

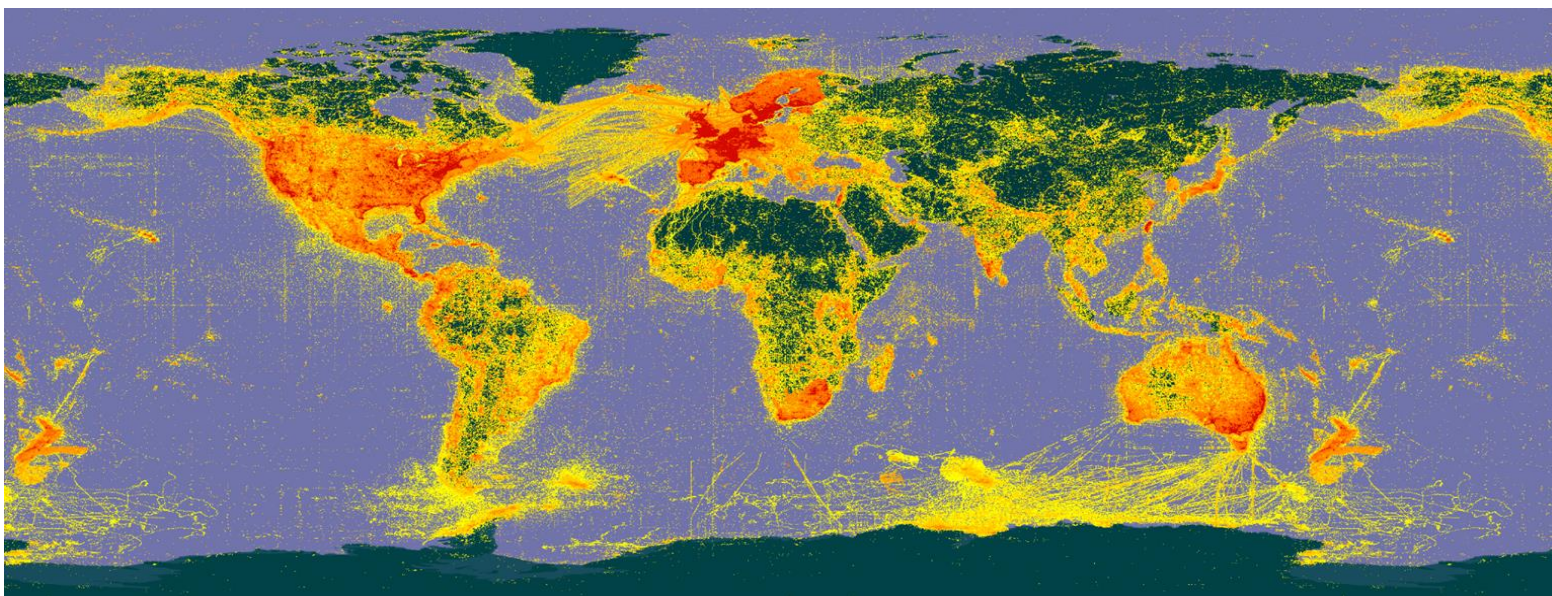
«Подготовка данных для публикации в Глобальной информационной системе о биоразнообразии GBIF»  
10 октября 2020 г., Екатеринбург

