



DATAPAPERS – TABLE RONDE

Journée science ouverte du CNRS

16 novembre 2020

Participant(e)s :

Cécile Callou – *MNHN*

Louis Manière – *CNRS*

Thierry Forveille – *UGA , Astronomy & Astrophysics*

Animation : Laurence El Khouri - CNRS

DONNEES OUVERTES DE LA RECHERCHE

- Aussi ouvertes que possible aussi fermées que nécessaire
- Ouvertes par défaut
 - Loi Pour une République numérique, articles 11, 30, 38 (7 oct 2016)
 - Révision de la directive PSI, considérant 4 et article 10 où elle introduit les principes FAIR dans la loi (20 juin 2019)
- Objectifs : Données réutilisables ET réutilisées

DONNEES OUVERTES DE LA RECHERCHE

Réutilisées → **F A I R**

→ Documentées, enrichies, explicitées, licences → **R**éutilisables

→ Métadonnées choisies, indexation, identifiants (PID) → **F**aciles à trouver

→ Ouvertes, signalées (bonnes métadonnées), accord de tous → **A**ccessibles

→ Bons formats, standards, licences et protocoles d'accès → **I**nteropérables

Ecosystème :

Entrepôts, PID, DMP

DATAPAPER : UN MOYEN DE FACILITER LA RÉUTILISATION

Article qui met en valeur certains jeux données

Enjeux

Reproductibilité

Gratifications

Freins

DATAPAPER : UN MOYEN DE FACILITER LA RÉUTILISATION

Article qui met en valeur certains jeux données

Enjeux

- Mise en valeur d'un jeu de données
- Rattachement ou pas à un résultat
- « Revue par les pairs »
- Transparence des méthodes et des résultats
- Soutenir les bonnes pratiques de gestion

Reproductibilité

- Qualification des données
- Description des données
- Mise à disposition
- Facilitation de la compréhension

Gratifications

- Citation des participants
- Intégrer le circuit de la publication scientifique
- Citabilité et valorisation plus facile
- Mise en valeur des métiers d'accompagnement de la recherche

Freins

- Gestion des données (data management) difficile et fastidieuse
- Difficulté à trouver le bon journal / le bon entrepôt
- Intérêt de garder ses données pour soi
- Coût des articles

Datapapers et publication des données en astronomie

Thierry Forveille, Observatoire de Grenoble, Université Grenoble-Alpes
Editeur en chef de Astronomy & Astrophysics

Astronomy & Astrophysics

- 1 des 4 revues qui publient ~90% de la recherche astronomique
- A&A est la revue européenne (continentale) + Amérique du Sud
- ~2000 articles/an
- Bras de levier significatif sur les auteurs

Publication des données en astronomie

- Très longue tradition de publication de (certaines) données: Hipparque (129 BC), Bayer (1603),...
- Beaucoup de données obtenues à travers des TGIR (ou missions spatiales) qui imposent la publicité des données


Beaucoup de données publiées, mais (bien sûr) aussi certains auteurs qui résistent...

Comment:

- Tout article [A&A](#) peut avoir des données associées
 - Une section du journal (Catalogs & Data) pour les « Datapapers » purs
-

- Directement dans l'article (souvent en appendice) pour les petits volumes
- Via un lien vers un entrepôt
 - Centre de Données Stellaire (UMR CNRS/U. Strasbourg)
 - Archives de haut niveau du TGIR producteur
 - Archives spécialisées (e.g. données moléculaires pour l'astronomie, spectres de supernovae,...)
 - Github pour des codes; occasionnellement Zenodo, etc...

Quoi?

- Uniquement des données de haut niveau (données brutes archivées par les TGIR/IR)
 - Choix important: la mise à disposition a un coût, à comparer à l'utilité potentielle
- 

Datapaper en archéologie

ERC Desert Networks

DÉMARCHE D'UN PREMIER DATAPAPER SUR DES DONNÉES ET
RÉSULTATS DE MODÉLISATION D'ITINÉRAIRES ANTIQUES DANS LE
DÉSERT ORIENTAL ÉGYPTIEN

Démarche de publication

Nouvelle méthodologie sur la reconstitution des réseaux de transport antiques

Explication de la démarche et mise à disposition des outils pour la reproduire

Article principal
(Journal of Computer Applications in Archaeology)
Démarche de modélisation des itinéraires entre les sites archéologiques et résultats

