

SUPPORTING DATA MOBILIZATION ACTIVITIES IN RUSSIA WEST OF URAL MOUNTAINS

FINAL ACTIVITY REPORT

Contents

1. Executive summary	1
2. Contact information.....	2
3. Project summary.....	2
3.1. Activities completed	2
3.2. Ongoing and post-project activities	4
4. Project objectives	5
5. Project deliverables	5
6. Project communications.....	9
7. Evaluation: findings and conclusions	9
8. Recommendations and lessons learned	9
9. Future plans	10

1. Executive summary

This project aims to increase the number of publishers and mobilize new national biodiversity data through GBIF.org. During the project implementation, a data publishing workshop (at the Shirshov Institute of Oceanology in Moscow) was held. Already digitized data of IMPB, MSU, and Sibecocenter LTD were standardized according Darwin Core and published. More than 50 000 specimens of already published MW collection was georeferenced and now available through GBIF.org. So, about 95 000 occurrences, and 86 records were published. Our efforts were also aimed at involving of new publishers from Russia. During our project, 28 new GBIF publishers were registered, and so the total number of Russian publishers has increased more than twice. Our efforts contribute to the creation of a Russian GBIF node, which will have greater capacity to publish and process biodiversity data. GBIF.ru data mobilization activities will be continued next year.

2. Contact information

Natalya Ivanova, Researcher

Maxim Shashkov, Senior Researcher, IPT administrator

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

3. Project summary

Our efforts during project implementation included three main activities: organisation of data publishing workshop, involvement of new GBIF publishers, and data standardization. Data publishing workshop was co-funded by MEDUSA project. In the data standardization process, we closely cooperated with Moscow university (MW) and Sibecocenter LTD. Additional direction of project implementation was a gathering metadata about Russian natural collections and archives of Russian protected areas. This continued activity is conducting jointly with Prioksko-Terrasny Nature Reserve.

3.1. *Activities completed*

Data publishing workshop

The main activity at the first step of the project was the data publishing workshop. It was held in Moscow, P.P. Shirshov Institute of Oceanology, Russian Academy of Sciences (SIO RAS). In the event, 22 participants from 18 Russian organizations took part. Our workshop was the first practical course for Russian researchers. 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). We combined two activities in this course: 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. According the analysis of the feedbacks the workshop was successful. The detailed description and evaluation of this activity were presented at the mid-term report.

Involvement of new GBIF publishers and data standardization process

At the second part of the project, we focused on data mobilization by existing Russian GBIF publishers and involvement new participants. Already digitized data (in XLS format) of IMPB, MSU and Sibecocenter LTD were standardized according Darwin Core and published. Moreover, we improved the quality of already published data of MW collection (doi [10.15468/cpnhcc](https://doi.org/10.15468/cpnhcc)). Initially this dataset was published without georeferenced occurrences. During this project, we made more than 50 000 georeferences based on the verbal descriptions from labels.

We also involved 28 new GBIF publishers (including 10 from data publishing workshop), some publishers are located in Asian part of Russia, but they have data collected in European part. Thus the total number of Russian publishers has increased has increased more than twice in 2018.

The review of Russian biodiversity data sources

We summarized Russian biodiversity data sources (DOI: [10.15468/hwv9pa](https://doi.org/10.15468/hwv9pa)). We found best known and largest Russian herbarium collections are stored in Komarov Botanical Institute (LE, more than 6 mil. sheets) and Moscow University (MW, more than 1 mil. sheets). The largest zoological collection is located in Zoological institute and counts more than 60 mil. specimens. There are also many different others biological collections in Russia, but unified list about them is absent and related information is scattered.

There are 81 biocollections from Russia registered on the Global Registry of Biodiversity Repositories (<http://grbio.org>). According to the portal Genetic and biological (zoological and botanical) collections of the Russian Federation (<http://www.sevin.ru/collections/>), 145 herbarium collections from 102 cities were present in Russia in 2004. During this work, we founded information about 160 regional herbarium collections with total storage more than 8 mil. specimens (excluding MW and LE). There are also 4 Russian live algae collection, total storage are 1258 living specimens (<http://www.wfcc.info/>). Available data on zoological collections are poorer. Large collections are stored in the Zoological Museum of Moscow University (6378700 specimens), the Museum of Institute of Plant and Animal Ecology (1150023 sp.), the Institute of Biology of Komi (118576 sp.) and the Museum of National Scientific Center of Marine Biology (about 1 mil. sp.).

