
Электронный гербарий БСИ ДВО РАН

Выпуск

Dmitry E. Kislov

янв. 18, 2018

1	Базовый поиск по электронному гербарию	3
1.1	Возможности	3
1.2	Поисковое меню	3
1.3	Фильтрация результатов поиска	6
2	Автоматизация поиска гербарных записей в базе данных	11
2.1	Введение	11
2.2	Параметры запросов	11
2.3	Параметры ответа сервера	13
2.4	Ограничения	18
2.5	Примеры	18
3	Администрирование электронного гербария БСИ ДВО РАН	19
3.1	Введение	20
3.2	Структура электронного гербария	21
3.3	Поля формы и их назначение	23
3.4	Персональная страница образца	28
3.5	Этикетка	29
3.6	Этикетка-Конверт	29
3.7	Штрихкодирование образцов	30
3.8	Известные виды	30
3.9	Изображения	32
4	Подготовка и загрузка гербарных изображений	33
4.1	Введение	33
4.2	Меню загрузки изображений	34
4.3	Доступ к обработанным файлам	36
4.4	Параметры сжатия изображений	37
4.5	Названия файлов изображений	37
4.6	Калибровочные ячейки	37
5	Цитирование образцов	41
	Алфавитный указатель	43

Разделы «*Базовый поиск*» и «*Автоматизация поиска*» посвящены вопросам пользовательского поиска в базе данных электронного гербария с использованием web-интерфейса и программных средств соответственно.



Версия для печати

[English version](#)

Базовый поиск по электронному гербарию

1.1 Возможности

Доступ к данным электронного гербария осуществляется либо через предоставляемый web-интерфейс на официальном сайте организации, либо посредством использования *сервиса* автоматизированной обработки поисковых запросов.

В поиске участвуют только опубликованные гербарные записи.

Основные возможности поиска через web-интерфейс:

- поиск в заданных временных интервалах сбора и определения образцов;
- поиск по синонимам видов;
- поиск по заданной прямоугольной географической области;
- поиск по дополнительным видам (для мультивидовых сборов) и истории определений;
- поиск по коду образца;
- поиск по стране места сбора образца;
- поиск по таксономической принадлежности (семейство, род, видовой эпитет)

1.2 Поисковое меню

Базовые поисковые операции возможны через поисковое меню на странице каталога электронного гербария (*Рис. 1*)

При задании нескольких условий в поисковом меню результатом поиска являются записи, для которых выполняются все условия одновременно. Таким образом, поисковое меню позволяет формировать поисковые запросы типа **И**. При необходимости выполнения поисковых запросов типа **ИЛИ** рекомендуется использовать предоставляемый *сервис* автоматизации поисковых запросов.

Найти

Очистить

Семейство:

Выберите Семейство

Род:

Выберите Род

Видовой эпитет:

☐ Учитывать синонимы

☐ Искать в дополнительных видах

Код:

Собрал(и):

Текстовое поле

Определил(и):

Текстовое поле

Страна:

Выберите Страну

Место сбора:

Место сбора

Дата сбора (начало):

Начало сбора

Дата сбора (окончание):

Конец сбора

Рис. 1. Вид стандартного поискового меню

Выбор значений полей **Семейство**, **Род** и **Страна** осуществляется из выпадающего списка, формируемого исходя из известных семейств, родов — у опубликованных образцов и всех известных стран соответственно.

Дата начала и окончания сбора задаются при помощи автоматически всплывающего календаря при попытке внесения значений в эти поля.

В случае задания только даты начала сбора — условие поиска считается выполненным для тех записей, у которых дата начала сбора больше указанной даты начала сбора.

В случае задания только даты окончания сбора — условие поиска считается выполненным для тех записей, у которых дата окончания сбора меньше указанной даты окончания сбора.

В случае задания дат начала и окончания сбора одновременно — условие поиска считается выполненным для тех записей, у которых интервал соответствующих дат пересекается с указанным временным интервалом сбора.

В отношении остальных текстовых полей — **видовой эпитет**, **код**, **собрали**, **определили**, **место сбора** — выполнение поискового условия предполагает включение введенной подстроки (без учёта регистра) в соответствующее поле записей таблицы опубликованных гербарных записей.

В случае поиска по полям **собрали** и/или **определили**, если значения полей записаны кириллическими символами, поисковый запрос также выполняется по транслитерированным значениям, обратная транслитерация при этом не выполняется. Таким образом, если, например, выполнить поиск с полем **собрали** равным «бакалин», то будут найдены все образцы, у которых в поле собрали значится подстрока «bakalin» либо «бакалин»; однако, если поиск изначально выполнить по строке «bakalin», то образцы, у которых в данном поле использованы кириллические символы («бакалин»), не будут найдены.

Булевы поля **Учитывать синонимы** и **Искать в дополнительных видах** указывают, что: в первом случае — дополнительно будет использоваться таблица синонимов видов при поиске образцов, а во втором случае — что поиск также будет распространяться на присоединённые к данному гербарному сбору другие виды (для мультивидовых сборов). В случае, когда отмечены оба поля поиск по известным системе синонимам видов будет проводится также и во всех присоединённых видах гербарных сборов.

Примечание: Функция учитывать синонимы работает только в случае точного указания пары (род, видовой эпитет), в противном случае — условие поиска по синонимам игнорируется и в результатах поиска выводится соответствующее предупреждение, что условие было проигнорировано.

Отдельно следует отметить особенности поиска по полю **Код**.

Электронные гербарные записи БСИ ДВО РАН имеют тройную систему кодирования — инвентарный номер образца (используемый в хранилище), уникальный числовой ID гербарной записи, назначаемый системой, а также полевой код, назначаемый сборщиком гербария.

В связи с этим, таблица результатов поиска имеет колонку **Код сбора (комбинированный)**, в которой отражены все три кода, если таковые заданы.

Комбинированный код имеет следующую структуру:

Примечание: Инвентарный номер или символ */ID образца/Полевой код образца, если задан

Таким образом, возможны следующие варианты:

- */27031/M.I.38 — означает, что инвентарный номер образца не задан, автоматически присваиваемый уникальный номер (ID записи) образца — 27031, а полевой код, присвоенный сборщиком гербария — M.I.38;

- 42/27029 — инвентарный номер — 47, уникальный ID записи — 27029, полевой код — не задан;
- возможны также ситуации, когда заданы все три кода, в этом случае комбинированный код имеет, например, вид: 132123/32032/F-3829-3k (*это, однако, гипотетический пример, так как действительный образец со всеми тремя заданными кодами я не нашел*)

Выполняя поиск по полю **Код** необходимо задавать какой-либо один из кодов; система автоматически выполнит поиск с учетом соответствия значения любому из трех кодовых полей записи. Например, если в поле **Код** задано «231» будут найдены все образцы у которых в каком-либо кодовом поле (либо инвентарном коде, либо ID, либо полевом коде) встречается подстрока «231».

1.3 Фильтрация результатов поиска

Стандартный поисковый интерфейс предоставляет возможность фильтрации результатов поиска по подразделам гербария, а также по гербарным акронимам. Панель фильтрации результатов поиска дана на *Рис. 2*.

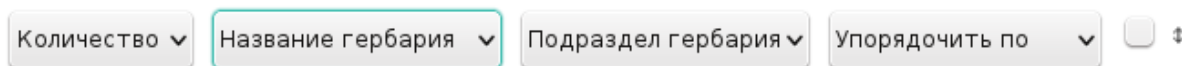


Рис. 2. Панель фильтрации результатов поиска

Панель фильтрации результатов поиска имеет следующие поля:

- **Количество** — количество найденных образцов, отображаемых на одной странице;
- **Название гербария** — ограничение результатов поиска акрониму;
- **Подраздел гербария** — ограничение результатов поиска по подразделу гербария;
- **Упорядочить** — упорядочивание результатов по какому-либо из полей; справа — булево поле для смены порядка отображаемых результатов.

Вид панели результатов выполнения поискового запроса дан на *Рис. 3*.

Во вкладке **Общая информация** выводится перечень удовлетворяющих текущему поисковому условию (в случае, если никаких поисковых условий не было задано — выводятся все опубликованные записи, внесённые в базу на текущий момент).

Вкладка **Информация об образце** активируется, при попытке посмотреть информацию о конкретной гербарной записи в таблице результатов. В этой вкладке отображается уменьшенная копия персональной страницы образца.

Вкладка **Карта** представляет собой результаты поиска с их отображением на карте. Как и в случае со вкладкой **Общая информация** на карте отображается только одна страница результатов поиска. Страницы результатов поиска можно листать, нажимая ссылки **Предыдущая** и **Следующая**, что приведет к синхронному изменению отображаемых записей как на **Карте**, так и во вкладке **Общая информация**.

Вкладка **Автоматизация доступа** содержит общую информацию об организации автоматизированного доступа посредством предоставляемого *сервиса HTTP API*.

Во вкладке **Карта** также имеется возможность фильтрации результатов поиска по прямоугольной области, определённой пользователем на карте. Если область географической фильтрации задана, то в поиске участвуют только те записи, у которых определены географические координаты сбора.

Общая информация				
Информация об образце				
Карта				
Автоматизация доступа				
Результаты поиска:				
Предыдущая			Страница 2 из 564	Следующая
			Всего: 11272	
Код объекта	Вид	Дата сбора	Коллектор(ы)	Определил(и)
115826/26881	<i>Dryopteris austriaca</i> (Jacq.) Woy. ex Schinz & Thell.	1978-08-12	Верхолат В.П.	Верхолат В.П.
111636/26880/5578	<i>Dryopteris mutica</i> (Franch. & Savat.) C. Chr.	1975-08-27	Верхолат В.П., Ворошилова Г.И., Кривенко Н., Коурова О.	Верхолат В.П.
111789/26879	<i>Dryopteris mutica</i> (Franch. & Savat.) C. Chr.	1975-08-27	Верхолат В.П., Ворошилова Г.И., Кривенко Н., Коурова О.	Верхолат В.П.
116079/26878/5578	<i>Dryopteris mutica</i> (Franch. & Savat.) C. Chr.	1970-08-27	Егорова Е.М.	Егорова Е.М.
111225/26877	<i>Dryopteris expansa</i> (C. Presl) Fraser-Jenk. & Jermy	1998-08-18	Галанин А.В.	Кузьмин А.А.
	<i>Dryopteris austriaca</i>			

Рис. 3. Панель результатов поиска

Для активации поиска по географической области необходимо нажать (или дважды кликнуть по окну карты) на кнопку поиска. При этом на карте появится редактируемая прямоугольная область поиска (Рис. 4, Рис. 5).

