



pensoft.net

# Biodiversity data papers

Lyubomir Penev  
CEO & Founder, Pensoft  
Publishers  
[l.penev@pensoft.net](mailto:l.penev@pensoft.net)



# PENSOFT: Science Publisher & Technology Provider



- Founded in 1992, ca. 80 permanent staff, 3 offices in Sofia
- Publisher of 80+ academic journals & 1000+ scientific books
- 40+ journals on biodiversity and ecology
- Developer of the full-featured, open-access ARPHA Publishing Platform
- Multiple innovations in data publishing & semantic annotation of text and data

# Pensoft's Open Access Journals



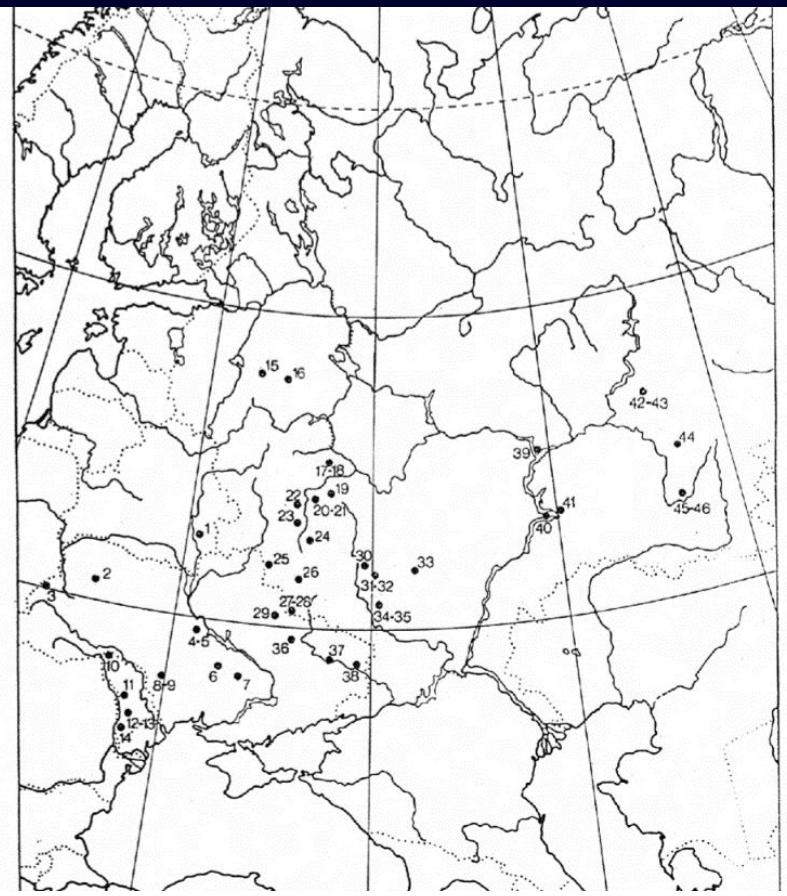


## Восемь лет в России

- Аспирант Института эволюционной морфологии и экологии животных им. А. И. Северцова (1985-1989)
- Старший научный сотрудник (1989-1992)
- Проект Почвенная мезофауна дубрав Русской равнины (1988-1991)

**ЮРИЙ ИВАНОВИЧ ЧЕРНОВ**

# Проект: Почвенная мезофауна дубрав Русской равнины





**Species diversity versus species composition in relation to climate and habitat variation: a case study on spider assemblages (Aranei) of the East European oak forests.<sup>1</sup>**

**Видовое разнообразие и видовой состав в связи с климатическими и биотопическим факторами: изучение группировок пауков (Aranei) в дубравах Восточно-Европейской равнины.<sup>1</sup>**

L.D. Penev\*, S.L. Esjunin\*\* and S.I. Golovatch\*

Л.Д. Пенев\*, С.Л. Есюнин\*\*, С.И. Головач\*

\* Institute of Evolutionary Morphology and Ecology of Animals, Russian Academy of Sciences, Leninsky prospekt 33, Moscow 117071 Russia.

\*\* Институт эволюционной морфологии и экологии животных РАН, Ленинский проспект, 33, Москва 117071 Россия.

\*\* Biological Faculty, Perm State University, ul. Bukireva 15, Perm GSP 614600 Russia.

\*\* Биологический факультет Пермского государственного университета, ул. Букирева, 15, Пермь ГСП 614600 Россия.

**KEY WORDS:** litter-dwelling spiders, biodiversity variation, environmental factors, oak forests, Russian Plain.

**КЛЮЧЕВЫЕ СЛОВА:** пауки подстилки, изменения биоразнообразия, факторы среды, дубравы, Русская равнина.

**ABSTRACT:** Two groups of parameters derived from both species diversity and species composition of spider assemblages of the East European Plain oak forests were tested in respect to large-scale variation in macroclimatic and habitat factors. The study is based on data equal in sample size and collecting efforts. Various diversity indices were taken as diversity measures. Spatial variation in the species composition was expressed by sample scores on the ordination axes derived from detrended correspondence analysis (DCA) and canonical correspondence analysis (CCA), both based on abundance and presence/absence data. Species composition is found to be a better indicator of spatial variation in macroclimate, giving higher and more stable values of correlation coefficients. Temperature parameters appear to be best explanatory variables for the patterns observed. The diversity measures vary with season and behave often in a different way for the spring, autumn or entire datasets. Only the total abundance shows a clear trend to a northerly and easterly decrease.

ких и биотопических факторов. Работа основана на данных, полученных по единой методике отбора проб и сбора материала. Для анализа биоразнообразия применялись различные индексы разнообразия. Пространственная изменчивость видового состава оценивалась по характеру расположения учетов обилия и присутствия-отсутствия на ординационных осиях с помощью детренного корреспондентного анализа (DCA) и канонического корреспондентного анализа (CCA). Отмечено, что видовой состав лучше отражает пространственную изменчивость макроклиматических факторов, обнаруживая более высокие и стабильные показатели коэффициентов корреляции. Температурные параметры оказались переменными, наилучшим образом объясняющие полученные данные. Меры биоразнообразия варьируют по сезонам, часто по разному отражая данные за весну, осень и весь полевой сезон. Лишь общее обилие обнаруживает четкую тенденцию к уменьшению к северу и востоку.