## Лекция 1.1

# Публикация данных в GBIF: основные этапы

**Наталья Иванова**

Институт математических проблем биологии РАН – филиал ИПМ им. М.В. Келдыша РАН



**Слайды CC BY:**

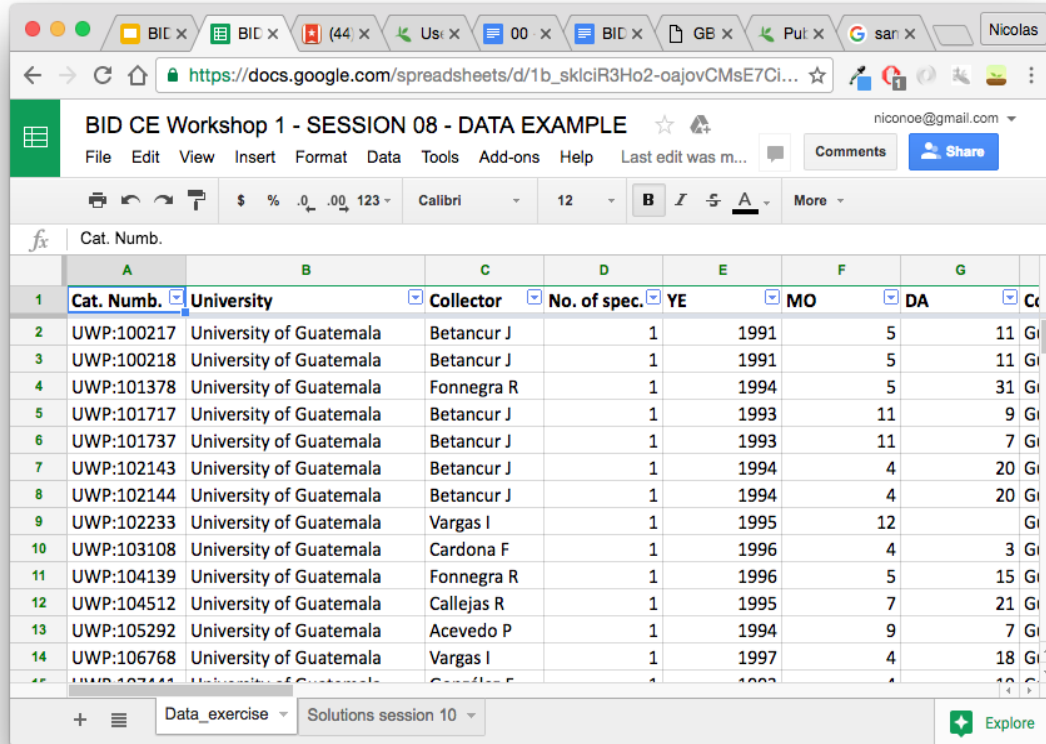
*Dag Endresen, GBIF Norway*

*Наталья Иванова,*

*Максим Шашков*

# Что такое публикация данных

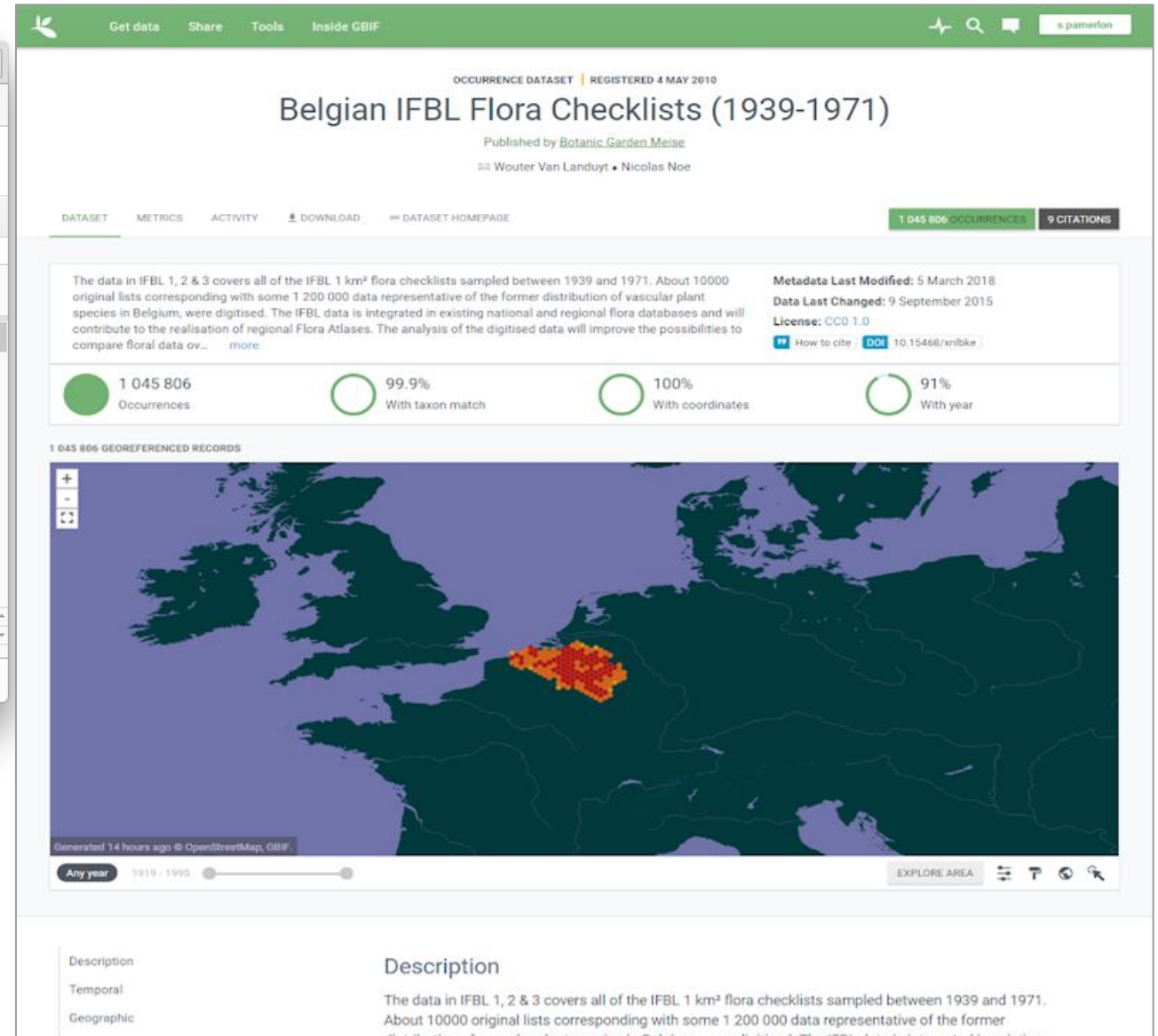
От электронной таблицы с данными к набору данных, доступному через глобальный портал.