Dataper
(Journal of open archaeology data)
Présente le jeu de données, sa construction et sa qualité

Dépôt
(Zenodo)
Données d'entrées, résultats, métadonnées et outils de traitement

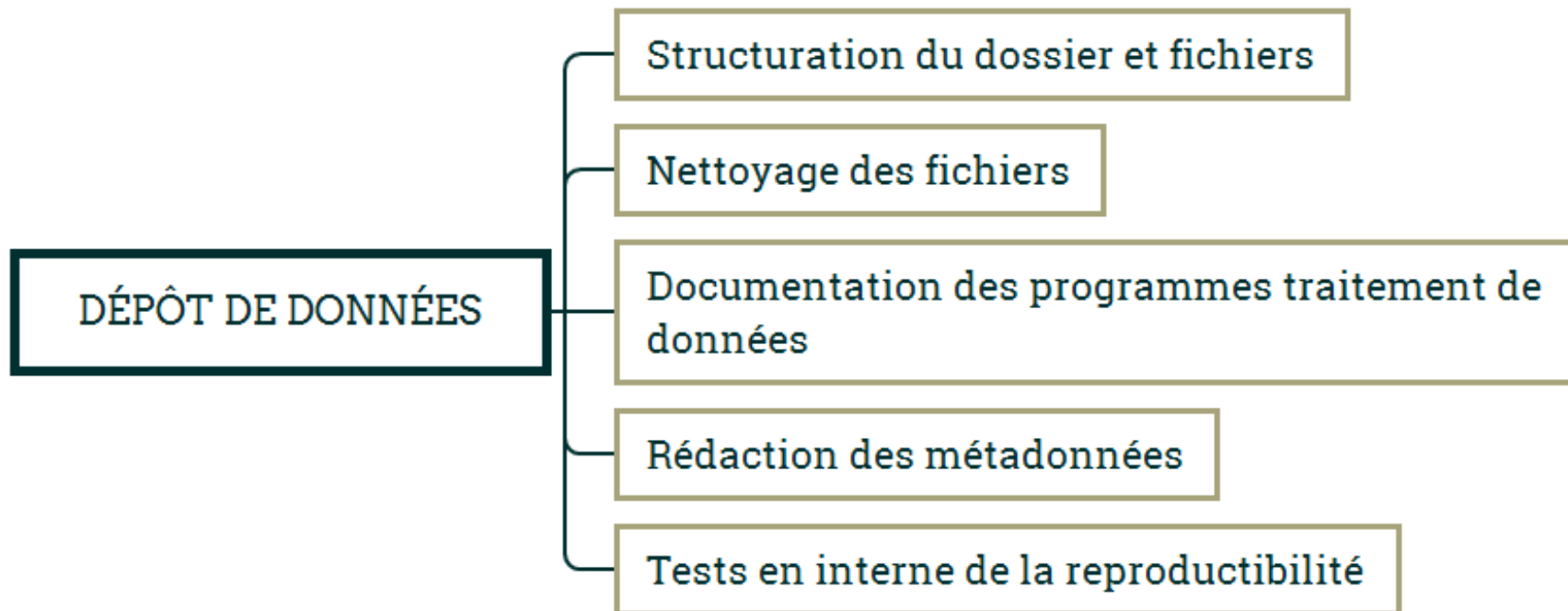
Motivations de la rédaction du datapaper

Datapaper en complément de la publication principale sur la démarche de modélisation

- Objectifs de recherche :
 - Diffusion des données auprès des chercheurs (archéologie, histoire, géographie, etc.) travaillant sur notre région d'étude
 - Permettre la reproductibilité mais surtout la répliquabilité sur d'autres terrains d'étude en fournissant les outils de la méthode
- Avoir une documentation du jeu de données publiée avec comité de lecture
- Détailler des aspects techniques facilitant la réutilisation des données et le transfert méthodologique que l'on peut facilement exclure d'une publication « traditionnelle » par manque de place

