



Гербарийді цифрлау және GBIF IPT арқылы деректерді жариялау
Оцифровка гербария и публикация данных через GBIF IPT
Herbarium data digitization and publishing through GBIF IPT

Оцифровка научных биологических коллекций

Наталья Иванова

Институт математических проблем биологии РАН -
филиал ИПМ им. М.В. Келдыша РАН, Пущино, Россия

Сколько образцов хранится в научных биологических коллекциях мира?

В 2018 году в мире функционировало **3 095** гербариев, общее число образцов - **387 000 000** (Thiers, 2018). Эта оценка получена на основе [Index Herbariorum](#), т.е. включает только коллекции, у которых есть акроним.



Сколько образцов хранится в научных биологических коллекциях мира?



ВИКИПЕДИЯ
Свободная энциклопедия

[Список гербариев России](#)

110 зарегистрированы в Index Herbariorum

65 не зарегистрированы в Index Herbariorum, фонды >5000 листов

Коллекции с фондом <5000 листов - ?

Herbarium Code	Institution	Location
<u>UKSPI</u>	S. Amanjolov East-Kazakhstan State University	Kazakhstan. Ust-Kamenogorsk.
<u>PPIU</u>	M. Utemisov Western Kazakhstanian State University	Kazakhstan. Uralsk.
<u>AA</u>	Ministry of Science, Academy of Sciences	Kazakhstan. Alma-Ata.
<u>KG</u>	International Phytochemistry Research and Production Institute	Kazakhstan. Karaganda.
<u>KSPI</u>	Kostanay State Pedagogical Institute	Kazakhstan. Kostanay.
<u>NUR</u>	Astana Botanical Garden	Kazakhstan. Nur-Sultan.
<u>QAR</u>	Buketov University	Kazakhstan. Karaganda Region. Karaganda.
<u>MANG</u>	Mangyshlak experimental botanical garden	Kazakhstan. Mangystau region. Aktau.
<u>ABG</u>	Altai Botanical Garden	Kazakhstan. East Kazakhstan. Ridder.

Herbarium Code	Institution	Location
<u>UKSPI</u>	S. Amanjolov East-Kazakhstan State University	Kazakhstan. Ust-Kamenogorsk.
<u>PPIU</u>	M. Utemisov Western Kazakhstanian State University	Kazakhstan. Uralsk.
<u>AA</u>	Ministry of Science, Academy of Sciences	Kazakhstan. Alma-Ata.
<u>KG</u>	International Phytochemistry Research and Production Institute	Kazakhstan. Karaganda.
<u>KSPI</u>	Kostanay State Pedagogical Institute	Kazakhstan. Kostanay.
<u>NUR</u>	Astana Botanical Garden	Kazakhstan. Nur-Sultan.
<u>QAR</u>	Buketov University	Kazakhstan. Karaganda Region. Karaganda.
<u>MANG</u>	Mangyshlak experimental botanical garden	Kazakhstan. Mangystau region. Aktau.
<u>ABG</u>	Altai Botanical Garden	Kazakhstan. East Kazakhstan. Ridder.



Некоторые оцифрованные гербарии, доступные через Интернет

The vascular plants collection (P) at the Herbarium of the Muséum national d'Histoire Naturelle (MNHN - Paris)	5 467 968
Naturalis Biodiversity Center (NL) - Botany	4 649 780
The New York Botanical Garden Herbarium (NY)	2 984 037
Meise Botanic Garden Herbarium (BR)	2 352 880
Moscow University Herbarium (MW)	965 468

Для чего оцифровывать коллекции

Оцифровка повышает доступность информации

Lepidium karataviense Regel & Schmalh. | Клоповник каратавский

В музее Естественной истории (Лондон) хранится 1 образец, собранный в Казахстане

Перелет Астана - Лондон ~ 490 000 ₸
Расстояние 5 900 км

Отсканированный образец и информация этикетки бесплатно доступны [на сайте музея](#) и на портале [GBIF](#)



<https://www.gbif.org/occurrence/1799005094>

Для чего оцифровывать коллекции



Национальный музей Бразилии, Рио-де-Жанейро

Учреждён королём Португалии Жуаном VI в 1818 году

20 миллионов образцов

Для чего оцифровывать коллекции



Огнем уничтожено 92.5% образцов

https://en.wikipedia.org/wiki/National_Museum_of_Brazil_fire



Для чего оцифровывать коллекции

Небольшая часть утраченных коллекций была оцифрована и доступна на портале GBIF

Coleção Entomológica do Museu Nacional / UFRJ

Occurrence dataset

"Museu Nacional/UFRJ está vinculado ao Ministério da Educação. É a mais antiga instituição científica do Brasil e o maior museu de história natural e antropológica da América Latina. Criado por D..."

