmd.CmdDefaultPara at org.globalbioticinteractions.elton.cmd.CmdTabularWrit at pico-cli.CommandLine.executeUserObject(CommandLine.jav at picocli.CommandLine.access1300(*CommandLine.java* 145)*atpicocli.CommandLine*RunLast.executeUserObje at picocli.CommandLine*RunLast.handle*(*CommandLine.* 2352)*atpicocli.CommandLine*RunLast.handle(Comma at picocli.CommandLine*AbstractParseResultHandler.ex* 2179)*atpicocli.CommandLine*RunLast.execute(Comma at pico-cli.CommandLine.execute(CommandLine.java:2078) at org.globalbioticinteractions.elton.Elton.run(Elton.java:1 at org.globalbioticinteractions.elton.Elton.main(Elton.java Caused by: java.io.IOException: failed to read [file:/tmp/dwca8457334304832273951tmp.zip]

reviewDate	reviewCommentType	reviewComment
2026-01-28T20:42:22Z	note	[] Caused by: org.eol.globi.data.StudyImporterException: failed to read archive [https://biorepo.neonscience.org/portal/collections/mis at org.eol.globi.data.DatasetImporterForDwCA.importStu at org.eol.globi.util.DatasetImportUtil.importDataset(Dat at org.eol.globi.util.DatasetImportUtil.importDataset(Dat at org.eol.globi.util.DatasetImportUtil.indexDatasets(Data at org.eol.globi.util.DatasetImportUtil.indexUnresolvedDe at org.eol.globi.util.DatasetImportUtil.resolveAndImportD at org.eol.globi.data.DatasetImporterForRSS.importStudy at org.eol.globi.util.DatasetImportUtil.importDataset(Dat at org.globalbioticinteractions.elton.cmd.CmdReview.revie at org.globalbioticinteractions.elton.cmd.CmdReview.revie at org.globalbioticinteractions.elton.cmd.CmdReview.doRu at org.globalbioticinteractions.elton.cmd.CmdDefaultPara at org.globalbioticinteractions.elton.cmd.CmdTabularWrit at pico- cli.CommandLine.executeUserObject(CommandLine.jav at picocli.CommandLine.access1300(*CommandLine.java* 145)*atpicocli.CommandLine*RunLast.executeUserObje at picocli.CommandLine*RunLast.handle*(*CommandLine.* 2352)*atpicocli.CommandLine*RunLast.handle(Comma at picocli.CommandLine*AbstractParseResultHandler.ex* 2179)*atpicocli.CommandLine*RunLast.execute(Comma at pico- cli.CommandLine.execute(CommandLine.java:2078) at org.globalbioticinteractions.elton.Elton.run(Elton.java:1 at org.globalbioticinteractions.elton.Elton.main(Elton.java Caused by: java.io.IOException: failed to read [file:/tmp/dwca5264248726656922586tmp.zip] at org.globalbioticinteractions.dataset.DwCAUtil.archiveF at
	27	

reviewDate	reviewCommentType	reviewComment
2026-01-28T20:42:31Z	note	[] Caused by: org.eol.globi.data.StudyImporterException: failed to read archive [https://biorepo.neonscience.org/portal/collections/mis at org.eol.globi.data.DatasetImporterForDwCA.importStu at org.eol.globi.util.DatasetImportUtil.importDataset(Dat at org.eol.globi.util.DatasetImportUtil.importDataset(Dat at org.eol.globi.util.DatasetImportUtil.indexDatasets(Data at org.eol.globi.util.DatasetImportUtil.indexUnresolvedDe at org.eol.globi.util.DatasetImportUtil.resolveAndImportD at org.eol.globi.data.DatasetImporterForRSS.importStudy at org.eol.globi.util.DatasetImportUtil.importDataset(Dat at org.globalbioticinteractions.elton.cmd.CmdReview.revie at org.globalbioticinteractions.elton.cmd.CmdReview.revie at org.globalbioticinteractions.elton.cmd.CmdReview.doRu at org.globalbioticinteractions.elton.cmd.CmdDefaultPara at org.globalbioticinteractions.elton.cmd.CmdTabularWrit at pico- cli.CommandLine.executeUserObject(CommandLine.jav at picocli.CommandLine.access1300(*CommandLine.java* 145)*atpicocli.CommandLine*RunLast.executeUserObje at picocli.CommandLine*RunLast.handle*(*CommandLine.* 2352)*atpicocli.CommandLine*RunLast.handle(Comma at picocli.CommandLine*AbstractParseResultHandler.ex* 2179)*atpicocli.CommandLine*RunLast.execute(Comma at pico- cli.CommandLine.execute(CommandLine.java:2078) at org.globalbioticinteractions.elton.Elton.run(Elton.java:1 at org.globalbioticinteractions.elton.Elton.main(Elton.java Caused by: java.io.IOException: failed to read [file:/tmp/dwca7145190097613054960tmp.zip] at org.globalbioticinteractions.dataset.DwCAUtil.archiveF at
	28	