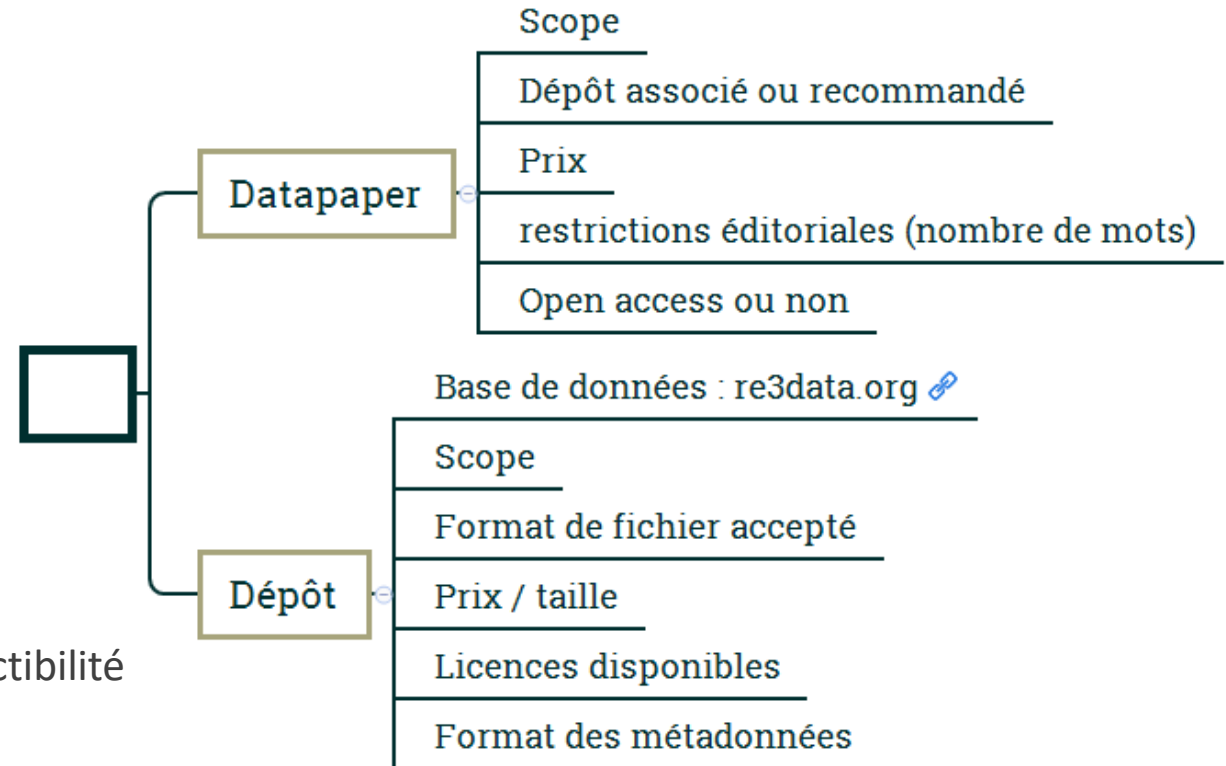
Une rédaction facile pour un dépôt bien préparé

- Datapaper = documentation générale du dépôt => rapide à rédiger



Contraintes et interrogations

- Etude de marché des éditeurs et dépôts
 - Choisir le journal avant le dépôt



- Peer review : test du jeu de données, traitement ?
 - Le datapaper n'est pas une certification de la reproductibilité

- Impact encore incertain sur la communauté, besoin de compétences techniques dans les laboratoires des équipes intéressées

Synthèse

Avantages

- Datapaper => laisse une trace publiée et révisée des données partagées
- Cadre pour bien organiser son dépôt de données

Contraintes

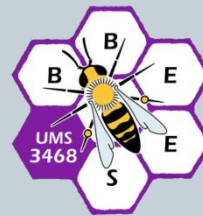
- Moyens humains pour préparer le dépôt et financier selon le journal et le dépôt choisis

DATA PAPER

Une incitation à la publication des données sur la Biodiversité



Cécile CALLOU (UMS BBEES – CNRS InEE/MNHN)
avec la Collaboration d’A.S. ARCHAMBEAU & S. PAMERLON
(GBIF France/UMS Patrinat)

