



**III Всероссийская научно-практическая конференция «Безопасный Север – чистая Арктика»
11-12 ноября 2020, Сургутский государственный университет**

Открытые данные о биоразнообразии для науки и практики

Наталья Иванова

*Институт математических проблем биологии РАН - филиал ИПМ им. М.В. Келдыша РАН
г. Пущино Московской обл.*



FIELD NOTES®

№	Входные данные	Результат анализа
11	1	менее востреб
12	1/4	чем результа
13	1р	анализа
14	1р	
15	1р	
16	1р 344	Береза
17	1р 475	Ву
18	1р 523	Лето
19	1р 466	Береза
20	1р 202	а

21 августа
Hortensia

A photograph of a plant specimen, likely a stem with several leaves, resting on a blue grid paper. The leaves are elongated and show signs of damage or decay, with some appearing brown and others green. The stem is dark and woody.

PCA 5 vars
princomp(x = data, cor = cor)

(1-3) 60%

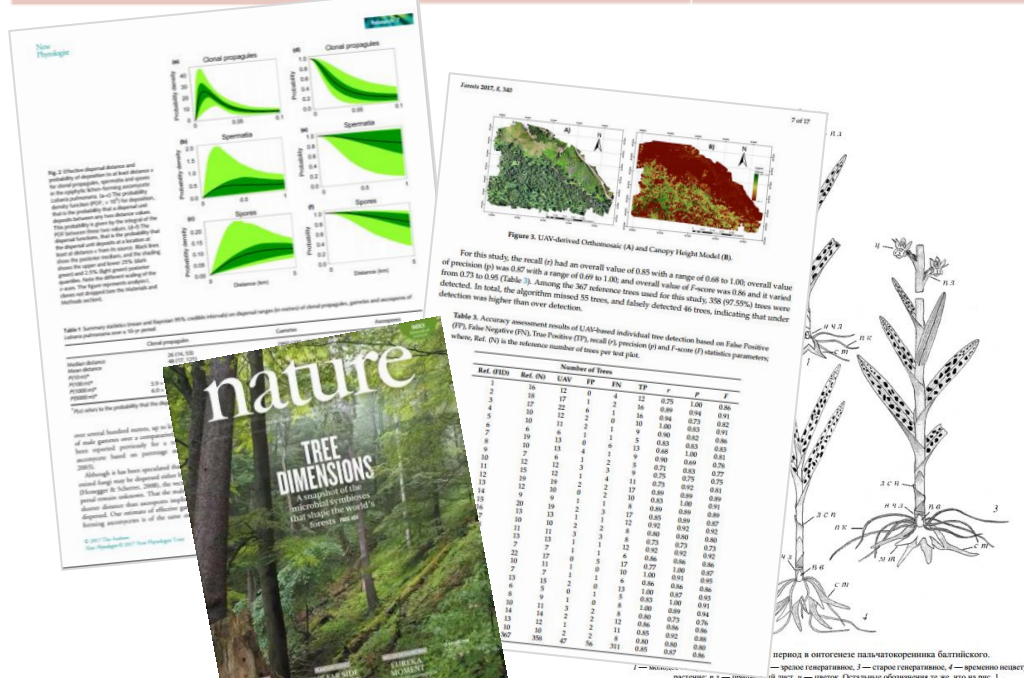
Clustering 4 groups

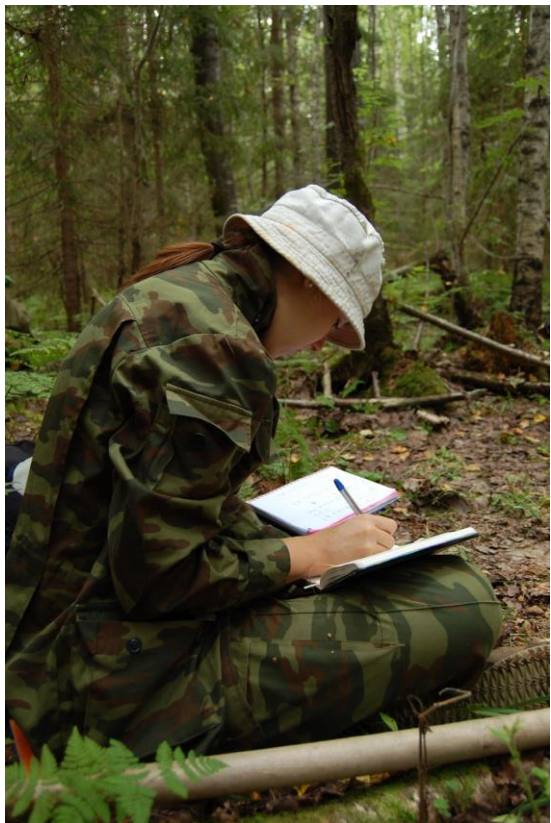
Factor 1 [41%]

Factor 3 [19%]

V. De Geneve

<https://www.r-project.org/>





Полевые данные ценны сами по себе и должны быть доступны для повторного анализа



Идея

Гипотеза

Дизайн
эксперимента



Полевые
исследования

Объединенные
данные

```
00101011
01101010
101110101
11011000
10100110
```

Данные из
репозитория

```
00101011
01101010
101110101
11011000
10100110
```

Данные из
репозитория

```
00101011
01101010
101110101
11011000
10100110
```


Анализ
объединенных
данных



FAIR принципы данных

Данные сами по себе являются научным продуктом


Для этого данные должны быть:

