

# A Review of Biotic Interactions and Taxon Names Found in globalbioticinteractions/ucsb-izc

by Nomer and Elton, two naive review bots  
review@globalbioticinteractions.org  
<https://globalbioticinteractions.org/contribute>  
<https://github.com/globalbioticinteractions/ucsb-izc/issues>

2024-07-23

## 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 process of such an openly accessible digital interactions dataset of known origin, and discuss its outcome. The dataset under review, named globalbioticinteractions/ucsb-izc, is 6.27MiB in size and contains 2,369 interaction with 8 unique types of associations (e.g., interactsWith) between 395 primary taxa (e.g., Araneidae) and 463 associated taxa (e.g., Lupinus bicolor). The report includes detailed summaries of interactions data as well as a taxonomic review from multiple catalogs.

## Contents

|                                |           |
|--------------------------------|-----------|
| <b>Introduction</b>            | <b>2</b>  |
| Data Review . . . . .          | 2         |
| <b>Methods</b>                 | <b>2</b>  |
| <b>Results</b>                 | <b>3</b>  |
| Biotic Interactions . . . . .  | 3         |
| Interaction Networks . . . . . | 6         |
| Taxonomic Alignment . . . . .  | 8         |
| Additional Reviews . . . . .   | 12        |
| GloBI Review Badge . . . . .   | 13        |
| GloBI Index Badge . . . . .    | 13        |
| <b>Discussion</b>              | <b>14</b> |
| <b>Acknowledgements</b>        | <b>14</b> |

|                      |    |
|----------------------|----|
| Author contributions | 14 |
| References           | 14 |

## 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:

University of California Santa Barbara Invertebrate Zoology  
Collection <https://github.com/globalbioticinteractions/ucsb-izc/archive/42a422d5369aafe3a41c6b4d4c9f905fe2a74b2f.zip> 2024-07-20T05:16:35.591Z 6d5c8c051a264b07a767631b35d2208dfa509ab332921a4394d06d5b726af3

For additional metadata related to this dataset, please visit <https://github.com/globalbioticinteractions/ucsb-izc> 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.

Table 1: Tools used in this review process

| tool name | version |
|-----------|---------|
| elton     | 0.13.4  |
| nomer     | 0.5.13  |
| mlr       | 6.0.0   |
| pandoc    | 3.1.6.1 |

The review process can be described in the form of the script below <sup>1</sup>.

```
# get versioned copy of the dataset (size approx. 6.27MiB) under review
elton pull globalbioticinteractions/ucsb-izc
```