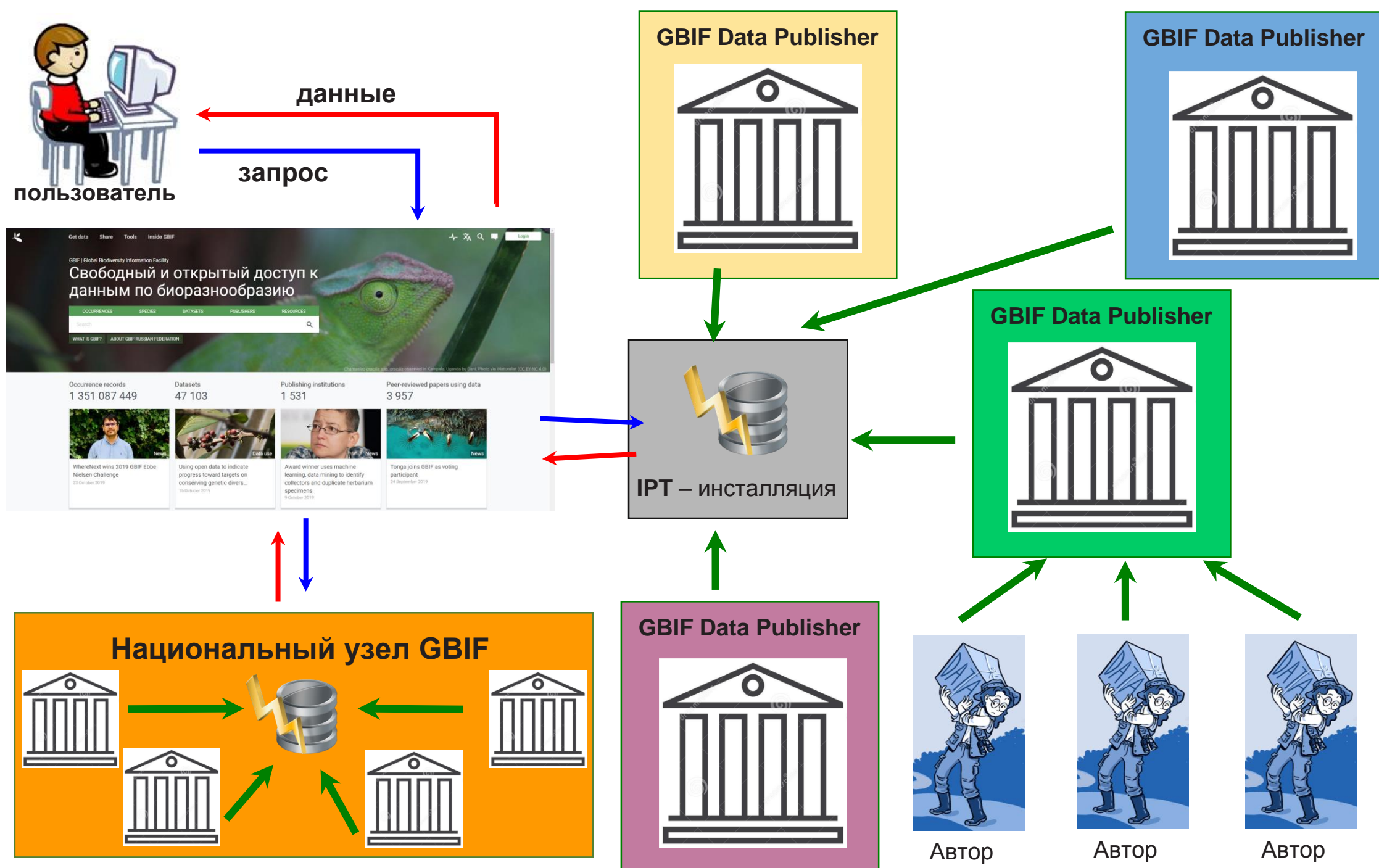
The screenshot shows a Google Sheet with the following data:

	A	B	C	D	E	F	G
1	Cat. Numb.	University	Collector	No. of spec.	YE	MO	DA
2	UWP:100217	University of Guatemala	Betancur J	1	1991	5	11
3	UWP:100218	University of Guatemala	Betancur J	1	1991	5	11
4	UWP:101378	University of Guatemala	Fonnegra R	1	1994	5	31
5	UWP:101717	University of Guatemala	Betancur J	1	1993	11	9
6	UWP:101737	University of Guatemala	Betancur J	1	1993	11	7
7	UWP:102143	University of Guatemala	Betancur J	1	1994	4	20
8	UWP:102144	University of Guatemala	Betancur J	1	1994	4	20
9	UWP:102233	University of Guatemala	Vargas I	1	1995	12	
10	UWP:103108	University of Guatemala	Cardona F	1	1996	4	3
11	UWP:104139	University of Guatemala	Fonnegra R	1	1996	5	15
12	UWP:104512	University of Guatemala	Callejas R	1	1995	7	21
13	UWP:105292	University of Guatemala	Acevedo P	1	1994	9	7
14	UWP:106768	University of Guatemala	Vargas I	1	1997	4	18

