

# Разработка информационного портала «Биоразнообразии Югры»

Филиппова Нина Владимировна

*Югорский государственный университет (Ханты-Мансийск), Россия*  
e-mail: [filippova.courlee.nina@gmail.com](mailto:filippova.courlee.nina@gmail.com)

Филиппов Илья Владимирович

*Югорский государственный университет (Ханты-Мансийск), Россия*

Ханты-Мансийский автономный округ обладает огромными биологическими ресурсами, значительная часть которых может быть использована в качестве сырья в различных сферах экономики округа, направленных на повышение качества жизни населения. С другой стороны, на территории округа обитает и произрастает большое число редких и исчезающих видов, представляющих высокую научную и практическую ценность в региональном, национальном и глобальном масштабе. Наиболее редкие и ценные из них включены в региональную Красную книгу. Однако, несмотря на усилия ученых, и курирующего Департамента недропользования и природных ресурсов ХМАО-Югры, биологическое разнообразие автономного округа до настоящего времени остается недостаточно изученным, а имеющиеся сведения разрознены по многочисленным учреждениям на территории округа (музеи, университеты, ООПТ) и за его пределами.

С развитием компьютерных и интернет-технологий актуальным является использование современных стандартов для интеграции данных о биологическом разнообразии в региональных, национальных и глобальных информационных системах. В связи с этим нами была поставлена задача реализации регионального информационного портала по биологическому разнообразию ХМАО-Югры (ИП ЮграБио) для оцифровки и менеджмента биологических коллекций и интеграции цифровой информации о биоразнообразии округа в национальные и глобальные узлы.

Постоянно пополняющаяся актуальная региональная информационная система по биоразнообразию необходима и будет востребована для принятия управленческих решений в области планирования мероприятий и развития стратегии охраны окружающей среды, оптимизации сети ООПТ, составления проектной документации и проведения государственной экспертизы, оценки биологических ресурсов и планирования научных исследований. Использование современных методов информатики биоразнообразия значительно повысит качество этих исследований и откроет новые перспективы для изучения биоразнообразия и биологических ресурсов на территории Югры.

На территории Югры история создания информационных систем по разнообразию находится на начальном этапе. Так, информационная система для регистрации краснокнижных видов UgraBio разрабатывалась в рамках работ по ведению Красной книги Ханты-Мансийского автономного округа и предназначалась для унифицированного сбора и первичного анализа информации о биоразнообразии округа [Филиппов, 2012]. В 2011 г. была разработана web-версия системы UgraBio ([ugrbio.ru](http://ugrbio.ru)) с дифференцированным на роли многопользовательским доступом. К

началу 2018 г. UgraBio содержит порядка 3000 точек встреч видов, собранных 180 пользователями (в основном это точки встреч видов, включенных в Красную книгу ХМАО) и также списки видов, приуроченные к ключевым участкам.

В 2016 г. была создана информационная система Биологической коллекции ЮГУ на основе открытого программного продукта – Specify [Филиппова и др., 2017а; Филиппова, Лапшина, 2016], <https://fungariumysu.org/ysu-bc>. Эта система обеспечивает хранение данных биологических коллекций в формате Darwin Core, их демонстрацию онлайн, функции менеджмента коллекции и потенциальный экспорт данных в другие системы, в том числе Глобальную информационную систему по биоразнообразию GBIF. В настоящее время в систему загружены данные 4,5 тыс. образцов коллекции Фунгария ЮГУ. Кроме информации об образцах, система позволяет хранить связанные с ними иллюстративные материалы (рисунки, фотографии), а также имеет ГИС модули для отображения на карте точек находок.

Однако инструменты и технологии, использованные при организации этих систем, в настоящий момент устаревают. Поэтому необходимо выполнить интеграцию данных из имеющихся систем в построенную на современных стандартах с более широким кругом задач.