No	Site	Province	Date	Forest type	Forest age (yr)
1	Struki 12 km NE Buda-Koshelevo	Gomel	7-9.09.1989	interfluve Querceto-Carpinetum nudum	120-130
2	Dernovo 30 km E Lutsk	Volyn	12-14.09.1989	interfluve Querceto-Carpinetum nudum	100-150
3	Rostochye Reserve ca. 40 km WNW Lvov	Lvov	16-20.09.1989	upland Quercetum nudum	80-120
4	Peruny nr. Mikhaylovka 10 km S Kanev	Cherkassy	28-29.05.1988	Quercetum aegopodioides on fluvial terrace	80-120
5	Peruny nr. Mikhaylovka 10 km S Kanev	Cherkassy	28-29.09.1990	Quercetum aegopodioides on fluvial terrace	100-120
6	Chernyi Les nr. Znamenka	Kirovograd	20-22.05.1988	interfluve Fraxineto-Quercetum aegopodioides	100-200
7	Komissarovski Les 12 km W Volnogorsk	Dnepropetrovsk	19-21.05.1990	interfluve Querceto-Fraxinetum mixtoherboso-urticosum in ravine	80-120
8	Savranski Les nr. Polyanetsko	Odessa	31.05-2.06.1988	interfluve Quercetum mixtoherboso-aegopodioides	120-200
9	Savranski Les nr. Polyanetsko	Odessa	4-7.10.1989	interfluve Quercetum mixtoherboso-aegopodioides	120-200
10	Rossoshanskii Les 3 km N Briceani	Moldova	29-31.05.1990	interfluve Quercetum mixtoherboso-urticosum	70-120
11	Kodry Reserve nr. Lozova	Moldova	7-9.06.1988	upland Quercetum petraeae aegopodioides	80-100
12	Kotovskii Les nr. Kotovsk	Moldova	26-29.09.1989	upland Quercetum pubescens cotinosum	40-70
13	Kotovskii Les nr. Kotovsk	Moldova	1-2.06.1990	upland Quercetum pubescens cotinosum	40-70
14	Tigechskii Les 5 km S Tigceci	Moldova	5-6.06.1990	upland Quercetum nudum	70-150
15	Sokolovo 15 km E Dubniki	Novgorod	31.08-3.09.1989	interfluve Quercetum mixtoherboso-oxalidosum	200-220
16	Dubniki 16 km NNE Udomlya (=Kalinin)	Tver	4-6.09.1990	floodland Quercetum caricosum	150-250
17	Sharapovo 50 km SW Moscow	Moscow	9-10.09.1990	interfluve Quercetum aegopodioides-caricosum pilosae	120-150
18	Sharapovo 50 km SW Moscow	Moscow	23-24.05.1991	interfluve Quercetum aegopodioides-caricosum pilosae	120-150
19	Vasilevskii 10 km N Venec	Tula	7-10.05.1988	interfluve Querceto-Tilietum equisetosum	50-70
20	Tulskiye Zascki nr. Krapivna	Tula	3-7.05.1989	interfluve Querceto-Tilietum aegopodioides	200-300
21	Tulskiye Zascki nr. Krapivna	Tula	12-13.09.1990	interfluve Querceto-Tilietum aegopodioides	200-300
22	Chernysheno ca. 30 km WSW Kozelsk	Kaluga	28.04-1.05.89	interfluve Quercetum aegopodioides-alliosum	-100150
23	Trud ca. 20 km SE Ulyanovo	Kaluga	18-22.08.1991	interfluve Quercetum aegopodioides-galeobdolosum	150-400
24	Setukha 50 km E Orel	Orel	15-16.09.1990	interfluve Quercetum mixtoherbosum in ravine	100-120
25	Khinel 25 km SW Sevsk	Bryansk	18-19.09.1990	Quercetum caricoso-aegopodioides on fluvial terrace	100-170
26	Kazatskii Les 25 km SW Kursk	Kursk	5-9.05.1990	interfluve Quercetum aegopodioides-convalariosum	50-100
27	Les-na-Vorskle Reserve	Belgorod	15-20.05.1988	upland Quercetum aegopodioides-caricosum pilosae	200-300

## The Fauna and Zoogeography of Spiders Inhabiting Oak Forests of the East European Plain<sup>1)</sup>

(Arachnida: Araneae)

by

S.L. ESJUNIN, S.I. GOLOVATCH & L.D. PENEV \*

Die Tierwelt und Tiergeographie von Spinnen in Eichenwäldern der osteuropäischen Tiefebene  
(Arachnida: Araneae)

**Synopsis:** The spider fauna of oak forests of the East European Plain (492 species) is analyzed in terms of both landscape-zonal distribution and historical zoogeography. The fauna has been shown to be largely composed of extremely widespread species (over 76 %), with only a minor proportion characteristic of the nemoral (= broadleaved forest) biome and displaying a classical (south)west-(north)east decline toward the Urals. The role of both nemoral Carpathian and southern Ural refuges in spider faunogenesis in the Holocene is emphasized as based on both present-day distribution patterns and abundance gradients.

OIKOS 63: 180–192. Copenhagen 1992

## Qualitative and quantitative spatial variation in soil wire-worm assemblages in relation to climatic and habitat factors

Ljubomir D. Penev

Penev, L. D. 1992. Qualitative and quantitative spatial variation in soil wire-worm assemblages in relation to climatic and habitat factors. — Oikos 63: 180–192.

This paper considers large-scale variation in both species composition and alpa diversity of wire-worm assemblages (Coleoptera, Elateridae) of oak forests along a latitudinal gradient in Central Russia. The study is based on original data equal in both sample size and collecting effort. Changes in assemblage characteristics are analyzed in terms of species and/or community parameters such as species richness, species evenness, species diversity, species abundance and species-area parameters as well as to local habitat conditions. Species composition based on presence/absence data is found to be the best indicator of macro-environmental changes caused by climate. The species abundances are affected by both climatic trends and local habitat characteristics. There is little evidence for a strong effect of the environmental factors on the qualitative assemblage parameters such as species richness and abundance, and on various alpha diversity indices. Species diversity does not depend on the age of the habitat and the level of disturbance. The patterns observed suggest that species and community responses to the environment may play a significant role in explaining variation among isolated continental communities in addition to or instead of theories of island biogeography and interspecific competition.

L. D. Penev, Inst. of Animal Evolutionary Morphology and Ecology, Leninsky pr. 33, Moscow 117000, Russia (permanent address: Slavyanska Street 9, 2700 Blagoveshchensk, Russia)

Distribution and assemblage classification of spiders of the East European oak forests (Arachnida Aranei)<sup>1)</sup>.

Распределение и классификация группировок пауков дубрав Восточной Европы (Arachnida Aranei)<sup>1)</sup>.

S.L. Esjunin\*, L.D. Penev\*\* and S.I. Golovatch\*\*  
С.А. Есюнин\*, Л.Д. Пенев\*\*, С.И. Головач\*\*

\* Biological Faculty, Perm State University, ul. Bukireva 15, Perm GSP 614600 Russia

\* Биологический факультет Пермского государственного университета, ул. Букирева, 15, Пермь ГСП 614600 Россия.

\*\* Institute for the Problems of Ecology and Evolution, Russian Academy of Sciences, Leninsky prospect 33, Moscow 117071 Russia

\*\* Институт проблем экологии и эволюции РАН, Ленинский проспект, 33, Москва 117071 Россия.

**KEY WORDS:** litter-dwelling spiders, oak forests, Russian Plain, fauna, assemblages, phenology, distribution, classification, history.

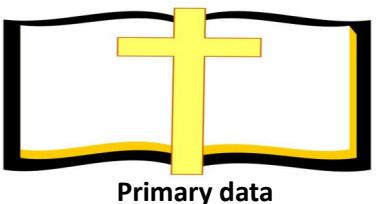
**КЛЮЧЕВЫЕ СЛОВА:** подстильочные пауки, дубравы, Русская равнина, фауна, группировки, фенология, распределение, классификация, история.

УДК 595.142.34

© 1994 г. Л. Д. ПЕНЕВ, А. И. ВАСИЛЕВ, С. И. ГОЛОВАЧ,  
Е. Ш. КВАВАДЗЕ

## ВИДОВОЙ СОСТАВ И КЛАССИФИКАЦИЯ ГРУППИРОВОК ДОЖДЕВЫХ ЧЕРВЕЙ (OLIGOCHAETA, LUMBRICIDAE) ДУБРАВ РУССКОЙ РАВНИНЫ

На основе 4-летних учетов дождевых червей в дубравах Русской равнины сделана классификация их группировок методом двухшагового анализа индикаторных видов (TWINSPAN). Сходство группировок в пределах полосы широколиственных лесов зависит прежде всего от региональных факторов. Зональность играет второстепенную роль. Как наиболее мощный меридиональный рубеж выступает Поволжье, затем — перечня Северского Донца и Днепра. Выделяются четыре основные группы дубрав по наследию дождевых червей: дубравы Заволжья и Приуралья, центра подзоны хвойно-широколиственных лесов, юга центральной лесостепи и запада Русской равнины. Полученные результаты обсуждаются в сравнении с традиционными ботанико-географическими классификациями дубрав и с аналогичными данными по паукам и диплонодам.



Drawings: slavenaneneya.com

© BioOne (www.bioone.org). Downloaded on [04/06/2016]. See the Terms and Conditions (http://www.bioone.org/page/termsconditions) on the BioOne website for rules of use; OA articles are governed by the applicable Creative Commons License

**Revision der kaukasischen Athous-Arten aus  
Untergattung *Heplatheca* RETTNER 1905  
(Coleoptera, Elachistidae)**

**Einführung**

Aren der Unterlassung: Hypothese KONTROLLE besagt, dass die sozialen und kulturellen Ressourcen der subdialektischen Regionen und gehören zu den wenig unterschätzten Grundvoraussetzungen für die Entwicklung von Sprach- und Kulturräumen (vgl. auch Dörr-David, Tatjana & Dietrich, 1998; sowie von Knefken (DÖRR, 1971, 1972; OLOFSSON, 1994). Es ist zu hoffen, dass die katalanische Region trotz langjähriger Forschung ähnlich neue Erkenntnisse her vorbringt, diesseit einer umfassenden Übersicht.