Программа IPT



# Как устроена сеть GBIF





## Как устроена сеть GBIF

GBIF это не одна база данных, а сеть национальных “узлов” (nodes), которые объединяют организации (научные институты, университеты, ботанические сады, музеи, заповедники и др.), публикующие данные. Национальные узлы взаимодействуют друг с другом и с Секретариатом GBIF по обмену опытом, навыками и техническим потенциалом. В странах, которые не являются официальными участниками GBIF функции национальных узлов неформально выполняют организации, предоставляющие хостинг для всех, желающих публиковать данные.

## Информационный сайт о Глобальной Системе о Биоразнообразии



Свободный и открытый доступ к данным о биоразнообразии

### Цифровые данные о биоразнообразии России

Информационные системы о распространении видов  
Оцифрованные коллекции  
Таксономические базы данных  
Библиография

### О GBIF на русском

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

### Конференции и семинары

Предстоящие  
Прошедшие  
Biodiversity\_next  
Презентации (разные)

### Информатика биоразнообразия: основные концепции

Методы оценки качества данных о биоразнообразии

### Руководства

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

### Веб-инструменты

### 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

2020-04-15 Новый веб-инструмент для проверки таксономии по ThePlantList

# GBIF.ru – сайт российского GBIF сообщества

Техническая поддержка в  
публикации данных  
E-mail: [gbif.ru@yandex.ru](mailto:gbif.ru@yandex.ru)

Информационная рассылка  
[gbif-in-russian@googlegroups.com](mailto:gbif-in-russian@googlegroups.com)  
344 подписчика на сентябрь 2020



# Publishers - организации, публикующие данные в сети GBIF

PUBLISHER | SINCE APRIL 4, 2018

Federal State Autonomous Educational Institution of Higher Education  
«Ural Federal University named after the first President of Russia  
B.N.Yeltsin»

ABOUTMETRICSHOME PAGE

29,106 OCCURRENCES5 DATASETS11 CITATIONS


**Description:** The Botanical Museum of the Department of Biodiversity and Bioecology was created in 1948 by Ph.D. V.P. Golubintseva. The main part of the herbarium consists of materials collected during the expedition work of the employees of the department, students at the Ural. There are materials collected by P.N. Krylov, O.E. Claire, B.P. Kolesnikov, A.A. Yelenkin. Herbarium consists of two parts: 1) stock herbarium, included collections characterizing the biological diversity of the plant world of the Ural; 2) educational herbarium used in the educational process. Currently, in the Botanical Museum (the international acronym UFU) there are about 30 thousand herbarium sheets and about 2000 species. Now one of priorities in the activity of the Botanical Museum is the digitization of herbarium funds.

**Endorsed by:** [Participant Node Managers Committee](#)

**Administrative contact:** [Victor Mukhin](#)

**Technical contact:** [Pavel Kondratkov](#)

**Country or area:** [Russian Federation](#)



© OpenStreetMap contributors

CONTACTS

Federal State Autonomous Educational Institution of Higher Education «Ural Federal University named after the first President of Russia B.N.Yeltsin»

Pavel Kondratkov

Technical point of contact

[pavel.kondratkov@urfu.ru](mailto:pavel.kondratkov@urfu.ru)

+7(912)2876059

Alyona Tretyakova

Point of contact

[alyona.tretyakova@urfu.ru](mailto:alyona.tretyakova@urfu.ru)

+7(912)6576504

Публикация данных в GBIF.org возможна вне зависимости от уровня официального участия страны.

Данные в сети GBIF публикуются только от имени организаций. Для этого организация должна быть зарегистрирована и одобрена участниками GBIF.

<https://www.gbif.org/become-a-publisher>

# Шаг 1. Регистрация новой организации в GBIF

Tools Community About

SHARE

## Become a publisher

*Organizations wishing to share data through GBIF can register here to request endorsement as a data publisher*

To prepare for publishing data through GBIF, we ask prospective new publishers to complete the online form below.

Your answers will help us to give proper credit and attribution for the datasets you share. They will also help users to understand more about the provenance of data shared through the GBIF network.

Before GBIF indexes your datasets, your institution must **receive endorsement as a data publisher** from one of the Participant nodes that coordinate activities of the **national and organizational participants in the GBIF network**. If your country is not yet participating in GBIF, we will seek endorsement on your behalf from the GBIF community.

<https://www.gbif.org/become-a-publisher>

[Инструкция на русском языке](#)

PUBLISHER | SINCE MAY 29, 2019

## Institute of Plant and Animal Ecology (IPAE)

ABOUT METRICS ∞ HOME PAGE

5,973 OCCURRENCES 2 DATASETS 6 CITATIONS


**Description:** The Institute of Plant and Animal Ecology of the Ural Branch of the Russian Academy of Sciences (IPAE UB RAS) is the largest research institute of the ecological and biological profile in the Ural region, one of the leaders in basic research in the field of ecology and environmental protection in Russia. The biological collections of the Museum of the IPAE contain over a 1.5 million specimens, including 700.000 fossil animal remains, 200.000 recent animals samples, 300.000 insects and 300.000 plant specimens (Herbarium SVER).