reviewDate	reviewCommentType	reviewComment
2026-01-28T20:42:31Z	note	[] Caused by: org.eol.globi.data.StudyImporterException: failed to read archive [https://biorepo.neonscience.org/portal/collections/mis at org.eol.globi.data.DatasetImporterForDwCA.importStu at org.eol.globi.util.DatasetImportUtil.importDataset(Dat at org.eol.globi.util.DatasetImportUtil.importDataset(Dat at org.eol.globi.util.DatasetImportUtil.indexDatasets(Data at org.eol.globi.util.DatasetImportUtil.indexUnresolvedDe at org.eol.globi.util.DatasetImportUtil.resolveAndImportD at org.eol.globi.data.DatasetImporterForRSS.importStudy at org.eol.globi.util.DatasetImportUtil.importDataset(Dat at org.globalbioticinteractions.elton.cmd.CmdReview.revie at org.globalbioticinteractions.elton.cmd.CmdReview.revie at org.globalbioticinteractions.elton.cmd.CmdReview.doRu at org.globalbioticinteractions.elton.cmd.CmdDefaultPara at org.globalbioticinteractions.elton.cmd.CmdTabularWrit at pico- cli.CommandLine.executeUserObject(CommandLine.jav at picocli.CommandLine.access1300(*CommandLine.java* 145)*atpicocli.CommandLine*RunLast.executeUserObje at picocli.CommandLine*RunLast.handle*(*CommandLine.* 2352)*atpicocli.CommandLine*RunLast.handle(Comma at picocli.CommandLine*AbstractParseResultHandler.ex* 2179)*atpicocli.CommandLine*RunLast.execute(Comma at pico- cli.CommandLine.execute(CommandLine.java:2078) at org.globalbioticinteractions.elton.Elton.run(Elton.java:1 at org.globalbioticinteractions.elton.Elton.main(Elton.java Caused by: java.io.IOException: failed to read [file:/tmp/dwca8089333683738027903tmp.zip] at org.globalbioticinteractions.dataset.DwCAUtil.archiveF at
	29	

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

reviewComment	count

Table 13: Most frequently occurring review notes, if any.

reviewComment	count
[] Caused by: org.eol.globi.data.StudyImporterException: failed to read archive [https://biorepo.neonscience.org/portal/collections/misc/collprofiles.php?collid=97] at org.eol.globi.data.DatasetImporterForDwCA.importStudy(DatasetImporterForDwCA.java:358) at org.eol.globi.util.DatasetImportUtil.importDataset(DatasetImportUtil.java:71) at org.eol.globi.util.DatasetImportUtil.importDataset(DatasetImportUtil.java:42) at org.eol.globi.util.DatasetImportUtil.indexDatasets(DatasetImportUtil.java:160) at org.eol.globi.util.DatasetImportUtil.indexUnresolvedDependencies(DatasetImportUtil.java:117) at org.eol.globi.util.DatasetImportUtil.resolveAndImportDatasets(DatasetImportUtil.java:84) at org.eol.globi.data.DatasetImporterForRSS.importStudy(DatasetImporterForRSS.java:45) at org.eol.globi.util.DatasetImportUtil.importDataset(DatasetImportUtil.java:71) at org.globalbioticinteractions.elton.cmd.CmdReview.review(CmdReview.java:248) at org.globalbioticinteractions.elton.cmd.CmdReview.reviewLocal(CmdReview.java:201) at org.globalbioticinteractions.elton.cmd.CmdReview.doRun(CmdReview.java:159) at org.globalbioticinteractions.elton.cmd.CmdDefaultParams.run(CmdDefaultParams.java:223) at org.globalbioticinteractions.elton.cmd.CmdTabularWriterParams.run(CmdTabularWriterParams.java:12) at pico-cli.CommandLine.executeUserObject(CommandLine.java:1939) at picocli.CommandLine.access1300(*CommandLine.java* : 145)*atpicocli.CommandLine*RunLast.executeUserObjectOfLastSubcommandWithSameParent(CommandLine at picocli.CommandLine*RunLast.handle*(*CommandLine.java* : 2352)*atpicocli.CommandLine*RunLast.handle(CommandLine.java:2314) at picocli.CommandLine*AbstractParseResultHandler.execute*(*CommandLine.java* : 2179)*atpicocli.CommandLine*RunLast.execute(CommandLine.java:2316) at pico-cli.CommandLine.execute(CommandLine.java:2078) at org.globalbioticinteractions.elton.Elton.run(Elton.java:103) at org.globalbioticinteractions.elton.Elton.main(Elton.java:94) Caused by: java.io.IOException: failed to read [file:/tmp/dwca8457334304832273951tmp.zip] at	1