Findable 


Легко обнаружимыми через Интернет

Accessible 

Доступными для исследователей

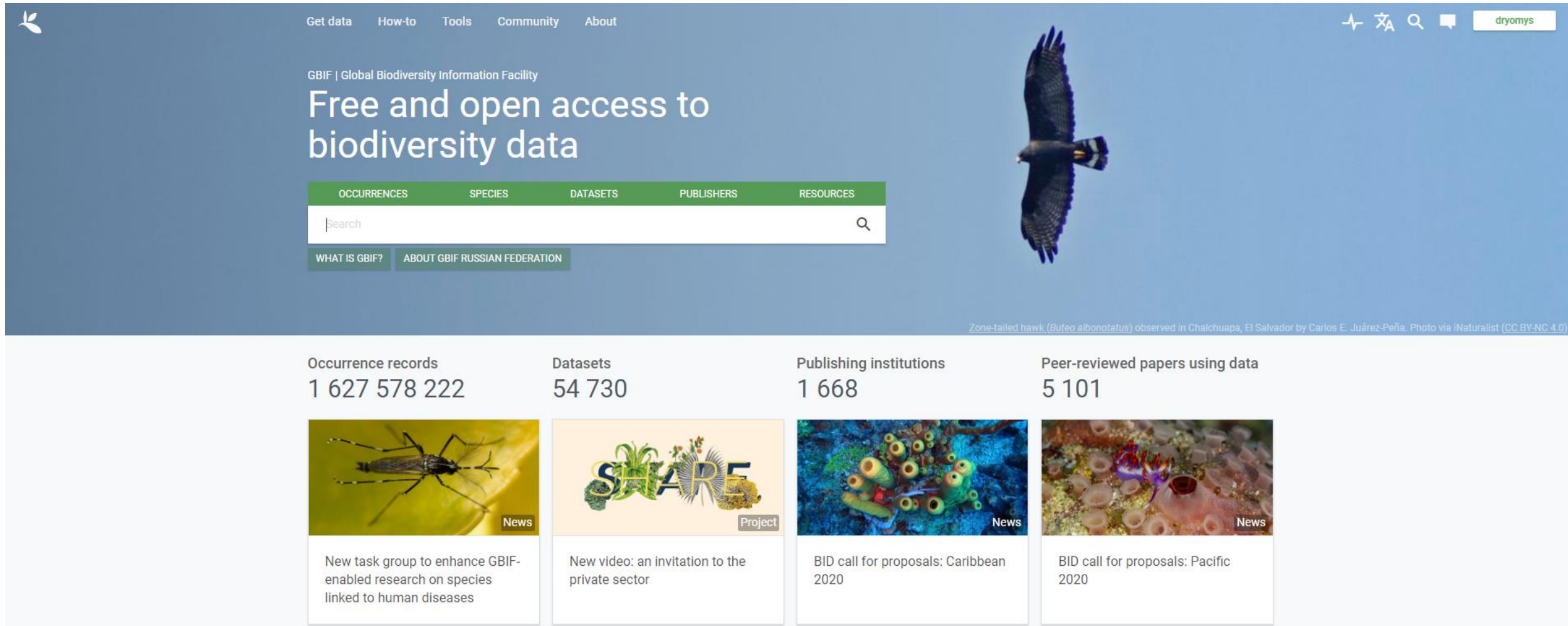
Interoperable 

Совместимыми друг с другом

Reusable 

Многократно используемыми
Однозначное указание авторства
Правила использования данных (лицензия)

Портал GBIF – крупнейший в мире источник открытых данных о биоразнообразии



GBIF | Global Biodiversity Information Facility

Free and open access to biodiversity data


OCCURRENCES SPECIES DATASETS PUBLISHERS RESOURCES

Search

WHAT IS GBIF? ABOUT GBIF RUSSIAN FEDERATION


Zone-tailed hawk (*Buteo albonotatus*) observed in Chalchuapa, El Salvador by Carlos E. Juárez-Peña. Photo via iNaturalist (CC BY-NC 4.0)

Occurrence records	Datasets	Publishing institutions	Peer-reviewed papers using data
1 627 578 222	54 730	1 668	5 101




News

New task group to enhance GBIF-enabled research on species linked to human diseases