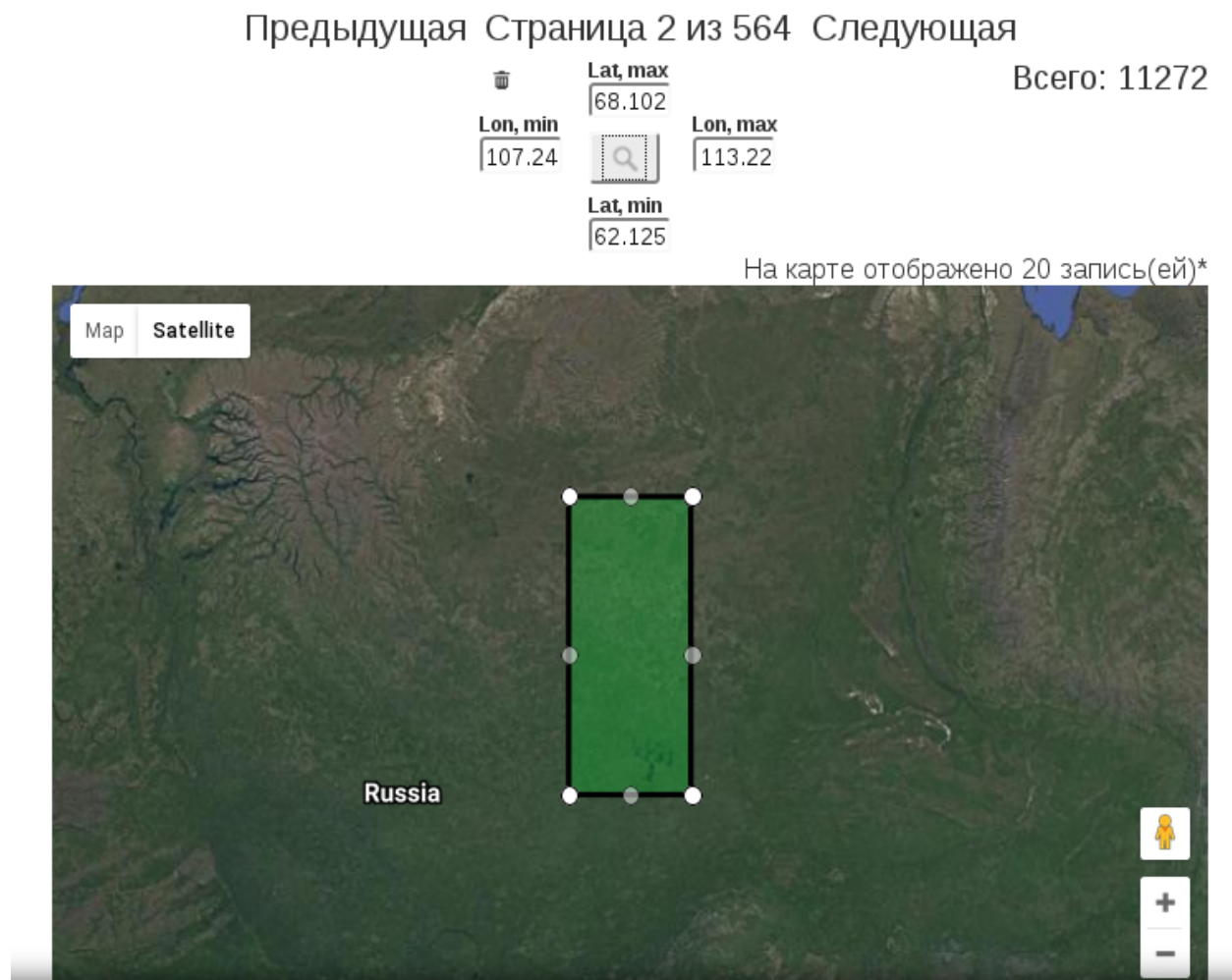
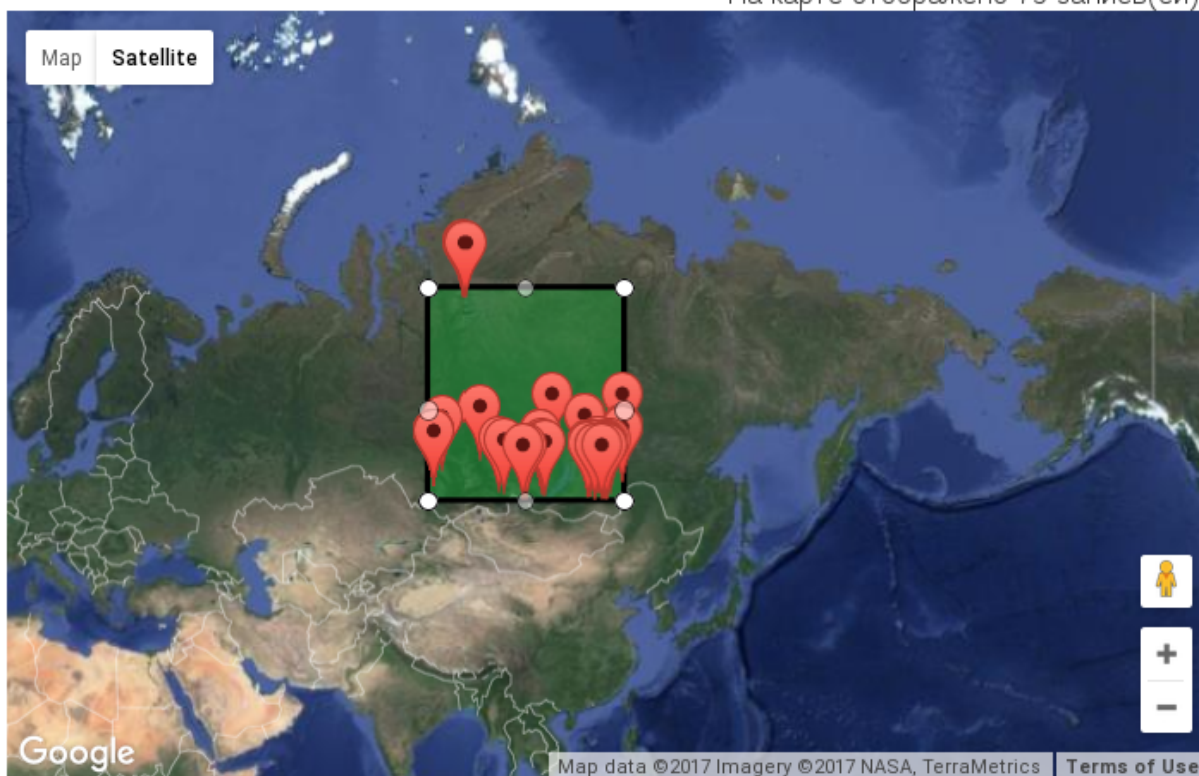


Рис. 4. Выделение области

Для выполнения поиска, после завершения процедуры редактирования поисковой области, необходимо повторно нажать на кнопку поиска. Это приведёт к ограничению результатов поиска только образцами, принадлежащей выделенной области. Для отмены поискового условия по области необходимо нажать на значок **урны** слева от отображаемых координат — границ заданной поисковой области.

Выполнение поиска по более сложным — полигональным поисковым областям в текущей реализации — на базе MySQL — не поддерживаются системой, однако, такой вид поисковых запросов может быть эмулирован программно, в частности, на базе использования предоставляемого сервиса автоматизации поисковых запросов (см. [раздел Дополнительно](#)).

На карте отображено 75 запись(ей)*



*Отображены только те записи, у которых заданы координаты сбора

Рис. 5. Результаты поиска по области

Автоматизация поиска гербарных записей в базе данных

- *Введение*
- *Параметры запросов*
- *Параметры ответа сервера*
 - *Структура массива данных*
 - * *История переопределений и дополнительные сборы*
- *Ограничения*
- *Примеры*

2.1 Введение

Данный документ представляет собой описание HTTP-API (программного интерфейса по протоколу HTTP) для работы с электронным гербарием, работающим под управлением данного web-приложения. HTTP-API не предназначено для изменения данных, а может использоваться только для их чтения.

2.2 Параметры запросов

При формировании поисковых запросов к серверу допустимы только GET-запросы (протокол HTTP(S)).

Одновременное задание полей при формировании GET-запроса приводит к поисковому запросу типа «И»: т.е. если задать одновременно *colstart=2016-01-01* и *collectedby=Бар*, то найдутся все записи, для которых начало сбора позже 1 января 2016 г и у которых в поле «Собрали» встречается последовательность символов „бар“ (без учета регистра).

Формирование запросов типа „ИЛИ“ не реализовано в текущем HTTP-API, но может быть эмулировано серией дополнительных запросов к базе данных.

Возможны следующие параметры GET-запроса, по которым осуществляется поиск:

- **family** — название семейства; название семейства с точностью до регистра записи должно совпадать с запрашиваемым;
- **genus** — название рода; название рода с точностью до регистра записи должно совпадать с запрашиваемым; в случае противоречивости значений параметров **family** и **genus** выводится сообщение об ошибке – отсутствии образцов, удовлетворяющих условиям поиска;
- **species_epithet** — видовой эпитет; поиск осуществляется без учета регистра; условие поиска считается выполненным, если запрашиваемая строка включается в поле «видовой эпитет» основного вида гербарного сбора;
- **place** — место сбора; выполнение условия поиска: включение запрашиваемой строки без учета регистра хотя бы в одно из полей данных о гербарном сборе, поля: место сбора, регион, район, примечание.
- **collectedby** — коллекторы; выполнение условия поиска: включение запрашиваемой строки без учета регистра (поиск также автоматически выполняется по транслитерированной на английский язык строке);
- **identifiedby** — определители; выполнение условия поиска: включение запрашиваемой строки без учета регистра (поиск также автоматически выполняется по транслитерированной на английский язык строке);
- **country** — название страны, места сбора образца; включение запрашиваемой строки без учета регистра в название страны на русском или английском языках (используются названия стран, принятые в стандартах **ISO3166-1-ru** или **ISO3166-1-en**; для Российской Федерации используется сокращенное название Россия);
- **colstart** — дата начала сбора; задается в формате **yyyy-mm-dd**
- **colend** — дата окончания сбора; формат аналогичен параметру **colstart**;
- **acronym** — название гербарного акронима; выполнение условия поиска – точное совпадение акронима гербария без учёта регистра;
- **subdivision** — название гербарного подраздела; выполнение условия поиска: включение запрашиваемой строки без учета регистра;
- **latl** — минимальное значение широты сбора; допустимый диапазон значений (-90.0, 90.0); используется для поиска по прямоугольной области сбора;
- **latu** — максимальное значение широты сбора; допустимый диапазон значений (-90.0, 90.0); используется для поиска по прямоугольной области сбора;
- **lonu** — максимальное значение долготы сбора; допустимый диапазон значений (-180.0, 180.0); используется для поиска по прямоугольной области сбора;
- **lonl** — минимальное значение долготы сбора; допустимый диапазон значений (-180.0, 180.0); используется для поиска по прямоугольной области сбора;
- **synonyms** — возможные значение false или true; отсутствие параметра в GET-запросе приравнивается к false; индикатор поиска с учетом синонимов; при поиске с учетом синонимов важно задать поля **genus** и **species_epithet**; true – означает выполнять поиск по синонимам;
- **additional** — возможные значения false или true; отсутствие параметра в GET-запросе приравнивается к false; индикатор поиска по дополнительным видам внутри гербарных сборов; true – означает искать совпадения в дополнительных видах;