Vorbericht der katalanischen Untergruppe

Die katalanische Untergruppe hat sich mit dem Thema der sozialen und kulturellen Ressourcen beschäftigt und eine entsprechende Arbeit erarbeitet. Außerdem wurde die Arbeit des Arbeitskreises Sprache und Kultur untersucht, welche die sozialen Ressourcen benötigt reicht.

Kontakt:

Thomas Erspecker, Universität Konstanz  
Kontakt: Thomas.Erspecker@uni-konstanz.de

Themen-Exemplar dieser Gattung von Katalanen

PDF

BDF





# How to publish biodiversity data?



Home Articles About

About Pensoft

Books

E-Books

Blog

Journals

My tasks



Lyubomir Penev



Guidelines

Research Ideas and Outcomes 3: e12431  
<https://doi.org/10.3897/rio.3.e12431> (28 Feb 2017)

Reviewed

v1



XML

PDF



0



Contents

Article info

Citation

Metrics

Reviews 2

Related

Figs

Tabs

Refs

Cited

## Strategies and guidelines for scholarly publishing of biodiversity data

▼ Lyubomir Penev, Daniel Mietchen, Vishwas Shravan Chavan, Gregor Hagedorn, Vincent Stuart Smith, David Shotton, Éamonn Ó Tuama, Viktor Senderov, Teodor Georgiev, Pavel Stoev, Quentin John Groom, David Remsen, Scott C. Edmunds

### Abstract ▲

The present paper describes policies and guidelines for scholarly publishing of biodiversity and biodiversity-related data, elaborated and updated during the Framework Program 7 EU BON project, on the basis of an earlier version published on Pensoft's website in 2011. The document discusses some general concepts, including a definition of datasets, incentives to publish data and licenses for data publishing. Further, it defines and compares several routes for data publishing, namely as (1) supplementary files to research articles, which may be made available directly by the publisher, or (2) published in a specialized open data repository with a link to it from the research article, or (3) as a data paper, i.e., a specific, stand-alone publication describing a particular dataset or a collection of datasets, or (4) integrated narrative and data publishing through online import/download of data into/from manuscripts, as provided by the Biodiversity Data Journal.

### Article metadata

#### Data Publishing in a Nutshell

- Introduction
- What Is a Dataset
- Why Publish Data
- How to Publish Data
- How to Cite Data

#### Data Publishing Policies

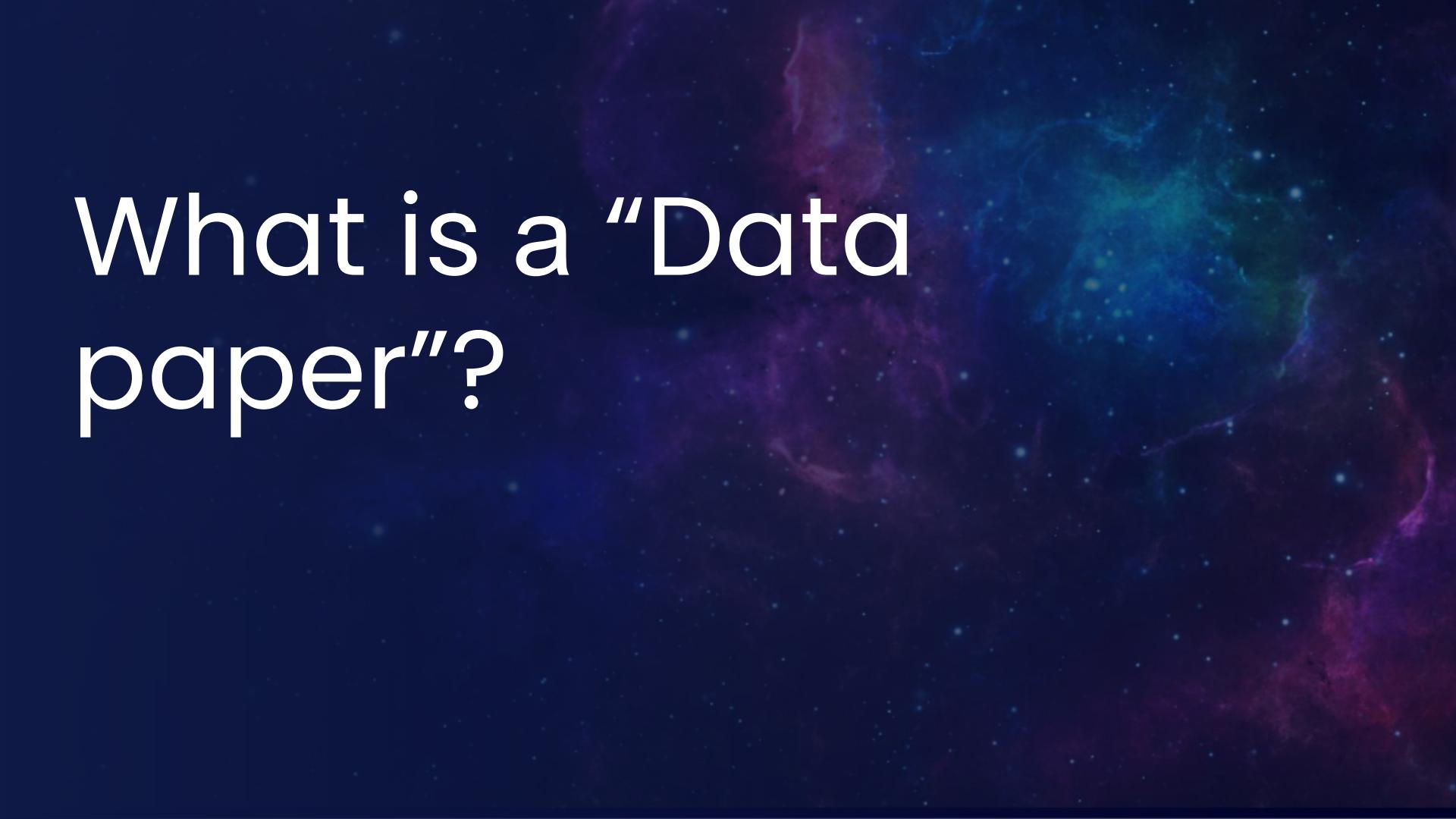
- General Policies for Biodiversity data
- Data Publishing Licenses

#### Data Deposition in Open Repositories

- General Information
- Taxonomy
- Species-by-Occurrence and Sample-Based data

# Three main models of data publishing

- Standalone data publishing (GBIF, GenBank, etc.)
- Data published together with a research article
  - within the article narrative (specimen records, tables)
  - supplementary file(s)
  - deposited in a repository and linked in the article
- Data published as **data papers**



What is a “Data  
paper”?

