

# SUPPORTING DATA MOBILIZATION ACTIVITIES IN RUSSIA WEST OF URAL MOUNTAINS

---

## MID-TERM ACTIVITY REPORT

### Contents

Executive summary .....	1
Contact information .....	1
Introduction.....	2
The project and its objectives.....	2
Project activities completed by mid-term.....	3
Project communications .....	4
Mid-term evaluation findings and recommendations for the remaining project implementation period .....	5

---

### Executive summary

The project aim is to support biodiversity data mobilization from European Russia. At this step of the project, the main activity was the data publishing workshop. The workshop was held in P.P. Shirshov Institute of Oceanology, Russian Academy of Sciences, Moscow. A total of 22 participants from 18 Russian organizations took part in the workshop. By now 12 new GBIF publishers have registered and 2 datasets were published. Mid-term evaluation was based on the original project proposal and the feedback from workshop participants. In general, we consider the workshop was successful. Our recommendations are rather about long-term GBIF activities in Russia, because it deal with methodical approaches and improving of learning courses.

---

### Contact information

**Natalya Ivanova**, Researcher, [Natalya.dryomys@gmail.com](mailto:Natalya.dryomys@gmail.com)

**Maxim Shashkov**, IPT administrator, [Max.carabus@gmail.com](mailto:Max.carabus@gmail.com)

Institute of Mathematical Problems of Biology RAS – the Branch of the Keldysh Institute of Applied Mathematics of the Russian Academy of Sciences (IMPB RAS)

Vitkevicha str. 1, 142290, Pushchino, Moscow Region, Russia

---

## Introduction

The first results of the project are presented in this report. At least one dataset must be published before this report according to the original project proposal. New GBIF publishers and published datasets are listed in the section *Project activities completed by mid-term*. The main activity was the Data publishing workshop at this step of the project. We estimated its results in the *Mid-term evaluation* section. The evaluation was conducted on the basis of feedbacks from participants. We received and analysed 16 responses.

The results of our project will contribute to filling a gap in the Russian data at the international GBIF biodiversity map. New occurrence data from Russia available through GBIF.org will allow doing some more accurate assessments in species distribution modelling, and contribute to development of international cooperation in the field of biodiversity study. It is also important for the national GBIF node establishment and the development of the national Russian portal based on ALA tools.

---

## The project and its objectives

This short-term project (1 January 2018 – 30 September 2018) aims at supporting of biodiversity data mobilization from European Russia. By now a huge amount of biodiversity data was collected by Russian scientists and naturalists, but most of this remain scattered, inconsistent and thereby unavailable for Russian or foreign researchers. The focus of this project is on already digitised (in different formats, such as DOC, XLS, DBF) data that can be easily and quickly mobilized through GBIF.org.

Implementation of the project provides 3 types of activities: i) data publishing workshop for beginners who would like to learn publishing biodiversity data through GBIF, but do not know how it works; ii) support of data standardization and data publishing for GBIF publishers from European Russia, who already have some experience in data publishing or theoretical understandings about it; and iii) conversion the data of IMPB RAS and Sibecocenter LTD into Darwin Core standard and publishing it through GBIF.org. The most of the work is carried out by IMPB RAS, which maintains Russian GBIF IPT. Currently

this IPT has associated 16 of 27 of the Russian publishers and actually (with GBIF.ru website) represented the informal national node.

We plan to involve 28 new publishers from European Russia and publish some number of datasets with about 105 000 occurrences in total as a result of these 3 activities.

---

## Project activities completed by mid-term

### 1. Data publishing workshop

The workshop was held in Moscow, P.P. Shirshov Institute of Oceanology, Russian Academy of Sciences (SIO RAS). We received total of 35 applications, and 22 participants from 18 Russian organizations were chosen among them. MEDUSA project co-funding allowed us to invite students from Asian part of Russia (Altai State University, Barnaul, Altai Krai and Ural Federal University, Ekaterinburg). These universities are big data holders and already have started the digitisation of collections. An intensive on-site training was comprised of general introduction to the world of global biodiversity data and GBIF, starting and managing of digitisation projects at the collection or institutional scale, general principles of digitisation and practical advices, data standardization according to Darwin Core standard, and data publishing through IPT. The program and other materials of workshop are available via GBIF.ru [http://gbif.ru/workshop\\_May\\_2018](http://gbif.ru/workshop_May_2018)

As the result of the workshop 10 new GBIF publishers were registered, mostly from European Russia, thus the total number of Russian publishers has increased by 30%.

