Title: Global Biotic Interactions: Benefits of Pragmatic Reuse of Species Interaction Datasets

Author: Jorrit H. Poelen

Affiliation: Independent Open Software/Data Engineer in Biodiversity Informatics, Oakland, CA, USA

Abstract: Global Biotic Interactions (GloBI, https://globalbioticinteractions.org) uses frugal and pragmatic methods to make openly available species interaction datasets (e.g., parasite-host, predator-prey, plant-pollinator) easier to find and reuse. Since 2013, GloBI increased the reach of existing datasets, facilitated research, improved data integration methods and provided dataset reviews. In this talk, GloBI is introduced and various reuse examples are presented to discuss the question: Why should we bother to reuse existing (species-interaction) datasets?

Global Biotic Interactions

Benefits of Pragmatic Reuse of Species-Interaction Datasets

by Jorrit Poelen at IZW-Berlin 2019-01-09

Jorrit H. Poelen, James D. Simons and Chris J. Mungall. (2014). Global Biotic Interactions: An open infrastructure to share and analyze species-interaction datasets. *Ecological Informatics*. http://dx.doi.org/10.1016/j.ecoinf.2014.08.005.

Why reuse existing species-interaction datasets?

topics

> Context
> What is pragmatic reuse?
> Pragmatic reuse examples
> Q&A

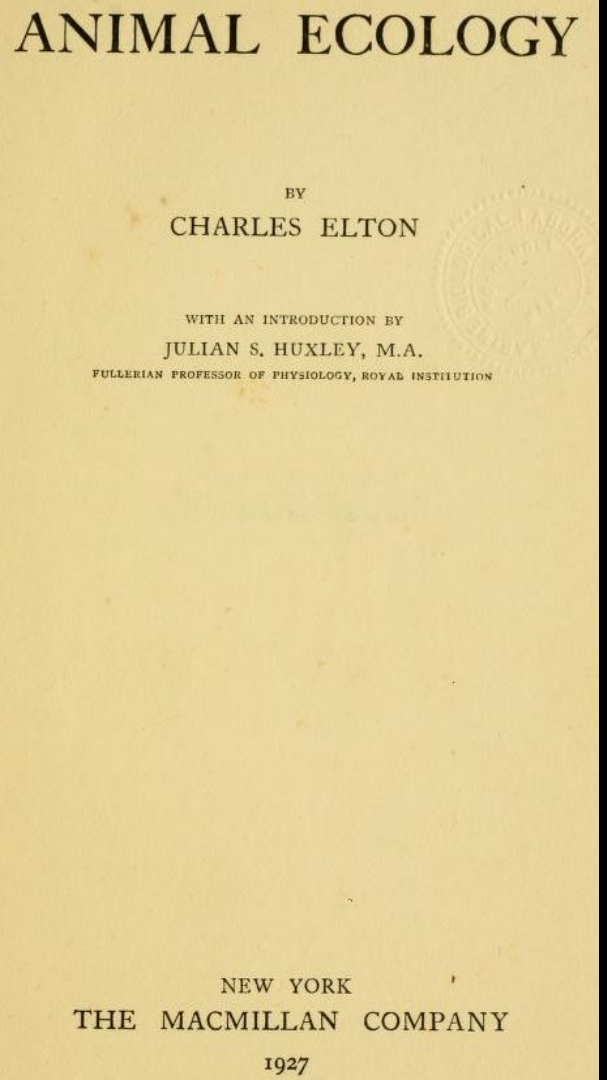
topics

> Context

> What is pragmatic reuse?

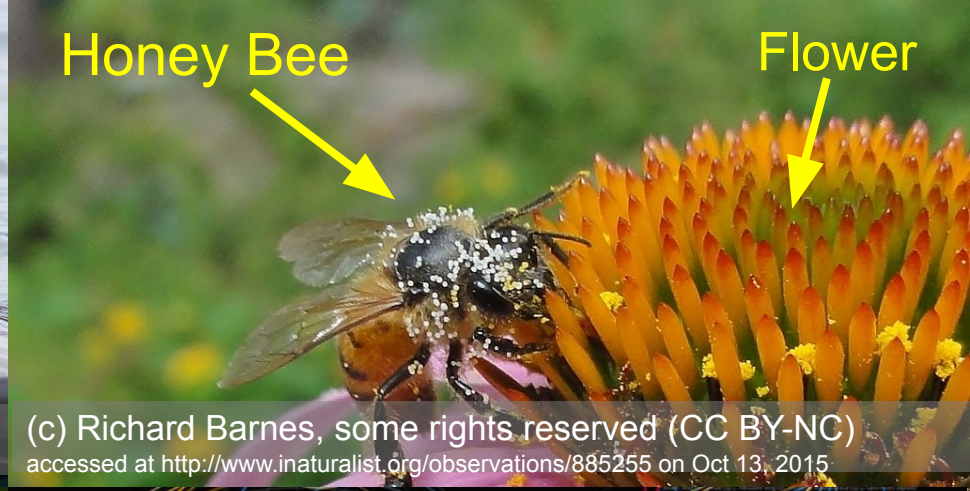
> Pragmatic reuse examples

> Q&A



“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.”

Charles Elton, **1927**, *Animal Ecology*.



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

Background image: Slyusarev et al. (2015): Global Biotic Interactions food web map. figshare. http://dx.doi.org/10.6084/m9.figshare.1297762

http://globalbioticinteractions.org

a brief history

2011 - Jorrit meets Jim at Texas A&M College Station for
. Ecological Integration Symposium.

2013 - Encyclopedia of Life Rubenstein Fellowship /
. GloBI / GoMexSI born

2014 - Jorrit, Jim and Chris publish GloBI paper

<many citations, workshops, conferences, integrations later>

2020 - Jorrit gives talk at IZW Berlin

in a nutshell

1. Existing openly accessible species interaction datasets in **any data format** are **registered** using https://globalbioticinteractions.org/contribute

2. **GloBI** continually and automatically **indexes and links** most recent species interaction datasets.

3. **Users discover, access and improve** datasets via GloBI's many integrations, search pages, data archives or APIs.

Location
Lat: 36.713851
Lon: -121.960949

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

Southern Sea Otter (Enhydra lutris nereis)
inaturalist.org/taxa/117520

collected_at

classified_as

classified_as

Specimen/ Individual/ Occurrence

eats
RO_0002470

Specimen/ Individual/ Occurrence

collected
2014-03-09 PDT

in_dataset

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

Dataset
https://github.com/globalbioticinteractions/inaturalist/archive/fae51d40d470e9adaa15b89bd2f1ba6c0a5c8fbd.zip

GloBI

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

Ecoregions

Environments

Location
Lat: 36.713851
Lon: -121.960949

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

Southern Sea Otter (Enhydra lutris nereis)
inaturalist.org/taxa/117520

Southern Sea Otter (Enhydra lutris nereis)
marinespecies.org/aphia.php?p=taxdetails&id=242601

Southern Sea Otter (Enhydra lutris nereis)
gbif.org/species/6163936

collected_at

classified_as

classified_as

Specimen/ Individual/ Occurrence

eats
RO_0002470

Specimen/ Individual/ Occurrence

collected
2014-03-09 PDT

in_dataset

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

Dataset
https://github.com/globalbioticinteractions/inaturalist/archive/fae51d40d470e9adaa15b89bd2f1ba6c0a5c8fbd.zip

GloBI

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

it takes a village to keep and share a record

Data Providers

Data Integration

GLOBI

Data Mobilizers

Data Users

Ontology & Standards Communities

Taxonomists

Ecologists

OBO Foundry

linking all the data

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, MEOW, TEOW, doi.org, ITIS, Open Tree of Life, NBN and ALA.

Link services include Global Names and CrossRef.

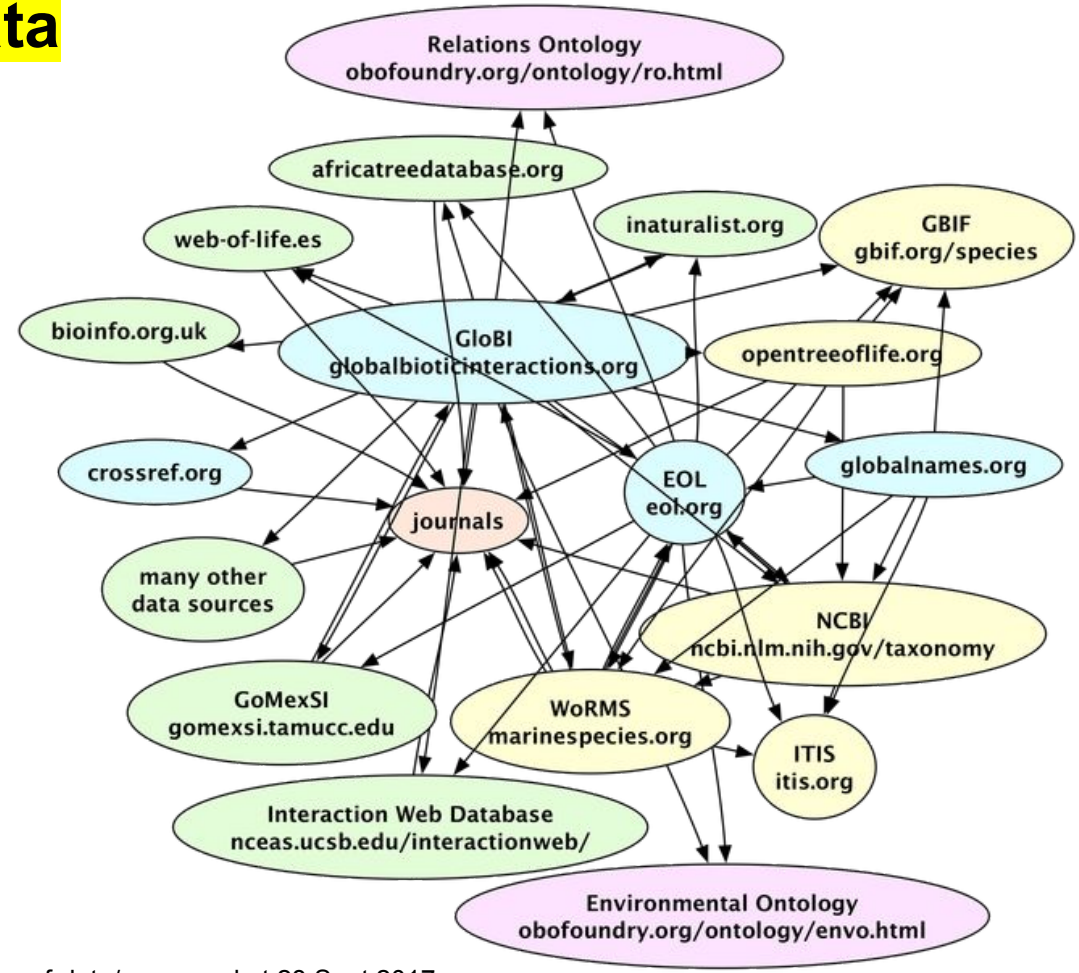


Figure from http://globalbioticinteractions.org/2016/02/08/ecology-of-data/ accessed at 28 Sept 2017

linking all the data and their curators

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, MEOW, TEOW, doi.org, ITIS, Open Tree of Life, NBN and ALA.

Link services include Global Names and CrossRef.