Published by Museu Nacional / UFRJ

117,269 occurrences | 78 citations

R - Herbario do Museu Nacional

Occurrence dataset

O acervo está sendo organizado em armários compactados, em ordem alfabética de famílias Consta do Index Herbariorum com a sigla R, e estima-se que a coleção possua cerca de 550.000 exemplares, sendo 9...

Published by Museu Nacional / UFRJ

59,183 occurrences | 469 citations

Mollusca Collection - Museu Nacional/UFRJ

Occurrence dataset

The Mollusca Collection of the Museu Nacional (acronym MN RJ), one of the most important of its kind in South America, holds more than 40,000 registered lots and about 15,000 unregistered lots, which a...

Published by Museu Nacional / UFRJ

25,065 occurrences | 82 citations

Coleção de Aves do Museu Nacional / UFRJ

Occurrence dataset

Criado por D. João VI em 6 de junho de 1818, o Museu Nacional constitui um dos maiores e mais tradicionais centros de pesquisa da América Latina, sendo detentor de um dos mais vastos e representativos...

Published by Museu Nacional / UFRJ

32,060 occurrences | 145 citations

Coleção Ictiológica (MNRJ), Museu Nacional (MN), Universidade Federal do Rio de Janeiro(UFRJ)

Occurrence dataset

A Coleção Ictiológica (MNRJ) do Museu Nacional, Universidade Federal do Rio de Janeiro, é uma das maiores e mais antigas coleções científicas de peixes do Brasil, contando com um acervo iniciado no fi...

Published by Museu Nacional / UFRJ

Home Tools Community About

Occurrence dataset

Search

Login

Baza compta Simone & Cunha, 2006

Triphora alyra De Jong & Coomane, 1988

Triphora portoricensis Rolán & Redfern, 2008

Bullata bullata (Born, 1778)

Eulimastoma franklini Pimenta, 2012

Decilla notialis Pimenta, Santos & Absalão, 2008

sotriphora orca M.Fernandes;Pimenta & Leal, 2013

Centhilia sandela M.Fernandes;Garofalo & Pimenta, 2015

Pseudotornaria jonasii Tenório, Barros, Francisco & Silva, 2011

Psilaxis clertoni Tenório, Barros, Francisco & Silva, 2011

Heliacus williamsensis Tenório, Barros, Francisco & Silva, 2011

Thaumastus parvus Pena, Salgado & Coelho, 2011

Bullata guerrini de Souza & Coovet, 2001

Astralium latispina (Philippi, 1844)

Anachis carloslrae Costa, 1996

Thaumastus caetensis Pena, Salgado & Coelho, 2011

Nanaphora leei M.Fernandes & Pimenta, 2015

Osella aquilonia Pimenta, Santos & Absalão, 2008

Eulimetta atlantica Souza & Pimenta, 2015

Tomigerus esamianus Salgado & Coelho, 1990

Tomigerus matthewsi Salgado & Leme, 1991

Terebra crassireticula Simone, 1999

niforis pseudothomae Rolán & Fernández-Garcés, 1993

Nototriphora decorata C.B.Adams, 1850

Latiriphora albida A.Adams, 1854

Mitra barbadensis (Gmelin, 1791)

Monophorus olivaceus Dall, 1889

niforis carmiae Rolán & Fernández-Garcés, 1993

Cosmotriphora melanura (C.B.Adams, 1850)

sotriphora tigrina M.Fernandes;Pimenta & Leal, 2013

Coriophora novae Nowell & Usticke, 1969

Nanomelon viperinus Leal & Bouchet, 1989

Melissa ciliolata Gmelin, 1791

https://www.gbif.org/dataset/search?type=OCCURRENCE&publishing_org=4205110f-3f0f-40d8-bd0f-2fa71bc827b5
https://www.gbif.org/occurrence/gallery?publishing_org=4205110f-3f0f-40d8-bd0f-2fa71bc827b5

Оцифровка - перенос данных с бумажного носителя на электронный



DataCleaningExample - Excel (Сбой активации продукта)

Файл Главная Вставка Разметка страницы Формулы Данные Рецензирование Вид Надстройки ? Что вы хотите сделать?