**Endorsed by:** Participant Node Managers Committee

**Administrative contact:** Denis Veselkin

**Technical contact:** Alex Vorobiev

**Country or area:** Russian Federation



**CONTACTS**

Institute of Plant and Animal Ecology (IPAE) 8 Marta str., 202 Yekaterinburg 620144 Sverdlovskaya Russian Federation <a href="mailto:common@ipae.uran.ru">common@ipae.uran.ru</a> +73432103853 <a href="https://ipae.uran.ru/">https://ipae.uran.ru/</a>	Alex Vorobiev Technical point of contact <a href="mailto:eng@ipae.uran.ru">eng@ipae.uran.ru</a> +73432103853+106	Nickolay Erokhin Point of contact <a href="mailto:eng@ipae.uran.ru">eng@ipae.uran.ru</a> +73432103853+106
Denis Veselkin Administrative point of contact <a href="mailto:veselkin_dv@ipae.uran.ru">veselkin_dv@ipae.uran.ru</a> +73432666484		

<https://www.gbif.org/publisher/23d95083-5f92-48f2-9556-27a5b0f2400d>



## Шаг 2. Выбор необходимого типа набора данных

### 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

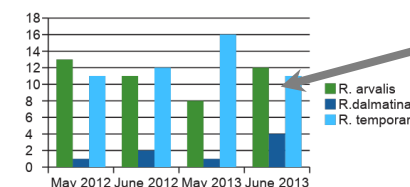


+

#### Описания, учеты или сборы SAMPLE EVENT DATA

2 таблицы: данные о сборе  
+ данные о находках

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





## Шаг 3. Стандартизация набора данных


### Darwin Core

Darwin Core is a standard maintained by the [Darwin Core maintenance group](#). It includes a glossary of terms intended to **facilitate the sharing of information about biological diversity** by providing identifiers, labels, and definitions. Darwin Core is primarily based on taxa, their occurrence in nature as documented by observations, specimens, samples, and related information.

technical specification

current standard

2009

[Darwin Core website](#) [Find us on GitHub](#) <http://rs.tdwg.org/dwc/terms/>

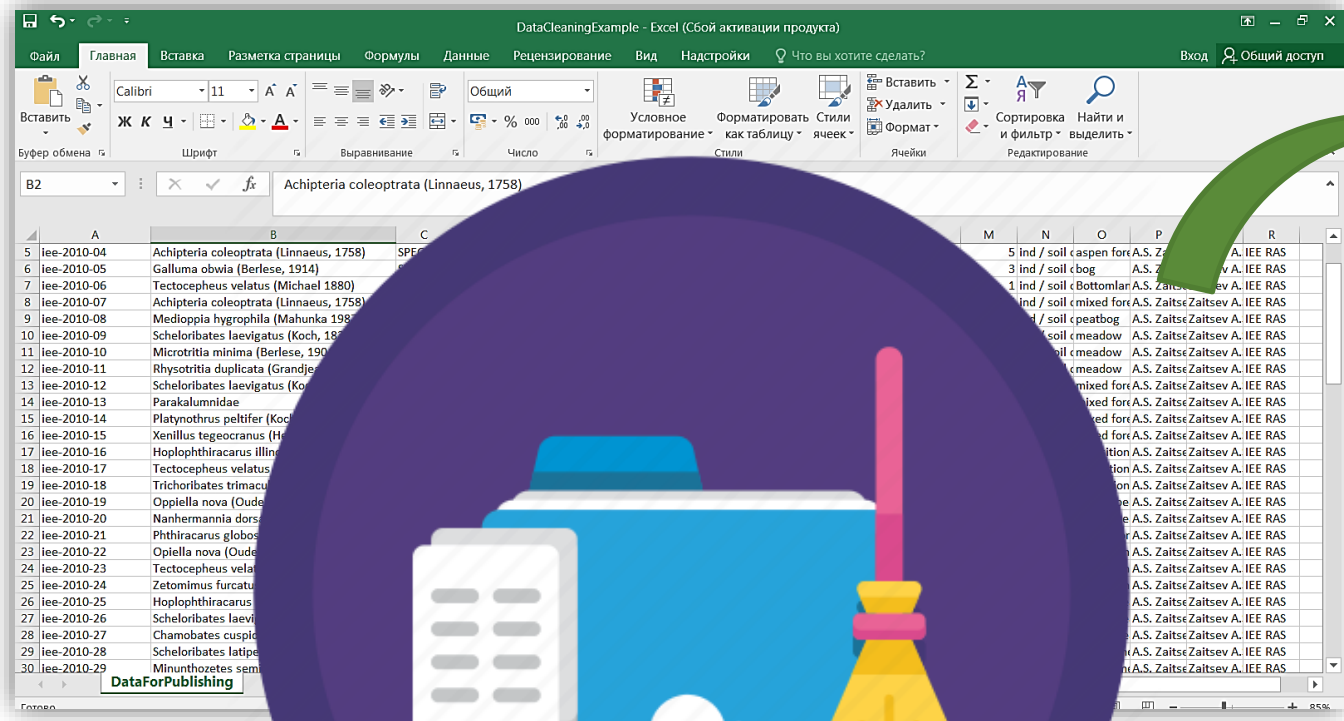
Система обмена информацией в GBIF построена на открытых стандартах, центральный из которых – **Darwin Core (DwC)**, стандарт, разработанный специально для хранения данных по биоразнообразию.