Для выполнения проекта по созданию ИП ЮграБио биологи, работающие на территории Югры, сотрудничают с ведущими учреждениями в области информатики биоразнообразия [Филиппова и др., 2017b]. Заявка на финансирование проекта была подана на региональный конкурс РФФИ 2018; первые работы по организации портала будут начаты вне зависимости от результатов конкурса. В ходе выполнения проекта ожидаются следующие результаты. Информация о локальных разрозненных коллекциях биологических образцов и других источников данных о биоразнообразии на территории округа будет интегрирована в базу метаданных. Основные региональные коллекции будут оцифрованы и их базы данных импортированы на ИП. Разработанные в ходе оцифровки этих коллекций алгоритмы и руководства будут положены в основу дальнейшей работы по оцифровке всех коллекций округа. Одновременно с интегрированием данных будет формироваться интерфейс, набор необходимых инструментов и структура баз данных исходя из региональных и национальных особенностей. Будет реализован регулярный периодический экспорт данных из регионального портала в национальный портал и GBIF. На основе полученного массива данных будет возможно выполнение таксономического, биографического, экологического и других видов анализов с использованием инструментов реализованных на ИП или после экспорта данных в стороннее программное обеспечение. Специалисты, занимающиеся изучением биоразнообразия на территории округа, в ходе семинаров, организованных в рамках проекта, пройдут обучение в сфере информатики биоразнообразия и использования ИП ЮграБио на практике.

В основе ИП по биологическому разнообразию Югры будет использован открытый свободно распространяемый программный пакет Symbiota, широко используемой в мире для создания порталов о биологическом разнообразии [Gries et al., 2014]. База данных пакета реализована в СУБД MySQL, программная часть интерфейса – при помощи языка PHP, а также JavaScript. Открытый программный код ресурса позволяет вносить изменения и присоединять дополнительные модули в зависимости от потребностей портала. Таким образом, продукт имеет гибкую настройку и возможности модификации с учетом национальных и региональных особенностей. Установка программного пакета будет осуществлена на виртуальном сервере в облачном сервисе Amazon. В установке пакета будут принимать уча-

стие техническая поддержка Symbiota и участники проекта, таким образом, будет обеспечено обучение локальных специалистов, и в дальнейшем портал будет поддерживаться силами участников проекта с частичным привлечением технической поддержки.

Индивидуальная настройка портала будет осуществляться с учетом национальных и региональных особенностей. Так, база географических названий будет создана с учетом организации национального Российского портала по биоразнообразию [Иванова, Шашков, 2014; Ivanova, Shashkov, 2017]. В этом же контексте будет осуществляться настройка таксономической базы данных. Будут разработаны модули, необходимые для интеграции данных об охраняемых видах Югры, в настоящее время хранящиеся в ИС UgraBio. Для удобства пользования порталом русскоязычным пользователем, основные поля интерфейса будут русифицированы. Настройка экспорта данных из ЮграБио в GBIF будет осуществлена через уже существующую инсталляцию IPT на базе Института математических проблем биологии РАН – филиала ИПМ им. М.В. Келдыша [Шашков, Иванова, 2016].

Оцифровка биологических коллекций существующих на территории округа будет проходить в несколько этапов. 1) На первом этапе будут созданы и импортированы в ЮграБио уже существующие базы данных образцов коллекций. На этапе создания баз данных отдельных коллекций будут определены обязательные и второстепенные поля для заполнения, а также написаны инструкции для стандартизированного выполнения операций (в частности, инструкция по географической привязке на основе алгоритмов в: [Charman, Wiczorek, 2006]. На этом этапе также будут определены условия правообладания на публикуемые данные и правила цитирования наборов данных [Muller, 2004; Penev et al., 2017]. 2) Второй этап будет включать собственно оцифровку коллекций и/или их частей, требующих перевода в электронную форму, т.е. сканирование или фотографирование хранящихся образцов и внесение данных этикеток в электронные базы. Для этого будут использованы специальные сканеры и фотокамеры с высоким разрешением. Сканирование, обработка и загрузка изображений в систему будут осуществляться на основе накопленного в этой области опыта крупных коллекций [Nelson et al., 2012; Tegelberg et al., 2014; Seregin, 2016].

Первичный анализ биоразнообразия Югры на основе данных, импортированных на ИП ЮграБио, будет осуществлен как внутри самого портала с помощью алгоритмов, включенных в пакет Symbiota [Gries et al., 2014], так и после экспорта данных с помощью стороннего ПО (R, QGIS). Анализ будет включать: 1) стандартный таксономический анализ по отдельным группам и отдельным географическим районам [Серегин, 2017]; 2) флористический анализ; 3) вычисление индексов биологического разнообразия; моделирование ареалов распространения видов и оценка уязвимости редких видов, и другие [Graham et al., 2004]. Для небольшого числа тестовых таксономических групп с помощью программного пакета, реализованного в Symbiota, будут созданы таксономические ключи [Gries et al., 2014].