# The Data paper is:

- A standard, already widely accepted type of scholarly article
- It does not analyse data, it describes the data
- Extended “metadata description” of the data
- A key prerequisite of Open Science

# Data Paper for biodiversity (2011)



BMC Bioinformatics

Home   About   Articles   Submission Guidelines

Abstract  
Background  
The data paper  
Discussion  
Conclusions  
Declarations  
References

Volume 12 Supplement 15

[Data publishing framework for primary biodiversity data](#)

Research | Open Access

## The data paper: a mechanism to incentivize data publishing in biodiversity science

Vishwas Chavan <sup>†</sup>✉ and Lyubomir Penev <sup>†</sup>

<sup>†</sup>Contributed equally

BMC Bioinformatics 2011 12 (Suppl 15) :S2

<https://doi.org/10.1186/1471-2105-12-S15-S2> | © Chavan and Penev; licensee BioMed Central Ltd. 2011

Published: 15 December 2011

[Download PDF](#)

[Export citations ▾](#)

Metrics

Article accesses: 16537

Citations: 67 [more information](#)

Altmetric Attention Score: 76



[Share This Article](#)

# Examples of data papers



## Occurrences of Neuroptera and Raphidioptera in some regions in European Russia

Alexander Ruchin, Vladimir Makarkin, Mikhail Esin, Leonid Egorov, Oleg Artaev, Evgeniy Lobachev, Sergey Lukyanov, Vasilii Anikin, Anatoliy Khapugin, Gennadiy Semishin

### Abstract ▾

### Background

The document presents an extensive set of data on the occurrence of **Neuroptera** and **Raphidioptera** in some regions of European Russia. The results of our own research, as well as scientific collections, have been processed. The data were collected in 17 regions. In our own research, we used different ways to obtain information, which allowed us to collect extensive material for the dataset. This dataset provides valuable information about the biodiversity of **Neuroptera** and **Raphidioptera**, the abundance of each taxon collected and the time of taxon collections.

### New information

Our dataset contains up-to-date information on the occurrence of **Neuroptera** and **Raphidioptera** in

Contents Article Info Cite Metrics Comment Related

Figs Tabs Taxa Refs Cited Nanopubs

### Article metadata

#### Introduction

#### General description

#### Sampling methods

#### Geographic coverage

#### Taxonomic coverage

#### Temporal coverage

#### Usage licence

#### Data resources

#### Additional information

##### Observers

#### Acknowledgements

#### References

Subscribe to email alerts for current Article's categories

Email

Sign up

## Data Paper

Biodiversity Data Journal 12: e117169  
<https://doi.org/10.3897/BDJ.12.e117169> (16 May 2024)

XML

PDF



# The InBIO Barcoding Initiative Database: DNA barcodes of Portuguese moths

▼ Sónia Ferreira, Martin F. V. Corley, João Nunes, Jorge Rosete, Sasha Vasconcelos, Vanessa A. Mata, Joana Veríssimo, Teresa L Silva, Pedro Sousa, Rui Andrade, José Manuel Grosso-Silva, Catarina J. Pinho, Cátia Chaves, Filipa MS Martins, Joana Pinto, Pamela Puppo, Antonio Muñoz-Mérida, John Archer, Joana Pauperio, Pedro Beja

## Abstract ▲

## Background

The InBIO Barcoding Initiative (IBI) Dataset - DS-IBILP08 contains records of 2350 specimens of moths (*Lepidoptera* species that do not belong to the superfamily *Papilioidea*). All specimens have been morphologically identified to species or subspecies level and represent 1158 species in total. The species of this dataset correspond to about 42% of mainland Portuguese *Lepidoptera* species. All specimens were collected in mainland Portugal between 2001 and 2022. All DNA extracts and over 96% of the specimens are deposited in the IBI collection at CIBIO, Research Center in Biodiversity and Genetic Resources.

## New information

Contents Article Info Cite Metrics Comment Related

Figs

Taxa

Data

Refs

Cited

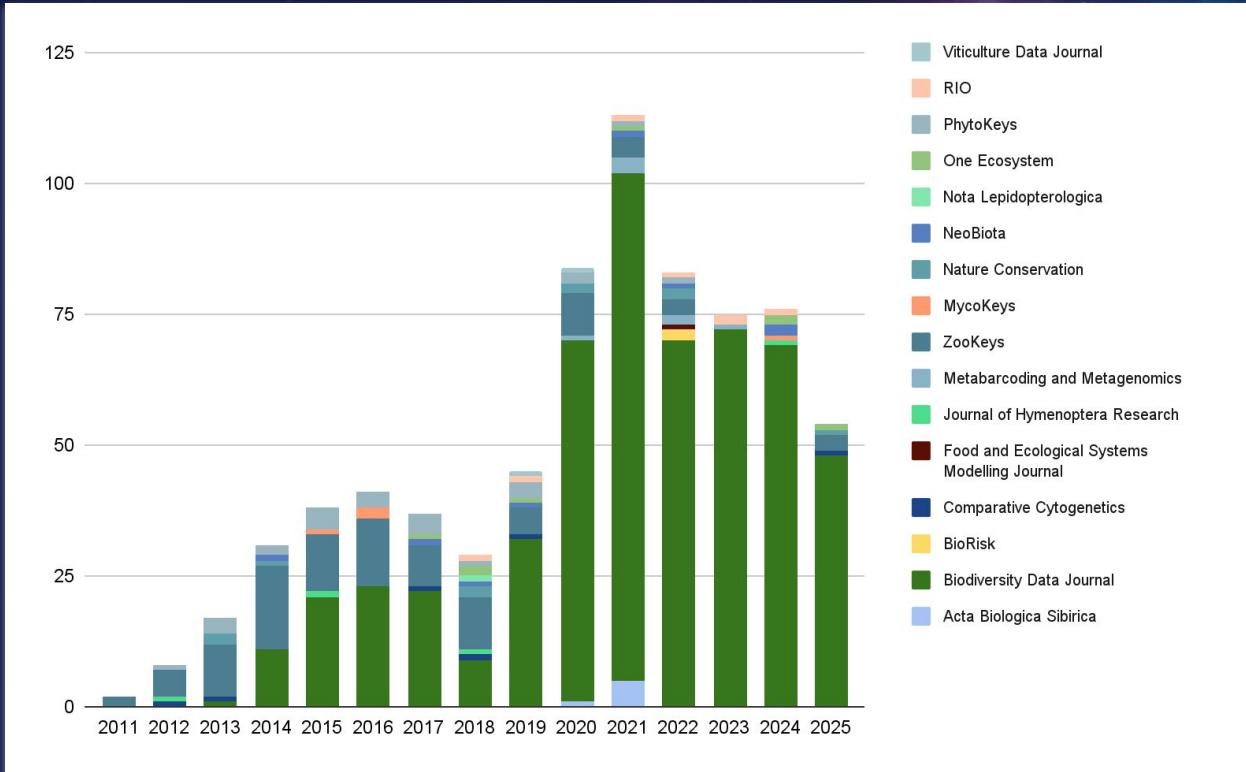
Nanopubs

[Article metadata](#)[Introduction](#)[General description](#)[Project description](#)[Sampling methods](#)[Geographic coverage](#)[Taxonomic coverage](#)[Temporal coverage](#)[Collection data](#)[Usage licence](#)[Data resources](#)[Acknowledgements](#)[References](#)[Supplementary files](#)

Subscribe to email alerts for current Article's categories

[Sign up](#)

# Data papers in Pensoft's journals



More than  
**700**  
Data papers  
published  
in Pensoft's  
journals

# Data publishing benefits all

- Data authors, data managers and their institutions: credit through citations, registration of priority and “data ownership”
- Science managers: measure impact of published data through the data paper citation metrics; proves the Open Science agenda
- Publishers: drive visits and citations of data to their journals
- Data users: facilitate data discovery, use and re-use
- Society in general: multiplies the public investments in data collecting and maintenance

# Data publishing benefits also AI

To understand the complexity of past, recent and future changes in biodiversity and natural environments

**the use of LLMs and AI tools**  
should be based on  
**adequately curated, semantically structured and interlinked biodiversity data**



## Data Paper

Biodiversity Data Journal 5: e11764  
<https://doi.org/10.3897/BDJ.5.e11764> (21 Mar 2017)

XML PDF



# A global map of saltmarshes

Chris J Mcowen, Lauren V Weatherdon, Jan-Willem Van Bochove, Emma Sullivan, Simon Blyth, Christoph Zockler, Damon Stanwell-Smith, Naomi Kingston, Corinne S Martin, Mark Spalding, Steven Fletcher

## Abstract ▾

## Background

Saltmarshes are extremely valuable but often overlooked ecosystems, contributing to livelihoods locally and globally through the associated ecosystem services they provide, including fish production, carbon storage and coastal protection. Despite their importance, knowledge of the current spatial distribution (occurrence and extent) of saltmarshes is incomplete. In light of increasing anthropogenic and environmental pressures on coastal ecosystems, global data on the occurrence and extent of saltmarshes are needed to draw attention to these critical ecosystems and to the benefits they generate for people. Such data can support resource management, strengthen decision-making and facilitate tracking of progress towards global conservation targets set by multilateral environmental agreements, such as the Aichi Biodiversity Targets of the United Nations' (UN's) Strategic Plan for Biodiversity 2011-2020, the Sustainable Development Goals of the UN's 2030 Agenda for Sustainable Development and the Ramsar Convention.