**Стандарт** – набор терминов (заголовки столбцов в таблице с данными) и правила использования этих терминов.

## Пример готовой для публикации в GBIF таблицы

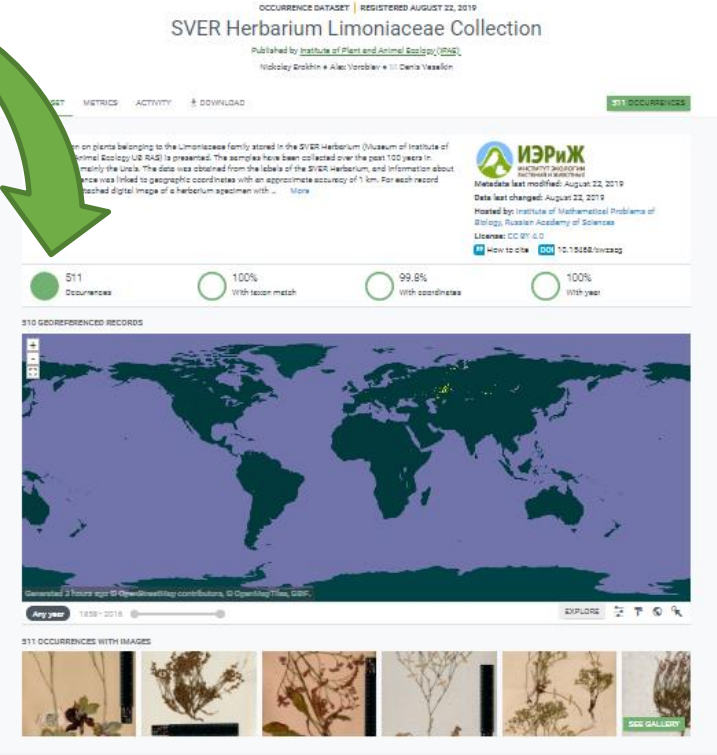
[illegible]

# Шаг 4. Проверка данных. Исправление ошибок и документация изменений.



The image shows a Microsoft Excel spreadsheet titled "DataCleaningExample - Excel (Сбой активации продукта)". The spreadsheet contains a list of botanical specimens with columns for date, species name, and collector. A large purple circular overlay with a broom and a checklist icon is positioned over the spreadsheet, symbolizing data cleaning. A green arrow points from this overlay to the right-hand image.

