



DATA
TERRA

Ecole Thématique DATA SDUE

**Guide de Survie dans la jungle des données
en Sciences de l'Univers et de l'Environnement (SDUE) :
Comment gérer les données pour les valoriser?**

Session «Valorisation – Data Paper »

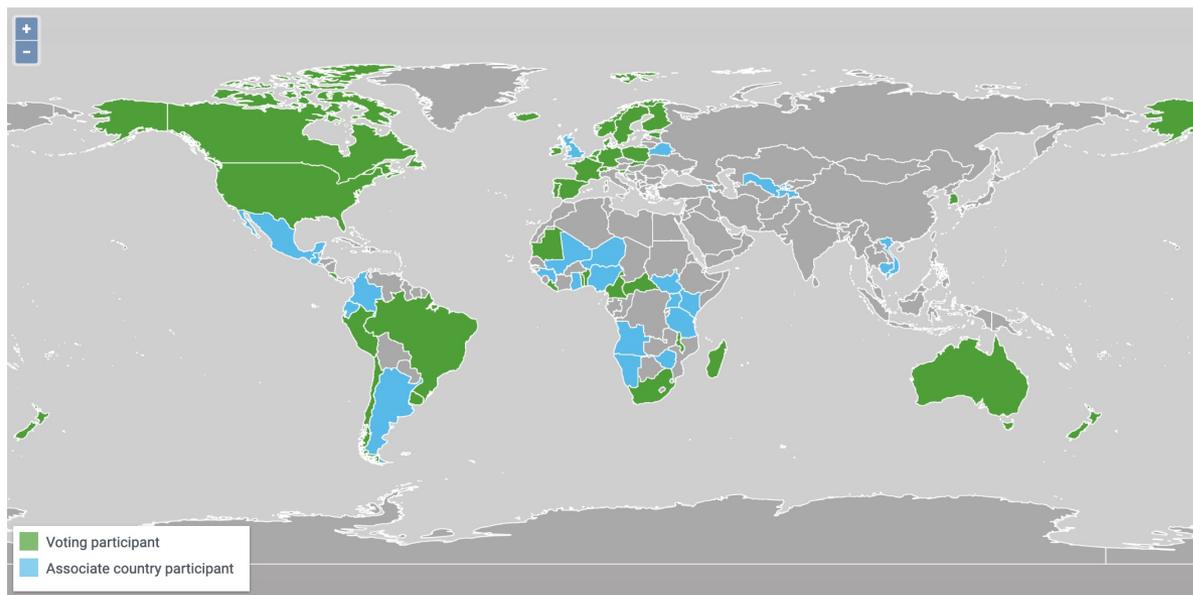
Sophie Pamerlon



Qu'est-ce que le GBIF?

- Programme intergouvernemental et infrastructure de données, créé en 2001, à l'initiative du comité scientifique de l'OCDE
- Promouvoir et faciliter l'accès libre et ouvert aux données sur la biodiversité

- Collaboration par le biais d'un protocole d'accord
- Réseau des Nœuds participants (63 pays, 42 organisations associées)
- Secrétariat basé à Copenhague, Danemark



39 VOTING PARTICIPANTS

24 ASSOCIATE COUNTRY PARTICIPANTS

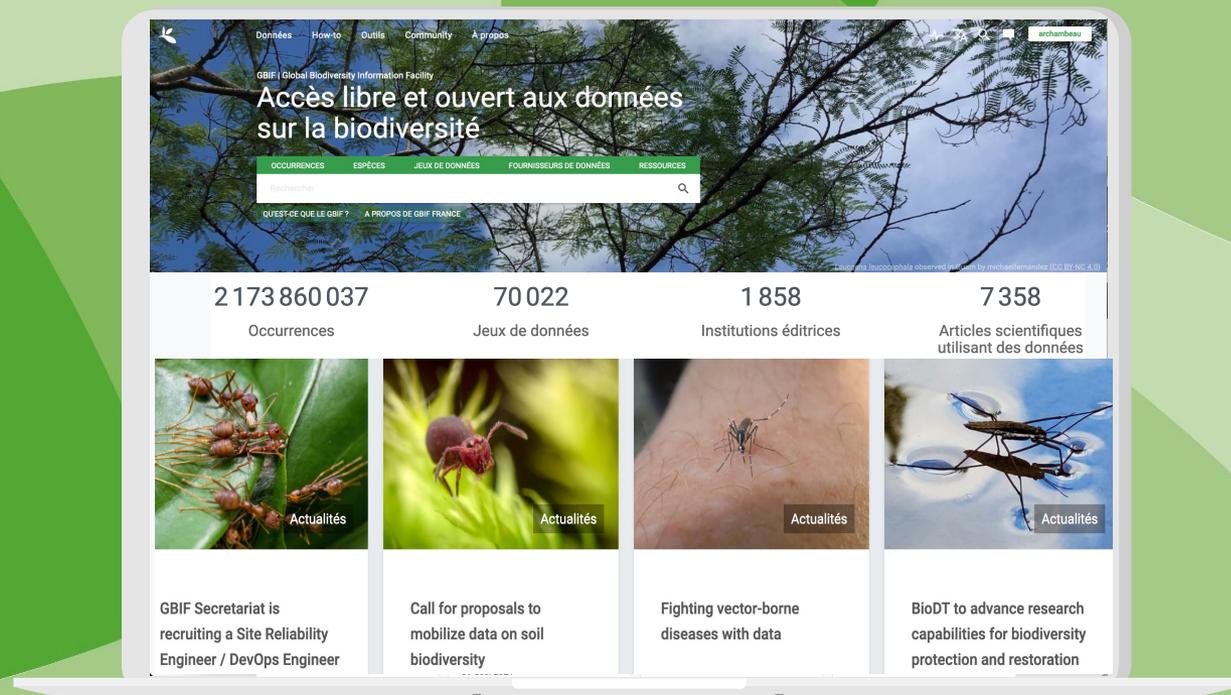
42 OTHER ASSOCIATE PARTICIPANTS

1 867 PUBLISHERS



Le GBIF en chiffres (octobre 2022)

