

Discovering known biodiversity

May 24, 2022

Empowering Biodiversity Research II conference

Africa Museum, Tervuren

Donat Agosti

ORCID: [0000-0001-9286-1200](https://orcid.org/0000-0001-9286-1200)

Plazi, Switzerland

Quiz

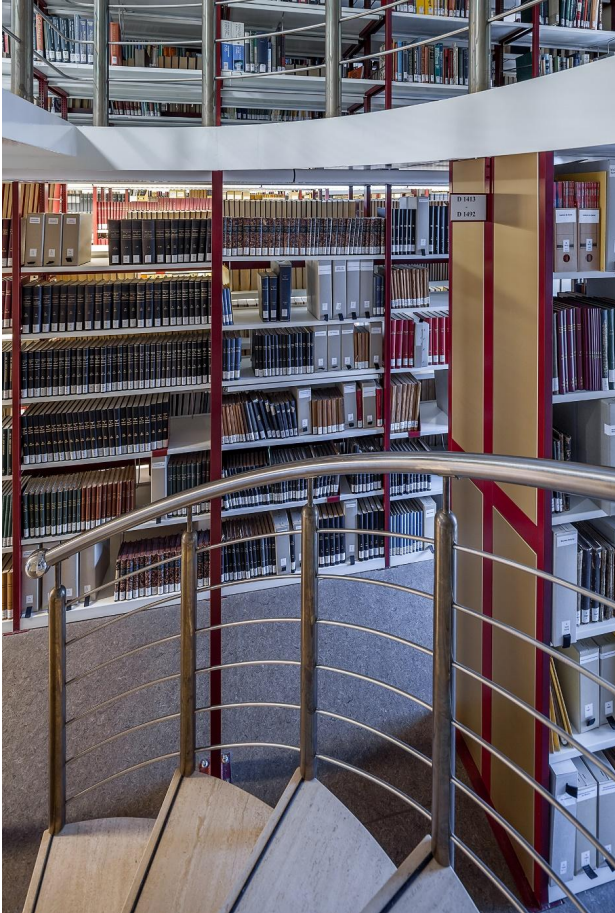
What is the taxonomic name of the Puer tea?

Provide as evidence:

- Taxonomic name
- Digital copy of the protologue of the taxon
- Evidence of holotype
- Digital copies of the protologues of five synonyms
- Digital copies of the documentation of the synonymy
- Digital copy of the phylogenetic position of the taxon
- Time sheet including processing time

Prize:





Known biodiversity knowledge

- Empiric science: all results published
- 500,000,000+ printed pages
- >> 1,000 journals publishing taxonomic content
- > 1,900,000 species described
- > 20,000,000+ taxonomic treatments
- approx. 18,000 new species discovered / year
- >> Millions of specimens identified by specialists (material citations)
- Billions of facts

BUT: only ca. 10-30% of the knowledge is digital.
Most is “unknown known knowledge”, not Digital Accessible Knowledge (DAK)



Take home points:

- Nothing in taxonomic publications makes sense except in the light of taxonomic treatments
- An identification of a specimen is only complete with a link to a taxonomic treatment
- A material citation in a publication is the gateway to the knowledge about the specimen
- Automate as much as possible, curate as much as needed
- Be aware of the power and use of your data beyond domain: make use of semantic publishing

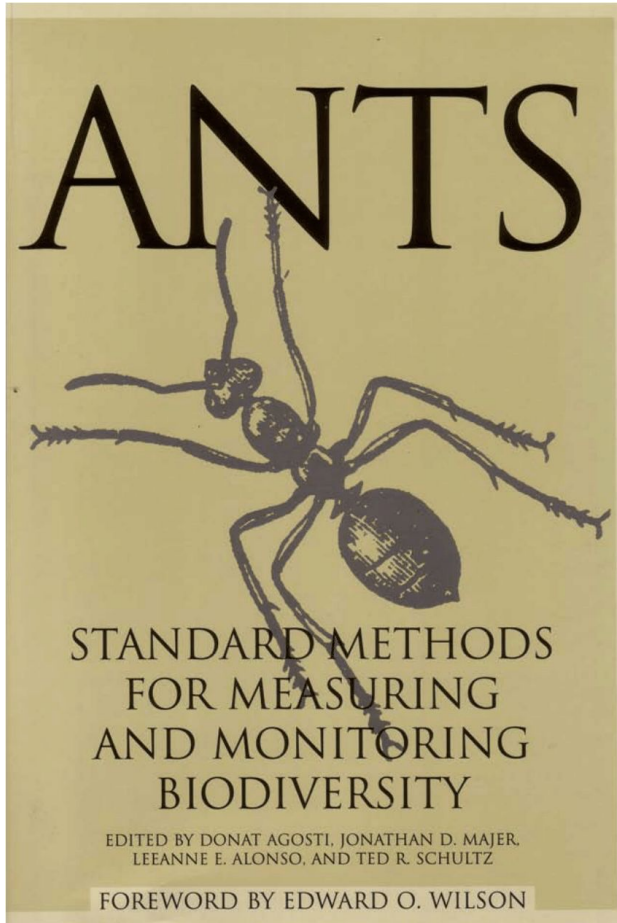


How many species do we lose?
How many species do we know?
How many species are on Earth?
What do we know about the species?



Basic questions in any monitoring program:

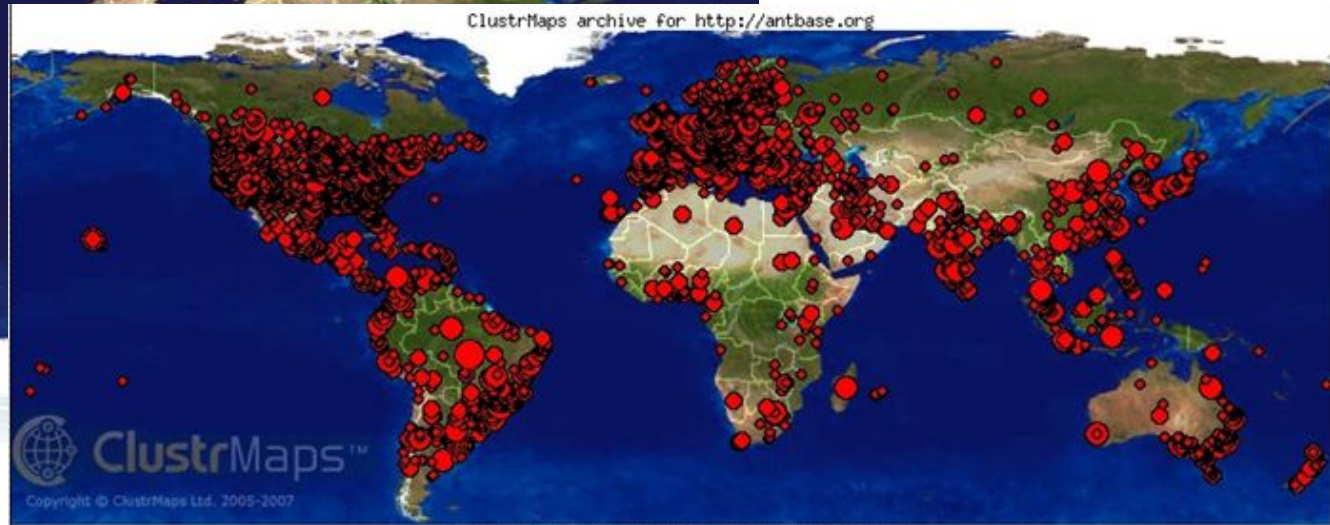
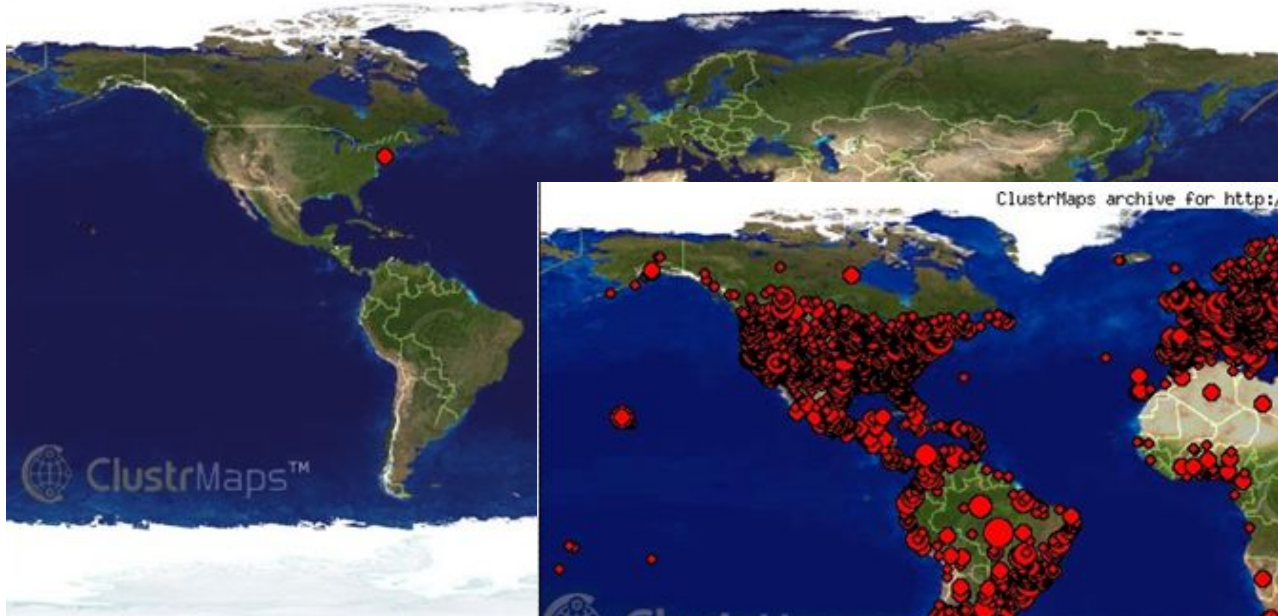
- What taxon is it?
- How does it fit into the phylogeny and classification?
- What do I know about it?
- What is its distribution?



The impact of WWW in 90ties



Before antbase.org, Harvard's Museum of Comparative Zoology could claim to be the only location with a complete set of ant systematics publications from 1758 - present.



