

БАЗА ДАННЫХ “ЛОКАЛЬНЫЕ ФЛОРЫ РОССИИ” В ОТКРЫТОМ ДОСТУПЕ

© 2024 г. А. П. Серёгин^{1,*}, Д. А. Бочков¹, К. Ю. Марченкова², Я. О. Магазов³,
С. В. Дудов¹, В. В. Чепинога^{4,5}

¹Московский государственный университет им. М. В. Ломоносова
Ленинские горы, 1, Москва, 119991, Россия

²Централизованная библиотечная система г. Сельцо, Брянской области
Ул. Куйбышева, 30, Сельцо, Брянская обл., 241550, Россия

³Центр детско-юношеского туризма “Космос” г. Челябинска
Ул. Кулибина, 54, Челябинск, 454085, Россия

⁴Ганноверский университет им. Лейбница
Нинбургерштрассе, 17, Ганновер, 30167, Германия

⁵Иркутский государственный университет
Ул. Карла Маркса, 1, Иркутск, 664003, Россия

*e-mail: botanik.seregin@gmail.com

Поступила в редакцию 17.04.2024 г.

Получена после доработки 04.08.2024 г.

Принята к публикации 13.08.2024 г.

Для целей создания “Атласа флоры России” мы перевели в электронный вид флористические списки, опубликованные в 636 источниках с 1975 по 2023 г. Предпочтение мы отдавали полным флористическим спискам, но для некоторых территорий с низкой плотностью данных источниками информации служили также парциальные флоры, маршрутные наблюдения, описания и обзоры растительности, а в отдельных случаях — разрозненные публикации с флористическими находками. На 04.04.2024 база данных “Локальные флоры России” (точка доступа в GBIF, doi:10.15468/rxtjt2) содержит 682130 записей о находках отдельных видов из 3297 географических пунктов. Для всех видов в пределах каждого флористического списка указаны общие географические координаты и точность геопривязки. Средняя точность геопривязки составляет 28 км. Открытый доступ к базе обеспечен через GBIF, что позволяет пользователям скачивать ее целиком или по частям в табличном виде. База продолжает наполняться новыми источниками.

Ключевые слова: пространственные данные, география растений, библиография, сосудистые растения

DOI: 10.31857/S0006813624080062, **EDN:** PBFLBK

Перевод сведений, ранее опубликованных в литературе, в формат баз данных получил широкое распространение в последнее время в разных областях науки, поскольку только электронные данные поддаются затем быстрой и эффективной обработке. Фактически, в подавляющем большинстве областей современной науки основное рабочее время исследователь проводит за монитором компьютера. В электронный вид переводятся архивы, библиотеки, музейные каталоги, научно-справочный аппарат, полевые записи. Процесс перевода данных (в частности, по био-

разнообразию) из аналогового вида в электронный получил название “мобилизация данных” (Nelson, Ellis, 2019).

Данные о биоразнообразии накапливают миллионы профессиональных исследователей и любителей природы по всему миру. Этот процесс затронул и флору России. В последнее время стало правилом хорошего тона выкладывать первичные электронные данные о встречах отдельных видов растений, грибов и животных в открытый доступ через интернет. В подавляющем большинстве случаев это так или иначе делается через воз-

возможности, предоставляемые Global Biodiversity Information Facility (GBIF, <https://www.gbif.org/>) (Global..., 2024).

В январе 2022 г. мы проанализировали информацию, имеющуюся в GBIF, на предмет возможности создания на ее основе сеточных карт для “Атласа флоры России” и их дальнейшего пространственно-статистического анализа. Стало ясно, что большие массивы имевшихся к тому моменту электронных данных – такие как iNaturalist (iNaturalist contributors, iNaturalist, 2024), Гербарий Московского университета (Seregin, 2024), оцифрованные значковые карты из “Флоры Сибири” (Artemov, Egorova, 2021) и др. – формируют диспропорциональную картину покрытия с преобладанием находок из регионов с высокой плотностью населения, которые охватывают центр и юг европейской части, Крым, Кавказ, юг Урала, неширокою полосу на юге Сибири и Дальнего Востока. Напротив, для обширных территорий Европейского Севера, таежной зоны Сибири, горных узлов Забайкалья, Охотии и Северо-Востока, большинства регионов Дальнего Востока покрытие электронными данными, имеющимися в GBIF, было чрезвычайно низким.

Массовая оцифровка флористических списков из опубликованных источников стала разумным ответом на этот вызов. Целью этой работы, прежде всего, стало закрытие белых пятен в электронных массивах данных с помощью сведений из литературы по флоре России. С другой стороны, целью этой работы стало создание единой точки доступа к унифицированным флористическим данным, опубликованным в разных форматах в сотнях статей, монографий, сборниках и диссертациях. Ключевыми задачами для достижения этой цели стали: (1) поиск источников и их сканирование (если это не было сделано ранее); (2) разработка протокола безошибочного распознавания списков видов и их перевод в формат электронных таблиц; (3) геопривязка списков; (4) публикация массивов оцифрованных данных в GBIF, отслеживание и исправление выявляемых ошибок; (5) создание библиографического списка оцифрованных источников.

В целом оцифровка литературы как отдельного типа данных пока заметно отстает от таких типов данных как оцифрованные гербарные образцы и фотонаблюдения. Так, ведущим массивом электронных данных по флоре России в GBIF

являются сведения из проекта “Флора России” на платформе iNaturalist (Seregin et al., 2020). На 17.04.2024 он охватывает 3.29 млн фотонаблюдений. Большинство верифицированных наблюдений в этом источнике происходит из Московской области, Москвы, Иркутской области, Крыма, Брянской области, Свердловской области и других регионов.

По оцифрованным гербарным образцам в GBIF с территории России выделяются Гербарий Московского университета (MW; 692 тыс.) и Гербарий Главного ботанического сада РАН (МНА; 140 тыс.) (Seregin, 2024; Seregin, Stepanova, 2024). Также в GBIF поступают данные из NS, NSK, ALTB, IRKU, VOR, KUZ, TUL и других гербариев. Оцифровка ведется также в LE, SVER, VBG, ABGI, но их данные в GBIF не отправляются.

Что касается GBIF-массивов, основанных на сведениях из литературы, то крупнейшими из них являются массив оцифрованных точек с карт “Флоры Сибири” (Artemov, Egorova, 2021) общим объемом 170 тыс. записей по 2569 видам. Позднее для целей “Атласа флоры России” был создан массив оцифрованных точек с карт многотомника “Сосудистые растения Советского Дальнего Востока” (Pospelov, Seregin, 2023a, b) общим объемом 157 тыс. записей по 4142 видам.

Ранее мы создали и опубликовали в GBIF пробный массив данных по двенадцати локальным флорам Европейской России, который был основан на опубликованных оригинальных флористических списках различного исходного формата (Seregin, 2021a, b). Это позволило отладить протоколы работы с текстовыми данными, протестировать автоматический и полуавтоматический парсинг текста для перевода списков в табличный вид, проверить возможности и ограничения GBIF в практической стандартизации номенклатуры.

МАТЕРИАЛЫ И МЕТОДЫ

1. Поиск источников и их сканирование

Корпус источников состоит из пяти неравных блоков, каждый из которых обладает своей спецификой: (1) статьи в журналах, (2) монографии, (3) электронные источники (базы данных), (4) главы и статьи в сборниках, (5) диссертации. Именно широкое развитие оцифровки (сканирования) литературы в последние годы и ее до-

ступность в интернете стала рациональной основой для скорого и эффективного выполнения нашего исследования. Не более 10% источников мы в итоге взяли в бумажном виде — все остальное было найдено в Интернете.

1.1. Статьи в журналах. Целенаправленно были целиком проработаны архивы основных тематических журналов, которые публикуют полные флористические списки: “Ботанический журнал” (в основном, списки по Арктике), “Комаровские чтения” (по Дальнему Востоку), “Фиторазнообразия Восточной Европы” (по Европейской России). Реже публикуют списки журналы “Turczaninowia”, “Растительный мир Азиатской России”, “Ботанические исследования Сибири и Казахстана”, а также “Труды Карельского научного центра РАН”. Архивы других журналов целенаправленно не прорабатывались, отдельные статьи из них мы включали в базу по результатам тематического поиска.

1.2. Монографии. Основой поиска монографий и сборников стала электронная библиотека “Флора и фауна”, основанная А.Б. Шипуновым (<https://herba.msu.ru/shipunov/school/sch-ru.htm>) (Fundamentalnaya..., 2024). Отсутствующие там книги, как правило, были в фондах ведущих библиотек Москвы (МГУ, РГБ) или были получены в электронном виде от коллег по нашим запросам. Списки литературы из обзорных монографий мы аккуратно изучали на предмет выявления редких и (или) подзабытых источников.

1.3. Электронные источники (базы данных). Списки видов отдельных заповедников, национальных парков и государственных заказников мы изначально взяли с портала-агрегатора “ООПТ России” (<http://www.oopt.aari.ru/>) (ООПТ, 2024). Общий объем исходных записей составил 70.5 тыс. единиц (строк), однако затем мы последовательно меняли их на изданные списки флор отдельных ООПТ. На 04.04.2024 на этот источник приходится 39.6 тыс. записей. Работа по замене данных агрегатора на списки из первоисточников продолжается.

Семь списков локальных флор для закрытия отдельных “белых пятен” мы взяли с сайта “Флора Таймыра” (<http://byrranga.ru/>) (Flora..., 2007–2024). Основой этого источника послужили материалы, собранные для монографии “Флора сосудистых растений Таймыра” (Pospelova, Pospelov, 2007) ее авторами в течение более чем

30-летних полевых работ, а также долгой и кропотливой работы с литературными источниками и гербарными материалами.

Кроме того, в нашем распоряжении имелась неопубликованная ранее база данных по флоре Крыма (около 10 тыс. записей), созданная ранее А. П. Серегиним на основе оцифровки десяти списков локальных флор отдельных урочищ и ООПТ.

1.4. Главы и статьи в сборниках, а также случайные статьи из непрофильных журналов и материалов конференций отслеживались преимущественно путем поиска по ключевым словам в специализированных поисковых системах, а также сплошным просмотром некоторых тематических сборников с подходящими заголовками. В целом именно корпус сборников, по всей видимости, может дать дополнительные источники, которые пока оказались вне поля нашего внимания.

1.5. Диссертации. Это ценнейший источник информации для “Атласа флоры России”. Некоторые диссертации по флористике и ботанической географии содержали обширный объем первичных сведений по флорам отдельных пунктов, который так и не был в дальнейшем опубликован. Диссертации последних лет находятся в открытом доступе на сайтах диссертационных советов. Тексты диссертаций 1995–2010 гг. мы заказывали в электронном виде, а более старые диссертации брали в диссертационном фонде РГБ.

По мере накопления данных и закрытия основных пространственных лакун на карте, мы проводили дополнительный поиск локальных и региональных источников для обеспечения равномерного покрытия данными.

Постепенно мы выработали основные принципы отбора источников для дальнейшей оцифровки. Общие требования к спискам таковы:

- список должен быть опубликован (диссертации — единственное исключение, когда мы брали источник с пометкой “На правах рукописи”);
- списки берутся строго из первоисточников (если автор присылал нам авторский набор статьи, мы сравнивали его с итоговым сканом работы);
- работы до 1974 г. пока не рассматриваются (причиной такого решения стало появление

сводов С. К. Черепанова (Czerpanov, 1973, 1981, 1995), ставших стандартом номенклатуры большинства работ, в то время как более старые публикации регулярно содержат архаичную номенклатуру);

- линейная протяженность выдела не более 200 км, т. е. радиус точности привязки должен быть менее 100 км (исключение делалось для работ, которые включали объекты как больше, так и меньше этой размерности);

- отсутствие кластерности (например, по заповедникам, которые состоят из нескольких участков, мы брали только списки по отдельным участкам, а не общий список для всего заповедника);

- предпочтение отдавалось полным флористическим спискам (однако в дальнейшем для территорий с небольшой плотностью данных источниками первичной информации служили парциальные флоры, описания и обзоры растительности, маршрутные наблюдения, а в исключительных случаях – публикации с флористическими находками).