- Plus de 2,2 milliards de données accessibles dont 99,6 millions avec objets multimédias associés
- 1917 institutions
- 7817 publications scientifiques
- 115 milliards de données téléchargées par mois en moyenne (2022)





Une fenêtre sur où, quand et par qui a été observé une espèce

• Observations



• Spécimens digitalisés

• Littérature



• Publication et indexation des données

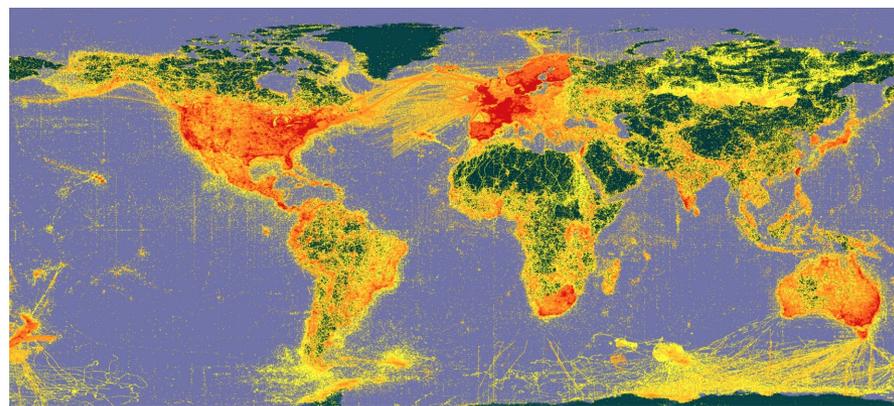


• Standard commun (DwC)

• Télédétection



• ADN environnemental (eDNA)

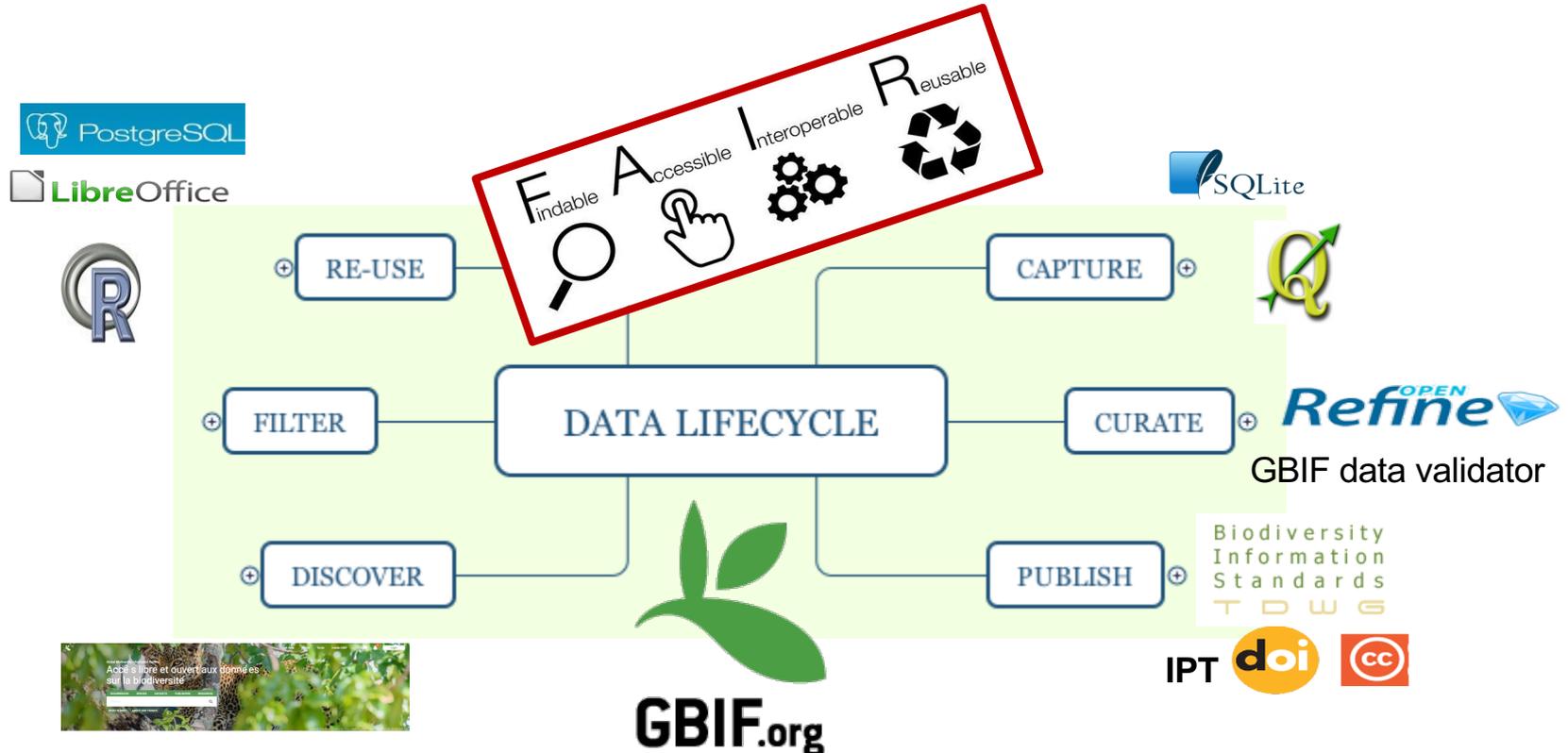


• Découverte et utilisation des données pour la recherche et l'appui aux politiques publiques



