

Процесс публикации статьи о данных (data papers) в Biodiversity Data Journal

Любомир Пенев, Теодор Георгиев

Pensoft Publishers § Bulgarian Academy of Sciences, Sofia



БОЛЕЕ 50 ЖУРНАЛОВ ОТКРЫТОГО ДОСТУПА НА ARPHA



ИЗДАТЕЛЬСКАЯ ПЛАТФОРМА ПОЛНОГО ЦИКЛА НА ОСНОВЕ XML



AUTHORING - REVIEWING - PUBLISHING - HOSTING - ARCHIVING

FULL-FEATURED PUBLISHING PLATFORM AND SERVICES



Как публиковать данные?

The screenshot displays the R10 website interface. At the top, the R10 logo is on the left, and navigation links (Home, Articles, About, About Pensoft, Books, E-Books, Blog, Journals) are in the center. On the right, there's a user profile for Lyubomir Penev with a 'My tasks' link. Below the navigation bar, a teal 'Guidelines' button is on the left, followed by the article title 'Research Ideas and Outcomes 3: e12431' and its DOI link. A red 'Reviewed' badge and a 'v1' version indicator are also present. The main content area on the left features the article title 'Strategies and guidelines for scholarly publishing of biodiversity data' and a list of authors. The abstract text describes the paper's focus on biodiversity data publishing policies. On the right, a sidebar contains a 'Contents' tab and a list of article sections, including 'Data Publishing in a Nutshell' and 'Data Publishing Policies'. The bottom right corner indicates the site is 'Powered by arpha'.

R10 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/r10.3.e12431> (28 Feb 2017) Reviewed v1

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.

Contents

- Article info
- Citation
- Metrics
- Reviews
- Related

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

Powered by arpha

2011: Концепция статей о данных

 **BMC** Part of Springer Nature

[Explore Journals](#)

[Get Published](#)

[About BMC](#)

Search 

[Login](#)

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 [†]  and Lyubomir Penev [†]

