

A Review of Biotic Interactions and Taxon Names Found in `globalbioticinteractions/kalro2023`

by Nomer and Elton, two naive review bots

review@globalbioticinteractions.org

<https://globalbioticinteractions.org/contribute>

<https://github.com/globalbioticinteractions/kalro2023/issues>

2024-02-11

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 describe a review process of such an openly accessible digital interactions dataset of known origin, and discuss their outcome. The dataset under review (aka `globalbioticinteractions/kalro2023`) has size 410KiB and contains 1023 interactions with 8 unique types of associations (e.g., pollinates) between 512 primary taxa (e.g., *Apis mellifera*) and 331 associated taxa (e.g., *Persea americana*). The report includes detailed summaries of interactions data as well as a taxonomic review from multiple perspectives.

Table of Contents

Introduction	2
Data Review	2
Methods	2
Results	3
Biotic Interactions	3
Interaction Networks	8
Taxonomic Alignment	9
Additional Reviews	13
GloBI Review Badge	14
GloBI Index Badge	14
Discussion	15
Acknowledgements	15
References	15

Introduction

Data Review

Data review can be a time consuming process, especially when done manually. This review report aims to help facilitate data review of species interaction claims made in datasets registered with Global Biotic Interactions (Poelen, Simons, and Mungall 2014). The review includes summary statistics of, and observations about, the dataset under review:

A review of the status of web-based African Plant-Pollinator Interaction data. file:///home/runner/work/kalro2023/kalro2023/./

For additional metadata related to this dataset, please visit <https://github.com/globalbioticinteractions/kalro2023> 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, Elton, Nomer combined with third-party tools like grep, mlr, tail and head.

Tools used in this review process

tool name	version
elton	0.13.2
nomer	0.5.6
mlr	6.0.0
pandoc	3.1.6.1

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

```
# get versioned copy of the dataset (size approx. 410KiB) under review  
  
elton pull globalbioticinteractions/kalro2023
```

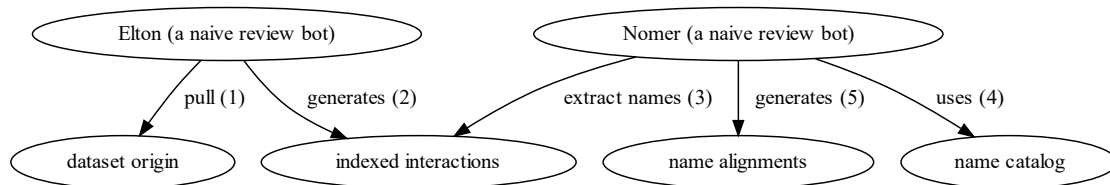
¹ Note that you have to first get the data (e.g., via `elton pull globalbioticinteractions/kalro2023`) before being able to generate reviews (e.g., `elton review globalbioticinteractions/kalro2023`), extract interaction claims (e.g., `elton interactions globalbioticinteractions/kalro2023`), or list taxonomic names (e.g., `elton names globalbioticinteractions/kalro2023`)

```
# generate review notes
elton review globalbioticinteractions/kalro2023\
> review.tsv

# export indexed interaction records
elton interactions globalbioticinteractions/kalro2023\
> interactions.tsv

# export names and align them with the Catalogue of Life using Nomer
elton names globalbioticinteractions/kalro2023\
| nomer append col\
> name-alignment.tsv
```

or visually, in a process diagram.



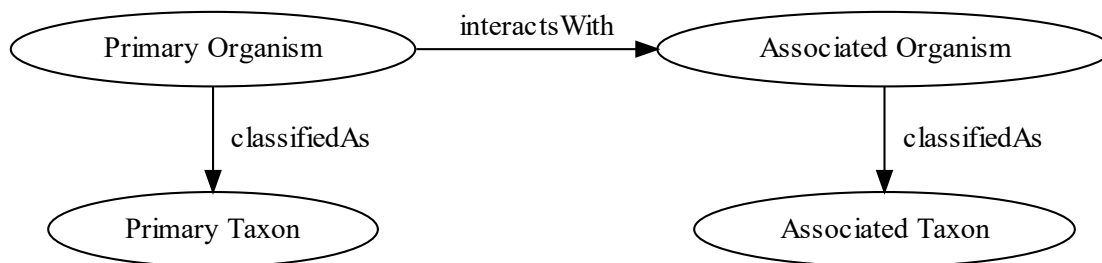
Review Process Overview

You can find a recent copy of the full review script at check-data.sh.

