

Научно-методический семинар «Естественно-научные коллекции Югры 2021»

30 марта — 2 апреля 2021 года, Музей Природы и Человека, Ханты-Мансийск

Публикация статьи формата Data Paper – технические навыки и возможности грантов GBIF

Чадин Иван

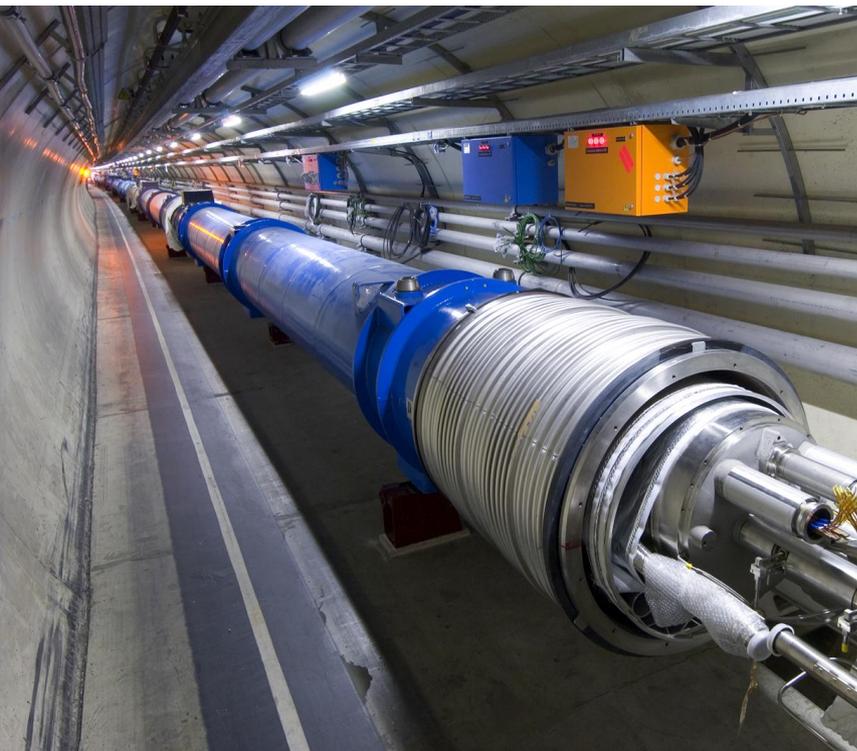
Институт биологии Коми научного центра Уральского отделения Российской академии наук

Биология – преимущественно эмпирическая наука

Результативность работы в ней напрямую зависит от скорости накопления данных

Физика - лидер по скорости накопления данных

Большой адронный
коллайдер



30 петабайт / год



≈ 1.2 млн дисков Blu-ray

Молекулярная биология, биомедицина – лидирующие отрасли биологии по скорости накопления данных



**ДНК секвенатор
Illumina NovaSeq6000**

до 4.35 терабайт / 44 часа

Козырь в рукаве исследователей биоразнообразия – данные биологические коллекций и полевой документации



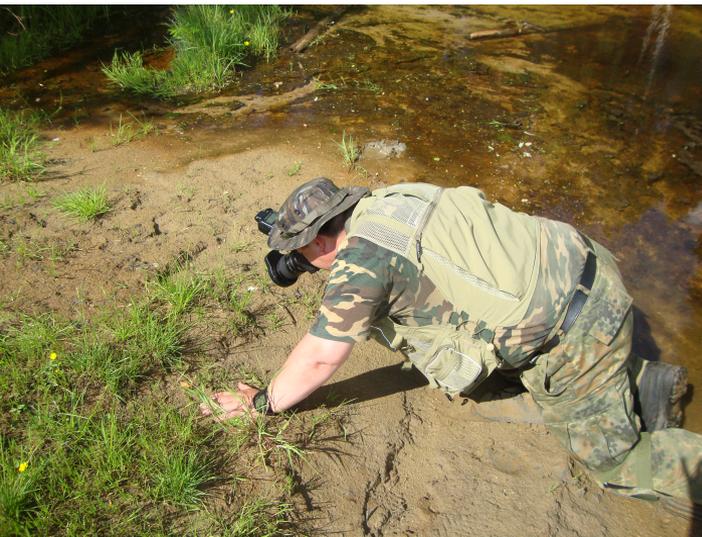
Классическая биология может резко, скачкообразно нарастить объем данных, необходимых для ее развития, если сможет в короткое время оцифровать биологические коллекции.

