

From Repository to a Data Paper

Melissa Liu, content manager



GBIF

Global Biodiversity
Information Facility

Biodiversity Information
Fund for Asia (BIFA)



If you publish a dataset to GBIF...

OCCURRENCE DATASET | REGISTERED JUNE 19, 2020

Bats of Bukit Barisan Selatan Landscape, Sumatra, Indonesia

Published by [Southeast Asian Bat Conservation Research Unit](#)

Joe Chun Chia Huang • Tigga Kingston • Elly Lestari Rustiati • Meyner Nusalawo • ✉ Gábor Csorba • Tamás Görföl

DATASET PROJECT METRICS ACTIVITY **DOWNLOAD**

1,776 OCCURRENCES

Bat diversity of Bukit Barisan
Indonesia

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

...ing robusta coffee farms in southwestern Sumatra,

Project ID: [BIFA04_24](#)

Metadata last modified: June 30, 2020

Hosted by: [GBIF Secretariat](#)

License: [CC BY-NC 4.0](#)

How to cite [DOI](#) 10.15468/emgg92

1,776
Occurrences

100%
With taxon match

100%
With coordinates

100%
With year

Why is publishing data papers important?



Citable with credits



Global/ regional research network



The lead of your research field



More cooperation opportunities



**Visible
Readable
Reusable**



More funding opportunities

Reasons to publish data papers (1/4)

If you are a **SCHOLAR/ PROFESSOR...**

- Increase citations of your research.
- Earn academic credits



Created by Gan Khoo Lay
from Noun Project

Reasons to publish data papers (2/4)

If you are a SCHOLAR/ PROFESSOR...

	Journal Impact factor 2019	Scopus CiteScore 2019
Biodiversity Data Journal	1.029	2.1
ZooKeys	1.143	2.0
BMC Ecology	2.381	3.5
PLoS One	2.776	5.2
Scientific Data	5.929	8.4

Reasons to publish data papers (3/4)

If you are a STUDENT...

Only the 'METHODS' & 'RESULTS', including the RAW DATA, are required for data paper publication!

On the other hand, for the advisors...

- Students would leave, but research shall last.
- Data would be well-organized and could be reused in the future.
- Easier to find your data



Created by Gan Khoon Lay
from Noun Project

Reasons to publish data papers (4/4)

If you are working in NGOs or other types of organizations...

If you leave the job...
Or if the project does not continue...

Getting credits from your work!
A solid reference for biodiversity conservation!



What is a data paper?

A **citable and peer-reviewed** description of your data
NOT a research investigation.

Main components of a data paper



Where you publish
the dataset

Considerations when picking up a data journal

- **Data Format effort**
- **Writing effort**
- **Requirements of data quantity and significance**
- **List of data journals - Index & publication fee**

Data format efforts

No need for reformatting your data!

No need for reformatting your data!

No need for reformatting your data!

Darwin core is acceptable

Data format efforts

Main components of a data paper



A red-bordered box contains a blue rounded rectangle. Inside this rectangle, on the left, is an icon of a server rack. To its right is the GBIF logo, which consists of a green leaf icon followed by the letters "GBIF" in a bold, black, sans-serif font. Below the blue rectangle, the text "Where you publish the dataset" is written in a bold, blue, sans-serif font.

Writing efforts

Main components of a data paper



**Where you publish
the dataset**

Writing efforts - General formats from GBIF

OCCURRENCE DATASET | REGISTERED JUNE 19, 2020

Bats of Bukit Barisan Selatan Landscape, Sumatra, Indonesia

Published by [Southeast Asian Bat Conservatic](#)

Joe Chun Chia Huang • Tigga Kingston • Elly Lestari Rustiati • Meyner N

DATASET PROJECT METRICS ACTIVITY **DOWNLOAD**

GBIF annotated archive *Recommended*

Source archive Darwin Core Archive

GBIF annotated metadata EML

Bat diversity of Bukit Barisan Selatan National Park, Sumatra, Indonesia
g robusta coffee farms in southwest