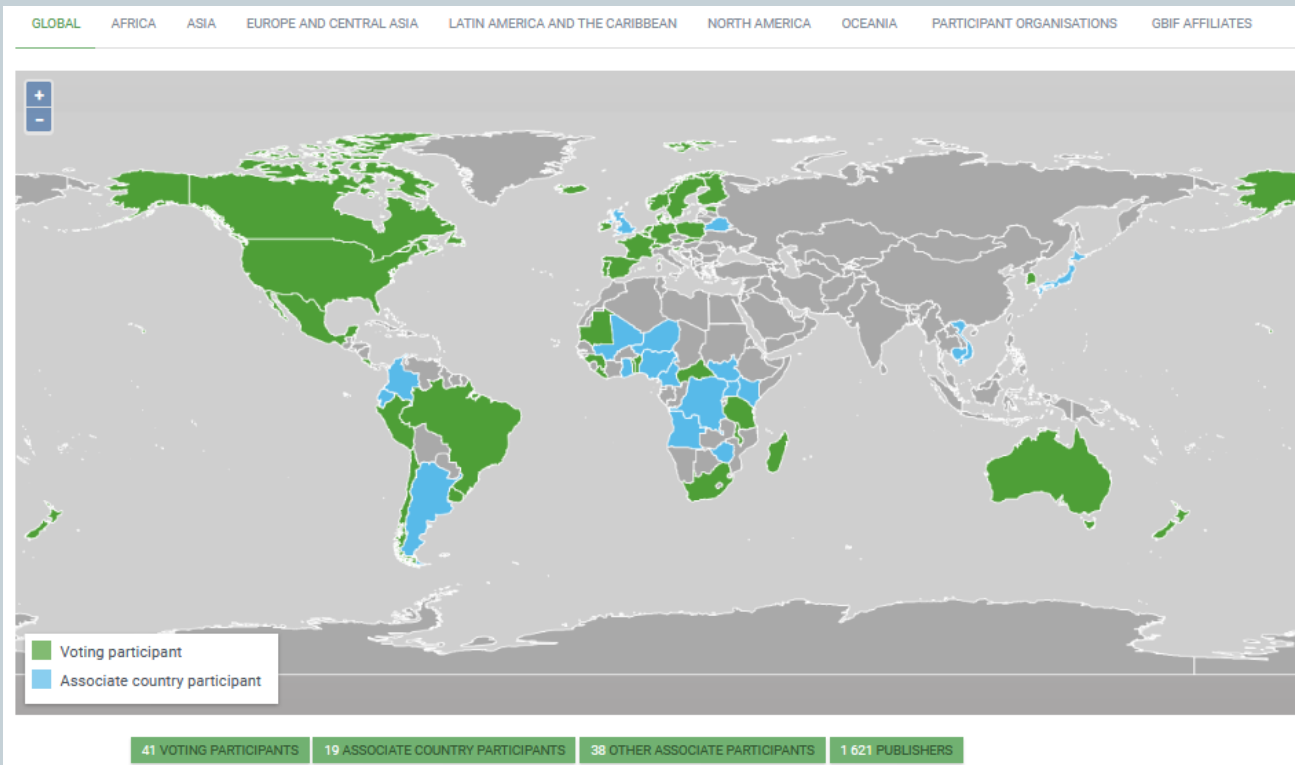




Global Biodiversity Information Facility Système Mondial d'Information sur la Biodiversité

> Accès libre et gratuit aux données de la biodiversité

- Programme intergouvernemental, créé en 2001, à l'initiative du comité scientifique de l'OCDE
- 60 pays et 38 organisations



79,8 milliards de données téléchargées par mois (moyenne mensuelle 2018)

www.gbif.org

Données de recherches, d'observations et de collections

pour

- mettre en valeur les données et leurs gestionnaires
- récompenser la publication des données