#### New GBIF publishers

1. Institute of Ecology of the Volga river basin of the Russian Academy of Sciences <https://www.gbif.org/publisher/2fe11fba-80ca-4c62-998c-4318b893d072>
2. Bashkir State University <https://www.gbif.org/publisher/c22a8193-fbe7-436f-8da9-480eb5b07b96>
3. Nikolai Pertsov White Sea Biological Station <https://www.gbif.org/publisher/602f70f5-2131-412c-8745-8dcf928686f5>
4. Institute of Oceanology P.P. Shirshov <https://www.gbif.org/publisher/9abe4949-29aa-450e-b767-b2788f66e04d>
5. Pskov State University <https://www.gbif.org/publisher/174615f7-de29-4892-bf5d-dd028737b5ab>
6. Petrozavodsk State University <https://www.gbif.org/publisher/9cafd438-1edd-441d-b542-631878d9a3be>
7. Institute of North Industrial Ecology Problems – Subdivision of the Federal Research Centre “Kola Science Centre of Russian Academy of Science”

8. Karelian Research Centre of the Russian Academy of Sciences

<https://www.gbif.org/publisher/b87dc34e-ff76-458a-b163-f37e5634d0fb>

9. All-Russian Research Institute of Medicinal and Aromatic Plants (VILAR)

<https://www.gbif.org/publisher/5f8232b3-776b-47ec-b3a4-8a84fde46c95>

10\*. Federal State Autonomous Educational Institution of Higher Education 'Ural Federal University named after the first President of Russia B.N. Yeltsin'

<https://www.gbif.org/publisher/21a491e7-546a-4922-80ef-66594ed9d90c>

\*This institute located in the Asian part of Russia, but has data from European Russia. Its participation was supported by the MEDUSA project.

Published datasets:

Khlyap L, Albov S (2018). Checklist of mammal species of Prioksko-Terrasny reserve. Prioksko-Terrasnyi Biosphere Reserve. Checklist Dataset <https://doi.org/10.15468/6ijm2g> accessed via GBIF.org on 2018-05-27. 61 records

## 2. Data standardization support

New GBIF publishers from European Russia

1. Pechora-Ilych state nature biosphere reserve  
<https://www.gbif.org/publisher/d2e71fea-9a8a-4037-8cc4-be24013f6046>
2. Center of scientific creativity of youth "На Донской"  
<https://www.gbif.org/publisher/e9536c84-401a-45a2-8280-e9126d7dded8>

## 3. Data standardization by IMPB and Sibecocenter LTD

Published datasets:

Ivanova N, Shanin V, Grozovskaya I, Khanina L (2018). Forest vegetation of the northeastern part of the Kostroma region (European Russia). Institute of Mathematical Problems of Biology, Russian Academy of Sciences. Sampling\_event Dataset <https://doi.org/10.15468/qemuyc> accessed via GBIF.org on 2018-05-27. 6441 occurrences

---

## Project communications

A few hundred years of the exploration of Russian flora, fauna, and mycota have generated a great body of biodiversity data. Some data are already digitised and arranged in local databases, but most of the data are disaggregated and presented in different formats, while a central national biodiversity system is missing. As a result, Russia is still a 'blank spot' on the international GBIF biodiversity map. The results of our project will contribute to filling this gap. New occurrence data from Russia available through GBIF.org will allow doing

more accurate assessments in species distribution modelling. They will contribute to development of international cooperation in the field of biodiversity study. It is also important for the national GBIF node establishment and the development of the national Russian portal based on ALA tools.