1,776
Occurrences

100%
With taxon match

10
Wi

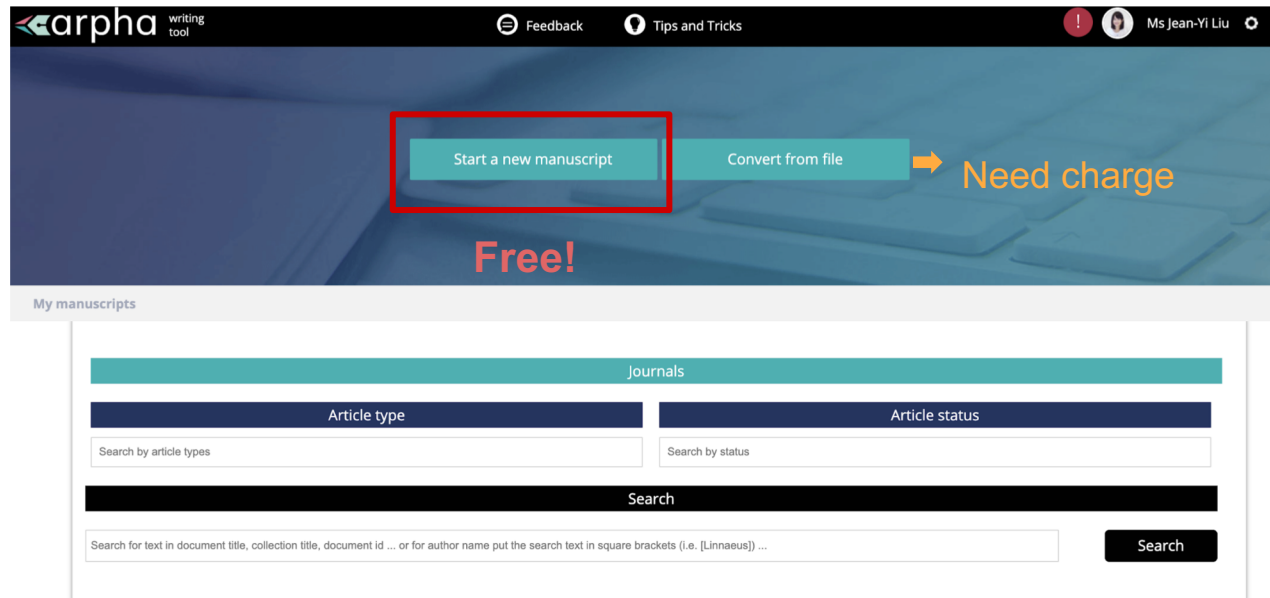
- Dataset Name
- Authors & Authors' information
- Citation
- Resource Citation
- Abstract & Keywords
- Project details
- Methods
- Method step description
- Dataset description

Metadata of your dataset

Writing efforts - After downloading the metadata from GBIF (as the authors)

Choose a data paper journal  Following their guideline and fit the contents in

Biodiversity data journal: using the [alpha Writing Tool](#)



The screenshot shows the alpha writing tool interface. At the top, there is a navigation bar with the logo 'alpha writing tool', a 'Feedback' button, a 'Tips and Tricks' button, and a user profile for 'Ms Jean-Yi Liu'. The main content area features a large blue banner with a keyboard background. A red box highlights the 'Start a new manuscript' button. To its right is a 'Convert from file' button with an arrow pointing to the text 'Need charge'. Below the banner, the word 'Free!' is displayed in red. Underneath, there is a 'My manuscripts' section with a 'Journals' header. Below this, there are two search filters: 'Article type' and 'Article status', each with a corresponding search input field. At the bottom, there is a 'Search' section with a large search input field and a 'Search' button. The search input field contains the text: 'Search for text in document title, collection title, document id ... or for author name put the search text in square brackets (i.e. [Linnaeus]) ...'.