109775 visits from 28 Aug 2006 to 29 Aug 2007

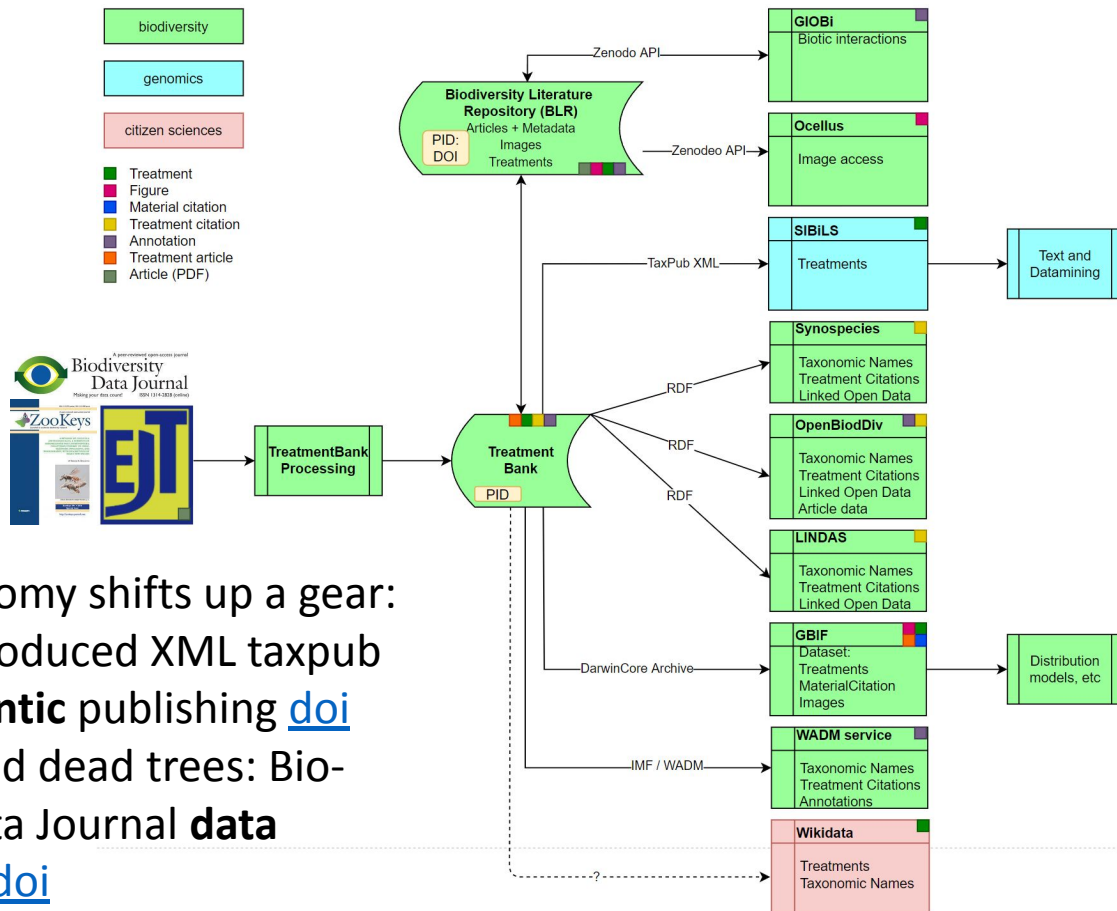
! distance in which individuals are clustered
Total number of visits depicted above = 109457

Dot sizes:

● = 1000 + ● = 100 - 999 ● = 10 - 99 ● = 1 - 9

Through antbase.org's digital library, access to this body of literature is worldwide, and it is actively used (>10,000 visits in one month only).

2022: Immediate reuse of published data



2010: Taxonomy shifts up a gear: Zookeys introduced XML taxpub based **semantic** publishing [doi](#)

2013: Beyond dead trees: Biodiversity Data Journal **data** publishing. [doi](#)

BUT: we still do not know what we know.

Publishing specimen data: Material citation: access to data about a specimen (occurrence)



zenodo

July 27, 2020

Review of South African Euryphyminae

Tshililo, Precious

Tshililo, Precious (2020). Review of South African Euryphyminae. Zootaxa 4820 (1): 70-104. DOI: <https://doi.org/10.11646/zootaxa.4820.1.4>

15 views 0 downloads

PDF

Biodiversity Literature Repository

OpenAIRE

Publication date: July 27, 2020

DOI: [10.11646/zootaxa.4820.1.4](https://doi.org/10.11646/zootaxa.4820.1.4)

Keywords: **Arthropoda** **Insecta** **Arachnida** **Collembola** **Collembola** **Collembola**

Published in: Zootaxa 4820 pp. 70-104 (1)

Related identifiers: DOI: 10.11646/zootaxa.3753.5.1 (Publication) 10.1111/1365-3113.11956.020270.x (Publication) 10.59619/zoo.10030 (Publication) 10.1002/1365-3113.11956.020270.x (Publication) 10.1111/1365-3113.11956.020270.x (Publication) 10.1002/1365-3113.11956.020270.x (Publication)

Remarks: This genus was erected by Doherty (1956a) on the basis of flat pronotum, narrow fastigium of vertex and truncate subgenital plate of a female. This genus resembles *Calliptamoides* by its shape of the male cercus. Its taxonomic status is not satisfactory. The SANCO collection contains a few specimens, which may be unique species.

Published as part of Tshililo, Precious, 2020: Review of South African Euryphyminae, pp. 70-104 in Zootaxa 4820 (1) on pages 85-86, DOI: 10.11646/zootaxa.4820.1.4. <http://zenodo.org/record/4397315>

Calliptamoides minimus: Dirsh 1956

5 views 3 downloads

PDF

Biodiversity Literature Repository

OpenAIRE

Publication date: July 27, 2020

DOI: [10.5281/zenodo.4397315](https://doi.org/10.5281/zenodo.4397315)

Keywords: **Arthropoda** **Insecta** **Arachnida** **Collembola** **Collembola** **Collembola**

Published in: Zootaxa 4820 pp. 70-104 (1)

Related identifiers: DOI: 10.5281/zenodo.4397323 (Figure) 10.5281/zenodo.4397323 (Figure) 10.5281/zenodo.4397323 (Figure) 10.5281/zenodo.4397345 (Figure)

Remarks: This genus was erected by Doherty (1956a) on the basis of flat pronotum, narrow fastigium of vertex and truncate subgenital plate of a female. This genus resembles *Calliptamoides* by its shape of the male cercus. Its taxonomic status is not satisfactory. The SANCO collection contains a few specimens, which may be unique species.

Published as part of Tshililo, Precious, 2020: Review of South African Euryphyminae, pp. 70-104 in Zootaxa 4820 (1) on pages 85-86, DOI: 10.11646/zootaxa.4820.1.4. <http://zenodo.org/record/4397315>

Calliptamoides minimus Dirsh, 1956

5 views 3 downloads

PDF

Biodiversity Literature Repository

OpenAIRE

Publication date: July 27, 2020

DOI: [10.5281/zenodo.4397315](https://doi.org/10.5281/zenodo.4397315)

Keywords: **Arthropoda** **Insecta** **Arachnida** **Collembola** **Collembola** **Collembola**

Published in: Zootaxa 4820 pp. 70-104 (1)

Related identifiers: DOI: 10.5281/zenodo.4397323 (Figure) 10.5281/zenodo.4397323 (Figure) 10.5281/zenodo.4397323 (Figure) 10.5281/zenodo.4397345 (Figure)

Remarks: This genus was erected by Doherty (1956a) on the basis of flat pronotum, narrow fastigium of vertex and truncate subgenital plate of a female. This genus resembles *Calliptamoides* by its shape of the male cercus. Its taxonomic status is not satisfactory. The SANCO collection contains a few specimens, which may be unique species.

Published as part of Tshililo, Precious, 2020: Review of South African Euryphyminae, pp. 70-104 in Zootaxa 4820 (1) on pages 85-86, DOI: 10.11646/zootaxa.4820.1.4. <http://zenodo.org/record/4397315>

Calliptamoides minimus Dirsh, 1956

5 views 3 downloads

PDF

Biodiversity Literature Repository

OpenAIRE

Publication date: July 27, 2020

DOI: [10.5281/zenodo.4397315](https://doi.org/10.5281/zenodo.4397315)

Keywords: **Arthropoda** **Insecta** **Arachnida** **Collembola** **Collembola** **Collembola**

Published in: Zootaxa 4820 pp. 70-104 (1)

Related identifiers: DOI: 10.5281/zenodo.4397323 (Figure) 10.5281/zenodo.4397323 (Figure) 10.5281/zenodo.4397323 (Figure) 10.5281/zenodo.4397345 (Figure)

Remarks: This genus was erected by Doherty (1956a) on the basis of flat pronotum, narrow fastigium of vertex and truncate subgenital plate of a female. This genus resembles *Calliptamoides* by its shape of the male cercus. Its taxonomic status is not satisfactory. The SANCO collection contains a few specimens, which may be unique species.

Published as part of Tshililo, Precious, 2020: Review of South African Euryphyminae, pp. 70-104 in Zootaxa 4820 (1) on pages 85-86, DOI: 10.11646/zootaxa.4820.1.4. <http://zenodo.org/record/4397315>

Calliptamoides minimus Dirsh, 1956

5 views 3 downloads

PDF

Biodiversity Literature Repository

OpenAIRE

Publication date: July 27, 2020

DOI: [10.5281/zenodo.4397315](https://doi.org/10.5281/zenodo.4397315)

Keywords: **Arthropoda** **Insecta** **Arachnida** **Collembola** **Collembola** **Collembola**

Published in: Zootaxa 4820 pp. 70-104 (1)

Related identifiers: DOI: 10.5281/zenodo.4397323 (Figure) 10.5281/zenodo.4397323 (Figure) 10.5281/zenodo.4397323 (Figure) 10.5281/zenodo.4397345 (Figure)

Remarks: This genus was erected by Doherty (1956a) on the basis of flat pronotum, narrow fastigium of vertex and truncate subgenital plate of a female. This genus resembles *Calliptamoides* by its shape of the male cercus. Its taxonomic status is not satisfactory. The SANCO collection contains a few specimens, which may be unique species.

Published as part of Tshililo, Precious, 2020: Review of South African Euryphyminae, pp. 70-104 in Zootaxa 4820 (1) on pages 85-86, DOI: 10.11646/zootaxa.4820.1.4. <http://zenodo.org/record/4397315>

GBIF species: <https://www.gbif.org/species/1703942>

GBIF occurrence: <https://www.gbif.org/occurrence/3015409563>

BLR Taxonomic treatment: <https://doi.org/10.5281/zenodo.4397319>

BLR Article: <https://zenodo.org/record/4397315>

PDF: <https://doi.org/10.11646/zootaxa.4820.1.4>

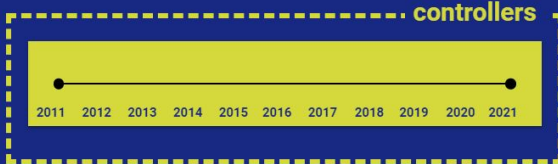


EUROPEAN JOURNAL OF TAXONOMY

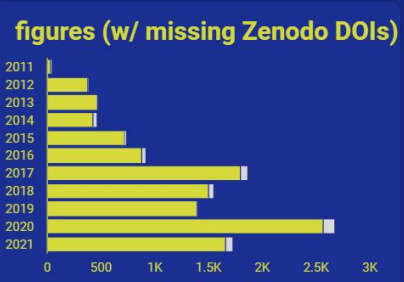
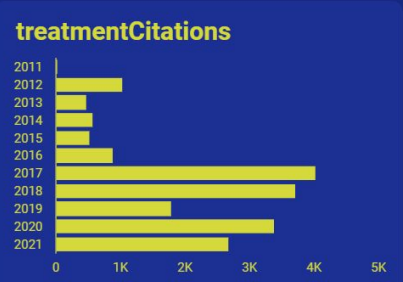
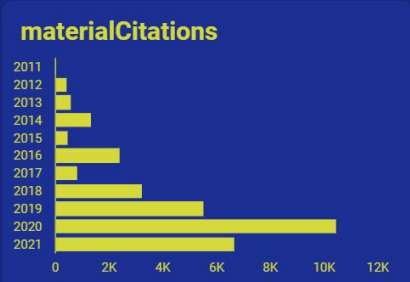
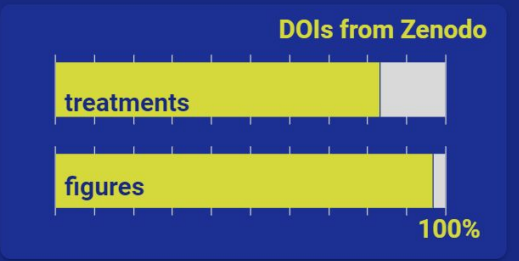
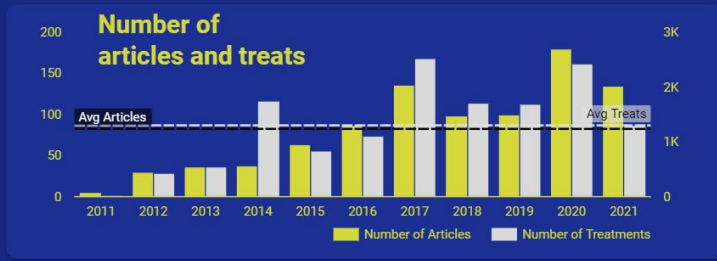
An Open Access journal for descriptive taxonomy of living and fossil eukaryotes