- **id** — **ID** образца; при указании в запросе данного поля, все остальные поля игнорируются; возвращается информация только об образце с заданным **ID**, если такой был найден, в противном случае – выводится сообщение об ошибке;
- **fieldid** — полевой номер образца;
- **itemcode** — инвентарный номер образца (используется в гербарном хранилище);
- **authorship** — авторство основного вида; условие выполняется, если строка авторства основного вида гербарного сбора содержит в себе значение передаваемое в данном параметре (сравнение производится без учета регистра записи);
- **imonly** — возможные значения **false** или **true**; отсутствие параметра в GET-запросе приравнивается к его **false** значению; фильтр **imonly=true** означает, что будут выведены только записи, имеющие изображения.

Замечание При поиске по полям **collectedby** и **identifiedby** выполняется только односторонняя транслитерация: т.е. если, например, задано условие поиска ****collectedby** = Бакалин**, то это эквивалентно поиску ****collectedby** = Бакалин** или ****collectedby** = Bakalin** (найдутся все записи на русском и английском языках). Если задано условие поиска ****collectedby** = Bakalin**, то найдутся только записи, где в поле «сборщики» встречается строка «bakalin» (без учёта регистра).

2.3 Параметры ответа сервера

Ответ сервера на поисковый запрос представляет собой **JSON-форматированный** текст, передаваемый по протоколу HTTP, и имеющий следующие параметры:

- **errors** — массив ошибок, возникших при обработке поискового запроса.
- **warnings** — массив предупреждений, возникших при обработке поискового запроса. Предупреждениями являются различные поисковые ситуации: например, отсутствие данных, удовлетворяющих текущему поисковому запросу, игнорирование тех или иных поисковых параметров, при их противоречивости и т.п.
- **data** — массив структурированных данных гербарных записей, удовлетворяющих текущему поисковому запросу.

2.3.1 Структура массива данных

Параметр **data** представляет собой массив данных, удовлетворяющих текущему поисковому запросу.

Он имеет следующую структуру, описывающую текущий гербарный сбор:

- **family** — название семейства (заглавными буквами, на латыни);
- **family_authorship** — автор семейства;
- **genus** — название рода;
- **genus_authorship** — автор рода;
- **species_epithet** — видовой эпитет;
- **species_id** — **ID** вида образца; не путать с **ID** текущей гербарной записи. **ID** текущей гербарной записи однозначно характеризует данную оцифрованную гербарную запись. **ID** вида образца, только вид. Гербарных записей, содержащих какой-либо вид может быть много.
- **species_authorship** — автор вида;

- **species_status** — текущий статус вида; определяет степень признанности данного вида, точнее триплета (род, видовой эпитет, авторство вида) в научном сообществе на настоящее время. Возможные значения данного параметра 1) «Recently added» — вид недавно добавлен и, скорее, не проверялся специалистом; название вида с таким статусом может быть устаревшим, либо содержать ошибки; 2) «Approved» — название вида подтверждено специалистом; 3) «Deleted» — вид имеет ошибку в записи, или его название устарело и не используется; 4) «From plantlist» — название импортировано из базы <http://theplantlist.org>.
- **type_status** — типовой статус гербарного сбора; для мультивидовых сборов — отражает типовой статус в отношении основного вида сбора;
- **infraspecific_rank** — подвидовой ранг (возможные значения: пусто, subsp., subvar., f., subf., var.);
- **infraspecific_epithet** — подвидовой эпитет;
- **infraspecific_authorship** — автор подвидового эпитета;
- **species_fullname** — полное название вида, т.е. Род + видовой эпитет + авторство;
- **short_note** — замечания к главному виду сбора (используется по необходимости в случае мультивидовых сборов);
- **significance** — неопределенность знаний относительно главного вида гербарного сбора (возможные значения: пусто, aff., cf.);
- **id** — уникальный идентификатор данной гербарной записи; всегда целое число;
- **gpsbased** — получены ли данные о географической привязки места сбора образца с помощью GPS (значение **true**), либо другим способом (**false**); следует иметь в виду, что у многих образцов, даже при **gpsbased** равном **false**, координаты, если таковые заданы, были получены при помощи GPS; это связано с тем, что не все отмечают соответствующее поле (**gpsbased**) при заполнении электронной формы образца;
- **latitude** — широта, градусы; географическая координата места сбора в системе WGS-84;
- **longitude** — долгота, градусы; географическая координата места сбора в системе WGS-84;
- **fieldid** — полевой номер образца;
- **itemcode** — инвентаризационный номер, используемый в гербарном хранилище;
- **acronym** — гербарный акроним, которому принадлежит данная гербарная запись (для большинства записей поле имеет значение **VBGI**);
- **branch** — подраздел гербария внутри акронима; иногда удобно иметь подразделы внутри общей гербарной базы: например, «гербарий грибов», «биоморфологический гербарий» и т.п.;
- **collectors** — текстовая строка: сборщики образца;
- **identifiers** — текстовая строка: те, кто определил вид гербарного сбора;
- **devstage** — стадия развития; определена для биоморфологического гербария; возможные значения: Development stage partly, Life form, или пустое поле;
- **updated** — дата последнего изменения записи в базе данных;
- **created** — дата создания записи (т.е. занесения её в электронную базу данных);
- **identification_started** — дата начала определения вида образца;
- **identification_finished** — дата окончания определения вида образца; дата определения вида задана в виде интервала, поскольку не всегда может быть указана точная дата, а например, только месяц, или время проведения какой-либо экспедиции;
- **collection_started** — дата начала сбора образца;

- **collection_finished** — дата окончания сбора образца; дата сбора задана в виде интервала, поскольку не всегда может быть указана точная дата, а например, только месяц, или время проведения какой-либо экспедиции;
- **country** — название страны сбора образца;
- **country_id** — числовой идентификатор страны сбора образца;
- **altitude** — высота над уровнем моря места сбора образца; значение представляется собой строку, не всегда однозначно определяющую реальную высоту сбора. Возможны, например, варианты: 100-300 м, 120 м, 400, 300-400 и т.п.
- **region** — регион сбора;
- **district** — район сбора;
- **details** — экологические условия сбора, дополнительные уточнения не вошедшие в поля регион и район;
- **note** — примечание; может содержать информацию о месте сбора, экологических условиях и т.п.;
- **dethistory** — представляет собой массив — историю переопределений вида гербарного сбора;
- **additional** — некоторые гербарные сборы могут содержать более одного вида; данный массив описывает характеристики каждого из них.
- **images** — перечень изображений, относящихся к гербарной записи ([] — пустой список — означает с каждым элементом списка — это путь до изображения, соответствующего данному гербарному сбору.

Примечание: Изображения в списке **images** доступны в нескольких разрешениях. В текущей реализации сервиса изображения различных разрешений размещены по директориям, имеющим названия: *ts* — thumbnail size (очень маленький размер); *ss* — small size (30% от оригинала); *ms* — medium size (60% от оригинала); *fs* — full size (оригинальное разрешение). Таким образом, в каждом url изображения есть одна из компонент пути '/ts/' , '/ss/' , '/ms/' или '/fs/' , указывающая изображение какого разрешения соответствует данному url.

Примечание: Все изображения сохранены в формате jpeg. Для этих целей используется редактор командной строки **ImageMagick** с параметрами:

```
'-strip', '-interlace', 'Plane',
'-sampling-factor', r'4:2:0',
'-quality',
r'90%'
```

Практика показывает, что используемый набор параметров не оказывает видимого эффекта на изображения. Вместе с тем, использование сжатия позволяет существенно сократить требуемое дисковое пространство.

Пример вида массива **images**:

```
['http://botsad.ru/herbarium/view/snapshots/VBGI/ss/VBGI32618_1.jpg',
'http://botsad.ru/herbarium/view/snapshots/VBGI/ts/VBGI32618_1.jpg',
'http://botsad.ru/herbarium/view/snapshots/VBGI/ms/VBGI32618_1.jpg',
'http://botsad.ru/herbarium/view/snapshots/VBGI/fs/VBGI32618_1.jpg',
...
]
```

Поля **region**, **district**, **details**, **note**, **altitude** могут быть заполнены с поддержкой двуязычности с использованием спецсимвола «|». Например, строка, возвращаемая в поле **region**, может быть такой «Russian Far East|Дальний Восток России». Это означает, что относительно символа «|» даётся русско- и англоязычный варианты строки. Дальнейшая обработка значений таких строк ложится на пользователя системы, которому решать какую из компонент строки относительно символа «|» оставить, а какую – удалить. Система HTTP-API не принимает таких решений.

Структура массивов **dethistory** и **additional**s приводится ниже.

История переопределений и дополнительные сборы

История переопределений

Каждый элемент массива «История переопределений» (**dethistory**) представляет собой описание попытки определения (переопределения) вида в текущем гербарном сборе и имеет следующие поля (значения полей, характеризующих вид, аналогично описанным выше):

- **valid_from** — дата валидности определения;
- **valid_to** — дата окончания валидности определения; поле может быть не задано, что означает, что предполагает, что определение актуально в настоящее время;
- **family** — название семейства;
- **family_authorship** — авторство семейства;
- **genus** — название рода;
- **genus_authorship** — автор рода;
- **species_epithet** — видовой эпитет;
- **species_id** — ID вида образца;
- **species_authorship** — автор вида;
- **species_status** — текущий статус вида;
- **species_fullname** — полное название вида;
- **infraspecific_rank** — подвидовой ранг (возможные значения: пусто, subsp., subvar., f., subf., var.);
- **infraspecific_epithet** — подвидовой эпитет;
- **infraspecific_authorship** — автор подвидового эпитета;
- **significance** — неопределенность знаний относительно вида (возможные значения: пусто, aff., cf.);

Сроки валидности вида (**valid_from**, **valid_to**) позволяют корректно описать любые его последующие переопределения.