De la capture des données à leur réutilisation, un cycle complet des données

Open Source Data Management/Visualisation Tools





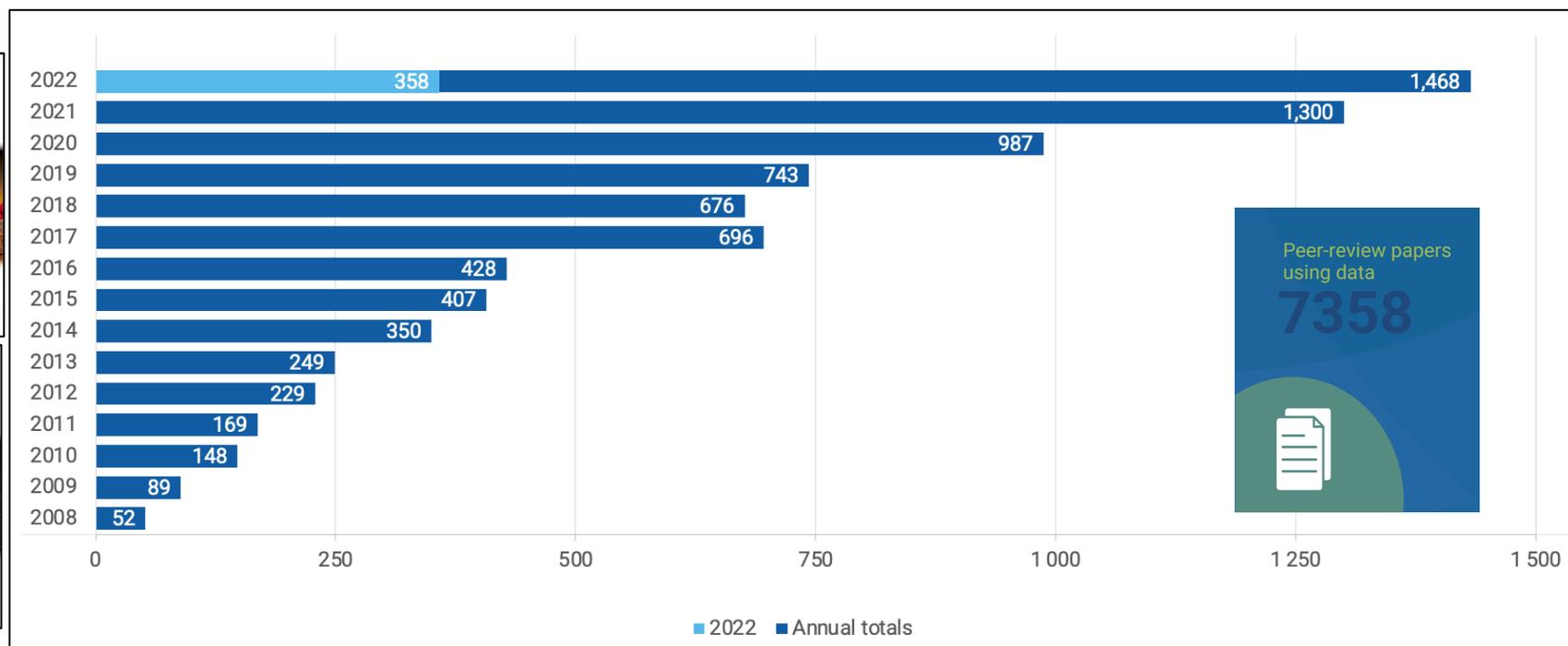
Processus mis en place

<p>Attribution</p> 	<p>Digital Object Identifiers assignés aux :</p> <ul style="list-style-type: none"> • Jeux de données publiés • Téléchargements des utilisateurs • Data papers
<p>Licences</p> 	<p>Les éditeurs choisissent une licence appropriée:</p> <ul style="list-style-type: none"> • CC-0 • CC-BY • CC-BY-NC
<p>Standards</p> 	<p>Données: DarwinCore, ABCD, DublinCore Metadonnées: Ecological Metadata Language (EML)</p>
<p>Accessibilité</p> 	<p>GBIF.org portail d'accès aux données</p> <ul style="list-style-type: none"> • Requêtes interopérables • Cartes interactives • Téléchargements gratuits • web services
<p>Publication</p>	<p>Integrated Publishing Toolkit (IPT), BioCASE, Spreadsheet templates</p>



Utilisation des données dans la recherche

Peer-reviewed publications using GBIF-mediated data. (march 2022)



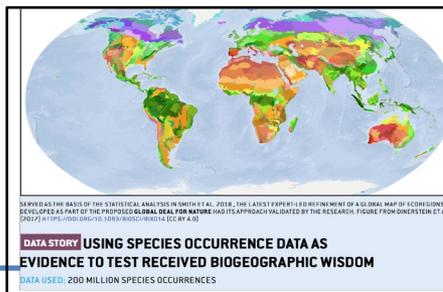
https://www.gbif.org/resource/search?contentType=literature&literatureType=journal&relevance=GBIF_USED&peerReview=true

GBIF Science Review :
<https://www.gbif.org/science-review>





Exemples d'utilisation pour le développement durable



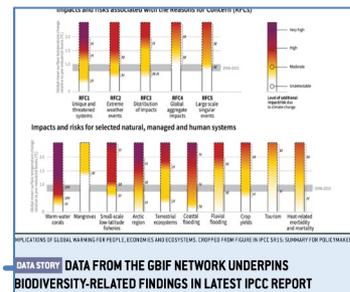
Conservation

- Zones protégées
- Espèces menacées
- Risques liés aux espèces envahissantes



Sécurité alimentaire

- Espèces sauvages apparentées
- Conservation *in situ*, *ex situ* de la diversité génétique
- Planification des pêches



Changement climatique

- Modélisation des impacts sur les aires de répartition des espèces
- Stratégies d'adaptation
- Avantages, risques d'atténuation



Santé humaine

- Risques de maladies basées sur la présence de vecteurs, d'hôtes, de réservoirs
- Plantes médicinales
- Dangers, par ex. morsure de serpent

<https://www.gbif.org/science-review>



Data Papers : qu'est-ce que c'est?

Une **publication scientifique** dont le but principal est de **décrire un ensemble de données** ou un groupe d'ensembles de données, plutôt que de rendre compte d'une enquête de recherche.