The majority of Russian biological collections are not digitized. In our assessment total amount of digitized samples is about 1 mil., mostly presented by MW herbarium (99.5%). This collection almost completely digitized and now available through GBIF.org (doi

[10.15468/cpnhcc](https://doi.org/10.15468/cpnhcc)) and thematic information system (<https://plant.depo.msu.ru/>). A small part of the LE and SVER herbaria are also digitized (1320 and 5031 sp. respectively). Generalized data of labels for herbarium collections of the Polar-Alpine Botanical Garden-Institute of N.A. Avrorin KPABG, N.I. Vavilov Institute of Plant Genetic Resources (VIR) and the Prioksko-Terrasnyi Biosphere Reserve also available via GBIF.org (total 39114 records; doi [10.15468/yxt7co](https://doi.org/10.15468/yxt7co), [10.15468/nctfm2](https://doi.org/10.15468/nctfm2), [10.15468/80tu83](https://doi.org/10.15468/80tu83), [10.15468/xtcciv](https://doi.org/10.15468/xtcciv), [10.15468/cjzloe](https://doi.org/10.15468/cjzloe), [10.15468/r8ybng](https://doi.org/10.15468/r8ybng)). Thus, at least 82 million specimens are stored in Russian biological collections, but available (on-line) data is only 1.2%.

Another important Russian biodiversity data source is the Chronicle of Nature (in Russian «летопись природы / Letopis' Prirody»). It is the official document of Russian Nature Reserves. The Chronicle of Nature includes phenological and others observations from protected areas. So, this source has a huge potential for data mobilization. Collaboratively with Prioksko-Terrasny Nature Reserve we have started the work on data and metadata mobilization of the Chronicles of Nature. We interviewed scientific staff of Russian protected areas about their amount of Chronicles and already received 6 feedbacks.

3.2. Ongoing and post-project activities

In 2018, we plan following post-project activities:

1. Support for data processing and data publishing for new GBIF publishers. By now 28 new Russian GBIF publishers were registered, but most of them have not yet published data. The additional efforts by GBIF.ru are required for mobilization of the data from these publishers. Currently we are working on data standardization of herbarium of Central Siberian Botanical Garden SB RAS. The collection contains about 800 000 specimens of high vascular plants, mosses, lichens and fungi sampled in Siberia, Russian Far East, Europe, Asia and America.
2. Aggregation metadata about Russian digital biodiversity data sources, including the assessment of the potential for data mobilization of Russian protected areas. We will continue our activity with Prioksko-Terrasny Reserve and create the Register of Chronicles of Nature and biological collections. The results will be available through GBIF.ru.
3. Improvement of Russian-language guides on data standardisation and data publishing. GBIF.ru team had got a big experience during data publishing workshop in Moscow (P.P. Shirshov Institute of Oceanology, May) and Baikal data mobilization course (https://www.forbio.uio.no/events/courses/2018/Data_mobilization_Baikal.html). Existing manuals will be improved according with gained experience.

4. Organisation of two round tables devoted to principles of data standardization and data publishing. These events will be conducted in course of the meetings Russian Soil Zoology conference (October, 22–26, <http://www.bio-soil.ru/28-kruglyy-stol.html>) and «Eurasian Chronicle of Nature – Large Scale Analysis of Changing Ecosystems (ECN)» (Kenozero National Park, October, 28 – November, 2) .

4. Project objectives

All of the objectives included in the original project proposal were achieved (see detail in the section 5). We made some changes in the original plans in part of data standardization process – focused on georeferencing of MW. On the one hand, data of project generated publishers was not digitised or have low data quality. On another hand, MW is a largest Russian digitised collection and improvement of the data quality is very important.

5. Project deliverables

The original list of the project proposal / Reporting	Obtained results
Two-day data publishing workshop / 10 new publishers and 20 000 published occurrences	<p>10 new publishers (see the List of new GBIF publishers, №1-10)</p> <p>Published datasets (for details see the List of Published datasets):</p> <p>doi 10.15468/zwfsle: 20389 occurrences</p> <p>doi 10.15468/6ijm2g: 86 records</p> <p>Total: 20389 occurrences</p>
Support in data standardization / 18 new publishers and 75 000 published occurrences	<p>18 new publishers (see the List of new GBIF publishers, №11-28)</p> <p>Published datasets (for details see the List of Published datasets):</p> <p>doi 10.15468/h3ee9d: 6278 occurrences</p> <p>doi 10.15468/qcfcxk: 8077 occurrences</p> <p>doi 10.15468/qemuyc: 6441 occurrences</p> <p>doi 10.15468/cpnhcc: 54200 georeferenced occurrences</p> <p>doi: 10.15468/ryctj 95 occurrences</p> <p>Total: 75 095 published occurrences</p>
Data processing of Sibecocenter LTD / 10 000 published occurrences	the dataset will be published as soon as possible
TOTAL: 28 new publishers and 105 000 published occurrences	TOTAL: 28 new publishers and 95 484 published occurrences