General Assessment of TreatmentBank Data



Number of Articles 905	Treatments 14,290	MatCitations 32,005	TreatCitations 19,134
on Zenodo 100.00%	on GBIF 100.00%	Number of Pages 31,499	ContFisCount 11,810
		ContTabCount 1,090	





The use of treatment citation

LESTES (LESTES) BARBARUS (Fabricius) Text fig. 11, 21, 22, pp. 347, 366

Agriion barbara Fabricius, 1798: 286; van der Linden, 1825: 36

Lestes barbara Hagen, 1849: 147 (type species); Hagen & Selys, 1850: 161; Selys, 1862: 318 (34 sep.); Hagen, 1863: 194; McLachlan, 1889: 348; Martin, 1910: 82; Fraser, 1933: 49 fig. 20

Lestes barbarus Selys, 1887: 67 (Egypt); Roster, 1888: 164 pl. 3 (early stages); Kirby, 1890: 162; Martin, 1910: 84, 87; Morton, 1924: 30; May, 1933: 31, 89 (adult and larva); Cowley, 1940: 174; Schmidt, 1951: 122 (quoting Hagen, 1849, *Lestes barbarus* as type species of *Lestes*); Conci & Nielsen, 1956: 65, 66 figs.; Aguesse & Pruja, 1958: 103 (Morocco); Robert, 1958: 82; Loibl, 1958: 55-80; Corbet, 1962: 178; Lieftinck, 1966: 10 (Morocco)

Agriion nympha Hanseemann, 1823: 161 (syn. Kirby, 1890)

Lestes barbata Belyshev, 1973: 509 fig. 213 (? laps. cal.)

Type from Barbary in North Africa.

Martin (1910) says (p. 82) “Enfin, à juger d’après un exemplaire rapporté de l’Afrique tropicale par M. Ch. Alluaud, *Lestes barbara*, l’espèce européenne, habite également l’Afrique chaude”. On p. 87, in placing *Lestes barbarus* near *obscurus* and *niger* he gave the specific distribution as central and southern Europe, Asia Minor and from there to India; throughout Mediterranean Africa; and states that Alluaud collected the only known record for tropical Africa. He said it was identical to the European form. It would seem probable that Alluaud’s *barbara* was either a misidentification, or the locality was incorrect.

Characters. Non pruinose. Pterostigma bicolorous. Wing apices broadly rounded, as in *dryas* and *ictericus*. Sectors of arcus well forward as in *sponsa*, *viridis*, and all Ethiopian species. Orbits below and sternites unmarked with black. Anal appendages yellow with black apex; inferior appendage over half as long as superior, with flat extension ending in an out-turned point. Vulvar scale simple.

General description. Non-pruinose.

Almost mature male (Morocco). Labrum greenish yellow. Postclypeus bronze-black, yellow at lateral margins. Head above bronze or metallic green, but antennal bases yellow. Orbits below yellow.

[doi](#)

Treatment Citation Data

Treatment Citation ID

Verbatim Treatment Citation

Verbatim Cited Taxon Name

Rank of Cited Taxon

Taxonomic Kingdom

Taxonomic Phylum

Taxonomic Class

Taxonomic Order

Taxonomic Family

Cited Taxon Genus

Cited Taxon Species

Verbatim Cited Taxon Authority

Cited Taxon Authority Name

Cited Taxon Authority Year

Cited Authors

Cited Year of Publication

Cited Journal / Publisher

Cited Volume Number

Verbatim Cited Volume Number

Cited Page

Verbatim Cited Reference

Cited Treatment HTTP URI

Countries

Collections

Materials Data

Materials Citation Data

Fields to Use in Statistics

Output?	Order? (Desc?)	Field Name	Filter on Values	Operation	Filter on Operation Result
<input checked="" type="checkbox"/>	<input type="checkbox"/>	Document UUID	B82187B6FA4FFF50F	Show Individual Value ▾	
<input checked="" type="checkbox"/>	<input type="checkbox"/>	Article UUID		Show Individual Value ▾	
<input checked="" type="checkbox"/>	<input type="checkbox"/>	Taxon Genus		Show Individual Value ▾	
<input checked="" type="checkbox"/>	<input type="checkbox"/>	Taxon Species		Show Individual Value ▾	
<input checked="" type="checkbox"/>	<input type="checkbox"/>	Cited Treatment HTTP URI		Show Individual Value ▾	
<input checked="" type="checkbox"/>	<input type="checkbox"/>	Verbatim Cited Taxon Name		Show Individual Value ▾	
<input checked="" type="checkbox"/>	<input type="checkbox"/>	Verbatim Treatment Citation		Show Individual Value ▾	
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	Cited Year of Publication		Show Individual Value ▾	
<input checked="" type="checkbox"/>	<input type="checkbox"/>	Year of Publication		Show Individual Value ▾	



data liberation infrastructure

Get publications:

- PDF (library access)
- Zenodo via microservice

Receive publications:

- PDF
- Born digital
- Scanned
- XML
- JATS
- TaxPub/JATS
- HTML / XHTML

Decode

- manual
- automated

Enhance

- semantics
- link

Data Quality Control

- Criteria
- Corrections

 **TreatmentBank**

Create

- Open FAIR data

FAIR Digital Objects

Access / Metrics

- User Interface
- API
- Dashboards

SynoSpecies

Taxonomic names
Treatments

Disseminate

- GBIF
- NCBI

Reuse



Research publications

DOI: 10.1111/j.1365-3113.2011.04501.x

The distribution of female angulated mandibles in *Hemacrosia* sp.

Published online in Wiley InterScience, DOI: 10.1111/j.1365-3113.2011.04501.x

© 2011 The Authors. Journal compilation © 2011 British Ecological Society, *Journal of Animal Ecology*, 80, 100–108

Policy support: Collaborative for Innovation with open access and supporting it

© 2011 British Ecological Society

Reuse

FAIR Digital Objects

Occurrence / Material Citation

Hemacrosia manganian Hu, Zhu & Du, 2021

Map showing distribution of Hemacrosia manganian in China.

Treatment

Hemacrosia manganian Hu, Zhu & Du, 2021

Figure 1: Illustrations of Hemacrosia manganian and Hemacrosia Enderlein.

Treatment Article Dataset

New illustrations, new species and new combination of *Hemacrosia Enderlein* (Plecoptera: Perlidae) from China

Map of China showing distribution of Hemacrosia manganian.

Description: The new species Hemacrosia manganian Hu, Zhu & Du, 2021 is described from China. It is characterized by the following features: ...

Contacts: Hu, Zhu, Du

zenodo

Figure

FIGURE 1 Hemacrosia violacea from Fujian, Mainland China. Illustrations, new species and new combination of Hemacrosia Enderlein (Plecoptera: Perlidae) from China

Treatment

Hemacrosia violacea Enderlein 1909

OpenAIRE

Illustrations, new species and new combination of Hemacrosia Enderlein (Plecoptera: Perlidae) from China

zenodo

Access / Metrics

- User Interface
- API
- Dashboards

ocellus⁴

A PLAZI PROJECT

zenodeo²

A PLAZI PROJECT

FAIR Digital Objects

Biodiversity Literature Repository

Publication



An example of fully automated processing

Life demonstration to the Flemish government, September 13, 2021

- Original article: <https://doi.org/10.11646/zootaxa.5032.4.6>
- **Zenodo** article deposition: <https://zenodo.org/record/5499090#.YTtdHJ0zb8A>
- Zenodo figure deposition: <https://zenodo.org/record/5500033#.YTtdMJ0zb8A>
- Zenodo/BLR treatment deposition: <https://doi.org/10.5281/zenodo.5500037> (see the custom metadata)
- **TreatmentBank** article: <https://treatment.plazi.org/GgServer/summary/C154DB54FFA9B65D2460FFED5F578933>
- TreatmentBank treatment (HTML): <http://treatment.plazi.org/id/3D6DA32C-FFA1-B655-24F7-FDAA5C4F8EC3>
- TreatmentBank treatment (JSON): <https://zenodo.org/record/5500037/export/json>
- TreatmentBank stats:
<https://tb.plazi.org/GgServer/dioStats/stats?outputFields=doc.articleUuid+doc.doi+doc.zooBankId+doc.gbifId+doc.zenodoDepId+bib.source+cont.pageCount+cont.treatCount+cont.treatCountDoi+cont.treatCitCount+cont.matCitCount+cont.figCount+cont.figCountZen+cont.bibRefCount&groupingFields=doc.articleUuid+doc.doi+doc.zooBankId+doc.gbifId+doc.zenodoDepId+bib.source&FP-doc.articleUuid=C154DB54FFA9B65D2460FFED5F578933&format=JSON>
- **GBIF** dataset ID: <https://www.gbif.org/dataset/8f239084-30f3-4a6c-ba97-3eb65356beb5>
- GBIF occurrence data set: https://www.gbif.org/occurrence/search?dataset_key=8f239084-30f3-4a6c-ba97-3eb65356beb5
- GBIF occurrence data set, holotypes only:
https://www.gbif.org/occurrence/search?dataset_key=8f239084-30f3-4a6c-ba97-3eb65356beb5&type_status=HOLOTYPE
- GBIF species pages:
https://www.gbif.org/species/search?dataset_key=8f239084-30f3-4a6c-ba97-3eb65356beb5&origin=SOURCE&status=ACCEPTED&advanced=1
- **Ocellus** (Bilder in BLR):
<https://ocellus.info/images.html?q=%2210.11646/zootaxa.5032.4.6%22&size=30&page=1&communities=biosyslit>
- **Synospecies** (Synonymy triple store): <https://synospecies.plazi.org/#Kiotina+spatulata>



Biodiversity Literature Repository



Repository for data liberated from publications

453,000 images

72,000 articles

400,000 taxonomic treatments

Collaboration with Zenodo /CERN

Recognized as EU research infrastructure

Mints DOI for treatments and figures

TreatmentBank



Data conversion and access service

73,000 articles

762,000 taxonomic treatments

1,118,000 materials citations

>50% of annually described new species

production in 2021: 25,000 articles, 224,000 treatments, 180,000 images



Global Biodiversity Information Facility



Reuse of treatment articles mediated by Plazi

39,700 treatment article data sets (58% of total data sets in GBIF)

377,000 taxonomic treatments (90,000 unique species)

224,000 figures

591,000 materials citations (occurrences)

