Global Biotic Interactions

A Catalyst for Integrating Existing Species-Interaction Datasets, Connecting Curators and Developing Data Exchange Methods

Global Biotic Interactions (GloBI) is a collaborative, open source, open data project that makes existing species-interaction datasets easier to discover and use.

http://globalbioticinteractions.org
“The advantage, and at the same time the difficulty, of ecological work is that it attempts to provide conceptions which can link up into some complete scheme the colossal store of facts about natural history which has accumulated up to date in this rather haphazard manner. […] Until more organised information about the subject is available, it is only possible to give a few instances of some of the more clear--cut niches which happen to have been worked out.”

Global Biotic Interactions

A Catalyst for Integrating Existing Species-Interaction Datasets, Connecting Curators and Developing Data Exchange Methods
Southern Sea Otter (Enhydra lutris nereis)
inaturalist.org/taxa/117520

Pacific rock crab (Romaleon antennarium)
inaturalist.org/taxa/202315

Specimen/Individual/Occurrence
classified_as

Specimen/Individual/Occurrence
classified_as

eats
RO_0002470
collected
2014-03-09 PDT

collected_at

Location
Lat: 36.713851
Lon: -121.960949

Study
https://www.inaturalist.org/observations/563486

in_dataset

inaturalist.org/taxa/117520

Dataset
https://github.com/globalbioticinteractions/inaturalist/archive/fae51d40d470e9adaa15b89bd2f1ba6c0a5c8fb0.zip

Simplified internal data model used by GloBI to integrate interaction data.
Southern Sea Otter (Enhydra lutris nereis)
inaturalist.org/taxa/117520

Pacific rock crab (Romaleon antennarium)
inaturalist.org/taxa/202315

Pacific rock crab (Romaleon antennarium)
gbif.org/species/5970317

Pacific rock crab (Romaleon antennarium)
marinespecies.org/aphia.php?p=taxdetails&id=440397

Simplified internal data model used by GloBI to integrate interaction data.
federated data registries

discover, import, aggregate and link

open read-only archives

zenodo

github social coding
data = code

an automated and continuous process
Bidirectional links include Encyclopedia of Life, Gulf of Mexico Species Interactions, NCBI Taxonomy, World Register of Marine Species, iNaturalist, Fishbase and SeaLifeBase.

Outgoing links include UBERON (body parts, life stage, physiological state), EnvO, GeoNames, CMECS, FEOW, MEO, TEO, doi.org, ITIS, Open Tree of Life, NBN and ALA.

Link services include Global Names and CrossRef.
federated data registries

peer review

an automated and continuous process
Just like organisms, datasets get born, grow up, reproduce and die. GloBI's mission is to help increase the productivity (or reuse) and lifespan of datasets before they meet their maker.
Eltonian shortfall*: a lack of species-interaction records

Eltonian shortfall*: a lack of species-interaction records

geospatial and temporal Eltonian shortfall

taxonomic Eltonian shortfall

Global Biotic Interactions

A Catalyst for Integrating Existing Species-Interaction Datasets, Connecting Curators and Developing Data Exchange Methods
Open discussions and reviews using GitHub issues

A life cycle database for parasitic acanthocephalans, cestodes, and nematodes by Benesh et al. #305

jhpoelen opened this issue on Aug 14 · 16 comments

jhpoelen commented on Aug 14 · edited

As suggested by @derele -

[...]
I am currently interacting with Dan Benesh who compiled a database on parasite life cycles (http://onlinelibrary.wiley.com/doi/10.1002/ecy.1680/full). Dan will work with my team in Berlin soon. He told me that he also wants to include this data in globi,

[...]
Open discussions and reviews using GitHub issues

(https://onlinelibrary.wiley.com/doi/10.1002/ecn.1680/full). Dan will work with my team in Berlin soon. He told me that he also wants to include this data in globi,

[...]

millerse self-assigned this on Aug 21

millerse added the new data label on Aug 21

dbenesh82 commented 28 days ago

Hi, I compiled this parasite life cycle database and have been meaning to get it into globi. If anything is unclear, I would be happy to help.

millerse commented 9 days ago

And done see here

https://github.com/jhpoelen/eol-globi-data/issues/305 accessed on 28 Sept 2017
Open discussions and reviews using GitHub issues

millerse commented 9 days ago
And done see here

millerse added the needs data review label 9 days ago

jhammock commented 9 days ago
Looks good! If you want to get fancy, you will probably find terms for some of the LocationinhostName values in UBERON. I wouldn't bother chasing elusive ones, but if you find an exact match to the string you have, that should be worth using.