Примечание: Если в гербарном сборе представлен не один вид, то массив «История переопределений» представляет собой историю переопределений основного вида.

Дополнительные виды

Каждый элемент массива «Дополнительные виды» (**additional**s) представляет собой описание вида, находящегося в данном гербарном сборе. Каждое из таких описаний имеет поля, аналогичные записям из **Истории переопределений**:

- **valid_from** — дата валидности определения;

- **valid_to** — дата окончания валидности определения; поле может быть не задано, что означает, что предполагает, что определение актуально в настоящее время;
- **family** — название семейства;
- **family_authorship** — авторство семейства;
- **genus** — название рода;
- **genus_authorship** — автор рода;
- **species_epithet** — видовой эпитет;
- **species_id** — ID вида образца;
- **species_authorship** — автор вида;
- **species_status** — текущий статус вида;
- **species_fullname** — полное название вида;
- **infraspecific_rank** — подвидовой ранг (возможные значения: пусто, subsp., subvar., f., subf., var.);
- **infraspecific_epithet** — подвидовой эпитет;
- **infraspecific_authorship** — автор подвидового эпитета;
- **note** — примечания о текущем дополнительном сборе;
- **significance** — неопределенность знаний относительно вида (возможные значения: пусто, aff., cf.);

Таким образом, массив «Дополнительные виды» позволяет хранить информацию о видах в гербарном сборе, сопутствующих данному основному виду (выделенному из экспертных соображений в качестве основного), а указание валидности позволяет описать переопределения (если таковые имеются) каждого из таких видов.

Примечание: Поле **note** поддерживает режим двуязычного заполнения, поэтому в отношении его справедливо замечание, указанное *выше*.

Пояснение и интерпретация

Рассмотрим для примера следующий массив «Дополнительных видов» (для краткости выписаны не все поля):

```
[
  {'genus': 'Quercus', 'species_epithet': 'mongolica', ... , 'valid_from': '2015-05-05', 'valid_to':
  ↪ '2016-01-01'},
  {'genus': 'Quercus', 'species_epithet': 'dentata', ... , 'valid_from': '2016-01-01', 'valid_to': ''}
  ↪ ,
  {'genus': 'Betula', 'species_epithet': 'manshurica', ... , 'valid_from': '2015-05-05', 'valid_to': ''
  ↪ },
  {'genus': 'Betula', 'species_epithet': 'davurica', ... , 'valid_from': '2015-05-05', 'valid_to': ''}
  ↪ ,
]
```

Если сегодня, например, 1 сентября 2015 года (2015-09-01), то массив дополнительных видов состоит из *Quercus mongolica*, *Betula manshurica* и *Betula davurica*, а *Quercus dentata* является неактуальным определением на данный момент времени.

Если сегодня 2017 год, например, 2017-01-01, то неактуальным оказывается *Quercus mongolica*, и, таким образом, актуальными видовыми составляющими сбора являются *Quercus dentata*, *Betula manshurica* и *Betula davurica*

2.4 Ограничения

Поскольку поисковому запросу пользователя может удовлетворять большой объём данных, для формирования ответа сервера может потребоваться значительное время.

Чтобы снизить нагрузку на сервер, вызванную вероятно долгими keep-alive HTTP-соединениями, действуют ограничения.

Количество одновременно возможных соединений для сервиса автоматизированного опроса гербарной базы определяется текущим значением параметра `JSON_API_SIMULTANEOUS_CONN`.

По превышении этого количества, сервер не обрабатывает поисковые запросы, а возвращает сообщение об ошибке.

На запросы, содержащие **id**, данное ограничение не действует, поскольку получение информации об объекте по его **ID** — не ёмкая в плане ресурсов операция.

Информация о **неопубликованных** образцах не выводится; при попытке получить информацию о неопубликованном образце по его **ID** выводится ошибка.

2.5 Примеры

Для проверки работы системы и получения json-ответа сервера достаточно передать поисковый запрос в url браузера.

Например, переход по ссылке

<http://botsad.ru/hitem/json/?genus=riccardia&collectedby=bakalin>

приведет к появлению на экране браузера json-ответа, содержащего информацию о всех сборах — представителей рода *Riccardia*, в строке, содержащей информацию о сборщиках которых встречается *bakalin*.

При указании **id** в **GET** запросе, все остальные поисковые поля игнорируются и выводится информация о гербарном образце с указанным **id**:

<http://botsad.ru/hitem/json?id=500>

<http://botsad.ru/hitem/json?id=44>

<http://botsad.ru/hitem/json?id=5>

См.также:

Работа с электронным гербарием из Python

Работа с электронным гербарием из R

Администрирование электронного гербария БСИ ДВО РАН

- *Введение*
 - *Аутентификация и безопасность*
- *Структура электронного гербария*
 - *Гербарные акронимы*
 - *Подразделы гербария*
 - *Группы пользователей и права*
 - * *Куратор*
 - * *Куратор подраздела гербария*
 - * *Регулярный пользователь (наборщик)*
- *Поля формы и их назначение*
 - *Поддержка альтернативного языка*
 - *Вид*
 - *Примечание к виду*
 - *Типовой статус*
 - *Сходство*
 - *Код образца*
 - *Полевой код*
 - *Акроним*
 - *Страна*
 - *Регион*

- *Район*
- *Место сбора*
- *Координаты*
- *Высота*
- *GPS-Based*
- *Собрали*
- *Начало и Конец сбора*
- *Определили*
- *Начало и конец определения*
- *Биоморфологический статус*
- *Подраздел гербария*
- *Заметки*
- *Опубликовано*
- *Переопределения*
- *Дополнительные виды*
- *Сохранить для добавления следующей записи*
- *Персональная страница образца*
- *Этикетка*
- *Этикетка-Конверт*
- *Штрихкодирование образцов*
- *Известные виды*
 - *Поля таблицы Известных видов*
 - * *Название*
 - * *Авторство*
 - * *Род*
 - * *Подвидовой ранг*
 - * *Подвидовой эпитет*
 - * *Статус*
 - * *Синоним вида*
- *Изображения*

3.1 Введение

В документе изложены основные принципы работы с формой администрирования, ее основные функции, права пользователей, описание полей, а также рекомендации и соглашения об их заполнении.

3.1.1 Аутентификация и безопасность

Аутентификация осуществляется по предоставленному имени пользователя и паролю.

Пароль рекомендуется сменить после первого успешного входа в панель администрирования. Это можно сделать, перейдя по ссылке в правом верхнем углу окна панели администрирования.

Пароли не хранятся в открытом виде в системе, у администратора системы (или сотрудника с правами суперпользователя) не существует (кроме прямого подбора) способа, чтобы узнать пароль; однако, суперпользователь может изменить пароль любого другого пользователя.

После ста неудачных попыток авторизации IP адрес, с которого были сделаны неудачные попытки, блокируется на час.

Если пользователь в течение 16 часов работы в системе не производит никаких действий – сохранений, переходов на другие страницы, то срок валидности его сессии истекает. По истечении срока валидности сессии пользователю необходимо пройти аутентификацию заново (ввести имя пользователя и пароль), чтобы иметь возможность редактировать образцы.

3.2 Структура электронного гербария

Электронный гербарий представляет собой многопользовательское web-приложение, которое позволяет организовать хранение\изменение\добавление данных с учетом разграничения пользовательских прав, с учетом названий гербария и гербарных подгрупп.

Разграничение гербарных записей в общей таблице осуществляется по двум основным полям – принадлежности определенному гербарии (гербарному акрониму) и подразделу гербария.

Например, может потребоваться необходимость внутри данного гербария (например, VBGI), выделить биоморфологический гербарий, или коллекции грибов и пр. Для этих целей могут использоваться подразделы гербария.

Права пользователей могут распространяться как на образцы, принадлежащие данному акрониму, и\или быть ограниченными определенными подразделами гербария.

Пользовательская поисковая форма (размещаемая на странице <http://botsad.ru/herbarium>) позволяет осуществлять фильтрацию результатов поиска по гербарным акронимам и\или подразделам гербария.

Каждая гербарная запись, независимо от принадлежности гербарному акрониму или подразделу гербария, получает при своем сохранении уникальный, привязанный только к этой записи, целочисленный номер ID. Он назначается системой автоматически. Даже при удалении образца, данный номер уже не будет занят.

3.2.1 Гербарные акронимы

Гербарные акронимы — идентификаторы принадлежности тому или иному гербарии (например, гербарии БСИ ДВО РАН, или гербарии Амурского филиала БСИ ДВО РАН) хранятся в отдельной таблице акронимов, редактирование которой разрешено только суперпользователю (см. о правах суперпользователя в разделе *Группы пользователей и права*).

Пользователи, авторизуемые в системе администрирования, привязаны к тому или иному акрониму. Когда они сохраняют заполненный гербарный образец, доступное только для чтения (изначально пустое) поле акронима, заполняется автоматически в соответствии с этой привязкой.

Каждая запись таблицы акронимов (в настоящий момент таблица хранит записи — VBGI, AmBGI, SAKH) хранит информацию о физическом размещении гербария (его адрес на англ. языке), привязку к пользователям, собственно сокращенное название.

Акронимы используются для разграничения прав кураторов. Куратор гербария привязывается к тому или иному акрониму, но не имеет доступа к гербарным записям принадлежащим другому акрониму.

Информация, содержащаяся в таблице акронимов (адрес гербария, международное обозначение гербария), используется при автоматизированном создании макетов этикеток.

Если требуется внести изменения в таблицу акронимов, например, исправить AmBGI на ABGI, или что-то подобное, изменить адрес гербария, нужно обращаться к сотруднику с правами суперпользователя системы.

3.2.2 Подразделы гербария

Подраздел гербария — необязательное автоматически назначаемое поле при добавлении гербарной записи. Оно может быть пустым, либо заполняется при сохранении образца с учетом привязки конкретного пользователя к подразделу гербария.

Если пользователь с правами куратора привязан к определенному подразделу гербария, то он является куратором только данного подраздела (см. также *Группы пользователей и права*).