The figures are lower in GBIF because of Quality Control measures

In comparison to what's in TreatmentBank



Data conversion and access service

73,000 articles

762,000 taxonomic treatments

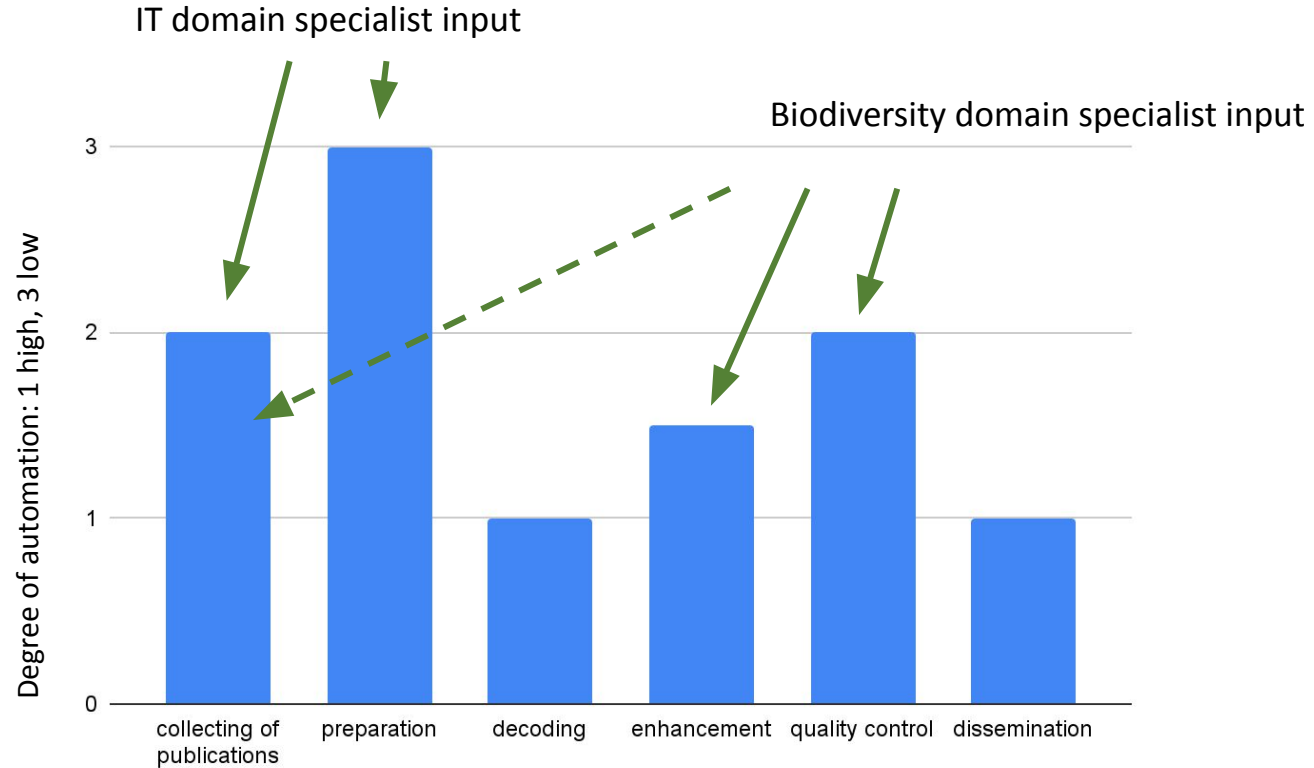
1,118,000 materials citations

>50% of annually described new species

production in 2021: 25,000 articles, 224,000 treatments, 180,000 images

The lower number in GBIF is due to quality control and minimal standards for data targeted for GBIF.

Next steps: getting the crowd involved



This is a very expensive effort, has to be shared by partners avoiding duplications as much as possible, and needs to be avoided by changing the way we publish.

Vision



zenodo

OpenAIRE

Access / Metrics

- User Interface
- API
- Dashboards

ocellus⁴
A PLAZI PROJECT

zenodeo²
A PLAZI PROJECT

Biodiversity Literature Repository



Get publications:

- PDF (library access)
- Zenodo via microservice

Receive publications:

- PDF
- Born digital
- Scanned
- XML
- JATS
- TaxPub/JATS
- HTML / XHTML

Decode

- manual
- automated

Enhance

- semantics
- links

Data Quality Control

- Criticism

Create

- Open FAIR data

Global biodiversity data liberation service

TreatmentBank

Species

Taxonomic names
Treatments

Reuse

GBIF

Research publications

Reuse

Occurrence /
Material Citation

Treatment

Treatment Article Dataset



Redescription and synonymies of *Diplura macrura* (C. L. Koch, 1841) and *D. lineata* (Lucas, 1857), with notes o

Annotations

Showing 54 annotations

ITEM: [PAGE0003.PNG]

Mygale macrura C. L. Koch , 1841 : 38 , fig. 715 (♂) .
Diplura bicolor Simon , 1889 : 215 (♀)...

subSubSection part 4

Diagnosis Both sexes of this species have a strongly contrasting color pattern , with a reddish brown...

subSubSection

Type material *Mygale macrura* : BRAZIL : ♂ , holotype , Minas Gerais , São João del Rei , [1830] ,

subSubSection part 1

Type material *Mygale*

European Journal of Taxonomy 210: 1–21 (2016)

Diplura unifর্মis – Platnick 1993: 89 (transfer) — Silva-Moreira et al., 2010: 32
Lincobele bicolor – Platnick 1998: 120 (transfer)

Diagnosis
Both sexes of this species have a strongly contrasting color pattern, with a reddish brown carapace and a dark brown abdomen shared only with *Diplura paraguayensis* (Gerschman & Schiapelli, 1942). In mature *D. macrura*, the abdomen is uniformly dark brown, without spots or any visible markings. However, according to its original description (Gerschman & Schiapelli 1942), *D. paraguayensis* has a blackish brown reticulate on abdomen dorsum. Moreover, the lyra of *D. paraguayensis* has 13 setae (Gerschman & Schiapelli 1942: pl. ix; Schiapelli & Gerschman 1968, fig. 7), in contrast with the 7–5 setae in *D. macrura*. The sub of *D. macrura* (Figs 6–7) has an embolic base much larger than in *D. paraguayensis* (Schiapelli & Gerschman 1968: figs 14–15). Also, the embolus is about 2x longer than the bulb in *D. macrura* and 3x longer in *D. paraguayensis*. The spermathecae of *D. macrura* has a longer and thinner stalk (Fig. 13) than the females of *D. paraguayensis* from Argentina (Goloboff 1982: 1). Also, the distal lobes are larger and spread over the distal third of the stalk in *D. macrura*, while they are smaller and concentrated in the apex in *D. paraguayensis*.

Type material
Mygale macrura: BRAZIL ♂, holotype, Minas Gerais, São João del Rei, [1830], Sellow, Friedrich W. MN 2083, examined.

Diplura bicolor: BRAZIL ♀♀, syntypes, Minas Gerais, Catas Altas, Caraça (coll. E. Simon, MNIN AR 4932 B337), examined (photos).

Thalerochele unifর্মis: BRAZIL ♂, holotype, São Paulo, E. of Garbe 'MZSP', MNRJ, not located.

Thalerochele minensis: BRAZIL ♂, holotype, Minas Gerais, Ouro Preto, Magalhães Gomes (Mello-Leitão collection 880, MNIN 1360), not located.

Thalerochele aurantiacal: BRAZIL ♀, holotype, Minas Gerais, Ouro Preto/Mariana: Itacolom O. Leonards (MNRJ 53945), not located.

Material examined
BRAZIL: 1 ♂, Minas Gerais, Parque Nacional Serra do Cipó, R. Bertan, R. Martins, C.S. Fukushima & M.P. Pavani leg. (MNRJ) 1 ♂, 2 juvs. Caté, 11–19 Feb. 1961, P. Vanzolini, H. Britski & N. Menezes

Fig. 1. Original (right) and added (left) labels from the holotype of *Diplura macrura* (C. L. Koch, 1841)

4 of 21 • page0003.png

[Source](#)



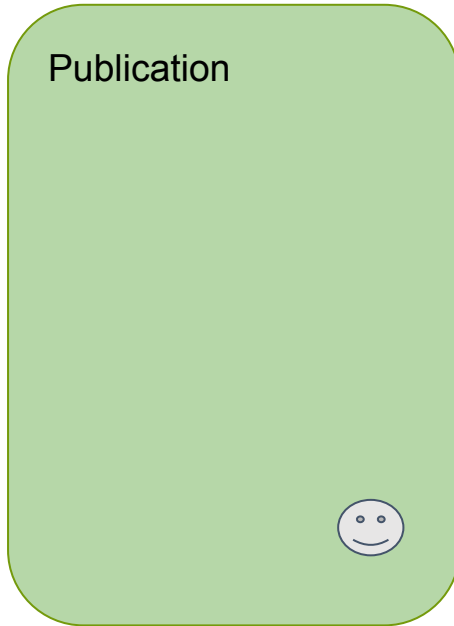
Discovering known biodiversity

Create a list of the Earth' known taxa, and related digital accessible knowledge embedded in scholarly publications as open findable, accessible, interoperable and reusable data about the Earth's species (FAIR digital objects), as input to the biodiversity knowledge graph, liberated from scholarly publications.

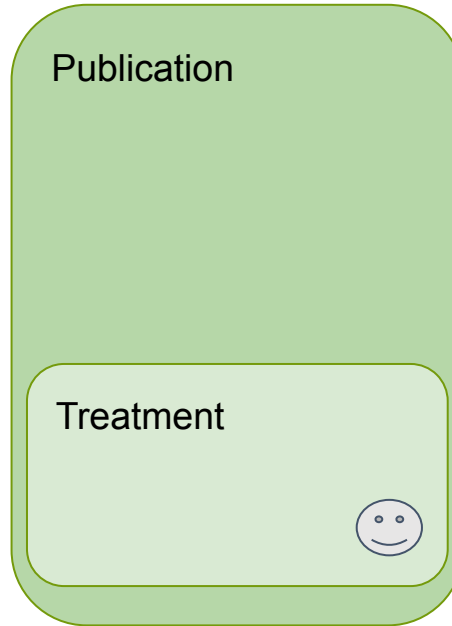
Digital accessible knowledge in biodiversity (DAK)

- Data understandable by human and actionable by machine reflecting the growth of our knowledge and interdependence.
- Editorial structure including textflow, paragraphs, sections such as title, authors and affiliations, materials and methods, etc., tables, figures, bibliographic references and their citations in the text.
- Semantic annotated (e.g. using TDWG standards) data at the base of the biodiversity knowledge graph:
 - data about a taxon (taxonomic treatments with their nomenclature section)
 - cited previous treatments (treatment citations)
 - cited specimens (material citations)
 - named entities (persons, taxonomic names, accession -, collection -, institution - or specimen codes)
 - attributes including their persistent identifiers

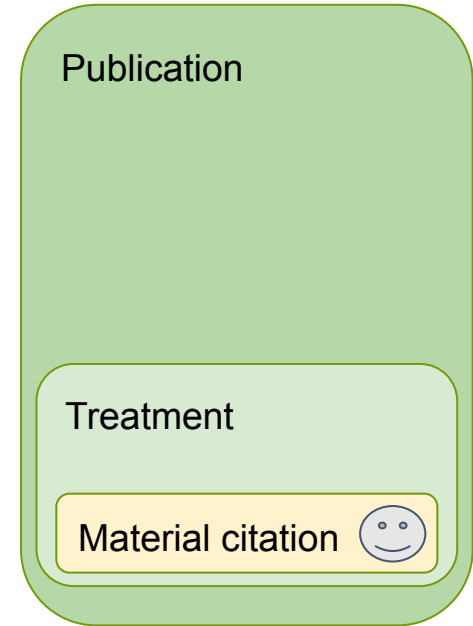
Annotations are context, defining the role of named entities



Person: **role**
Public. author, authority, collector

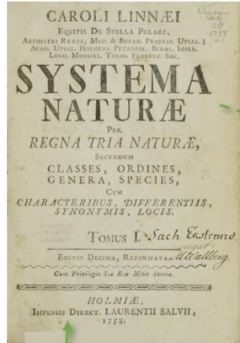


Person: **role**
Publication
Treatment authority, collector



Person: **role**
Publication
Treatment
Material citation collector

Digital Accessible Knowledge: Treatment



Apis mellifera L.