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

```
# generate review notes
elton review globalbioticinteractions/ucsb-izc\
> review.tsv

# export indexed interaction records
elton interactions globalbioticinteractions/ucsb-izc\
> interactions.tsv

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

or visually, in a process diagram.

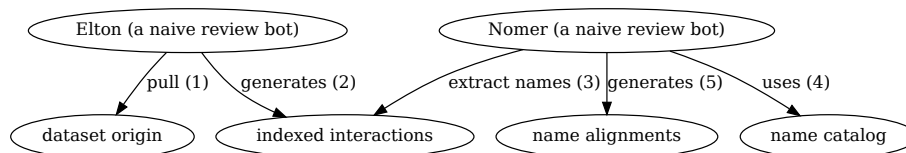


Figure 1: Review Process Overview

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 <sup>2</sup>. Then, links to the detailed review reports are provided.

### Biotic Interactions

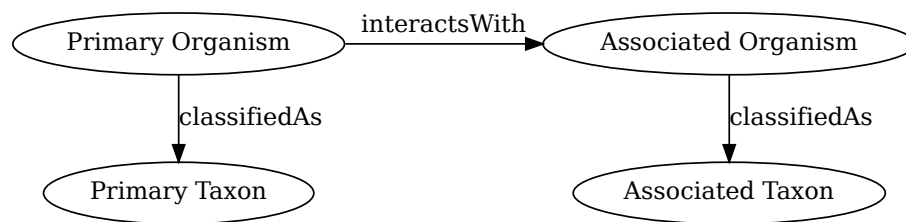


Figure 2: Biotic Interaction Data Model

<sup>2</sup>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, named globalbioticinteractions/ucsb-izc, is 6.27MiB in size and contains 2,369 interaction with 8 unique types of associations (e.g., interactsWith) between 395 primary taxa (e.g., Araneidae) and 463 associated taxa (e.g., *Lupinus bicolor*).

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.

Table 2: Sample of Indexed Interaction Claims

| sourceTaxonName | interactionTypeNam | targetTaxonName                         | referenceCitation  |
|-----------------|--------------------|---|--|
| Lasioglossum    | interactsWith      | inside flower of<br>Eschscholzia        | UCSB-<br>IZC00038170   |
| Diptera         | adjacentTo         | californica<br>flower of Mimulus        | https://ecdysis.org/collections/individual/index.p<br>UCSB-<br>IZC00038207 |
| Diptera         | adjacentTo         | auranticus<br>flower of Mimulus         | https://ecdysis.org/collections/individual/index.p<br>UCSB-<br>IZC00038169 |
| Diptera         | adjacentTo         | auranticus<br>flower of<br>Eschscholzia | https://ecdysis.org/collections/individual/index.p<br>UCSB-<br>IZC00038248 |
|                 |                    | californica                             | https://ecdysis.org/collections/individual/index.p                         |

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

| interactionTypeNam | count |
|--------------------|-------|
| interactsWith      | 1574  |
| adjacentTo         | 573   |
| visits             | 151   |
| visitsFlowersOf    | 44    |
| hasHost            | 18    |
| hostOf             | 4     |
| eats               | 3     |
| coOccursWith       | 2     |

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

| sourceTaxonName          | count |
|--------------------------|-------|
| Araneidae                | 133   |
| Lasioglossum             | 106   |
| Apis mellifera           | 105   |
| Linepithema humile       | 87    |
| Cicadellidae             | 72    |
| Augochlorella pomoniella | 63    |
| Agapostemon texanus      | 47    |
| Solenopsis               | 46    |
| Hemiptera                | 39    |
| Heleomyzidae             | 39    |
| Syrphidae                | 38    |
| Halictus tripartitus     | 36    |
| Ceratina acantha         | 36    |
| Bombus vosnesenskii      | 35    |
| Temnothorax andrei       | 35    |
| Colletes                 | 30    |
| Aphididae                | 27    |
| Lygus                    | 26    |
| Anyphaenidae             | 25    |

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

| targetTaxonName         | count |
|-------------------------|-------|
| Lupinus bicolor         | 282   |
| Lupinus nipomensis      | 99    |
| Quercus agrifolia       | 96    |
| Venegasia carpesioides  | 63    |
| Erigonium               | 59    |
| Arctostaphylos          | 49    |
| Marrubium               | 46    |
| Populus                 | 42    |
| Brassica                | 40    |
| Salix                   | 39    |
| Encelia californica     | 37    |
| Foeniculum              | 36    |
| Atriplex lentiformis    | 32    |
| Cupressus macrocarpa    | 31    |
| Heteromeles arbutifolia | 28    |

| targetTaxonName                    | count |
|------------------------------------|-------|
| flower: Malocothamnus fasciculatus | 28    |
| Lupinus succulentus                | 26    |
| Convolvulus arvensis               | 26    |
| flower: Calystegia macrostegia     | 24    |

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

| sourceTaxonName       | interactionTypeName | targetTaxonName          | count |