Условные форматирования - как таблицу - ячейк - Стили Формат Удалить Вставить Сумма Сортировка Найти и фильтр - Редактирование

Вход Общий доступ

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
5	lee-2010-04	Achiptera coleoptrata (Linnaeus, 1758)	SPECIES	2010	2010-06-0RU	Bezhaniit!Polstovsk	57.0568	30.6463	WGS1984	10	5 ind / soil casper for A.S. Zaitsev A.IEE RAS						
6	lee-2010-05	Galluma obvia (Berlese, 1914)	SPECIES	2010	2010-06-0RU	Bezhaniit!Polstovsk	57.0993	30.3815	WGS1984	10	3 ind / soil bog A.S. Zaitsev A.IEE RAS						
7	lee-2010-07	Tectocephus velutinus (Michael 1880)	SPECIES	2010	2010-06-0RU	Bezhaniit!Polstovsk	57.0993	30.3815	WGS1984	10	1 ind / soil (Bottomanlar A.S. Zaitsev A.IEE RAS						
8	lee-2010-08	Achiptera coleoptrata (Linnaeus, 1758)	SPECIES	2010	2010-06-0RU	Bezhaniit!Polstovsk	57.1019	30.3891	WGS1984	10	25 ind / soil (mixed for A.S. Zaitsev A.IEE RAS						
9	lee-2010-09	Medonius hygrophilus (Muhaka 1987)	SPECIES	2010	2010-06-0RU	Bezhaniit!Polstovsk	57.1019	30.3907	WGS1984	10	92 ind / soil (peeing A.S. Zaitsev A.IEE RAS						
10	lee-2010-09	Salpingophora sp. (Linnaeus 1758)	SPECIES	2010	2010-06-0RU	Bezhaniit!Polstovsk	57.4417	30.6801	WGS1984	10	20 ind / soil (peeing A.S. Zaitsev A.IEE RAS						
11	lee-2010-10	Microtritta minima (Berlese, 1904)	SPECIES	2010	2010-06-0RU	Bezhaniit!Polstovsk	57.4417	30.6801	WGS1984	10	26 ind / soil meadow A.S. Zaitsev A.IEE RAS						
12	lee-2010-11	Rhysotritta duplicita (Grandjean 1953)	SPECIES	2010	2010-06-0RU	Bezhaniit!Polstovsk	57.4417	30.6801	WGS1984	10	4 ind / soil meadow A.S. Zaitsev A.IEE RAS						
13	lee-2010-12	Schelorhabtes laevigatus (Koch 1835)	SPECIES	2010	2010-06-0RU	Bezhaniit!Polstovsk	57.1019	30.4306	WGS1984	10	21 ind / soil (mixed for A.S. Zaitsev A.IEE RAS						
14	lee-2010-13	Paracatolaccus sp. (Linnaeus 1758)	SPECIES	2010	2010-06-0RU	Bezhaniit!Polstovsk	57.0993	30.3815	WGS1984	10	15 ind / soil (mixed for A.S. Zaitsev A.IEE RAS						
15	lee-2010-14	Phylodrepa peltifer (Koch, 1840)	SPECIES	2010	2010-06-0RU	Bezhaniit!Polstovsk	57.097	30.6418	WGS1984	10	20 ind / soil (mixed for A.S. Zaitsev A.IEE RAS						
16	lee-2010-15	Xenillus tegeocrenatus (Hermann 1804)	SPECIES	2010	2010-06-0RU	Bezhaniit!Polstovsk	57.057	30.6418	WGS1984	10	2 ind / soil (mixed for A.S. Zaitsev A.IEE RAS						
17	lee-2010-16	Hoplolithescarus illinoiensis (Ewing, 1909)	SPECIES	2010	2010-06-0RU	Bezhaniit!Polstovsk	57.171	30.6404	WGS1984	10	24 ind / soil (Transition A.S. Zaitsev A.IEE RAS						
18	lee-2010-17	Tectocephus velutinus (Michael 1880)	SPECIES	2010	2010-06-0RU	Bezhaniit!Polstovsk	57.171	30.6404	WGS1984	10	3 ind / soil (Transition A.S. Zaitsev A.IEE RAS						
19	lee-2010-18	Trichorhithes trimaculatus (C.L.Koch, 1836)	SPECIES	2010	2010-06-0RU	Bezhaniit!Polstovsk	57.171	30.6404	WGS1984	10	1 ind / soil (Transition A.S. Zaitsev A.IEE RAS						
20	lee-2010-19	Uroderes sp. (Linnaeus 1758)	SPECIES	2010	2010-06-0RU	Bezhaniit!Polstovsk	57.1019	30.3815	WGS1984	10	28 ind / soil (mixed for A.S. Zaitsev A.IEE RAS						
21	lee-2010-20	Manhannia dorsalis (Barliss, 1896)	SPECIES	2010	2010-06-0RU	Bezhaniit!Polstovsk	57.1042	30.39	WGS1984	10	3 ind / soil (raised for A.S. Zaitsev A.IEE RAS						
22	lee-2010-21	Phthiracarus globosus (Koch, 1841)	SPECIES	2010	2010-06-0RU	Bezhaniit!Polstovsk	57.1019	30.3891	WGS1984	10	1 ind / soil Spruce for A.S. Zaitsev A.IEE RAS						
23	lee-2010-22	Opilla nova (Oudemans, 1902)	SPECIES	2010	2010-06-0RU	Bezhaniit!Polstovsk	57.098	30.3891	WGS1984	10	2 ind / soil (Transition A.S. Zaitsev A.IEE RAS						
24	lee-2010-23	Tectocephus velutinus (Michael 1880)	SPECIES	2010	2010-06-0RU	Bezhaniit!Polstovsk	57.098	30.3891	WGS1984	10	11 ind / soil (Transition A.S. Zaitsev A.IEE RAS						
25	lee-2010-24	Zelotes sp. (Linnaeus 1758)	SPECIES	2010	2010-06-0RU	Bezhaniit!Polstovsk	57.098	30.3891	WGS1984	10	1 ind / soil (Transition A.S. Zaitsev A.IEE RAS						
26	lee-2010-25	Hoplolithescarus illinoiensis (Ewing, 1909)	SPECIES	2010	2010-06-0RU	Bezhaniit!Polstovsk	57.6452	57.1743	WGS1984	10	12 ind / soil (raised for A.S. Zaitsev A.IEE RAS						
27	lee-2010-26	Schelorhabtes laevigatus (Koch, 1835)	SPECIES	2010	2010-06-0RU	Bezhaniit!Polstovsk	57.1743	30.6452	WGS1984	10	1 ind / soil (raised for A.S. Zaitsev A.IEE RAS						
28	lee-2010-27	Chamobates cupidulus (Michael, 1884)	SPECIES	2010	2010-06-0RU	Bezhaniit!Polstovsk	57.3531	30.8125	WGS1984	10	1 ind / soil (Upland m.A.S. Zaitsev A.IEE RAS						
29	lee-2010-28	Schelorhabtes latipes (C.L.Koch, 1846)	SPECIES	2010	2010-06-0RU	Bezhaniit!Polstovsk	57.3531	30.8125	WGS1984	10	6 ind / soil (Upland m.A.S. Zaitsev A.IEE RAS						
30	lee-2010-29	Minuthroctes seminulus (Koch, 1841)	SPECIES	2010	2010-06-0RU	Bezhaniit!Polstovsk	57.3531	30.8125	WGS1984	10	4 ind / soil (Upland m.A.S. Zaitsev A.IEE RAS						