## New information

Contents Article Info Cite Metrics Comment Related

Tabs

Refs

Cited

Nanopubs



Crossref (296)  
Cited-by Linking

Scopus (317)



Europe PM

M. Aranda, G. Peralta, J. Montes, F.J. Gracia, G.S. Flavash, T.J. Bouma, D. van der Wal (2022)

**Salt marsh fragmentation in a mesotidal estuary: Implications for medium to long-term management.** Science of The Total Environment 846: 157410.

Crossref DOI: [10.1016/j.scitotenv.2022.157410](https://doi.org/10.1016/j.scitotenv.2022.157410)

Daniel M. Alongi (2020)

**Carbon Balance in Salt Marsh and Mangrove Ecosystems: A Global Synthesis.** Journal of Marine Science and Engineering 8: 767.

Crossref DOI: [10.3390/jmse8100767](https://doi.org/10.3390/jmse8100767)

Ralph J. M. Temmink, Leon P. M. Lamers, Christine Angelini, Tjeerd J. Bouma, Christian Fritz, Johan van de Koppel, Robin Lexmond, Max Rietkerk, Brian R. Silliman, Hans Joosten, Tjisse van der Heide (2022) **Recovering wetland biogeomorphic feedbacks to restore the world's biotic carbon hotspots.** Science 376: .

Crossref DOI: [10.1126/science.abn1479](https://doi.org/10.1126/science.abn1479)

Subscribe to email alerts for current Article's categories

Sign up

# Open Data Publishing at Pensoft

**DATA**

publishing

audit & peer review

deposition

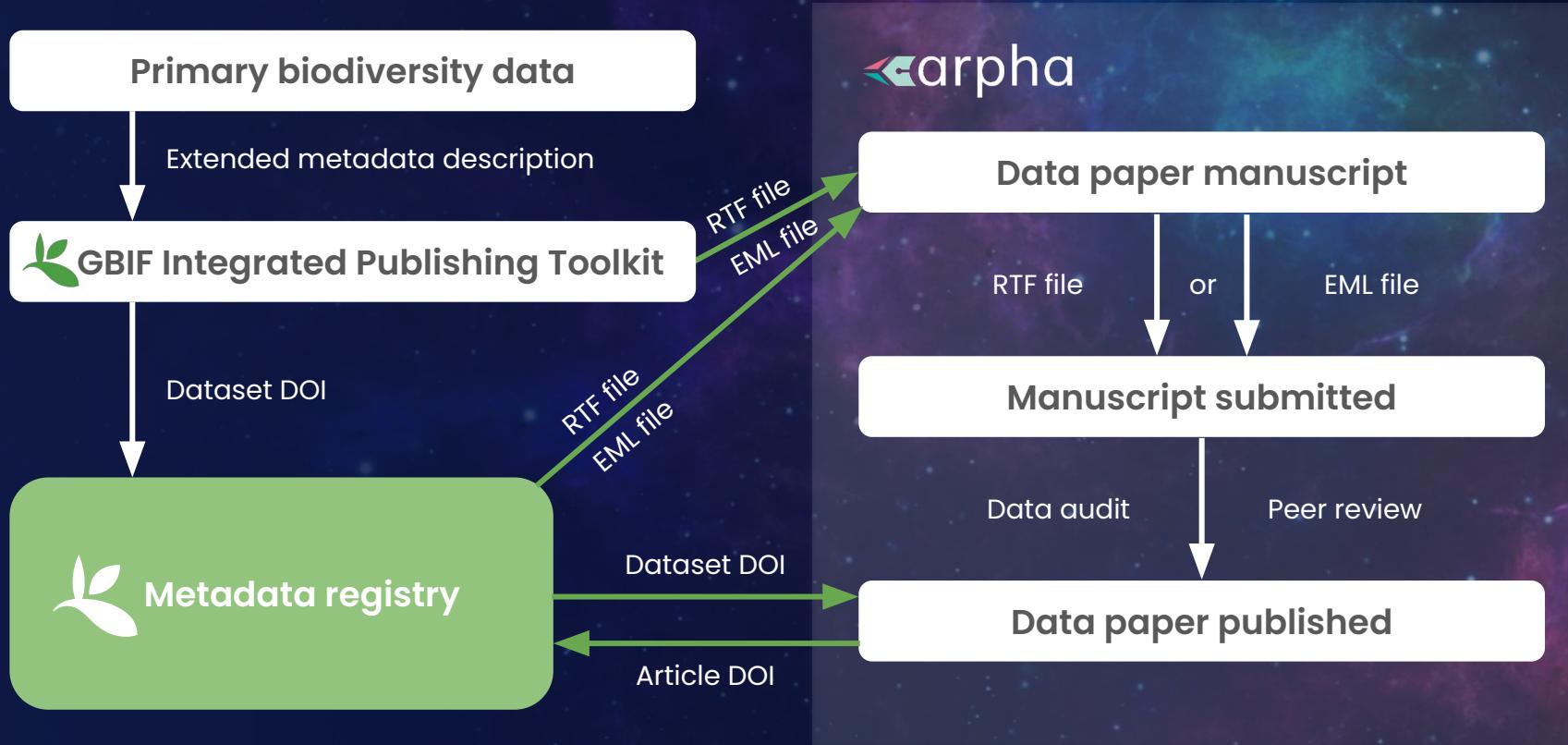
reuse

mobilization

# Publishing GBIF data with Pensoft

Improving data quality

# Data Paper workflows



# Download EML or RTF metadata file

A checklist to the wasps of Peru (Hymenoptera, Aculeata)

Latest version published by ZooKeys on 17 February 2011

The first checklist to the 225 genera and 1169 reported species of Peruvian entomological collections and include locality information provided when available. The occurrence data are published in 10.3897/zookeys.15.196.app.3.ds, and 10.3897/zookeys.15.196.app.4.ds. Following new combinations are proposed: Ancistrocerus (Ancistrocerus) 1912.

GBIF [DWC-A](#) [EML](#) [RTF](#) [Versions](#)

Data Records

The data in this occurrence resource has been published in more data tables. The core data table contains 2,910 records.

This IPT archives the data and thus serves as the data source. It lists other versions of the resource that have been made available by its providers.

OCCURRENCE DATASET | REGISTERED JANUARY 27, 2022

A dataset of bird inventory records at Cloudberry Nature Reserve, Talamanca Mountains, Costa Rica, between March 2016 and May 2020.

Published by [Check List](#)  
Powell J

DATASET PROJECT METRICS ACTIVITY DOWNLOAD

40,263 OCCURRENCES 32 CITATIONS

GBIF annotated archive Recommended  
Source archive Darwin Core Archive  
GBIF annotated metadata EML

A compilation of bird inventory records from Cloudberry Nature Reserve on the Pacific slope of the Talamanca Mountains, Costa Rica. The reserve consists of primary forest patches in montane cloud forest.