[†]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



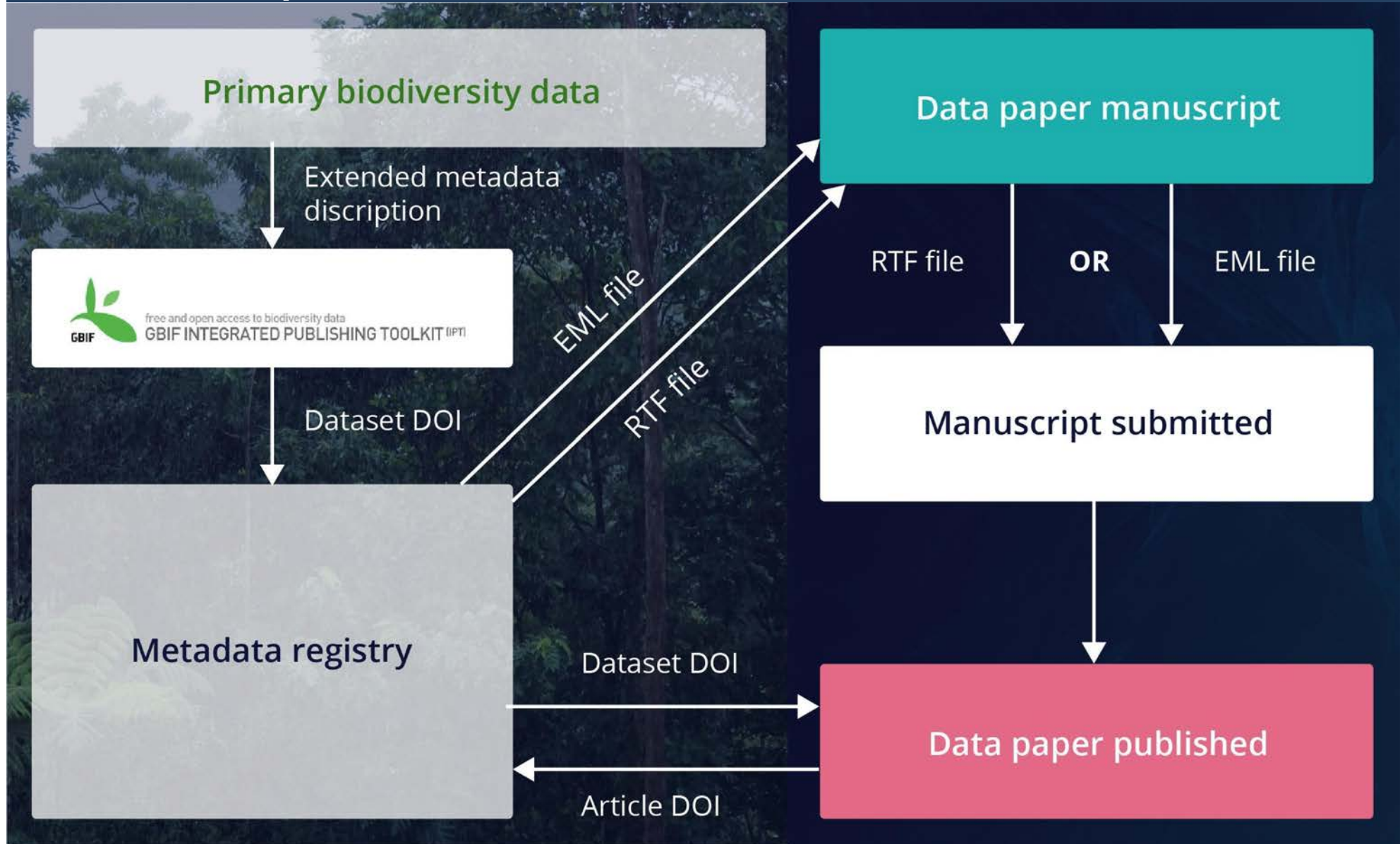
Share This Article



[Get shareable link](#)

Abstract

Почему Open Data: процесс публикации статей о данных при помощи ARPHA



Создайте статью о данных из Ecological Metadata Language (EML) метаданных

Наборы данных из GBIF EML метаданных



Canadensys

Français

explorer ▾

repository

tools

vascan

community ▾

login

Database of Vascular Plants of Canada (VASCAN)

Latest version published by Université de Montréal Biodiversity Centre on Aug 19, 2019



The Database of Vascular Plants of Canada or VASCAN (<http://data.canadensys.net/vascan>) is a comprehensive and curated checklist of all vascular plants reported in Canada, Greenland (Denmark), and Saint Pierre and Miquelon (France). VASCAN was developed at the Université de Montréal Biodiversity Centre and is maintained by a group of editors and contributors. For every core taxon in the checklist (species, subspecies, or variety), VASCAN provides the accepted scientific name, the accepted French and English vernacular names, and their synonyms/alternatives in Canada, as well as the distribution status (native, introduced, ephemeral, excluded, extirpated, doubtful or absent) of the plant for each province or territory, and the habit (tree, shrub, herb and/or vine) of the plant in Canada. For reported hybrids (nothotaxa or hybrid formulas) VASCAN also provides the hybrid parents, except if the parents of the hybrid do not occur in Canada. All taxa are linked to a classification. VASCAN refers to a source for all name, classification and distribution information. All data have been released to the public domain under a CC0 waiver and are available through Canadensys and the Global Biodiversity Information Facility (GBIF).

DwC-A



EML



RTF



GBIF



Database of Vascular Plants of Canada (VASCAN): a community contributed taxonomic checklist of all vascular plants of Canada, Saint Pierre and Miquelon, and Greenland

Peter Desmet¹, Luc Brouillet¹

¹ *Université de Montréal Biodiversity Centre, 4101 rue Sherbrooke est, H1X2B2, Montreal, Canada*