И

Источник: Книга А.С. Зайцева и А.Н. Баранова «Материалы к биологическому изучению насекомых из коллекции А.С. Зайцева» (1995). Книга включает в себя описание 100 видов насекомых из коллекции А.С. Зайцева, а также обзор литературы по насекомым Кызылординской области. В книге приведены описания видов, их систематическое положение, описание экологических особенностей, ареалы распространения и т.д.

© Баранов А.Н., Зайцев А.С. 1995

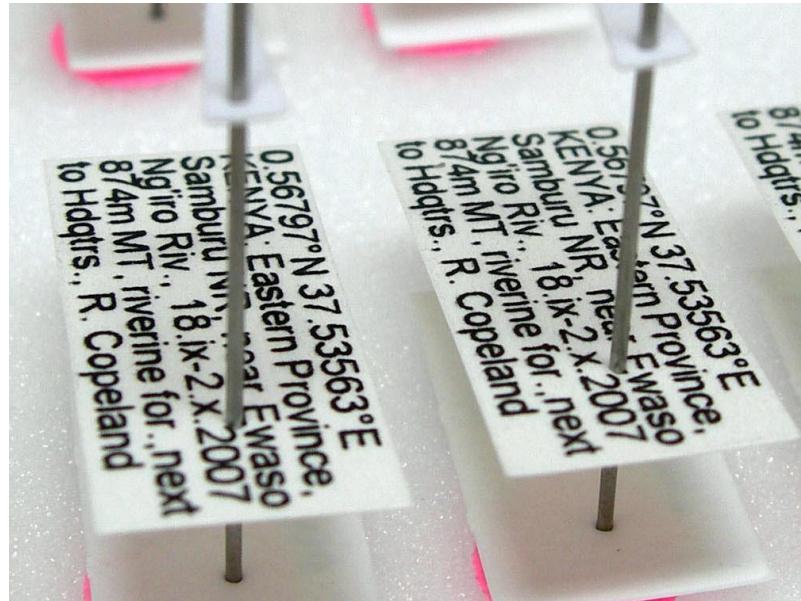
Оцифровка научных биологических коллекций