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



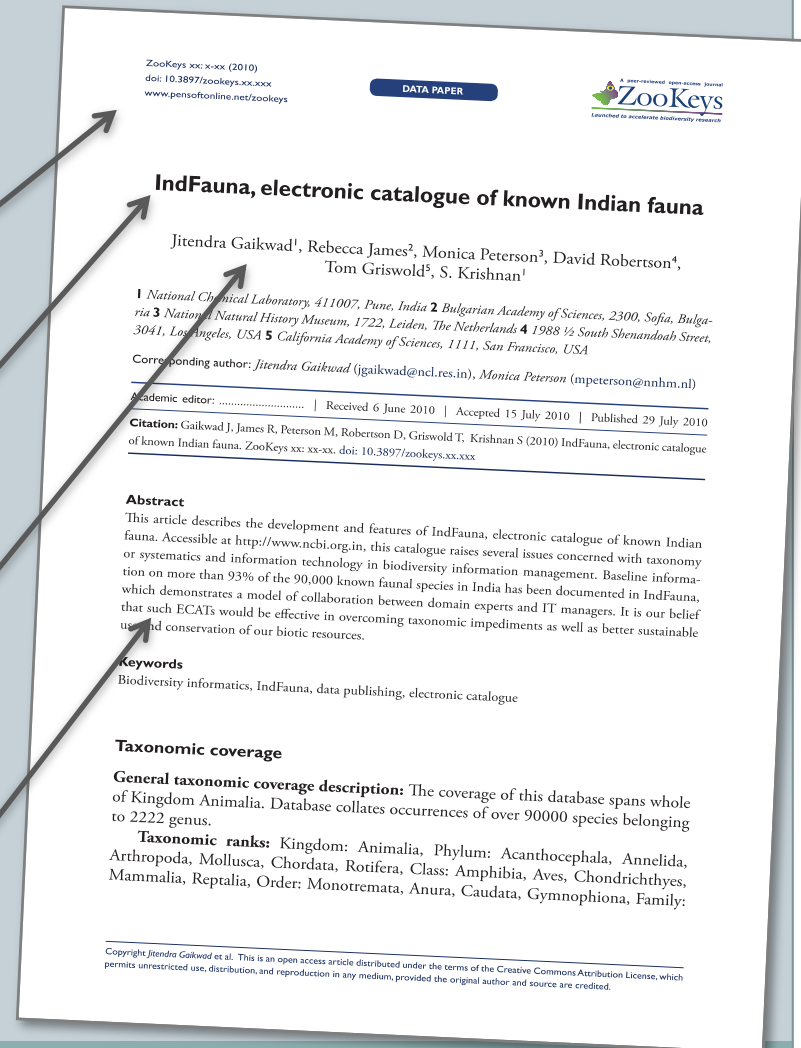
*Promouvoir et faire
connaître les données*



Reconnaissance des
éditeurs de données via
une publication
scientifique



Décrit les données sous
forme structurée et lisible
par un humain



Un outil (IPT) facilite le remplissage des métadonnées et la production automatisée d'un manuscrit de Data Paper



GBIF INTEGRATED PUBLISHING TOOLKIT (IPT)
free and open access to biodiversity data

email password login ENGLISH

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

Section

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 l'environnement (INRAE) on Feb 28, 2020

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 preferred habitat of young of the year salmon. The data consist of abundance indices and biological traits measured on the fish sampled: sex, maturity status, length, weight and age. The latter is ascertained from scale samples taken from the fish which size does not allow to determine their age unambiguously. The survey is carried out under the Research Observatory on Diadromes Fishes in Coastal Streamss (ERO DiAPFC) program. The data are stored in the database of the ERO. They are used to develop predictive models and tools for providing scientific advice to improve the management of this heritage species.



Home GBIF DwC-A EML RTF Versions Rights Cite this

GBIF INTEGRATED PUBLISHING TOOLKIT (IPT)
free and open access to biodiversity data

Logged in as pamerlon@gbif

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*

Abundances and biological traits of the juveniles salmon sampled in the survey of Salmon abundance Indices in

Publishing Organisation*

Institut national de recherc

Type*

Occurrence

Metadata Language*

English

Update Frequency*

Unknown

Subtype

Select a subtype

Data Language*

English

Data Licence*

Creative Commons Attribution (CC-BY) 4.0



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

Basic Metadata

[Geographic Coverage](#)

[Taxonomic Coverage](#)

[Temporal Coverage](#)

[Keywords](#)

[Associated Parties](#)

[Project Data](#)

[Sampling Methods](#)