Таблица гербарных подразделов с их наименованиями и привязкой к конкретным пользователям редактируется сотрудником с правами суперпользователя.

Если текущий пользователь, редактирующий гербарий, не привязан ни к какому гербарному подразделу, при сохранении поле поле, отвечающая за подраздел, остается пустым.

3.2.3 Группы пользователей и права

Суперпользователь — имеет права на все; за исключением удаления\изменения опубликованных гербарных образцов. Тем не менее, он может изменить статус опубликованного образца на неопубликованный.

Куратор

Куратор гербария — осуществляет мониторинг всех гербарных записей в рамках акронима, к которому он привязан.

Кроме того:

- куратор может быть привязан только к одному гербарному акрониму;
- куратор осуществляет публикацию (и\или снятие с публикации) гербарных образцов;
- куратор может редактировать и просматривать любые гербарные образцы в рамках своего акронима;
- куратор не может удалять виды, рода или семейства, но может их добавлять;
- куратор может исключить вид из результатов поиска в всплывающих подсказках, присвоив ему статус «Deleted»;
- куратор может добавлять\изменять названия уже добавленных видов (при этом изменения скажутся сразу на всех образцах, у которых указан данный вид: например, если вид *Betula mandshurica* в таблице видов переименовать на *Betula davurica*, то все образцы, где был указан до переименования вид *Betula mandsurica*, после переименования будут иметь вид *Betula davurica*);
- куратор может удалять любые образцы в рамках своего акронима

Куратор подраздела гербария

Если имя пользователя с правами куратора привязано к какому-либо гербарному подразделу, то все права куратора ограничиваются только данным подразделом. Остальные права куратора гербарной группы идентичны правам куратора гербария.

Регулярный пользователь (наборщик)

Осуществляет набор данных от своего имени. Акроним гербарной записи при этом назначается автоматически, исходя из привязки пользователя определенному гербарному акрону.

Регулярный пользователь может быть дополнительно привязан к подразделу гербария. В этом случае, сохранение образца влечет автоматическое назначение подраздела данной гербарной записи.

- регулярный пользователь может просматривать и редактировать\удалять только собственно созданные записи;
- может создавать новые виды;
- не может создавать рода и семейства;
- не может изменять статус видов;
- не может публиковать\снимать с публикации гербарные записи;

Существует еще дополнительный набор прав — «Редактор мультивидовых сборов», который подразумевает, что форма редактирования образца включает поля, отражающие присутствие дополнительных видов в данном гербарном сборе; в обычной гербарии сосудистых растений такая функция не нужна, поэтому она включается только при присвоении пользователю данного набора прав.

3.3 Поля формы и их назначение

Любые поля формы, выделенные жирным шрифтом, обязательны для заполнения.

В форме редактирования гербарных образцов такое поле одно — это поле **Вид**.

3.3.1 Поддержка альтернативного языка

Заполняя данные в цифровой гербарий, прежде всего следует ориентироваться на международное научное сообщество. Таким образом, предпочтительным языком заполнения является английский.

Тем не менее, некоторые поля, а именно *Страна, Регион, Район, Место сбора, Высота, Заметки, Примечание к виду* поддерживают эмуляцию двуязычного заполнения при помощи спецсимвола «|».

Эмуляция двуязычности работает следующим образом. Разграничение языков осуществляется при помощи специального символа «|». Например, поле *Регион* может содержать: «Дальний Восток России|Russian Far East» (*кавычки в форме администрирования не ставятся*). В этом случае, система автоматически будет использовать русский вариант (Дальний Восток России), если пользовательский язык просмотра страницы «русский», и английский вариант (Russian Far East) — в случае, если язык просмотра «английский».

Кроме того, при выборе языка в строках с символом «|» учитываются следующие правила:

- если в строке символ «|» встречается более одного раза, система отображает строку как есть (никакого выбора языка не происходит);

- если в строке символ отделяет пустую подстроку (например, «Владивосток|»), то строка отображается как есть (никакого выбора языка не происходит);
- порядок русско- и англоязычных вариантов в строке относительно символа «|» не имеет значения: т.е. «Дальний Восток России|Russian Far East» и «Russian Far East|Дальний Восток России» являются эквивалентными записями с точки зрения системы;
- система считает русскоязычной ту подстроку относительно символа «|», в которой больше встретилось кириллических символов;
- если число кириллических символов в обеих частях строки относительно «|» одинаково, или они вообще отсутствуют, англоязычной считается правая подстрока относительно символа «|».

Выбор языка для полей – **Дат сбора\определения** осуществляется автоматически системой, исходя из текущего языка браузера пользователя.

Поля **Определили** и **Собрали** будут использовать автоматическую транслитерацию на английский язык, если они заданы на русском, а язык просмотра пользователем страницы отличен от русского. Данные поля не нужно переводить, если они заполнены на русском, но можно сразу заполнять на английском.

3.3.2 Вид

Заполнить поле вид можно только элементом из всплывающей подсказки. Подсказка формируется по уже добавленным видам в базу, а также более 500k (по состоянию на конец 2016 г) видам из базы данных theplantlist.org.

Поиск выполняется как только набрано 3 и более символов в поле **Вид**; при этом полагается, что первые символы — должны состоять в названии рода, а последующие, если они идут через пробел – видового эпитета. Иными словами, если мы вводим в поисковое поле **Вид**: *Tra*, то в поиске появятся все виды с родами, начинающиеся на *Tra*, при этом количество видимых вариантов будет ограничено 50 вариантами; если в поисковом поле **Вид** будет, например, *Tra ps*, то найдутся все виды, рода у которых начинаются на *Tra*, и, кроме того, видовой эпитет содержит *ps*, т.е., например, *Trapa pseudoincisa* и т.п.

Если требуемый вид отсутствует, нужно нажать рядом кнопку в виде «зеленого плюса» и добавить недостающий вид. Если отсутствует необходимый род и\или семейство, необходимо обратиться к сотруднику с правами куратора и\или суперпользователя, чтобы добавить недостающий род и\или семейство.

У вида можно задать синоним. Добавление синонимов доступно на странице редактирования вида. Подробно об этом можно прочитать в разделе *Известные виды*.

3.3.3 Примечание к виду

Строка; максимальная длина — 300 символов. Поддерживает режим двуязычности с использованием символа «|». Видна только пользователям с правами «редактора мультивидовых сборов». Характеризует специфические особенности основного вида данного гербарного сбора в случае мультивидовых сборов.

3.3.4 Типовой статус

Поле, доступное для редактирования только пользователям с правами куратора.

Возможные значения: пустое значение, HOLOTYPUS, ISOTYPUS, PARATYPUS, LECTOTYPUS.

Значение типового статуса отражается на генерируемых этикетках (данная функциональность еще не реализована).

3.3.5 Сходство

Характеризует степень уверенности в определении вида данного образца. Возможные значения «affinis» (aff.), «confertum» (cf.).

Данное поле присутствует также во вкладках *Переопределения* и *Дополнительные виды*.

3.3.6 Код образца

Уникальный в рамках данного акронима числовой код. Этот код опциональный и может не заполняться. Однако, в некоторых публикациях могут быть ссылки на гербарные образцы с указанием этого кода, поэтому он может быть важен.

Если данный код задан, то он отражается на этикетке; если не задан — вместо него используется символ «*».

Регулярный пользователь (наборщик) не имеет прав заполнять данное поле, однако его права могут быть расширены на этот случай.

По умолчанию, права на заполнение данного поля имеют кураторы и суперпользователи.

3.3.7 Полевой код

Опциональный код. Его назначает сборщик гербария; его максимальная длина 20 символов, при этом допустимо использовать любые символы в рамках кодировки utf-8 (т.е. там могут быть и японские иероглифы).

Также может быть важным для ссылок.

Если существует, то отражается на этикетке.

3.3.8 Акроним

Автозаполняемое поле. Оно доступно только для чтения для всех пользователей, за исключением суперпользователя. Суперпользователь может самостоятельно назначить принадлежность образца любому акрониму. Автозаполнение осуществляется на основе привязки пользователей к акронимам.

Поле используется при формировании заголовка этикетки.

3.3.9 Страна

Рекомендуемое к заполнению поле. Необходимо выбрать страну происхождения гербарного сбора. Поиск осуществляется по русскоязычным и англоязычным общепринятым в рамках стандарта ISO перечнем стран (вместо «Российская Федерация» для компактности формы при создании электронных макетов этикеток используется «Россия»).

Данное поле отображается на английском языке на этикетке.

3.3.10 Регион

Отражается на этикетке. Это поле с возможным автозаполнением из того, что уже было введено в базу.

Поддерживает двуязычное заполнение, описанное в разделе *Поддержка альтернативного языка*.

3.3.11 Район

Не отражается на этикетке. Опциональное поле.

Поддерживает двуязычное заполнение, описанное в разделе *Поддержка альтернативного языка*.

3.3.12 Место сбора

Максимальная длина этого поля 600 символов. В этом поле следует также размещать важную информацию об экологических особенностях места сбора.

Поддерживает двуязычное заполнение, описанное в разделе *Поддержка альтернативного языка*.

3.3.13 Координаты

Для заполнения можно использовать флажок на прилагаемой карте google. При изменении позиции флажка, автоматически изменяются и координаты. В правом верхнем углу карты есть и поисковое поле, в которое можно ввести здесь название населенного пункта и флажок переместится в центр этого пункта, если, конечно, такой будет найден (т.е. если название известно google).

3.3.14 Высота

Высота над уровнем моря в метрах.

Поддерживает двуязычное заполнение, описанное в разделе *Поддержка альтернативного языка*.

3.3.15 GPS-Based

Отмечается, если координаты сбора были получены при помощи GPS; это характеристика точности позиционирования сбора; поскольку координаты сбора могут быть получены исходя из описания сбора ручным указанием положения флажка на google-карте.

3.3.16 Собрали

Поле-автоподсказка. Автоподсказка формируется из уже известных уникальных записей, внесенных в базу.

3.3.17 Начало и Конец сбора

Для заполнения может быть использован всплывающий календарик (кнопка справа). Начало и конец указываются если не известна точная дата сбора, но известны, например, даты проведения экспедиции, в ходе которой был осуществлен сбор.

Если дата известна точно, то можно заполнить только одно поле — начало сбора; также можно указать конец сбора, таким же как и начало сбора, либо оставить пустым.

Дата сбора отражается на этикетке, в виде, например, таком: 15 Jul 1998.