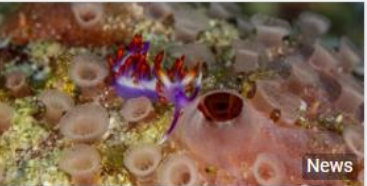
Project

New video: an invitation to the private sector



News

BID call for proposals: Caribbean 2020



News

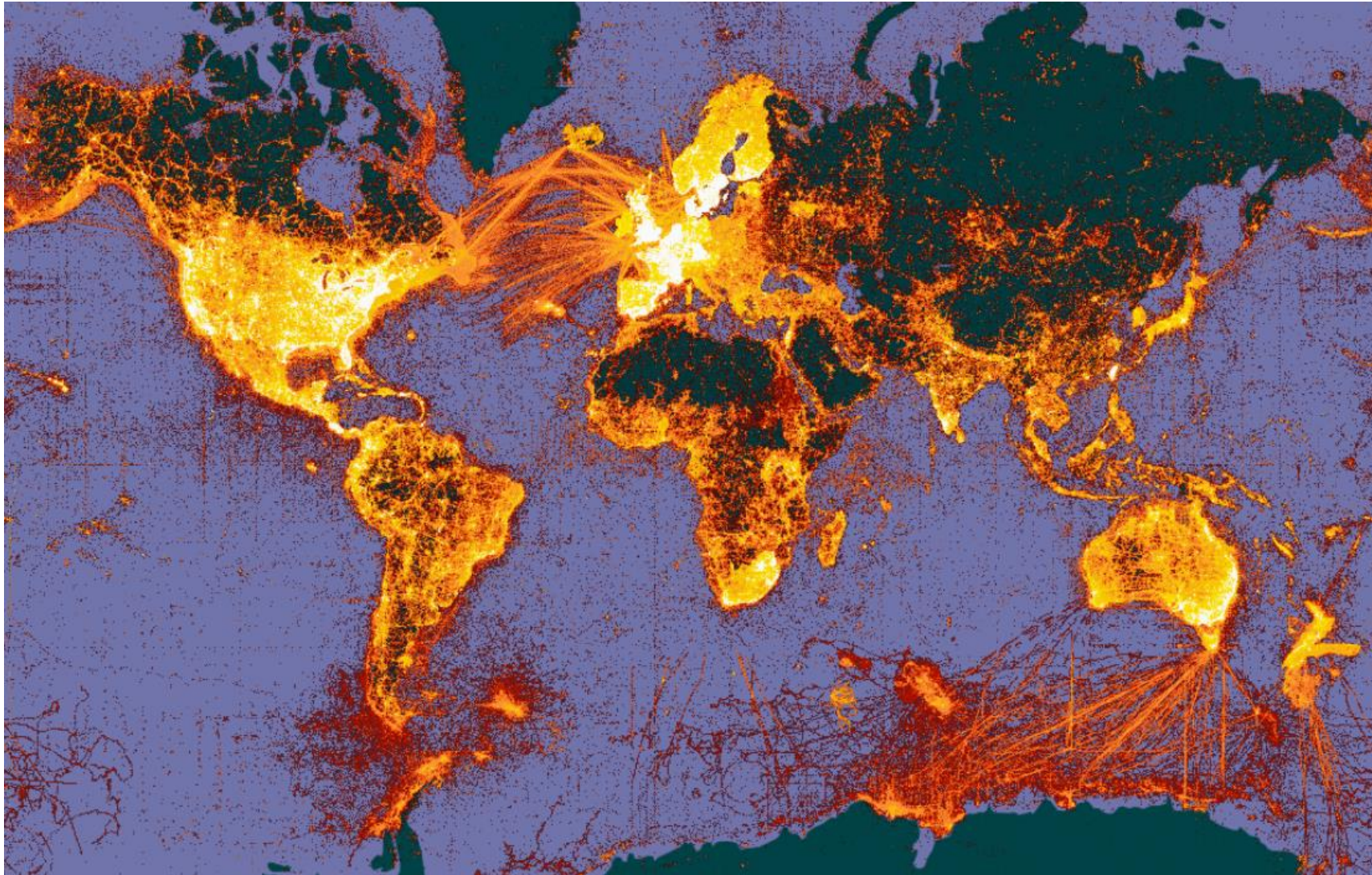
BID call for proposals: Pacific 2020

Global Biodiversity Information Facility

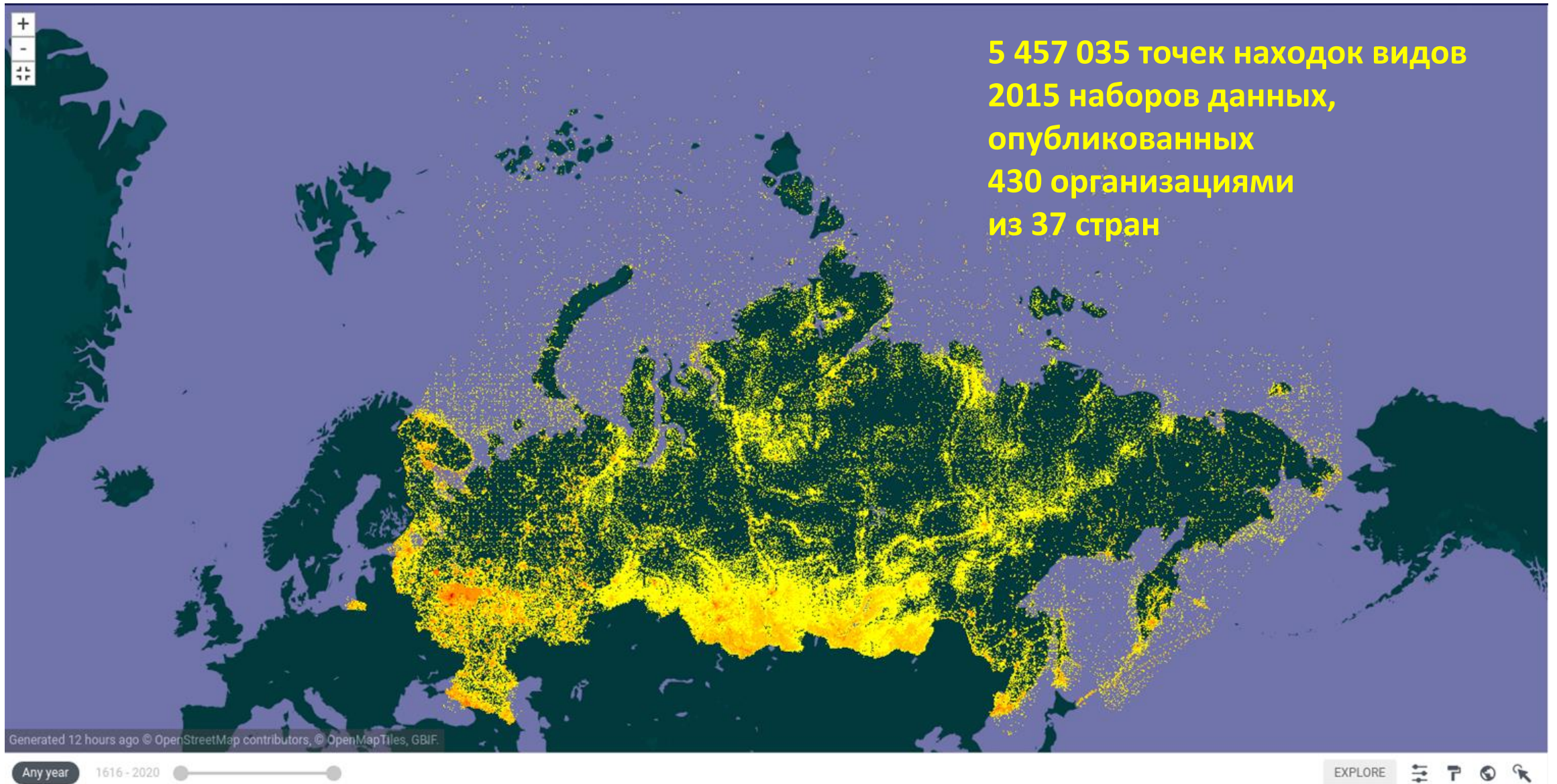
Глобальная информационная система о биоразнообразии

- Международная инфраструктура открытых данных
- Финансируется правительствами стран-участниц
- **Сеть для открытого и свободного доступа к объединенным данным о биоразнообразии**
- 100 официальных участников: 62 страны, 38 международных организаций, 1668 зарегистрированных организаций, публикующих данные

Карта глобального биоразнообразия GBIF



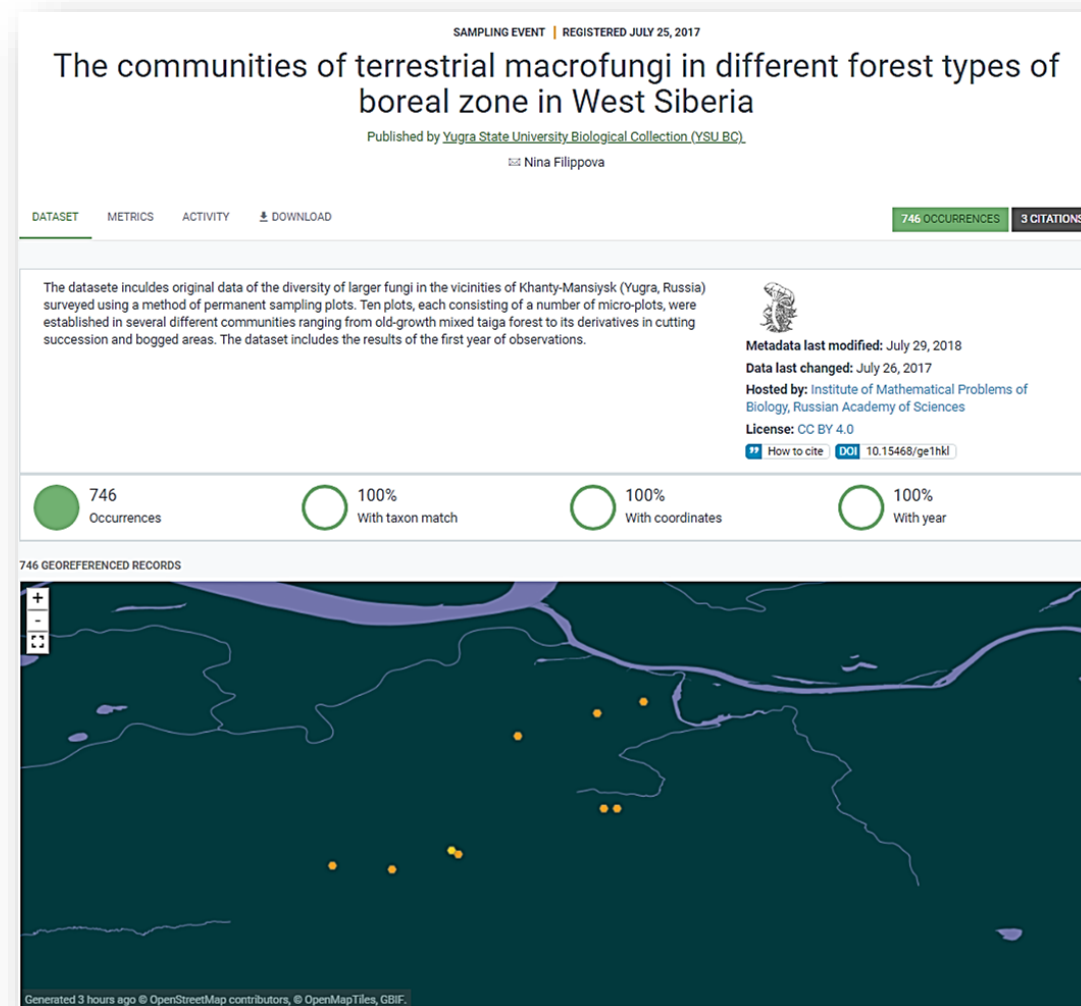
GBIF - крупнейший ресурс о биоразнообразии России