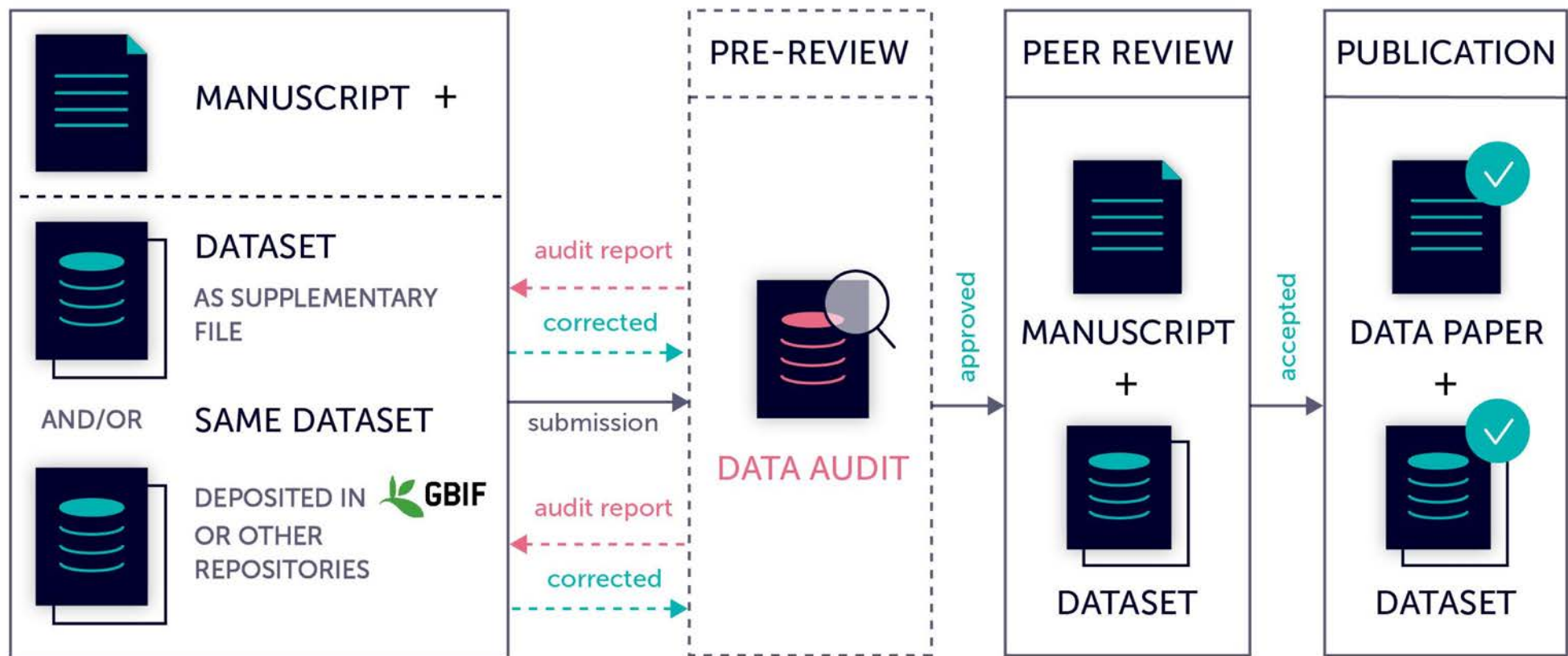
Corresponding author: *Peter Desmet* (peter.desmet@umontreal.ca)

Academic editor: *Vishwas Chavan* | Received 19 March 2012 | Accepted 17 July 2013 | Published 24 July 2013

Citation: Desmet P, Brouillet L (2013) Database of Vascular Plants of Canada (VASCAN): a community contributed taxonomic checklist of all vascular plants of Canada, Saint Pierre and Miquelon, and Greenland. *PhytoKeys* 25: 55–67. doi: 10.3897/phytokeys.25.3100 Resource ID: GBIF key: 3f8a1297-3259-4700-91fc-acc4170b27ce

Resource citation: Brouillet L, Desmet P, Coursol F, Meades SJ, Favreau M, Anions M, Bélisle P, Gendreau C, Shorthouse D and contributors* (2010+). Database of Vascular Plants of Canada (VASCAN). 27189 records. Online at <http://data.canadensys.net/vascan>, <http://dx.doi.org/10.5886/Y7SMZY5P>, and <http://www.gbif.org/dataset/3f8a1297-3259-4700-91fc-acc4170b27ce>, released on 2010-12-10, version 24 (last updated on 2013-07-22). GBIF key: 3f8a1297-3259-4700-91fc-acc4170b27ce. Data paper ID: <http://dx.doi.org/10.3897/phytokeys.25.3100>

Аудит данных / процесс очистки



Отчет об аудите

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.au>)
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 passes or has only minor problems, the data paper manuscript is referred to reviewers. If the dataset has major problems, a review of the paper is postponed until 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 whether the correct author is given for a 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>).