Publication date: January 27, 2022  
Metadata last modified: January 27, 2022  
Hosted by: ZooKeys  
Licence: CC BY-NC 4.0  
How to cite: DOI 10.15468/73nqqr

40,263 Occurrences 100% With taxon match 100% With coordinates 100% With year

40,263 GEOFERENCED RECORDS

# Upload the GBIF EML file and create a data paper manuscript

The screenshot shows the carphatool interface for a Data Paper (Biosciences). The left sidebar contains navigation links for Authors, Contributors, Article metadata (Title, Abstract & Keywords, Classifications, Funder, Nanopublications), Introduction, General description, Project description, Sampling methods, Geographic coverage, Taxonomic coverage, Traits coverage (Data coverage of traits), Temporal coverage, Collection data, Usage licence, Data resources, Additional information, Acknowledgements, and Author contributions. The main content area displays the title "A checklist to the wasps of Peru (Hymenoptera, Aculeata)" by Teodor Georgiev. It includes a box for Corresponding author (Teodor Georgiev, preprint@pensoft.net), a copyright notice (© Teodor Georgiev), and an "OPEN ACCESS" logo. Below the title are sections for Abstract, Background, and Keywords. The Abstract section describes the first checklist for 225 genera and 1169 species. The Background section provides context about the survey and collections. The Keywords section lists Aculeata, Biodiversity, Gazetteer, Hymenoptera, Sampling coverage, and Peru. A toolbar at the top provides various editing and viewing options.

# Data audit & peer review

Improving data quality

# Author-performed data check



A peer-reviewed open-access journal

Biodiversity  
Data Journal

ISSN 1314-2828 (online)

Full Text Author Title

Submit manuscript

My drafts

My tasks

About

Articles

Topical collections

Nanopublications

Author Guidelines

Editorial Team

Contacts

## About

Author Guidelines

Data Publishing Guidelines

Data Quality Checklist and  
Recommendations

FAIR Data Checklist

Linked Data Table for Primary  
Biodiversity Data

Data Review Guidelines

Omics Data Papers

What is "omics" data?

Where do I deposit my omics  
data and metadata?

## CHECKLIST

### Characters

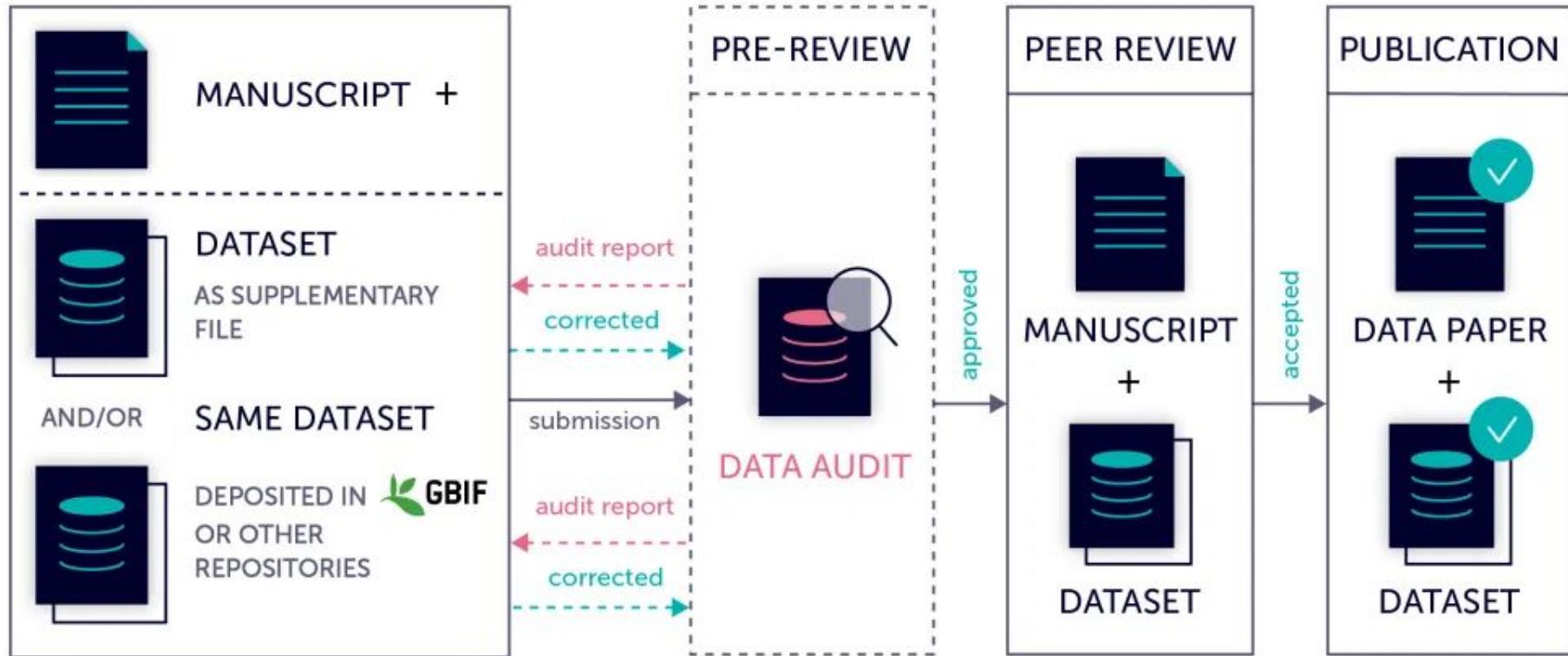
- The dataset is UTF-8 encoded
- The only characters used that are not numbers, letters or standard punctuation, are tabs and whitespaces
- Each character has only one encoding in the dataset
- No line breaks within data items
- No field-separating character within data items (tab-separated data preferred)
- No "?" or replacement characters in place of valid characters
- No Windows carriage returns
- No leading, trailing, duplicated or unnecessary whitespaces in individual data items

### Records

- No broken records, i.e. records with too few or too many fields
- No blank records



# Editor-performed data check



# Data audit report: list of errors



- inappropriately used fields
- non-compliance with the Darwin Core recommendations
- pseudo-duplicates
- data inconsistency

(7) *municipality* has "\_" for EC42F49A-68D5-4504-8F9E-0010859712A1.

(8) *locality* needs cleaning for the many pseudo-duplicates, e.g.

2 casco urbano, avda. del Brillante, nº 187, carril de la Huerta de los Arcos  
7 casco urbano, avda. del Brillante, nº 187, carril Huerta de los Arcos  
7 casco urbano, avda. del Brillante, nº 187, Carril Huerta los Arcos

and the many unnecessarily quoted entries, e.g.

"casa ""Rompealbardas"""  
""Villa Carmen"", ""El Calvario""

Also, *locality* is "\_" for CC465E40-9868-4B01-8D2B-5CB9AC747674 and 8547AA0D-682B-4848-B31F-0399427D51FA

(9) *decimalLatitude* errors:

1 30S266977.44  
1 37,91560°  
1 40.9449°  
1 41.9425N

Also, several entries have too many significant figures and should be rounded off, e.g. "37.0233172796695"

# Data audit report: recommendations



Data audit for technical evaluation of

## Vascular plants dataset of the COFC herbarium (University of Cordoba, Spain) (associated GBIF dataset)

Downloaded on 2019-06-19 from <https://www.gbif.org/dataset/837c0162-f762-11e1-a439-00145eb45e9a>

Dr Robert Mesibov (robert.mesibov@gmail.com; <https://www.datafix.com>)  
2019-06-20

### About this evaluation

Pensoft does a technical evaluation of the dataset (or datasets) referred to in the data paper. If the dataset has major problems, the data paper manuscript is referred to reviewers. If the dataset has minor problems, the dataset has been corrected.

To see what features of a dataset are checked in a technical evaluation, please go to

<https://zookeys.pensoft.net/about#DataQualityChecklistandRecommendations>

Please note that Pensoft does not check the details of the content of a dataset, for example, the scientific name, or whether the correct latitude/longitude is given for a locality.

**Recommendation.** The dataset associated with the manuscript has been processed by GBIF and the data paper could go on to review. However, there are many data problems in the GBIF upload, and I recommend to the authors that these problems be fixed and the data re-uploaded to GBIF for processing. The problems are detailed below by Darwin Core field in the field order in the dataset.

Many of the problems are not trivial and are causing data loss. For example, the decimalLatitude in FF92A873-601C-4360-86C7-5C9D483D6DAE is "30S266977.44". GBIF has rejected the location as "Coordinate invalid" (<https://www.gbif.org/occurrence/2235670578>).

**Recommendation.** The dataset associated with the manuscript has been processed by GBIF and the data paper could go on to review. However, there are many data problems in the GBIF upload, and I recommend to the authors that these problems be fixed and the data re-uploaded to GBIF for processing. The problems are detailed below by Darwin Core field in the field order in the dataset.