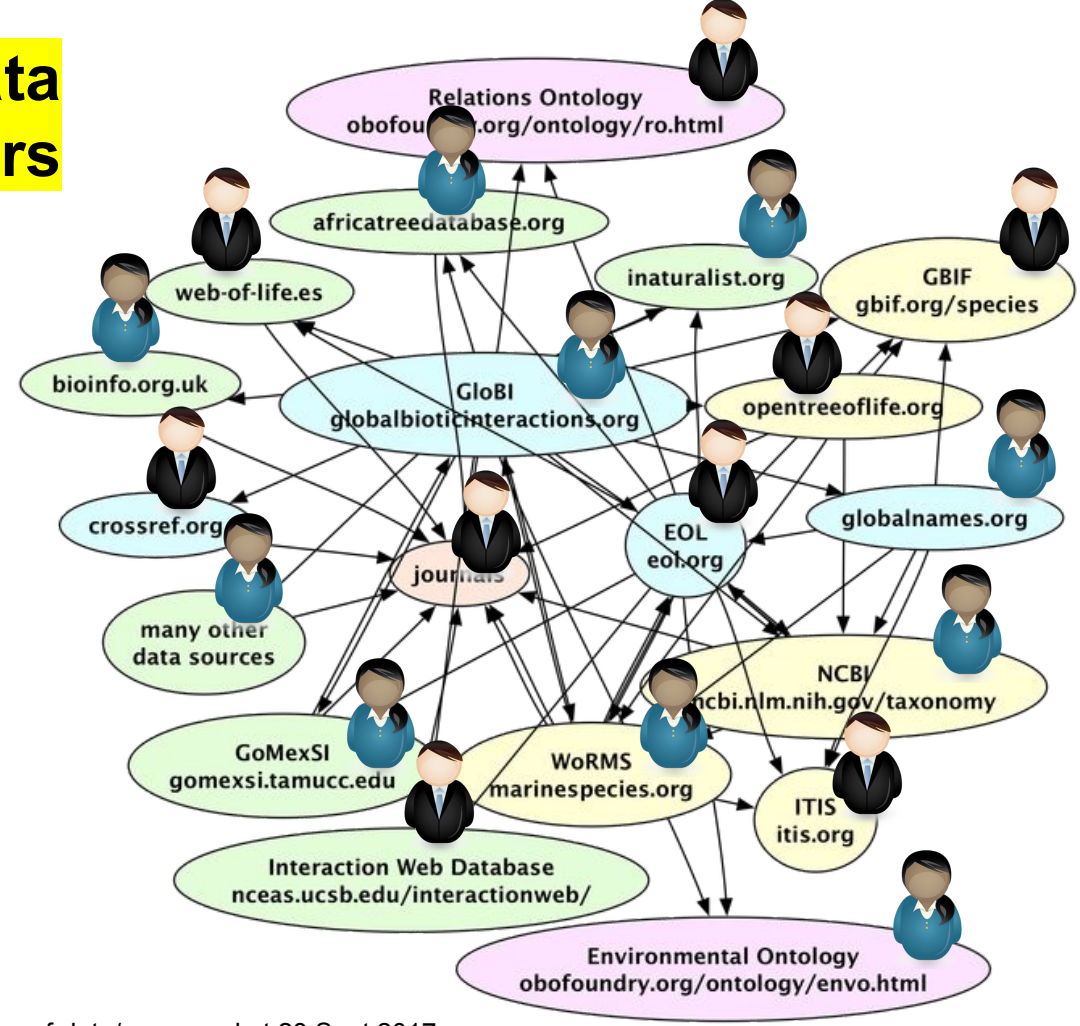
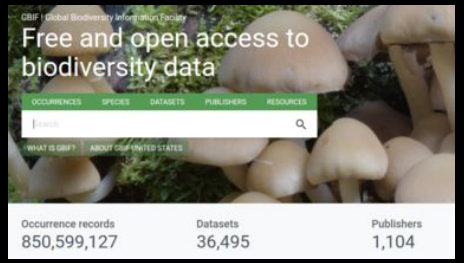


Figure from http://globalbioticinteractions.org/2016/02/08/ecology-of-data/ accessed at 28 Sept 2017

http://globalbioticinteractions.org/references
Accessed at 28 Sept 2017

GBIF Global Biodiversity Information Facility



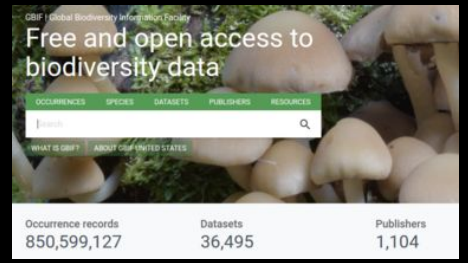
https://gbif.org
Accessed at 28 Sept 2017

2.8M records	950.6M records
0.1k datasets	36.5k datasets
~100k **taxa**	~1-2M **species**

Eltonian shortfall*: a lack of species-interaction records

*Hortal, J. et al., 2015. Seven Shortfalls that Beset Large-Scale Knowledge of Biodiversity. Annual Review of Ecology, Evolution, and Systematics, 46(1). Available at: http://dx.doi.org/10.1146/annurev-ecolsys-112414-054400.

http://globalbioticinteractions.org/references
Accessed at 26 Feb 2018



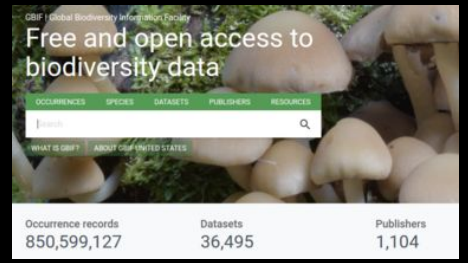
https://gbif.org
Accessed at 26 Feb 2018

3.4M records	972.7M records
0.1k datasets	38.1k datasets
~100k **taxa**	~1-2M **species**

Eltonian shortfall*: a lack of species-interaction records

*Hortal, J. et al., 2015. Seven Shortfalls that Beset Large-Scale Knowledge of Biodiversity. Annual Review of Ecology, Evolution, and Systematics, 46(1). Available at: http://dx.doi.org/10.1146/annurev-ecolsys-112414-054400.

http://globalbioticinteractions.org/references
Accessed at 9 Jan 2020



https://gbif.org
Accessed at 9 Jan 2020

4.4M records	**1.4B records**
0.2k datasets	50.0k datasets
~0.2M **taxa**	~1-2M **species**

Eltonian shortfall*: a lack of species-interaction records

*Hortal, J. et al., 2015. Seven Shortfalls that Beset Large-Scale Knowledge of Biodiversity. Annual Review of Ecology, Evolution, and Systematics, 46(1). Available at: http://dx.doi.org/10.1146/annurev-ecolsys-112414-054400.

topics

> Context

> What is pragmatic reuse?

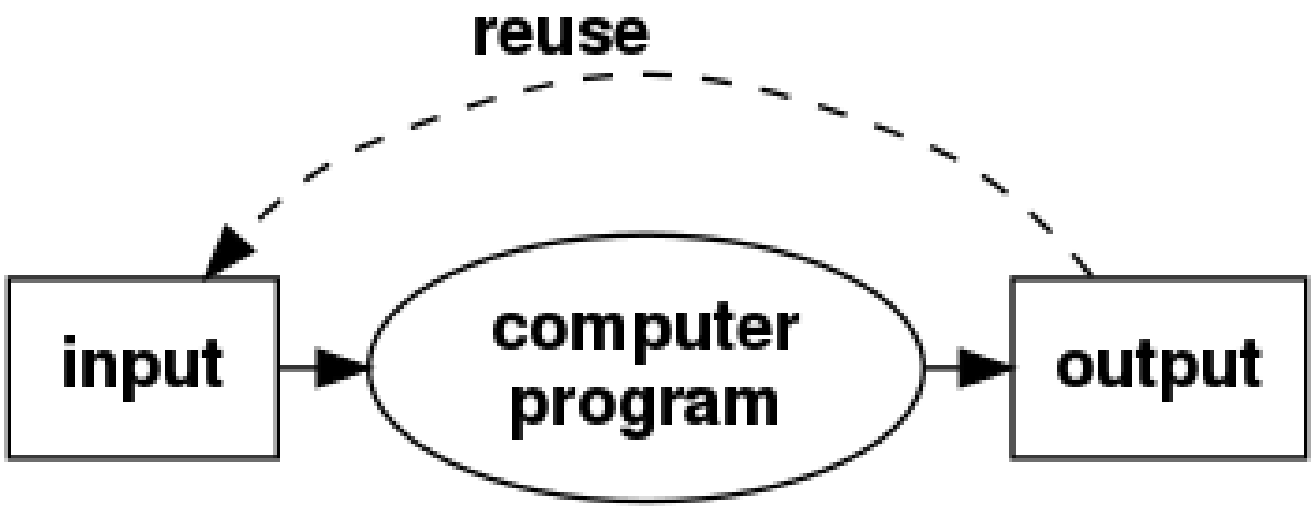
> Pragmatic reuse examples

> Q&A

Unix philosophy

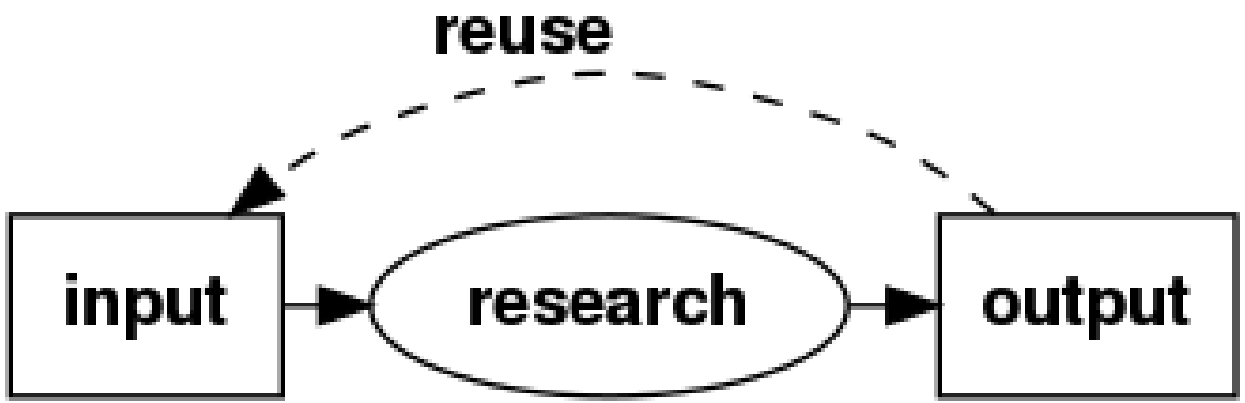
Expect the output of every program to become the input to another, as yet unknown, program.

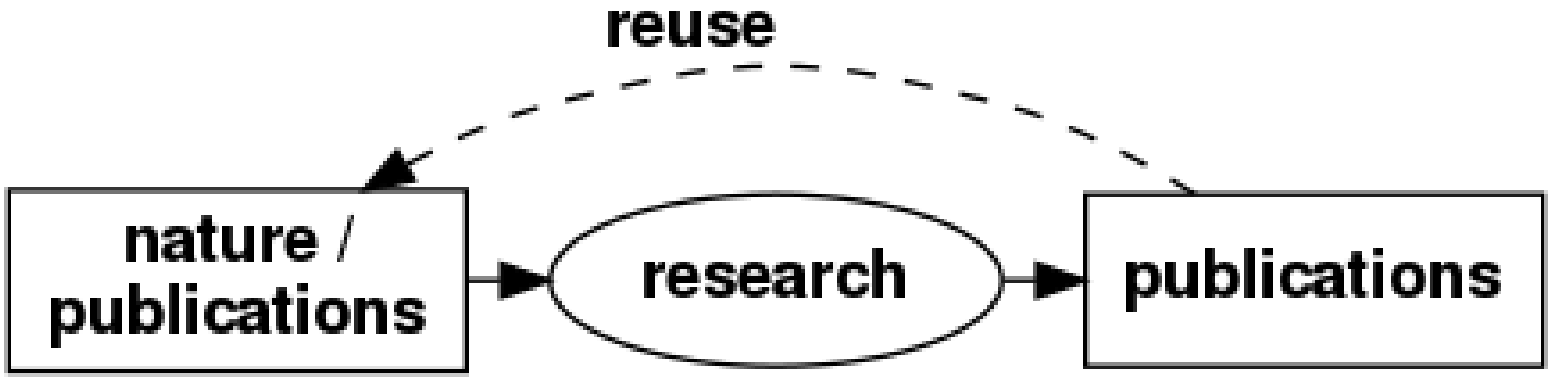
Doug McIlroy, E. N. Pinson, B. A. Tague (8 July 1978). "Unix Time-Sharing System: Foreword". The Bell System Technical Journal. Bell Laboratories. pp. 1902–1903.