	A	B	C	M	N	O	P	R
5	lee-2010-04	Achipteria coleoptrata (Linnaeus, 1758)	SPE	5 ind / soil caspen for	A.S. Zaitsev A. IEE RAS			
6	lee-2010-05	Galluma obvia (Berlese, 1914)		3 ind / soil cbog	A.S. Zaitsev A. IEE RAS			
7	lee-2010-06	Tectocephus velatus (Michael 1880)		1 ind / soil c Bottomlan	A.S. Zaitsev A. IEE RAS			
8	lee-2010-07	Achipteria coleoptrata (Linnaeus, 1758)		1 ind / soil c mixed for	A.S. Zaitsev A. IEE RAS			
9	lee-2010-08	Medioplia hygrophila (Mahunka 198)		1 ind / soil c peatbog	A.S. Zaitsev A. IEE RAS			
10	lee-2010-09	Schelorbates laevigatus (Koch, 18)		1 ind / soil c meadow	A.S. Zaitsev A. IEE RAS			
11	lee-2010-10	Microtritia minima (Berlese, 190)		1 ind / soil c meadow	A.S. Zaitsev A. IEE RAS			
12	lee-2010-11	Rhysotritia duplicata (Grandje		1 ind / soil c meadow	A.S. Zaitsev A. IEE RAS			
13	lee-2010-12	Schelorbates laevigatus (Koc		1 ind / soil c mixed for	A.S. Zaitsev A. IEE RAS			
14	lee-2010-13	Parakalumnidae		1 ind / soil c mixed for	A.S. Zaitsev A. IEE RAS			
15	lee-2010-14	Platynocheilus peltifer (Koc		1 ind / soil c mixed for	A.S. Zaitsev A. IEE RAS			
16	lee-2010-15	Xenillus tegeocranus (H		1 ind / soil c mixed for	A.S. Zaitsev A. IEE RAS			
17	lee-2010-16	Hoplophthiracarus illino		1 ind / soil c mixed for	A.S. Zaitsev A. IEE RAS			
18	lee-2010-17	Tectocephus velatus		1 ind / soil c mixed for	A.S. Zaitsev A. IEE RAS			
19	lee-2010-18	Trichoribates trimacu		1 ind / soil c mixed for	A.S. Zaitsev A. IEE RAS			
20	lee-2010-19	Opiella nova (Oude		1 ind / soil c mixed for	A.S. Zaitsev A. IEE RAS			
21	lee-2010-20	Nanhermannia dors		1 ind / soil c mixed for	A.S. Zaitsev A. IEE RAS			
22	lee-2010-21	Phthiracarus globos		1 ind / soil c mixed for	A.S. Zaitsev A. IEE RAS			
23	lee-2010-22	Opiella nova (Oude		1 ind / soil c mixed for	A.S. Zaitsev A. IEE RAS			
24	lee-2010-23	Tectocephus vela		1 ind / soil c mixed for	A.S. Zaitsev A. IEE RAS			
25	lee-2010-24	Zetomimus furcatus		1 ind / soil c mixed for	A.S. Zaitsev A. IEE RAS			
26	lee-2010-25	Hoplophthiracarus		1 ind / soil c mixed for	A.S. Zaitsev A. IEE RAS			
27	lee-2010-26	Schelorbates laevi		1 ind / soil c mixed for	A.S. Zaitsev A. IEE RAS			
28	lee-2010-27	Chamobates cuspid		1 ind / soil c mixed for	A.S. Zaitsev A. IEE RAS			
29	lee-2010-28	Schelorbates latipe		1 ind / soil c mixed for	A.S. Zaitsev A. IEE RAS			
30	lee-2010-29	Minunthozetes sem		1 ind / soil c mixed for	A.S. Zaitsev A. IEE RAS			



The image shows the SVER Herbarium Limoniaceae Collection website. The page displays the collection name, publication information, and a world map showing the distribution of the collection. A green arrow points from the Excel spreadsheet to this website, indicating the flow of data from the spreadsheet to the online collection.

SVER Herbarium Limoniaceae Collection  
Published by [Institute of Plant and Animal Ecology \(RAS\)](#)  
Vladimir Golobin & Alex Vorobiev & Denis Vassilov

511 Occurrences

100% With better match

99.8% With coordinates

100% With year

510 GEOREFERENCED RECORDS

Any year 1838 - 2019

511 OCCURRENCES WITH IMAGES



## Шаг 5. Подготовка метаданных

OCCURRENCE DATASET | REGISTERED AUGUST 22, 2019

→ SVER Herbarium Limoniaceae Collection

Published by [Institute of Plant and Animal Ecology \(IPAE\)](#)

→ Nickolay Erokhin • Alex Vorobiev • ✉ Denis Veselkin

DATASET


METRICS

ACTIVITY

↓ DOWNLOAD

511 OCCURRENCES6 CITATIONS

Information on plants belonging to the Limoniaceae family stored in the SVER Herbarium (Museum of Institute of Plant and Animal Ecology UB RAS) is presented. The samples have been collected over the past 100 years in Europe and mainly the Urals. The data was obtained from the labels of the SVER Herbarium, and information about each occurrence was linked to geographic coordinates with an approximate accuracy of 1 km. For each record there is an attached digital image of a herbarium specimen with ... [More](#)



**ИЭРиЖ**  
ИНСТИТУТ ЭКОЛОГИИ  
РАСТЕНИЙ И ЖИВОТНЫХ


Metadata last modified: August 22, 2019


Hosted by: [Institute of Mathematical Problems of Biology RAS – the Branch of Keldysh Institute of Applied Mathematics of Russian Academy of Sciences](#)


→ License: CC BY 4.0


” How to cite

DOI 10.15468/xwzszg

511  
Occurrences

100%  
With taxon match

99.8%  
With coordinates

100%  
With year

12

# Лицензии Creative Commons

**CC-0**



**CC-BY**



**CC-BY-NC**



**Не продается!**

# Подготовка метаданных

511 OCCURRENCES WITH IMAGES



- Description
- Geographic scope
- Taxonomic scope
- Contacts
- Data description
- GBIF registration
- Citation

## Description

Information on plants belonging to the Limoniaceae family stored in the SVER Herbarium (Museum of Institute of Plant and Animal Ecology UB RAS) is presented. The samples have been collected over the past 100 years in Europe and mainly the Urals. The data was obtained from the labels of the SVER Herbarium, and information about each occurrence was linked to geographic coordinates with an approximate accuracy of 1 km. For each record there is an attached digital image of a herbarium specimen with label made in high resolution 600 dpi.

Представлена информация о растениях семейства Limoniaceae, хранящихся в гербарии SVER (Музей Института экологии растений и животных УрО РАН). Эти образцы были собраны в Европе и, главным образом, на Урале в течение последних 100 лет. Данные были получены из гербарных этикеток, и информация о каждом образце была привязана к географическим координатам с приблизительной точностью 1 км. К каждой записи прилагается цифровое изображение гербарного образца вместе с этикеткой, выполненное в высоком разрешении 600 dpi.