ИСТОЧНИКИ:

<https://blog.pensoft.net/2019/10/17/could-biodiversity-data-be-finally-here-to-last/>

<https://blog.pensoft.net/2019/10/17/case-study-data-audit-for-the-vascular-plants-dataset-of-the-cofc-herbarium-university-of-cordoba-spain-a-data-paper-in-phytokeys/>

Список ошибок

(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

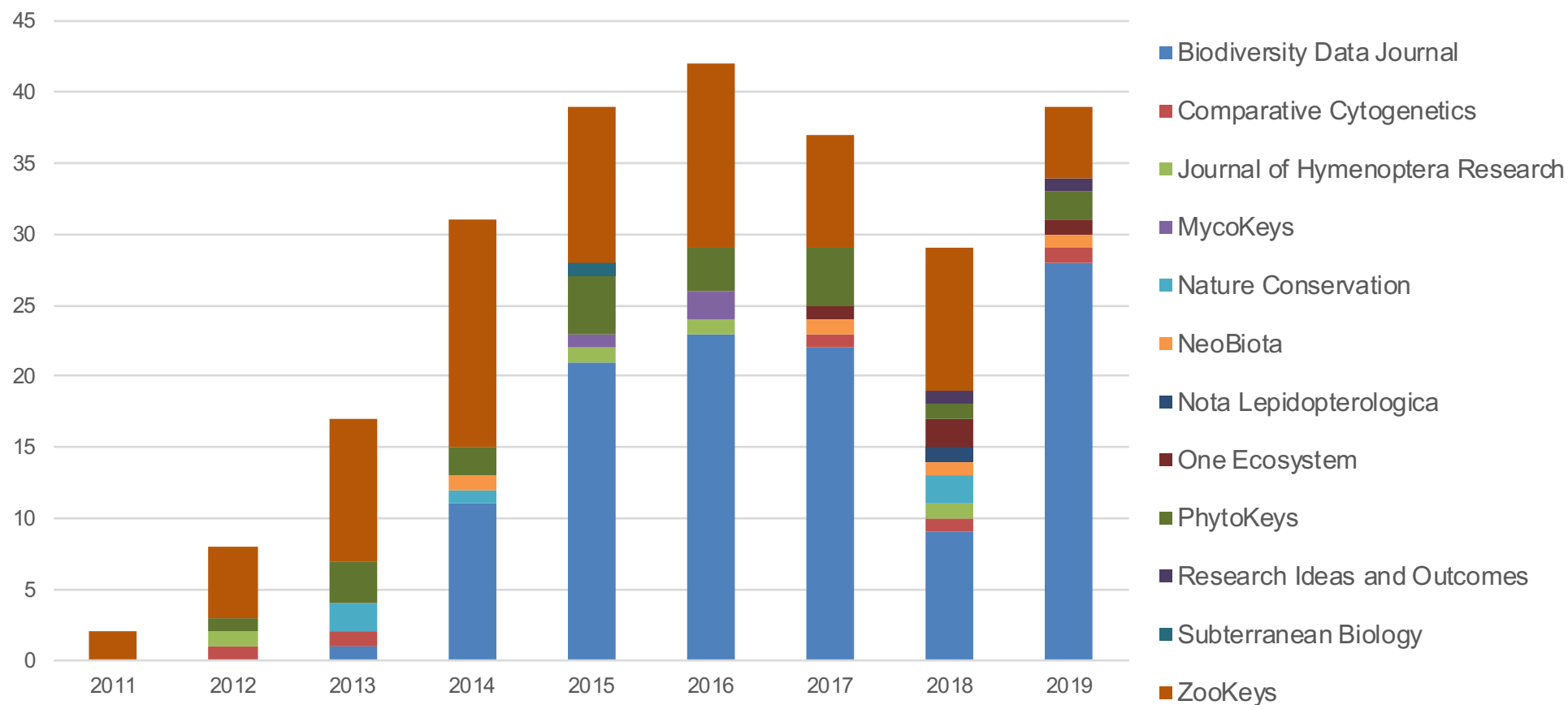
Also, several entries have too many significant figures and should be rounded off, e.g. "37.0233172796695"

ИСТОЧНИКИ:

<https://blog.pensoft.net/2019/10/17/could-biodiversity-data-be-finally-here-to-last/>

<https://blog.pensoft.net/2019/10/17/case-study-data-audit-for-the-vascular-plants-dataset-of-the-cofc-herbarium-university-of-cordoba-spain-a-data-paper-in-phytokeys/>

Статистика статей о данных



Подпись: Статьи о данных публикуются в 12 журналах открытого доступа из эксклюзивного портфолио Pensoft.