---

## **Mid-term evaluation findings and recommendations for the remaining project implementation period**

Our workshop was the first practical course for Russian researchers. We combined two activities: our data publishing workshop and Digitisation of collections workshop organized by the MEDUSA project (Multidisciplinary EDUcation and reSearch in mARine biology in Norway and Russia project funded by the Norwegian Centre for International Cooperation in Education SIU). Thus, the program was covered all steps of data processing from labels digitising to data publishing through IPT. This format allowed us to invite collection's curators, researchers, university teachers, PhD students, Nature Reserve workers.

### ***List of participants***

1. Aleksandr Kokorin (LMSU Marine Research Center, Moscow)
2. Alexey Mishin (Institute of Oceanology, Moscow)
3. Alyona Tretyakova (Ural Federal University named after the first President of Russia B.N. Yeltsin, Ekaterinburg)
4. Anna Zhadan (M.V. Lomonosov Moscow State University (ББС), Moscow)
5. Antonina Kremenetskaia (Shirshov Institute of Oceanology RAS, Moscow)
6. Daria Portnova (Institute of Oceanology, Moscow)
7. Diana Malikova (Institute of Oceanology, Moscow)
8. Dim Akhmedyanov (Bashkir State University, Ufa)
9. Kirill Minin (Institute of Oceanology, Moscow)
10. Kristina Medvedeva (Altai State University, Barnaul)
11. Lyudmila Khlyap (A.N. Severtsov Institute of Ecology and Evolution, Moscow)
12. Maria Korneykova (Institute of North Industrial Ecology Problems – Subdivision of the Federal Research Centre 'Kola Science Centre of Russian Academy of Science', Apatity)
13. Nikolay Fadeev (All-Russian Research Institute of Medicinal and Aromatic Plants (VILAR), Moscow)
14. Nikolay Neretin (Lomonosov Moscow State University (ББС), Moscow)
15. Olga Ilina (Petrozavodsk State University, Petrozavodsk)
16. Roman Ufimov (Komarov Botanical Institute of RAS, Saint-Petersburg)
17. Stepan Senator (Institute of Ecology of the Volga river basin of RAS, Togliatty)
18. Tatyana Berezina (Institute of Steppe UB RAS, Orenburg)
19. Tatyana Drozdenko (Pskov State University, Pskov)
20. Vera Timofeeva (Forest Research Institute of the Karelian Research Centre of the

- RAS, Petrozavodsk)  
21. Yury Buyvolov (Prioksko-Terrasniy Biosphere Reserve, Moscow)  
22. Zorigto Namsaraev (NRC "Kurchatov Institute", Moscow)

Six teachers were invited, 3 from Europe and 3 from Russia.

- Digitisation* Sophie Pamerlon (GBIF France),  
Larissa Smirnova (Royal Museum for Central Africa, Botanic Garden  
Meise, Belgium),  
Susan Matland (University of Oslo, Norway)
- Publishing data* Natalya Ivanova (IMPB RAS, GBIF.ru, Russia),  
Maxim Shashkov (IMPB RAS, GBIF.ru, Russia),  
Nina Filippova (Yugra State University, GBIF.ru, Russia)

The participation of Sophie Pamerlon and Larissa Smirnova was very important not only for digitisation section but also for data publishing course. They made introduction to the Darwin Core standard in their reports. This is absolutely new topic for many Russian biologists and, to our opinion, it was a good practice for participants to hear about similar, intersected and adjoined terms from different point and from more than one speaker. According to the survey form, Darwin Core usage and data cleaning were attracted a great interest.

We received 16 responses through the Evaluation Form. Participants completed the form soon after returning home. In the future, it makes sense to do it in the last 10 minutes of the workshop, because this only makes sense if asked soon after the event.

The students appreciated the Course contents (mean evaluation was 4.6 of 5.0), Course structure and schedule (4.6), Course materials (4.8), Trainers (4.9), Interaction with the peers (4.7), Preparatory activities (4.8), Practical organization (4.9), Use of digital and online resources (4.7), Applicability of what they learned to their work place (4.9). Not all participants were satisfied by the support for languages (4.3), although 4 out of 6 teachers spoke in Russian. Apparently, we need in more Russian-language manuals on different stages and directions on both data publishing and usage.