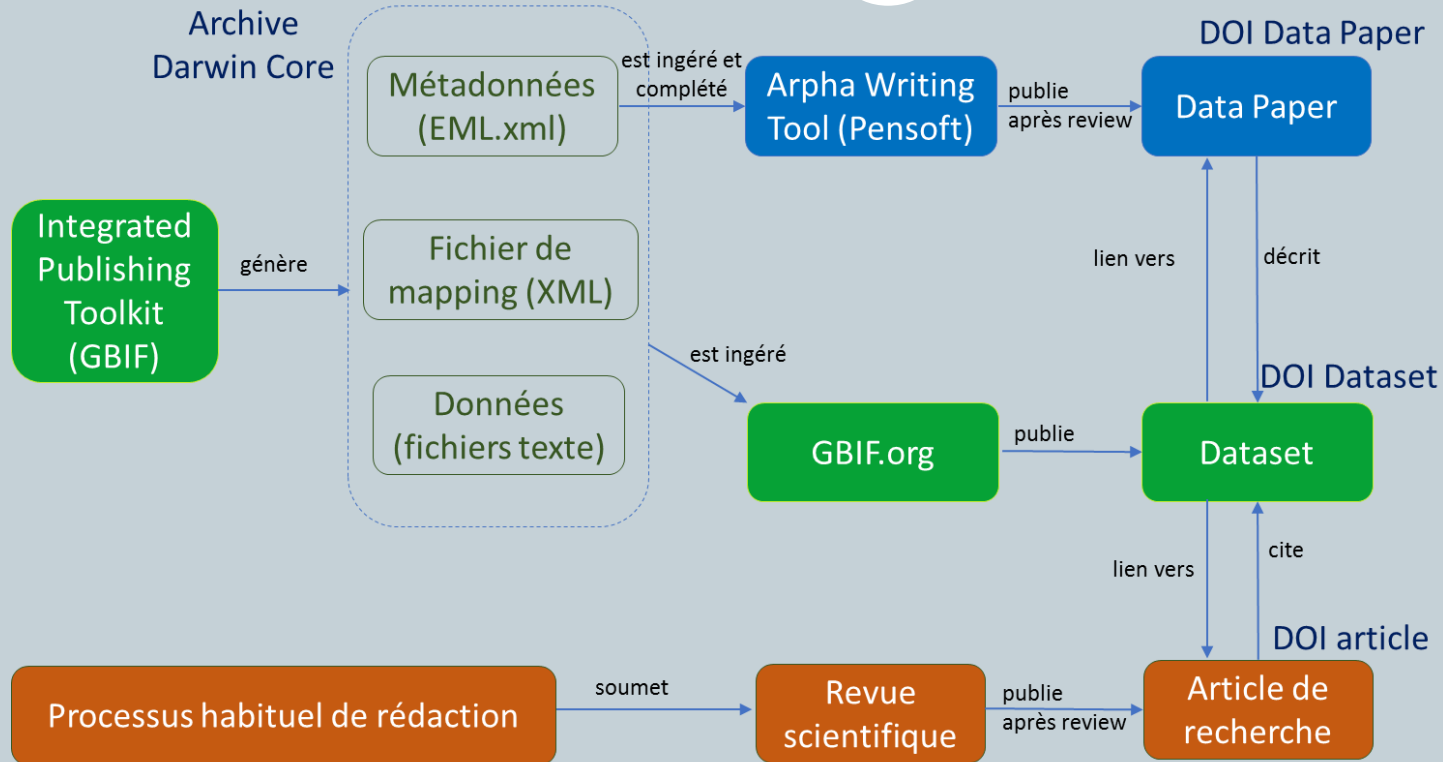
[Citations](#)

[Collection Data](#)

[External links](#)

[Additional Metadata](#)

Un outil (ARPHA Writing Tool) facilite la mise en page, la soumission, le processus de relecture, la publication, l'hébergement et l'archivage d'articles scientifiques (actuellement soumission directe à *Biodiversity Data Journal*)



<http://www.gbif.fr/page/contrib/publier-un-datapaper>

REVUES PUBLIANT DES DATA PAPERS LIÉS À LA BIODIVERSITÉ

Merci de communiquer toute correction ou mise à jour à l'adresse gbif@gbif.fr

Revue	Maison d'édition	Open Access (licence)	Estimation des coûts de publication	Facteur d'impact de la revue (2017)	Scopus CiteScore (2017)
Arxius de Miscel·lània Zoològica	Nat Hist Museum of Barcelona	CC BY	?	-	0.31



An occurrence records database of French Guiana harvestmen (Arachnida, Opiliones)

Sébastien Cally[†], Pierre Solbès[†], Bernadette Grosso[†], Jérôme Murienne[†]

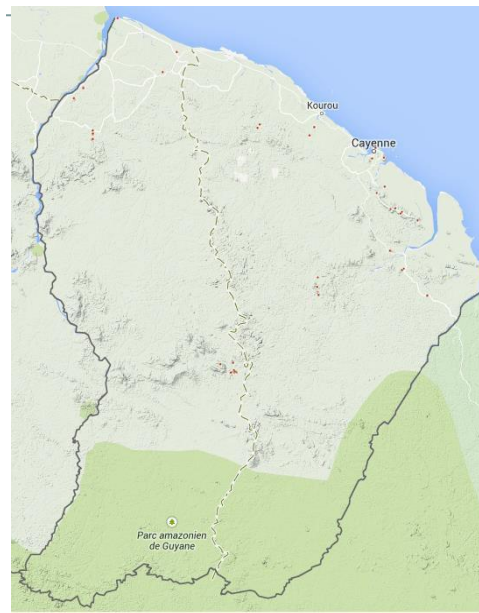
[†]UMR5174 CNRS/UPS/ENFA, Toulouse, France