Впрочем, исходя из плотности имевшихся данных, подход к этим пунктам общих требований был гибким. Так, например, для Якутии порог возраста публикаций был снижен до 1963 г. (Lukicheva, 1963; Buks, 1964; Kildyushevsky, 1964).

2. Перевод списков в формат электронных таблиц

Ключевым техническим этапом стала разработка протокола безошибочного распознавания и перевода списков в электронный вид. Мы располагали тремя типами исходных данных: (1) pdf-файлы публикаций (макеты от издателей); (2) сканы (pdf или djvu) со слоем распознанного текста; (3) сканы (pdf или djvu) без такого слоя. Для третьего типа распознанный слой мы создавали самостоятельно методами OCR (оптического распознавания символов), поэтому дальнейшая работа велась либо с pdf-макетами, либо с распознанным текстом.

Кроме того, публикации с флористическими списками зачастую имеют разную структуру представления данных: список видов через запятую; список видов в столбик; аннотированный список; таблица. Каждый из этих типов может быть дополнительно снабжен русскими названиями, указаниями экологических и географических групп,

распределением видов по отдельным пунктам, описанием местообитаний, ссылками на образцы и литературу и т. п. Все это создает довольно пеструю массу способов изложения флористической информации в публикациях, а потому единых (шаблонных) решений при переводе текста в электронный табличный формат у нас не было.

Парсинг – это процесс перевода исходного текста в таблицу (Jain et al., 2020). В целом каждая строка делилась пробелами на отдельные слова. После этого легко вычленились такие обязательные элементы как название рода (или его сокращение до первой буквы), видовой эпитет, авторы таксона, а также (при наличии) ранг внутривидового таксона и внутривидовой эпитет. В аннотированных списках латинское название с авторами таксона, как правило, отделено от самой аннотации тире, двоеточием или другими повторяющимися и легко вычленимыми элементами. После разделения текста на смысловые элементы они записывались в стандартные поля базы, а сокращенные родовые названия заменялись на полные.

Несколько иначе обстояло дело с парсингом опубликованных таблиц. Как правило, столбик с названиями переводился в поля таблицы отдельно от географического блока. Затем оцифровывались столбики с распространением таксонов – автоматически, полуавтоматически или вручную в зависимости от формата и качества исходного набора таблицы в скане публикации.

В отечественных публикациях с флористическими списками имеется как латиница (названия таксонов), так и кириллица (аннотации). При OCR регулярно происходит неконтролируемая замена сходных латинских символов на кириллические и обратно (*a, e, o, p, c* и т. п.). Та же проблема есть и в pdf-макетах статей. Итоговая таблица с латинскими названиями автоматически проверялась на отсутствие кириллических элементов, сходные символы менялись на латиницу.

Кроме того, в обязательном порядке мы отмечали заносные и (или) культурные виды, если это в явном виде отмечали авторы исходной публикации.

3. Геопривязка списков

Для всего списка каждой описанной флоры указываются одни и те же координаты, соответ-

ствующие центроиду работ. Как правило, каждая публикация содержит исчерпывающие описания географических границ выявленной флоры, длины радиальных выходов при работе на стационаре или карту района работ. Ключевым параметром, который также учитывается в нашей базе для каждой записи, является “точность”, т. е. радиус круга, охватывающего без остатка территорию исследования. Например, для всех видов из списка флоры Сохондинского заповедника (Vasilchenko, 2003) указаны координаты 49.7216° с.ш., 111.0230° в.д. и точность привязки 35000 м.

Описанное разнообразие представления флористических списков в публикациях усложняется и всевозможным характером представления географических данных. Если в публикации суммируются сведения по нескольким локальным флорам или отдельным районам исследований, то указание на встречу вида в том или ином пункте (районе) может быть дано разными абзацами аннотации, цифровым кодом, явным текстовым указанием пункта, значком, шрифтом (курсив, полужирный, подчеркнутый) или плюсиком в таблице. Таким образом, одна публикация может давать несколько итоговых списков. Например, во “Флоре Курильских островов” (Barkalov, 2009) в тексте аннотаций указано распространение 1411 видов сосудистых растений по 31 острову Курильской гряды. Это в общей сложности 7813 отдельных указаний видов с точностью до острова. При этом точность привязки различна – от 300 м на острове Сторожевой до 107000 м на острове Итуруп.

В некоторых случаях авторы не приводили описания мест сбора материала, однако давали карту с пунктами работ. Например, 108 пунктов работ в окрестностях Магадана (Belikovich, Galanin, 1992) или 34 пункта в Аяно-Майском районе Хабаровского края (Charkevich et al., 1983). Растровые карты из оригинальных публикаций мы привязывали в ГИС-пакете к реальным координатам с учетом исходных проекций, а затем в базу записывали координаты каждого пункта работ в отдельности.

Если в публикации имелась “матрешка” из общего списка какой-либо флоры и детализация распространения по отдельным стандартным пунктам, то мы, по возможности, использовали именно подробные указания. Например,

для флоры окрестностей Диксона (Matveyeva, Zanokha, 2017) в табличном виде приводится распространение видов как для острова Диксон, так и для нескольких пунктов на материке, отмеченных на спутниковом снимке. Для базы мы оцифровали флористический список для каждого из этих пунктов. Сходным образом отмечено распространение видов для 35 пунктов в пределах локальной флоры острова Русский в Приморье (Nedoluzhko, Denisov, 2001), 28 пунктов Верхнего Анадыря на Чукотке (Polezhayev, Berkutenko, 1982) и ряда других.

Помимо списков локальных флор и территорий сходной размерности, маршрутных точек, мы постепенно стали включать в базу сведения о распространении видов по административным и природным районам, принятым в различных региональных флорах. Например, это сведения о распространении видов по 13 природным районам Дагестана (Murtazaliev, 2009), 21 административному району Ивановской области (Shcherbakov et al., 2022), четырем отрезкам Черноморского побережья Кавказа (Zernov, 2002) и проч. Список привязывается к центроиду такого района, а параметр точность “накрывает” весь район кругом соответствующего радиуса. Это позволило получить равномерную флористическую информацию с давно освоенных территорий, на которых метод локальных флор не получил распространения (Кавказ, Средняя Россия, юг Европейской России, Урал, горы юга Сибири).

4. Размещение массива оцифрованных данных в GBIF

Мы использовали инструментарий GBIF для решения нескольких взаимосвязанных задач: (1) автоматическая стандартизация номенклатуры по внешнему источнику; (2) картографическая визуализация данных; (3) открытый доступ к базе; (4) стандартный формат выходных данных для дальнейшей работы над “Атласом флоры России”. Первый вариант массива данных (Bochkov, Seregin, 2024) был опубликован 16.01.2022 г., а затем многократно обновлялся по мере добавления новых оцифрованных списков.

4.1. Стандартизация номенклатуры. Технологии искусственного интеллекта и мягкого поиска семантических соответствий позволяют надежно связывать разнообразные орфографические варианты названий (и особенно авторов таксонов) со

стандартным названием в GBIF. Это позволило нам не заниматься длительной сверкой вариантов написания, которые автоматически привязывались к единственной записи.

Кроме того, наличие таксономического хребта GBIF (GBIF Secretariat, 2023), включающего как принятые названия, так и синонимы, позволило в автоматическом режиме добиться структурирования массива названий. “Невставшие” названия в случае необходимости мы исправляли вручную. Ключевой проблемой таксономии GBIF остается случай с гибридными формулами, которые не привязываются к родительским таксонам и не имеют отдельной точки входа (например, *Potentilla elegans* × *P. gelida* или *Rhinanthus groenlandicus* × *R. minor*). Для гибридов с бинарными названиями (например, *Calamagrostis* × *andrejewii* Litv. или *Drosera* × *obovata* Mert. et W.D.J. Koch) такой проблемы нет.

Небольшое число списков (например, Matveyeva, 1998) опубликованы без указания авторов таксонов. В редких случаях при наличии омонимов автоматическая привязка некоторых названий становилась невозможной — здесь мы добавляли авторов таксонов полуавтоматически, используя другие списки.

4.2. Картографическая визуализация данных. Ценный инструмент GBIF заключается в автоматизации создания карт. Это позволило нам контролировать и устранять случайные ошибки в геопривязке отдельных списков (например, по отлетам в ареалах отдельных видов), постепенно пополнять источниками незаполненные квадраты на разном масштабе агрегации данных, а также намечать места будущих полевых работ.

4.3. Открытый доступ к базе. База данных изначально задумывалась как открытый источник информации. Весь объем оцифрованных сведений доступен пользователям как по отдельным записям (поиск по таксонам, географическим выделам, отдельным точкам, произвольным выделам на карте и пр.), так и в виде полноценных табличных выгрузок в формате Darwin Core. За два года GBIF-выгрузки из нашей базы были процитированы с помощью doi 121 раз.

Доступ к базе на данный момент осуществляется только через GBIF (doi:10.15468/rxtjt2). Зарегистрированный пользователь может скачать табличный вариант данных (формат csv) с переменными и константами, описанными в разделе

“Результаты”. Незарегистрированный пользователь также может использовать для поиска поля поисковых форм GBIF в двух вариантах: простой поиск (14 полей) или расширенный поиск (92 поля). Выборки, полученные в результате заполнения поисковых форм, могут быть скачаны в табличном формате.

Кроме того, онлайн-работа с фильтрами позволяет получить разнообразные варианты картографической визуализации оцифрованных нами данных как отдельно по базе данных “Локальные флоры России”, так и в целом по всему массиву GBIF-данных.

4.4. Работа со стандартным форматом выходных данных. “Пропуская” базу данных через GBIF, мы добиваемся стыковки наших электронных данных с другими GBIF-источниками (например, Artemov, Egorova, 2021; iNaturalist contributors, iNaturalist, 2024; Seregin, 2024; Seregin, Stepanova, 2024 и множеством других). Именно GBIF-выгрузки в итоге положены в основу “Атласа флоры России”.

5. Создание библиографического списка оцифрованных источников

Эта работа шла параллельно оцифровке списков из самих источников. К сожалению, формат журнальной статьи не позволяет привести список 636 источников (на 04.04.2024), использованных при создании базы данных “Локальные флоры России”. В самой базе ссылка дается кратко (автор — год) для каждой записи.

Библиографический список оцифрованных источников ведется нами в табличном виде и содержит стандартные для библиографической записи поля: название публикации (статьи, главы), название источника (журнал, монография, сборник), авторы, год, место публикации, издатель, doi, страницы и пр. Информация дублируется на английском языке при наличии перевода в самой публикации, а при его отсутствии заполняется транслитом автоматически. Полный библиографический список доступен по запросу.

Авторы базы данных Д.А. Бочков и А.П. Серегин. Также в ее создании приняли участие Я.О. Магазов (источники по Челябинской области), К.Ю. Марченкова (оцифровка списков в 2024 г.), С.В. Дудов (геопривязка оригинальных карт с пунктами работ).

РЕЗУЛЬТАТЫ

На 04.04.2024 база данных “Локальные флоры России” содержит 682130 записей о находках отдельных видов из 3297 географических пунктов, флора которых опубликована в 636 источниках. При этом в 1781 пункте отмечено 100 и более видов, а в 1136 пунктах — от 10 до 99 видов. Еще 380 географических пунктов введены на основании разрозненных сведений из дополнительных источников для закрытия крупных белых пятен (в каждом из них отмечено не более девяти видов). Средняя точность геопривязки составляет 28014 м.

Структура записи. Каждая запись в GBIF-массиве “Local floras of Russia” (Bochkov, Seregin, 2024) содержит 42 поля стандарта Darwin Core — 20 констант и 22 переменные. Еще несколько полей GBIF генерирует автоматически после загрузки данных (например, название семейства, порядка и класса на основании предоставленного названия рода).

Константы — это общие для всего массива технические характеристики записей, которые индексируются в GBIF. Это, например, язык массива; лицензия; организация-владелец массива; идентификаторы организации, коллекции и массива данных в разных системах; код организации; название массива; тип записи (у нас — сведения из литературы); протокол обобщения данных; страна; геодезический эллипсоид; царство и отдел; юрисдикция кодекса (Turland et al., 2018) и пр.