ЗАЧЕМ публиковать свои данные через глобальные порталы?



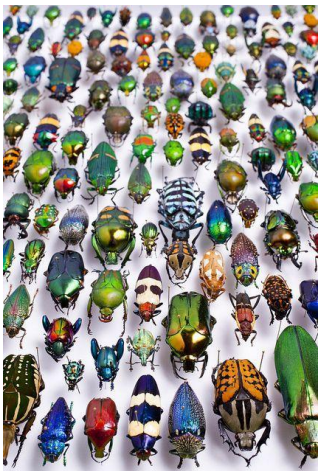
или



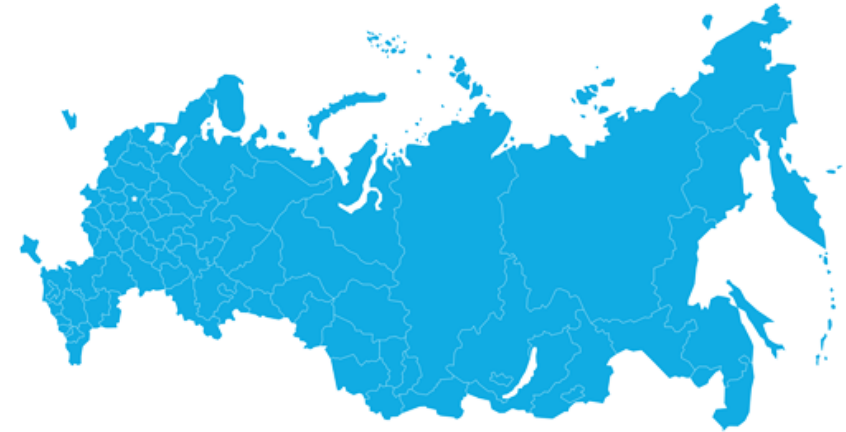
ЗАЧЕМ публиковать свои данные через глобальные порталы?



- Проверяемость результатов научных исследований
- Закрепление авторства собственных данных через DOI
- Повышение обнаружимости данных и исследований
- Возможности совместных исследований



- Повышение научного использования коллекций
- Лучшая сохранность фондовых образцов
- Систематизация и улучшение качества данных



- Выполнение обязательств в рамках международных Конвенций.
- Достижение Целей в области устойчивого развития, в том числе связанных с здоровьем человека, продовольственной безопасностью и глобальным изменением климата, а также сохранением биоразнообразия на суше и в океане.



Данные о коллекционных образцах



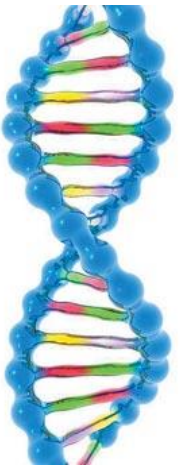
Данные об окаменелостях



Литературные данные



Биоматериалы



Наблюдения



Данные автоматических регистраторов



Данные площадных и маршрутных учетов

Какие данные доступны через GBIF

Типы наборов данных, доступных через GBIF

Метаданные (METADATA only)

+

Таксономический список CHECKLIST

для публикации
таксономических данных:
списки охраняемых
видов, тематические
видовые списки и др.

1 строка – 1 таксон



+

Находки OCCURRENCE DATA

Находка - простое полевое
наблюдение или
коллекционный образец

1 строка = 1 особь или 1 группа
особей

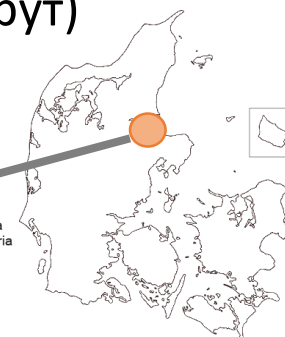
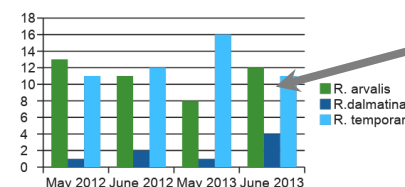