Writing efforts

Biodiversity Data Journal: using the [arpha Writing Tool](#)

The screenshot shows the arpha writing tool interface. At the top, there is a navigation bar with the arpha logo, a 'writing tool' label, and links for 'Feedback', 'Tips and Tricks', and 'Tutorial'. A user profile for 'Ms Jean-Yi Liu' is visible in the top right. Below the navigation bar, a blue banner contains the text 'To start a manuscript select a journal and an article type'. The main content area displays a grid of journal selection options, each with a radio button and an icon. The 'Biodiversity Data Journal' option is highlighted with a red rectangular box. Below the journal selection, a red text annotation reads 'The only data paper journal here'. Underneath, there are three columns of article types: 'Research ideas', 'Early research outcomes', and 'Brief research outcomes'. Each article type is preceded by a radio button and a magnifying glass icon. The 'Data Paper (Biosciences)' option in the 'Early research outcomes' column is highlighted with a red rectangular box.

arpha writing tool

Feedback Tips and Tricks Tutorial

Ms Jean-Yi Liu

To start a manuscript select a journal and an article type

Biodiversity Data Journal

Research Ideas and Outcomes

One Ecosystem

BioDiscovery

BISS Biodiversity Information Science and Standards

Food Modelling Journal

ARPHA Conference Abstracts

Viticulture Data Journal

The only data paper journal here

Research ideas	Early research outcomes	Brief research outcomes
<input type="radio"/> Data Management Plan	<input type="radio"/> Applied Study	<input type="radio"/> Commentary
<input type="radio"/> Grant Proposal	<input type="radio"/> Case Study	<input type="radio"/> Conference Abstract
<input type="radio"/> PhD Project Plan	<input type="radio"/> Clinical Case Studies	<input type="radio"/> Correspondence
<input type="radio"/> PostDoc Project Plan	<input type="radio"/> Data analytics	<input type="radio"/> Ecosystem Inventory
<input type="radio"/> Research Idea	<input type="radio"/> Data Paper (Biosciences)	<input type="radio"/> Ecosystem Service Mapping
<input type="radio"/> Small Grant Proposal	<input type="radio"/> Data Paper (Generic)	<input type="radio"/> Ecosystem Service Models
<input type="radio"/> Software Management Plan	<input type="radio"/> Emerging Technique	<input type="radio"/> Monitoring Schema
	<input type="radio"/> Forum Paper	<input type="radio"/> Research Poster
	<input type="radio"/> FSKY (Food Safety Knowledge)	<input type="radio"/> Research Presentation

Writing effort - Biodiversity Data Journal: using the [arpha Writing Tool](#)



Import from EML metadata

Screenshot of the Biodiversity Data Journal 'Import a manuscript' interface. The 'Import from EML metadata' section is highlighted with a red box.

Feedback Tips and Tricks

Import a manuscript

Import from EML metadata

Supported EML versions: 2.1.1, 2.1.0 (e.g. generated from GBIF IPT, DataONE and LTER)

OR

Import an European Nucleotide Archive (ENA) Study ID or Project ID

Resource study or project id

Note: You can take the identifier from URL
<https://www.ebi.ac.uk/ena/data/view/P/RJDB2900&display=xml>

OR

Import from FSKX archive

Upload from file [Upload from url](#)

OR

Import an AGINFRA resource item

Writing efforts

Biodiversity Data Journal: using the [arpha Writing Tool](#)

Edit the contents
in detail by
clicking each
section

The screenshot shows the arpha writing tool interface. On the left, a sidebar menu is highlighted with a red border and a red arrow pointing to it. The menu items are: Authors, Contributors, Article metadata (with sub-items: Title, Abstract & Keywords, Classifications, Funder), Introduction (with sub-item: Subsection), General description, Project description, Sampling methods, Geographic coverage, Taxonomic coverage, Traits coverage, Temporal coverage, Collection data, Usage rights, Data resources, Additional information, and Acknowledgements. The main content area displays the manuscript draft for Biodiversity Data Journal. The title is "Records of Halieutichthys aculeatus complex (Ogcocephalidae, Lophiiformes)". The authors are Jean-Yi Liu[†] and Hsuan-Ching HO[§]. The journal name is "Biodiversity Data Journal : Data Paper (Biosciences)". A large blue diagonal banner across the bottom of the manuscript area reads "Ta-da! The basic manuscript is formed!".

arpha writing tool

Data Paper (Biosciences)

Biodiversity Data Journal : Data Paper (Biosciences) [Print](#)

Records of *Halieutichthys aculeatus* complex (Ogcocephalidae, Lophiiformes)

Jean-Yi Liu[†], Hsuan-Ching HO[§]

[†] Taiwan Biodiversity Information Facility (TaiBIF), Taipei, Taiwan
[§] National Museum of Marine Biology & Aquarium, Pingtung, Taiwan

Corresponding author: Jean-Yi Liu (melissaliu0520@gmail.com)

© Jean-Yi Liu, Hsuan-Ching HO

Citation:

Abstract

Background

The original data in the taxonomy review by Ho et al. (2010) were mainly organized for taxonomic purpose, of which were not the format for bioinformation and openaccessible online. In order to maximize the interoperability of these species occurrence data, the original data has been compiled, reformatted following the Darwin Core and published as a database to GBIF.