|-----------------------|---------------------|--------------------------|-------|
| Cicadellidae          | interactsWith       | Lupinus bicolor          | 52    |
| Solenopsis            | interactsWith       | Quercus agrifolia        | 32    |
| Heleomyzidae          | interactsWith       | Lupinus bicolor          | 30    |
| Lasioglossum          | interactsWith       | Venegasia carpesioides   | 30    |
| Linepithema humile    | interactsWith       | Populus                  | 29    |
| Lygus                 | interactsWith       | Lupinus bicolor          | 24    |
| Linepithema humile    | interactsWith       | Salix                    | 24    |
| Aphididae             | interactsWith       | Lupinus bicolor          | 22    |
| Hemiptera             | interactsWith       | Lupinus bicolor          | 21    |
| Apis mellifera        | interactsWith       | Arctostaphylos           | 19    |
| Temnothorax andrei    | interactsWith       | Quercus agrifolia        | 17    |
| Apis mellifera        | interactsWith       | Marrubium                | 16    |
| Araneidae             | interactsWith       | Cylindropuntia prolifera | 15    |
| Araneidae             | interactsWith       | Opuntia cf. littoralis   | 15    |
| Closterocoris amoenus | adjacentTo          | Lupinus nipomensis       | 15    |
| Apis mellifera        | interactsWith       | Erigonium                | 15    |
| Syrphidae             | interactsWith       | Lupinus bicolor          | 14    |
| Andrena principalis   | interactsWith       | Arctostaphylos           | 13    |
| Thrips                | adjacentTo          | Lupinus nipomensis       | 13    |

## 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.

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.

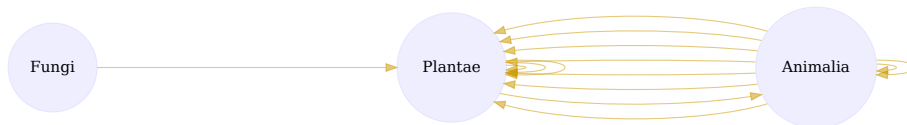


Figure 3: Interactions on taxonomic kingdom rank as interpreted by the Catalogue of Life download svg

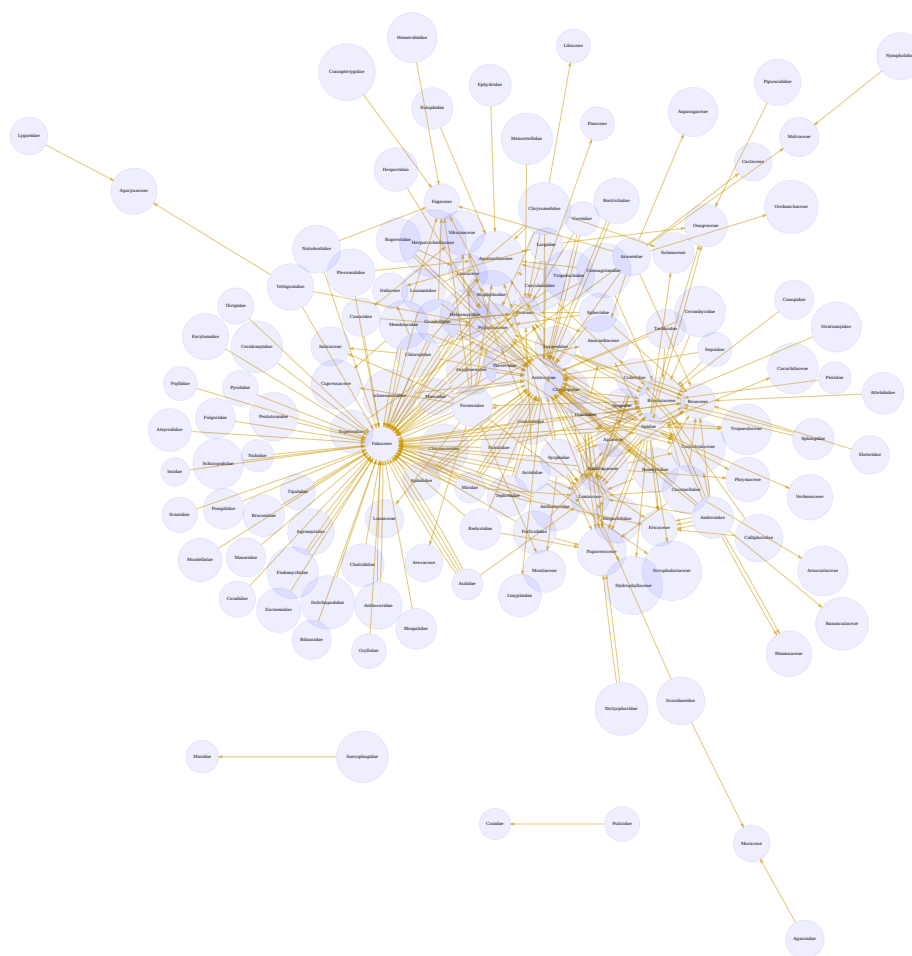


Figure 4: Interactions on the taxonomic family rank as interpreted by the Catalogue of Life. download svg

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            |
|-------------------------|-------------------|---------------------|-------------------------|
| Distichlis spicata      | HAS_ACCEPTED_NAME | col                 | Distichlis spicata      |
| Cryptonevra nigratarsis | HAS_ACCEPTED_NAME | col                 | Cryptonevra nigratarsis |
| Eulonchus               | HAS_ACCEPTED_NAME | col                 | Eulonchus               |
| Meromyza                | HAS_ACCEPTED_NAME | col                 | Meromyza                |

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           | 280   |
| col                 | class        | 1     |
| col                 | family       | 105   |
| col                 | genus        | 116   |
| col                 | kingdom      | 1     |
| col                 | order        | 12    |
| col                 | species      | 229   |
| col                 | subfamily    | 9     |
| col                 | subgenus     | 4     |
| col                 | suborder     | 2     |
| col                 | subspecies   | 8     |
| col                 | subtribe     | 1     |
| col                 | superfamily  | 4     |
| col                 | tribe        | 2     |
| col                 | variety      | 1     |
| discoverlife        | NA           | 697   |
| discoverlife        | species      | 64    |
| gbif                | NA           | 252   |
| gbif                | class        | 1     |



| resolvedCatalogName | resolvedRank | count |
|---------------------|--------------|-------|
| gbif                | family       | 108   |
| gbif                | genus        | 133   |
| gbif                | kingdom      | 1     |
| gbif                | order        | 12    |
| gbif                | species      | 250   |
| gbif                | subspecies   | 16    |
| gbif                | variety      | 8     |
| itis                | NA           | 263   |
| itis                | class        | 1     |
| itis                | family       | 105   |
| itis                | genus        | 103   |
| itis                | infraorder   | 2     |