Many of the problems are not trivial and are causing data loss. For example, the decimalLatitude in FF92A873-601C-4360-86C7-5C9D483D6DAE is "30S266977.44". GBIF has rejected the location as "Coordinate invalid" (<https://www.gbif.org/occurrence/2235670578>).

In addition, the number of records in the dataset uploaded to GBIF (**verbatim.txt** in the user download) does not always agree with the number of records given in the data paper:

Family	Data paper	verbatim.txt
Asteraceae	8625	8625
Fabaceae	7929	7929
Poaceae	6324	5513 (as stated in paper, without 811 <i>Festuca</i> records)
Lamiaceae	3105	3087
Caryophyllaceae	2156	2156
Plantaginaceae	2023	750
Brassicaceae	1851	1852
Apiaceae	1707	1707
Ranunculaceae	1319	1318
Boraginaceae	1111	1289

Genus	Data paper	verbatim.txt
Centaurea	1538	1537
Trifolium	1406	1406
Erophila	769	756

# Data submitted to GBIF after publication



Home Articles About

About Pensoft

Books

Journals

Blog

My tasks Edit article



Teodor Georgiev



Taxonomy & Inventories

Biodiversity Data Journal 12: e124006

<https://doi.org/10.3897/BDJ.12.e124006> (25 Apr 2024)

XML

PDF



## A new species of *Otacilia* Thorell, 1897 (Araneae, Phrurolithidae) from Yintiaoling National Nature Reserve, Chongqing, China

▼ Changbin Zheng, Yannan Mu

### Abstract ▲

### Background

Phrurolithidae is a family of spiders with 395 species belonging to 26 genera distributed worldwide, of which 205 species belong to 17 genera was recorded in China.

### New information

Contents Article Info Cite Metrics Comment Related

Figs

Map

Taxa

Data

Refs

Cited Nanopubs

💡 Tables and Figures, if present, can be downloaded from the article.

Explore this dataset on GBIF

Download all occurrences as Darwin Core Archive

Download all treatments as Darwin Core Archive



# Data visualised on GBIF after publication

Get data How-to Tools Community About

TREATMENT ARTICLE | REGISTERED MAY 3, 2024

A new species of *Otacilia* Thorell, 1897 (Araneae, Phrurolithidae) from Yintiaoling National Nature Reserve, Chongqing, China

Meditated by [Biodiversity Data Journal](#)

Mu Y

DATASET TAXONOMY METRICS ACTIVITY DOWNLOAD

1 MATERIAL EXAMINED 1 RECORD

Phrurolithidae is a family of spiders with 395 species belonging to 26 genera distributed worldwide, of which 205 species belong to 17 genera recorded in China. A new species of the genus *Otacilia* Thorell, 1897 is described from Yintiaoling National Nature Reserve, Chongqing, China. Diagnosis, morphological description, photos of the habitus and genitalia of the new species are provided.

Publication date: April 24, 2024  
Metadata last modified: May 3, 2024  
Hosted by: Biodiversity Data Journal  
Licence: CC BY 4.0  
[How to cite](#) DOI: 10.3897/bdj.12.e124006

1 Occurrences	100% With taxon match	100% With coordinates	100% With year
1 Accepted names	0 Synonyms	85% Overlap with GBIF Backbone	71% Overlap with Catalogue of Life

1 GEOFERENCED RECORD



# Data visualised on GBIF after publication

Get data How-to Tools Community About

SPECIES | ACCEPTED

## Otacilia wuxi Zheng & Mu

In: Zheng C, Mu Y (2024) A new species of *Otacilia* Thorell, 1897 (Araneae, Phrurolithidae) from Yintiaoling National Nature Reserve, Chongqing, China. Biodiversity Data Journal 12: e124006. <https://doi.org/10.3897/BDJ.12.e124006>

Mediated through: Biodiversity Data Journal

TREATMENT VERBATIM SOURCE

Male: total length 5.01, carapace 2.19 long, 1.91 wide; abdomen 2.59 long, 1.67 wide. Eye sizes and interdistances: AME 0.14, ALE 0.15, PME 0.11, PLE 0.13, AME–AME 0.04, AME–ALE 0.02, PME–PME 0.16, PME–PLE 0.09, ALE–PLE 0.15. MOA 0.37 long, anterior width 0.30, posterior width 0.42. Clypeal height 0.19. Chelicerae with three promarginal and eight retromarginal teeth. Measurements of legs: I 8.86 (2.34+3.44+2.00+1.08), II 7.05 (1.87 +2.59+1.55+1.04), III 6.18 (1.70+1.93+1.58+0.97), IV 9.46 (2.52+2.95+2.62+1.37). Spination: tibia I pv 8 rv 8, tibia II pv 7 rv 7, metatarsus I pv 4 rv 4, metatarsus II pv 3 rv 3. Legs yellow. Carapace yellow, with several indistinct shapes resembling flowing water droplets beside fovea. Abdomen yellow, with a small, thin dorsal scutum and irregular black pattern anterior and four black chevron stripes posterior (Fig. 1A).

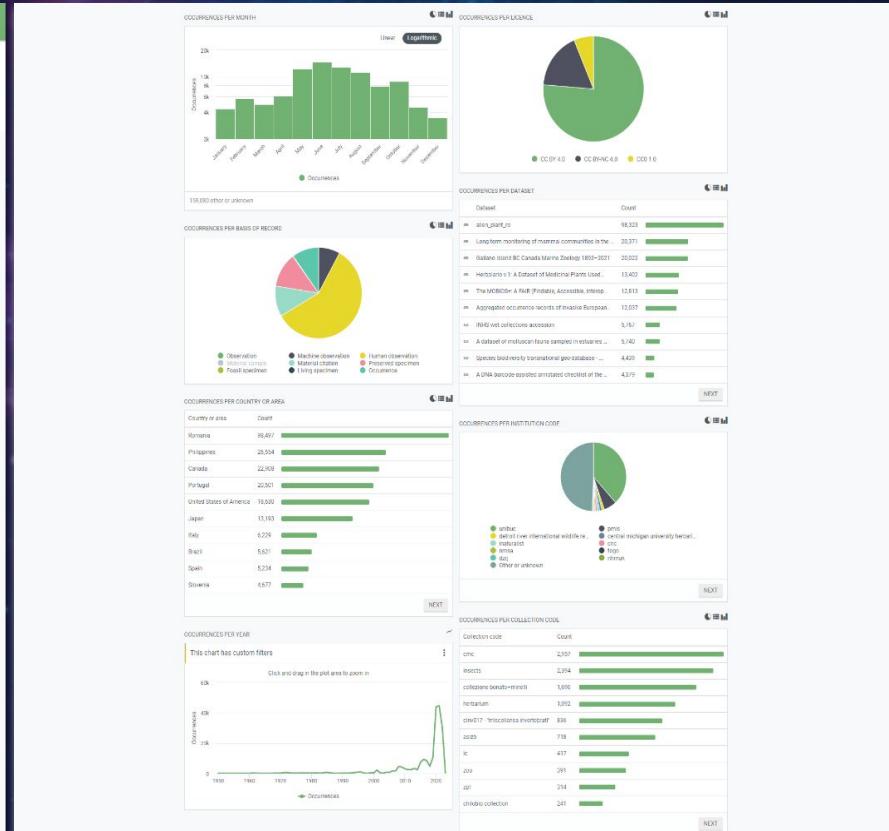
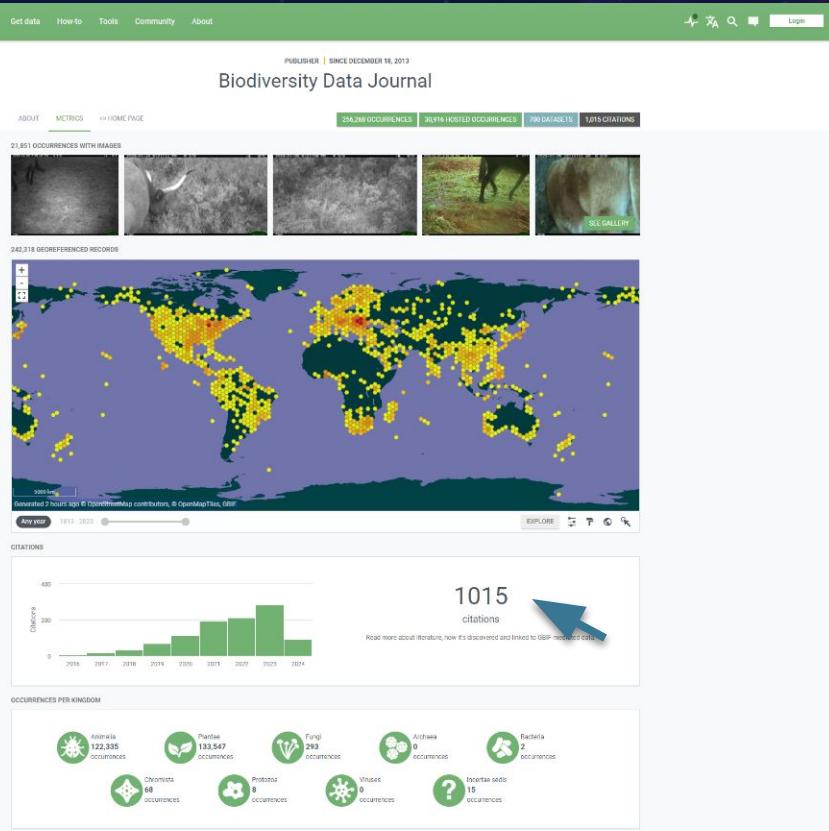
**Palp.** Femoral apophysis high, located at middle part of femur, well-developed (Fig. 1C and D). Dorsal tibial apophysis long and large, strongly curved as semi-elliptic, base wide, tapering from middle to tip (Fig. 1C and D); prolateral tibial apophysis distinct (Fig. 1B). Tegulum bean-shaped, wider than cymbium; tegular apophysis semicircular. Conductor membranous (Fig. 1B). Sperm duct obvious, tapering from retrolateral of tegulum to embolus. Embolus long, needle-like, strongly curved retrolaterally from basal part (Fig. 1B).