⇒ Reconnaissance et crédit pour les éditeurs de données via une publication scientifique

⇒ DOI : indexation et citation

les Data Papers sont indexés par Web of Knowledge (ISI), PubMedCentral, Scopus, Zoological Record, Google Scholar, CAB Abstracts, Directory of Open Access Journal (DOAJ), EBSCO.

⇒ Une incitation à publier des données de biodiversité

www.nature.com/scientificdata

SCIENTIFIC DATA

OPEN Data Descriptor: The French Muséum national d'histoire naturelle vascular plant herbarium collection dataset

Received: 07 June 2016
Accepted: 04 January 2017
Published: 14 February 2017

Gwenaël Le Bras¹, Marc Pignal^{1,2}, Marc L. Jeanson¹, Serge Muller², Cécile Aupic¹, Benoît Carré³, Grégoire Flament¹, Myriam Gaudeul², Claudia Gonçalves¹, Vanessa R. Invernón¹, Florian Jabbour², Elodie Lerat¹, Porter P. Lowry^{2,4}, Bérandère Offroy¹, Eva Pérez Pimparé¹, Odile Poncy², Germinal Rouhan² & Thomas Haevermans²

We provide a quantitative description of the French national herbarium vascular plants collection dataset. Held at the *Muséum national d'histoire naturelle*, Paris, it currently comprises records for 5,400,000 specimens, representing 90% of the estimated total of specimens. Ninety nine entries are linked to one or more images and 16% have field-collecting information available. This major

<https://www.nature.com/articles/sdata201716#citeas>



Data Paper : avantages



DOI : indexation et citation
- indexés par Web of Knowledge (ISI), PubMedCentral, Scopus, Zoological Record, Google Scholar, CAB Abstracts, DOAJ, EBSCO.



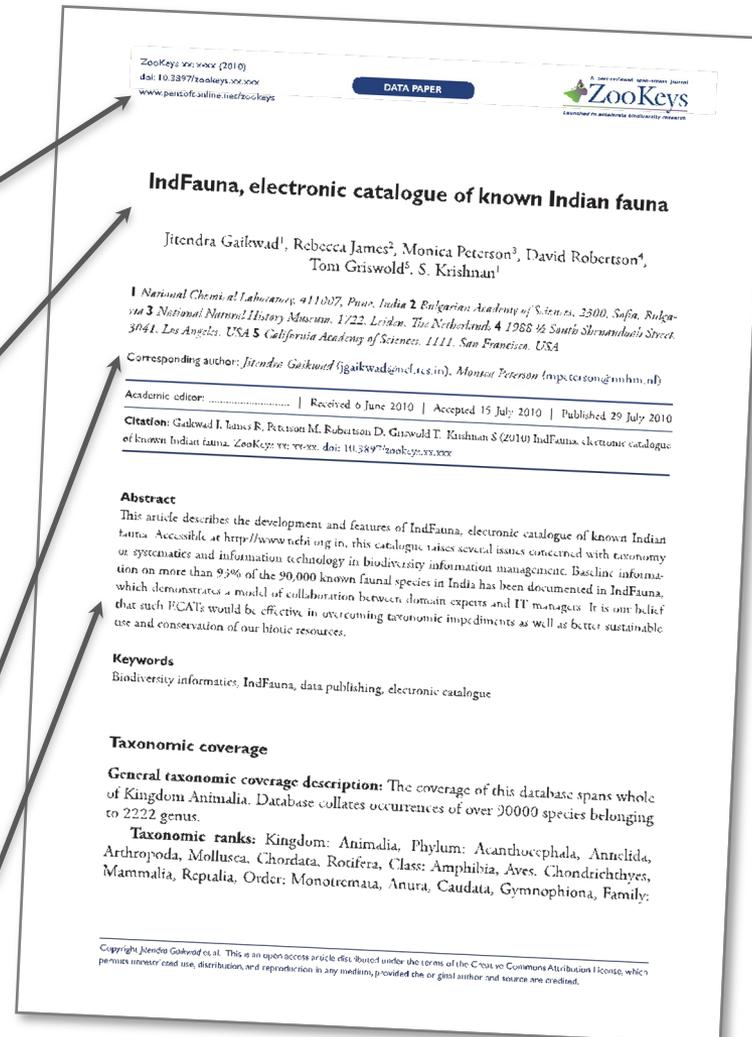
Promouvoir et faire connaître les données



Reconnaissance des éditeurs de données via une publication scientifique



Décrit les données sous forme structurée et lisible par un humain





Sur le site GBIF.org

Get data

Share

Tools

Inside GBIF

Journals that publish biodiversity-related data papers

Please notify us at communication@gbif.org with corrections or updates.

Journal	Publisher	Open Access	APC estimate	Impact factor
Arxius de Miscel·lània Zoològica	Nat Hist Museum of Barcelona	Yes	€0	-
Biodiversity Data Journal	Pensoft	Yes	€300	-
BioInvasions Records	REABIC Journals	Yes	€600	-
BioRisk	Pensoft	Yes	€300	-
Biota Colombiana	Humboldt Institute, Colombia	Yes	€0	-
BMC Ecology	Biomed Central	Yes	€1,745	2.724
BMC Plant Biology	Biomed Central	Yes	€1,745	3.631
Botanical Studies	SpringerOpen	Yes	€600	1.159
Check List	Biotaxa	Yes	€27	-



ZooKeys 489: 15–24 (2015)
doi: 10.3897/zookeys.489.9292
<http://zookeys.pensoft.net>

DATA PAPER

The Jean Gutierrez spider mite

Alain Migeon¹

¹ INRA, UMR 1062 CBGP, F-34988 Montpellier-sur-Lez, France

Corresponding author: *Alain Migeon* (alain.migeon@supagro.inra.fr)

Academic editor: *V. Pesic* | Received 28 January 2015 | Accepted 10 May 2015

<http://zoobank.org/0719C382-988D-4DCC-8B24-00E>

Citation: Migeon A (2015) The Jean Gutierrez spider mite collection. ZooKeys 489: 15–24