cmungall commented 9 days ago
Scanned the 78 distinct values, fairly sure we have them all, I can do the mapping if you like

https://github.com/jhpoelen/eol-globi-data/issues/305 accessed on 28 Sept 2017
Open discussions and reviews using GitHub issues

jhammock commented 9 days ago

Fantastic! Yes please

cmungall commented 9 days ago

consider also mapping these to ENVO:

- Host.habitat: freshwater
- Host.habitat: marine
- Host.habitat: terrestrial

These to PATO:

- Shape: NA
- Shape: coiled
- Shape: cylinder
- Shape: ellipsoid

https://github.com/jhpoelen/eol-globi-data/issues/305 accessed on 28 Sept 2017
Open discussions and reviews using GitHub issues

jhammock commented 6 days ago

using endoparasite is easy enough. @dbenesh82 , can you verify that the definition and context displayed at http://purl.obolibrary.org/obo/RO_0002634 meet your needs? I just came from an ontology workshop and have discovered how easily definitions can be found incompatible...

dbenesh82 commented 4 days ago

Yes @jhammock, endoparasite is more appropriate for the interactions in this dataset.

millerse commented 2 days ago

Okay everyone. I have added the endoparasite link and the body terms to the sheet.

@dbenesh82 To your comment about the extra numbers: You had multiple body locations in a single record (i.e. gut, body cavity, etc.). I took each of those terms and made a single full record for each term. That is where the records come from.

https://github.com/jhpoelen/eol-globi-data/issues/305 accessed on 28 Sept 2017
Global Biotic Interactions

A Catalyst for Integrating Existing Species-Interaction Datasets, Connecting Curators and Developing Data Exchange Methods
Supports ~40 existing species-interaction data formats
Exports rdf, neo4j, tsv and darwin core (-ish) archives
Integrates with various ontologies and taxonomies

Rebuilt from sources automatically and continuously
Offline workflows via elton, elton-archive and nomer*

Offers open source and open access data
Encourages to share structured datasets
Provides a community hub for data curators and users

*see https://github.com/globalbioticinteractions/
Acknowledgments / funding
an incomplete list in no particular order

GloBI is not possible without the many contributions (big and small) of folks like Jen Hammock, Katja Schulz, Pepper Luboff, Chris Mungall, Katja Seltmann, Brian Hayden, Ken-ichi Ueda, Mariana Cains, Nuria Altimir, Srini Anand, William Liao, Sean Shiverick, Jim Simons, Theresa Mitchell, Emanuel Heitlinger, Marius Bäsler, Kathy Kwan, Deng Palomares, Josephine “Skit” Barile, Anne Thessen, Allen Hurlbert, Malcolm Storey… and thousands of others that have collected and shared species-interaction data.

GloBI has received funding from several projects since 2013. Those funding sources include, but are not limited to, the Encyclopedia of Life, EOL Rubenstein Fellows Program (CRDF EOL-33066-13/F33066, 2013) and the David M. Rubenstein Grant (FOCX-14-60988-1, 2014), and the Smithsonian Institution (SI) (T15CC10297-002, 2016).
Workshop Prompt

What can you do to help your colleagues and friends to share interactions datasets and make sure that they remain accessible (even after GloBI ceases to exist)?
Extra
AfricaTreeDatabase

```json
{
  "@context": ["http://www.w3.org/ns/csvw", {"@language": "en"}],
  "rdfs:comment": ["inspired by https://www.w3.org/TR/2015/REC-tabular-data-model-20151217/"],
  "headerRowCount": 1,
  "null": ["NULL", ""],
  "tableSchema": {
    "columns": [
      {
        "name": "interactionType");
      "titles": "intertype",
      "datatype": "string"
      },
      {
        "name": "basisOfRecordName",
        "titles": "obstype",
        "datatype": "string"
      },
      {
        "name": "effunit",
        "titles": "effunit",
        "datatype": "string"
      }
    ]
  }
}
```
install:
  - wget "https://raw.githubusercontent.com/content.com
  - chmod +x check-dataset.sh

script:
  - ./check-dataset.sh ${TRAVIS_REPO_SLUG}

AfricaTreeDatabase