Results

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

Biotic Interactions



Biotic Interaction Data Model

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.

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 (aka globalbioticinteractions/kalro2023) has size 410KiB and contains 1023 interactions with 8 unique types of associations (e.g., pollinates) between 512 primary taxa (e.g., *Apis mellifera*) and 331 associated taxa (e.g., *Persea americana*).

An exhaustive list of indexed interaction claims can be found in [csv](#) and [tsv](#) archives. To facilitate discovery, the first 500 claims available on the html page at [indexed-interactions.html](#) are shown below.

The exhaustive list was used to create the following data summaries below.

Sample of Indexed Interaction Claims

sourceTaxonName	interactionTypeName	targetTaxonName	referenceCitation
Leptoconchus incycloseris	parasiteOf	Fungia (Cycloseris) costulata	Gittenberger, A., Gittenberger, E. (2011). Cryptic, adaptive radiation of endoparasitic snails: sibling species of Leptoconchus (Gastropoda: Coralliophilidae) in corals. Org Divers Evol, 11(1), 21–41. doi:10.1007/s13127-011-0039-1
Leptoconchus infungites	parasiteOf	Fungia (Fungia) fungites	Gittenberger, A., Gittenberger, E. (2011). Cryptic, adaptive radiation of endoparasitic snails: sibling species of Leptoconchus (Gastropoda: Coralliophilidae) in corals. Org Divers Evol, 11(1), 21–41.

sourceTaxonName	interactionTypeNa me	targetTaxonName	referenceCitation
			doi:10.1007/s13127-011-0039-1
Leptoconchus ingrandifungi	parasiteOf	Sandalolitha dentata	Gittenberger, A., Gittenberger, E. (2011). Cryptic, adaptive radiation of endoparasitic snails: sibling species of Leptoconchus (Gastropoda: Coralliophilidae) in corals. Org Divers Evol, 11(1), 21–41. doi:10.1007/s13127-011-0039-1
Leptoconchus ingranulosa	parasiteOf	Fungia (Wellsofungia) granulosa	Gittenberger, A., Gittenberger, E. (2011). Cryptic, adaptive radiation of endoparasitic snails: sibling species of Leptoconchus (Gastropoda: Coralliophilidae) in corals. Org Divers Evol, 11(1), 21–41. doi:10.1007/s13127-011-0039-1

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

interactionTypeName	count
pollinates	338
acquiresNutrientsFrom	294
visits	197
visitsFlowersOf	184
parasiteOf	11
preysOn	2

interactionTypeName	count
parasitoidOf	1
hostOf	1

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

sourceTaxonName	count
Apis mellifera	121
Apis melifera	38
Megachile sp. 2	16
Agrius convolvuli	14
Macrogalea candida	12
Pseudapis sp.	12
Ceratina sp. 3	10
Xylocopa sp.	10
Apis melliferra	9
Apis mellifera Linnaeus	9
Amegilla sp. 1	8
Xylocopa caffra	7
Xylocopa nigrita	7
Xylocopa flavorufa	7
Hypotrigona gribodoi	7
Meliponula ferruginea	6
Lasioglossum sp.	6
Braunsapis sp.	6
Catopsilia florella	6