Список переменных уточнялся в ходе оцифровки массива литературы и, собственно, основан на результатах нашей работы. Характеристики переменных (с примерами) даны ниже, в скобках приведен шифр переменной в стандарте Darwin Core:

1. Идентификатор записи (*Occurrence ID*). Это машиночитаемый уникальный код вида urn: lsid: biocol.org: col:15550:16:0746:0001, где urn: lsid: biocol.org: col:15550 — код МГУ, 16 — код массива, 0746 — код географического пункта, 0001 — порядковый номер вида в списке растений из этого пункта.

2. Номер в каталоге (*Catalog Number*). Это более удобный для восприятия код вида Syunaisale:0001, где Syunaisale — это краткое название географического пункта. В данном случае это локальная флора Сюнайсале, опубликованная

О.В. Ребристой (Rebristaya, 2013), а 0001 — порядковый номер вида в списке растений из этого пункта.

3. Характер натурализации (*Establishment Means*). Указания на то, является ли вид заносным, интродуцированным, культурным и проч. (стандартные значения: “alien”, “cultivated”, “alien or cultivated”, “introduced”, “garden escape / relic of the former cultivation” и ряд других).

4. Библиографическая ссылка (*Associated References*). Краткая библиографическая ссылка на латинице такого вида: Rebristaya (2013).

5. Часть света (*Continent*). В нашем массиве используются варианты “Europe” и “Asia”.

6. Код страны / территории по ISO 3166-1-alpha-2 (*Country Code*). Как правило, RU.

7. Административная единица первого порядка (*State Province*). Нормированное название субъекта РФ на английском языке (например, Yamalo-Nenets Autonomous Okrug).

8. Место (*Locality*). Краткое название географического пункта — например, “Syunaisale” (см. пояснения в пункте 2). В случае необходимости пункт может быть назван более подробно, чем номер в каталоге (пункт 2).

9 и 10. Географические координаты (*Decimal Latitude* и *Decimal Longitude*). В десятичном виде с точностью до четвертого знака после запятой (например, 66.9064N, 71.2702E).

11. Точность геопривязки (*Coordinate Uncertainty In Meters*). Например, 5500 м для локальной флоры Сюнайсале.

12. Автор геопривязки (*Georeferenced By*). Автор оцифровки списка и его геопривязки к координатам. Если это разные люди, указано два человека через запятую.

13. Дата геопривязки (*Georeferenced Date*). Фактически, дата добавления полностью обработанного источника в базу в формате ГГГГ-ММ-ДД (например, 2022-02-04).

14. Примечания к геопривязке (*Georeference Remarks*). Для публикаций, в которых точные указания для отдельных пунктов (районов) чередуются с общими указаниями “повсеместно”, “во всех районах” и пр., в этом поле вводятся дополнительные пометки. Например, “Precise indication for a definite division” и “Indicated for

all divisions of Chelyabinsk Oblast” при оцифровке определителя П.В. Куликова (Kulikov, 2010).

15. Примечание к определению (*Identification Qualifier*). Пометки вида “sensu lato”, “sensu stricto”, “aggregate”, “cfr.” и пр.

16. Название таксона (*Scientific Name*). Дано по первоисточнику. Производится следующая нормировка: сокращенное название рода раскрывается (*T.* до *Trifolium* и т. п.); у подвидов и разновидностей указываются только авторы внутривидового таксона, а не вида; стандартизируются ранги (например, “subsp.”, а не “ssp.” для подвида).

17. Источник номенклатуры (*Name According To*). Та же ссылка, что и в пункте 4 (библиографическая ссылка), поскольку номенклатура берется строго по опубликованному источнику.

18, 19 и 20. Род (*Genus*), видовой эпитет (*Specific Epithet*), внутривидовой эпитет (*Infraspecific Epithet*). Берутся по опубликованному источнику.

21. Ранг таксона (*Taxon Rank*). Как правило, вид или подвид, но встречаются также рода, разновидности, формы, нотоподвиды и пр.

22. Автор(ы) таксона (*Scientific Name Authorship*). Полное указание авторства таксона согласно опубликованному источнику.

Географический охват базы изначально формировался для целенаправленного заполнения пропусков для “Атласа флоры России”. Таким образом, мы максимально полно оцифровывали источники по Дальнему Востоку, Арктике, таежным и горным районам Сибири, северу европейской части России. Напротив, многие источники, охватывающие, например, Среднюю Россию, пока отложены в сторону. Число источников, пунктов и записей в базе данных по субъектам РФ приведено в табл. 1.

Поскольку площадь регионов различна, то нормированный на площадь объем данных можно показать на сеточной карте (рис. 1). Видны

Таблица 1. Число источников литературы, географических пунктов и записей в базе данных “Локальные флоры России” по субъектам РФ (на 04.04.2024)

Table 1. Number of references, locations and records in the “Local Floras of Russia” dataset by regions of the Russian Federation (as of 04 Apr 2024)

Регион Region	Источники References	Пункты Locations	Записи Records
Европейская часть / European Russia			
Архангельская область / Arkhangelsk Oblast	24	163	21664
Астраханская область / Astrakhan Oblast	3	12	4760
Башкортостан / Republic of Bashkortostan	13	68	27362
Белгородская область / Belgorod Oblast	2	2	731
Брянская область / Bryansk Oblast	4	4	1752
Волгоградская область / Volgograd Oblast	9	21	3703
Вологодская область / Vologda Oblast	6	17	4872
Воронежская область / Voronezh Oblast	3	4	2666
Ивановская область / Ivanovo Oblast	2	22	13250
Калининградская область / Kaliningrad Oblast	2	3	787
Калмыкия / Kalmyk Republic	2	19	1417
Калужская область / Kaluga Oblast	1	1	693
Карелия / Republic of Karelia	20	49	10365
Кировская область / Kirov Oblast	2	2	1194
Коми / Komi Republic	15	64	14556

Таблица 1 (продолжение)

Table 1 (continued)

Регион Region	Источники References	Пункты Locations	Записи Records
Костромская область / Kostroma Oblast	1	1	584
Ленинградская область / Leningrad Oblast	9	56	12890
Липецкая область / Lipetsk Oblast	1	5	3214
Мари Эл / Mari El Republic	1	2	754
Мордовия / Republic of Mordovia	4	5	2066
Московская область / Moscow Oblast	3	3	832
Мурманская область / Murmansk Oblast	11	59	8256
Ненецкий АО / Nenets Autonomous Okrug	12	53	9220
Нижегородская область / Nizhny Novgorod Oblast	1	1	598
Новгородская область / Novgorod Oblast	3	10	1101
Оренбургская область / Orenburg Oblast	6	8	2322
Пензенская область / Penza Oblast	8	39	7044
Пермский край / Perm Krai	18	62	7756
Псковская область / Pskov Oblast	4	11	2840
Республика Крым / Republic of Crimea	12	15	8019
Ростовская область / Rostov Oblast	5	14	1660
Рязанская область / Ryazan Oblast	1	1	716
Самарская область / Samara Oblast	21	44	12391
Санкт-Петербург / City of Saint Petersburg	1	1	621
Саратовская область / Saratov Oblast	3	14	1697
Свердловская область / Sverdlovsk Oblast	8	51	6377
Севастополь / City of Sevastopol	2	9	2039
Смоленская область / Smolensk Oblast	1	1	927
Татарстан / Republic of Tatarstan	3	4	2534
Тверская область / Tver Oblast	3	4	3353
Тульская область / Tula Oblast	1	1	359
Удмуртская Республика / Republic of Udmurtia	6	12	7746
Ульяновская область / Ulyanovsk Oblast	19	49	14334
Челябинская область / Chelyabinsk Oblast	12	75	23095
Чувашская Республика / Chuvash Republic	1	1	352
Ярославская область / Yaroslavl Oblast	2	2	1044
Кавказ / Caucasus			
Дагестан / Dagestan	5	18	12494
Ингушетия / Ingushetia	1	1	162

Таблица 1 (окончание)

Table 1 (end)

Регион Region	Источники References	Пункты Locations	Записи Records
Кабардино-Балкария / Kabardino-Balkaria	1	5	3868
Карачаево-Черкесия / Karachay-Cherkessia	2	2	1751
Краснодарский край / Krasnodar Krai	4	71	8286
Республика Северная Осетия / Republic of North Ossetia	1	2	453
Ставропольский край / Stavropol Krai	1	2	485
Чеченская Республика / Chechnya	2	9	7596
Сибирь / Siberia			
Алтайский край / Altai Krai	8	89	27318
Бурятия / Buryatia	9	34	7553
Забайкальский край / Zabaikalsky Krai	11	53	6258
Иркутская область / Irkutsk Oblast	11	105	18702
Кемеровская область / Kemerovo Oblast	13	37	17297
Красноярский край / Krasnoyarsk Krai	71	240	54685
Курганская область / Kurgan Oblast	3	35	19664
Новосибирская область / Novosibirsk Oblast	1	1	581
Омская область / Omsk Oblast	3	3	1056
Республика Алтай / Altai Republic	7	17	5175
Томская область / Tomsk Oblast	3	7	2007
Тува / Tuva	7	22	9414
Тюменская область / Tyumen Oblast	7	34	10949
Хакасия / Khakassia	3	17	3980
ХМАО – Югра / Khanty-Mansi Autonomous Okrug – Yugra	15	27	4576
Якутия / Yakutia	70	277	34570
ЯНАО / Yamalo-Nenets Autonomous Okrug	16	57	8478
Дальний Восток / Russian Far East			
Амурская область / Amur Oblast	14	98	7177
Еврейская АО / Jewish Autonomous Oblast	2	4	1040
Камчатский край / Kamchatka Krai	23	98	17364
Магаданская область / Magadan Oblast	18	185	17684
Приморский край / Primorsky Krai	18	164	29549
Сахалинская область / Sakhalin Oblast	10	70	20170
Хабаровский край / Khabarovsk Krai	22	114	21355
Чукотский АО / Chukotka Autonomous Okrug	60	294	43891

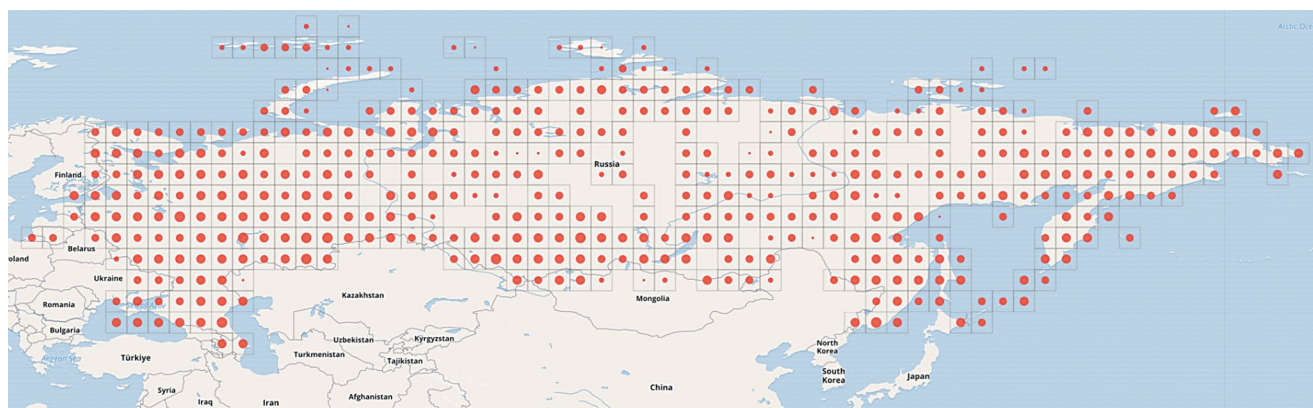


Рис. 1. Пространственное покрытие базы данных “Локальные флоры России” по квадратам 250 × 250 км (на 04.04.2024) (источник: Bochkov, Seregin, 2024).

Fig. 1. Spatial coverage in the “*Local Floras of Russia*” dataset across 250×250 km grid squares (as of 04.04.2024) (source: Bochkov, Seregin, 2024).