В биологических коллекциях содержится трудно извлекаемая, но бесценная информация, накопленная за сотни тысяч человеко-лет труда профессиональных исследователей.

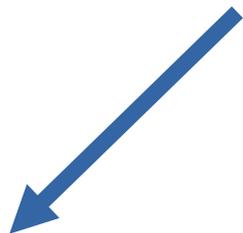
Открытый доступ к первичным данным о биологическом разнообразии необходим для:

- развития фундаментальных исследований в областях систематики, эволюции, экологии;
- решения природоохранных задач;
- принятия управленческих решений;
- общественного контроля за решениями, в области природопользования;
- ...

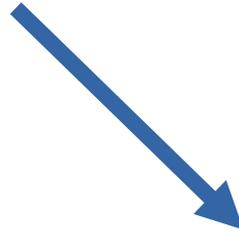
Исследователь – ключевой элемент в процессах получения, оцифровки и мобилизации данных о биоразнообразии



Два препятствия на пути мобилизации данных



Проблема признания
вклада лиц,
опубликовавших данные



Проблема качества
опубликованных
данных

Проблема признания вклада лиц, опубликовавших данные



Database, 2018, 1–6
doi: 10.1093/database/bay033
Perspective/Opinion



Perspective/Opinion

The tragedy of the biodiversity data commons: a data impediment creeping nigher?

Nora Escribano*, David Galicia and Arturo H. Ariño

Universidad de Navarra, Department of Environmental Biology, Biodiversity and Environmental Quality Data Analysis Group, 31008, Pamplona, Spain

*Corresponding author: Tel: +34 948 425 600 Ext 806281; Fax number: +34 948 425 649; Email: nescribano@alumni.unav.es

Citation details: Escribano, N., Galicia, D., and Ariño, A. H. The tragedy of the biodiversity data commons: a data impediment creeping nigher? *Database* (2018) Vol. 2018: article ID bay033; doi:10.1093/database/bay033

Received 10 January 2018; Revised 20 February 2018; Accepted 13 March 2018

Abstract

Researchers are embracing the open access movement to facilitate unrestricted availability of scientific results. One sign of this willingness is the steady increase in data freely

Признание значительных усилий, связанных с созданием, управлением и публикацией наборов данных, по-прежнему остается непрочной, не вполне устоявшейся практикой в научном сообществе.

Отсутствие норм, позволяющих авторам наборов данных получить признание, соответствующее их вкладу в науку может привести к сокращению числа качественных наборов данных, свободно доступных в сети Интернет»

Database, Volume 2018, 2018, bay033,
<https://doi.org/10.1093/database/bay033>

Проблема качества опубликованных данных

Global Ecology and Biogeography / Volume 24, Issue 8 / p. 973-984

Research Papers

 Open Access



Estimating species diversity and distribution in the era of Big Data: to what extent can we trust public databases?

Carla Maldonado, Carlos I. Molina, Alexander Zizka, Claes Persson, Charlotte M. Taylor, Joaquina Albán, Eder Chilquillo, Nina Rønsted , Alexandre Antonelli

First published: 25 May 2015

<https://doi.org/10.1111/geb.12326>

Citations: 126

Проблема качества опубликованных данных



Ecological Informatics

Volume 19, January 2014, Pages 10-15



Spatial bias in the GBIF database and its effect on modeling species' geographic distributions

Jan Beck ^a  , Marianne Böller ^{a, b}, Andreas Erhardt ^a, Wolfgang Schwanghart ^{b, 1}