Abstract

The family Tetranychidae (spider mites) currently comprises 1,275 species and includes some of the most important agricultural pest families among the Acari with approximately 100 species which are considered major pests. The dataset presented in this document includes the spider mites composing the Jean Gutierrez Collection hosted at the CBGP (Montpellier) which was created from 1963 to 1999 during his career at the Institut de Recherche pour le Développement. It consists of 5,262 specimens corresponding to 1,564 occurrences (combining

Keywords

Acari, Tetranychidae, World, Madagascar, Western Indian Ocean, New Caledonia

Data published through GBIF

<http://www.gbif.org/dataset/ac60a288-fcc9-43fe-a7d4-e732b748a981>

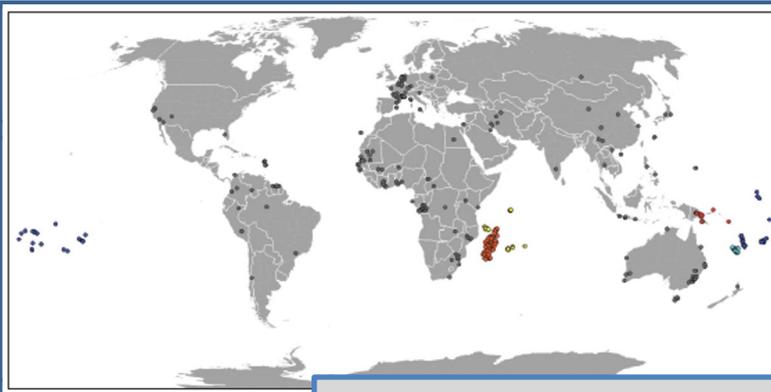


Figure 1. World map representing all the areas of particular interest. Grey spots represent all the species recorded in the dataset. Coloured spots represent the areas of particular interest: ● Madagascar, ● Western Indian Ocean (incl. Madagascar), ● Papuaia, ● New Caledonia, ● South Pacific). Grey spots represent all the species recorded in the dataset.

L'ajout d'analyses statistiques ou de représentations graphiques est possible...

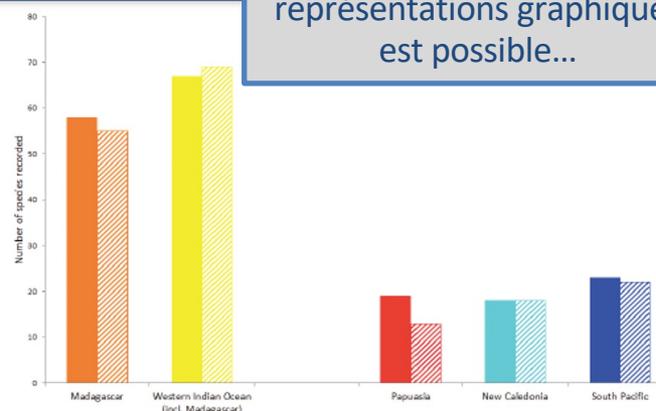


Figure 2. Number of species recorded in Jean Gutierrez collection dataset (solid bar) and in the literature (dashed bar) compiled in Spider Mites Web (<http://www1.montpellier.inra.fr/CBGP/spmweb/>) for the areas of particular interest. Colour scheme same as in Figure 1.



[Home](#)

[Manage Resources](#)

[Administration](#)

[About](#)

Resource Title [Abundances and biological traits of the juveniles salmon sampled in the survey of Salmon abundance Indices in the Scorff river \(France\)](#)

Basic Metadata

Please enter all the mandatory properties on the Basic Metadata page, and then continue entering metadata in the other pages that are applicable to your resource. The more metadata you provide, the greater the chance that your resource will be found, reused by other researchers, and cited.

Title*

Publishing Organisation*

Type*

Metadata Language*

Update Frequency*

Subtype

Data Language*

Data Licence*



This work is licensed under a [Creative Commons Attribution \(CC-BY\)](#)

Section

[Basic Metadata](#)

[Geographic Coverage](#)

[Taxonomic Coverage](#)

[Temporal Coverage](#)

[Keywords](#)

[Associated Parties](#)

[Project Data](#)

[Sampling Methods](#)

[Citations](#)

[Collection Data](#)

[External links](#)

[Additional Metadata](#)

L'Integrated Publishing Toolkit (IPT) facilite le remplissage des métadonnées et la production automatisée d'un manuscrit de Data Paper



> Accès libre et gratuit aux données de la biodiversité

Home

About



Summary

Data Records

Downloads

Versions

How to cite

Rights

GBIF Registration

Keywords

Contacts

Geographic Coverage

Taxonomic Coverage

Temporal Coverage

Project Data

Sampling Methods

Abundances and biological traits of the juveniles salmon sampled in the survey of Salmon abundance Indices in the Scorff river (France)

Latest version published by Institut national de recherche pour l'agriculture, l'alimentation et la forêt (INRAE) on Feb 28, 2020