пропуски отдельных квадратов в Якутии, Красноярском крае, отдельных местностях севера Дальнего Востока и Западной Сибири, по которым нам не удалось найти ни одного опубликованного источника. Подробнее о “белых пятнах” написано в разделе “Обсуждение”.

Таксономический охват базы установлен по автоматически нормированным названиям в GBIF-выгрузке. В ней номенклатура стандартизирована согласно GBIF Backbone Taxonomy (GBIF Secretariat, 2023). В общей сложности для флоры России учтены находки 10605 видов из 1529 родов и 127 семейств.

Больше всего отдельных находок учтено для следующих видов: *Deschampsia cespitosa* (L.) P.Beauv. (1711 записей), *Equisetum arvense* L. (1613), *Caltha palustris* L. (1595), *Calamagrostis purpurea* (Trin.) Trin. (1589), *Vaccinium vitis-idaea* L. (1574), *Chamaenerion angustifolium* (L.) Scop. (1488), *Rhododendron tomentosum* (Stokes) Harmaja [= *Ledum palustre* L.] (1327), *Urtica dioica* L. (1278), *Poa pratensis* L. (1215), *Galium boreale* L. (1202), *Cotmarum palustre* L. (1192), *Galium verum* L. (1155), *Eriophorum angustifolium* L. (1126), *Elymus repens* (L.) Gould (1125), *Orthilia secunda* (L.) House (1106), *Tanacetum vulgare* L. (1100), *Polygonum aviculare* L. (1097), *Epilobium palustre* L. (1094), *Betula pendula* Roth (1087), *Vaccinium uliginosum* L. (1087). Это в основном широко распространенные бореальные виды. Их преобладание связано с решением основной задачи – покрыть оцифрованными данными таежные и горно-таежные территории с низкой плотностью информации.

Ведущими родами по числу находок являются: *Carex* L. (33800), *Salix* L. (16289), *Poa* L. (10327), *Artemisia* L. (10290), *Ranunculus* L. (9891), *Potentilla* L. (8378), *Silene* L. (7435), *Viola* L. (7020), *Equisetum* L. (6878), *Galium* L. (6737), *Stellaria* L. (6709), *Juncus* L. (6380), *Pedicularis* L. (6004), *Rumex* L. (5916), *Veronica* L. (5873), *Calamagrostis* Adans. (5648), *Festuca* L. (5142), *Trifolium* L. (5125), *Astragalus* L. (5065), *Eriophorum* L. (4805).

На уровне семейств максимальное число находок учтено для: Asteraceae (78401), Poaceae (68089), Cyperaceae (44281), Rosaceae (38995), Fabaceae (30567), Caryophyllaceae (30287), Brassicaceae (29277), Ranunculaceae (28882), Salicaceae (18771), Apiaceae (17688), Ericaceae (17424), Lamiaceae (16873), Polygonaceae (16665), Plantaginaceae (13669), Orobanchaceae (12325), Juncaceae (11271), Amaranthaceae (9890), Boraginaceae (9594), Saxifragaceae (9202), Orchidaceae (8298).

Точность привязки менялась по мере включения новых данных. К апрелю 2022 г. она составила около 21000 м (среднее значение для 561 тыс. записей), а спустя два года достигла 28000 м (среднее значение для 682 тыс. записей) за счет привязки списков с точностью до административных районов по ряду регионов. Инструментарий GBIF позволяет фильтровать находки по точности привязки, отсекая те из них, которые, по мнению пользователя, обладают невысокой точностью.

Для создания карт “Атласа флоры России”, который использует квадраты 100 × 100 км, мы в дальнейшем берем только географические пункты с точностью привязки менее 100000 м.

Охват по публикациям. Проработка флористической литературы за последние 50 лет выявила несколько публикаций, в которых суммирован максимальный объем флористической информации (табл. 2). Охарактеризуем типовые массивы данных, полученные в результате оцифровки книг, диссертаций и статей на примере десяти публикаций с наибольшим числом полученных записей. Их можно разделить на две группы.

С одной стороны, распространение видов дано по пунктам (Naumenko, 2003; Antipova, 2012) или небольшим речным бассейнам (Sheremetova, 2011; Prokopenko, 2014) в публикациях, основанных преимущественно на оригинальных данных.

С другой стороны, в обобщающих публикациях распространение видов дано по природным или

административным районам, которые без остатка покрывают всю изученную территорию региональной размерности (Alekseev et al., 1988, 1989; Krasnoborov et al., 2003; Murtazaliev, 2009; Kulikov, 2010; Stepanov, 2016; Shcherbakov et al., 2022).

В диссертации Н.И. Науменко (Naumenko, 2003) в виде таблицы в приложении перечислено распространение видов по 44 локальным флорам *Южного Зауралья* (в том числе 41 флора в пределах России – Курганская, Тюменская, Челябинская, Свердловская области). Сведения о распространении 1234 видов в общей сложности дали 23850 отдельных находок, в том числе по 306 видам, известным во всех локальных флорах.

Определитель, охватывающий флору *Алтайского края* (Krasnoborov et al., 2003), содер-

Таблица 2. Источники, внесшие максимальный вклад по числу оцифрованных находок (на 04.04.2024)
Table 2. References with the largest contribution by the number of digitized records (as of 04 Apr 2024)

Источник / Reference	Территория / Area	Записей / Records
Naumenko (2003)	Степное Зауралье / Steppe Trans-Urals	23850
Krasnoborov et al. (2003)	Алтайский край / Altai Krai	22357
Alekseev et al. (1988, 1989)	Башкортостан / Republic of Bashkortostan	19156
Antipova (2012)	Юг Красноярского края / Krasnoyarsk Krai (south)	13125
Kulikov (2010)	Челябинская область / Chelyabinsk Oblast	12663
Shcherbakov et al. (2022)	Ивановская область / Ivanovo Oblast	12651
Stepanov (2016)	Западные Саяны / Western Sayans	12459
Murtazaliev (2009)	Дагестан / Dagestan	11549
Sheremetova (2011)	Бассейн р. Томь / Tom River Catchment	10342
Prokopenko (2014)	Находка (Приморье) / Nakhodka (Primorsky Krai)	9481
Shmidt (2005)	Архангельская область / Arkhangelsk Oblast	9338
Martynenko et al. (2008)	Республика Коми / Komi Republic	9081
Barkalov (2009)	Курильские острова / Kuril Islands	7813
Barkalov & Taran (2004)	Сахалин / Sakhalin	7577
Verkhovina (2004)	Иркутская область / Irkutsk Oblast	7502
Taysumov & Omarkhadzhieva (2012)	Чеченская Республика / Chechnya	7249
Iminskikh (2014)	Города Поволжья / Middle Volga cities	6815
Drachev (2010)	Тюменская область / Tyumen Oblast	5856
Tzvelev (2000)	Северо-Запад / European Russia North-West	5569
Yakubov & Chernyagina (2004)	Камчатка / Kamchatka	5251
Busyina (2009)	Центральная Якутия / Central Yakutia	5058

жит в информацию о распространении растений по 61 административному району края со средней точностью геопривязки к центроидам районов равной 42700 м. Это позволило мобилизовать из этого источника сведения по 22357 находкам с точностью до района: 11072 точных указаний для отдельных районов и 11285 указаний для 185 видов, указанных для всех районов края. Эти сведения (точное указание для района или для видов, распространенных повсеместно) также хранятся в базе в стандартном поле “Georeference remarks”.

Важным источником стал определитель растений *Башкирии* (Alekseev et al., 1988, 1989), в котором распространение видов дано по 25 природным районам, обозначенным цифрами. Районы имеют довольно большую площадь, поэтому средняя точность привязки этого массива составляет 85160 м. Всего в этой работе содержится 19156 указаний для отдельных районов: 14606 точных указаний для отдельных районов по 129 видам и 4550 указаний по 182 видам, которые указаны для всех районов республики. Еще 70 видов, указанных в этом источнике, по разным причинам не привязаны к сетке районов.

В монографии по участкам островной лесостепи юга *Красноярского края* (Antipova, 2012) аннотации видов содержат указание на распространение в пределах 26 локальных флор, насчитывающих от 300 до 1095 видов. В общей сложности оцифровано 13125 находок по 1570 таксонам. В дальнейшем список одной из локальных флор (Красноярск) взят по более позднему источнику (Stepanov, 2016).

П.В. Куликов подвел итоги изучения флоры *Челябинской области* в конспекте (Kulikov, 2005), в котором приводятся обширные фактические материалы по сосудистым растениям региона. Однако для наших целей мы оцифровали данные о распространении видов по 17 природным выделам (14 районов и 4 подрайона) по более позднему определителю автора (Kulikov, 2010), содержащему множество дополнений и исправлений. Средняя точность привязки по области составила 86170 м, а общий объем оцифрованных данных — 12663 находки (8804 точных указаний по отдельным выделам и 3859 указаний по 227 видам для всех районов Челябинской области).

Отдельно следует сказать о таком табличном источнике как флора *Ивановской области*

(Shcherbakov et al., 2022). Эта работа — пока единственная в серии однотипных работ по регионам бассейна Оки, которая была оцифрована для “Атласа флоры России”. В Средней России площадь административных районов обычно составляет около 1000 км², поэтому средняя точность привязки к центроиду районов в этом массиве составила 25600 м.

Идеологической основой этой публикации стала таблица распространения видов по административным районам, опубликованная в приложении к “Флоре Владимирской области” (Seregin, 2012). Начиная с Московской области (Shcherbakov, Lyubeznova, 2018), региональные группы исследователей под общим руководством А.В. Щербакова опубликовали таблицы с распространением видов по административным районам Орловской, Ивановской, Калужской, Пензенской, Нижегородской областей. Эта работа продолжается. К сильным сторонам этих таблиц относится высокая степень структурированности материала, которая идеально подходит для наших целей, и большой объем оригинального полевого материала, который ранее не публиковался. К слабым сторонам — отсутствие отметок для адвентивных видов (и степени их натурализации), отсутствие таксономических авторов в ряде таблиц. В списках для Москвы, Орла и Нижнего Новгорода исключительно много сугубо культурных видов, которые также никак не обозначены.

Монография Н.В. Степанова (Stepanov, 2016) охватывает *Приенисейские Саяны* в пределах юга Красноярского края и приграничных районов Тувы. Распространение видов дано по 12 выделам (10 больших природных районов и отдельно для национального парка “Столбы” и города Красноярска) со средней точностью геопривязки массива 91170 м. Всего по этому источнику учтено 12459 записей.

В четырехтомнике по флоре *Дагестана* (Murtazaliev, 2009) распространение видов дано по 13 природным районам (иногда используются характеристики “все районы” или “все горные районы” для широко распространенных видов). Флора республики — самая богатая региональная флора России (Seregin et al., 2020), однако тут было описано всего две локальные флоры (Stepanova, 2012; Adjieva, 2015). В связи с этим необходимость оцифровки конспекта с сотнями видов, которые в России известны только в Дагестане,

не вызывала сомнения. В общей сложности из конспекта оцифровано 11549 записей со средней точностью 64900 м.

В обзорной статье С.А. Шереметовой (Sheremetova, 2011) список сосудистых растений *бассейна р. Томи* приведен в виде таблицы с находками видов по бассейнам малых рек (Кемеровская и Томская области). Это самая полная из учтенных публикаций, основанная на бассейновом подходе. Она содержит находки 1522 видов по 19 притокам Томи в табличном виде. Это 10342 отдельных записи без учета находок, сделанных на самой Томи.

В статье С.В. Прокопенко (Prokopenko, 2014) территория площадью 187 км² в *окрестностях Находки* подробно охарактеризована в пространственном отношении. “Матрешкой” дано распространение 1056 видов по 20 бассейновым районам, объединенным в три группы районов. Для целей “Атласа флоры России” важно и то, что автор дал характеристики того, какие виды на юге Приморья являются заносными, а какие — элементом местной флоры.