All theoretical topics were very interesting for students and were appreciated: GBIF, role and mission (mean 5.0), Introduction to digitisation, institutional and global value (5.0), Starting and running a digitisation project in a (wet) collection (4.9), Principles of digitisation, what and why (4.9), Digitisation: presentations (4.9), IPT: presentations (4.8). Practical sessions were evaluated a little bit lower. Digitisation: practices was 4.7 points and IPT

practice 4.8. Digitisation practice was focused on marine zoological wet collections, but only 8 students were marine biologists. We understand that the digitisation principles are unified, but marine specimens can be somewhat exotic for the terrestrial biologists. Some participants would like to know more about software for collection management. We agreed, this would be an important topic for the next workshop. It also can be useful to pay attention to the technical side of digitisation, photographing and scanning techniques, ways of obtained data volume storing, services and hosting that can be used for these tasks, and so on.

The practical session on publishing through IPT became the first experience for GBIF.ru teachers. We used sandbox IPT that was deployed for EU BON program (<http://eubon-ipt.gbif.org/>). For our opinion, IPT session was successful because most of the students published the test dataset. Most of the difficulties in this part were due to the insufficient experience in PC using. Some participants expressed the willingness to obtain more practical skills in IPT working. Alyona Tretyakova (student) expressed her opinion about the possible program of further intensive IPT training. The program should include short introduction, session 1 Preparing Metadata only; session 2 Preparing Checklist Data and Occurrence Data; session 3 Preparing Sampling Event Data. In each session requirements, samples, and analyse errors are to be considered. We think that this is a good idea, but two full study days are needed at least for its implementation. On this way, a Russian-language handbook on data publishing by means of IPT could be written. Team of GBIF.ru already has an extensive experience and able to implement it with support of GBIF.

According to the feedbacks, the georeference topics were of great interest. Maxim Shashkov briefly described the theoretical basis, and Sophie Pamerlon showed some resources for geographic data validation. It is not news for us that insufficient awareness in the field of geographic informatics becomes a source of many inaccuracies and errors when it deals with spatial information associated with occurrences. The situation cannot be improved without basic knowledge on geographic coordinate systems, ways of coordinate representation and cartographic projections. It was a surprise, but Russian-language georeference resources (such as [gis-lab.info](http://gis-lab.info) and others) still not widely known for Russian biologists. We consider it would be useful to organize the georeference workshop with theoretical session, GPS using practice and calculating uncertainties using point-radius method.

In general, we consider the workshop was successful. All participants were interested. The discussion took place not only in class, but also during breaks and coffee breaks. Practical use of the acquired skills (data publishing through GBIF.org) by students is very important for successful project implementation. We are in contact with workshop

participants and support them in the data standardization. By now 9 new GBIF publishers already registered, 7 out of 9 new publishers have been associated with GBIF.ru IPT.

Workshop materials were posted on the GBIF.ru ([http://gbif.ru/workshop\\_May\\_2018](http://gbif.ru/workshop_May_2018)), also the small piece of news was published on the Bashkir State University website (<http://www.bashedu.ru/rnews/sotrudnik-bashqu-predstavil-universitet-v-seminare-otsifrovka-kollektsii-i-publikatsiya-dannyk>). Prioksko-Terrasny Reserve offered to organize a data publishing workshop for employees of Russian nature reserves and national parks in 2019.

The workshop program did not cover getting and using data from GBIF.org. This topic is not directly related to the data mobilization, but many potential publishers want firstly to get familiar with the already published data in the GBIF before joining themselves. In our opinion, it is necessary to write a Russian-language guide to working with GBIF data.

We will focus on the data mobilization process in the second part of the project. A total of 16 new publishers from European Russia will be registered in the GBIF. We will support the data standardization and publishing by new GBIF publishers. Also we will continue the IMPB vegetation data processing and start Sibecocenter LTD data standardisation. It is expected by the end of the project that the total amount of Russian publishers will be more than 40. Thus, the Russian GBIF community will become the largest national biodiversity information web. This will be a compelling argument for the national GBIF node creation.