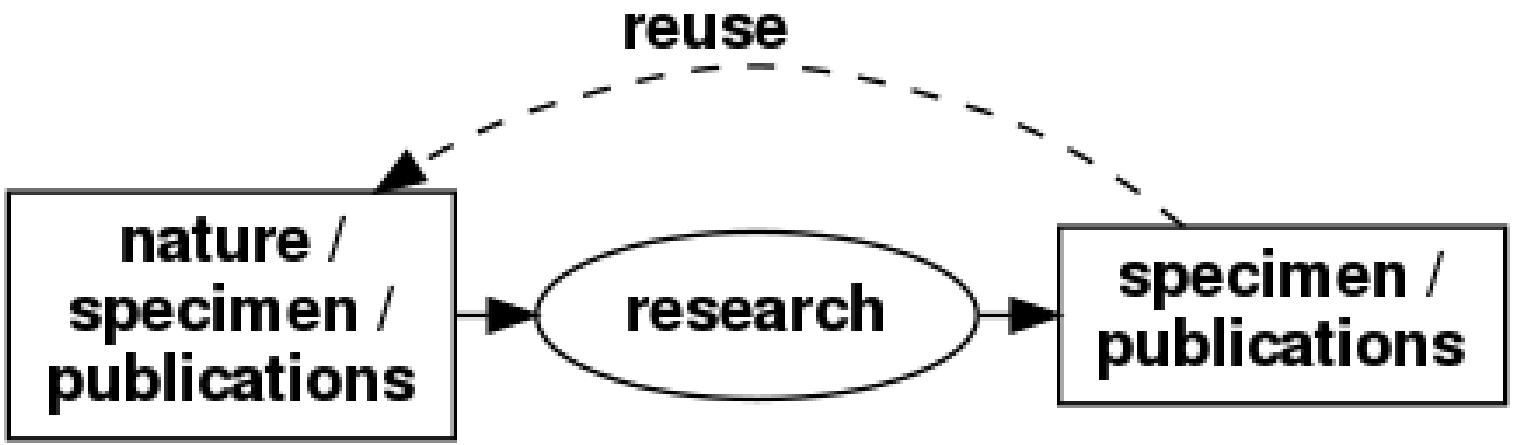

reuse
input
computer program
output

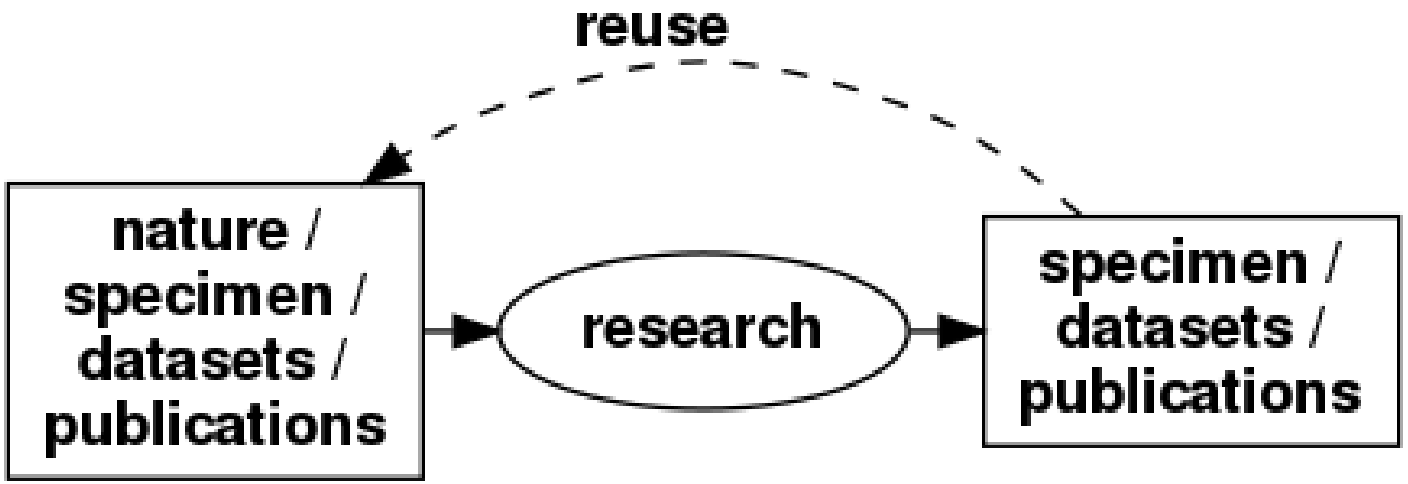
a research philosophy

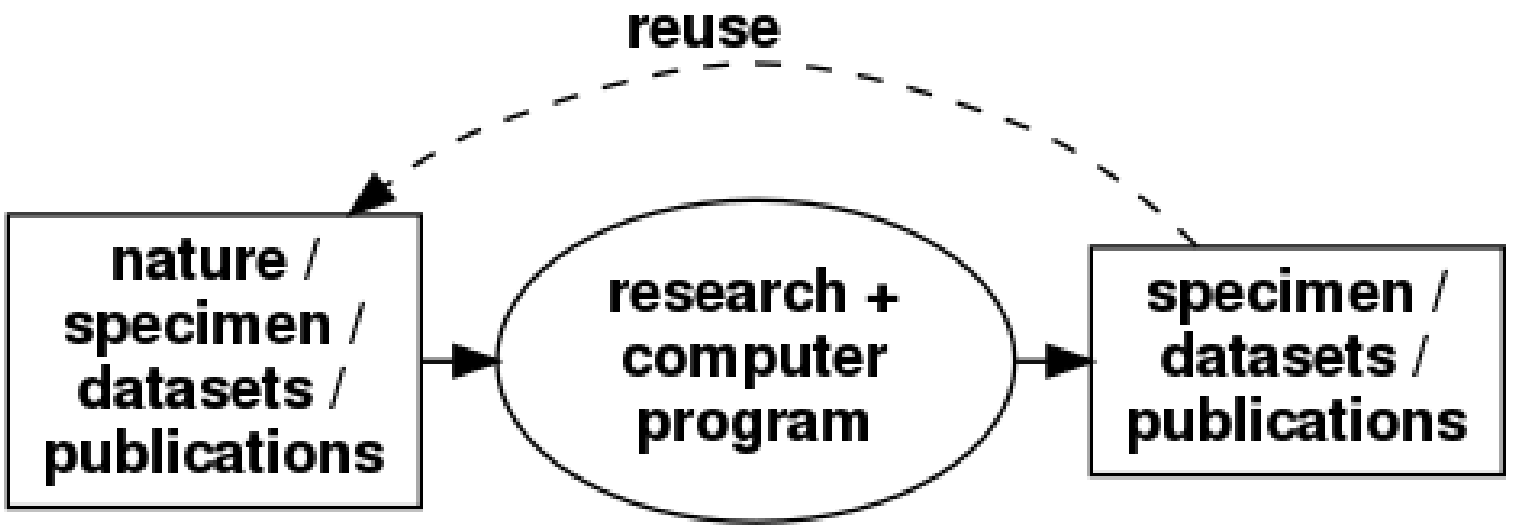
Expect the output of every ~~program~~ research project to become the input to another, as yet unknown, ~~program~~ research project.


reuse
input
research
output


reuse
nature / publications
research
publications


reuse
nature / specimen / publications
research
specimen / publications


reuse
nature / specimen / datasets / publications
research
specimen / datasets / publications


reuse
nature / specimen / datasets / publications
research + computer program
specimen / datasets / publications

topics

> Context

> What is pragmatic reuse?

> Pragmatic reuse examples

> Q&A

pragmatic reuse examples

> Research and Reviews
> Terrestrial Parasite Tracker TCN
> CCBER at UC Santa Barbara
> Arctos Collection Management
> Automated Dataset Reviews
> Reliable Dataset References

pragmatic reuse examples

> Research and Reviews

> Terrestrial Parasite Tracker TCN

> CCBER at UC Santa Barbara

> Arctos Collection Management

> Automated Dataset Reviews

> Reliable Dataset References

MENU SCIENTIFIC REPORTS

Article | Open Access | Published: 06 August 2019

Biological and environmental drivers of trophic ecology in marine fishes – a global perspective

B. Hayden, M. L. D. Palomares, B. E. Smith & J. H. Poelen

Scientific Reports **9**, Article number: 11415 (2019) | Cite this article

ANNUAL REVIEWS
JOURNALS A-Z | JOURNAL INFO | PRICING & SUBSCRIPTIONS | SUBS

Seven Shortfalls that Beset Large-Scale Knowledge of Biodiversity

Annual Review of Ecology, Evolution, and Systematics
Vol. 46:523-549 (Volume publication date December 2015)
First published online as a Review in Advance on October 28, 2015
https://doi.org/10.1146/annurev-ecolsys-112414-054400

Joaquín Hortal,[1,2,3,*] Francesco de Bello,[4,5] José Alexandre F. Diniz-Filho,[2] Thomas M. Lewinsohn,[6] Jorge M. Lobo,[1] and Richard J. Ladle[7,8,*]

ECOGRAPHY
AN OPEN ACCESS JOURNAL FROM 1 JANUARY 2020

Research | Open Access

Exploring the temporal variability of a food web using long-term biomonitoring data

Pierre Olivier, Romain Frelat, Erik Bonsdorff, Susanne Kortsch, Ingrid Kröncke, Christian Möllmann, Hermann Neumann, Anne F. Sell, Marie C. Nordström

First published: 20 September 2019 | https://doi.org/10.1111/ecog.04461

ECOLOGY LETTERS

Review and Synthesis | Free Access

Pyramids and cascades: a synthesis of food chain functioning and stability

Matthieu Barbier, Michel Loreau

First published: 17 December 2018 | https://doi.org/10.1111/ele.13196 | Citations: 5

MENU nature communications

Article | Open Access | Published: 29 November 2019

7-Deazaguanine modifications protect phage DNA from host restriction systems

Geoffrey Hutinet, Witold Kot, Liang Cui, Roman Hillebrand, Seetharamsingh Balamkundu, Shanmugavel Gnanakalai, Ramesh Neelakandan, Alexander B. Carstens, Chuan Fa Lui, Denise Tremblay, Deborah Jacobs-Sera, Mandana Sassanfar, Yan-Jiun Lee, Peter Weigele, Sylvain Moineau, Graham F. Hatfull, Peter C. Dedon, Lars H. Hansen & Valérie de Crécy-Lagard

Nature Communications **10**, Article number: 5442 (2019) | Cite this article

pragmatic reuse examples

> Research and Reviews

> Terrestrial Parasite Tracker TCN

> CCBER at UC Santa Barbara

> Arctos Collection Management

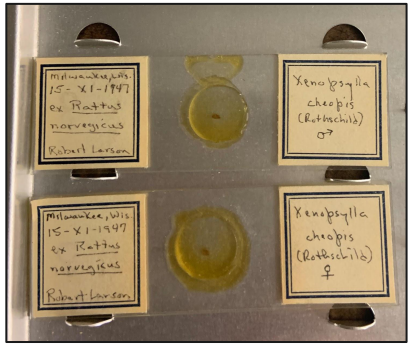
> Automated Dataset Reviews

> Reliable Dataset References

Digitization TCN: Terrestrial Parasite Tracker: Digitizing collections to trace parasite-host associations and predict the spread of vector-borne disease

Kathryn Sullivan*
Parasite Tracker TCN Project Manager
Invertebrate Zoology, Milwaukee Public Museum

Jennifer M. Zaspel
Associate Curator and Head of Zoology Milwaukee Public Museum
Director Puelicher Butterfly Wing
Adjunct Associate Professor Purdue University



- **Project Title:** Collaborative Research: Digitization TCN: Digitizing collections to trace parasite-host associations and predict the spread of vector-borne disease
- **PIs:** Jen Zaspel & Stephen Cameron
- **Project Start Date:** September 1st 2019
- **Project Period:** 3 years
- **Participating Institutions:** 27
- **Co-PIs/Leads on Subs:** 34 (*12/17)
- **Participants:** 60+