Female: unknown.

FIGURES

Figure 1  
*Otacilia wuxi* sp. nov., male, holotype. **A** habitus; **B** left palp, ventral view; **C** same, prolateral view; **D** same, retrolateral view. Abbreviations: C—conductor; DTA—dorsal tibial apophysis; E—embolus; FA—femoral

# Datasets from the Biodiversity Data Journal



# Import of occurrence records from GBIF into manuscripts

carpha writing tool

View dashboard Messages Collections Reviewers Email contributors Helpdesk Tips and tricks Tutorial Revision history

Comments Resolved Filter

Taxonomy & Inventories

Authors Contributors Article metadata Title Abstract & Keywords Classifications Funder Nanopublications

Introduction Materials and methods Data resources

Taxon treatments Cyclamen hederifolium Taxon name External Links Nomenclature Materials Treatment sections

Checklists Identification keys

Corresponding author: Teodor Georgiev (preprint@penssoft.net)

Teodor Georgiev OPEN ACCESS

Citation:

## Taxon treatment

*Cyclamen hederifolium* Aiton

Material Download as CSV

```
a. scientificName: Cyclamen hederifolium Aiton; taxonConceptID: NBNSYS0000003933; taxonomicStatus: ACCEPTED; taxonID: NBNSYS0000003933; kingdom: Plantae; phylum: Tracheophyta; class: Magnoliopsida; order: Ericales; family: Primulaceae; taxonRank: SPECIES; vernacularName: Sowbread; genus: Cyclamen; specificEpithet: hederifolium; continent: EUROPE; country: United Kingdom of Great Britain and Northern Ireland; countryCode: GB; stateProvince: England; locality: Copse Hill, Brighton and Hove, Brighton BN1 5EG, UK; decimalLatitude: 50.858986; decimalLongitude: -0.155622; geodeticDatum: WGS84; coordinateUncertaintyInMeters: 50; eventID: 24740609; eventDate: 2024-01-13; startDayOfYear: 13; endDayOfYear: 13; year: 2024; month: 1; day: 13; individualCount: 1; occurrenceDetails: http://api.gbif.org/v1/occurrence/4594204304; recordedBy: Rudling, M. Milly; occurrenceStatus: PRESENT; collectionCode: iNaturalist | UK and IOM data; basisOfRecord: HUMAN_OBSERVATION; occurrenceID: D4572408-6EA4-5F3B-9801-947D19864F23
```

# Parsing occurrence records using AI

Owner Contributors Untitled

SA

Print

Import

B I U X  $X^2$   $X^3$   $\Xi$   $\Xi^2$   $\Xi^3$   $\Sigma$   $\Delta V$   $\leftarrow$   $\rightarrow$   $\nabla$   $\nabla^2$   $\nabla^3$   $\nabla \Xi$   $\nabla \Xi^2$   $\nabla \Xi^3$   $\nabla \Sigma$   $\nabla \Delta V$   $\nabla \leftarrow$   $\nabla \rightarrow$   $\nabla \nabla$   $\nabla \nabla^2$   $\nabla \nabla^3$   $\nabla \nabla \Xi$   $\nabla \nabla \Xi^2$   $\nabla \nabla \Xi^3$   $\nabla \nabla \Sigma$   $\nabla \nabla \Delta V$   $\nabla \nabla \leftarrow$   $\nabla \nabla \rightarrow$

Template

[PS] Taxonomy & Inventories

Search sections...

Title

Abstract

Keywords

Introduction

Materials and methods

Data resources

No subsection

Taxon treatments

Taxon

- External Links
- Nomenclature
- Materials

## Data resources

Enter section content

## Taxon treatments

Materials [Download as CSV or XLSX](#)

Holotype :

a. catalogNumber: HBUMM08381-spec. 1; recordedByID: Chen, Tian; locationID: Zhangjiadi, Yunhe County, Lishui, Zhejiang Province; locality: around oaks in remote forest; verbatimElevation: c. 820 m a.s.l.; decimalLatitude: 27.974; decimalLongitude: 119.379; eventDate: 2019-08; basisOfRecord: PreservedSpecimen;

Paratype :

a. catalogNumber: HBUMM08381-spec. 2; recordedByID: Chen, Tian; locationID: Zhangjiadi, Yunhe County, Lishui, Zhejiang Province; locality: around oaks in remote forest; verbatimElevation: c. 820 m a.s.l.; decimalLatitude: 27.974; decimalLongitude: 119.379; eventDate: 2019-08; basisOfRecord: PreservedSpecimen;  
b. catalogNumber: HBUMM08370-spec. 1; recordedByID: Ye, Shi-Han; locationID: Mihougu, Fengyangshan, Longquan County, Lishui, Zhejiang Province; locality: Mihougu; verbatimElevation: 1100 m a.s.l.; decimalLatitude: 27.897; decimalLongitude: 119.159; eventDate: 2019-08-26; basisOfRecord: PreservedSpecimen;

# Data mobilization

Filling the gaps

# Special GBIF Data papers collections in BDJ



## Biota of Russia (2020-2021)

Papers published: **59**  
Total pages: **1204**  
Unique views: **130627**  
Total views: **202481**



## Biota of Northern Eurasia (2022-2023)

Papers published: **11**  
Total pages: **202**  
Unique views: **16991**  
Total views: **25287**



## Soil biodiversity (2023-2024)

Papers published: **3**  
Total pages: **86**  
Unique views: **4009**  
Total views: **5644**

# More than a decade together GBIF & Pensoft

Incentivizing data publishing

Improving data quality

Gathering more data

Saving time, minimizing errors

Filling the gaps

# Share your data through GBIF!

Publish its metadata  
as data papers!



Lyubomir Penev,  
[l.penev@pensoft.net](mailto:l.penev@pensoft.net)  
Pensoft Publishers  
[pensoft.net](http://pensoft.net)