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

targetTaxonName	count
Persea americana	48
Helianthus annuus L.	30
Acacia tortilis	29
Arctotheca calendula(Cape Marigold)	29
Vernonia cinerea	25
Sesamum indicum	22
Citrullus lanatus	21

targetTaxonName	count
Agathisanthemum bojeri Klotzsch	20
Mangifera indica	20
Coffee arabica	17
Phaseolus vulgaris	16
Phaseolus coccineus	15
Helianthus annuus L	14
Trifolium alexandrinum	13
Cucumis sativus	13
Phaseolus vulgaris L	12
Agathisanthemum bojeri	11
Zygophyllum sp.	11
Cucumis melo	11

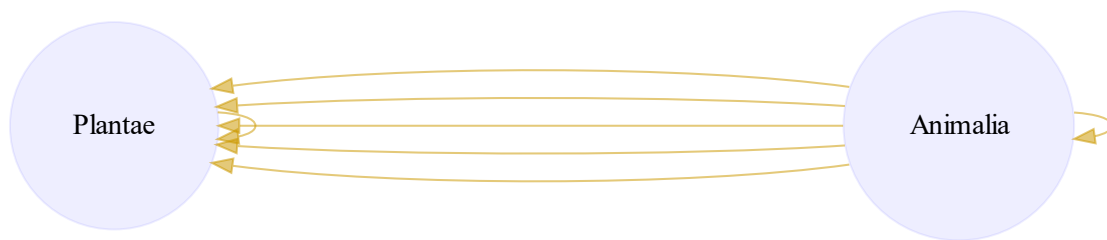
Most Frequent Interactions between Primary and Associate Taxa (up to 20 most frequent)

sourceTaxonName	interactionTypeName	targetTaxonName	count
Asclepias cucullata	pollinates	Sisyranthus trichostomus	3
Xylocopa caffra	acquiresNutrientsFrom	Rhynchosia velutina	2
Steganomus sp.	acquiresNutrientsFrom	Crotalaria emarginata Benth	2
Megachile sp. 2	acquiresNutrientsFrom	Indigofera paniculata	2
Xylocopa flavicollis DeGeer	acquiresNutrientsFrom	Rhynchosia velutina Wight & Arn.	2
Apis mellifera scutellata	pollinates	Mangifera indica	2
Apis mellifera scutellata	pollinates	Helianthus annuus	2
Braunsapis spp.	visits	Lantana camara	2
Belenois aurota Fabricius	visitsFlowersOf	Helianthus annuus L.	2
Hemipepsis hilaris Smith	pollinates	Xysmalobium involucratum	2