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”
<https://www.gbif.org/publisher/78447c11-2b3b-46e6-9409-5eeb5518a646>
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. Lomonosov Moscow State University Marine Research Center
<https://www.gbif.org/publisher/ea7e0ccc-44e1-4655-b2f3-e11715f75690>
11. Pechora-Ilych state nature biosphere reserve
<https://www.gbif.org/publisher/d2e71fea-9a8a-4037-8cc4-be24013f6046>
12. Center of scientific creativity of youth "На Донской"
<https://www.gbif.org/publisher/e9536c84-401a-45a2-8280-e9126d7dded8>
13. Federal State Budget Institution KENOZERO National Park
<https://www.gbif.org/publisher/99fe7290-d54a-4f1e-b0dd-3989337a22eb>
14. I.I. Sprygin Zhiguli State Nature Biosphere Reserve
<https://www.gbif.org/publisher/36a57305-7751-49d3-9752-d2225120f32d>
15. Yu. A. Izrael Institute of Global Climate and Ecology (IGCE)
<https://www.gbif.org/publisher/9b4c8128-ecf3-4645-b014-3cd2fc26fa18>

16. State Nature Reserve "Kaluzhskie Zaseki"
<https://www.gbif.org/publisher/01b2fe32-68d6-4c0d-98fc-a4e31390ae92>
17. Nature museum of Kostroma region
<https://www.gbif.org/publisher/80de37d1-0b27-4da3-877b-21d7ac020253>
18. Kaluga Regional Eco-biological Center
<https://www.gbif.org/publisher/e3661514-2a08-495e-928a-070572bf70c9>
19. VORONEZHISKY STATE NATURE BIOSPHERE RESERVE NAMED AFTER V. PESKOV
<https://www.gbif.org/publisher/e56a7137-e0b9-491b-99ae-1dd03b8b0292>
20. National park "Ugra"
<https://www.gbif.org/publisher/d2ac4728-21f9-40e8-bb14-fa295957cd1e>
21. State nature biosphere reserve "Chernye zemli"
<https://www.gbif.org/publisher/55a21f97-04d4-4e7a-a317-786e4ad0b324>
22. CRIS
<https://www.gbif.org/publisher/0b00f562-9702-4d6d-b936-406996c60869>
23. Kostroma Division of Russian Geographical Society
<https://www.gbif.org/publisher/155915ea-1e12-423c-ac44-23a18cdf83a>
24. 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>
25. Federal State Budgetary Institution "United Directorate of Reserves of Taimyr"
<https://www.gbif.org/publisher/50d3275c-9e9d-406c-a7c2-99d0beb808ca>
26. Denezhkin Kamen Federal Nature Preserve
<https://www.gbif.org/publisher/42796302-9a3f-4a85-a5a5-c61f35ffba3d>
27. Beringia National Park
<https://www.gbif.org/publisher/efa7f97b-ba80-409e-b778-d596903a5aa3>
28. Central Siberian Botanical Garden SB RAS
<https://www.gbif.org/publisher/8f6232c3-5523-44ad-825c-e066fc6b15ad>

Published datasets

Khlyap L, Albov S (2018). Checklist of mammal species of Prioksko-Terrasny reserve. Prioksko-Terrasnyi Biosphere Reserve. Checklist Dataset <https://doi.org/10.15468/6ijm2q> accessed via GBIF.org on 2018-05-27. **86 records**

Senator S, Saksonov S, Vasjukov V (2018). Flora of the Volga River basin. Institute of Ecology of the Volga river basin of Russian Academie of Sciences. Occurrence dataset <https://doi.org/10.15468/zwfsle> accessed via GBIF.org on 2018-08-24. **20389 occurrences**