| itis                | kingdom      | 1     |
| itis                | order        | 12    |
| itis                | species      | 236   |
| itis                | subclass     | 1     |
| itis                | subfamily    | 10    |
| itis                | suborder     | 8     |
| itis                | subspecies   | 13    |
| itis                | superfamily  | 4     |
| itis                | superorder   | 1     |
| itis                | variety      | 1     |
| mdd                 | NA           | 760   |
| ncbi                | NA           | 279   |
| ncbi                | class        | 1     |
| ncbi                | cohort       | 1     |
| ncbi                | family       | 106   |
| ncbi                | genus        | 115   |
| ncbi                | infraorder   | 3     |
| ncbi                | order        | 12    |
| ncbi                | species      | 217   |
| ncbi                | subclass     | 1     |
| ncbi                | subfamily    | 16    |
| ncbi                | subgenus     | 8     |
| ncbi                | suborder     | 3     |
| ncbi                | subspecies   | 3     |
| ncbi                | superfamily  | 6     |
| ncbi                | varietas     | 1     |
| pbdb                | NA           | 543   |
| pbdb                | class        | 2     |
| pbdb                | family       | 106   |
| pbdb                | genus        | 59    |
| pbdb                | infraclass   | 1     |
| pbdb                | infraorder   | 1     |

| resolvedCatalogName | resolvedRank   | count |
|---------------------|----------------|-------|
| pbdb                | kingdom        | 1     |
| pbdb                | order          | 13    |
| pbdb                | species        | 10    |
| pbdb                | subfamily      | 15    |
| pbdb                | suborder       | 5     |
| pbdb                | superfamily    | 6     |
| pbdb                | unranked clade | 3     |
| tpt                 | NA             | 756   |
| tpt                 | genus          | 1     |
| tpt                 | species        | 3     |
| wfo                 | NA             | 603   |
| wfo                 | family         | 1     |
| wfo                 | genus          | 48    |
| wfo                 | phylum         | 1     |
| wfo                 | species        | 106   |
| wfo                 | subspecies     | 2     |
| wfo                 | variety        | 1     |
| worms               | NA             | 551   |
| worms               | family         | 83    |
| worms               | genus          | 52    |
| worms               | kingdom        | 1     |
| worms               | order          | 11    |
| worms               | species        | 52    |
| worms               | subclass       | 2     |
| worms               | suborder       | 4     |
| worms               | subspecies     | 1     |
| worms               | superfamily    | 3     |

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 | 520   |
| col                 | SYNONYM_OF        | 79    |
| col                 | NONE              | 342   |
| discoverlife        | NONE              | 804   |
| discoverlife        | SYNONYM_OF        | 23    |

| resolvedCatalogName | relationName       | count |
|---------------------|--------------------|-------|
| discoverlife        | HAS_ACCEPTED_NAME  | 61    |
| discoverlife        | HOMONYM_OF         | 2     |
| gbif                | HAS_ACCEPTED_NAME  | 643   |
| gbif                | SYNONYM_OF         | 139   |
| gbif                | NONE               | 314   |
| itis                | HAS_ACCEPTED_NAME  | 520   |
| itis                | NONE               | 324   |
| itis                | SYNONYM_OF         | 35    |
| mdd                 | NONE               | 864   |
| mdd                 | HAS_ACCEPTED_NAME  | 1     |
| ncbi                | SAME_AS            | 527   |
| ncbi                | SYNONYM_OF         | 22    |
| ncbi                | NONE               | 344   |
| ncbi                | COMMON_NAME_OF     | 1     |
| pbdb                | NONE               | 622   |
| pbdb                | HAS_ACCEPTED_NAME  | 250   |
| pbdb                | SYNONYM_OF         | 16    |
| tpt                 | NONE               | 859   |