Stephen Cameron, TPT Co-Lead

Jen Zaspel, TPT Co-Lead

PARASITE TRACKER

DOCUMENTING ARTHROPOD VERTEBRATE PARASITES

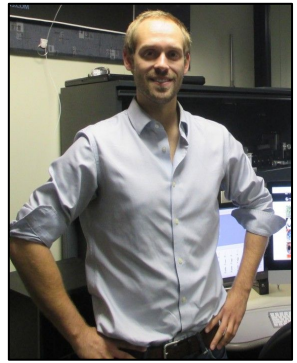
Neil Cobb, NAU

Rob Guralnick, UF

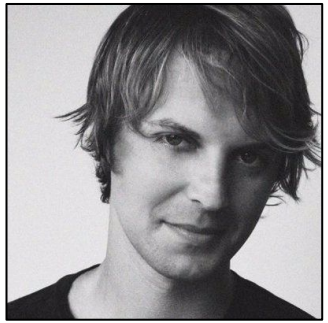
Katja Seltmann, UCSB

Julie Allen, UNR

Jorrit Poelen, GloBI



Mark Smith, MS LLC



André Poremski, Fg

Anna Monfils, BLUE

Alyssa Caywood, MPM

Chris Tyrell, MPM

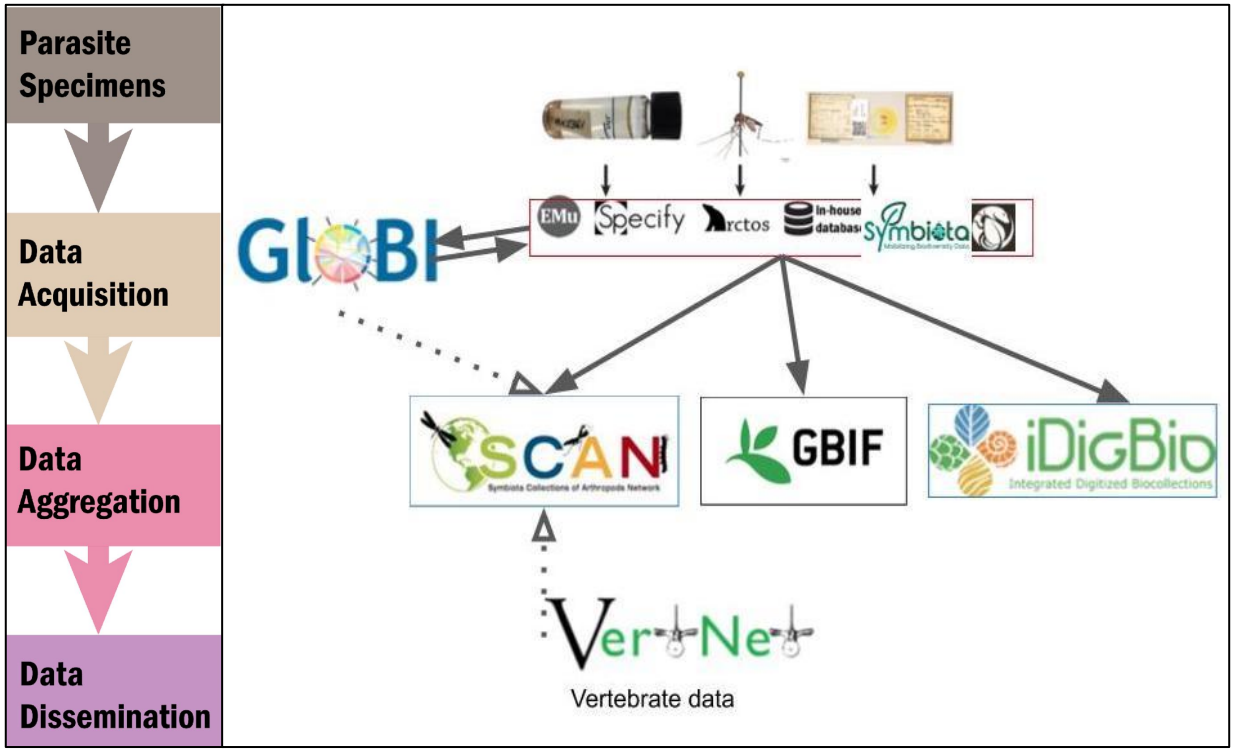
Julia Colby, MPM

Barb Thiers, NYBG

Matt Yoder, TaxonWorks

Specimen Digitization

- Transcribe and georeference label data from **1.2+** million arthropod parasite specimens from 22 collections across North America (U.S. and territories) including ~55,000 specimens from biotic-association collections

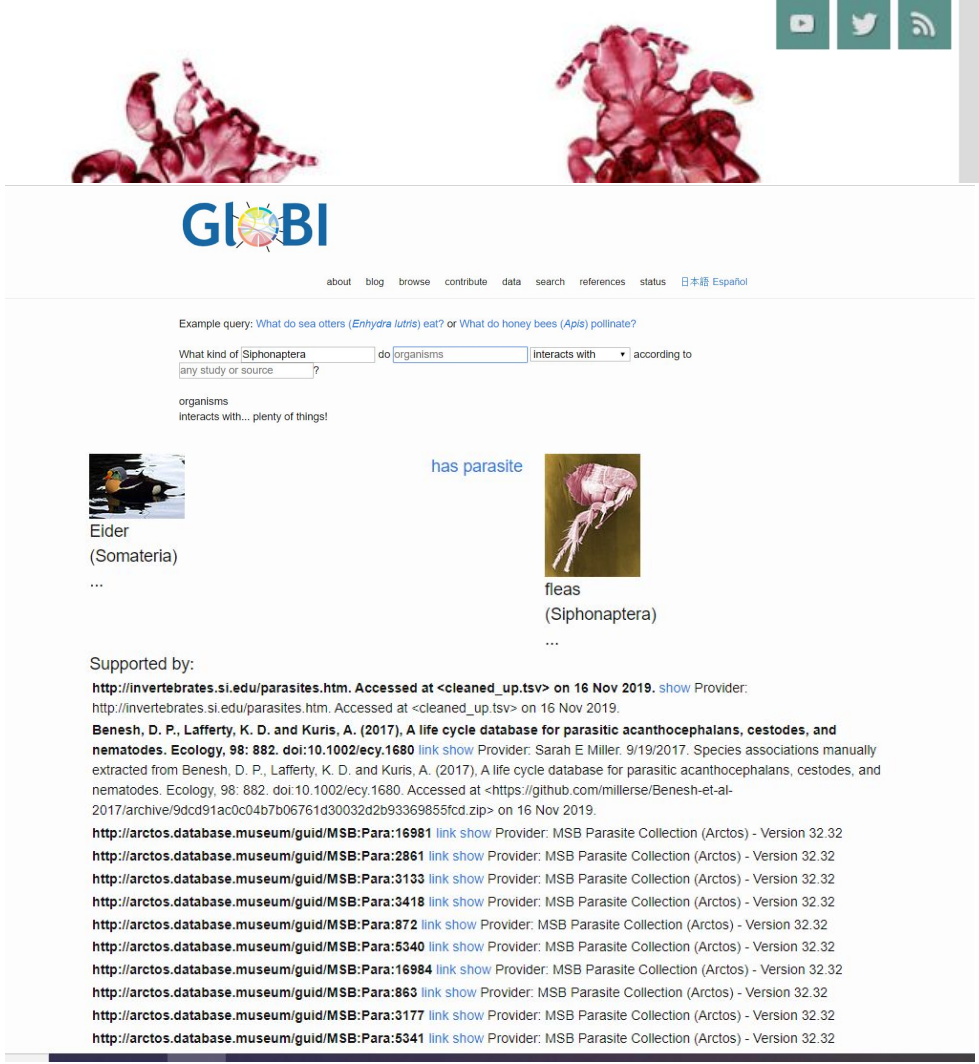


Data Integration

Global Biotic Interactions (GloBI) is a data integration tool that indexes existing species interaction datasets, literature, and specimen records in collections.

https://www.globalbioticinteractions.org/

Specimen data transcribed for the TPT project will generate 500,000 new parasite-host association records in GloBI.



pragmatic reuse examples

> Research and Reviews
> Terrestrial Parasite Tracker TCN
> CCBER at UC Santa Barbara
> Arctos Collection Management
> Automated Dataset Reviews
> Reliable Dataset References

North Campus Open Space Restoration Project

NCOS NEWS *December 2019*

Recent rains have filled up the Slough at NCOS. Photo taken on December 2nd.

UCSB Natural History Collection News

Introduction to the Identification of California Bees Workshop

Introduction to the Identification of California Bees

Learn how to identify our diverse native bees in this three day workshop.

January 18-20th with instructor Jaime Pawelek, CA Native Bees, Consultant specializing in the taxonomic identification of native bees

CARPINTERIA SALT MARSH

Protects a critically important estuarine ecosystem that supports many sensitive plan and animal species.

COAL OIL POINT

Includes coastal dune and wetland ecosystems that provide critical habitat for migratory birds and threatened and endangered species.

K.S.N. RANCHO MARINO

Has some of the most spectacular stretches of coast in central California and one of the few remnants of Monterey pine forest left in the state.

SANTA CRUZ ISLAND

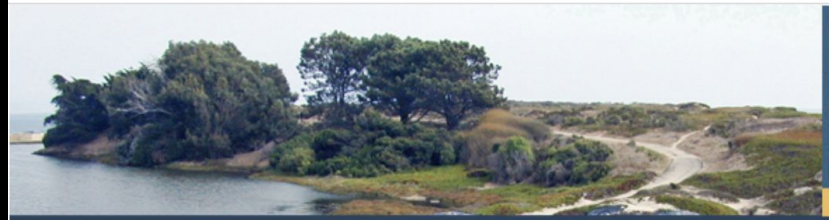
The largest of the Channel Islands and the site of important archaeological resources and many endemic plant and animal species including the island fox.

SEDGWICK

Over nine square miles at the base of Figueroa Mt. encompassing a diverse array of geological features and a wide variety of vegetation types and wildlife.

VALENTINE EASTERN SIERRA

Comprised of two reserves, the Sierra Nevada Aquatic Research Laboratory and Valentine Reserve, with excellent facilities to support research and teaching throughout the region.



University of California Santa Barbara Collection Network

- Insects of UCSB and Nearby
- Insects of North Campus Open Space
- Bees of UCSB and Nearby
- Ichneumonoidea of UCSB and Nearby
- Ants of Santa Barbara County
- Ophioninae of Coastal California

Santa Cruz Islan

Games

Authors:

Families: 67
Genera: 187
Species: 253 (species rank)
Total Taxa: 258 (including subsp. and var.)

ACRIDIDAE

Schistocerca nitens (Thunberg, 1815)
Liam O'Brien 2015-07-21 [SCIR]

Trimerotropis pallidipennis (Burmeister, 1838)
Alex Bairstow 2017-05-20 [SCIR] , jmaley 2017-10-21 [SCIR] , James Bailey 2017-05-18 [SCIR]

Trimerotropis pseudofasciata Scudder, S.H., 1876
James Bailey 2017-05-18 [SCIR]

AESHNIDAE

Rhionaeschna multicolor (Hagen, 1861)
Joan Ball 2012-08-19 [SCIR] , Liam O'Brien 2015-07-24 [SCIR] , Joan Ball 2012-08-19 [SCIR] , Liam O'Brien 2015-07-22 [SCIR] , more...

ANDRENIDAE

Andrena angustitarsata Viereck, 1904
D. R. Miller 1969-04-29 [SCIR] , B. J. Donovan 1969-04-27 [SCIR] , B. J. Donovan 1969-04-27 [SCIR] , R. W. Thorp 1971-03-08 [SCIR] , more...

Andrena candida Smith, 1879
J. S. Ascher, R. W. Thorp 2008-04-09 [SCIR] , J. S. Ascher, R. W. Thorp 2008-04-09 [SCIR] , J. S. Ascher, R. W. Thorp 2008-04-08 [SCIR] , J. S. Ascher, R. W. Thorp 2008-04-08 [SCIR] , more...

Andrena chlorogaster Viereck, 1904
UCSBees Survey 2019-04-15 [SCIR] , R. O. Schuster 1971-03-11 [SCIR] , R. O. Schuster 1971-03-11 [SCIR] , R. O. Schuster 1971-03-11 [SCIR] , more...

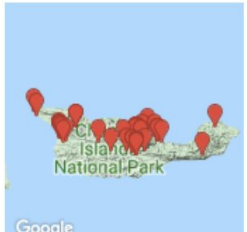
Andrena oenotherae Timberlake, 1937
J. S. Ascher, R. W. Thorp 2008-04-10 [SCIR] , J. S. Ascher, R. W. Thorp 2008-04-10 [SCIR] , J. S. Ascher, R. W. Thorp 2008-04-10 [SCIR] , J. S. Ascher, R. W. Thorp 2008-04-10 [SCIR] , more...

Options

Search:
- Common Names
- Synonyms

Filter: Original Checklist
- Common Names
- Display as Images
- Notes & Vouchers
- Taxon Authors
- Show Taxa Alphabetically

Rebuild List



Santa Cruz Island Reserve Invertebrates *Games*

Authors:

Families: 67
Genera: 187
Species: 253 (species rank)
Total Taxa: 258 (including subsp. and var.)