ОБСУЖДЕНИЕ

Анализ флор. Одной из ключевых проблем, которая встала перед нами, стало отсутствие публикаций со списками видов по тем пунктам и территориям, по которым опубликованы так называемые анализы флор. Эти краткие публикации (как правило, в журналах и материалах конференций) однообразного содержания содержат информацию о том, сколько видов в ведущих семействах, какие преобладают географические элементы, жизненные формы и проч. Иногда может быть упомянуто несколько видов в кратком обзоре растительности района исследований или среди наиболее примечательных находок. Ничего кроме досады, что человек работал (и зачастую хорошо работал) в каком-нибудь интересном месте, но не удосужился при этом опубликовать итоговый список локальной флоры, это не вызывает. Именно в списках видов, на наш взгляд, содержится ценная фактическая информация о встречах как редких, так и обычных видов, в то время как безликие анализы флор имеют нулевые цитирования.

К счастью, эта ситуация постепенно меняется. Журналы стали более многочисленными, а их перевод на электронные носители привел к увели-

чению объема возможных статей для публикации флористических списков. Впрочем, ряд журналов все еще имеет жесткие требования по объему публикаций, из-за чего, например, бедные арктические локальные флоры регулярно публикуются в “Ботаническом журнале”, а более богатые флоры уже не влезают в объем статей. В любом случае, мы считаем публикацию каких-либо анализов флоры при отсутствии исходных данных бессмысленным занятием, поскольку научный результат должен быть верифицируемым, что невозможно сделать без исходных материалов.

Флористические находки. Другим бичом современных журнальных публикаций по флористике стало тотальное преобладание среди статей флористов сведений о всевозможных флористических находках. Это могут быть как редкие и охраняемые виды природной флоры, так и заносные растения.

Более того, сформировались целые научные группы, которые, по сути, занимаются “кладоискательством”, т. е. целенаправленно (хотя и бессистемно) ищут редкие заносные виды вдоль путей сообщения, на свалках, в городах, публикуя лишь некоторые находки новых сорняков для того или иного региона. Экспедиционные исследования носят в основном маршрутный характер и по их итогам тоже публикуются лишь редкости. Многочисленные работы на ООПТ ограничиваются, как правило, выявлением охраняемых видов.

Напротив, обобщающих флористических работ или полных флористических списков по результатам полевых исследований публикуется совсем немного, несмотря на многочисленность сообщества. Возникает ощущение, что научная деятельность часто сводится к публикации находок нескольких видов для очередного ежегодного отчета.

Белые пятна. Наиболее значительным белым пятном в России, по которому совсем нет флористических данных, является территория площадью около 250000 км² в таежной зоне на границе Красноярского края, Иркутской области и Якутии (между 61° и 70° с.ш. и между 100° и 108° в.д.). На востоке она ограничена серией локальных флор, описанных вдоль 108° в.д. (Vodopianova, 1984), и маршрутными данными с Верхнего Вилюя (Kildyushevskii, 1964). На юге — флорой Тунгусского заповедника (ООРТ, 2024, по:

Vasilyev et al., 2003; Timoshok et al., 2008). На западе – списками для окрестностей Туры (Shemberg, 2017), оз. Чирингда – р. Ланга (Mironenko et al., 1971). На севере – Хайя-Кюель (Flora..., 2007–2024, по: Flora..., 1976) и Верхней Меркю (Flora..., 2007–2024). Судя по описанному соседним флорам, это флористически бедный район с преобладанием зональной листовенничной тайги. Труднодоступность территории и отсутствие интересных объектов делает ее малопривлекательной, однако сам по себе размер этого белого пятна впечатляет.

Отдельные белые пятна меньшего размера отмечены в Якутии (Индигирка, Средняя Колыма, Средний Оленек, Средний Вилюй), на севере Хабаровского края, в таежно-болотных районах Томской и Свердловской областей.

Недоступность отдельных источников – нечастая, но типовая проблема. Система обязательных экземпляров в 1990-е гг. дала сбой, и часть региональных малотиражных изданий не представлены в центральных библиотеках, а система оцифровки краеведческих фондов налажена далеко не в каждой областной библиотеке. На данный момент наличие многочисленных научных журналов в интернете во многом снимает проблему ограниченного доступа к научной информации.

Частичное огрубление данных – к сожалению, распространенная проблема. В некоторых работах региональной размерности (флоры, конспекты, определители) распространение части видов дано по четко перечисленным районам, а для растений массовых дано указание “во всех районах” или “повсеместно”.

Зачастую такая пометка ставится безосновательно также для малоизвестных и (или) случайно пропущенных видов. Кроме того, это способ скрыть отсутствие данных из малоизученных районов. В итоге до половины видов якобы присутствуют во всех районах. Статистический анализ работ, где распространение видов дано по районам для всех видов (в том числе для обычных), показывает, что даже в небольших регионах со слабой ландшафтной дифференциацией число повсеместных видов редко превышает 30% из общего состава региональной флоры – до 31.8% по данным сеточного картирования флоры Владимирской области (Seregin, 2012). В регионах с зональными границами растительности и горными массивами эта доля падает – вплоть до 2.7%

в Дагестане (Murtazaliev, 2009), где менее 100 видов зарегистрированы во всех 13 природных районах республики.

Проблема доверия к данным. Данные литературы – это такой источник информации о распространении растений, где, по сути, единственным мериллом его надежности является авторство публикации. Да, хорошая флористическая работа сопровождается обильными гербарными сборами, которые затем передаются (или не передаются) в большие и малые гербарные коллекции. Где-то они затем переводятся в электронный вид и исходные данные автора становятся общедоступными. Но в целом о работе того или иного флориста мы знаем по его публикациям.

Обработанный массив флористических публикаций, безусловно, содержит ошибки и неточности, аккуратно оцифрованные из первоисточников. Тем не менее после работы над редактированием созданных нами карт “Атласа флоры России” мы вынуждены были целиком исключить несколько работ, которые содержали большое число видов, в принципе не произрастающих в изученных местностях. На наш взгляд, основной причиной этих “лишних” указаний стала неверная интерпретация в локальных работах более крупных флор региональной размерности.

Отсутствие пометок, какие виды были реально встречены авторами на основании оригинальных исследований, а какие растения указаны лишь по литературе (из самых общих соображений), привело к тому, что нам пришлось целиком убрать из базы четыре списка (Chimonina, 2004; Fedyaeva, 2004; Ivanov, 2004; Turchin et al., 2004).

Что осталось сделать? Мы продолжаем работу по переводу массива флористической литературы в электронный вид. Предстоит последовательно заменить данные из агрегатора по биоразнообразию ООПТ (ООРТ, 2024) на опубликованные списки по отдельным заповедникам и национальным паркам; добавить публикации по Средней России и Поволжью, в том числе таблицы А.В. Щербакова и др.; получить сканы некоторых недоступных малотиражных публикаций и диссертаций; отслеживать новейшие публикации. Архив имеющихся публикаций позволит создать массив флористических данных общим объемом свыше 1 млн записей.

ЗАКЛЮЧЕНИЕ

Проработка обширного массива публикаций позволяет нам дать несколько советов авторам работ с полными флористическими списками.

1. Не тратьте время на долгие подсчеты “анализа флоры”, когда полный флористический список не опубликован; сам список содержит гораздо больше информации, чем любая статистика на его основе.

2. Не стоит сокращать названия родов. В эпоху, когда информацию пользователи ищут через интернет, при запросе “*Trifolium repens*” вряд ли будет найден упомянутый вами “*T. repens*”.

3. Не допускайте случайного использования кириллицы в латинских названиях (как в публикациях, так и в электронных массивах информации).

4. Приводите карту района исследований и пунктов работ.

5. Не смешивайте оригинальные данные со сведениями из литературы, онлайн-источников и гербарных коллекций; любой фрагмент заимствованной информации должен иметь четкую и ясную ссылку в списке литературы.

6. Обязательно указывайте, какие виды в вашем списке являются заносными, а какие — элементами природной флоры.

7. Если в одной публикации обобщаются данные по нескольким пунктам или районам, то стоит аккуратно приводить каждый из районов — пометка “во всех районах” не приветствуется.

8. Крупные региональные работы (аннотированные конспекты, определители) стоит дополнять таблицами с распространением видов по районам.

9. Своевременно обрабатывайте, этикетировать и определяйте гербарные коллекции; передавайте их в крупные гербарии, где они будут своевременно оцифрованы и включены в основной фонд.

10. Размещайте фотоматериалы, документирующие находки отдельных видов, в открытом доступе и делайте бэкапы, иначе ваши фотонаблюдения безвозвратно исчезнут вместе с носителем данных.

БЛАГОДАРНОСТИ

Работа поддержана грантом РНФ № 21-77-20042. Отдельные источники, отсутствующие в Интернете,

нам в разное время присылали: В.А. Канев (Сыктывкар), Н.В. Филиппова (Ханты-Мансийск), Д.Н. Шауло (Новосибирск), С.А. Шереметова (Кемерово), И.Н. Поспелов (Москва), Н.Н. Панасенко (Брянск), И.В. Кузьмин (Тюмень). Выражаем им глубокую благодарность.

СПИСОК ЛИТЕРАТУРЫ

- [Adjieva] Аджиева А.И. 2015. Конспект флоры сосудистых растений массива Сарыкум (Дагестан). — Бот. журн. 100(12): 1298–1310.
<https://doi.org/10.1134/S0006813615120054>
- [Alekseev et al.] Алексеев Ю.Е., Алексеев Е.Б., Габбасов К.К., Горчаковский П.Л., Губанов И.А., Гуфранова И.Б., Кузяхметов Г.Г., Кулагин Ю.З., Кучеров Е.В., Минибаев Р.Г., Наумова Л.Г., Назирова З.М., Шурова Е.А., Хайретдинов С.С. 1988. Определитель высших растений Башкирской АССР. Ч. 1. М. 316 с.
https://herba.msu.ru/shipunov/school/books/opred_vys_rast_bashkirii1988_1.djvu
- [Alekseev et al.] Алексеев Ю.Е., Галеева А.Х., Губанов И.А., Гуфранова И.Б., Жирнова Т.В., Князев М.В., Кулагин Ю.З., Кулаковская Л.А., Культиасов И.М., Кучеров Е.В., Михайлова Т.П., Мулдашев А.А., Скворцов А.К., Тихомиров В.Н., Шурова Е.А. 1989. Определитель высших растений Башкирской АССР. Ч. 2. М. 375 с.
https://herba.msu.ru/shipunov/school/books/opred_vys_rast_bashkirii1989_2.djvu
- [Antipova] Антипова Е.М. 2012. Флора внутриконтинентальных островных лесостепей Средней Сибири. Красноярск. 662 с.
https://herba.msu.ru/shipunov/school/books/antipova2012_flora_vnutrikont_ostrovn_lesost_sred_sibiri.pdf
- Artemov I., Egorov A. 2021. Locations of plants on dot distribution maps in the Flora of Siberia (Flora Siberiae, 1987–1997). Version 1.2. Central Siberian Botanical Garden SB RAS. Occurrence dataset accessed via GBIF.org on 2024-04-04.
<https://doi.org/10.15468/jb84wg>
- [Barkalov] Баркалов В.Ю. 2009. Флора Курильских островов. Владивосток. 468 с.
https://herba.msu.ru/shipunov/school/books/barkalov2009_flora_kuril.djvu
- [Barkalov, Taran] Баркалов В.Ю., Таран А.А. 2004. Список видов сосудистых растений острова Сахалин. — В сб.: Растительный и животный мир острова Сахалин (Материалы Международного сахалинского проекта). Ч. 1. Владивосток. С. 39–66.
<https://www.biosoil.ru/storage/entities/publication/309/00000309.pdf>
- [Belikovich, Galanin] Беликович А.В., Галанин А.В. 1992. Растительный покров Магаданского геоботанического района. — Комаровские чтения. 39: 26–79.
<https://www.biosoil.ru/storage/entities/fscpublication/701/3713f41c-9312-4036-926a-68f9d2ed04ef.pdf>