+

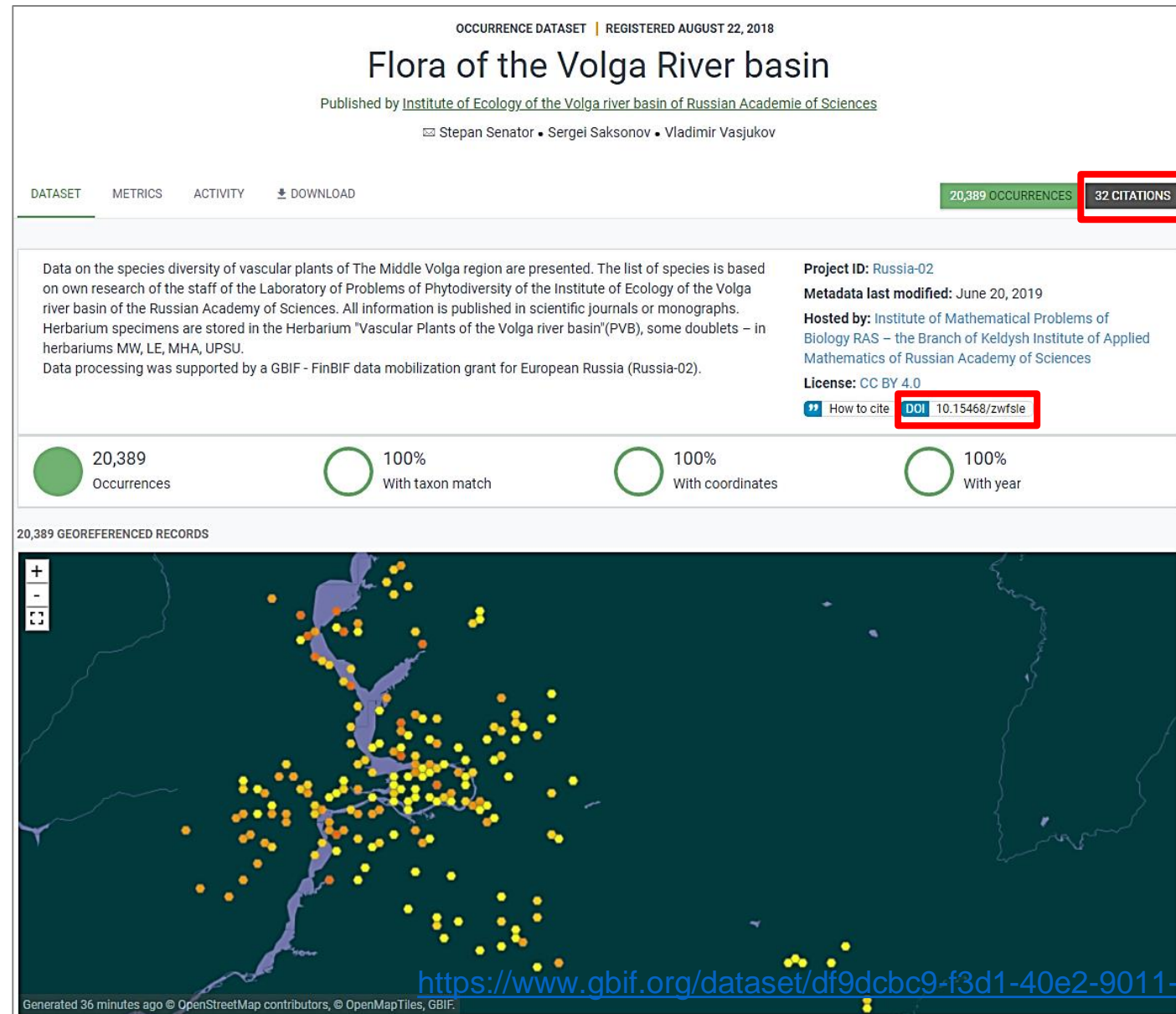
Описания, учеты или сборы SAMPLE EVENT DATA

2 таблицы: данные о сборе +
данные о находках

1 запись на листе event =
1 событие (площадка,
маршрут)



Набор данных (dataset), доступный через GBIF



[Get data](#)
[How-to](#)
[Tools](#)
[Community](#)
[About](#)

dryomys

Resources

3

Search

Country or area of researcher

Country or area of coverage

Literature type

Journal article

Relevance

GBIF used

Year

Topic

Agriculture

253

Biodiversity science

144

Biogeography

356

Citizen science

30

Climate change

784

Conservation

654

Data management

54

Data paper

31

Ecology

1 206

Ecosystem services

40

Evolution

717

Freshwater

31

Human health

181

Invasives

737

Marine

61

Phylogenetics

260

Species distributions

200

Taxonomy

318

Dataset

Publisher

SEARCH RESOURCES | 5,101 RESULTS

ALL LITERATURE

Standardized data to support conservation prioritization for sharks and batoids (Elasmobranchii)

Literature

Nielsen, R. da Silva, R. Juergens, J. Staerk, J. Sorensen, L. Jackson, J. ... - (2020) Data in Brief
#N/A

Aquariums • Aquatic management • CITES • Captive husbandry • International treaties • Life history traits

Journal article Open access Peer-reviewed

Data referenced in study
DOI 10.15468/dl.54oesv DOI 10.15468/dl.59ncfb DOI 10.15468/dl.5uuncg DOI 10.15468/dl.5yut6b DOI 10.15468/dl.8phc9c DOI 10.15468/dl.kfr6u1 DOI 10.15468/dl.kqmg3 DOI 10.15468/dl.llxcxj DOI 10.15468/dl.obnvcs DOI 10.15468/dl.rwwqee DOI 10.15468/dl.slwae DOI 10.15468/dl.un0pej

Trophic behavior of specialist predators from a macroecological approach: the case of the Magellanic woodpecker in South American temperate forests

Literature

Alaniz, A. Carvajal, M. Vergara, P. Fierro, A. Moreira-Arce, D. Rojas-Osorio, A. ... - (2020) Global Ecology and Conservation
Theoretical and empirical studies have addressed predator-prey relationships, but the relevance of individual prey species, prey species richness and environment on the niche and the distribution of specialist predators have been scarcely tested with a macroecological approach. Here we aim to analyz...

Journal article Open access Peer-reviewed

Data referenced in study
DOI 10.15468/dl.71qpiy DOI 10.15468/dl.b7jlr0 DOI 10.15468/dl.cgazdt DOI 10.15468/dl.uevwgh DOI 10.15468/dl.vllkvz DOI 10.15468/dl.zpbeqk

A Disjunctive Marginal Edge of Evergreen Broad-Leaved Oak (Quercus gilva) in East Asia: The High Genetic Distinctiveness and Unusual Diversity of Jeju...

Literature

Han, E. Cho, W. Park, J. Choi, I. Kwak, M. Kim, B. ... - (2020) Genes
Jeju Island is located at a marginal edge of the distributional range of East Asian evergreen broad-leaved forests. The low genetic diversity of such edge populations is predicted to have resulted from genetic drift and reduced gene flow when compared to core populations. To test this hypothesis, we...