Одно решение двух проблем

Editorial

Data Papers – Peer Reviewed Publication of High Quality Data Sets

The International Journal of Robotics Research

First Published May 1, 2009 Editorial

<https://doi.org/10.1177/0278364909104283>

This issue marks the launch of an undertaking by *The International Journal of Robotics Research* (IJRR) to solicit and publish a new genre of journal paper: a “data paper”. Our prime goal is to facilitate and encourage the release of high-quality, peer-reviewed datasets to the robotics community. We hope to implicitly create benchmark tests for use across many fields of robotics research. We recognize that acquiring high-calibre data is a substantial technical endeavour and one that need not be

the data itself but on easing access to it to ensure its widespread adoption. Note also that data paper authors will retain the copyright and ownership of their data.

It is the intent of IJRR to solicit archival quality data, which means the provision of an archival service to authors. This will be an ongoing process, and while we aim ultimately to host all datasets and their supporting websites, it is anticipated that initially many data papers will refer to websites hosted author-

Одно решение двух проблем

Chavan and Penev *BMC Bioinformatics* 2011, **12**(Suppl 15):S2
<http://www.biomedcentral.com/1471-2105/12/S15/S2>



RESEARCH

Open Access

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

Vishwas Chavan^{1,2*†}, Lyubomir Penev^{1,2†}

Background: Free and open access to primary biodiversity data is essential for informed decision-making to achieve conservation of biodiversity and sustainable development. However, primary biodiversity data are neither easily accessible nor discoverable. Among several impediments, one is a lack of incentives to data publishers for publishing of their data resources. One such mechanism currently lacking is recognition through conventional scholarly publication of enriched metadata, which should ensure rapid discovery of 'fit-for-use' biodiversity data resources.

Discussion: We review the state of the art of data discovery options and the mechanisms in place for incentivizing

Сравнение с традиционными научными статьями

Статья о данных

Описывает методы получения, структуру и объем данных

Публикация и внешняя верификация набора данных обязательно предшествует публикации статьи

Подлежит рецензированию и текст статьи, и набор данных

Статья о результатах исследования

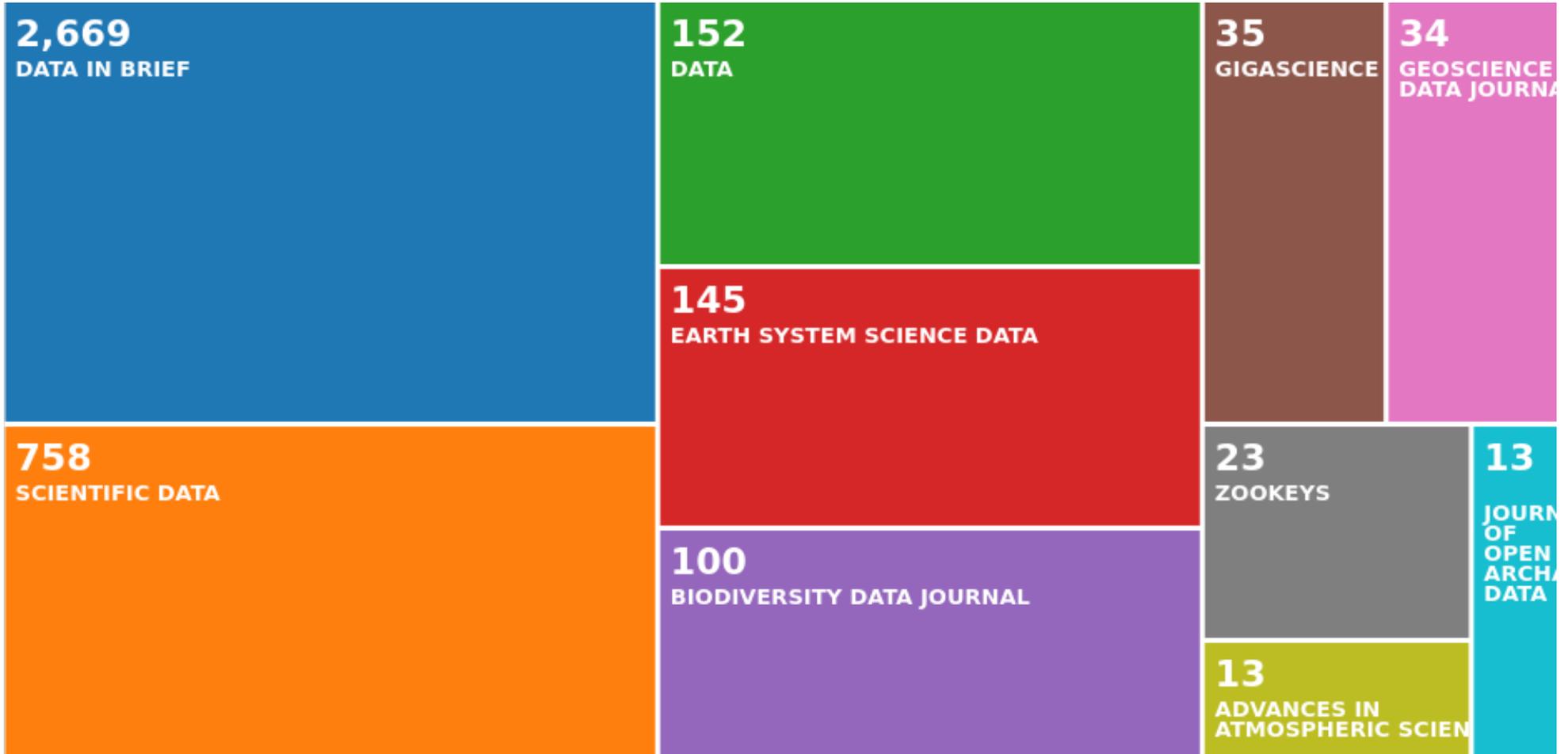
Описывает проблему, гипотезу, методы, результаты и их обсуждение