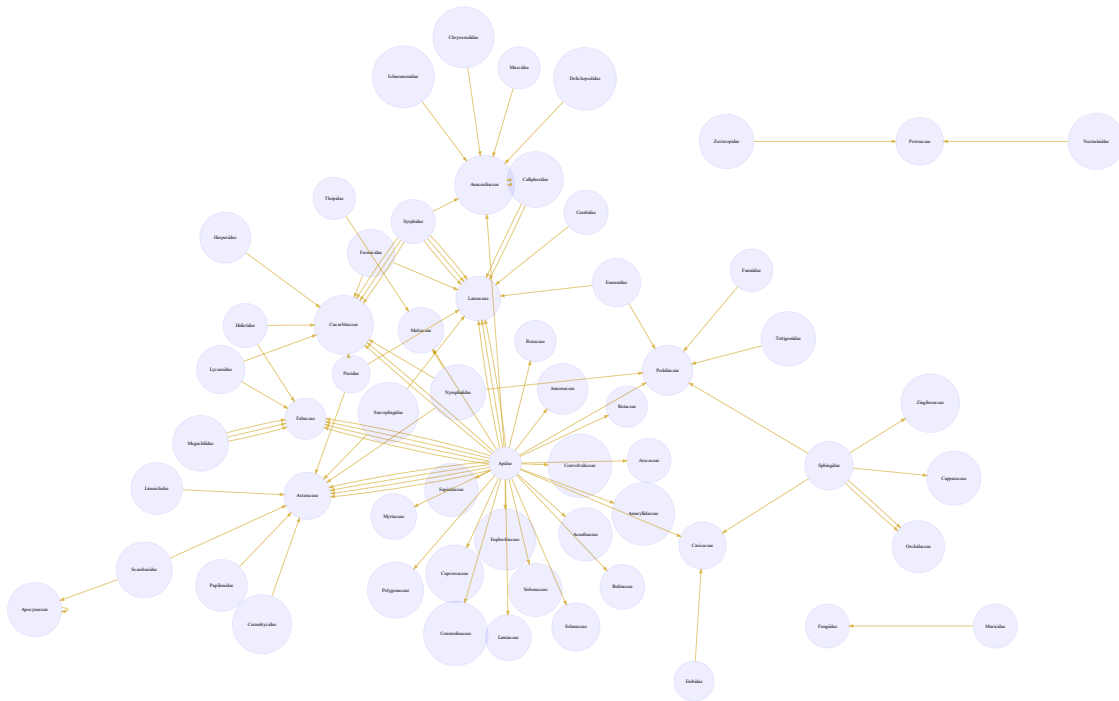
sourceTaxonName	interactionTypeNa me	targetTaxonName	count
Leptoconchus incycloseris	parasiteOf	Fungia (Cycloseris) costulata	1
Leptoconchus infungites	parasiteOf	Fungia (Fungia) fungites	1
Leptoconchus ingrandifungi	parasiteOf	Sandalolitha dentata	1
Leptoconchus ingranulosa	parasiteOf	Fungia (Wellsofungia) granulosa	1
Leptoconchus inlimax	parasiteOf	Herpolitha limax	1
Leptoconchus inpileus	parasiteOf	Halomitra pileus	1
Leptoconchus inpleuractis	parasiteOf	Fungia (Pleuractis) gravis	1
Leptoconchus inpleuractis	parasiteOf	Fungia (Pleuractis) moluccensis	1
Leptoconchus inpleuractis	parasiteOf	Fungia (Pleuractis) paumotensis	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. Note that both network graphs were first aligned taxonomically via the Catalogue of Life. Please refer to the original (or verbatim) taxonomic names for a more original view on the interaction data.



Interactions on taxonomic kingdom rank as interpreted by the Catalogue of Life
[download svg](#)



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](#). A tab-separated file can be found at [indexed-interactions.tsv](#)

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, globi, mdd, tpt, and pbdb). These alignments may serve as a way to review name usage or aid in selecting of a suitable taxonomic name resource to use.

Sample of Name Alignments

providedName	relationName	resolvedCatalogName	resolvedName
Dialictus	NONE	col	Dialictus
Pseudalpis	NONE	col	Pseudalpis
Brevispica	NONE	col	Brevispica
Etbaica	NONE	col	Etbaica

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
tpt	NA	642
tpt	genus	1
tpt	species	1
pbdb	NA	568
pbdb	genus	47
pbdb	family	15
pbdb	species	7
pbdb	order	3
pbdb	subfamily	3
pbdb	class	1
ncbi	species	284
ncbi	NA	240
ncbi	genus	90
ncbi	subgenus	15
ncbi	family	14
ncbi	subspecies	4
ncbi	subfamily	4
ncbi	order	3
ncbi	species group	1
ncbi	class	1
mdd	NA	644
itis	NA	336
itis	species	217
itis	genus	68
itis	family	14
itis	subfamily	3
itis	order	3
itis	variety	1
itis	subspecies	1
itis	class	1
globi	NA	442
globi	species	314

resolvedCatalogName	resolvedRank	count
globi	genus	150
globi	family	19
globi	subgenus	14
globi	subspecies	10
globi	order	7
globi	variety	5
globi	subfamily	5
globi	form	3
globi	superfamily	2
globi	subclass	1
globi	species group	1
globi	class	1
gbif	species	353
gbif	NA	168
gbif	genus	103
gbif	family	15
gbif	subspecies	8
gbif	order	3
gbif	variety	2
gbif	class	1
discoverlife	NA	559
discoverlife	species	85
col	species	337
col	NA	194
col	genus	91
col	family	14
col	subspecies	7
col	order	3
col	subfamily	2
col	section	1
col	variety	1
col	class	1