Физический образец
(само насекомое)



Этикетка

Что важнее - физический образец или этикетка?



ГЕРБАРИЙ КОМИ НАУЧНОГО ЦЕНТРА
УРАЛЬСКОГО ОТДЕЛЕНИЯ РАН (SYKO)

Dieramum drumondii Müll. Hal.

Республика Коми: Княжпогостский р-н,
р. Бычава, приокр р. Киселовка,
6 км до п. Киселово. У водог.

(62°20' - 50°50')

16.39

21.07.2010

Leg.: Тегерюк Б.Ю.

№ 12

№ общий 52388

Det.: Жемэлова Г.В.

Минимум данных, желательный для
регистрации находки в GBIF

Оцифрованные коллекции не
всегда предоставляют доступ к
изображениям образцов, но
обязательно - к этикеточным
данным

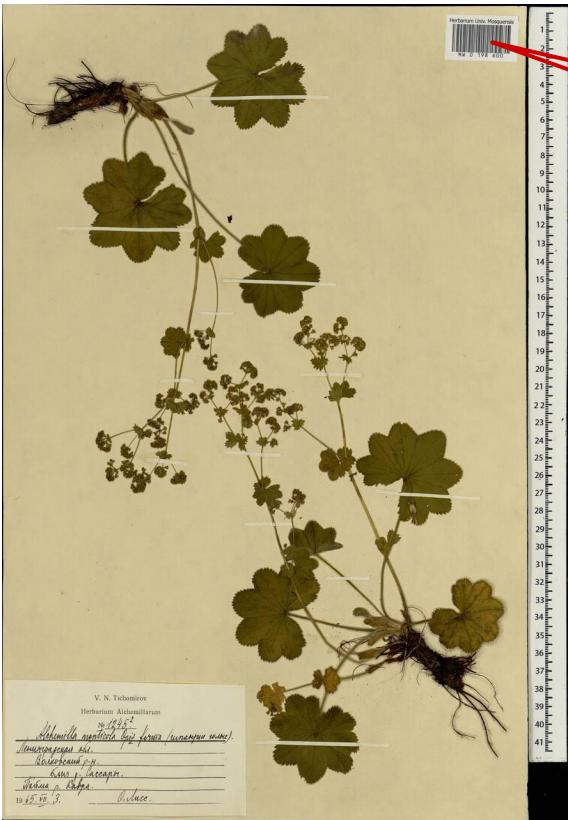
Оцифровка этикеточных данных без изображений образцов



Оцифровка этикеточных данных без изображений образцов

- Подготовка образцов к оцифровке (проверка, баркоды)
- Перенос информации с этикеток в таблицы Excel, отформатированные в соответствии требованиям GBIF (Darwin Core)
- Географическая привязка образцов
- Проверка и верификация данных (data cleaning)
- Публикация набора данных в GBIF

Баркод (штрихкод) | Barcode



Идентификатор
(уникальный номер, ID)
образца в коллекции для связи
физического образца и записи о
нем в таблице (базе данных)

OCCURRENCE DATASET | REGISTERED APRIL 19, 2023

Herbarium of Karaganda Buketov University

Published by [Karaganda Buketov University](#)

Ishmuratova M • Ageev D • Gavrilikova E • Mussina R • Madieva A • Ramazanov A • Tieukanova S

DATASET METRICS ACTIVITY [DOWNLOAD](#)

70 OCCURRENCES

The flora of the Karkaraly SNNP includes 198 species of angiosperms, 3 species of gymnosperms, 2 species of ferns, 27 species of mosses, 14 species of lichens currently registered, although according to literary sources their number can reach about 800 species. This is three times more than in the adjacent steppe zone. Of these 244 species, 5 are listed in the Red Book of the Republic of Kazakhstan: Karkaraly barberry, smooth sphagnum, Kyrgyz birch, thin poppy, spring adonis.