Публикация и верификация первичных данных на усмотрение авторов / редколлегии

Подлежит рецензированию текст статьи

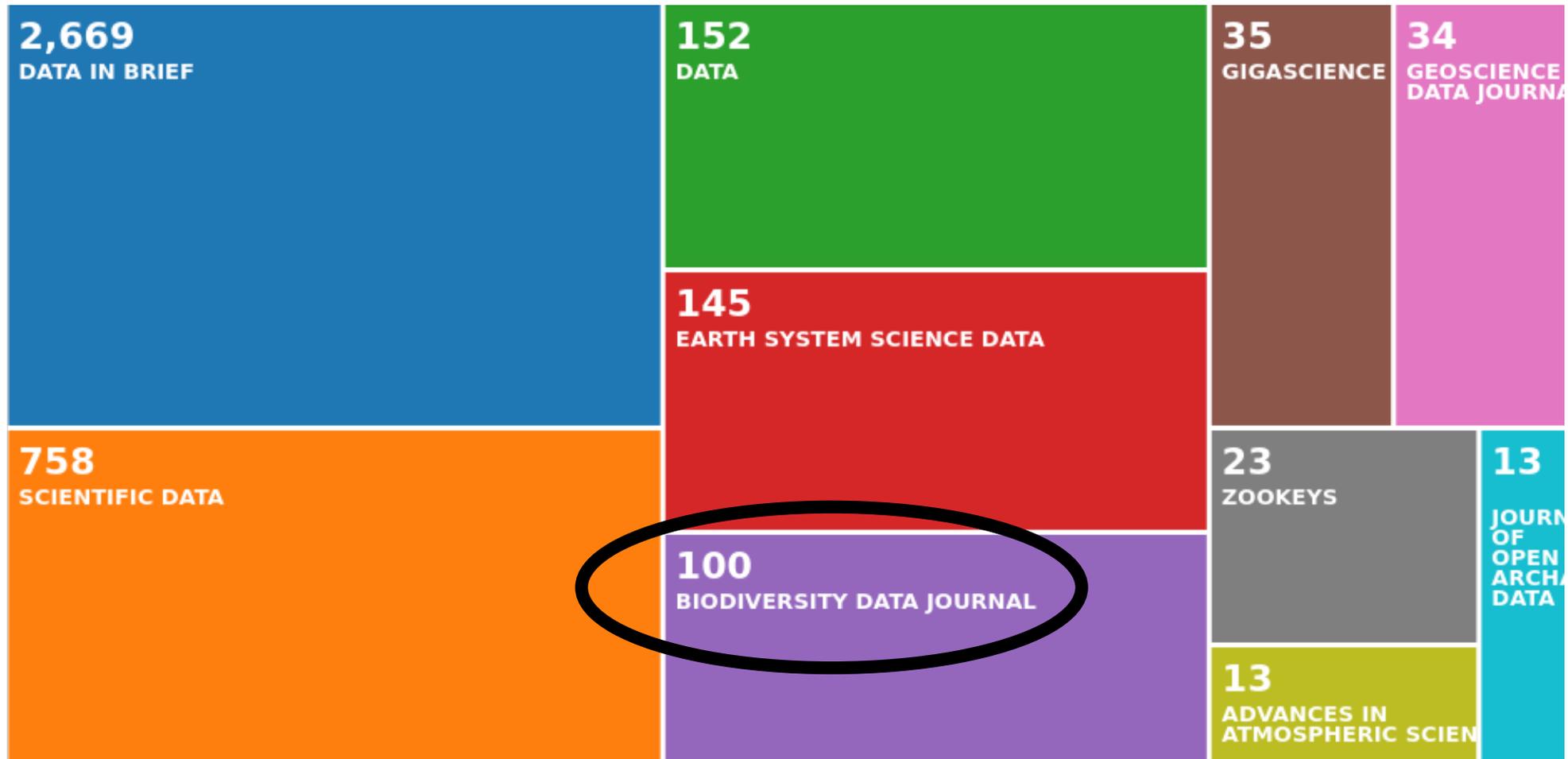
Распределение числа статей о данных среди журналов WoS