Name relationship types per catalog. Name relationship type "NONE" means that a name was not recognized by the associated catalog. "SAME_AS" indicates either a

“HAS_ACCEPTED_NAME” or “SYNONYM_OF” name relationship type. We recognize that “SYNONYM_OF” encompasses many types of nomenclatural synonymies (ICZN 1999) (e.g., junior synonym, senior synonyms).

resolvedCatalogName	relationName	count
col	NONE	208
col	HAS_ACCEPTED_NAME	592
col	SYNONYM_OF	163
discoverlife	NONE	731
discoverlife	HAS_ACCEPTED_NAME	94
discoverlife	SYNONYM_OF	29
discoverlife	HOMONYM_OF	5
gbif	HAS_ACCEPTED_NAME	720
gbif	NONE	180
gbif	SYNONYM_OF	300
globi	SAME_AS	8056
globi	NONE	260
itis	NONE	381
itis	HAS_ACCEPTED_NAME	427
itis	SYNONYM_OF	39
mdd	NONE	841
ncbi	SAME_AS	601
ncbi	NONE	262
ncbi	SYNONYM_OF	34
pbdb	NONE	690
pbdb	HAS_ACCEPTED_NAME	149
pbdb	SYNONYM_OF	8
tpt	NONE	839
tpt	HAS_ACCEPTED_NAME	2

List of Available Name Alignment Reports

catalog name	alignment results
col	associated names alignments (first 500, full csv/tsv)
ncbi	associated names alignments (first 500, full csv/tsv)
discoverlife	associated names alignments (first 500, full csv/tsv)

catalog name	alignment results
gbif	associated names alignments (first 500, full csv/tsv)
itis	associated names alignments (first 500, full csv/tsv)
globi	associated names alignments (first 500, full csv/tsv)
mdd	associated names alignments (first 500, full csv/tsv)
tpi	associated names alignments (first 500, full csv/tsv)
pbdb	associated names alignments (first 500, full csv/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.

First few lines in the review notes.

reviewDate	reviewCommentType	reviewComment
2024-02-11T20:23:14Z	note	found unsupported interaction type with name: [Honey production]
2024-02-11T20:23:14Z	note	found malformed doi [doi.org/10.1093/aob/mc v137]
2024-02-11T20:23:14Z	note	found malformed doi [doi.org/10.1093/aob/mc v138]
2024-02-11T20:23:14Z	note	found malformed doi [doi.org/10.1093/aob/mc v139]

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

Most frequently occurring review notes, if any.

reviewComment	count
found unsupported interaction type with name: [Honey production]	1

reviewComment	count
found malformed doi [doi.org/10.1093/aob/mcv137]	1
found malformed doi [doi.org/10.1093/aob/mcv138]	1
found malformed doi [doi.org/10.1093/aob/mcv139]	1

For addition information on review notes, please have a look at the first 500 [Review Notes](#) or the download full [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.



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



Sample 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.

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 (2024-02-11) the version of the GloBI dataset index was available at [\[https://globalbioticinteractions.org/datasets\]](https://globalbioticinteractions.org/datasets)(<https://globalbioticinteractions.org/datasets>)

Discussion

This review is intended to provide a perspective on the dataset to aid understanding of species interaction claims discovered. However, this review should *not* be considered as fitness of use or other kind of quality assessment. Instead, the review may be used as an indication of the open-ness⁶ and FAIRness (Wilkinson et al. 2016; Trekels et al. 2023) of the dataset: in order to perform this review, the data was likely openly available, **F**indable, **A**ccessible, **I**nteroperable and **R**eusable. Currently, this Open-FAIR assessment is qualitative, and with measurement units specified, a more quantitative approach can be implemented.

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 means that it can be more precisely processed 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 to facilitate seamless data exchanges, i.e., 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.

References

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

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>.

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

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

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>.