Если известен только месяц сбора, то этот факт следует отражать указав начало сбора — первое число месяца, а конец сбора — последнее число месяца. Например, если сбор выполнен в марте, 1999 года, то начало сбора будет 1 марта 1999 г, а конец сбора — 31 марта 1999 г.

Если время сбора указано с точностью до года, следует поступать аналогичным образом — указать первое и последнее числа года — 1 января и 31 декабря.

3.3.18 Определили

Поле-автоподсказка. Работает по аналогии с полем «Собрали». Отражается на этикетке. Если ученых, участвующих в определении много, на этикетке будет указан сокращенный вариант — первые одна, две фамилии (сколько удастся автоматически разместить).

3.3.19 Начало и конец определения

Аналогично началу и концу сбора. Поле не отражается на этикетке.

3.3.20 Биоморфологический статус

Отражается на этикетке, если не пусто. Возможные значения «Dev.stage partly» или «life form». Эти словосочетания и печатаются на этикетке. Специально для биоморфологического гербария БСИ ДВО РАН.

Вполне возможно, оно будет строго привязано к гербарному подразделу «Биоморфологический гербарий», и не будет появляться у пользователей, не привязанных к этой группе.

3.3.21 Подраздел гербария

Автоматически назначаемое поле и доступное только для чтения для регулярных пользователей и кураторов. Может быть пустым. Суперпользователь может редактировать данное поле и указывать подраздел гербария явно. В других случаях поле назначается исходя из привязки пользователя подразделу. Назначается при сохранении образца. До сохранения — не определена.

3.3.22 Заметки

Все что еще мы хотим сообщить о сборе. Для этого здесь доступно 1000 символов.

Поддерживает двуязычное заполнение, описанное в разделе *Поддержка альтернативного языка*.

3.3.23 Опубликовано

Если отмечено, то образец опубликован. Публиковать образцы (как и снимать их с публикации) могут только кураторы гербария\гербарной группы, а также суперпользователь.

3.3.24 Переопределения

Переопределения заполняются, если первоначально определенный вид, потом переопределили. На этикетке, однако, при этом сохраняется первоначальные данные. История переопределений не отражается на этикетке.

В разделе «Переопределения» можно добавить несколько определений, указав соответственно начало (и при необходимости конец) валидности определения. Последним полем блока «Переопределения» является вид, то на что текущий вид был переопределен.

Если этот раздел заполнен, то он отображается на персональной странице образца.

Если поле «Определили» пусто, а история переопределений имеется, то на этикетке будет отображена последняя запись из истории переопределений.

3.3.25 Дополнительные виды

Раздел доступен для редактирования только пользователям с правами «Редактор мультивидовых сборов» (пользователей со специальными правами, у которых в сборах может быть больше одного вида).

Дополнительные виды заполняются по аналогии с полем *Переопределения*, за исключением того, что для каждого дополнительного вида имеется возможность указать индивидуальное примечание. Данные примечания ограничены объемом 300 символов, также как и поле *Примечание к виду*.

3.3.26 Сохранить для добавления следующей записи

Если отметить данную позицию, то следующий добавляемый гербарный объект будет иметь уже заполненные поля, как у текущего. Данная функция удобна, когда необходимо добавить несколько гербарных объектов, имеющих однотипное описание: собранных в одинаковом месте, в одинаковых условиях и т.п.

Прежде чем отмечать позицию «Сохранить для добавления следующей записи», важно предварительно сохранить заполненную форму. Поэтому, рекомендуется всегда сначала нажимать «Сохранить и продолжить редактирование», а уже потом отмечать «Сохранить для добавления следующей записи».

Отмечая позицию «Сохранить для добавления следующей записи» пользователь сообщает системе, чтобы она запомнила ID текущего образца и использовала данные его полей для добавления следующих записей.

При добавлении последующих образцов надпись позиции «Сохранить для добавления следующей записи» дополнится фразой «(не этот образец)», которая указывает, что информация для предварительного заполнения полей берется из какой-то другой записи (на которой позиция была отмечена).

Снятие галочки с данной позиции выключает данную функциональность.

Функциональность работает в рамках данной сессии пользователя. Если пользователь завершит работу с системой и потом снова авторизуется, функциональность запоминания полей будет отключена.

Следует иметь в виду, что в текущей реализации данная функциональность не запоминает поля форм *Переопределения* и *Дополнительные виды* (вполне возможно, что в будущем это будет изменено).

3.4 Персональная страница образца

Детальная информация об опубликованном образце доступна по адресу: <http://botsad.ru/hitem/ID>, где «ID» это уникальный код образца, назначаемый системой.

Также, в целях соответствия общим стандартам, валидными ссылками на персональную страницу образца являются ссылки вида: <http://botsad.ru/hitem/ACRONYMXXXXXX>, где ACRONYM – акроним гербария, XXXXX – уникальный числовой код. Примеры: <http://botsad.ru/hitem/VBGI133>, <http://botsad.ru/hitem/VBGI120>

Аналогичный формат доступа к персональным страницам используется в электронном гербарии KEW: <http://apps.kew.org/herbcat/gotoCiteUs.do>

На этой странице указывается история определений, заметки и прочая информация, не вошедшая на этикетку.

Адрес персональной страницы не зависит от акронима и\или гербарной группы.

3.5 Этикетка

Для генерации этикеток в общем списке гербарных записей панели администрирования необходимо выделить образцы, выбрать действие – «Создать этикетки» и нажать «Выполнить».

За один запрос можно сгенерировать не более 100 этикеток; они автоматически размещаются оптимально на странице формата А4.

QR-код, размещаемый на этикетке, представляет собой URL персональной страницы образца, также указываемый мелким шрифтом сразу под изображением QR-кода.

Если среди опубликованных образцов для генерации этикетки будут выбраны и неопубликованные – последние будут проигнорированы, а этикетки будут созданы только для опубликованных записей.

URL для генерации этикеток можно ввести вручную, указывая через запятую ID тех гербарных записей, для которых необходимо сгенерировать этикетки.

Примечание: Пример: <https://botsad.ru/hitem/pdf/33682,33682,33682,33680>

Такой подход может быть полезен в случае, если необходимо сгенерировать несколько одинаковых этикеток, соответствующих определенной гербарной записи.

Генерирование этикеток путем ввода URL возможно только после авторизации пользователя в системе.

Пример этикетки

3.6 Этикетка-Конверт

Некоторые гербарные сборы предпочтительно хранить в конвертах. С этой целью система предлагает специальный формат этикетки «Этикетка-Конверт».

Этикетка-Конверт представляет собой лист формата А4, в нижней трети которого, представляющей лицевую сторону конверта, располагается подробная информация о сборе, а также штрих-код, оформленный в соответствии с правилами, описанными в разделе *Штрихкодирование образцов*; в центре страницы, соответствующей оборотной стороне конверта, размещается QR-код, кодирующий ссылку на персональную страницу образца;

Для генерации этикеток в общем списке гербарных записей панели администрирования необходимо выделить образцы, выбрать действие – «Создать этикетки-конверты» и нажать «Выполнить».

За один запрос можно сгенерировать не более 100 этикеток-конвертов.

Первый вид, указываемый на этикетке-конверте – основной вид сбора, далее, каждый на новой строке, идут дополнительные (сопутствующие) виды.

После располагается блок информации о сборе; в этом блоке отображаются:

- Страна
- Регион
- Район
- Координаты сбора
- Статус координат (если координаты получены при помощи gps, то это отмечается строкой „[GPS-based]“;
- Кто собрал и дата сбора
- Заметки
- Место сбора

Далее, идут примечания к видам, а также к основному сбору; Переопределения основного вида интегрируются с примечаниями к нему.

Пример этикетки-конверта

3.7 Штрихкодирование образцов

Назначение инвентаризационных номеров важный этап систематизации гербарных накоплений. Штрихкод размещается на гербарном листе перед его сканированием и представляет собой уникальный идентификатор данной гербарной записи, формат которого в текущий момент принимается мировым научным сообществом.

Для генерации страницы штрихкодов в общем списке гербарных записей панели администрирования необходимо выделить образцы, выбрать действие – «Создать штрихкоды» и нажать «Выполнить».

За один запрос можно сгенерировать не более 100 штрихкодов; они размещаются автоматически оптимальным образом на странице формата А4. Если штрихкоды не помещаются на одной странице, генерируется многостраничный pdf-документ.

Штрихкоды можно генерировать для всех (не обязательно опубликованных) образцов, внесенных в базу.

В качестве алгоритма для создания штрихкодов используется [CODE39](#), а кодируемая строка имеет вид ACRONYMXXXX, где XXXX – уникальный числовой код (ID) образца внутри данного акронима. Под штрихкодом дублируется кодируемая им строка (размер шрифта фиксирован). Над штрихкодом приводится название организации (размер шрифта динамический, выбирается так, чтобы название организации не выходило за границы штрихкода). Оформление штрихкода сделано по образцу электронного гербария [KEW](#).

Пример документа со штрихкодами

3.8 Известные виды

Все известные виды представлены в трех таблицах – таблице семейств, таблице родов и, собственно, названий видов. Названия видов с авторами привязаны к таблице родов, записи таблицы родов – привязаны к таблице семейств.

Таблица названий видов используется для формирования подсказок при заполнении поля **Вид** формы гербарного образца.

Каждая запись таблицы видов имеет дополнительный статус – «From plantlist» (из базы theplantlist.org), «Approved» (проверенный), «Deleted» (удаленный) и «Recently added» (новый, недавно добавленный).

Названия видов, имеющие статус «From plantlist» или «Approved» считаются доверенными, и образцы, в которых участвуют такие виды, могут быть беспрепятственно опубликованы куратором гербария.

В случае, если название вида имеет статус «Recently added», и оно участвует в гербарной записи, такую гербарную запись опубликовать не получится.

Чтобы опубликовать такую запись необходимо, чтобы куратор (или суперпользователь) изменил статус вида (проверил вид) на «Approved».

Ни куратор, ни регулярный пользователь не может полностью удалить вид из таблицы известных видов. Вместо этого, куратор может изменить статус вида на «Удаленный» («Deleted»). Виды, имеющие статус «Deleted», не участвуют во всплывающих списках-подсказках при заполнении полей формы гербарного образца. В таблице видов отображаются все виды, в том числе и имеющие статус «Deleted». Таким образом, статус «Deleted» должен использоваться чтобы ограничить результаты поиска во всплывающих подсказках, что может быть полезным чтобы исключить устаревшие и/или неправильные названия видов.