The re-description and open of the data are crucial for biodiversity study and conservation for the study region. For example, the Deepwater Horizon oil spill in 2010 was happened in northern Gulf of Mexico, which covers the major distribution range of one of the target species, *Halieutichthys intermedius*. The present dataset would provide opportunities to assess the potential impacts of the pollution to marine dweller systems with environmental fluctuation, such as water current.

Writing efforts

What else you should include

Biodiversity Data Journal: using arpha Writing Tool

- New information
- Introduction
- Purpose
- Usage rights: **choose a license** (CC-zero or other)
- Acknowledgements
- Author contributions
- References
- Figures / Tables

Click it and it will
tell you the
missing fields

Ready for
submission!

The screenshot shows the 'arpha writing tool' interface for a 'Data Paper (Biosciences)'. A list of sections is displayed on the right side of the page, including: Project description, Sampling methods, Geographic coverage, Taxonomic coverage, Traits coverage, Temporal coverage, Collection data, Usage rights, Data resources, Additional information, Acknowledgements, Author contributions, References, Supplementary files, Figures, Tables, and Endnotes. At the bottom of the list, there are two buttons: a teal 'Validate' button and a dark blue 'Submit for technical review' button. A blue arrow points from the text 'Click it and it will tell you the missing fields' to the 'Validate' button. A red arrow points from the text 'Ready for submission!' to the 'Submit for technical review' button.



Dataset for mosquitoes (Diptera: Culicidae) from Vaca Key, Monroe County, Florida USA

Lawrence J. Hribar †

† Florida Keys Mosquito Control District, Florida, United States of America

Corresponding author: Lawrence J. Hribar (lhribar@keysmosquito.org)

Academic editor: Gunnar Kvitte

Received: 03 Jun 2020 | Accepted: 22 Jun 2020 | Published: 27 Jul 2020

Citation: Hribar LJ (2020) Dataset for mosquitoes (Diptera: Culicidae) from Vaca Key, Monroe County, Florida USA. Biodiversity Data Journal 8: e55059. <https://doi.org/10.3897/BDJ.8.e55059>

Abstract

Background

The Florida Keys Mosquito Control District has used dry ice-baited light traps to monitor mosquito populations on Vaca Key since 1998. The first site sampled was monitored continuously for almost 20 years until all vegetation was removed.

New information

This paper describes a dataset compiled over almost 20 years of continuous trapping along Manor Lane on Vaca Key, Florida.

Keywords

Diptera, Culicidae, seasonal distribution, species composition, relative abundance

Introduction

Background

The Florida Keys Mosquito Control District conducted adult mosquito surveillance along Manor Lane in Vaca Key for almost twenty years. Surveillance was accomplished by use of dry ice-baited light traps. Traps were set and retrieved weekly except for interruptions due to unavoidable situations, such as storms and illness. Traps were provisioned with 2 pounds (ca. 1 kg) of dry ice, deployed in the late afternoon and retrieved the following morning. Traps were hung from the same tree limb each time they were set. Mosquitoes were returned to the laboratory, killed by freezing and identified to species. This paper reports data pertaining to 32 mosquito species.

General description

Purpose: These data were collected to document the species composition, seasonal distribution and relative abundance of mosquitoes on Vaca Key, Florida.

Sampling methods

Study extent: Monitoring of the mosquito fauna on Vaca Key began in the late summer of 1998 and continued until the winter of 2017. Collections were intended to be made weekly although, due to storms, illness and vacations, this was not always possible.