Neretin N, Zhadan A, Kolbasova G, Neretina T, Kotenev A, Tzetlin A, Kalyakin M, Klyukvin A (2018). ZMMU MSU, White Sea Branch. Version 1.8. Lomonosov Moscow State University. Occurrence dataset <https://doi.org/10.15468/h3ee9d> accessed via GBIF.org on 2018-08-19. **6278 occurrences**

Shovkun M, Moskalenko S, Ivanova N, Khanina L (2018). Occurrence Data of Vascular Plants from Abandoned Agricultural Fields (Moscow and Tula Regions, Russia). Version 1.2. Institute of Mathematical Problems of Biology, Russian Academy of Sciences. Occurrence dataset <https://doi.org/10.15468/qcfcxk> accessed via GBIF.org on 2018-08-19. **8077 occurrences**

Ivanova N, Shanin V, Grozovskaya I, Khanina L (2018). Forest vegetation of the northeastern part of the Kostroma region (European Russia). Version 1.1. Institute of Mathematical Problems of Biology, Russian Academy of Sciences. Sampling event dataset <https://doi.org/10.15468/qemuyc> accessed via GBIF.org on 2018-08-19. **6441 occurrences**

Seregin A (2018). Moscow University Herbarium (MW). Version 1.47. Lomonosov Moscow State University. Occurrence dataset <https://doi.org/10.15468/cpnhcc> accessed via GBIF.org on 2018-09-26. **54204 georeferenced occurrences**

Zubkova E, Ermolaev A, Frolov P (2018). Vegetation relevés of pine forests (Moscow region). Institute of physicochemical and biological problems in soil science of the Russian Academy of Sciences. Sampling event dataset <https://doi.org/10.15468/rycktj> accessed via GBIF.org on 2018-09-30. **95 occurrences**

Ivanova N, Shashkov M, Buyvolov Y, Filippova N (2018). The Register of Russian Biodiversity Sources. Version 1.2. Institute of Mathematical Problems of Biology, Russian Academy of Sciences. Metadata dataset <https://doi.org/10.15468/hwv9pa> accessed via GBIF.org on 2018-09-25. **Meta data only**

Important notice. We included project ID in the all dataset metadata, but for 2 datasets published through National Depository Bank of Live Systems IPT (doi [10.15468/cpnhcc](https://doi.org/10.15468/cpnhcc), [10.15468/h3ee9d](https://doi.org/10.15468/h3ee9d)) we include project ID only in Basic metadata description. Moscow University already thanked their project funding using the IPT fields. But now add another project as a reference is not possible.

6. 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 digitized 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. Our efforts contribute to the creation of a Russian GBIF node, which will have greater capacity to publish and process biodiversity data. It is also important for the development of the national Russian portal based on ALA tools.

7. Evaluation: findings and conclusions

This biodiversity information project is a first GBIF grant for Russia. During the project implementation, Russian GBIF-community has increased significantly. New national biodiversity data from European Russia were published through GBIF.org. These efforts contribute to the creation of a Russian GBIF node, which will have greater capacity to publish and process biodiversity data.

8. Recommendations and lessons learned

Recommendations for data publishing workshops preparing:

- Prepare the step-by-step instructions for exercises in the local language
- Use materials and examples from areas of student's investigations
- Pay attention to features of student's data. Good practice are presentations of student's research projects
- Special attention should be given to the occurrence georeference, the ways of representing the coordinates and coordinate reference systems

9. Future plans

During 2018 the largest number of publishers from Russia joined the GBIF in comparison with previous years, more than half of the total actual number - 44. Several more organization are going to become a publisher soon or we communicate with them about this. Nevertheless, most of new publishers have not yet published any datasets. In the next year we are going to focus our activities on these organizations, so that each will publish at least one dataset and, if possible by their staff. Thus in many of the new publishers a specialist will trained who has the skills to work with IPT.

The Russian language description of Darwin Core terms will be introduced into the specification (https://terms.tdwg.org/wiki/Darwin_Core). Another direction of our activity will be start of development of prototype of national biodiversity portal. After participating in ALA workshop in Madrid in February 2018 we have some basic knowledge in this field. Moreover, we installed server with blank ALA system <http://gbif.impb.ru/>. Further, using this server, we will study technologies for the developing and filling of portal with content.

Moreover, Prioksko-Terrasny Nature Reserve plans to organise data publishing workshop in autumn 2019. It will be the event focused on the Chronicle of Nature data digitization, standardisation and publishing through GBIF.org. GBIF.ru members will be among the course trainers.

Signed on behalf of the project partners

Date