Регулярный пользователь (наборщик гербария) не может изменять статус вида.

На странице редактирования вида можно задать его синоним. К данному виду можно присоединить только один вид-синоним. Если вид имеет несколько синонимов, можно поступить следующим образом. Допустим вид А имеет синонимы В, С, D. Тогда, редактируя страницы видов В, С, D можно в них указать, что они являются синонимами вида А. При выполнении поиска с учетом синонимов, информация об образуемых этим способом классах эквивалентности (синомичности) видов будет использована при формировании запроса к базе гербарных образцов.

В качестве дополнительной защиты от случайного редактирования уже проверенные (имеющие статус «Approved») виды по истечении определенного количества дней «замораживаются». Количество дней с момента последнего редактирования вида до «заморозки» возможности его редактирования определяется параметром `APPROVED_SPECIES_FREEZE`. Текущее значение этого параметра 30 дней. Виды, имеющие другие статусы, в том числе статус «From plantlist», не замораживаются. «Замороженные» виды могут быть отредактированы только сотрудником с правами суперпользователя.

3.8.1 Поля таблицы Известных видов

Название

Название вида – это видовой эпитет. Значение поля хранится в нижнем регистре. Если Вы введете в данное поле, например, **Davurica**, значение будет автоматически переведено в нижний регистр, т.е. **davurica**. Недопустимо включать в данное поле авторов вида (хотя бы потому, что авторы записываются с учётом регистра). Данное поле может включать информацию о вариациях вида или подвидах, например, **yokogurensis subsp. fragilifolia**.

Авторство

Авторство вида. Примеры (через точку с запятой): Maxim; L.; Kom.; (Moench) Mold.; Stephani Данное поле хранится с учетом регистра.

В случае, если для видовой записи заданы подвидовой ранг и подвидовой эпитет авторство относится к объекту, состоящему из (названия рода, видового эпитета, подвидового ранга, подвидового эпитета).

Род

Название рода. Должно выбираться из выпадающего списка предложенных названий. Если нужное наименование рода отсутствует, его необходимо добавить в систему нажав «Добавить/add» (кнопка «зеленый плюс» справа от поля; кнопка может отсутствовать, если Ваш уровень прав не позволяет добавлять **Рода**);

Подвидовой ранг

Возможные значения данного поля: «subsp.», «subvar.», «var.», «f.», «subf.».

Подвидовой эпитет

строка; используется совместно с полем *Подвидовой ранг*. Должна быть пустой, если не задано поле *Подвидовой ранг*.

Статус

Значение поля описано выше.

Синоним вида

Если синоним у вида отсутствует, поле должно быть пустым. Синоним выбирается из выпадающего списка известных системе видов. Если требуемый вид отсутствует, его можно добавить используя кнопку «Добавить/add» («зеленый плюс» справа от поля);

3.9 Изображения

Подготовка изображений для привязки их к гербарным образцам регламентируется отдельным *документом*.

Подготовка и загрузка гербарных изображений

- *Введение*
- *Меню загрузки изображений*
- *Доступ к обработанным файлам*
- *Параметры сжатия изображений*
- *Названия файлов изображений*
- *Калибровочные ячейки*

Предупреждение: Данный раздел документации был существенно переработан, в результате чего были установлены новые правила задания имен гербарных изображений.

4.1 Введение

Поскольку каждое гербарное изображение предполагается представлять в высоком разрешении, что повлечет за собой увеличение требуемого места на сервере, хранение таких изображений будет отдельной задачей, минимально связанной с заполнением электронного гербария.

Формирование финальных web-страниц гербарных образцов будет в этом случае осуществляться с подгрузкой соответствующих изображений, которые должны быть строго структурированными и привязанными к конкретным гербарным образцам, учитывая их уникальные номера и принадлежность к акронимам.

С целью такой привязки, а также удобства последующей автоматизированной обработки изображений (создания малых их копий для предпросмотра и т.п.) ниже приводится система соглашений о размещении и наименовании изображений.

4.2 Меню загрузки изображений

Загрузка гербарных изображений в текущей реализации системы возможна двумя способами: централизованно, из локальной сети Ботанического сада-института ДВО РАН, или посредством Сервиса загрузки изображений, который доступен по адресу.

Сервис загрузки изображений доступен только пользователям, обладающим правами куратора гербария или суперпользователя.

Для пользователя, обладающего необходимыми правами для загрузки изображений, страница администрирования, где представлен перечень гербарных записей будет иметь вид, как на (Рис. 1).

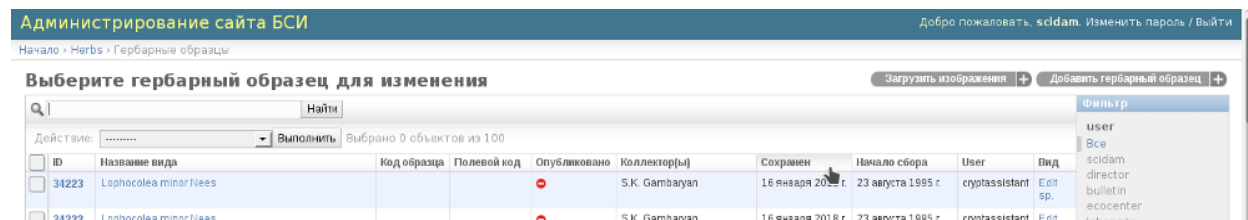


Рис. 1. Фрагмент списка образцов и дополнительное меню «Загрузить изображения» (кнопка)

При переходе по ссылке «Загрузить изображения» появится окно (Рис. 2).

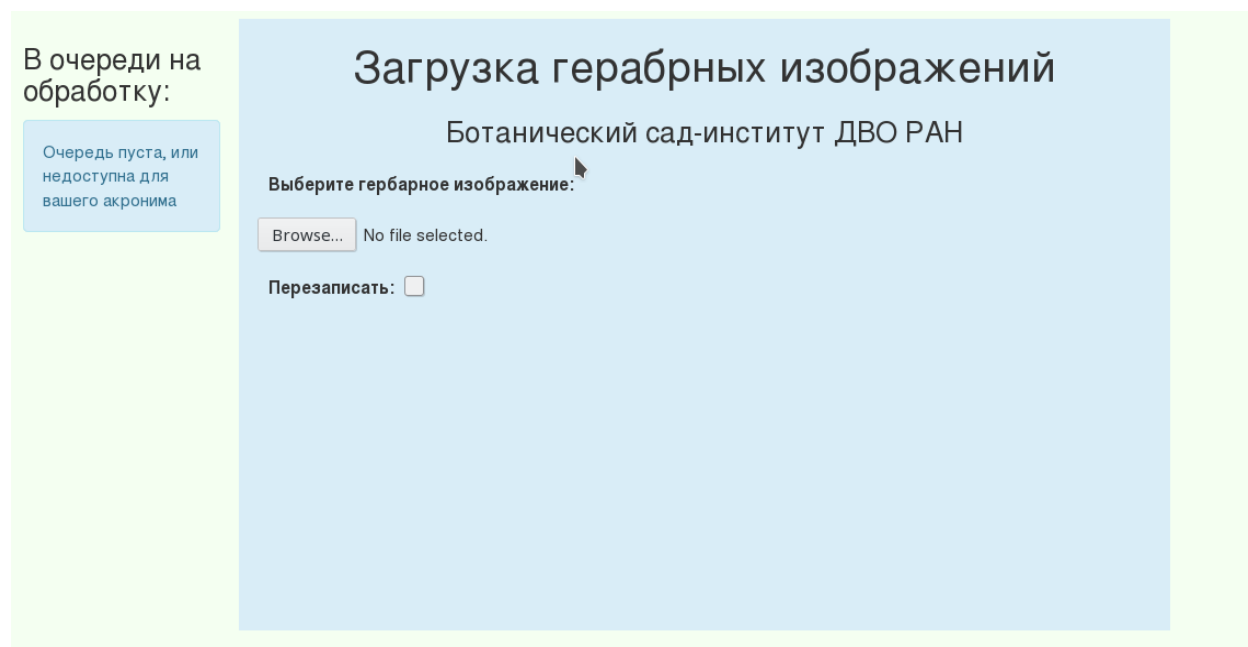


Рис. 2. Сервис загрузки изображений

При нажатии кнопки **Browse...** и выборе файла изображения происходит валидация имени изображения на предмет:

- Соответствия имени файла установленному формату;
- Соответствия гербарного акронима, извлеченного из имени изображения, и акронима,

которому принадлежит пользователь, выполняющий загрузку изображения; * Наличие файла с таким же уникальным номером и принадлежащим тому же акрониму;

Загрузки файла на удаленный сервер на этапе валидации не происходит.

При возникновении ошибок в процессе валидации система выведет сообщения об ошибке, см. например (Рис. 3).

The screenshot shows a web form titled "Выберите гербарное изображение:" (Select herbarium image:). It contains a "Browse..." button followed by the text "3.jpg". Below this is a checkbox labeled "Перезаписать:" (Overwrite:). At the bottom, there is a red rectangular error message box containing the text "Неправильное имя или формат файла" (Incorrect name or file format).

Рис. 3. Пример ошибки валидации: имя файла не соответствует принятому формату.

В случае, если файл с таким же именем уже загружен в систему, или находится в очереди на обработке, чтобы его заменить новым, необходимо отметить флаг <Перезаписать>.

Файлы, находящиеся в очереди на обработке, с определенной периодичностью (ориентировочно, раз в три часа), автоматически конвертируются и публикуются. С этого момента они являются доступными всем в сети Интернет.

После того, как валидация имени файла пройдена, будет доступна кнопка «Загрузить» (Рис. 4), нажатие которой приведет к загрузке выбранного файла на удаленный сервер и постановку его в очередь на обработку.