Jeju Island • Quercus gilva • conservation • genetic diversity • marginal edge • massive colonization

Journal article Open access Peer-reviewed

Modelled distributions and conservation priorities of wild sorghums (Sorghum Moench)

Literature

Myrans, H. Diaz, M. Khoury, C. Carver, D. Henry, R. Gleadow, R. (2020) Diversity and Distributions
Aim: To fill knowledge gaps regarding the distributions, ecogeographic niches and conservation status of sorghum's wild relatives (Sorghum Moench). Location: The study covered the potential native ranges of wild Sorghum taxa worldwide, including Australia, New

dryomys

Zone-tailed hawk (*Buteo albonotatus*) observed in Chaichuapa, El Salvador by Carlos E. Juárez-Peña. Photo via iNaturalist (CC BY-NC 4.0)

Peer-reviewed papers using data

5 101

ibbean

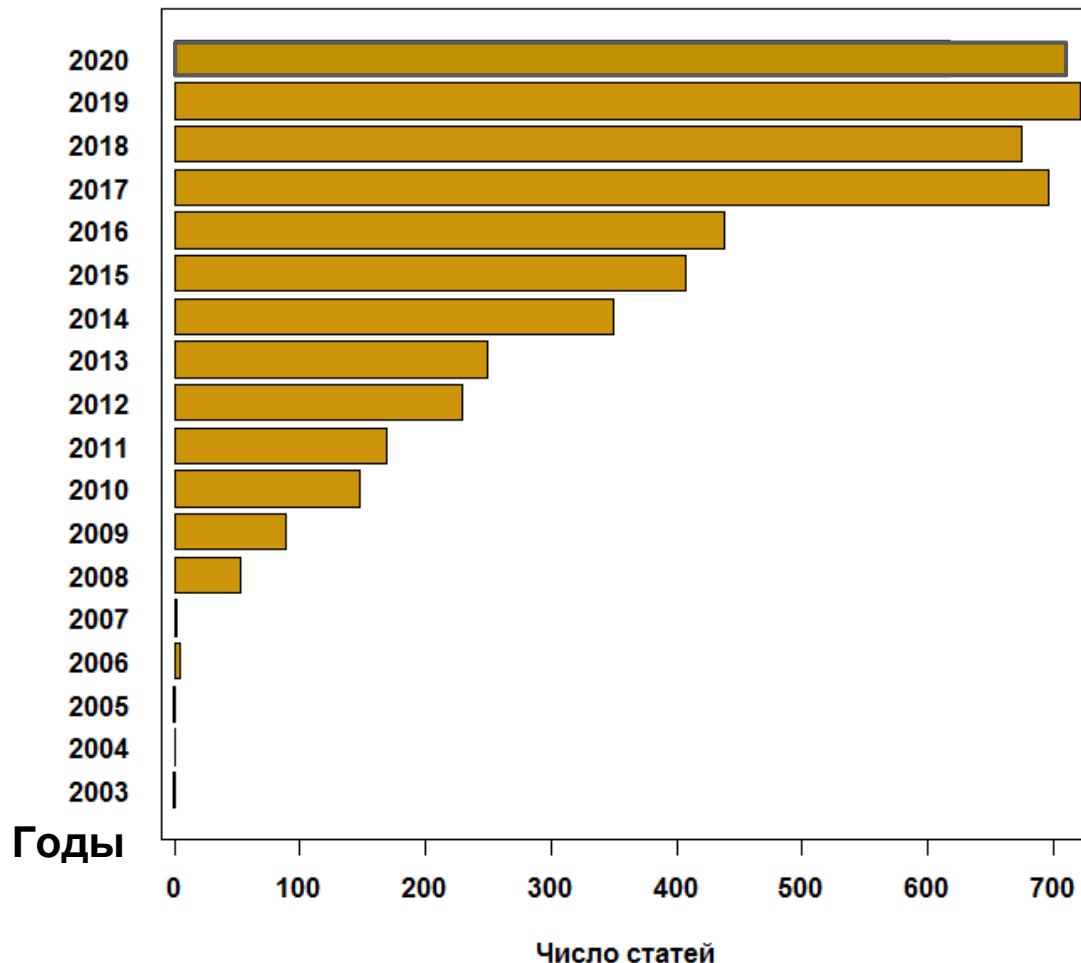
BID call for proposals: Pacific 2020

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

15

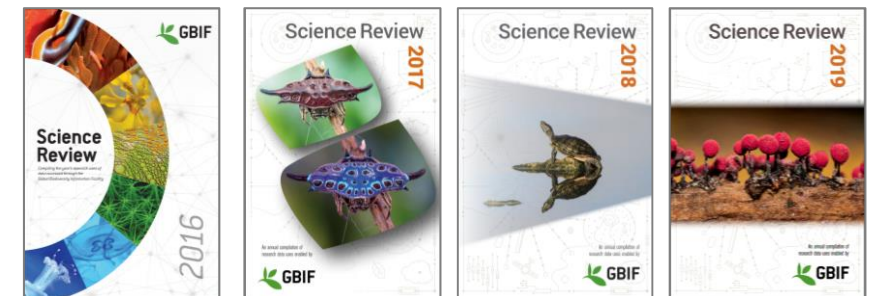
Использование данных, доступных через GBIF, в научных исследованиях

Число научных статей, основанных на анализе данных GBIF, опубликованных в рецензируемых журналах



Основная сфера использования данных GBIF в экологии – моделирование распространения видов

- Оценка глобальных изменений климата на распространение видов
- Факторы, влияющие на расселение инвазивных видов
- Прогнозы изменения ареалов переносчиков опасных для человека заболеваний
- и т.д.