## Шаг 6. Публикация набора данных в GBIF

### **IPT: The Integrated Publishing Toolkit**

Непосредственно процедура публикации данных в GBIF осуществляется с помощью специального программного обеспечения IPT. Все опубликованные данные хранятся непосредственно на сервере с установленным IPT, на глобальном портале размещаются только метаданные. Вы можете установить IPT на собственном сервере, или получить учетную запись в одном из [существующих IPT](#), связавшись с ее администратором. В любом случае данные будут опубликованы от имени вашей организации. Опубликованные данные в любое время можно будет обновить или отозвать.

Nickolay Erokhin • Alex Vorobiev • ✉ Denis Veselkin

511 OCCURRENCES

**ИЭРиЖ**  
ИНСТИТУТ ЭКОЛОГИИ  
РАСТЕНИЙ И ЖИВОТНЫХ

**Hosted by:** Institute of Mathematical Problems of Biology RAS – the Branch of Keldysh Institute of Applied Mathematics of Russian Academy of Sciences

How to cite DOI 10.15468/xwzszg

100%  
With fa

99.8%  
With coordinates

100%  
With year

# IPT

## Таблица с данными в Darwin Core

## Описание набора данных (метаданные)



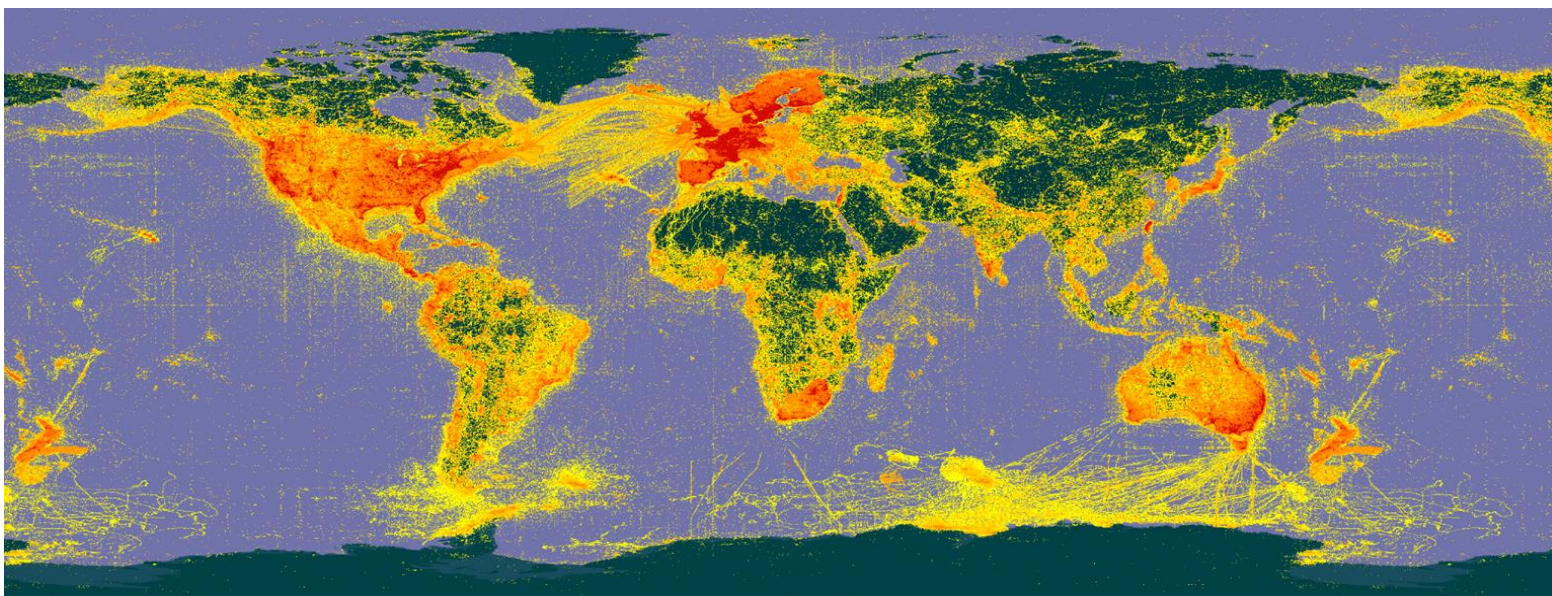
«Подготовка данных для публикации в Глобальной информационной системе о биоразнообразии GBIF»  
10 октября 2020 г., Екатеринбург

## Лекция 1.1

# Публикация данных в GBIF: основные этапы

**Наталья Иванова**

Институт математических проблем биологии РАН – филиал ИПМ им. М.В. Келдыша РАН



**Слайды CC BY:**

*Dag Endresen, GBIF Norway*

*Наталья Иванова,*

*Максим Шашков*