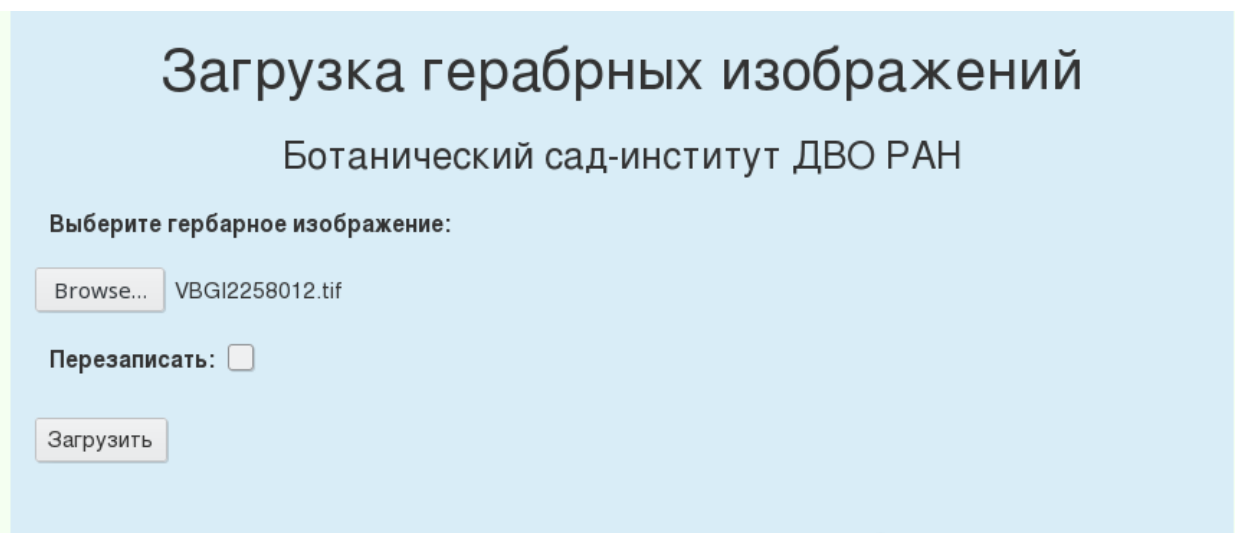
The screenshot shows a web form titled "Загрузка гербарных изображений" (Upload herbarium images) with the subtitle "Ботанический сад-институт ДВО РАН" (Botanical Garden-Institute DVO RAS). It contains the label "Выберите гербарное изображение:" (Select herbarium image:). Below it is a "Browse..." button followed by the text "VBGI2258012.tif". There is a checkbox labeled "Перезаписать:" (Overwrite:). At the bottom, there is a button labeled "Загрузить" (Upload).

Рис. 4. Кнопка «Загрузить» доступна при успешной валидации имени файла

Поскольку файлы изображений могут быть очень большого объема (более 100 МБ), поэтому их загрузка на сервер может занять определенное время.

По завершении загрузки, файл будет добавлен в очередь, а также появится уведомление (на светло-зеленом фоне), что «Файл загружен» (Рис. 5).

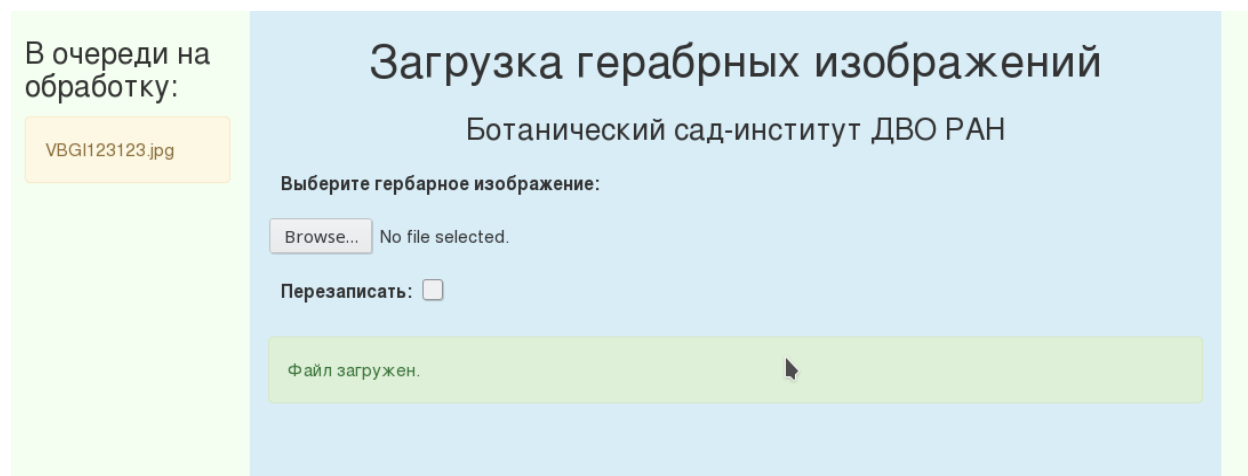


Рис. 5. Пример успешной загрузки файла и постановки его в очередь на обработку.

В левой колонке «В очереди на обработку» отображаются загруженные файлы, но не прошедшие обработку. При этом отображаются только те файлы, которые соответствуют тому же акрониму, что и акроним, которому принадлежит текущий пользователь, работающий в системе. Суперпользователь может видеть все файлы, находящиеся в данный момент в очереди на обработку.

4.3 Доступ к обработанным файлам

После обработки файлы доступны для свободной загрузки по сети Интернет.

URI файла при этом формируется из его названия, акронима и пути, указанном в параметре `SOURCE_IMAGE_URL`. Текущее значение этого параметра

`SOURCE_IMAGE_URL = http://botsad.ru/herbarium/view/snapshots`

Например, если файл *VBGI30124.tiff* был загружен в систему, тогда он доступен по протоколу (HTTP/HTTPS) по адресам:

- <http://botsad.ru/herbarium/view/snapshots/VBGI/ts/VBGI30124.jpg>
- <http://botsad.ru/herbarium/view/snapshots/VBGI/ss/VBGI30124.jpg>
- <http://botsad.ru/herbarium/view/snapshots/VBGI/ms/VBGI30124.jpg>
- <http://botsad.ru/herbarium/view/snapshots/VBGI/fs/VBGI30124.jpg>

Фрагменты пути *ts*, *ss*, *ms*, *fs* — используются, чтобы обозначить размер соответствующего файла (конкретные значения размеров, определены в файле спецификации: *ts* (thumbnail size) – минимальный размер, *ss* (small size) – маленький размер, *ms* (medium size) – средний размер, *fs* (full size) – полный размер.

Предупреждение: Веб-интерфейс (<http://botsad.ru/herbarium>) и HTTP API интерфейс (<http://botsad.ru/hitem/json>) «узнают» о наличии новых гербарных снимков в течение суток. Однако, файл изображения становится доступным (по указанным в примере выше ссылкам) для загрузки по сети Интернет уже после своей обработки и удаления из очереди.

4.4 Параметры сжатия изображений

Для конвертирования изображений используется набор утилит обработки изображений [ImageMagick](#), при этом параметры, используемые для сжатия, определяются в переменной конфигурации: `IMAGE_CONVERSION_OPTS`.

Примечание: Все изображения конвертируются в формат **.jpg*. Текущие параметры сжатия:

```
'-strip', '-interlace', 'Plane',
'-sampling-factor', r'4:2:0',
'-quality', r'90%'
```

Практика показывает, что используемый набор параметров не оказывает видимого эффекта на изображения. Вместе с тем, использование сжатия позволяет существенно сократить требуемое дисковое пространство.

4.5 Названия файлов изображений

Файлы загружаемых изображений должны быть установочного формата:

ACRONYMXXXX<_XX>[.tif][.jpg][.jpeg]

- ACRONYM – название гербарного акронима (заглавными буквами), которому принадлежит гербарный образец; например, VBG1, SAKH, ABGI и др.
- XXXX – уникальный числовой код образца (**ID** образца, назначаемый системой);
- <_XX> – **опциональный код; указывается после символа _**, задается пользователем, используется если данному гербарному сбору соответствуют несколько изображений;
- [.tif][.tif][.jpg][.jpeg] – расширение файла (регистр расширения не имеет значения), указывающее на его формат (одно из перечисленных в квадратных скобках значений);

Примеры валидных имен файлов изображений:

- VBG1123.jpg
- VBG1123_1.jpg
- ABGI9933.tif
- SAKH3842.jpeg

4.6 Калибровочные ячейки

Крайне рекомендуется на сканах гербарных образцов представлять калибровочные ячейки, и желательно, различных цветов.

Калибровочные ячейки используются для автоматизированного пересчета соответствия между «количеством пикселей» и принятыми единицами длины, а также для автоматической цветокоррекции изображений.

Калибровочные ячейки должны быть строго одинаковыми для всех изображений, размером — 1 см x 1 см.

Установленный формат калибровочных ячеек можно найти по [ссылке](#)

Пример гербарного изображения с калибровочными ячейками дан ниже.



Цитирование образцов

Раздел в разработке.

Нужно разработать систему цитирования, которая покрывает возможные варианты ссылки на гербарный объект: по его инвентарному номеру, по его уникальному ID, по полювому коду.

Symbols

акроним (поле), 25
акронимы гербария, 21
аутентификация, 20
автоматизация поиска, 11
безопасность, 20
биоморфологический статус (поле), 27
цитирование, 39
длительность сессии, 20
дополнительные виды, 28
экоусловия, 26
этикетка, 29
этикетка-конверт, 29
фильтр поиска, 6
гербарная запись, 21
группы, 22
история переопределений вида, 27
карта, 8
код образца, 25
конец определения, 27
конец сбора, 26
координаты сбора, 26
куратор, 22
куратор подраздела, 22
локализация, 26
меню загрузки изображений, 33
место сбора, 26
мультивидовые сборы, 28
мультиязычность, 23
наборщик, 23
начало определения, 27
начало сбора, 26
определили, 27
опубликовать запись, 27
персональная страница образца, 28
подраздел гербария (поле), 27
подразделы гербария, 22
поиск, 3
поиск по области, 8

полевой код, 25
пользователи, 22
поля формы, 23
примечание к виду, 24
пример гербария с калибровочными ячейками, 38
работа из Python, 18
работа из R, 18
район, 26
регион, 25
сессия, 20
сходство, 24
синоним вида, 32
сканирование образцов, 33
собрали, 26
сохранение изображений образцов, 37
создание этикетки, 29
создание этикетки-конверта, 29
статус вида, 30
страна, 25
структура, 21
структура JSON-ответа, 13
суперпользователь, 22
штрихкодирование, 30
таблица известные виды, 30
таблица известные виды (поля), 31
типовой статус, 25
высота сбора, 26
вид, 24
загрузка изображений, 33
заметки о сборе, 27
запомнить текущую запись, 28

G

gps (поле), 26

H

HTTP API, 11
HTTP API ограничения, 18
HTTP API ответ, 13
HTTP API проверка работоспособности, 18

HTTP API запрос, [11](#)

J

JSON структура дополнительных видов, [16](#)

JSON структура переопределений вида, [16](#)

JSON-ответ, [13](#)

W

web-форма, [3](#)