Publication date: April 19, 2023
Metadata last modified: April 19, 2023
Hosted by [Karaganda Buketov University](#)
Licence: CC0 1.0

[How to cite](#) [DOI](#) 10.15468/3asvbz 70 Occurrences 100% With taxon match 100% With coordinates 89% With year

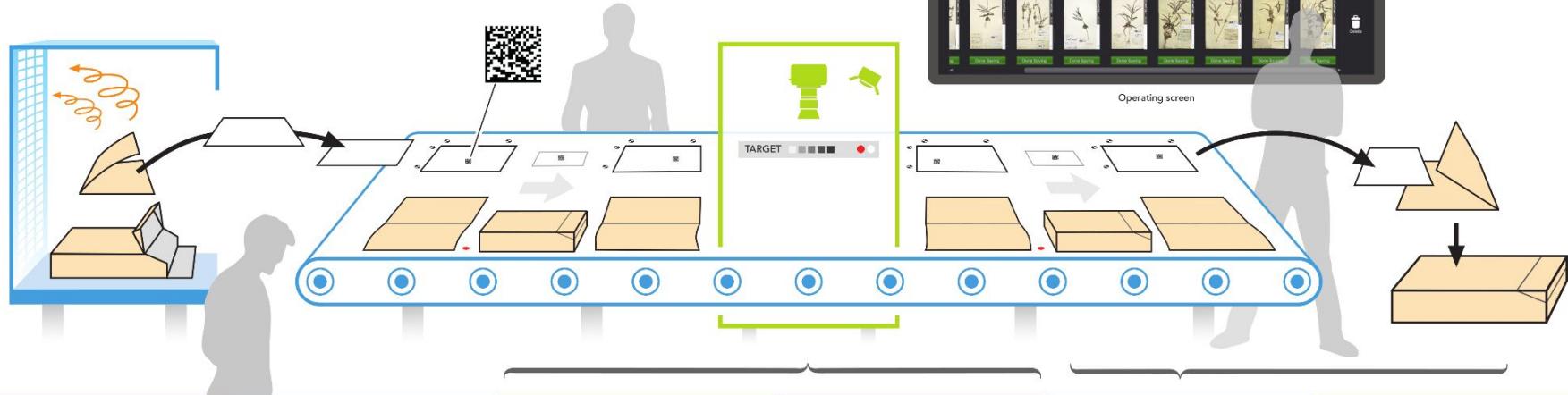
70 GEOREFERENCED RECORDS



**Набор данных,
подготовленный в
результате оцифровки
этикеток и
опубликованный в GBIF**



PICTURAE



1 Preparation

- Selection of boxes
- Extract vapor
- Apply box barcode
- Apply cover barcode

2 Place material on conveyor belt

- Spread on the conveyor belt
- Multisheet token
- Straighten
- Apply sheet barcode

3 Digitizing

- Read barcode
- Multisheet yes/no
- Apply ICC profile
- Readout color
- Readout sharpness
- Feedback → retake
- Rotating
- Cropping
- Readout target
- Multisheet color code
- Merge metadata to CSV file format
- Save deliverables

5 Packing

- Order remains intact
- Logistic management
- Return material

6 Metadata entry

- Cover & label description
- Look-up lists
- Linking to databases
- Multisheet processing

Обязательные элементы на оцифрованном гербарном листе



- (1) Цветовая палитра | Colour Chart
- (2) Масштабная линейка | Scale Bar
- (3) Баркод (штрихкод) | Barcode
- (4) Этикетка | Labels
- (5) Название института | Institution Name

Nieva de la Hidalga et al., 2020

<https://bdj.pensoft.net/article/47051/list/9/>

Оцифровка гербария



Установка для фотографирования
образцов в гербарии Arkansas
State University (STAR)
Harris & Marsico, 2017
[doi: 10.3732/apps.1600125](https://doi.org/10.3732/apps.1600125)

Оцифровка гербария



Установка для конвейерной
оцифровки образцов

Preparation of specimen prior to image capture



<https://doi.org/10.1002/tax.12767>

Где хранить изображения?