ACRIDIDAE

Schistocerca nitens (Thunberg, 1815)
Liam O'Brien 2015-07-21 [SCIR]

Trimerotropis pallidipennis (Burmeister, 1838)
Alex Bairstow 2017-05-20 [SCIR] , jmaley 2017-10-21 [SCIR] , James Bailey 2017-05-18 [SCIR]

Trimerotropis pseudofasciata Scudder, S.H., 1876
James Bailey 2017-05-18 [SCIR]

AESHNIDAE

Rhionaeschna multicolor (Hagen, 1861)
Joan Ball 2012-08-19 [SCIR] , Liam O'Brien 2015-07-24 [SCIR] , Joan Ball 2012-08-19 [SCIR] , Liam O'Brien 2015-07-22 [SCIR] , more...

ANDRENIDAE

Andrena angustitarsata Viereck, 1904
D. R. Miller 1969-04-29 [SCIR] , B. J. Donovan 1969-04-27 [SCIR] , B. J. Donovan 1969-04-27 [SCIR] , R. W. Thorp 1971-03-08 [SCIR] , more...

Andrena candida Smith, 1879
J. S. Ascher, R. W. Thorp 2008-04-09 [SCIR] , J. S. Ascher, R. W. Thorp 2008-04-09 [SCIR] , J. S. Ascher, R. W. Thorp 2008-04-08 [SCIR] , J. S. Ascher, R. W. Thorp 2008-04-08 [SCIR] , more...

Andrena chlorogaster Viereck, 1904
UCSBees Survey 2019-04-15 [SCIR] , R. O. Schuster 1971-03-11 [SCIR] , R. O. Schuster 1971-03-11 [SCIR] , R. O. Schuster 1971-03-11 [SCIR] , more...

Andrena oenotherae Timberlake, 1937
J. S. Ascher, R. W. Thorp 2008-04-10 [SCIR] , J. S. Ascher, R. W. Thorp 2008-04-10 [SCIR] , J. S. Ascher, R. W. Thorp 2008-04-10 [SCIR] , J. S. Ascher, R. W. Thorp 2008-04-10 [SCIR] , more...

Andrena piperi Viereck, 1904
UCSBees Survey 2019-04-15 [SCIR] , R. W. Thorp 1971-03-11 [SCIR] , R. O. Schuster 1971-03-11 [SCIR] , R. O. Schuster 1971-03-11 [SCIR] , more...

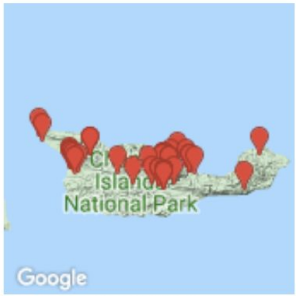
Andrena principalis LaBerge, 1986

Options

Search:
- [ ] Common Names
- [x] Synonyms

Filter:
Original Checklist
- [ ] Common Names
- [ ] Display as Images
- [x] Notes & Vouchers
- [x] Taxon Authors
- [ ] Show Taxa Alphabetically

Rebuild List



Simple Map
Advanced Map

Datasets 1

Search

Publisher

- [x] Cheadle Center for Biodiversity and Ecological Restoration

Host

Publishing country or area

Project

License

University of California Santa Barbara Algae Collection

Occurrence dataset

The UCSB botanical collections, housed at the Cheadle Center for Biodiversity and Ecological Restoration, include over 250,000 taxa of terrestrial and marine species. The vascular plant herbarium incl...

Published by Cheadle Center for Biodiversity and Ecological Restoration

6,498 occurrences

University of California Santa Barbara Invertebrate Zoology Collection

Occurrence dataset

University of California Santa Barbara Invertebrate Zoology Collection, Cheadle Center for Biodiversity and Ecological Restoration

Published by Cheadle Center for Biodiversity and Ecological Restoration

27,758 occurrences 3 citations

UCSB - University of California Santa Barbara Herbarium

Occurrence dataset

The University of California, Santa Barbara (UCSB) Herbarium has approximately 120,000 herbarium specimens of vascular plants, lichens, bryophytes, and marine macroalgae. The herbarium is housed at th...

Published by Cheadle Center for Biodiversity and Ecological Restoration

44,838 occurrences

Bird observations of UC Santa Barbara North Campus Open Space

Occurrence dataset

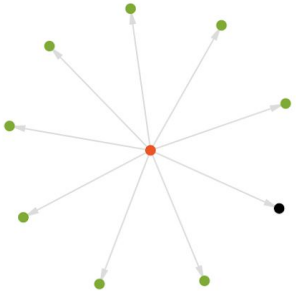
No description available

Published by Cheadle Center for Biodiversity and Ecological Restoration

Colletes ×

interacts with

		download details
taxon	interacts with	taxon
(1 distinct)	(9 interactions)	(9 distinct)
Colletes	interacts with	Phacelia tanacetifolia
Colletes	interacts with	Eriogonum
Colletes	interacts with	Foeniculum
Colletes	visits flowers of	Eriogonum parvifolium



GloBI

Berkeley Natural History Museums ReserveMapper

Map | Table | Statistics

Select University of California Natural Reserve

Santa Cruz Island Reserve

Query Type

(•) Occurrence ( ) Calphotos ( ) Checklists

Search Any Term

Colletes hyalinus gaudialis

Search by Taxonomy

SPECIES

Colletes hyalinus Provancher, 1888

Colletes hyalinus Provancher, 1888

Year Collected

YYYY to YYYY

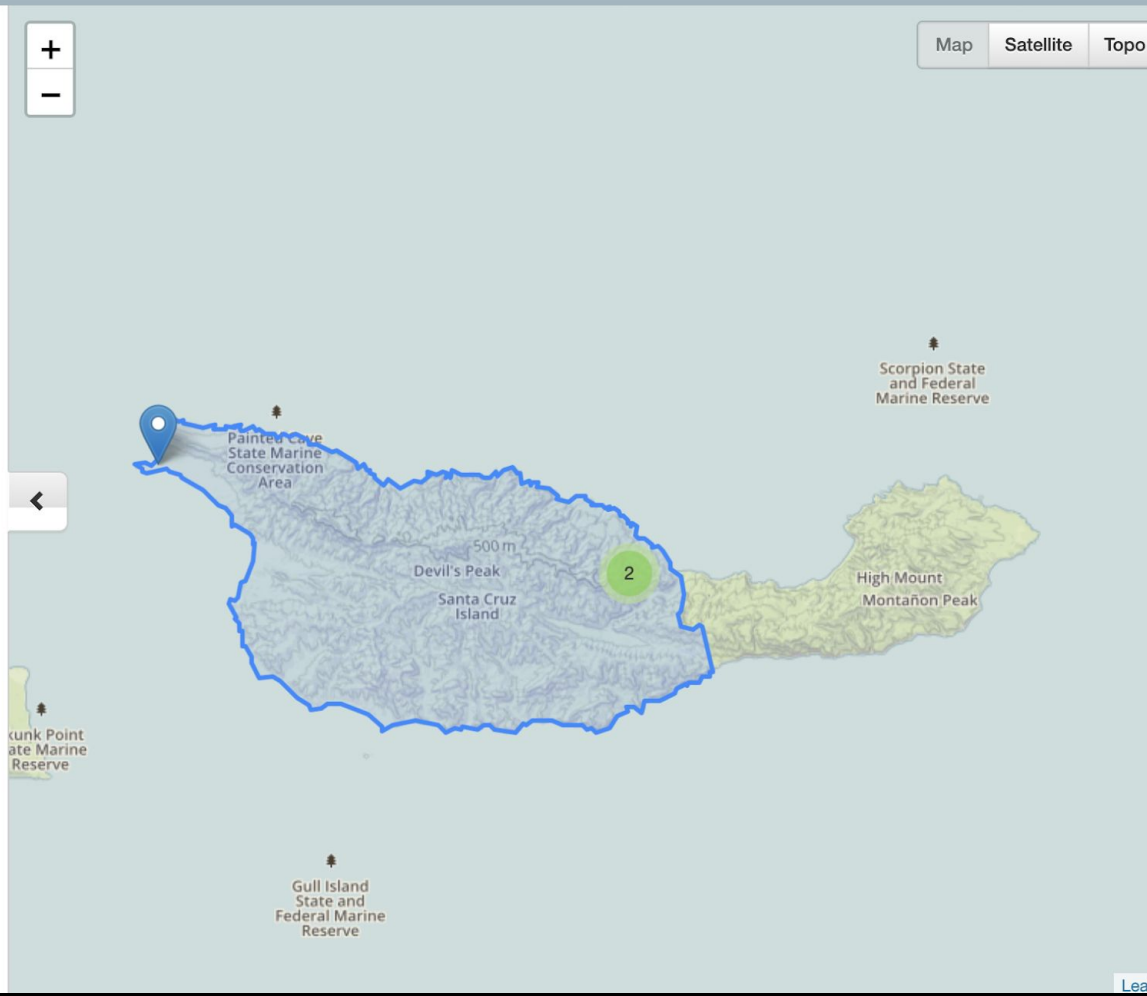
Darwin Core Terms

institutionCode collectionCode

-- Basis of Record --
PRESERVED_SPECIMEN
FOSSIL_SPECIMEN
LIVING_SPECIMEN

Search | Clear | Download

ReserveMapper performs a spatial-bounding box search on coordinates acquired through harvested data. Observations without coordinates, or that are purposefully obscured (e.g. for endangered species) or that have low accuracy may not appear in ReserveMapper results. Occurrence data are harvested from



Resources 1

Search

Country or area of researcher

Country or area of coverage

Literature type

Relevance

Year

Topic

Dataset

- [x] University of California Santa Barbara Invertebrate Zoology Collection

Publisher

Peer-reviewed

Open access

Download key

SEARCH RESOURCES | 3 RESULTS

ALL LITERATURE

Read more about literature, how it's discovered and linked to GBIF-mediated data.

Global Bee Decline

Literature

Zattara, E. Aizen, M. (2019) bioRxiv

Wild and managed bees are key pollinators, providing ecosystem services to a large fraction of the world's flowering plants, including ~85% of all cultivated crops. Recent reports of wild bee decline and its potential consequences are thus worrisome. However, evidence is mostly based on local or reg...

Working paper Open access

Data referenced in study DOI 10.15468/dl.h52qyh DOI 10.15468/dl.o73fzx

Evaluating the data quality of iNaturalist termite records

Literature

Hochmair, H. Scheffrahn, R. Basille, M. Boone, M. (2019) BioRxiv

Citizen science (CS) contributes to the combined knowledge about species distributions, which is a critical foundation in the studies of invasive species, biological conservation, and response to climatic change. In this study, we assessed the value of CS for termites worldwide. First, we compared t...

Working paper Open access

Data referenced in study DOI 10.15468/dl.0qi0rl

Insights into 260 years of taxonomic research gained from the Catalogue of Afrotropical Bees

Literature

Colletes hyalinus gaudialis Cockerell, 1905

Go to Encyclopedia of Life...
Source: ITIS, Bugguide
Family: Colletidae

Description #1

Host plants: Apiaceae Sanicula marilandica

Asteraceae Aster sp., Erigeron speciosus, Grindelia squarrosa, Hypochaeris radicata

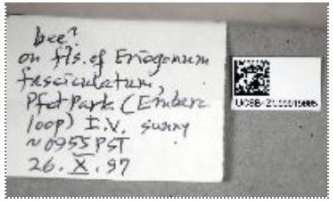
Caprifoliaceae Symphoricarpos occidentalis

Fabaceae Psorothamnus polydenius

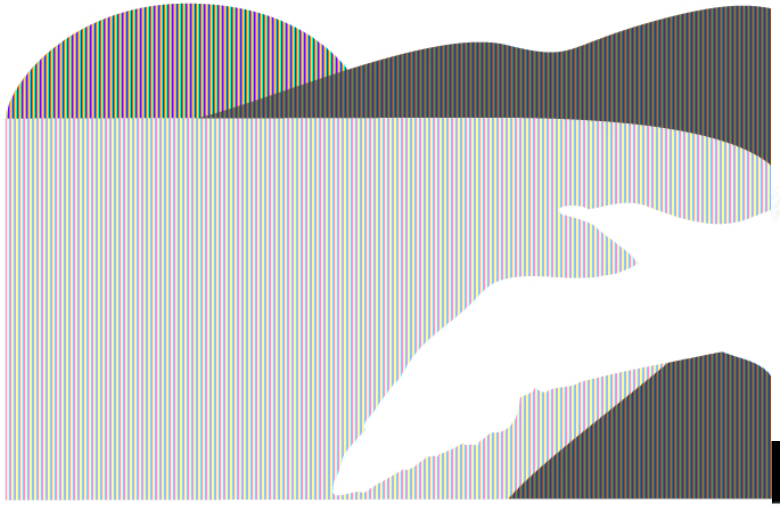
Geraniaceae Geranium richardsonii

Rosaceae Pentaphylloides floribunda

Tamaricaceae Tamarix sp.