- Bochkov D.A., Seregin A.P. 2024. Local floras of Russia: records from literature. Version 1.75. Lomonosov Moscow State University. Occurrence dataset accessed via GBIF.org on 2024-04-04. <https://doi.org/10.15468/rxtjt2>
- [Buks] Букс И.И. 1964. О некоторых новых местонахождениях фрагментов степной растительности в Якутском Заполярье. — Бюллетень МОИП. Отд. биол. 69(2): 135–137.
- [Bysyina] Бысыина М.Ф. 2009. Флора аласной части Лено-Амгинского междуречья (Центральная Якутия): Дисс. ... канд. биол. наук. Томск. 228 с.
- [Charkevicz et al.] Харкевич С.С., Буч Т.Г., Якубов В.В., Яшенкова Г.Ф. 1983. Материалы к изучению флоры Аяно-Майского района Хабаровского края. — Новости систематики высших растений. 20: 203–224. http://herba.msu.ru/shipunov/school/books/novosti_sistematiki_visshih_rasteniy_1983_20.djvu
- [Chimonina] Чимонина И.В. 2004. Флора Прикалауского флористического района (Центральное Предкавказье) и ее анализ. Дисс. ... канд. биол. наук. Ставрополь. 221 с.
- [Czerepanov] Черепанов С.К. 1973. Свод дополнений и изменений к “Флоре СССР” (тт. I–XXX). Л. 668 с. https://herba.msu.ru/shipunov/school/books/czerepanov1973_svod.djvu
- [Czerepanov] Черепанов С.К. 1981. Сосудистые растения СССР. Л. 510 с. https://herba.msu.ru/shipunov/school/books/czerepanov1981_sosud_rast_ssr.djvu
- [Czerepanov] Черепанов С.К. 1995. Сосудистые растения России и сопредельных государств (в пределах бывшего СССР). Русское издание. СПб. 992 с. https://herba.msu.ru/shipunov/school/books/czerepanov1995_sosud_rast_rus.djvu
- [Drachev] Драчев Н.С. 2010. Флора подзоны южной тайги в пределах Тюменской области. Дисс. ... канд. биол. наук. Новосибирск. 331 + 307 с.
- [Fedyayeva] Федяева В.В. 2004. Травяная флора сосудистых растений. — В сб.: Флора, фауна и микобиота государственного музея-заповедника М.А. Шолохова. Ростов-на-Дону. С. 24–74. http://herba.msu.ru/shipunov/school/books/fl_fauna_mikobiota_muz_zapov_sholokhova_2004.djvu
- [Flora...] Флора Путорана. Материалы к познанию особенностей состава и генезиса горных субарктических флор Сибири. 1976. Новосибирск. 245 с. https://herba.msu.ru/shipunov/school/books/flora_putorana_1976.djvu
- [Flora...] Флора Таймыра. 2007–2024. <http://byrranga.ru/>
- [Fundamentalnaya...] Фундаментальная электронная библиотека “Флора и фауна”: растения, животные, грибы и водоросли, теория эволюции и систематики. 2024. <https://herba.msu.ru/shipunov/school/sch-ru.htm>
- GBIF Secretariat. 2023. GBIF Backbone Taxonomy. Checklist dataset accessed via GBIF.org on 2024-04-05. <https://doi.org/10.15468/39omei>
- Global Biodiversity Information Facility. 2024. <https://www.gbif.org/>
- [Ilminskikh] Ильминских Н.Г. 2014. Флорогенез в условиях урбанизированной среды. Екатеринбург. 470 с.
- iNaturalist contributors, iNaturalist. 2024. iNaturalist Research-grade Observations. Occurrence dataset accessed via GBIF.org on 2024-04-04. <https://doi.org/10.15468/ab3s5x>
- [Ivanov] Иванов А.А. 2004. Флора Ставропольских высот и ее анализ: Дисс. ... канд. биол. наук. Ставрополь. 224 с.
- Jain S., De Buitler A., Fallon E. 2020. A review of unstructured data analysis and parsing methods. — International Conference on Emerging Smart Computing and Informatics (ESCI). Pune, India. P. 164–169.
- [Kildyushevskii] Кильдюшевский И.Д. 1964. К флоре верховьев Вилюя. — В сб.: Леса Южной Якутии. М. С. 148–193.
- [Krasnoborov et al.] Красноборов И.М., Ломоносова М.Н., Шауло Д.Н., Куцев М.Г., Красников А.А. и др. 2003. Определитель растений Алтайского края. Новосибирск. 634 с. https://herba.msu.ru/shipunov/school/books/opred_rast_altajsk_kraja_2003_v2.djvu
- [Kulikov] Куликов П.В. 2005. Конспект флоры Челябинской области (сосудистые растения). Миасс. 537 с. https://herba.msu.ru/shipunov/school/books/kulikov2005_konsp_fl_chel_obl.pdf
- [Kulikov] Куликов П.В. 2010. Определитель сосудистых растений Челябинской области. Екатеринбург. 969 с. https://herba.msu.ru/shipunov/school/books/kulikov2010_opred_sosud_rast_cheljabinsk_obl.djvu
- [Lukicheva] Лукичева А.Н. 1963. Растительность северо-запада Якутии и ее связь с геологическим строением местности. М., Л. 168 с.
- [Martynenko et al.] Мартыненко В.А., Груздев Б.И., Канев В.А. 2008. Локальные флоры таежной зоны Республики Коми. Сыктывкар. 76 с. https://herba.msu.ru/shipunov/school/books/martynenko2008_lokal_flory_taezhn_zony_resp_komi.pdf
- [Matveyeva] Матвеева Н.В. 1998. Зональность в растительном покрове Арктики. СПб. 220 с. https://herba.msu.ru/shipunov/school/books/matveyeva1998_zonaln_rastit_arkt.pdf
- [Matveyeva, Zanakha] Матвеева Н.В., Заноха Л.Л. 2017. Изменения во флоре сосудистых растений в окрестностях пос. Диксон (Западный Таймыр) с 1980 по 2012 г. — Ботанический журнал. 102(6): 812–846. <https://doi.org/10.1134/S0006813617060072>

- [Mironenko et al.] Мироненко О.Н., Петровский В.В., Юрцев Б.А. 1971. К познанию флоры центральной части плато Путорана (Средне-Сибирское плоскогорье). — Бот. журн. 56(7): 982–988.
- [Murtazaliev] Муртазалиев Р.А. 2009. Конспект флоры Дагестана. Махачкала. Т. 1. 320 с.; Т. 2. 248 с.; Т. 3. 304 с.; Т. 4. 231 с.
- [Naumenko] Науменко Н.И. 2003. Флора южного Зауралья. Дисс. ... докт. биол. наук. СПб. 812 с.
- [Nedoluzhko, Denisov] Недолужко В.А., Денисов Н.И. 2001. Флора сосудистых растений острова Русский (залив Петра Великого в Японском море). Владивосток. 98 с.
https://herba.msu.ru/shipunov/school/books/nedoluzhko2011_flora_ostr_russkij.pdf
- Nelson G., Ellis S. 2019. The history and impact of digitization and digital data mobilization on biodiversity research. — *Philosophical Transactions of the Royal Society B*. 374 (1763): 20170391.
<https://doi.org/10.1098/rstb.2017.0391>
- [ООПТ] ООПТ России. 2024.
<http://www.oopt.aari.ru/>
- [Polezhayev, Berkutenko] Полежаев А.Н., Беркутенко А.Н. 1982. Сосудистые растения верховий реки Анадырь. — Бот. журн. 67(10): 1400–1408.
- [Pospelov, Seregin] Поспелов И.Н., Серегин А.П. 2023a. Методика создания набора данных точечных карт находок в монографии “Сосудистые растения Советского Дальнего Востока”. — *Проблемы ботаники Южной Сибири и Монголии*. 22(2): 262–267.
<https://doi.org/10.14258/pbssm.2020138>
- Pospelov I.N., Seregin A.P. 2023b. Vascular plants of the Soviet Far East (1985–2006): georeferenced records from dot maps. Version 1.3. Lomonosov Moscow State University. Occurrence dataset accessed via GBIF.org on 2024-04-04.
<https://doi.org/10.15468/pa3dv8>
- [Pospelova, Pospelov] Поспелова Е.Б., Поспелов И.Н. 2007. Флора сосудистых растений Таймыра и сопредельных территорий. Ч. 1. Аннотированный список флоры и ее общий анализ. М. 457 с.
https://herba.msu.ru/shipunov/school/books/pospelova2007_fl_sosud_rast_tajmyra_1.pdf
- [Прокопенко] Прокопенко С.В. 2014. Флора полуострова Трудный (город Находка и окрестности). — *Комаровские чтения*. 62: 106–228.
<https://www.biosoil.ru/storage/entities/fscpublication/842/a8af05b3-1f18-4f58-9224-56f7b13809c9.pdf>
- [Rebristaya] Ребристая О.В. 2013. Флора полуострова Ямал. Современное состояние и история формирования. СПб. 312 с.
https://herba.msu.ru/shipunov/school/books/rebristaya2013_fl_poluostr_jamal.djvu
- [Seregin] Серегин А.П. 2012. Флора Владимирской области: конспект и атлас. Тула. 620 с.
<https://istina.msu.ru/download/58403700/1rwd1L:ih-pCqCSliU54aVUNh1ONWTi2dYo/>
- Seregin A.P. 2021a. Contribution to the “Atlas of the Russian Flora”: Twelve local floras of European Russia. — *Biodiversity Data Journal*. 9: e73013.
<https://doi.org/10.3897/bdj.9.e73013>
- Seregin A. 2021b. Twelve local floras of European Russia. Version 1.4. Lomonosov Moscow State University. Occurrence dataset accessed via GBIF.org on 2024-04-04.
<https://doi.org/10.15468/35rwhv>
- Seregin A. 2024. Moscow University Herbarium (MW). Version 1.326. Lomonosov Moscow State University. Occurrence dataset accessed via GBIF.org on 2024-04-04.
<https://doi.org/10.15468/cpnhcc>
- Seregin A.P., Bochkov D.A., Shner J.V., Garin E.V., Pospelov I.N. et al. 2020. “Flora of Russia” on iNaturalist: a dataset. — *Biodiversity Data Journal*. 8: e59249.
<https://doi.org/10.3897/BDJ.8.e59249>
- Seregin A., Stepanova N. 2024. МНА Herbarium: collections of vascular plants. Version 1.233. Tsitsin Main Botanical Garden Russian Academy of Sciences. Occurrence dataset accessed via GBIF.org on 2024-04-04.
<https://doi.org/10.15468/827lk2>
- [Shcherbakov, Lyubeznova] Щербаков А.В., Любезнова Н.В. 2018. Список сосудистых растений московской флоры. М. 160 с.
https://herba.msu.ru/shipunov/school/books/scherbakov2018_spis_sosud_rast_mosk_fl.pdf
- [Shcherbakov et al.] Щербаков А.В., Любезнова Н.В., Борисова Е.А., Курганов А.А., Шилов М.П. 2022. Список сосудистых растений Ивановской области. М. 73 с.
- [Shemberg] Шемберг А.М. 2017. Структура растительного покрова и особенности почв листовенничных экосистем северной тайги Средней Сибири (бассейн р. Нижняя Тунгуска). Бакалаврская работа. Красноярск. 44 с.
https://elib.sfu-kras.ru/bitstream/handle/2311/34088/diplom_shemberg.pdf
- [Sheremetova] Шереметова С.А. 2011. Список сосудистых растений бассейна реки Томи. — *Ботанические исследования Сибири и Казахстана*. 17: 43–94.
- [Shmidt] Шмидт В.М. 2005. Флора Архангельской области. СПб. 346 с.
https://herba.msu.ru/shipunov/school/books/shmidt2005_flora_arch_obl.djvu
- [Stepanov] Степанов Н.В. 2016. Сосудистые растения Приенисейских Саян. Красноярск. 252 с.
https://herba.msu.ru/shipunov/school/books/stepanov2016_sosud_rast_prienisejsk_sajan.pdf
- [Stepanova] Степанова Н.Ю. 2012. Флора Кума-Маньчской впадины: Дисс. ... канд. биол. наук. М. 308 с.
- [Taisumov, Omarkhadzhieva] Тайсумов М.А., Омархаджиева Ф.С. 2012. Анализ флоры Чеченской республики. Грозный. 320 с.
https://herba.msu.ru/shipunov/school/books/tajsumov2012_analiz_flory_tchetchensk_resp.djvu
- [Timoshok et al.] Тимошок Е.Е., Скороходов С.Н., Райская Ю.Г., Тимошок Е.Н. 2008. Видовое раз-