reviewComment	count
[] Caused by: org.eol.globi.data.StudyImporterException: failed to read archive [https://biorepo.neonscience.org/portal/collections/misc/collprofiles.php?collid=82] at org.eol.globi.data.DatasetImporterForDwCA.importStudy(DatasetImporterForDwCA.java:358) at org.eol.globi.util.DatasetImportUtil.importDataset(DatasetImportUtil.java:71) at org.eol.globi.util.DatasetImportUtil.importDataset(DatasetImportUtil.java:42) at org.eol.globi.util.DatasetImportUtil.indexDatasets(DatasetImportUtil.java:160) at org.eol.globi.util.DatasetImportUtil.indexUnresolvedDependencies(DatasetImportUtil.java:117) at org.eol.globi.util.DatasetImportUtil.resolveAndImportDatasets(DatasetImportUtil.java:84) at org.eol.globi.data.DatasetImporterForRSS.importStudy(DatasetImporterForRSS.java:45) at org.eol.globi.util.DatasetImportUtil.importDataset(DatasetImportUtil.java:71) at org.globalbioticinteractions.elton.cmd.CmdReview.review(CmdReview.java:248) at org.globalbioticinteractions.elton.cmd.CmdReview.reviewLocal(CmdReview.java:201) at org.globalbioticinteractions.elton.cmd.CmdReview.doRun(CmdReview.java:159) at org.globalbioticinteractions.elton.cmd.CmdDefaultParams.run(CmdDefaultParams.java:223) at org.globalbioticinteractions.elton.cmd.CmdTabularWriterParams.run(CmdTabularWriterParams.java:12) at pico-cli.CommandLine.executeUserObject(CommandLine.java:1939) at picocli.CommandLine.access1300(*CommandLine.java* : 145)*atpicocli.CommandLine*RunLast.executeUserObjectOfLastSubcommandWithSameParent(CommandLine at picocli.CommandLine*RunLast.handle*(*CommandLine.java* : 2352)*atpicocli.CommandLine*RunLast.handle(CommandLine.java:2314) at picocli.CommandLine*AbstractParseResultHandler.execute*(*CommandLine.java* : 2179)*atpicocli.CommandLine*RunLast.execute(CommandLine.java:2316) at pico-cli.CommandLine.execute(CommandLine.java:2078) at org.globalbioticinteractions.elton.Elton.run(Elton.java:103) at org.globalbioticinteractions.elton.Elton.main(Elton.java:94) Caused by: java.io.IOException: failed to read [file:/tmp/dwca5264248726656922586tmp.zip] at org.globalbioticinteractions.dataset.DwCAUtil.archiveFor(DwCAUtil.java:29) at org.eol.globi.data.DatasetImporterForDwCA.importStudy(DatasetImporterForDwCA.java:331)	1