Sampling description: A battery-powered light trap (American Biophysics Company, Clarke, John Hock) was baited with approximately two pounds (ca. 1 kg) of dry ice and hung from the same tree limb once per week for over 19 years. The trap was deployed in the late afternoon and retrieved the following morning. The trap collection was taken to the laboratory, frozen, and all mosquitoes separated, identified, and counted. Data were recorded in spreadsheets. Voucher specimens of all taxa are retained in the synoptic mosquito reference collection maintained by the Florida Keys Mosquito Control District. Data may be found in Suppl. material 1.

Geographic coverage

Description: The trap was deployed on Vaca Key, Monroe County, Florida, USA. The coordinates of the trap location trap were: 24°42'78"N, 81°04'56"W.

Taxonomic coverage

Description: Thirty-two mosquito species were documented during the sampling period.

Writing effort

Traits coverage

The study site was described by Hribar (2002). Hribar et al. (2018) described population declines for the four most commonly collected species

Temporal coverage

Notes: Data collection began on 17 August 1998 and ended on 27 December 2017.

Usage rights

Use license: Creative Commons Public Domain Waiver (CC-Zero)

Data resources

Data package title: CSV

Number of data sets: 1

Data set name: Manor Lane BDJ

Data format: Data in spreadsheet, csv.

Description: Number of adult female mosquitoes collected per trap per night. NA indicates that the species was not reported and so no numbers are available.

Acknowledgements

Sincere appreciation and thanks are extended to all Florida Keys Mosquito Control District employees who set and retrieved traps for all the years that collections were made at the Manor Lane site. Heidi L. Murray and David J. DeMay counted, sorted, and identified many thousands of the mosquitoes.

References

- Hribar LJ (2002) Mosquito (Diptera: Culicidae) collections in the Florida Keys, Monroe County, Florida, USA. *Studia Dipterologica* 9: 679-6921.
- Hribar LJ, DeMay DJ, Murray HL (2018) Life and death of a trap site. *Wing Beats* 29 (1): 33-38.

Supplementary material

Suppl. material 1: Manor Lane BDJ [doi](#)

Authors: Hribar, L.J.

Data type: Count data in a csv spreadsheet.

Brief description: Approximately 19.5 years of near [BDJ_article_55059.pdf](#) from Vaca Key, Monroe County, Florida.

[Download file](#) (107.29 kb)

Already done in GBIF metadata!

Writing efforts

General formats of a data paper

Just read the journal guideline

Title & Authors & Affiliations

Abstract & keywords

Background / Introduction

Project details

Study area description/ Design description/
Geographic coverage/ Taxonomic coverage/
Temporal Coverage...

Methods

Data records / Data resources

Overview of your data files and their formats /
explain each data record associated with this work

Technical Validation / Quality control

Citations of Public Resource Databases

Acknowledgements

Author contributions

Figures & Tables

References

Complete metadata as more as you can!

Stories should be addressed in your data paper

- No claims regarding new scientific findings
- Potentials of **data reuse**
- Importance of the data to biodiversity science
- **Contributions** of each researcher and other participants, e.g. the metadata provider
- Figures: the **spatial information** of data, the experimental workflow
- Tables: taxonomic information of species, explanation fields of data

**The representative of your data is
more important than the data quantity**

New records of species

New taxa

Filling the temporal or spatial gaps

List of data journals (example) - Index & Publication fee

Animalia

Occurrence

Sampling event

Distribution

Biodiversity

Data journal	Publisher	Impact factor 2019	Scopus 2019	Charge
Check List Taxonomy	Biotaxa	-	0.9	EUR 150
Data in Brief	Elsevier	-	1.5	USD 500
Biodiversity Data Journal	Pensoft	1.029	2.1	EUR 100
ZooKeys	Pensoft	1.143	2.0	EUR 300
Ecological Research	Wiley	1.546	3.0	Free
BMC Ecology	Biomed Central	2.381	3.5	EUR 1,870
PLoS One	Plos	2.776	5.2	USD 1,300
Scientific Data	Nature Publishing Group	5.929	8.4	USD 1,675

Thank you

**A data paper guideline
will be released in this
October!**

Hopefully...

Feel free to contact us
taibif.brcas@gmail.com



Coral Reef, from New Taipei City, Taiwan, by Eloise Fan