
**ИНТРОДУКЦИЯ И ПЕРСПЕКТИВЫ ИСПОЛЬЗОВАНИЯ
РАСТЕНИЙ В ИНДУСТРИАЛЬНОМ РЕГИОНЕ**

УДК 004(62+65)+582

И.И. Стрельников, С.А. Приходько, А.З. Глухов, А.В. Николаева

**РАЗРАБОТКА ИНФОРМАЦИОННОЙ СИСТЕМЫ
ДОНЕЦКОГО БОТАНИЧЕСКОГО САДА**

Государственное учреждение «Донецкий ботанический сад»

Статья посвящена аспектам разработки Информационной системы Донецкого ботанического сада, основной целью которой является управление таксономическими списками коллекций и учет живых особей растений. На рекогносцировочном этапе проанализированы текущие подходы к ведению баз метаданных коллекций. Выявлен ряд потенциальных ограничений и недостатков: дублирование данных, накопление ошибок, отсутствие стабильного подхода к учету отдельных особей. Предложенная модель Информационной системы призвана устранить обозначенные недостатки и унифицировать процедуры обращения с данными для всех отделов ботанического сада. Система реализована как веб-приложение с применением современных технологий. Задача повышения надежности и точности таксономических данных решена с использованием таксономической основы. Важной особенностью системы является использование ГИС-технологий для учета пространственного положения особей.

Ключевые слова: Информационная система, база данных, Донецкий ботанический сад, Python, Catalogue of Life

Цитирование: Стрельников И.И., Приходько С.А., Глухов А.З., Николаева А.В. Разработка Информационной системы Донецкого ботанического сада // Промышленная ботаника. 2021. Вып. 21, № 1. С. 36–52.

Введение

В последние годы в Донецком ботаническом саду реализуется масштабный проект по систематизации и подведению итогов 50-летнего опыта интродукционных экспериментов. Эта работа потребовала применения процедур, принципиально отличающихся от исторически сложившихся подходов. Прежде всего, возникла необходимость полной унификации и объединения метаданных коллекций разных научных отделов. Кроме того, при вовлечении в анализ исторических записей необходимо учитывать отсутствие единой точки зрения на систематическое положение отдельных таксонов. Например, для некоторых образцов ранее использованные названия к настоящему моменту могут быть признаны синонимами.

Сотрудниками Донецкого ботанического сада накоплен большой опыт в решении задач учета коллекционного фонда и формирования сопутствующих баз метаданных. Все коллекции сопровождаются таксономическими списками, ин-

формацией о происхождении образцов, ведутся записи фенологических наблюдений и прочих биологических характеристик. Специфическим элементом метаданных являются этикетки на образцах растений. Все же реализация нового проекта выявила ряд недостатков в текущих подходах, что привело к формулированию задачи реорганизации процедур управления базами сопутствующей информации.

Мировой опыт демонстрирует, что процедуры управления метаданными коллекциями могут быть оптимизированы путем внедрения компьютеризированных информационных систем [2, 3]. Такие системы обеспечивают надежность хранения и доступа к информации, а также позволяют автоматизировать ряд рутинных процедур [18]. Вместе с тем очевидно, что применяемые системы должны в первую очередь соответствовать целям и задачам конкретной коллекции. Поэтому разработка информационной системы должна опираться

ся на сложившиеся процедуры работы с информацией и учитывать ранее накопленный опыт [5, 6].

Цель и задачи исследований

Цель – разработать унифицированную систему учета коллекций и реализовать ее в виде компьютеризированной Информационной системы.

Для достижения цели предусмотрено решение ряда теоретических и практических задач:

1) проанализировать текущие подходы к формированию метаданных коллекций и выявить их проблемные аспекты. Предложить пути решения выявленных осложнений;

2) разработать процедуру нормализации таксономических данных, которая позволит минимизировать вероятность внесения ошибочных записей и упростит поиск синонимических названий таксонов;

3) реализовать геоинформационную систему для учета фактического местоположения растений на коллекционных участках;

4) практически реализовать Информационную систему в виде компьютерного программного продукта.