reviewComment	count
[] Caused by: org.eol.globi.data.StudyImporterException: failed to read archive [https://biorepo.neonscience.org/portal/collections/misc/collprofiles.php?collid=110] at org.eol.globi.data.DatasetImporterForDwCA.importStudy(DatasetImporterForDwCA.java:358) at org.eol.globi.util.DatasetImportUtil.importDataset(DatasetImportUtil.java:71) at org.eol.globi.util.DatasetImportUtil.importDataset(DatasetImportUtil.java:42) at org.eol.globi.util.DatasetImportUtil.indexDatasets(DatasetImportUtil.java:160) at org.eol.globi.util.DatasetImportUtil.indexUnresolvedDependencies(DatasetImportUtil.java:117) at org.eol.globi.util.DatasetImportUtil.resolveAndImportDatasets(DatasetImportUtil.java:84) at org.eol.globi.data.DatasetImporterForRSS.importStudy(DatasetImporterForRSS.java:45) at org.eol.globi.util.DatasetImportUtil.importDataset(DatasetImportUtil.java:71) at org.globalbioticinteractions.elton.cmd.CmdReview.review(CmdReview.java:248) at org.globalbioticinteractions.elton.cmd.CmdReview.reviewLocal(CmdReview.java:201) at org.globalbioticinteractions.elton.cmd.CmdReview.doRun(CmdReview.java:159) at org.globalbioticinteractions.elton.cmd.CmdDefaultParams.run(CmdDefaultParams.java:223) at org.globalbioticinteractions.elton.cmd.CmdTabularWriterParams.run(CmdTabularWriterParams.java:12) at pico-cli.CommandLine.executeUserObject(CommandLine.java:1939) at picocli.CommandLine.access1300(*CommandLine.java* : 145)*atpicocli.CommandLine*RunLast.executeUserObjectOfLastSubcommandWithSameParent(CommandLine at picocli.CommandLine*RunLast.handle*(*CommandLine.java* : 2352)*atpicocli.CommandLine*RunLast.handle(CommandLine.java:2314) at picocli.CommandLine*AbstractParseResultHandler.execute*(*CommandLine.java* : 2179)*atpicocli.CommandLine*RunLast.execute(CommandLine.java:2316) at pico-cli.CommandLine.execute(CommandLine.java:2078) at org.globalbioticinteractions.elton.Elton.run(Elton.java:103) at org.globalbioticinteractions.elton.Elton.main(Elton.java:94) Caused by: java.io.IOException: failed to read [file:/tmp/dwca7145190097613054960tmp.zip] at org.globalbioticinteractions.dataset.DwCAUtil.archiveFor(DwCAUtil.java:29) at org.eol.globi.data.DatasetImporterForDwCA.importStudy(DatasetImporterForDwCA.java:331)	1

reviewComment	count
[] Caused by: org.eol.globi.data.StudyImporterException: failed to read archive [https://biorepo.neonscience.org/portal/collections/misc/collprofiles.php?collid=111] at org.eol.globi.data.DatasetImporterForDwCA.importStudy(DatasetImporterForDwCA.java:358) at org.eol.globi.util.DatasetImportUtil.importDataset(DatasetImportUtil.java:71) at org.eol.globi.util.DatasetImportUtil.importDataset(DatasetImportUtil.java:42) at org.eol.globi.util.DatasetImportUtil.indexDatasets(DatasetImportUtil.java:160) at org.eol.globi.util.DatasetImportUtil.indexUnresolvedDependencies(DatasetImportUtil.java:117) at org.eol.globi.util.DatasetImportUtil.resolveAndImportDatasets(DatasetImportUtil.java:84) at org.eol.globi.data.DatasetImporterForRSS.importStudy(DatasetImporterForRSS.java:45) at org.eol.globi.util.DatasetImportUtil.importDataset(DatasetImportUtil.java:71) at org.globalbioticinteractions.elton.cmd.CmdReview.review(CmdReview.java:248) at org.globalbioticinteractions.elton.cmd.CmdReview.reviewLocal(CmdReview.java:201) at org.globalbioticinteractions.elton.cmd.CmdReview.doRun(CmdReview.java:159) at org.globalbioticinteractions.elton.cmd.CmdDefaultParams.run(CmdDefaultParams.java:223) at org.globalbioticinteractions.elton.cmd.CmdTabularWriterParams.run(CmdTabularWriterParams.java:12) at pico-cli.CommandLine.executeUserObject(CommandLine.java:1939) at picocli.CommandLine.access1300(*CommandLine.java* : 145)*atpicocli.CommandLine*RunLast.executeUserObjectOfLastSubcommandWithSameParent(CommandLine at picocli.CommandLine*RunLast.handle*(*CommandLine.java* : 2352)*atpicocli.CommandLine*RunLast.handle(CommandLine.java:2314) at picocli.CommandLine*AbstractParseResultHandler.execute*(*CommandLine.java* : 2179)*atpicocli.CommandLine*RunLast.execute(CommandLine.java:2316) at pico-cli.CommandLine.execute(CommandLine.java:2078) at org.globalbioticinteractions.elton.Elton.run(Elton.java:103) at org.globalbioticinteractions.elton.Elton.main(Elton.java:94) Caused by: java.io.IOException: failed to read [file:/tmp/dwca8089333683738027903tmp.zip] at org.globalbioticinteractions.dataset.DwCAUtil.archiveFor(DwCAUtil.java:29) at org.eol.globi.data.DatasetImporterForDwCA.importStudy(DatasetImporterForDwCA.java:331)	1

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

[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 (2026-01-28) the version of the GloBI dataset index was available at https://globalbioticinteractions.org/datasets

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

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

Acknowledgements

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

Author contributions

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

References

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

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,

[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."

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.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.