Файл с изображением этого образца занимает 521.5 kB

В настоящий момент в Цифровом гербарии ЦСБС СО РАН* 21 655 оцифрованных гербарных листов

В целом в гербарии высших растений, лишайников и грибов (NS, NSK)" хранится 808 550 листов и пакетов



*Центральный сибирский ботанический сад, Новосибирск, Россия

<http://www.csbg.nsc.ru/gerbarij/tematika-14.html>

Собственная разработка

Разработка структуры БД и
составление справочных
таблиц

Разработка интерфейса БД

Тестирование

Готовое ПО

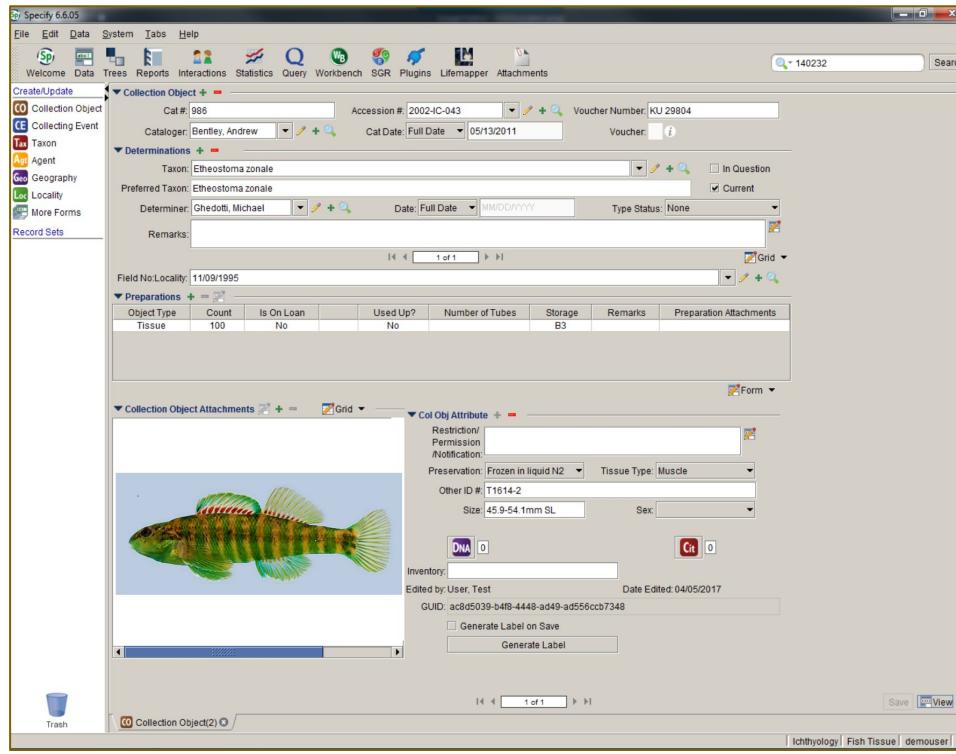
Установка и настройка
готового ПО
(+готовые справочные
таблицы)

Наполнение БД

Поддержание базы данных после окончания проекта

Техническая поддержка

Готовые программные продукты для внесения данных и ведения цифровых коллекций



<https://www.specifysoftware.org/>

Технический семинар по проекту ПЦФ Флора



КарУ, 15-17 марта 2023

Г

Этикеточные данные:

Excel таблицы

Данные в Darwine Core

Баркод: акроним+№листа
готовый скрипт для R



**Цветовая шкала X-Rite
ColorChecker Classic mini**



Отсканированные изображения:

Формат файла: TIFF,

Разрешение: 600 dpi

Название файла = баркод

(ABG000001.tiff)