Corresponding author: Sébastien Cally (sebastien.cally@univ-tlse3.fr), Jérôme Murienne (jerome.murienne@univ-tlse3.fr)

Academic editor: Adriano Kury

Received: 13 Nov 2014 | Accepted: 21 Dec 2014 | Published: 25 Dec 2014

Citation: Cally S, Solbès P, Grosso B, Murienne J (2014) An occurrence records database of French Guiana harvestmen (Arachnida, Opiliones). Biodiversity Data Journal 2: e4244. doi: [10.3897/BDJ.2.e4244](https://doi.org/10.3897/BDJ.2.e4244)



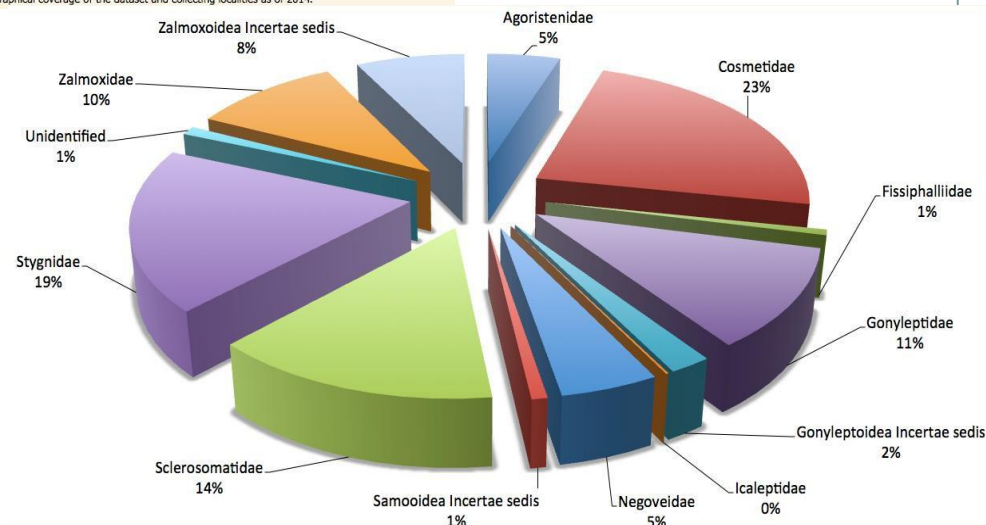
Geographical coverage of the dataset and collecting localities as of 2014.

Abstract

This dataset provides information on specimens of harvestmen (Arthropoda, Arachnida, Opiliones) collected in French Guiana. Field collections have been initiated in 2012 within the framework of the CEnter for the Study of Biodiversity in Amazonia (CEBA: www.labex-ceba.fr/en/). This dataset is a work in progress. Occurrences are recorded in an online database stored at the EDB laboratory after each collecting trip and the dataset is updated on a monthly basis. Voucher specimens and associated DNA are also stored at the EDB laboratory until deposition in natural history Museums. The latest version of the dataset is publicly and freely accessible through our Integrated Publication Toolkit at http://130.120.204.55:8080/ipt/resource.do?r=harvestmen_of_french_guiana or through the Global Biodiversity Information Facility data portal at <http://www.gbif.org/dataset/3c9e2297-bf20-4827-928e-7c7eefd9432c>.

Keywords

Occurrence, French Guiana, Neotropics, Opiliones.



1,126 OCCURRENCES WITH IMAGES



Harvestmen of French Guiana

Published by [Laboratoire EDB "Evolution et Diversité Biologique"](#)

Sébastien Cally

DATASET PROJECT METRICS ACTIVITY DOWNLOAD HOME PAGE

3,405 OCCURRENCES 19 CITATIONS

This dataset provides information on specimens of harvestmen (Arthropoda, Arachnida, Opiliones) collected in French Guiana. Field collections have been initiated in 2012 within the framework of the CEnter for the Study of Biodiversity in Amazonia (CEBA: www.labex-ceba.fr/en/). This dataset is a work in progress. Occurrences are recorded in an online database stored at the EDB laboratory after each collecting trip and the dataset is updated on a monthly basis. Voucher specimens and associated DN... [More](#)