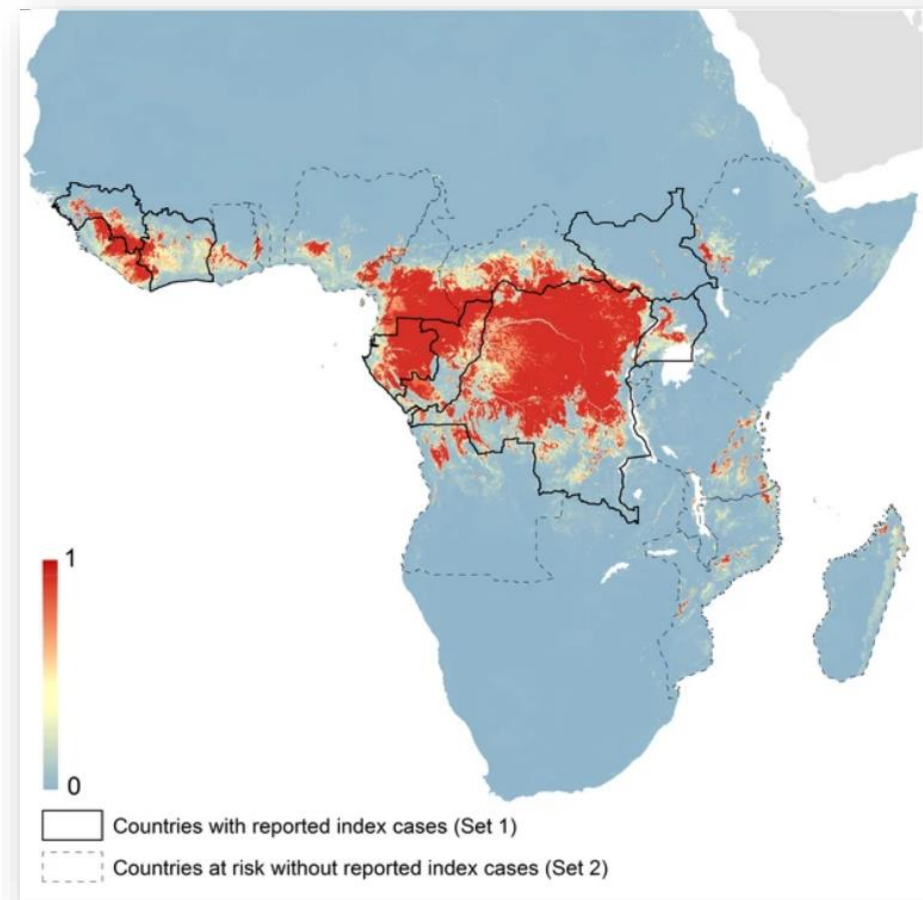
<https://www.gbif.org/science-review>

Mapping the zoonotic niche of Ebola virus disease in Africa



(c) jakob, some rights reserved (CC BY-NC)

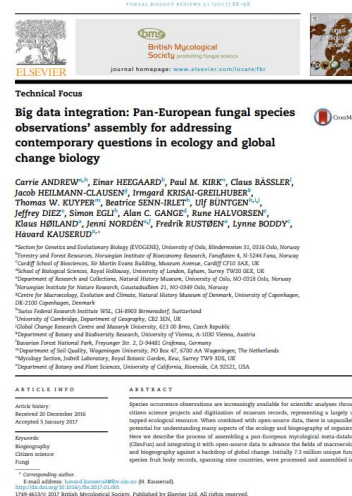
- Смоделированы экологические ниши трех видов летучих мышей, связанных с передачей вируса Эбола
- Данные получены из глобальной базы по биоразнообразию GBIF
- Результаты показали, что зоны риска охватывают 22 страны, население 22 миллиона чел.
- Результаты помогают расставить приоритеты в области эпидем. надзора и диагностики в зонах повышенного риска



**Распространение вируса Эбола,
предсказанное моделью**

Big data integration: Pan-European fungal species observations assembly for addressing contemporary questions in ecology and global change biology

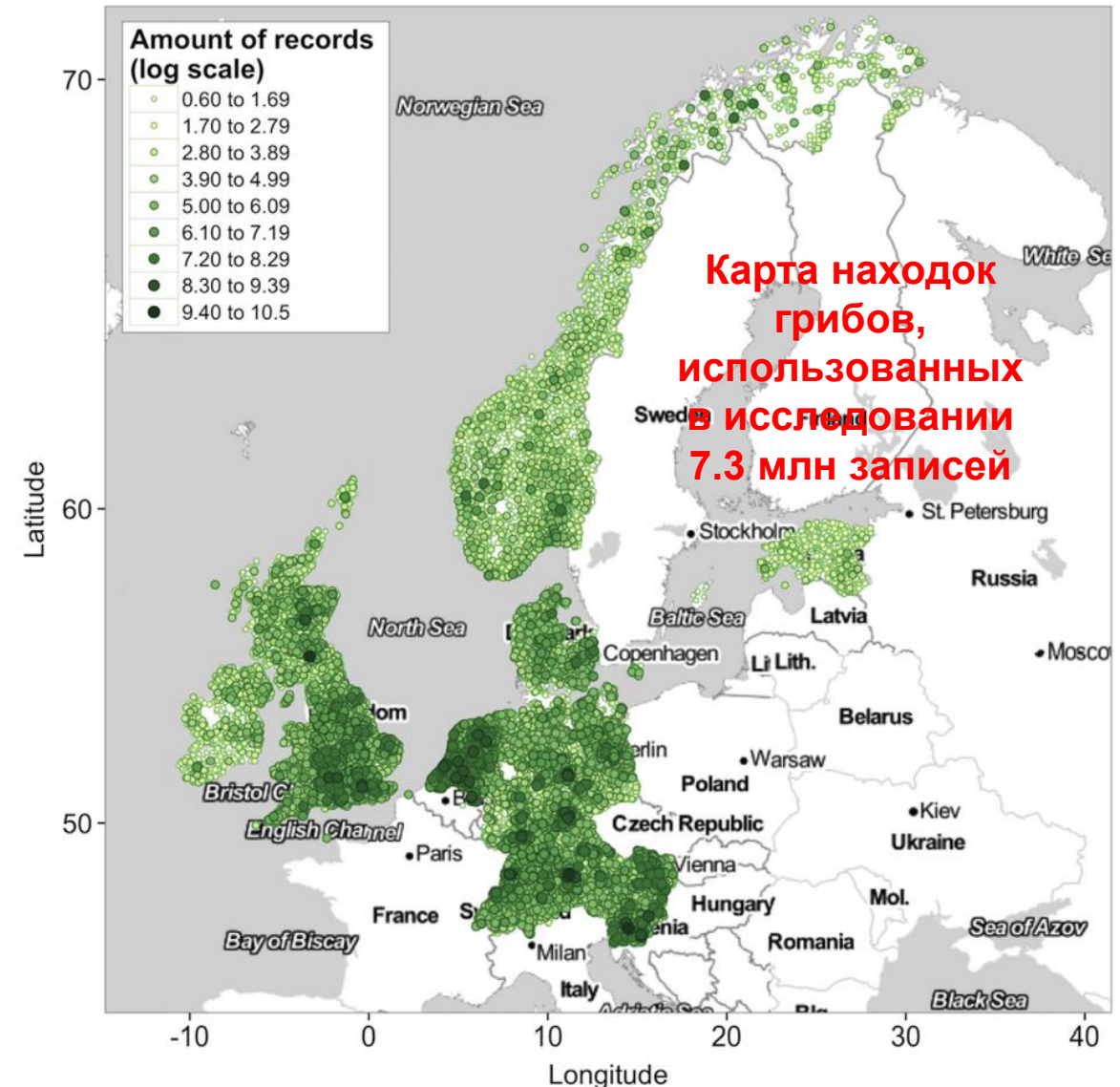
- Таксономический анализ
- Анализ сходства сообществ грибов в разных регионах
- Фенология появления плодовых тел
- Научный потенциал собранного набора данных
- Анализ данных с позиций глобальных изменений климата



Andrew et al.