Open Interactive Map



COASTAL FU
UC SANTA BARB
ASSOCIATED STUD
WWW.COASTALFUND.

INSTITUTE of
Museum and Library
SERVICES

ERI
Earth Research Institute

UNIVERSITY OF CALIFORNIA
SANTA BARBARA
LET THERE BE LIGHT

pragmatic reuse examples

> Research and Reviews
> Terrestrial Parasite Tracker TCN
> CCBER at UC Santa Barbara
> Arctos Collection Management
> Automated Dataset Reviews
> Reliable Dataset References

More than a data management system for museum collections, we are a community of curators and researchers dedicated to responsible curation and education.

Arctos

https://arctosdb.org

MSB:Mamm:259343

Tamias speciosus

Tahoe National Forest; Yuba Pass
North America, United States, California, Sierra County
18 June 2012

skin; skull; liver (frozen); heart, kidney, lung, spleen (frozen); nematode (95% ethanol); skeleton, postcranial

Comment or report bad data [0][0]

MSB Mammals

[expand all] [collapse all]

Identifications [expand]

Tamias speciosus
Animalia; Chordata; Mammalia; Rodentia; Sciuridae; Sciurinae; Tamias speciosus
lodgepole chipmunk
Identified by John R. Demboski on 2013-09-23
Nature of ID: features
Confidence: high
Remarks: Former nature_of_id: expert.

Tamias

Citations

voucher of Tamias speciosus in Bell et al. 2018 DOI:10.1086/698300

Media

Location (1 Events)

BerkeleyMapper

highlight linked components

Determination Type: collection
assigned by Kayce C. Bell on 2012-11-20
Higher Geography: North America, United States, California, Sierra County more
Verbatim Locality: USA: California; Sierra; Tahoe National Forest; Yuba Pass
Specific Locality: Tahoe National Forest; Yuba Pass
Collecting Method: Sherman trap

Collector(s)

Kayce C. Bell, Randle D. McCain

Preparator(s)

Kayce C. Bell, Randle D. McCain

Identifiers

DZTM: Denver Zoology Tissue Mammal: 2752
NK: 215135
original identifier: DZTM 1752

Relationships [expand]

[Find All] [Find host of] [Find same individual as]

locality@2019-12-05: North America, United States, California, Sierra County: Tahoe National Forest; Yuba Pass

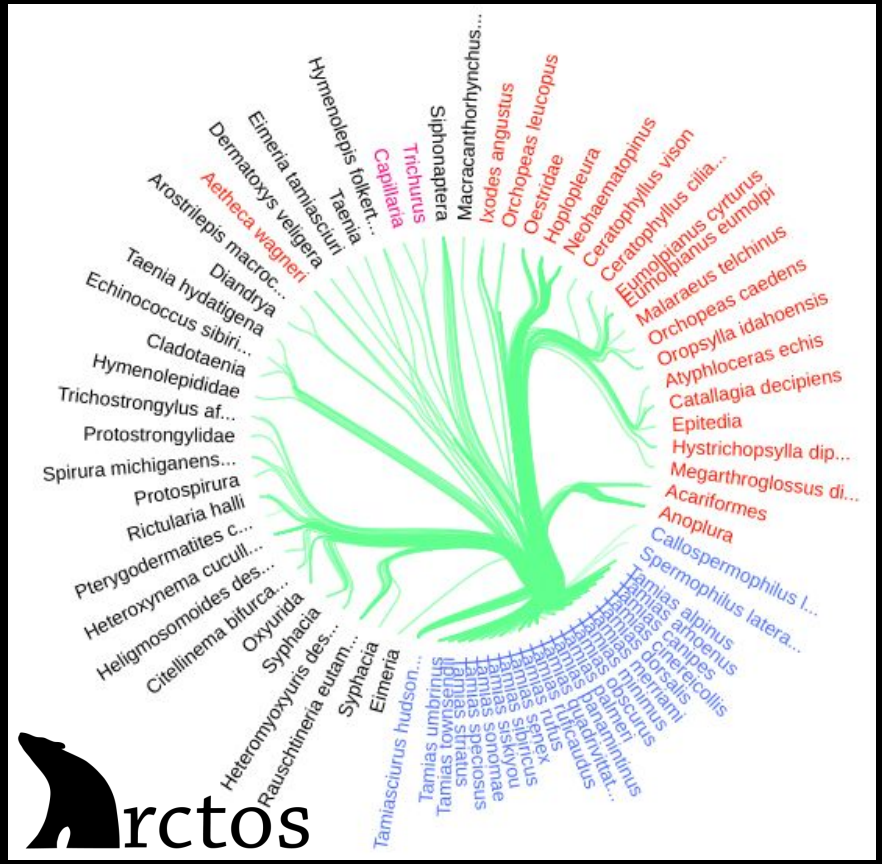
(*same individual as*) DMNS:Mamm:12990
family@2019-12-17: Sciuridae
identification@2019-12-17: Tamias speciosus
locality@2019-12-17: North America, United States, California, Sierra County: Tahoe National Forest; Yuba Pass

Links

GBIF Occurrence
iDigBio Occurrence
GloBI

Parts

Details



Arctos

GLOBI

about blog browse contribute data search references status Español 日本語

Example query: What do sea otters (*Enhydra lutris*) eat? or What do honey bees (*Apis*) pollinate?

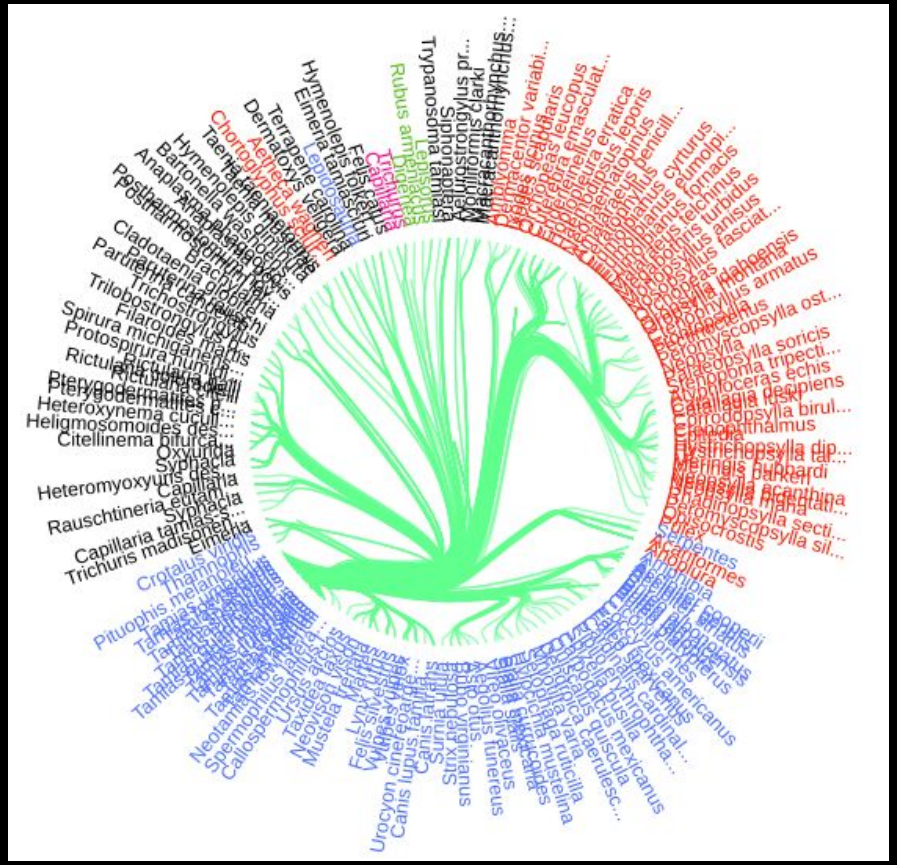
What kind of organisms do Tamias get parasitized by according to globi:globalbioticinteractions/vert?

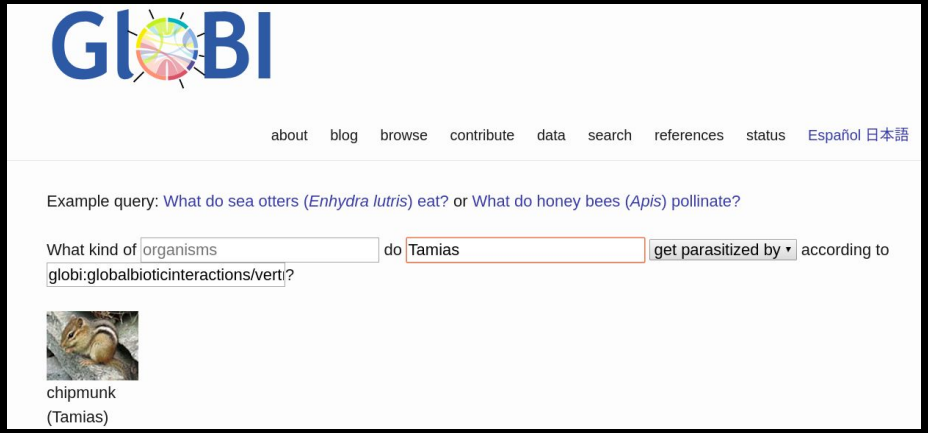
chipmunk
(Tamias)

Some chipmunk (Tamias sp.) biotic associations **provided by Arctos** as indexed and linked by GloBI on 26 September 2019

Background image: Slyusarev et al. (2015): Global Biotic Interactions food web map. figshare. http://dx.doi.org/10.6084/m9.figshare.1297762

https://globalbioticinteractions.org





Some chipmunk (Tamias sp.) biotic associations
by <u>all</u> providers
as indexed and linked by GloBI on 26 September 2019

Background image: Slyusarev et al. (2015): Global Biotic Interactions food web map. figshare. http://dx.doi.org/10.6084/m9.figshare.1297762

https://globalbioticinteractions.org

pragmatic reuse examples

> Research and Reviews
> Terrestrial Parasite Tracker TCN
> CCBER at UC Santa Barbara
> Arctos Collection Management
> Automated Dataset Reviews
> Reliable Dataset References

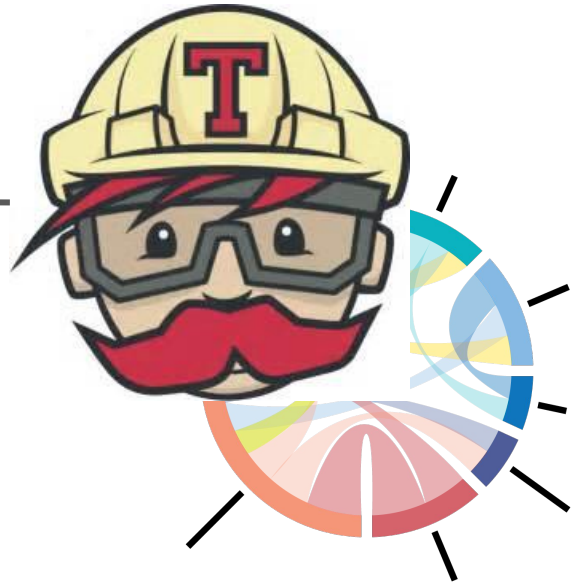
federated
data registries

zenodo

github
SOCIAL CODING

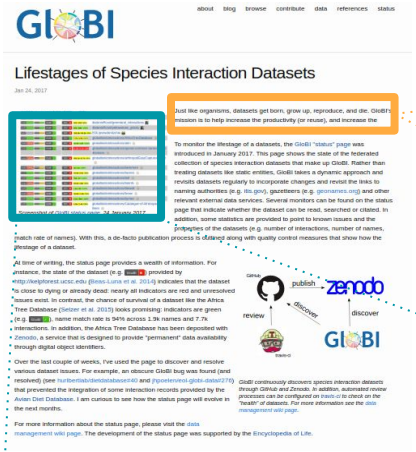
data = code

peer review



an automated and continuous process

http:/globalbioticinteractions.org/2017/01/24/lifestages-of-species-interaction-datasets/