Metadata last modified: July 2, 2020

Hosted by: [Laboratoire EDB "Evolution et Diversité Biologique"](#)

License: CC BY-NC 4.0

How to cite DOI: 10.15468/4nkd22

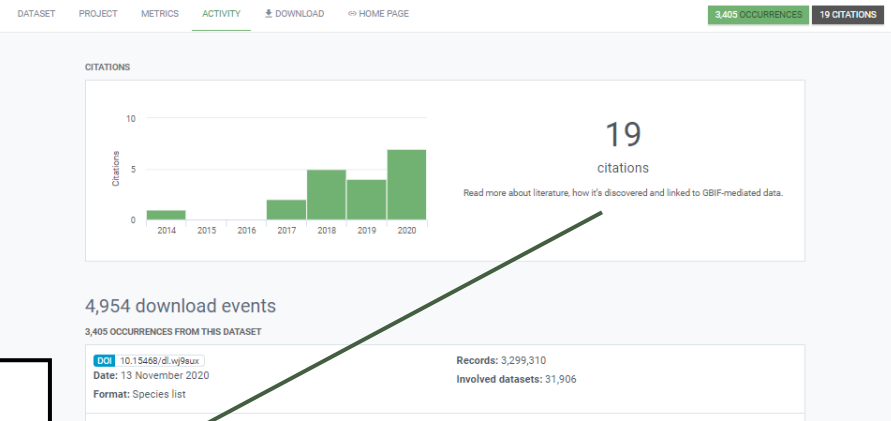
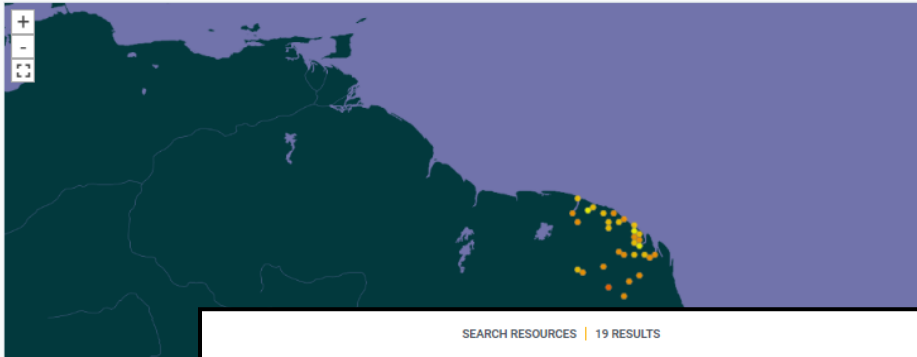
3,405 Occurrences

100% With taxon match

100% With coordinates

100%

3,405 GEOREFERENCED RECORDS



SEARCH RESOURCES | 19 RESULTS

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

Tropical and Mediterranean biodiversity is disproportionately sensitive to land-use and climate change [Literature](#)

Newbold, T., Oppenheimer, P., Etard, A., Williams, J. (2020) *Nature Ecology & Evolution*
Global biodiversity is undergoing rapid declines, driven in large part by changes to land use and climate. Global models help us to understand the consequences of environmental changes for biodiversity, but tend to neglect important geographical variation in the sensitivity of biodiversity to these ...

[Journal article](#) | [Peer-reviewed](#)

Data referenced in study DOI: 10.15468/dl.jt7f2u DOI: 10.15468/dl.kzyvzh DOI: 10.15468/dl.nmqqvj

DOI: 10.15468/dl.mfzxi DOI: 10.15468/dl.yjrrt5

Landscape Analysis for the Specimen Data Refinery [Literature](#)

Walton, S., Livermore, L., Bánki, O., Cubey, R., Drinkwater, R., Englund, M. ... - (2020) *Research Ideas and Outcomes*
This report reviews the current state-of-the-art applied approaches on automated tools, services and workflows for extracting information from images of natural history specimens and their labels. We consider the potential for repurposing existing tools, including workflow management systems; and ar...

collections digitisation • data reconciliation • data refinery • digitisation • linked open data • machine learning

[Journal article](#) | [Open access](#) | [Peer-reviewed](#)

Data referenced in study DOI: 10.15468/dl.8pg57z

- Connaissance de l'activité autour des données
- Intérêt pour la gestion des entrepôts de données