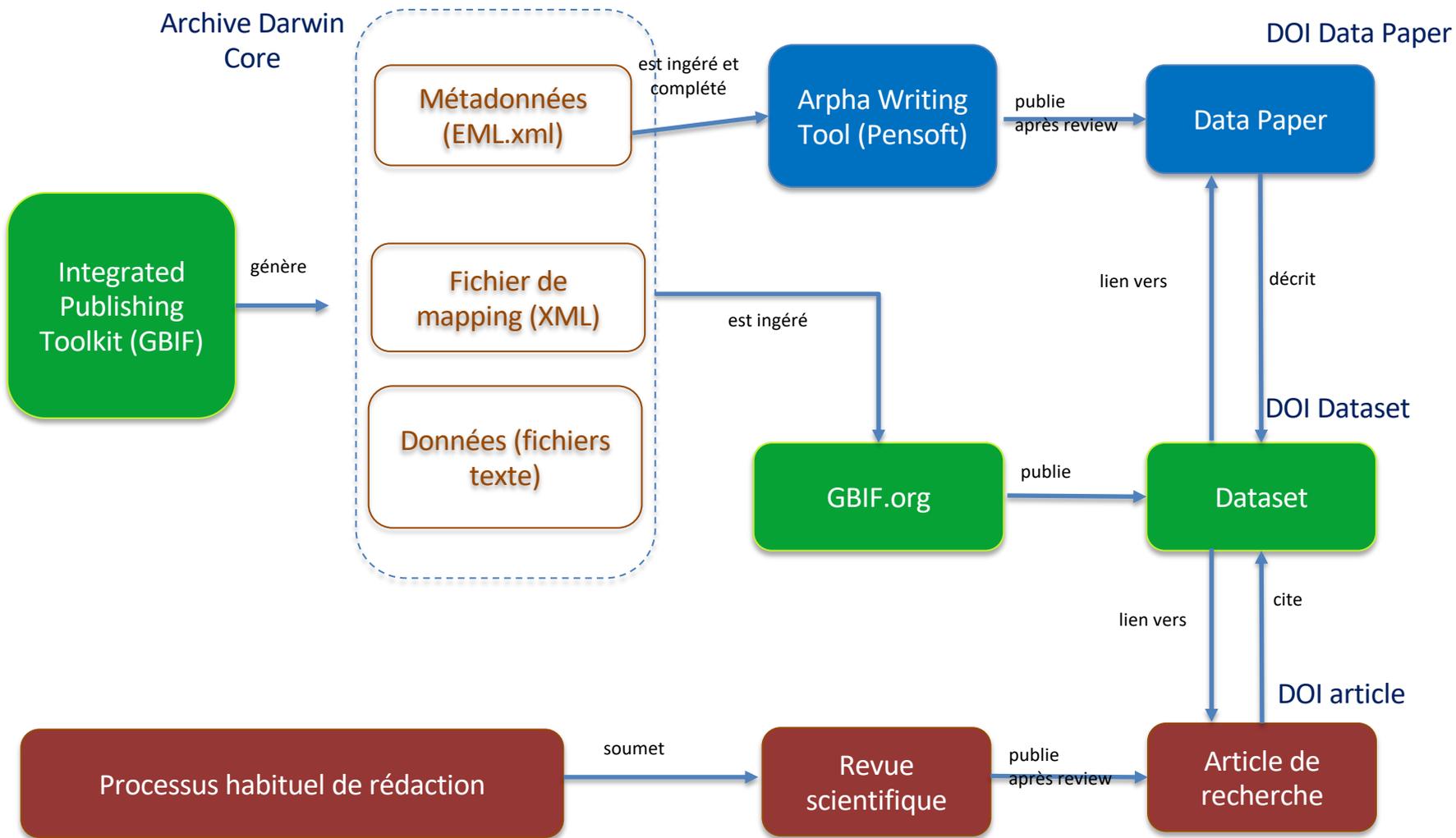
A survey started in 1993 is conducted every year in early autumn (October) to quantify the abundance of juvenile Atlantic salmon in the Scorff river. The electric fishing protocol of Prévost and Baglinière (1995) is used. It is carried out in October of the year (0+ parr), but older fish (juvenile salmon $\geq 1+$) are also included in the dataset. Sampling is restricted to areas with shallow running water and soft substrate, i.e. the preferred habitat of young of the year salmon. The following biological traits measured on the fish sampled: sex, maturity status, length, weight, scale samples taken from the fish which size does not allow to be determined. The data are stored in the database of the ERO. They are used to provide scientific advice to improve the management of this heritage.

[Home](#) [GBIF](#) [DwC-A](#) [EML](#) [RTF](#) [Data](#)

```
<eml:eml xsi:schemaLocation="eml://ecoinformatics.org/eml-2.1.1 http://rs.gbif.org/schema/eml-gbif-profile/1.1/eml.xsd" id="b236-092e1dae7042/v1.13" system="http://gbif.org" scope="system" xml:lang="eng">
  <!--dataset-->
  <alternateIdentifier>89064e3a-aa3c-495d-b236-092e1dae7042</alternateIdentifier>
  <alternateIdentifier>http://ipt.gbif.fr/resource/?r=inra_rennes_scorff</alternateIdentifier>
  <!--title xml:lang="eng"-->
  <title>Abundances and biological traits of the juveniles salmon sampled in the survey of Salmon abundance Indices in the Scorff river (France)</title>
  <!--creator-->
  <!--individualName-->
  <givenName>Nicolas</givenName>
  <surName>Jeannot</surName>
  <!--individualName-->
  <!--organizationName-->
  <organizationName>INRAE U3E, Pôle OFB-INRAE-Agrocampus Ouest-UPPA pour la gestion des migrateurs amphihalins dans le Scorff</organizationName>
  <positionName>Researcher</positionName>
  <!--address-->
  <deliveryPoint>Moulin des Princes</deliveryPoint>
  <city>Pont-Scorff</city>
  <postalCode>56620</postalCode>
  <country>FR</country>
  </address>
  <phone>+33297325429</phone>
  <electronicMailAddress>Nicolas.jeannot@inrae.fr</electronicMailAddress>
  <!--creator-->
  <!--individualName-->
  <givenName>Didier</givenName>
  <surName>Azam</surName>
  <!--individualName-->
  <!--organizationName-->
  <organizationName>INRAE U3E, Pôle OFB-INRAE-Agrocampus Ouest-UPPA pour la gestion des migrateurs amphihalins dans le Scorff</organizationName>
  </organizationName>
</eml>
```



En résumé





Pour en savoir plus

- <http://www.gbif.fr/page/ressources/data-papers>
- <https://www.gbif.org/data-papers>
- https://pensoft.net/J_FILES/Pensoft_Data_Publishing_Policies_and_Guidelines.pdf
- <https://coop-ist.cirad.fr/gerer-des-donnees/publier-un-data-paper/1-qu-est-ce-qu-un-data-paper>



Contacts

Equipe GBIF France
gbif@gbif.fr

Site web

<http://www.gbif.fr>

Twitter

@gbiffrance



L'outil Arpha Writing Tool (Pensoft)

Un outil - parmi d'autres - d'aide à la rédaction de Data papers



Pensoft : maison d'édition de littérature scientifique, fondée en 1992 en Bulgarie

A développé les Data Papers sur les données de biodiversité en partenariat avec le GBIF (Global Biodiversity Information Facility) en 2010.