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.

readable	issues	search citable		stats	
build unknown	issues 0 open	GloBI ✓	DOI ✖	9.1k / 4.3k / 87% / 6d	cmungall/Benesh-et-al-2017
build unknown	issues 0 open	GloBI ✓	DOI ✖	16 / 25 / 100% / 6d	cmungall/dinosaur-biotic-interactions
build passing	issues 0 open	GloBI ✓	DOI ✖	331 / 138 / 98% / 6d	diatomsRcool/greenland_interactions
build passing	issues 0 open	GloBI ✓	DOI ✖	254 / 252 / 99% / 6d	diatomsRcool/yellowstone_grizzly
build unknown	issues 0 open	GloBI ✖	DOI ✖	6d	Dryu0003/dietdatabase
build unknown	issues 2 open	GloBI ✓	DOI ✖	183.8k / 34.2k / 99% / 6d	EOL/pseudonitzchia

http://www.globalbioticinteractions.org/status.html accessed at 28 Sept 2017

Review included:
- 153989 interaction(s)
- 0 error(s)
- 55035 warning(s)

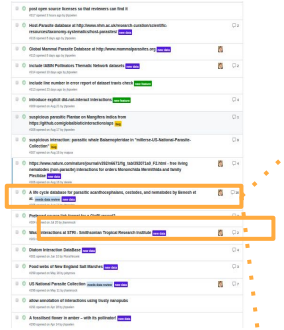
[globalbioticinteractions/scan] has the following reviewer comments:

48452	no source taxon defined
6477	found interaction type name [canis latrans (canidae)]
2	date [2919-08-06T00:00:00Z] is in the future

[...]

http://www.globalbioticinteractions.org/status.html accessed at 9 Jan 2020

Open discussions and reviews using GitHub issues



Code | Issues 62 | Pull requests 0 | Projects 0 | Wiki | Settings | Insights

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

Open **jhpoelen** opened this issue on Aug 14 · 16 comments

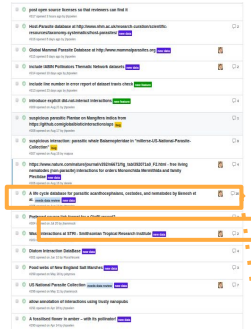
jhpoelen commented on Aug 14 • edited — Owner

As suggested by **@derele** -

[...]
I a 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,
[...]

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

Open discussions and reviews using GitHub issues



Code | Issues 62 | Pull requests 0 | Projects 0 | Wiki | Settings | Insights

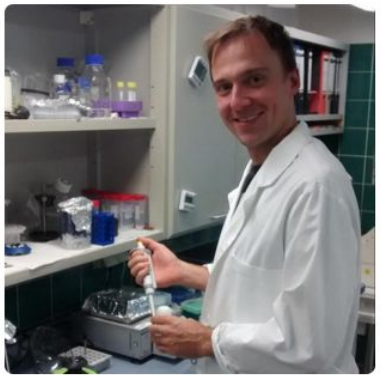
A life cycle database for parasitic acanthocephalans, cestodes, and nematodes by Benesh et al. #305

Open **jhpoelen** opened this issue on Aug 14 · 16 comments

jhpoelen commented on Aug 14 • edited

As suggested by **@derele** -

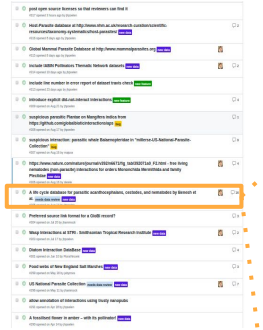
[...]
I a currently interacting with Dan Benesh who compiled a database on parasi
(http://onlinelibrary.wiley.com/doi/10.1002/ecy.1680/full)). Dan will work with r
me that he also wants to include this data in globi,
[...]



Emanuel Heitlinger
derele

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

Open discussions and reviews using GitHub issues



(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,
[...]

millerse self-assigned this on Aug 21

millerse added the **new data** label on Aug 21

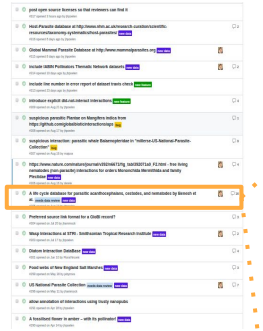
dbenesh82 commented 28 days ago

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 — Collaborator

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 — Collaborator

And done see here

millerse added the **needs data review** label 9 days ago

jhammock commented 9 days ago — Collaborator

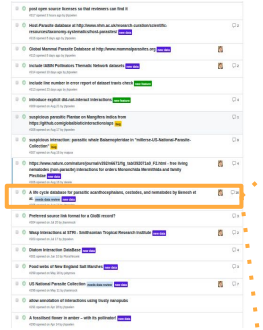
Looks good! If you want to get fancy, you will probably find terms for some of the LocationinhostName values in UBERON. I wouldn't bother chasing elusive ones, but if you find an exact match to the string you have, that should be worth using.