(результаты запроса по ключевому слову «**dataset**» в базе данных Web of Science 27/03/2021)



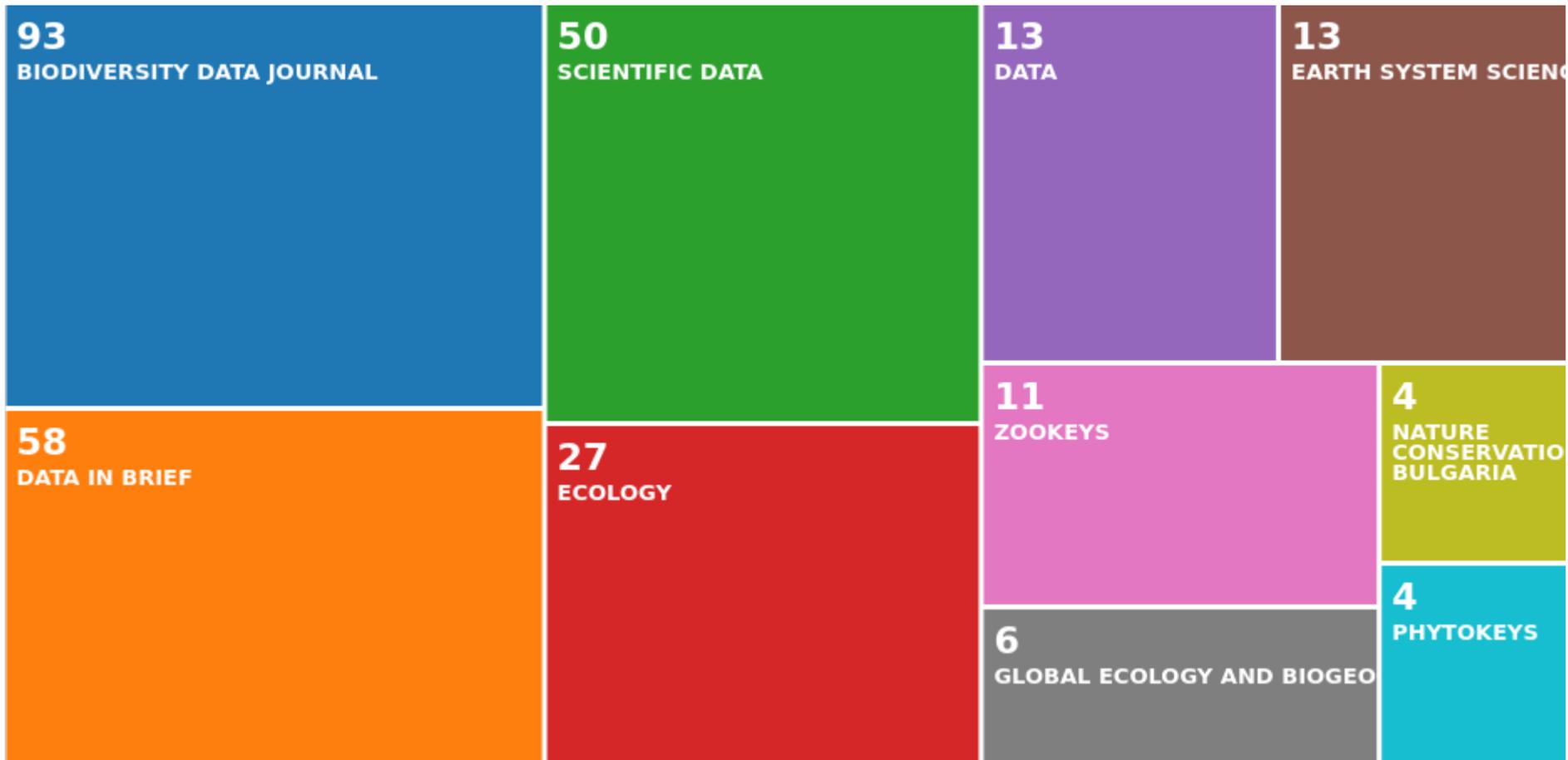
Распределение числа статей о данных среди журналов WoS