Growing family of Pensoft open access journals

<https://pensoft.net>



 **Check List**
the journal of biodiversity data

 **RIO**
Research Ideas and Outcomes

 **IMBMG**
Metabarcoding & Metagenomics

 **BioDISCOVERY**
powered by the Global Biodiversity Information Facility

 **One Ecosystem**
Ecology and Sustainability Data Journal

 **BISS** Biodiversity
Information
Science and
Standards



Outil de rédaction ARPHA : facilite la mise en page, la soumission, le processus de relecture, la publication, l'hébergement et l'archivage d'articles scientifiques.

ARPHA

an end-to-end journal publishing platform

Authoring, Reviewing
Publishing, Hosting, Archiving



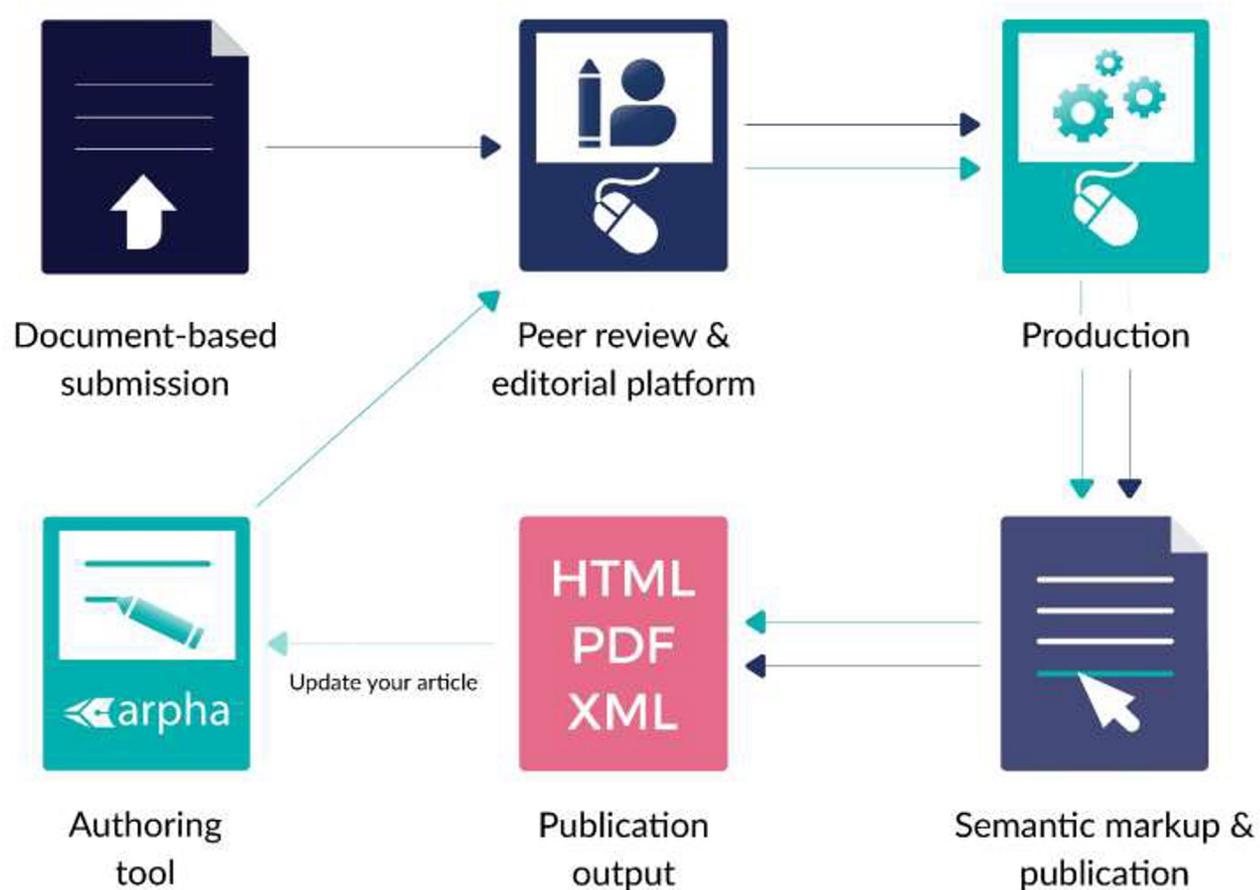
Interface utilisateur

The screenshot displays the GBIF user interface for editing a 'Data Paper'. The interface is divided into several sections:

- Top Navigation Bar:** Includes the 'arpha' logo, a 'writing tool' icon, and various utility icons such as 'Email co-authors and contributors', 'helpdesk', 'Tips and tricks', 'tutorial', 'Revision history', and a user profile for 'Ms Sophie Pamerlon'.
- Left Sidebar:** A navigation menu with categories like 'Authors', 'Contributors', 'Article metadata' (with sub-items: Title, Abstract & Keywords, Classifications, Funder), 'Introduction' (with sub-item: First section), 'General description', 'Project description', 'Sampling methods', 'Geographic coverage', 'Taxonomic coverage', 'Traits coverage', 'Temporal coverage', 'Collection data', 'Usage rights', 'Data resources', 'Additional information', and 'Acknowledgements'.
- Main Content Area:**
 - Header: 'Biodiversity Data Journal : Data Paper (Biosciences)' with a 'Print' icon.
 - Title: 'Data Paper' in a large, bold font, enclosed in a text box with a pencil icon for editing and a speech bubble icon for comments.
 - Author: 'Sophie Pamerlon' with a superscripted asterisk, followed by '† GBIF France, Paris, France'.
 - Corresponding author section: A box containing 'Corresponding author:', '© 2018 Sophie Pamerlon', and 'Citation: () . <https://doi.org/>'.
 - Open Access: An 'OPEN ACCESS' logo.
 - Section Headers: 'Abstract', 'Background', 'New information', and 'Introduction' are listed. Under 'Background', there is a link to 'Marchand 2017'. Under 'New information', the text 'test' is visible.
 - Section Headers: 'Introduction' and 'First section' are also listed.
- Right Sidebar:** Contains a 'Resolved' status indicator and a 'Filter' button.