Fungal Biology Reviews, 2017

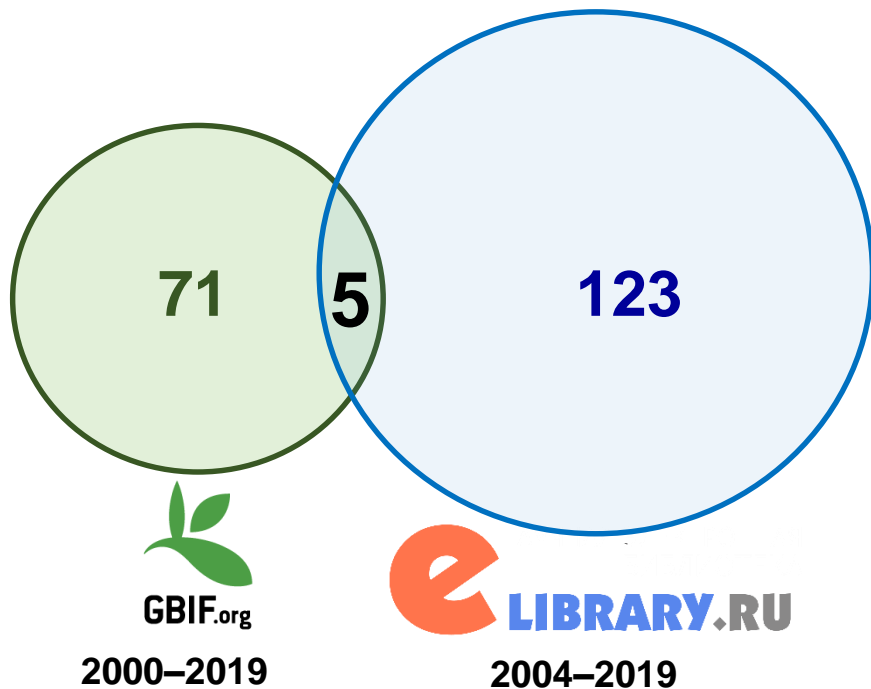
<https://doi.org/10.1016/j.fbr.2017.01.001>



Карта находок грибов, использованных в исследовании 7.3 млн записей

Использование данных, доступных через GBIF, российскими исследователями

Число статей, опубликованных
с участием российских
соавторов



На 9 ноября 2020 г.

- Всего статей в рецензируемых журналах: 5 101
- Из них с участием соавторов из России: 98
- Исследований, выполненных на территории РФ: 13

Несмотря на то, что для территории РФ через GBIF доступно >5.4 млн. записей о находках видов, репрезентативность информации для большинства таксономических групп и регионов страны остается низкой (Иванова, Шашков, 2021).

Плотность записей о находках сосудистых растений (а) и птиц (б) в России по данным GBIF



Кто публикует данные о биоразнообразии России в GBIF

Всего 430 организаций из 37 стран, первые 10 по числу опубликованных записей:

Publisher	Count	
☞ Lomonosov Moscow State University	1,300,001	<div></div>
☞ iNaturalist.org	1,017,408	<div></div>
☞ Cornell Lab of Ornithology	282,265	<div></div>
☞ Central Siberian Botanical Garden SB RAS	206,271	<div></div>
☞ Finnish Biodiversity Information Facility	195,696	<div></div>
☞ European Nucleotide Archive (EMBL-EBI)	192,397	<div></div>
☞ PlutoF	166,091	<div></div>
☞ Conservation of Arctic Flora and Fauna	122,753	<div></div>
☞ Sibecocenter LLC	117,344	<div></div>
☞ I.I. Sprygin Zhiguli State Nature Biosphere Reserv...	86,524	<div></div>



Свободный и открытый доступ к данным о биоразнообразии

Цифровые данные о биоразнообразии России

Информационные системы о распространении видов
Оцифрованные коллекции
Таксономические базы данных
Библиография

Информатика биоразнообразия: основные концепции

Методы оценки качества данных о биоразнообразии

Руководства

Оцифровка Летописей природы
Подготовка данных
Публикация данных через IPT
Создание проекта в iNaturalist

Веб-инструменты

Проверка таксономии по ThePlantList
Стандартизация данных о почвенных беспозвоночных

Контакты

Сообщество GBIF.ru
Техническая поддержка публикации данных
Администраторы сайта
Новостная рассылка

О GBIF на русском

Что такое GBIF
Как опубликовать данные в GBIF
Спецификация Darwin Core
Данные о России в GBIF
Зачем открывать свои данные
Использование данных GBIF

Конференции и семинары

Предстоящие
Прошедшие
Workshop, лекции
Презентации (разные)

Добро пожаловать!

Вы находитесь на сайте, посвященном открытым данным о биоразнообразии. Крупнейшим источником таких данных в мире является Глобальная информационная система о биоразнообразии GBIF.org – независимая межправительственная организация и растущее международное сообщество, которое обеспечивает свободный доступ к данным о распространении видов. Все данные публикуются в GBIF с сохранением авторства и распространяются по открытым лицензиям Creative Commons.

Россия пока не является официальным членом GBIF, но объем российских данных, доступных через GBIF.org, увеличивается с каждым годом. На нашем сайте вы найдете базовые сведения и русскоязычные материалы об открытых данных о биоразнообразии и их роли в современной науке, а также о том, как опубликовать свои данные в сети GBIF и правильно использовать данные других исследователей.

Новые материалы на сайте

2020-04-28 Материалы вебинара о публикации статей о данных
2020-04-26 Новый раздел Использование данных GBIF в научных исследованиях
2020-04-15 Новый веб-инструмент для проверки таксономии по ThePlantList
2020-04-13 Инструкция о том как правильно сохранить CSV файл
2019-12-18 Список статей о данных, опубликованных российскими исследователями
2019-12-14 Иван Чадин Что такое статья о данных и как ее опубликовать

Новостная рассылка

Подпишитесь на рассылку **GBIF на русском**, чтобы быть в курсе событий в России и у наших ближайших соседей.

Присоединившись к рассылке, вы также сможете делиться своими новостями с участниками сообщества через адрес gbif-in-russian@googlegroups.com

Сейчас в рассылке 343 подписчика. Подписаться

Российское Specify сообщество

О нас
О Specify на русском

iNaturalist.org

БУКЛЕТ
КАК СОЗДАТЬ ПРОЕКТ
Публикация наблюдений

Data paper: статья о данных

Что такое статья о данных?
new! Материалы вебинара

GBIF.ru – сайт российского GBIF сообщества

Техническая поддержка в публикации данных
E-mail: gbif.ru@yandex.ru

Информационная рассылка
gbif-in-russian@googlegroups.com
343 подписчика на ноябрь 2020
Группа GBIF.ru в Facebook



ХМАО-Югра в GBIF: 75,259 точек находок видов



**The future is already here —
it's just not very evenly distributed**

*Уильям Форд Гибсон,
писатель-фантаст*

Спасибо за внимание!