Хостинг изображений: НЦБ

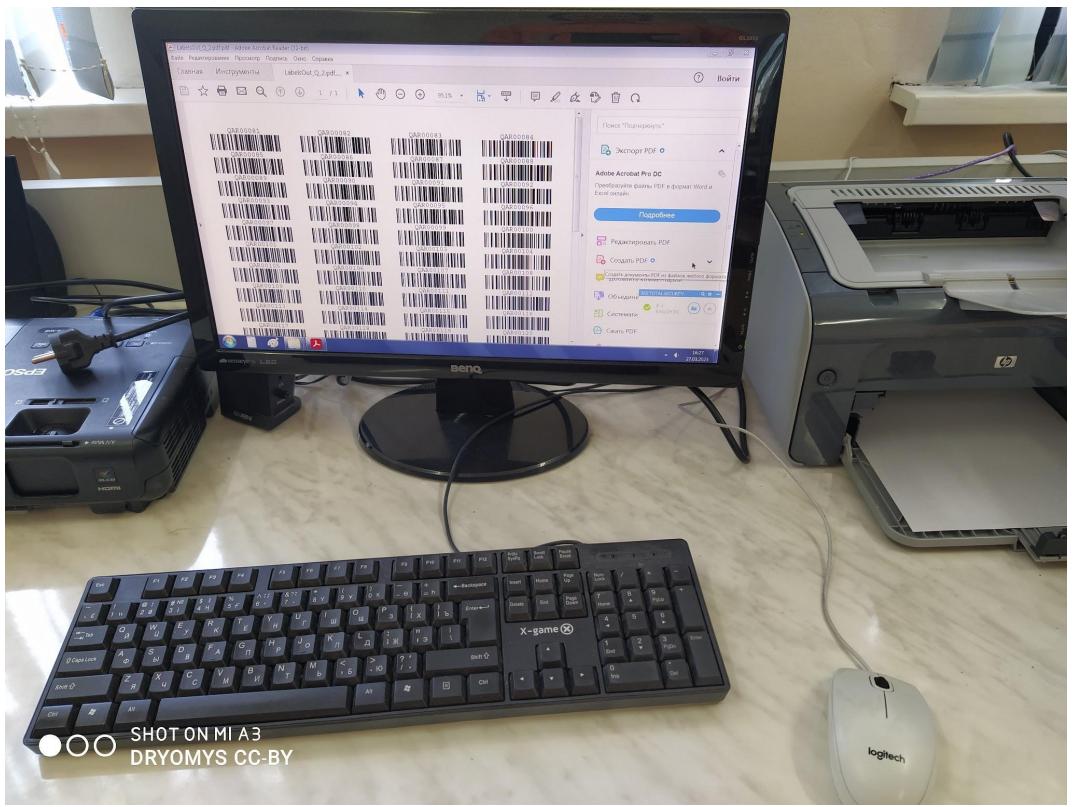
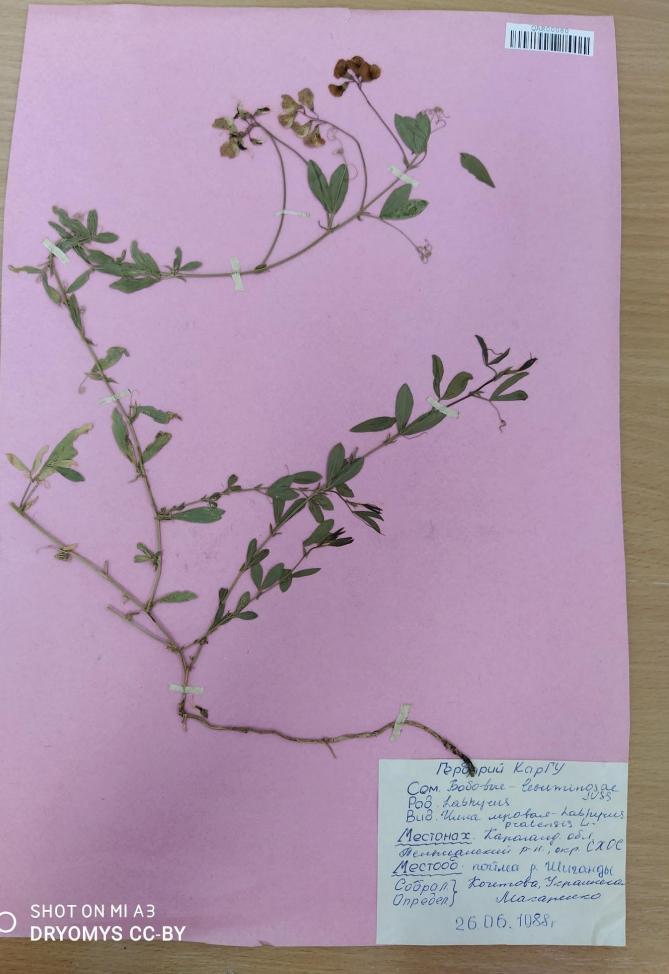
(Астана)

Основные активности по оцифровке этикеток

- Перенос этикеточных данных в таблицы (форматирование готовых таблиц) согласно стандартам GBIF
- Проверка таксономии
- Географическая привязка образцов
- Проверка и верификация данных
- Публикация данных через IPT

Дополнительные темы для обсуждения

- Получение и использование данных GBIF для науки и практики
- Вклад волонтеров в данные GBIF (citizen science)
- Статьи о данных (data paper)
- ГИС и R



Назарларыңызға рахмет! Спасибо за внимание!