Deux voies de publication possibles : document ou XML

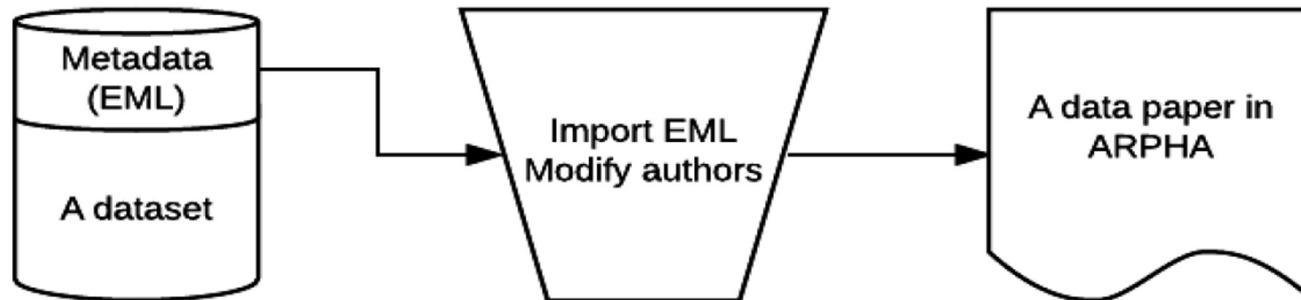




Outil et fonctionnalités interoperables.

Exemple : métadonnées GBIF (IPT) → Data Paper (ARPHA)

Generate and import an **entire manuscript**



DataONE



GBIF

ILTER

Name	Email	Include	Is Submitting
Viktor Senderov	datascience@pensoft.net	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
NoName PleaseChange	NoName@PleaseChange.This	<input type="checkbox"/>	<input type="checkbox"/>
Oscar Schofield	oscar@marine.rutgers.edu	<input type="checkbox"/>	<input type="checkbox"/>

Submit

Biodiversity Information Facility

Photosynthetic pigments of water column samples analyzed using High Performance Liquid Chromatography (HPLC), sampled during Palmer LTER field season at Palmer Station Antarctica, 1991 - 2009.

Oscar Schofield

Rutgers University

Corresponding author: Oscar Schofield, oscar@marine.rutgers.edu

© 2010 Oscar Schofield

October 2, 2010



ARPHA



Exemple de Data Paper publié via Arpha Writing Tool :



Data Paper

Biodiversity Data Journal 5: e15125 (09 Nov 2017)
<https://doi.org/10.3897/BDJ.5.e15125>



Abundance indices and biological traits of juvenile salmon (*Salmo salar*) sampled in three rivers on the Atlantic and Channel coasts (France)

▼ Frédéric Marchand, Laurent Beaulaton, Etienne Prévost, Richard Delanoë, Jean-Pierre Destouches, François Gueraud, Yoann Guilloux, Nicolas Jeannot, Emmanuel Huchet, Frédéric Lange, Jacques Rives, Julien Tremblay, Nadine Herrard, Didier Azam

Abstract ▲

Background

Atlantic Salmon (*Salmo salar*) is an anadromous migratory species adapted to cool temperatures. It is protected by the Bern convention and by the European Habitats Directive. It has been listed as vulnerable by the French IUCN Red List. Salmon decline is the result of combined and cumulated, mainly anthropic, causes: climate change, increasingly high number of impoundments, degradation of water quality and habitat and over-exploitation by fisheries. Monitoring of this species has been carried out on three rivers in France (Southern part of the distribution area) to produce data and knowledge (growth, precocious maturity, survival) for stock management.

For 24 years, a specific and standardised electric fishing protocol has been used to target young-of-the-year (0+ parr) Atlantic salmon. Sampling was restricted to areas with shallow running water that flows over a coarse bottom substrate, i.e. the preferred habitat of young salmon. This monitoring and inventory of growing areas thus allows assessment of juvenile recruitment and provides baseline data required to calculate total allowable catches (TACs).

Contents	Article info	Citation	Metrics	Comment	Related
Figs	Taxa	Data	Refs	Cited	

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



Un des jeux de données correspondant sur www.gbif.org :



OCCURRENCE DATASET | REGISTERED 6 MARCH 2017

Abundances indices and biological traits of juvenile salmon sampled in the Scorff river (France)

Published by [Institut National de la Recherche Agronomique \(INRA\)](#)

Nicolas Jeannot • Didier Azam • Yoann Guilloux • Etienne Prévost • Frédéric Marchand

[DATASET](#) [PROJECT](#) [STATS](#) [ACTIVITY](#) [DOWNLOAD](#) [DATASET HOMEPAGE](#)

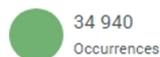
34 940 OCCURRENCES 1 CITATION

A survey started in 1993 is conducted every year in early autumn (late September to early October) to quantify the abundance of juvenile Atlantic salmon in the Scorff in Brittany. The electric fishing protocol of Prévost and Baglinière (1995) is used. It targets Atlantic salmon young of the year (0+ parr), but older fish (juvenile salmon $\geq 1+$) are also caught and included in this dataset. Sampling is restricted to areas with shallow running water flowing on coarse bottom substrate, i.e. the prefe... [more](#)

Last Modified: 28 September 2017

License: CC BY 4.0

[How to cite](#) [DOI](#) 10.15468/mz4lyw



34 940 GEOREFERENCED RECORDS