Linnaeus, 1758
4,819 taxa

LINNAEUS(2818) 1758
Type of
Apis mellifera L.
European honeybee

Type specimen = Standard

Natural Science Collections:

3.5 billion objects

10 million types

2 million standards



576 INSECTA HYMENOPTERA. Apis.

fucinata. 14. A. thorace flavescente subvillosa, abdomine nigro; cingulis quatuor albis.
Habitat in Terris australibus.
Rostrum subulatum, corvum, bivalve.

notata. 15. A. subpubescens fusca, abdomine cingulis quatuor curculis. *M. L. U.*
Habitat in Indiis.

Taxonomic treatment

taxonomic name *mellifera*. 17. A. pubescens, thorace subgrisco, abdomine fusco, pedibus posticis glabris utriusque margine ciliatis.
Fr. Juss. 1003.
Musf. inf. 2. *Swamm. bibl. t. 17. f. 1. operaria.*
Aldr. inf. 20. *f. 3. femina.*
Fonf. inf. 1. t. 2.
Raf. inf. 240.
Reaum. inf. 5. t. 2. f. 1. *operaria.*
f. 2. mas.
f. 3. femina.

Treatment citations

Distribution *Habitat* in Europæ arboribus cævis, frequentius culta.
REGINA (Femina) nitida, altilor, oblongior, aculeata.
FUCI (Mares) ad 1600, incermes, antennis 11-articulatis.
OPERARIE (Spadones) ad 20000, antennis 15-articulatis, ventriculis 2 mellis ceræque, aculeata.
Euse retrocurvum serrato venenato late intra vaginam cuspidatam.
RESPUBLICA Alvearii gynæceocratica est, dum Femina imperans semper meliosa, operariorum custodum satellitio assidue stipata, antennis obtusam Marem quæcumque salubrem ne comprimit cum illarum vitæ dispen. **Description.** vte ad 40000 sepe quotannis in aaupeatis jant cævis: primam noctem, deum mascula, demum novnulla feminea; his exclusis grada 2da supra calorem æstatis ordinaria, adolescent Larvæ sextidno, nent ad parietes cellæ, operuntur usque dum adultæ evolvant, atterogne die mellificant. Mares, ignotum pecus, incertis operantur ferentis diebus, galea dedita. Spadones operarii, acutis, indefessis, molle melle, ceram e pollice foram sedulo legunt: in miliaris spatio, quoties per calum licet, reducunt, fucus cereus

INSECTA HYMENOPTERA. Apis. 577

ctd. *cereus* firmus sexangulari-prismaticis, coque melle replens, feminam nutrituram, cellas inhabitatas repurgans. **Description** operantur, vigilas & stationes ad portas instituant, hostes arcent aculeis illis venenato sub ipso licet levis, marisque domum transactis mellei expellunt. **Biology & Ecology** operarii jant Alvearicæ, Hydrandines, Pavovori, Bæjovori, Mores, Crabrones, Velpæ, Pediculi, Apes jares, Mellouelle, Evmas &c. Flores mellificæ non omnes, præcipui sunt Echium, Boraginaceæ, Scrypyllum alii; imprimis Erica Securi, Thymus Sæcæ, Tilia Polonis, Rosmarinus, Absinthium Sardinis, Aconitum Ponticis &c. hinc prætium varium mellis: Vide Reaumurium, Adfonton.

subterita. 16. A. pubescens, thorace grisco, abdomine fusco, pedibus undique villosis.
Habitat in Terra sabulosa siccare, quam foraminibus pluribus parva remotis penetrant & distictæ adificiant plures.

variegata. 19. A. thorace abdomineque albo variegato.
Habitat in Europa.

rostrata. 20. A. labio superiore conico inflexo, abdominis fasciis glaucis repandis.
Fr. Juss. 999. Apis pedibus maxillisque flavis apice nigris, incisuris abdominis glabris margine nigro.
It. zool. 3:6.
Habitat in Europæ monticulis arenosis.

manicata. 21. A. nigra, pedibus anticis hirsutissimis, ano multidentato, abdomine maculis flavis.
Habitat in Europa.

4-dentata. 22. A. fusca abdominis cingulis quinque albidis, ano quadridentato: intermediis bidentis.
Swamm. bibl. t. 26. f. 4.
Habitat in Europa.

florifera. 23. A. nigra, abdomine subcylindrico incurvo apice bidentato, tibus posticis apice spinosis.
Habitat in Europa, per noctes floribus inhærens.

O o 24. A.

Each type has a published taxonomic treatment
Each taxonomic treatment is multiple times augmented

Tens of millions exist as part of ca 500 Million published pages of biodiversity literature
Each includes a plethora of facts

Treatment citation: Catalogue of life



PLAZI
TAXON CARE AT PRESSION

FactsMission

Home Advanced About Settings

SynoSpecies

Input Genus and species here:

2019 2021

- 2019
- 2021

Kiotina spatulata Wu, 1948
Hemacroneuria spatulata Li, 2019
Hemacroneuria spatulata Wu, 1948

Kiotina spatulata Wu, 1948

Defining treatment not yet on Plazi

Augmenting Treatments:

- Du, Yu-Zhou; Zhu, Bin-Qing; Huo, Qing-bo (2021) 3D6DA32CFFA5B65124F7FB445E7FBF50
 - Deprecates Hemacroneuria spatulata Li, 2019

Deprecating Treatments:

- Murányi, Dávid; Li, Weihai; Mo, Raorao (2019) 0384001F4228DD396287FB84FB04FA8E
 - Deprecates by Hemacroneuria spatulata Wu, 1948

Kiotina spatulata Wu, 1948, status revised

Hemacroneuria spatulata

Defining treatment not yet on Plazi

Deprecating Treatments:

- Huo, Qing-bo; Zhu, Bin-Qing; Du, Yu-Zhou (2021) 3D6DA32CFFA5B65124F7FB445E7FBF50
 - Deprecates by Kiotina spatulata Wu, 1948

Hemacroneuria spatulata

Defining Treatments:

- Murányi, Dávid; Li, Weihai; Mo, Raorao (2019) 0384001F4228DD396287FB84FB04FA8E
 - Deprecates Kiotina spatulata Wu, 1948

Remarks: The original description and illustrations of Wu (1948) show that this species has large paraprocts and paired sensilla basiconica patches on the male 10th tergum. However, these characteristics are not unique to *Hemacroneuria* and also exist in the genus *Sinacroneuria* . In

Kingdom Animalia
Phylum Arthropoda
Class Insecta
Order Plecoptera
Family Perlidae
Genus Kiotina
Species spatulata

Family Perlidae
Genus Hemacroneuria
Species spatulata

Wikidata Resource: <http://www.wikidata.org/entity/Q6387012>

- Taxon Name Kiotina spatulata
- Is subject of: https://ca.wikipedia.org/wiki/Kiotina_spatulata
- Is subject of: https://ceb.wikipedia.org/wiki/Kiotina_spatulata
- Is subject of: https://nl.wikipedia.org/wiki/Kiotina_spatulata