| tpt                 | HAS_ACCEPTED_NAME  | 6     |
| wfo                 | HAS_ACCEPTED_NAME  | 167   |
| wfo                 | NONE               | 687   |
| wfo                 | HAS_UNCHECKED_NAME | 9     |
| wfo                 | SYNONYM_OF         | 28    |
| worms               | HAS_ACCEPTED_NAME  | 223   |
| worms               | NONE               | 634   |
| worms               | SYNONYM_OF         | 19    |

Table 10: List of Available Name Alignment Reports

| catalog name | alignment results  |
|--------------|--|
| col          | associated names alignments (first 500,<br>full csv/tsv) |
| ncbi         | associated names alignments (first 500,<br>full csv/tsv) |
| discoverlife | associated names alignments (first 500,<br>full csv/tsv) |
| gbif         | associated names alignments (first 500,<br>full csv/tsv) |
| itis         | associated names alignments (first 500,<br>full csv/tsv) |
| wfo          | associated names alignments (first 500,<br>full csv/tsv) |

| catalog name | alignment results                                     |
|--------------|---|
| mdd          | associated names alignments (first 500, full csv/tsv) |
| tpt          | associated names alignments (first 500, full csv/tsv) |
| pbdb         | associated names alignments (first 500, full csv/tsv) |
| worms        | 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.

Table 11: First few lines in the review notes.

| reviewDate           | reviewCommentType | reviewComment   |
|----------------------|-------------------|---|
| 2024-07-23T01:26:39Z | note              | source taxon name missing: using institution-Code/collectionCode/collectionId/catalogNumber/occurrence as placeholder |
| 2024-07-23T01:26:39Z | note              | found unsupported interaction type with name: [flying]  |
| 2024-07-23T01:26:40Z | note              | source taxon name missing: using institution-Code/collectionCode/collectionId/catalogNumber/occurrence as placeholder |
| 2024-07-23T01:26:41Z | note              | found unsupported interaction type with name: [11]  |

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: using institution-Code/collectionCode/collectionId/catalogNumber/occurrenceId as placeholder | 12    |
| found unsupported interaction type with name: [11]  | 3     |
| found unsupported interaction type with name: [Hovering over]   | 3     |
| found unsupported interaction type with name: [flying]  | 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.



Figure 5: Picture of a GloBI Review Badge <sup>3</sup>

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

## GloBI Index Badge

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



Figure 6: Picture of a GloBI Index Badge <sup>4</sup>

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

<sup>4</sup>Up-to-date status of the GloBI Index Badge can be retrieved from GloBI's API

If you’d like to keep track of reviews or index status of the dataset under review, please visit [GloBI’s dataset index ^[At time of writing (2024-07-23) the version of the GloBI dataset index was available at <https://globalbioticinteractions.org/datasets> for badge examples.

## Discussion

This review aims to provide a perspective on the dataset to aid in understanding of species interaction claims discovered. However, it is important to note that this review does *not* assess the quality of the dataset. Instead, it serves as an indication of the open-ness<sup>5</sup> and FAIRness (Wilkinson et al. 2016; Trekels et al. 2023) of the dataset: to perform this review, the data was likely openly available, **F**indable, **A**ccessible, **I**nteroperable and **R**eusable. The current Open-FAIR assessment is qualitative, and a more quantitative approach can be implemented with specified measurement units.

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

## Author contributions

Nomer was responsible for name alignments. Elton carried out dataset extraction, and generated the review notes.

## 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/>.

---

<sup>5</sup>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.”

- 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 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>.