(результаты запроса по ключевому слову «**dataset**» в базе данных Web of Science 27/03/2021)



Распределение числа статей о данных среди журналов WoS

(результаты запроса по ключевому слову «**biodiversity**» в базе данных Web of Science 27/03/2021)



Распределение числа статей о данных среди журналов WoS

(результаты запроса по ключевому слову «**biodiversity**» в базе данных Web of Science 27/03/2021)



Journal name	Publisher	Subject	Impact factor (2019)	Publishing model	No. of data papers	No. of articles	% (data papers/articles)
<i>Data in Brief</i>	Elsevier	Multidisciplinary sciences	NA	OA	5,044	5,238	96.30
<i>Scientific Data</i>	Springer Nature	Multidisciplinary sciences	5.541	OA	990	1,226	80.75
<i>Human Genome Variation</i>	Springer Nature	Genetics, heredity	NA	OA	78	116	67.24
<i>Earth System Science Data</i>	Copernicus Publications	Geosciences, multidisciplinary	9.197	OA	285	511	55.77
<i>Geoscience Data Journal</i>	Wiley	Geosciences, multidisciplinary	2.714	OA	40	74	54.05
<i>Journal of Open Archaeology Data</i>	Ubiquity Press	Archaeology	NA	OA	11	22	50.00
<i>Data</i>	MDPI	Computer science information systems	NA	OA	134	272	49.26
<i>Gigascience</i>	Oxford University Press	Multidisciplinary sciences	5.993	OA	149	594	25.08
<i>Biodiversity</i>	Pensoft	Biodiversity	1.331	OA	125	562	22.24