Установка, настройка и развитие программного обеспечения ИП будет осуществляться силами участников проекта с привлечением технической поддержки Symbiota. На первых этапах работы участники проекта пройдут обучение и будут наделены правами администраторов системы. После создания баз данных отдельных коллекций, их владельцы будут наделены правами администрирования соответствующих коллекций и/или соответствующих разделов в таксономической базе данных. К наполнению системы данными будут привлечены пользователи, которые получают минимальный набор прав для администрирования. Таким образом, иерархическая структура управления информационной системой позволит

распределить права и обязанности и обеспечить стабильность ее функционирования в будущем.

Список источников и литературы:

1. Иванова Н.В., Шашков М.П. Перспективы создания открытого всероссийского информационного ресурса по биоразнообразию на основе международного стандарта GBIF // Математическая биология и биоинформатика. 2014. Т. 9. С. 396–405.
2. Серегин А.П. Гербарий Московского университета (MW) сегодня: фонды, онлайн доступ и научная работа // Ботанический журнал. 2017. Т. 102. № 3. С. 281–308.
3. Филиппов И.В. Итоговый отчет по договору № 038-2012: «Информационная система по объектам животного и растительного мира Ханты-Мансийского автономного округа». Москва, 2012. 31 с.
4. Филиппова Н.В., Бульонкова Т.М., Карпов Д.В., Лапшина Е.Д. Фунгарий Югорского университета и его база данных // Международная научно-практическая конференция «Использование современных информационных технологий в ботанических исследованиях». Апатиты, Мурманская область 28-31 марта 2017 г.: Тезисы докладов. Апатиты, 2017а. С. 137–144.
5. Филиппова Н.В., Лапшина Е.Д. Коллекция ваучерных образцов грибов ЮГУ: значение и основы функционирования // Биологические коллекции Югры: сбор, фиксация, хранение, введение в научный оборот. Материалы научно-методологического семинара. Ханты-Мансийск, 27 марта 2015 года. Ханты-Мансийск, 2016.
6. Филиппова Н.В., Филиппов И.В., Щигель Д.С., Иванова Н.В., Шашков М.П. Информатика биоразнообразия: мировые тенденции, состояние дел в России и развитие направления в Ханты-Мансийском Автономном Округе // Динамика окружающей среды и глобальные изменения климата. 2017б. Т. 8. №. 2. С. 46–56.
7. Шашков М.П., Иванова Н.В. Участие ИМПБ РАН – филиала ИПМ им. М.В. Келдыша РАН в развитии Глобальной Информационной Системы по Биоразнообразию (GBIF) // Математическая биология и биоинформатика: VI Международная конф., г. Пуццино, 16-21 октября 2016 г.: Доклады. М.: МАКС Пресс, 2016. С. 173–174.
8. Chapman A.D., Wiczorek J. Guide to Best Practices for Georeferencing. Copenhagen: Global Biodiversity Information Facility, 2006.
9. Graham C.H., Ferrier S., Huettman F., Moritz C., Peterson A.T. New developments in museum-based informatics and applications in biodiversity analysis // Trends in Ecology & Evolution. 2004. V. 19. № 9. P. 497–503.
10. Gries C., Gilbert E., Franz N. Symbiota – A virtual platform for creating voucher-based biodiversity information communities // Biodiversity Data Journal. 2014. V. 2. P. e1114.
11. Ivanova N.V., Shashkov M.P. Biodiversity Databases in Russia: Towards a National Portal // Arctic Science. 2017. doi: 10.1139/AS-2016-0050
12. Muller M.R. An Analysis of the Implications of Intellectual Property Rights (IPR) on the Global Biodiversity Information Facility (GBIF). GBIF, 2004. 43 p.

13. Nelson G., Paul D., Riccardi G., Mast A.R. Five task clusters that enable efficient and effective digitization of biological collections // *Zookeys*. 2012. № 209. С. 19–45.
14. Penev L., Mietchen D., Chavan V., Hagedorn G., Smith V., Shotton D., Tuama É.Ó., Senderov V., Georgiev T., Stoev P., Groom Q., Remsen D., Edmunds S. Strategies and guidelines for scholarly publishing of biodiversity data // *Research Ideas and Outcomes*. 2017. V. 3. P. e12431. doi:10.3897/rio.3.e12431
15. Seregin A.P. Making the Russian Flora Visible: Fast Digitisation of the Moscow University Herbarium (MW) in 2015 // *Taxon*. 2016. V. 65. № 1. P. 205–207.
16. Tegelberg R., Mononen T., Saarenmaa H. High-performance digitization of natural history collections: Automated imaging lines for herbarium and insect specimens // *Taxon*. 2014. V. 63. № 6. P. 1307–1313.