Currently accepted name

Synonymized taxonomic name

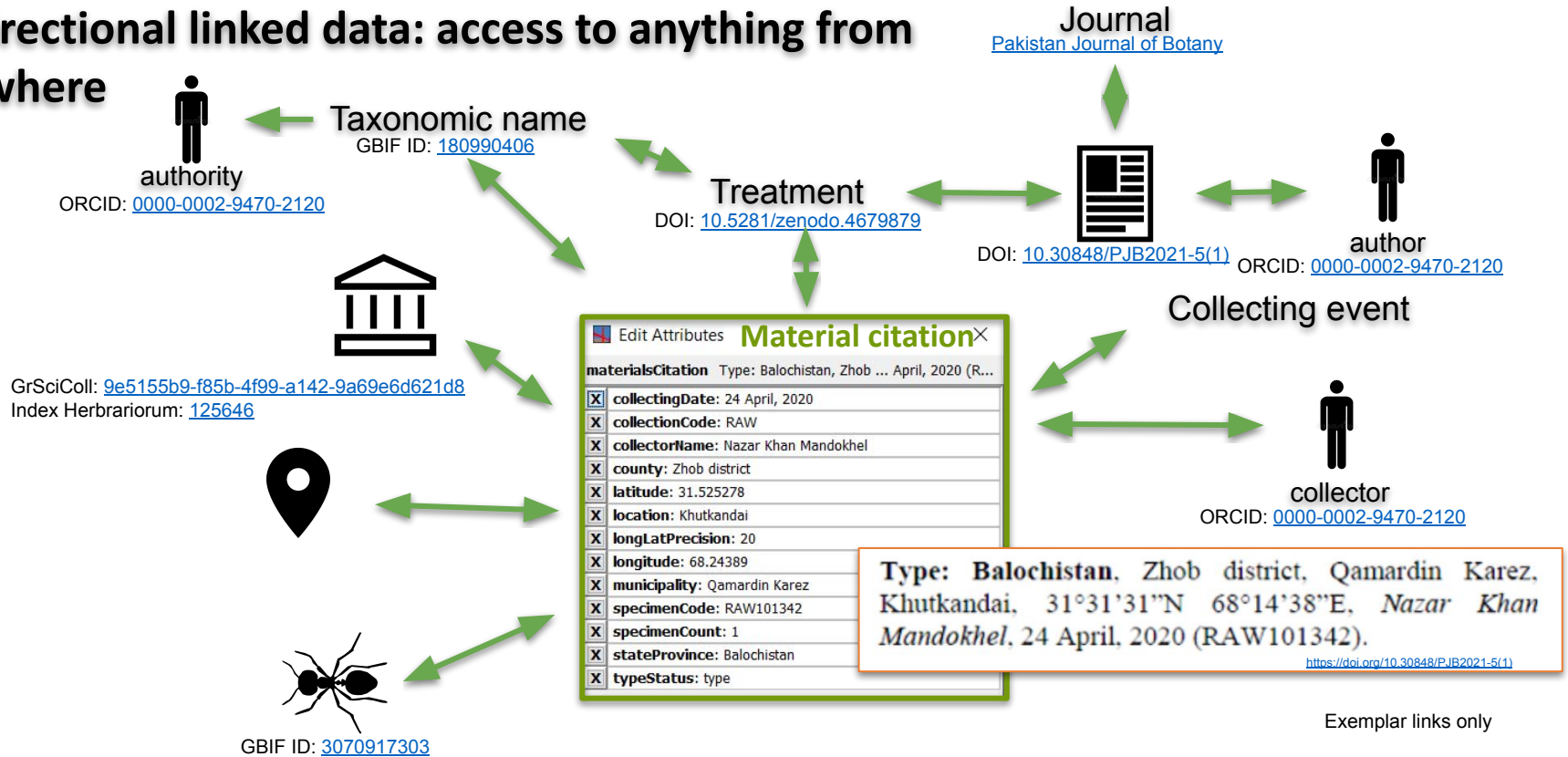
Original name or new combination

Missing treatment

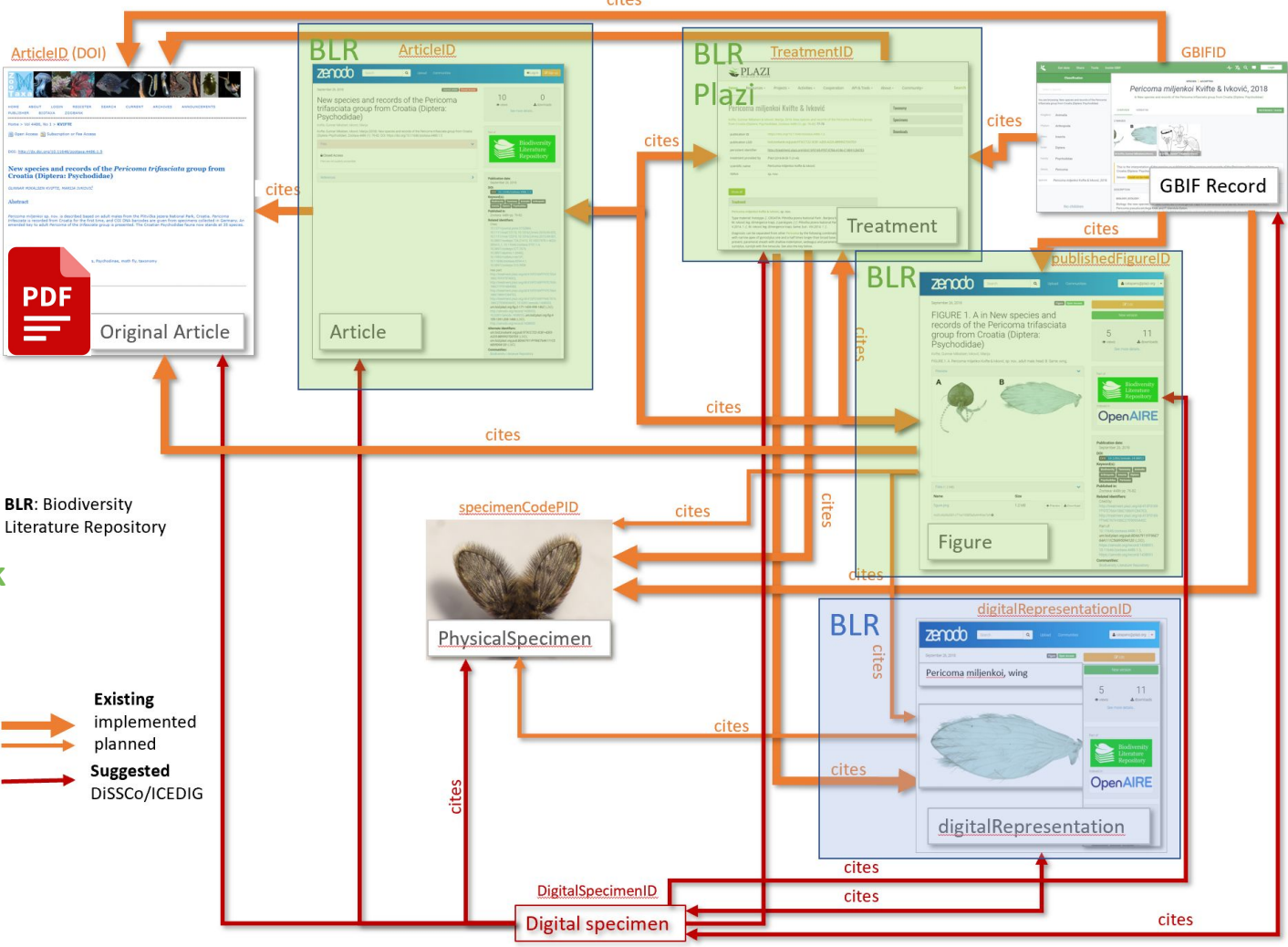
<https://synospecies.plazi.org/#Kiotina+spatulata>



Bi-directional linked data: access to anything from anywhere



Imagine the possible applications enabling making use of this big data?
Imagine the time saved if all these links are hyperlinks?

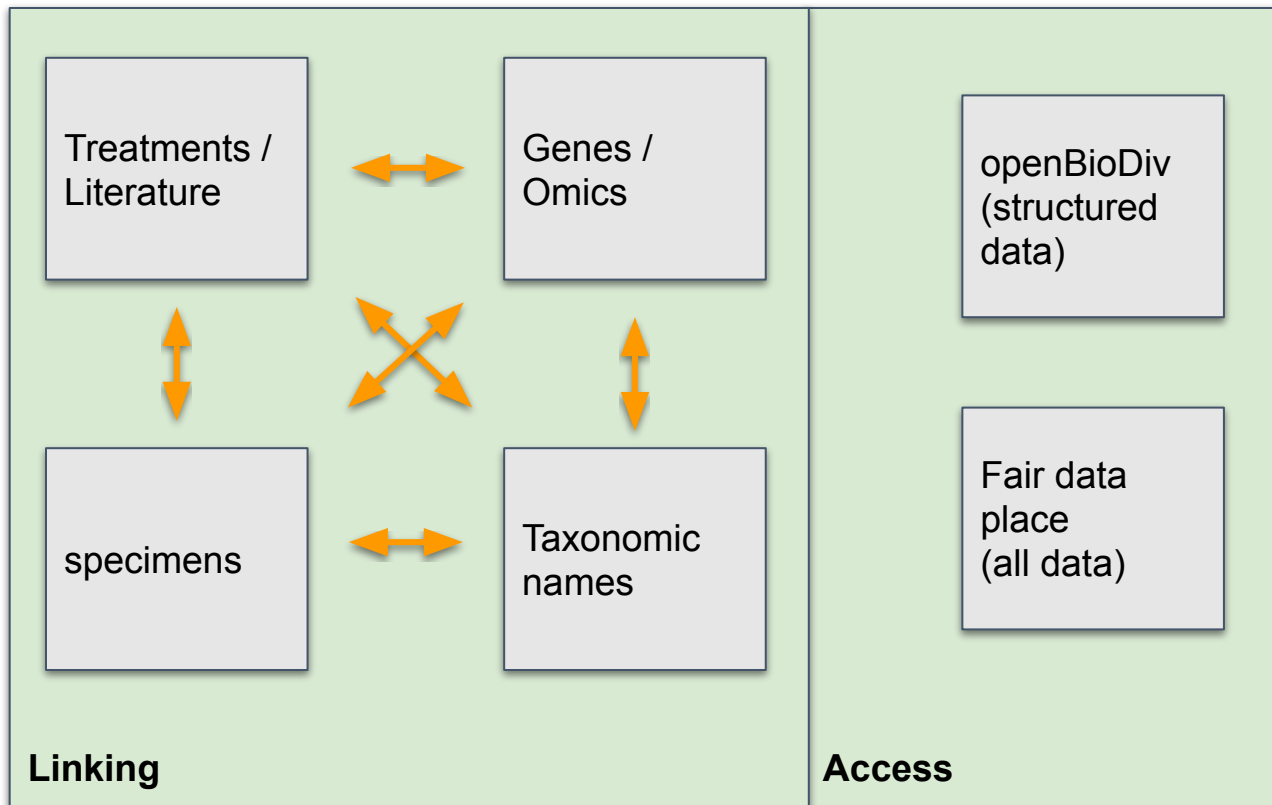


Focus on data in publications:
What's inside?
Citation network

BLR: Biodiversity Literature Repository

- Existing implemented
- planned
- Suggested
- DiSSCo/ICEDIG

BiCIKL: Linking and access

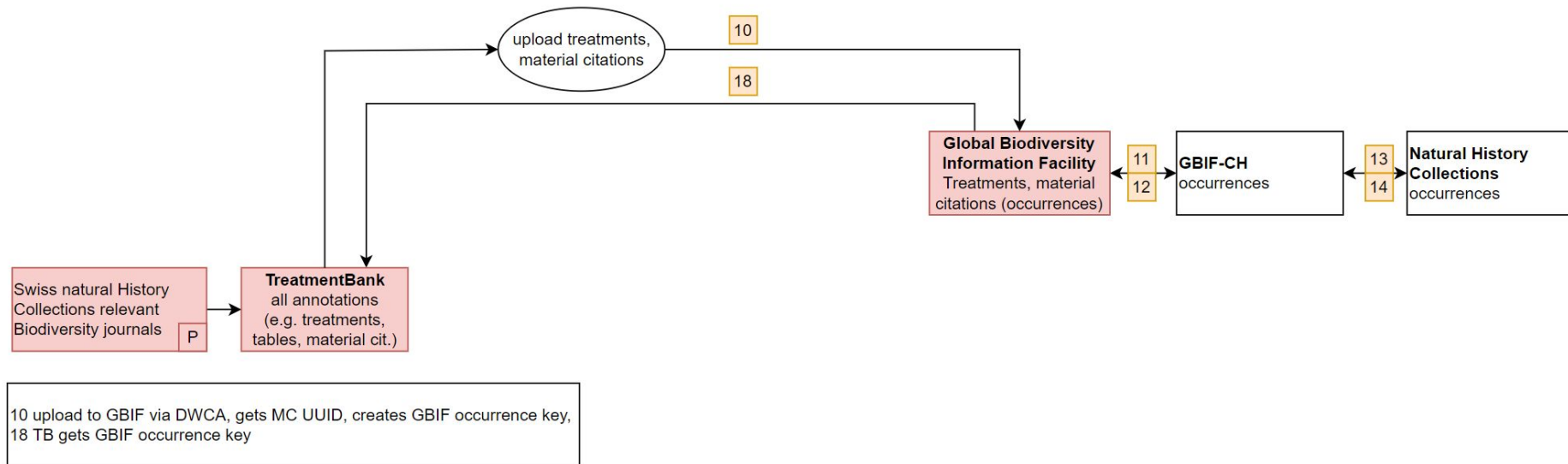


Linking specimens and literature via material citations

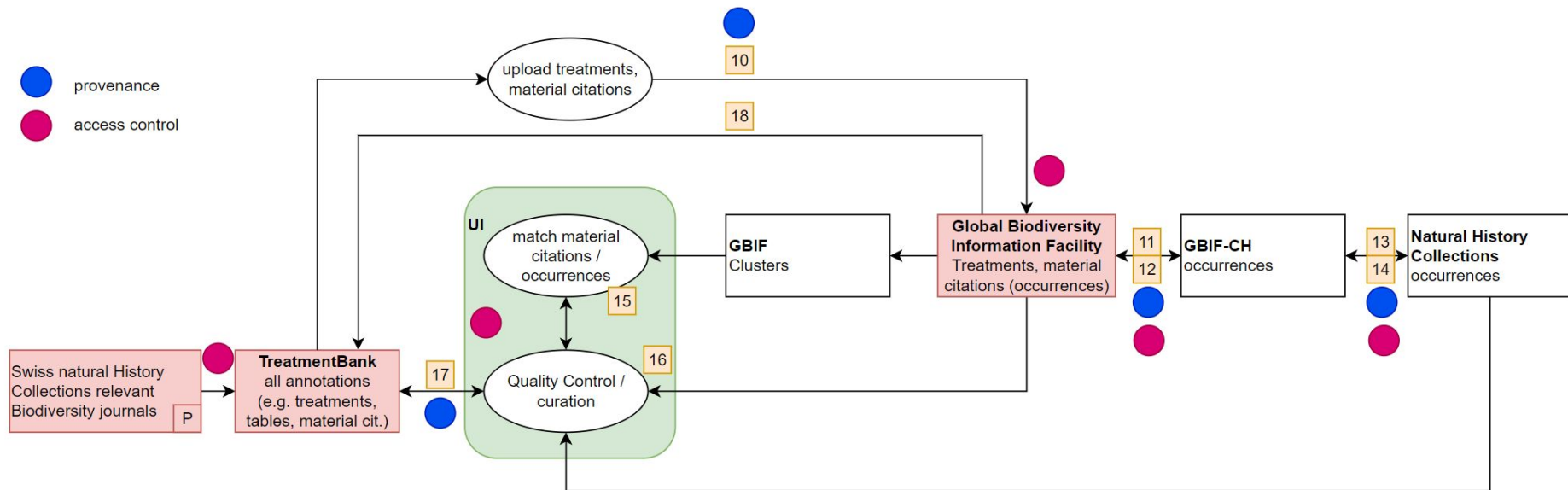


Bidirectional linking material citations

specimens / occurrences



Linking specimens and literature via material citations



Linking specimens with material citations



Occurrence 371220218

Caracladus zamoniensis Frick & Muff, 2009

Collected in Switzerland

Animalia > Arthropoda > Arachnida > Araneae > Linyphiidae > Caracladus

DETAILS CLUSTER

Species: *Caracladus zamoniensis* Frick & Muff, 2009
 Location: Switzerland
 Elevation: 1900m
 Basis of record: Preserved specimen
 Specimen type: Paratype