Journal name	Publisher	Subject	Impact factor (2019)	Publishing model	No. of data papers	No. of articles	% (data papers/articles)
<i>Data in Brief</i>	Elsevier	Multidisciplinary sciences	NA	OA	5,044	5,238	96.30
<i>Scientific Data</i>	Springer Nature	Multidisciplinary sciences	5.541	OA	990	1,226	80.75
<i>Human Genome Variation</i>	Springer Nature	Genetics, heredity	NA	OA	78	116	67.24
<i>Earth System Science Data</i>	Copernicus Publications	Geosciences, multidisciplinary	9.197	OA	285	511	55.77
<i>Geoscience Data Journal</i>	Wiley	Geosciences, multidisciplinary	2.714	OA	40	74	54.05
<i>Journal of Open Archaeology Data</i>	Ubiquity Press	Archaeology	NA	OA	11	22	50.00
<i>Data</i>	MDPI	Computer science information systems	NA	OA	134	272	49.26
<i>Gigascience</i>	Oxford University Press	Multidisciplinary sciences	5.993	OA	149	594	25.08
<i>Biodiversity</i>	Pensoft	Biodiversity	1.331	OA	125	562	22.24



A peer-reviewed open-access journal
**Biodiversity
Data Journal**
Making your data count! ISSN 1314-2828 (online)



[Submit manuscript](#)

[Full Text](#) [Author](#) [Title](#)

[Articles](#)

[About](#)

[Guidelines for Authors](#)

[Data Publishing Guidelines](#)

[Data Review Guidelines](#)

[Open Topical Collections](#)

[Editorial Team](#)

[Contacts](#)





A peer-reviewed open-access journal
**Biodiversity
Data Journal**
Making your data count! ISSN 1314-2828 (online)



[Submit manuscript](#)

[Full Text](#) [Author](#) [Title](#)

[Articles](#)

[About](#)

[Guidelines for Authors](#)

[Data Publishing Guidelines](#)

[Data Review Guidelines](#)

[Open Topical Collections](#)

[Editorial Team](#)

[Contacts](#)



*Scopus — Q2
Web of Science — Q3*

Критерии для принятия рукописи к публикации в BDJ

<https://bdj.pensoft.net/about#Criteriaforpublication>

1. Новизна
2. Объем
3. Данные должны быть опубликованы до публикации статьи
4. Данные должны быть опубликованы без ограничений на использование, кроме обязательства ссылаться на их источник
5. Данные прошли проверку на отсутствие технических ошибок

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

**Если Вы смогли опубликовать набор данных*,
то Вы сможете подготовить статью о данных**

** достаточно большой объем новых данных*

NEWS | 23 MARCH 2021

Call for data papers describing datasets from Russia

GBIF partners with FinBIF and Pensoft to support publication of new datasets about biodiversity from across Russia



Siberian buggloss (*Brunnera sibirica*) observed in Kirovo-Chepetsk, Kirov Oblast, Russia by Вотинцева Елена Александровна (CC BY-NC 4.0)

<https://www.gbif.org/news/HdDwtyv9jjrvG99OSHsP6/call-for-data-papers-describing-datasets-from-russia>

**Первые 36 рукописей статей о данных, которые:
присланы в между 23.03.2021 и 15.09.2021
и соответствуют критериям конкурса
будут опубликованы в Biodiversity Data Journal
бесплатно**

Критерии отбора статей:

- более 5000 новых записей, опубликованных в GBIF в 2021 году*
- высокое качество данных и метаданных*
- данные привязаны к территории России*

<https://www.gbif.org/news/HdDwtyv9jjrvG99OSHsP6/call-for-data-papers-describing-datasets-from-russia>

**Издание тематического выпуска
"Biodiversity Informatics for Nature Conservation"
журнала
"Nature Conservation Research. Заповедная наука"
<http://ncr-journal.bear-land.org/page/about-journal>
поддержано GBIF
публикация для авторов бесплатная**

Напоминаем, что в выпуск принимаются рукописи исследовательских работ, основанные на сведениях, полученных из открытых информационных систем или собственных данных, опубликованных в GBIF. Работа может быть посвящена описанию опубликованного набора данных. Согласно правилам журнала и условиям финансирования (проект GBIF N-Eurasia-Russia2021) большая часть данных должна быть посвящена ООПТ или охраняемым видам и относиться к территории России.

Подробная информация в анонсе выпуска:

https://drive.google.com/file/d/1Gtnv_K8hbilmDz5AbJ3pdoNJaSVobpZX

Рукописи следует присылать на электронную почту журнала до 1 июня 2021 года.