cmungall commented 9 days ago — Collaborator

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 — Collaborator

Fantastic! Yes please

cmungall commented 9 days ago — Collaborator

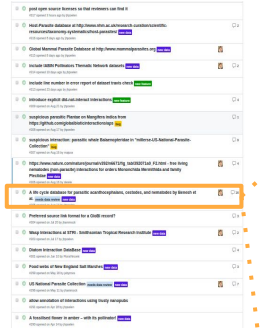
consider also mapping these to ENVO:

```
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 — Collaborator

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 — Collaborator

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

pragmatic reuse examples

> Research and Reviews
> Terrestrial Parasite Tracker TCN
> CCBER at UC Santa Barbara
> Arctos Collection Management
> Automated Dataset Reviews
> Reliable Dataset References

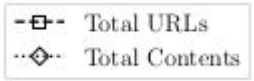
Toward Reliable Biodiversity Dataset References

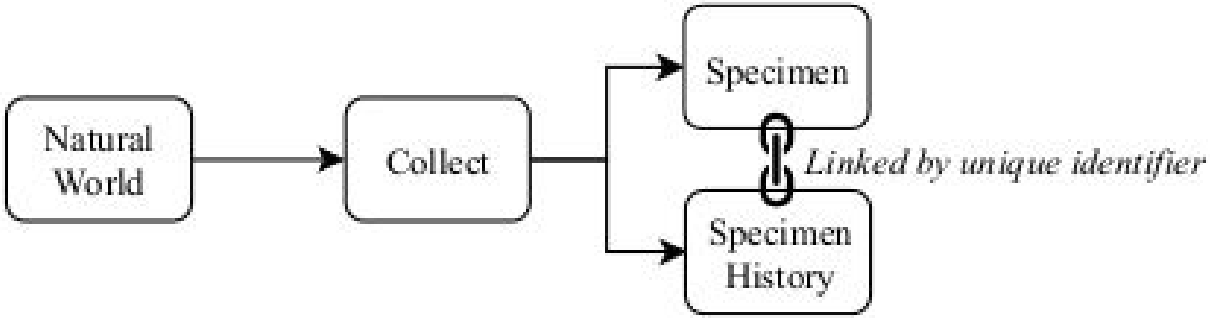
Michael J. Elliott[1†], Jorrit H. Poelen[2†*], José A.B. Fortes[1]

Abstract

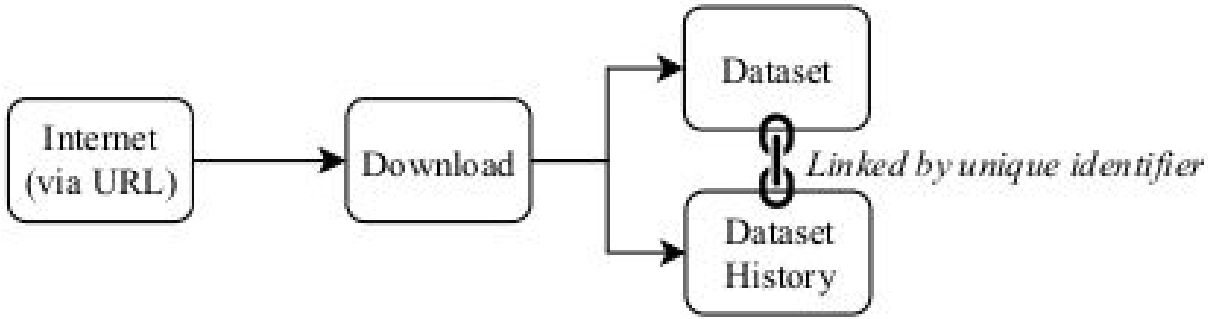
No systematic approach has yet been adopted to reliably reference and provide access to digital biodiversity datasets. Based on accumulated evidence, we argue that location-based identifiers such as URLs are not sufficient to ensure long-term data access. We introduce a method that uses dedicated data observatories to evaluate long-term URL reliability.

Elliott, M., Poelen, J.H. & Fortes, J., (submitted). Toward Reliable Biodiversity Dataset References. Preprint available at: http://dx.doi.org/10.32942/osf.io/mysfp . Funded by NSF OAC 1839201 .



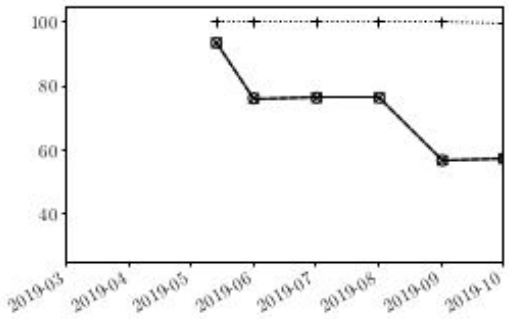


(a) Physical specimen collection

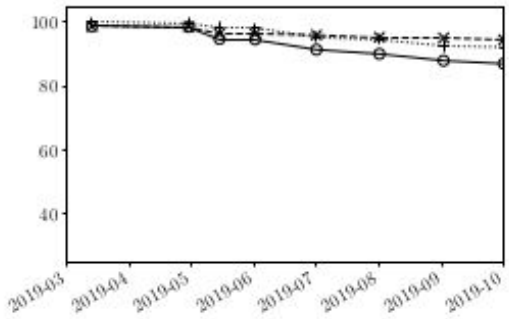


(b) Digital data collection

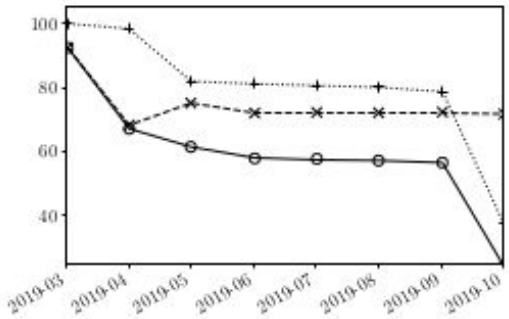
Elliott, M., Poelen, J.H. & Fortes, J., (submitted). Toward Reliable Biodiversity Dataset References. Preprint available at: http://dx.doi.org/10.32942/osf.io/mysfp . Funded by NSF OAC 1839201 .



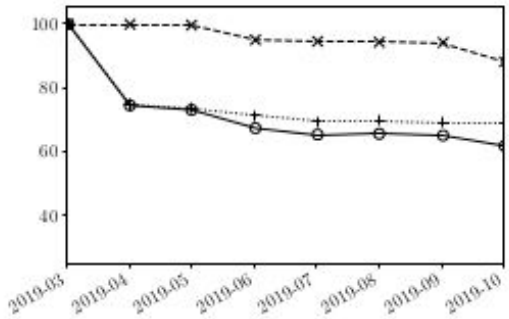
(a) BHL



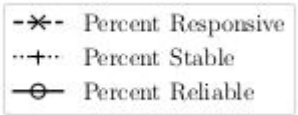
(b) DataONE



(c) GBIF



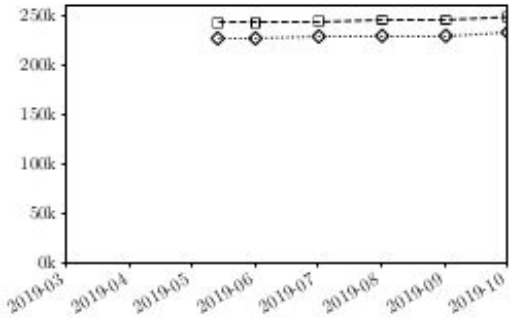
(d) iDigBio



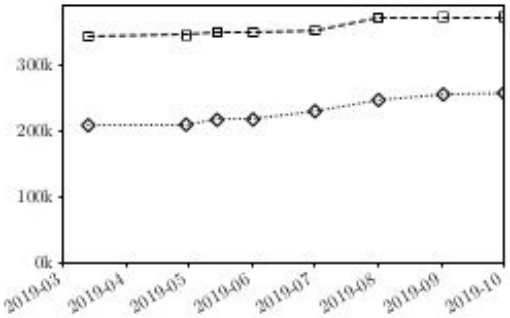
Elliott, M., Poelen, J.H. & Fortes, J., (submitted). Toward Reliable Biodiversity Dataset References. Preprint available at: http://dx.doi.org/10.32942/osf.io/mysfp . Funded by NSF OAC 1839201 .

Data Network	Responsive URLs	Stable URLs*	Reliable URLs
BHL[a]	57.41% (142,672)	99.97% (232,996)	57.39% (142,633)
DataONE[b]	94.55% (352,438)	92.27% (339,109)	87.09% (324,641)
GBIF[c]	71.72% (49,707)	37.35% (20,094)	24.05% (16,669)
iDigBio[c]	88.04% (5,477)	68.69% (4,251)	61.68% (3,837)
All observed URLs**	78.94% (546,645)	90.43% (593,469)	70.07% (485,203)

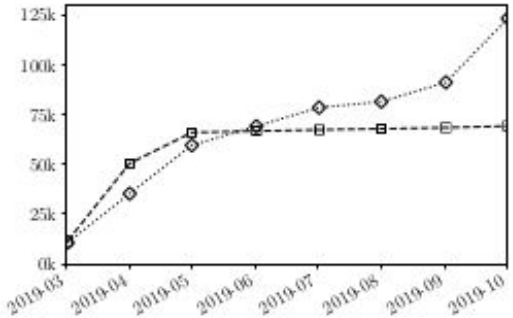
Elliott, M., Poelen, J.H. & Fortes, J., (submitted). Toward Reliable Biodiversity Dataset References. Preprint available at: http://dx.doi.org/10.32942/osf.io/mysfp . Funded by NSF OAC 1839201 .



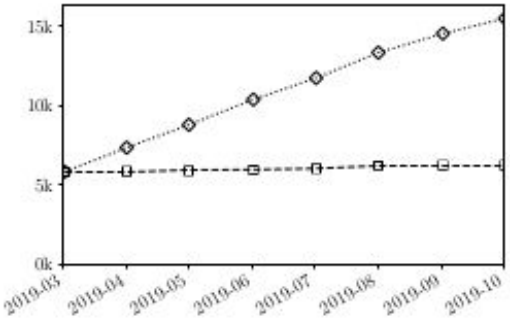
(a) BHL



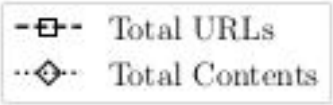
(b) DataONE



(c) GBIF



(d) iDigBio



Elliott, M., Poelen, J.H. & Fortes, J., (submitted). Toward Reliable Biodiversity Dataset References. Preprint available at: http://dx.doi.org/10.32942/osf.io/mysfp . Funded by NSF OAC 1839201 .

topics

> Context
> What is pragmatic reuse?
> Pragmatic reuse examples
> Q&A

it takes a village to keep and share a record

Data Providers

Data Integration

GLOBI

Data Mobilizers

Data Users

Ontology & Standards Communities

Taxonomists

Ecologists

OBO Foundry

https://globalbioticinteractions.org
info@globalbioticinteractions.org
@GlobalBiotic

Please cite GloBI using;
Jorrit H. Poelen, James D. Simons and Chris J. Mungall. (2014). Global Biotic Interactions: An open infrastructure to share and analyze species-interaction datasets. *Ecological Informatics*. http://dx.doi.org/10.1016/j.ecoinf.2014.08.005.

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).

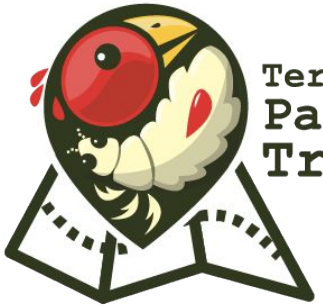
partners

GloBI

TEXAS A&M UNIVERSITY-CORPUS CHRISTI
GoMexSI
CENTER FOR COASTAL STUDIES

eol
Encyclopedia of Life

Heitlinger lab
Ecology
Evolution
molecular Mechanism
Interaction
parasites
of host and


Terrestrial Parasite Tracker

topics

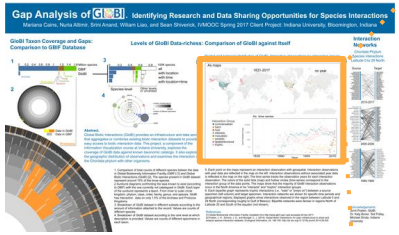
> Context
> What is pragmatic reuse?
> Pragmatic reuse examples
> Q&A

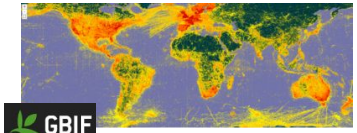
Why reuse existing species-interaction datasets?

> Increase Data Quality
> Avoid Rework
> Facilitate Research
> ...

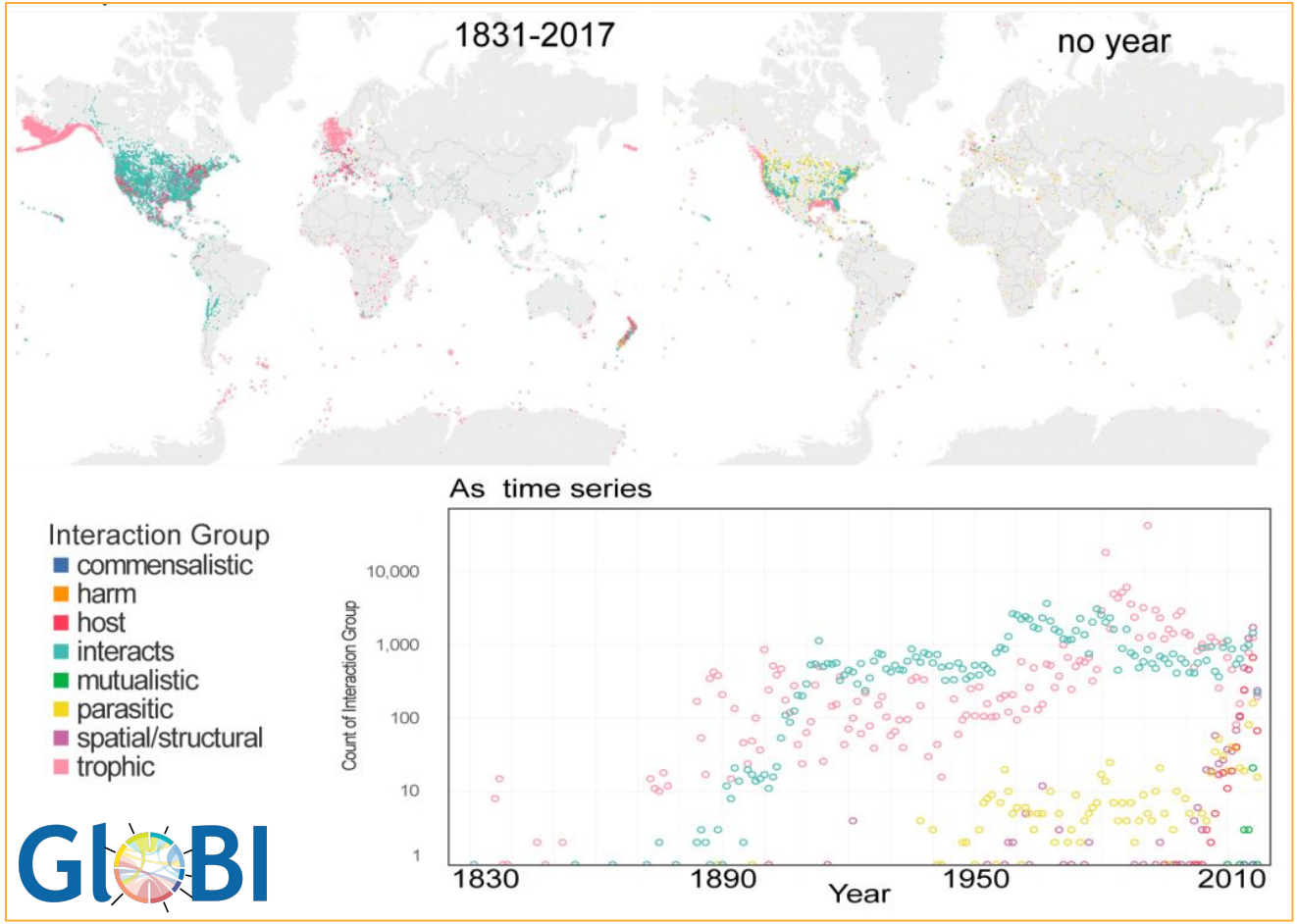
Why reuse existing species-interaction datasets?

Extra

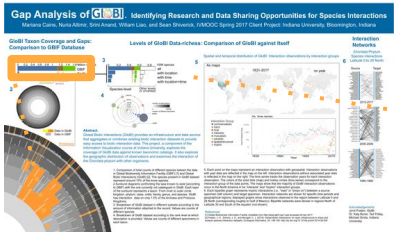


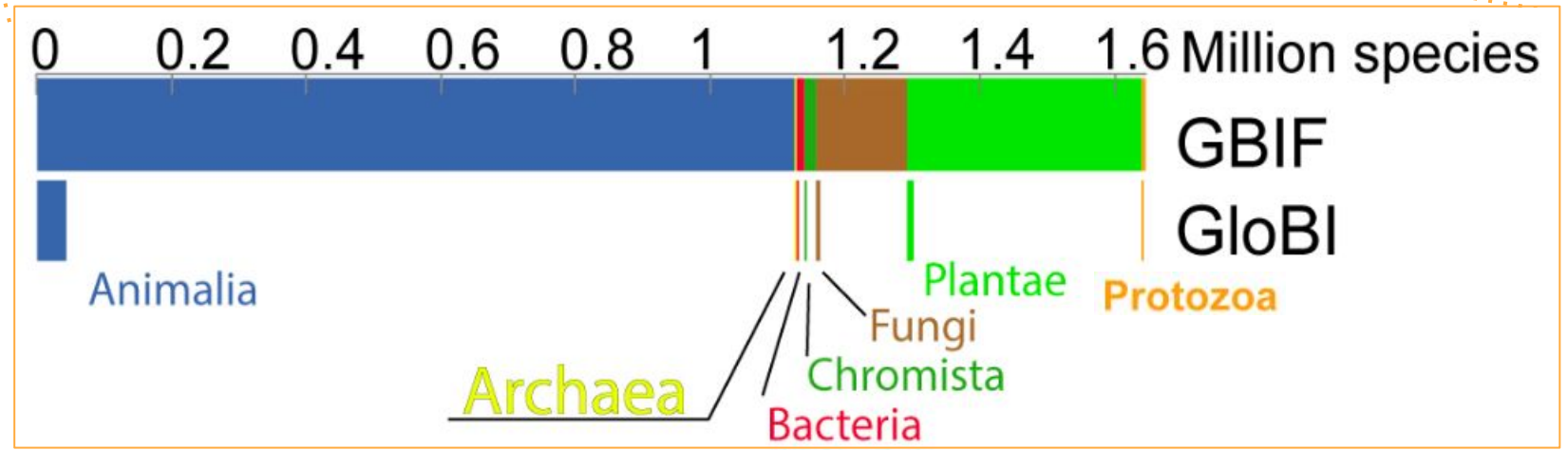


geospatial and temporal Eltonian shortfall



Cains, Mariana, Altimir, Nuria, Anand, Srini, Liao, William, & Shiverick, Sean. (2017). IVMOOC 2017 - Gap Analysis of GloBI Visualizations. Zenodo. http://doi.org/10.5281/zenodo.814922





taxonomic Eltonian shortfall

Cains, Mariana, Altimir, Nuria, Anand, Srini, Liao, William, & Shiverick, Sean. (2017). IVMOOC 2017 - Gap Analysis of GloBI Visualizations. Zenodo. http://doi.org/10.5281/zenodo.814922