- нообразии сосудистых растений южной части государственного природного заповедника “Тунгусский”. – В сб.: Труды ГПЗ “Тунгусский”. Вып. 2. Томск. С. 19–55.
<https://clck.ru/3A5i2h>
- [Turchin et al.] Турчин Т.Я., Турчина Т.А., Федяева В.В., Миронова Н.В. 2004. Древесная флора. – В сб.: Флора, фауна и микобиота государственного музея-заповедника М.А. Шолохова. Ростов-на-Дону. С. 14–23.
http://herba.msu.ru/shipunov/school/books/fl_fauna_mikobiota_muz_zapov_sholokhova_2004.djvu
- Turland N.J., Wiersema J.H., Barrie F.R., Greuter W., Hawksworth D.L., Herendeen P.S., Knapp S., Kubler W.-H., Li D.-Z., Marhold K., May T.W., McNeill J., Monro A.M., Prado J., Price M.J., Smith G.F. (Eds.) 2018. International Code of Nomenclature for algae, fungi, and plants (Shenzhen Code) adopted by the Nineteenth International Botanical Congress Shenzhen, China, July 2017. Glashütten. 254 p.
<https://doi.org/10.12705/Code.2018>
- [Tzvelev] Цвелев Н.Н. 2000. Определитель сосудистых растений северо-западной России (Ленинградская, Псковская и Новгородская области). СПб. 781 с.
https://herba.msu.ru/shipunov/school/books/tzvelev2000_opred_rast_sz.djvu
- [Vasilchenko] Васильченко З.А. 2003. Сосудистые растения Сохондинского заповедника (Аннотированный список видов). М. 87 с.
https://herba.msu.ru/shipunov/school/books/vasiljchenko2003_sosud_rast_sokhondinskogo_zapov_vyp_105.djvu
- [Vasilyev et al.] Васильев Н.В., Львов Ю.А., Плеханов Г.Ф., Логунова Л.Н., Мульдияров Е.Я., Бибилова В.В., Волков А.Е., Кузьмин С.Л., Лапшина Е.Д., Папанотиди А.И., Сергиева З.М., Сидоров К.С., Травинский И.В., Шефтель Б.И., Щербина С.С. 2003. Государственный природный заповедник “Тунгусский” (очерк основных данных). – В сб.: Труды ГПЗ “Тунгусский”. Вып. 1. Томск. С. 33–89.
<https://clck.ru/3A5i2B>
- [Verkhovina] Верховина А.В. 2004. Флора сосудистых растений Присяянья в пределах Иркутской области: Дисс. ... канд. биол. наук. Иркутск. 281 с.
- [Vodoriana] Водопьянова Н.С. 1984. Зональность флоры Среднесибирского плоскогорья. Новосибирск. 157 с.
https://herba.msu.ru/shipunov/school/books/vodopjanova1984_zon_flory_vost_sib.pdf
- [Yakubov, Chernyagina] Якубов В.В., Черныгина О.А. 2004. Каталог флоры Камчатки (сосудистые растения). Петропавловск-Камчатский. 160 с.
https://herba.msu.ru/shipunov/school/books/jakubov_chernyagina2004_katalog_fl_kamchatki.pdf
- [Zernov] Зернов А.С. 2002. Определитель сосудистых растений севера Российского Причерноморья. М. 283 с.
https://herba.msu.ru/shipunov/school/books/zernov2002_opred_sosud_rasd_sev_ross_pricher-nomorja.djvu

“LOCAL FLORAS OF RUSSIA” DATASET IN OPEN ACCESS

A. P. Seregin^{1, *}, D. A. Bochkov¹, K. Y. Marchenkova², Y. O. Magazov³,
 S. V. Dudov¹, V. V. Chepinoga^{4, 5}

¹*Lomonosov Moscow State University
 Leninskiye Gory, 1, Moscow, 119991, Russia*

²*MBCI “CLS” of the City of Seltso, Bryansk Region
 Kuibysheva Str., 30, Seltso, Bryansk Region, 241550, Russia*

³*Chelyabinsk Children and Youth Camping Trip Centre “Kosmos”
 Kulibina Str., 54, Chelyabinsk, 454085, Russia*

⁴*Leibniz University
 Nienburger Str., 17, Hannover, 30167, Germany*

⁵*Irkutsk State University
 Karl Marks Str., 1, Irkutsk, 664003, Russia*

*e-mail: botanik.seregin@gmail.com

For the purpose of the *Atlas of the Russian Flora* production, we converted data from floristic checklists published in 636 sources from 1975 to 2023 into e-form. We gave preference to complete floristic checklists, however for some areas with low data density partial lists, route records, vegetation relevés and reviews, and in some cases, scattered publications with floristic records also served as data sources. As of April 4, 2024, the *Local Floras of Russia* dataset (doi: 10.15468/rxtjt2) contains 682,130 plant records from 3,297 geographic locations. The geographic coordinates and georeferencing accuracy are indicated for all species

within each floristic list. The average georeferencing accuracy is 28 km. Open access to the database is provided through the GBIF, which allows users to download it either as a whole or in parts, in tabular form. New sources are continuously incorporated into the database.

Keywords: spatial data, plant geography, bibliography, vascular plants

ACKNOWLEDGEMENTS

The work was supported by the Russian Science Foundation grant No. 21-77-20042. Some references unavailable on the web were sent us by V.A. Kanev (Syktyvkar), N.V. Filippova (Khanty-Mansiysk), D.N. Shaulo (Novosibirsk), S.A. Sheremetova (Kemerovo), I.N. Pospelov (Moscow), N.N. Panasenko (Bryansk), and I.V. Kuzmin (Tyumen). We express our deep gratitude to them.

REFERENCES

- Adjieva A.I. 2015. Checklist of vascular plant flora of the Sarykum Massif (Dagestan). – *Bot. Zhurn.* 100(12): 1298–1310 (In Russ.).
<https://doi.org/10.1134/S0006813615120054>
- Alekseev Y.E., Alekseev E.B., Gabbasov K.K., Gorchakovskii P.L., Gubanov I.A., Gufranova I.B., Kuzyakhmetov G.G., Kulagin Y.Z., Kucherov E.V., Minibaev R.G., Naumova L.G., Nazirova Z.M., Shurova E.A., Khairetdinov S.S. 1988. Key to higher plants of the Bashkir Autonomous Soviet Socialist Republic. Part 1. Moscow. 316 p. (In Russ.).
https://herba.msu.ru/shipunov/school/books/opred_vys_rast_bashkirii1988_1.djvu
- Alekseev Y.E., Galeeva A.Kh., Gubanov I.A., Gufranova I.B., Zhirnova T.V., Knyazev M.V., Kulagin Y.Z., Kulakovskaya L.A., Kultiasov I.M., Kucherov E.V., Mikhailova T.P., Muldashev A.A., Skvortsov A.K., Tikhomirov V.N., Shurova E.A. 1989. Key to higher plants of the Bashkir Autonomous Soviet Socialist Republic. Part 2. Moscow. 316 p. (In Russ.).
https://herba.msu.ru/shipunov/school/books/opred_vys_rast_bashkirii1989_2.djvu
- Antipova E.M. 2012. Flora of inland island forest-steppes of Central Siberia. Krasnoyarsk. 662 p. (In Russ.).
https://herba.msu.ru/shipunov/school/books/antipova2012_flora_vnutrikont_ostrovn_lesost_sred_sibiri.pdf
- Artemov I., Egorova A. 2021. Locations of plants on dot distribution maps in the Flora of Siberia (Flora Sibiraea, 1987–1997). Version 1.2. Central Siberian Botanical Garden SB RAS. Occurrence dataset accessed via GBIF.org on 2024-04-04.
<https://doi.org/10.15468/jb84wg>
- Barkalov V.Y. 2009. Flora of the Kuril Islands. Vladivostok. 468 p. (In Russ.).
https://herba.msu.ru/shipunov/school/books/barkalov2009_flora_kuril.djvu
- Barkalov V.Y., Taran A.A. 2004. A checklist of vascular plants of Sakhalin Island. – *Flora and fauna of Sakhalin Island* (Materials of the International Sakhalin Project). Part 1. Vladivostok, 39–66 p. (In Russ.).
<https://www.biosoil.ru/storage/entities/publication/309/00000309.pdf>
- Belikov A.V., Galanin A.V. 1992. Vegetation cover of the Magadan Geobotanical Region. – *Komarovskie chteniya*. 39: 26–79 (In Russ.).
<https://www.biosoil.ru/storage/entities/fscpublication/701/3713f41c-9312-4036-926a-68f9d2ed04ef.pdf>
- Bochkov D.A., Seregin A.P. 2024. Local floras of Russia: records from literature. Version 1.75. Lomonosov Moscow State University. Occurrence dataset accessed via GBIF.org on 2024-04-04.
<https://doi.org/10.15468/rxtjt2>
- Buks I.I. 1964. On some new fragments of steppe vegetation in Yakut Arctic. – *Byulleten MOIP. Otd. biol.* 69(2): 135–137 (In Russ.).
- Bysyina M.F. 2009. Flora of the alass part of the Lena-Amga Interfluve (Central Yakutia). Doct. Diss. Tomsk. 228 p. (In Russ.).
- Charkevicz S., Bucz T., Jakubov V., Jaschenkova G. 1983. *Materies ad stadium florum districti Ajano-Majanensis provinciae Chabaroviensis*. – *Novosti sistematiki vysshikh rastenii*. 20: 203–224 (In Russ., Latin).
http://herba.msu.ru/shipunov/school/books/novosti_sistematiki_vishshih_rasteniy_1983_20.djvu
- Chimonina I.V. 2004. Flora of the Prikalaus Floristic Region (Central Ciscaucasia) and its analysis. Cand. Diss. Stavropol. 221 p. (In Russ.).
- Czerepanov S.K. 1973. *Additamenta et corrigenda ad "Floram URSS" (tomi I–XXX)*. Leningrad. 668 p. (In Russ.).
https://herba.msu.ru/shipunov/school/books/czerepanov1973_svod.djvu
- Czerepanov S.K. 1981. *Plantae vasculares URSS*. Leningrad. 510 p. (In Russ., Latin).
https://herba.msu.ru/shipunov/school/books/czerepanov1981_sosud_rast_ssr.djvu
- Czerepanov S.K. 1995. *Plantae vasculares Rossicae et civitatum collimitanearum*. St. Petersburg. 992 p. (In Russ., Latin).
https://herba.msu.ru/shipunov/school/books/czerepanov1995_sosud_rast_rus.djvu
- Drachev N.S. 2010. Flora of the southern taiga subzone within the Tyumen region. Cand. Diss. Novosibirsk. 331 + 307 p. (In Russ.).
- Fedyayeva V.V. 2004. Herbal flora of vascular plants. – *Flora, fauna and mycobiota of the State Museum-Reserve M.A. Sholokhov*. Rostov-on-Don. P. 24–74 (In Russ.).