Dataset: NMDE - Arachnological collection
 Publisher: Naturhistorisches Museum Bern - NMDE

Record

Term	Interpreted	Original	Remarks
Basis of record	Preserved specimen	PreservedSpecimen	
Dataset ID	NMDE-AR	NMDE-AR	
Dataset name	NMDE - Arachnological collection	NMDE - Arachnological collection	
Institution code	NMDE	NMDE	
Institution ID	08020167-446-4023-8b55-7a7c429c9193	08020167-446-4023-8b55-7a7c429c9193	
Owner institution code	NMDE	NMDE	

Occurrence

Term	Interpreted	Original	Remarks
Catalogue number	NMDE-AR-6742	NMDE-AR-6742	
Individual count	7	7	
Occurrence ID	NMDE-AR-6742	NMDE-AR-6742	
Occurrence status	PRESENT	Present	
Recorded by	Frick, Muff, Klopfstein	Frick, Muff, Klopfstein	
Sex	3M+4F	3M+4F	Locality

Interpreted Original

Occurrence 3012376301

Caracladus zamoniensis Frick & Muff, 2009

Collected in Switzerland

Animalia > Arthropoda > Arachnida > Araneae > Linyphiidae > Caracladus

DETAILS CLUSTER

Species: *Caracladus zamoniensis* Frick & Muff, 2009
 Location: Switzerland
 Elevation: 1900m
 Basis of record: Preserved specimen
 Specimen type: Paratype

Dataset: Revision of the genus *Caracladus* with the description of *Caracladus zamoniensis*
 Publisher: http://treatment.plazi.org/id/6F3EED3970E0FF1ECA...
 Reference: http://treatment.plazi.org/id/6F3EED3970E0FF1ECA...
 Issues: Occurrence status inferred from material record

Specimens for the material citation 3012376301

PARATYPES, Switzerland, Olsons, Sur, Alp Fix, Salatagnas, 1900 m (46° 31'11.00"N, 9° 30'49.00"E), 3 2 4 5 24 X, 2007, killer skiving, close to the edge of a subalpine forest of Norway spruce (*Picea abies*), leg. H. Frick, P. Muff, G. Klopfstein, det. H. Frick (NMDE-AR-6742).

Key	Family	Genus	Species epithet	Latitude, longitude	Elevation	Locality	Country	Date	Coll. code	Catalog nb	Individual nb	Collector (recorded by)	Yes	No	Save
3012376301	Linyphiidae	Caracladus	zamoniensis	46.519722; 9.646111	1900	Alp Fix	Switzerland	24/10/2007	NMDE	AR6742	7	H. Frick & P. Muff & S. Klopfstein	<input type="checkbox"/>	<input type="checkbox"/>	<input type="button" value="Save"/>

3 specimens to curate

Key	Family	Genus	Specific epithet	Latitude, longitude	Elevation	Locality	Country	Date	Coll. code	Catalog nb	Individual nb	Collector (recorded by)	Yes	No	Save
371220218	Linyphiidae	Caracladus	zamoniensis	46.518359; 6.9467	1900	Sur, Alp Fix, Salatagnas, Lungeburg	Switzerland	24/10/2007		NMDE-AR-6742	7	Frick, Muff, Klopfstein	<input type="checkbox"/>	<input type="checkbox"/>	<input type="button" value="Save"/>
371220287	Linyphiidae	Caracladus	zamoniensis	46.518359; 6.9467	1900	Sur, Alp Fix, Salatagnas, Lungeburg	Switzerland	9/5/2006		NMDE-AR-8736	1	Patrick Muff	<input type="checkbox"/>	<input type="checkbox"/>	<input type="button" value="Save"/>
371220288	Linyphiidae	Caracladus	zamoniensis	46.518359; 6.9467	1900	Sur, Alp Fix, Salatagnas, Lungeburg	Switzerland	16/10/2005		NMDE-AR-8736	1	Patrick Muff	<input type="checkbox"/>	<input type="checkbox"/>	<input type="button" value="Save"/>

Back to list Save

link

Occurrence 3012376301

Caracladus zamoniensis Frick & Muff, 2009

Collected in Switzerland

Animalia > Arthropoda > Arachnida > Araneae > Linyphiidae > Caracladus

DETAILS CLUSTER

Species: *Caracladus zamoniensis* Frick & Muff, 2009
 Location: Switzerland
 Elevation: 1900m
 Basis of record: Preserved specimen
 Specimen type: Paratype

Dataset: Revision of the genus *Caracladus* with the description of *Caracladus zamoniensis*
 Publisher: http://treatment.plazi.org/id/6F3EED3970E0FF1ECA...
 Reference: http://treatment.plazi.org/id/6F3EED3970E0FF1ECA...
 Issues: Occurrence status inferred from material record

Record

Term	Interpreted	Original	Remarks
Basis of record	Preserved specimen	PreservedSpecimen	
Collection code	NMDE	NMDE	

Occurrence

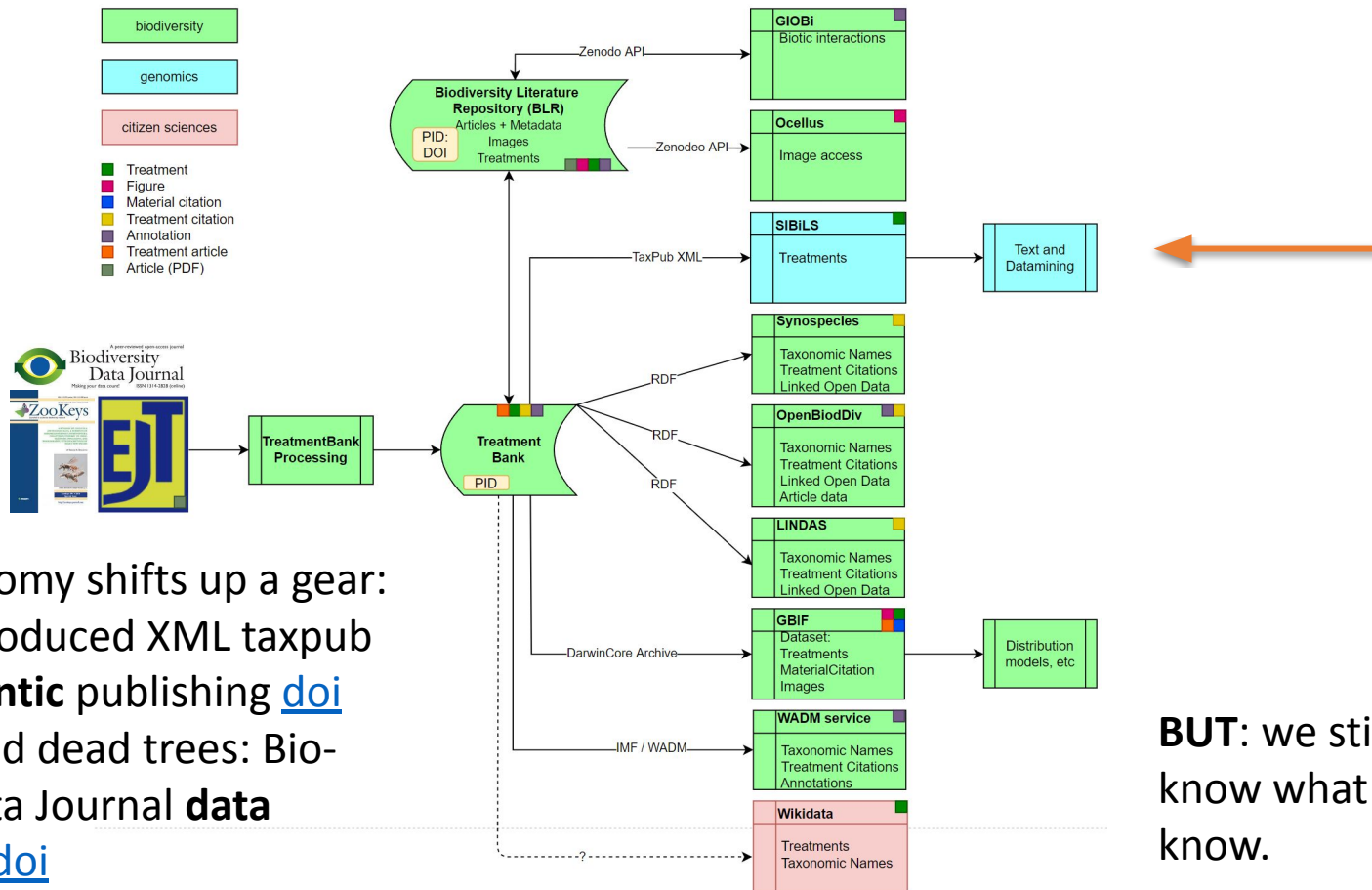
Term	Interpreted	Original	Remarks
Catalogue number	AR6742	AR6742	
Individual count	7	7	
Occurrence ID	FR3E-E35970E-FD11-CA025-A0FD-10C146-mc:CH1587E970E1FD8E-802795970E0364F088	FR3E-E35970E-FD11-CA025-A0FD-10C146-mc:CH1587E970E1FD8E-802795970E0364F088	
Occurrence status	PRESENT	PRESENT	Occurrence status inferred from material record
Recorded by	H. Frick & P. Muff & S. Klopfstein	H. Frick & P. Muff & S. Klopfstein	
Sex	MALF	male	

Event

Term	Interpreted	Original	Remarks
Day	24	Inferred	
Month	10	Inferred	
Year	2007	Inferred	
Event date	2007-10-24T00:00:00	2007-10-24	Altered

Identify as false

2022: Building a bridge between biodiversity and bioinformatics



2010: Taxonomy shifts up a gear:
Zookeys introduced XML taxpub
based **semantic** publishing [doi](#)

2013: Beyond dead trees: Bio-
diversity Data Journal **data**
publishing. [doi](#)

BUT: we still do not
know what we
know.