- http://herba.msu.ru/shipunov/school/books/fl_fauna_mikobiota_muz_zapov_sholokhova_2004.djvu
- Flora Putorana. Materials for understanding the characteristics of the composition and genesis of mountain subarctic floras of Siberia. 1976. Novosibirsk. 245 p. (In Russ.). https://herba.msu.ru/shipunov/school/books/flora_putorana_1976.djvu
- Flora Taimyra. 2007–2024. <http://byrranga.ru/>
- Fundamentalnaya elektronnaya biblioteka “Flora i fauna”: rasteniya, zhivotnye, griby i vodorosli, teoriya evolyutsii i sistematiki. 2024. <https://herba.msu.ru/shipunov/school/sch-ru.htm>
- GBIF Secretariat 2023. GBIF Backbone Taxonomy. Checklist dataset accessed via GBIF.org on 2024-04-05. <https://doi.org/10.15468/39omei>
- Global Biodiversity Information Facility. 2024. <https://www.gbif.org/>
- Ilminskikh N.G. 2014. Florogenesis in the urban environment. Ekaterinburg. 470 p. (In Russ.).
- Ivanov A.A. 2004. Flora of the Stavropol Heights and its analysis. Cand. Diss. Stavropol. 224 p. (In Russ.).
- Jain S., De Buitelir A., Fallon E. 2020. A review of unstructured data analysis and parsing methods. – International Conference on Emerging Smart Computing and Informatics (ESCI). Pune, India. P. 164–169.
- iNaturalist contributors, iNaturalist. 2024. iNaturalist Research-grade Observations. Occurrence dataset accessed via GBIF.org on 2024-04-04. <https://doi.org/10.15468/ab3s5x>
- Kildyushevskii I.D. 1964. To the flora of the upper reaches of the Vilyuy. – Forests of South Yakutia. Moscow. P. 148–193 (In Russ.).
- Krasnoborov I.M., Lomonosova M.N., Shaulo D.N., Kutsev M.G., Krasnikov A.A. et al. 2003. Key to plants of the Altai region. Novosibirsk. 634 p. (In Russ.). https://herba.msu.ru/shipunov/school/books/opred_rast_altajsk_kraja_2003_v2.djvu
- Kulikov P.V. 2005. Checklist of the flora of the Chelyabinsk Region (vascular plants). Miass. 537 p. (In Russ.). https://herba.msu.ru/shipunov/school/books/kulikov2005_konsp_fl_chel_obl.pdf
- Kulikov P.V. 2010. Key to vascular plants of Chelyabinsk Region. Yekaterinburg. 969 p. https://herba.msu.ru/shipunov/school/books/kulikov2010_opred_sosud_rast_cheljabinsk_obl.djvu
- Lukicheva A.N. 1963. Vegetation of the north-west of Yakutia and its connection with the geological structure of the area. Moscow, Leningrad. 168 p. (In Russ.).
- Martynenko V.A., Gruzdev B.I., Kanev V.A. 2008. Local floras of the taiga zone of the Komi Republic. Syktyvkar. 76 p. (In Russ.). https://herba.msu.ru/shipunov/school/books/martynenko2008_lokal_flory_taezhn_zony_resp_komi.pdf
- Matveyeva N.V. 1998. Zonation in plant cover of the Arctic. St. Petersburg. 220 p. (In Russ.).
- https://herba.msu.ru/shipunov/school/books/matveeva1993_zonaln_rastit_arkt.pdf
- Matveyeva N.V., Zanakha L.L. 2017. Changes in vascular flora in Dickson settlement vicinity (Western Taymyr) in between 1980 and 2012 years. – Bot. Zhurn. 102(6): 812–846 (In Russ.). <https://doi.org/10.1134/S0006813617060072>
- Mironenko O.N., Petrovskii V.V., Yurtsev B.A. 1971. A contribution to the knowledge of the flora in the central part of Putorana Plateau (Central Siberian Plateau). – Bot. Zhurn. 56(7): 982–988 (In Russ.).
- Murtazaliev R.A. 2009. Checklist of the flora of Dagestan. Makhachkala: Vol. 1. 320 p.; Vol. 2. 248 p.; Vol. 3. 304 p.; Vol. 4. 231 p.
- Naumenko N.I. 2003. Flora of the southern Trans-Urals. Doct. Diss. St. Petersburg. 812 p.
- Nedoluzhko V.A., Denisov N.I. 2001. Vascular flora of the Russky Island (Peter the Great Gulf, Sea of Japan). Vladivostok. 98 p. (In Russ.). https://herba.msu.ru/shipunov/school/books/nedoluzhko2011_flora_ostr_russkij.pdf
- Nelson G., Ellis S. 2019. The history and impact of digitization and digital data mobilization on biodiversity research. – Philosophical Transactions of the Royal Society B. 374(1763): 20170391. <https://doi.org/10.1098/rstb.2017.0391>
- OOPT Rossii. 2024. <http://www.oopt.aari.ru/>
- Polezhayev A.N., Berkutenko A.N. 1982. Vascular plants of the upper reaches of the Anadyr River. – Bot. Zhurn. 67(10): 1400–1408 (In Russ.).
- Pospelov I.N., Seregin A.P. 2023a. Methodology for creating a dataset of dot maps of records in the monograph “Vascular Plants of the Soviet Far East”. – Problemy botaniki Yuzhnoi Sibiri i Mongolii. 22(2): 262–267 (In Russ.). <https://doi.org/10.14258/pbssm.2020138>
- Pospelov I.N., Seregin A.P. 2023b. Vascular plants of the Soviet Far East (1985–2006): georeferenced records from dot maps. Version 1.3. Lomonosov Moscow State University. Occurrence dataset accessed via GBIF.org on 2024-04-04. <https://doi.org/10.15468/pa3dv8>
- Pospelova E.B., Pospelov I.N. 2007. Flora of vascular plants of Taimyr and adjacent territories. Part 1. Annotated checklist of flora and its general analysis. Moscow. 457 p. (In Russ.). https://herba.msu.ru/shipunov/school/books/pospelova2007_fl_sosud_rast_tajmyra_1.pdf
- Prokopenko S.V. 2014. Vascular flora of the Trudnyi Peninsula (Nakhodka City and its vicinity). – Komarovskie chteniya. 62: 106–228 (In Russ.). <https://www.biosoil.ru/storage/entities/fscpublication/842/a8af05b3-1f18-4f58-9224-56f7b13809c9.pdf>
- Rebristaya O.V. 2013. Flora of the Yamal Peninsula. Modern state and history of formation. St. Petersburg. 312 p. (In Russ.). https://herba.msu.ru/shipunov/school/books/rebristaya2013_fl_poluostr_jamal.djvu

- Shcherbakov A.V., Lyubeznova N.V. 2018. List of vascular plants of Moscow flora. Moscow. 160 p. (In Russ.). https://herba.msu.ru/shipunov/school/books/scherbakov2018_spis_sosud_rast_mosk_fl.pdf
- Shcherbakov A.V., Lyubeznova N.V., Borisova E.A., Kurganov A.A., Shilov M.P. 2022. List of vascular plants of the Ivanovo region. Moscow. 73 p. (In Russ.).
- Shemberg A.M. 2017. Structure of vegetation cover and soil characteristics of larch ecosystems in the northern taiga of Central Siberia (Lower Tunguska River basin). Bachelor thesis. Krasnoyarsk. 44 p. (In Russ.). https://elib.sfu-kras.ru/bitstream/handle/2311/34088/diplom_shemberg.pdf
- Sheremetova S.A. 2011. List of the vascular plants of a river basin of Tom. — *Botanicheskie issledovaniya Sibiri i Kazkhstana*. 17: 43–94 (In Russ.).
- Shmidt V.M. 2005. Flora of the Arkhangelsk Region. St. Petersburg. 346 p. (In Russ.). https://herba.msu.ru/shipunov/school/books/shmidt2005_flora_arch_obl.djvu
- Seregin A.P. 2012. Flora of Vladimir Oblast, Russia: Checklist and atlas. Tula. 620 p. (In Russ.). <https://istina.msu.ru/download/58403700/1rwd1L:ihpCqC5liU54aVUNh1ONWTi2dYo/>
- Seregin A.P. 2021a. Contribution to the “Atlas of the Russian Flora”: Twelve local floras of European Russia. — *Biodiversity Data Journal*. 9: e73013. <https://doi.org/10.3897/bdj.9.e73013>
- Seregin A. 2021b. Twelve local floras of European Russia. Version 1.4. Lomonosov Moscow State University. Occurrence dataset accessed via GBIF.org on 2024-04-04. <https://doi.org/10.15468/35rwhv>
- Seregin A. 2024. Moscow University Herbarium (MW). Version 1.326. Lomonosov Moscow State University. Occurrence dataset accessed via GBIF.org on 2024-04-04. <https://doi.org/10.15468/cpnhcc>
- Seregin A.P., Bochkov D.A., Shner J.V., Garin E.V., Pospelov I.N. et al. 2020. "Flora of Russia" on iNaturalist: a dataset. — *Biodiversity Data Journal*. 8: e59249. <https://doi.org/10.3897/BDJ.8.e59249>
- Seregin A., Stepanova N. 2024. MHA Herbarium: collections of vascular plants. Version 1.233. Tsitsin Main Botanical Garden Russian Academy of Sciences. Occurrence dataset accessed via GBIF.org on 2024-04-04. <https://doi.org/10.15468/827lk2>
- Stepanov N.V. 2016. Vascular plants of the Yenisei Sayans. Krasnoyarsk. 252 p. (In Russ.). https://herba.msu.ru/shipunov/school/books/stepanov2016_sosud_rast_prienisejsk_sajan.pdf
- Stepanova N.Y. 2012. Flora of the Kuma-Manych Depression. Cand. Diss. Moscow. 308 p. (In Russ.).
- Taisumov M.A., Omarkhadzhieva F.S. 2012. Analysis of the flora of the Chechen Republic. Grozny. 320 p. (In Russ.). https://herba.msu.ru/shipunov/school/books/tajsumov2012_analiz_flory_tchetchensk_resp.djvu
- Timoshok E.E., Skorokhodov S.N., Raiskaya Y.G., Timoshok E.N. 2008. Species diversity of vascular plants in the southern part of the Tungusky State Nature Reserve. — *Proceedings of the Tungusky State Reserve*. Vol. 2. Tomsk. P. 19–55. (In Russ.).
- Tzvelev N.N. 2000. Manual of the vascular plants of North-West Russia (Leningrad, Pskov and Novgorod provinces). St. Petersburg. 781 p. (In Russ.). https://herba.msu.ru/shipunov/school/books/tzvelev2000_opred_rast_sz.djvu
- Turchin T.Ya., Turchina T.A., Fedyayeva V.V., Mironova N.V. 2004. Tree flora. — *Flora, fauna and mycobiota of the State Museum-Reserve M.A. Sholokhov*. Rostov-on-Don, 14–23 pp. (In Russ.). http://herba.msu.ru/shipunov/school/books/fl_fauna_mikobiota_muz_zapov_sholokhova_2004.djvu
- Turland N.J., Wiersema J.H., Barrie F.R., Greuter W., Hawksworth D.L., Herendeen P.S., Knapp S., Kuster W.-H., Li D.-Z., Marhold K., May T.W., McNeill J., Monro A.M., Prado J., Price M.J., Smith G.F. (Eds.) 2018. International Code of Nomenclature for algae, fungi, and plants (Shenzhen Code) adopted by the Nineteenth International Botanical Congress Shenzhen, China, July 2017. Glashütten. 254 p. <https://doi.org/10.12705/Code.2018>
- Vasilchenko Z.A. 2003. Vascular plants of the Sokhondinsky Nature Reserve (Annotated list of species). Moscow. 87 p. (In Russ.). https://herba.msu.ru/shipunov/school/books/vasilchenko2003_sosud_rast_sokhondinskogo_zapov_vyp_105.djvu
- Vasilyev N.B., Lvov Y.A., Plekhanov G.F., Logunova L.N., Muldiyarov E.Ya., Bibikova V.V., Volkov A.E., Kuzmin S.L., Lapshina E.D., Papanotidi A.I., Sergieva Z.M., Sidorov K.S., Travinskii I.V., Sheftel B.I., Shcherbina S.S. 2003. State Nature Reserve “Tungusky” (essay on basic data). — *Proceedings of the Tungusky State Reserve*. Vol. 1. Tomsk. P. 33–89 (In Russ.).
- Verkhovina A.V. 2004. Flora of vascular plants of the Sayan Area within the Irkutsk Region. Cand. Diss. Irkutsk. 281 p. (In Russ.).
- Vodopyanova N.S. 1984. Zoning of the flora of the Central Siberian Plateau. Novosibirsk. 157 p. (In Russ.). https://herba.msu.ru/shipunov/school/books/vodopyanova1984_zon_flory_vost_sib.pdf
- Yakubov V.V., Chernyagina O.A. 2004. Catalog of Flora of Kamchatka (Vascular Plants). Petropavlovsk-Kamchatsky. 160 p. (In Russ.). https://herba.msu.ru/shipunov/school/books/jakubov_chernyagina2004_katalog_fl_kamchatki.pdf
- Zernov A.S. 2002. Key to vascular plants of the north of the Russian Black Sea Region. Moscow. 283 p. (In Russ.). https://herba.msu.ru/shipunov/school/books/zernov2002_opred_sosud_rasd_sev_ross_prichernomorja.djvu