Bridging domains: Get treatments into the bioinformatics world



Table	JSON
 SIBiLS_upload_time	May 4, 2022 @ 16:22:54.560
f _id	http://treatment.plazi.org/id/011D87C1FFC1CD711FBFFB9DE3D5FE4D
f _index	baseline_alpha
# _score	1
f _type	_doc
f article-title	Observations on non-didemnid ascidians from Australian waters (1)
f material_citation_distribution	> Previously recorded (see Kott 1985): Western Australia (Cape Jaubert, Shark Bay, Cockburn Sound); South Australia (Upper Spencer Gulf); Vic Mile Beach, Bass Strait, Warnambool); New South Wales (Byron Bay); Queensland (Maroochydhore, Hervey Bay, Gladstone, southern Great Barrier il); Indonesia, Singapore, Sri Lanka, Japan. New records: Queensland (17.165 -17.935 ° S, 146.535- 146.8 ° E, 35-68 m)
f nomenclature-taxon-name	Polycarpa procera (Sluiter, 1885)
f publication-doi	http://dx.doi.org/10.1080/00222930600621601
f taxon_name_distribution	Polycarpa chinensis (Tokioka, 1967) P. procera Polycarpa chinensis
f taxon_name_reference_group	Styela procera Sluiter 1885, p 196 Polycarpa procera: Kott 1985, p 196
f text	> Styela procera Sluiter 1885, p 196 . Polycarpa procera: Kott 1985, p 196 and synonymy. Distribution Previously recorded (see Kott 1985): Western Australia (Cape Jaubert, Shark Bay, Cockburn Sound); South Australia (Upper Spencer Gulf); Vic Mile Beach, Bass Strait, Warnambool); New South Wales (Byron Bay); Queensland (Maroochydhore, Hervey Bay, Gladstone, southern Great Barrier il); Indonesia, Singapore, Sri Lanka, Japan. New records: Queensland (17.165 -17.935 ° S, 146.535- 146.8 ° E, 35-68 m) . Remarks
f text_distribution	> Distribution Previously recorded (see Kott 1985): Western Australia (Cape Jaubert, Shark Bay, Cockburn Sound); South Australia (Upper Spencer Gulf); Vic Mile Beach, Bass Strait, Warnambool); New South Wales (Byron Bay); Queensland (Maroochydhore, Hervey Bay, Gladstone, southern Great Barrier il); Indonesia, Singapore, Sri Lanka, Japan. New records: Queensland (17.165 -17.935 ° S, 146.535- 146.8 ° E, 35-68 m) . Remarks The species is plentiful in inter-reefal locations in northeastern Queensland and a wide range in the Indo-West Pacific between Japan in th i Lanka and around the southern coast of the Australian continent. The species range supports the view that the Australian continental shel
f text_reference_group	Styela procera Sluiter 1885, p 196 . Polycarpa procera: Kott 1985, p 196 and synonymy.
f treatment-bank-uri	http://treatment.plazi.org/id/011D87C1FFC1CD711FBFFB9DE3D5FE4D
f treatment_title	Polycarpa procera
f zenodo-doi	10.1080/00222930600621601

Treatment imported as
TaxPub/JATS to Swiss
Institute of Bioinformatics
Library Service (SIBiLS).

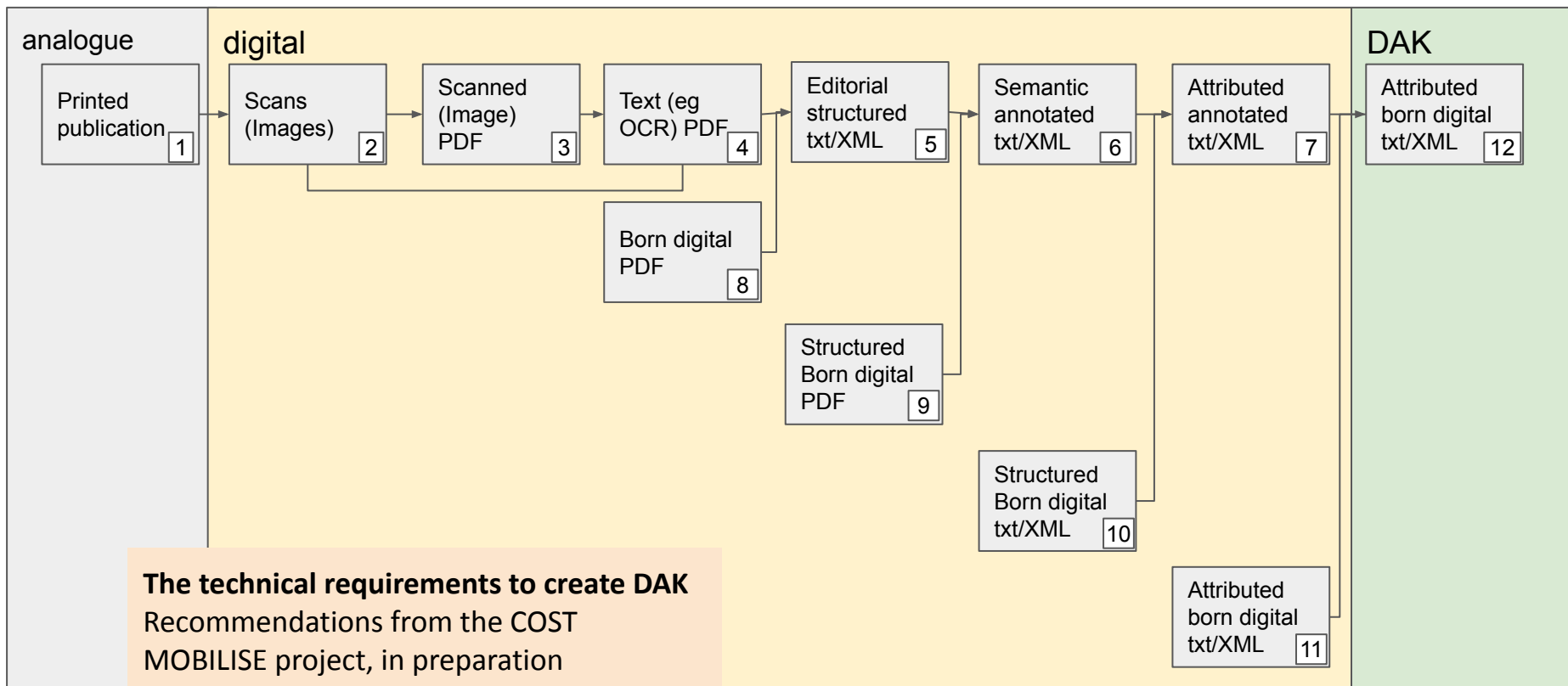
Bioinformatics: Access to
biodiversity data

Biodiversity Informatics:
Access to tools to annotated
and text and data mine
treatments (e.g. biotic
interactions terms)

[source](#)



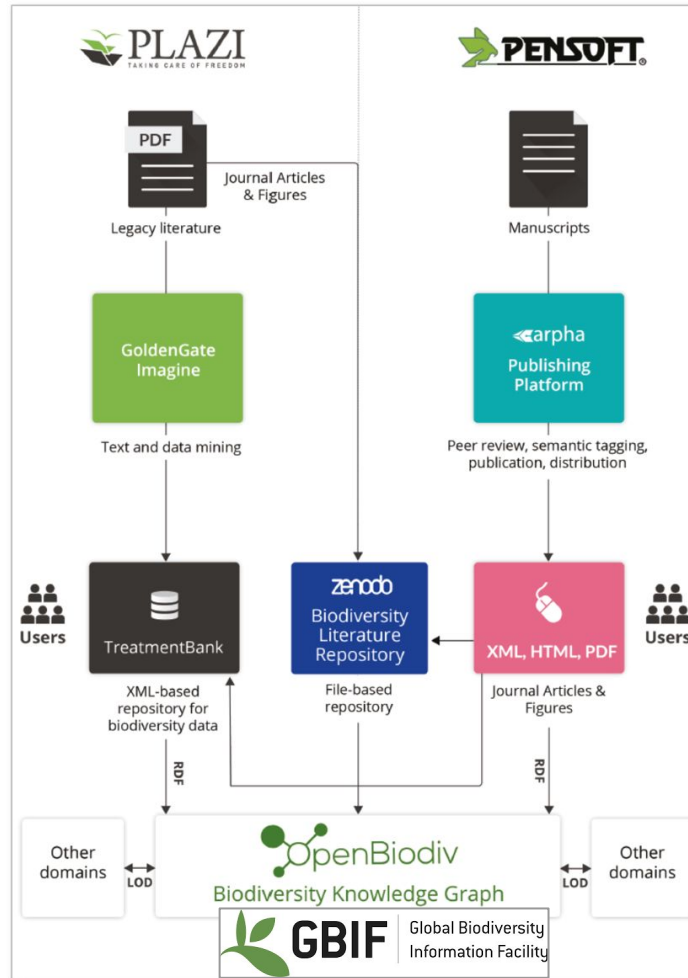
Origin of Digital Accessible Knowledge (DAK)





Legacy publications

Prospective publishing





Thank you!

Questions, answers, participation <https://github.com/plazi/community>
Introduction to digitizing taxonomic literature with Plazi [DOI](#)

Donat Agosti

agosti@plazi.org



swissuniversities



eBioDiv



datafutures



ARCADIA

SIBLS

SWISS INSTITUTE OF BIOINFORMATICS
LITERATURE SERVICES



PLAZI
TAKING CARE OF FREEDOM



Biodiversity
Literature
Repository



TreatmentBank





Further reading:

- Plazi: [further reading](#)
- Pensoft: doi: [10.3897/zookeys.50.543](https://doi.org/10.3897/zookeys.50.543) (e.g. Zookeys, BDJ)
- CETAF: doi: [10.5252/adansonia2018v40a1](https://doi.org/10.5252/adansonia2018v40a1) (e.g. European Journal of Taxonomy)

Data usage:

- Rivera-Quiroz et al. 2020, doi: [10.1038/s41598-020-72549-8](https://doi.org/10.1038/s41598-020-72549-8)
- Dikow & Agosti, 2015, doi: [10.3897/BDJ.3.e5707](https://doi.org/10.3897/BDJ.3.e5707)

Data access:

- Brief introduction into Treatmentbank stats: [PDF](#)
- Treatment statistics: <https://tb.plazi.org/GgServer/srsStats>
- Article statistics: <https://tb.plazi.org/GgServer/dioStats>
- Biodiversity Literature Repository API introduction: <https://developers.zenodo.org/>
- Biodiversity Literature Repository: <https://zenodo.org/communities/biosyslit/search?q=>

Applications based on and reuse of TreatmentBank and BLR data:

- Images via Ocellus: <https://ocellus.info/>
- Treatment citations via Synospecies: <https://synospecies.plazi.org/>
- TreatmentBank data in GBIF: <https://www.gbif.org/publisher/7ce8aef0-9e92-11dc-8738-b8a03c50a862>



[Plazi](#) is a Swiss based international association supporting and promoting the development of persistent and openly accessible scholarly digital taxonomic publications

NGO, SME owned by the NGO; Founded in 2008 as spin-off from a former US/DFG binational digital library award (2003-06); Supported by service contracts, EU-research funding, philanthropic funds, voluntary contributions. Plazi GmbH SME as service provider.

13 persons working for Plazi in Brazil, France, Germany, Spain, Switzerland, USA

Collaborations with Global Biodiversity Information Facility (GBIF), Zenodo at CERN, Pensoft Publishers Ltd, Consortium of European Taxonomic Facilities (CETAF), Swiss Institute of Bioinformatics (SIB), National Center for Biotechnology Informatics (NCBI), Muséum nationale d'Histoire Naturelle, Paris, Data Futures.

Support from Arcadia Fund, EU projects, service contracts, and voluntary contributions.

A mission of Plazi is to **discover, make accessible, and disseminate known biodiversity data**, not publications *per se* and to promote semantic enhanced publishing (TaxPub/JATS).