Объекты и методики исследований

Разработка Информационной системы Донецкого ботанического сада проводилась с использованием свободного программного обеспечения. Операционная система – Ubuntu 18. Преимущественная интегрированная среда разработки (IDE) PyCharm Community Edition. Версионирование кода реализовано с применением git с удаленным хранением репозитория в инфраструктуре Atlassian Bitbucket. При разработке реализованы техники модульного тестирования с применением библиотеки pytest. Управление проектом осуществлялось на основе метода «канбан» с реализацией досок в сервисе Trello.

Результаты исследований и их обсуждение

В результате проведенных исследований разработана унифицированная система учета коллекций в виде компьютеризированной Информационной системы.

Анализ текущих подходов ведения баз метаданных и сопряженные с ними проблемы

Первичное ознакомление с процедурами формирования и поддержания баз метаданных в раз-

ных отделах Донецкого ботанического сада выявило ряд потенциальных проблемных аспектов.

Прежде всего, каждый отдел и сотрудники, ответственные за отдельные коллекции, выработали в ходе работы схемы обращения с данными, удобные для выполнения их конкретных задач. В большинстве случаев метаданные хранятся в электронных таблицах, созданных с помощью табличного редактора Microsoft Office Excel. Кроме того, электронные таблицы могут дублироваться картотеками на бумажных носителях.

Недостатки ведения базы метаданных на основе табличных редакторов

Следует отметить, что электронные таблицы имеют большую популярность для организации разного рода структурированных данных, поэтому их использование для хранения метаданных коллекции растений, само по себе, не представляет проблему. Фактическими недостатками является:

1. Отсутствие единой для всех отделов и коллекций структуры таблиц для хранения метаданных. Например, в разных таблицах указывалось разное количество таксонов высокого уровня: отдел, класс, порядок, семейство. В ряде случаев присутствовали дополнительные колонки для обозначения нототаксонов. Кроме того, в разных подразделениях таксономические записи сопровождались разными блоками дополнительной информации: год поступления в коллекцию, экологическая группа, указание природного распространения и т.д.

2. Структура таблиц не оптимизирована под автоматизацию компьютерной обработки, что становится действительной проблемой при объединении данных в одну большую базу. Например, к началу описываемой работы в большинстве таблиц научные названия растений (родовой, видовой и внутривидовые эпитеты, метки внутривидовых таксонов и аббревиатура автора) подавались в одной колонке единой текстовой строкой, что противоречит одному из базовых принципов организации табличных данных – одна колонка должна представлять только одну категорию данных [22]. Имеющийся формат практически исключал возможность автоматического подсчета таксономического состава, например, определение суммарного количества представленных родов, видов, культиваров и прочее.

3. Ручной ввод и обновление данных приводят к накоплению ошибок. Особенности создания и учета коллекций живых организмов требуют регулярного обновления, в частности таксономических данных. В некоторых случаях организм может быть переопределен, в результате чего его название в таблице также требует изменения. Нередки случаи изменения принятого таксономического положения некоторого организма. Например, объединение или дробление родов и более высоких таксонов, признание названия синонимичным другому названию и тому подобное. Учитывая, что научные названия в основном представлены латынью, нередко во вносимых записях допускаются опечатки. В дальнейшем опечатки и прочие ошибки накапливаются. Особенно остро эта проблема стоит для записей авторов таксонов. Этот блок информации является чувствительным к регистру (необходимость верного сохранения заглавных и прописных символов), он может содержать знаки пунктуации и прочие символы, например скобки. Кроме того, в аббревиатуру автора могут входить не только латинские буквы, но и специфические символы латиницы, включая комбинации с диакритическими знаками. Например, запись автора *É. Morgen* содержит диакритический символ. В таком виде строка может попасть в таблицу, например, если она была скопирована из первоисточника. При ручном же наборе наиболее вероятно, что запись будет подана без диакритического знака – *E. Morgen*. Это не создаст трудностей при визуальном инспектировании таблицы, но при компьютерной обработке такие записи будут распознаны как разные. В результате один и тот же вид может быть неверно посчитан дважды.

Проблема дублирования данных

В ряде случаев данные разных коллекций хранятся и поддерживаются в отдельных таблицах. При этом имеются общие таблицы, объединяющие данные о коллекциях всего подразделения. Причина формирования такого подхода очевидна – проблематичность совместного редактирования единой таблицы группой сотрудников. Именно поэтому кураторы отдельных коллекций формируют индивидуальные наборы данных, которые периодически сводятся в единую таблицу. Теоретически такая организация работы вполне оправдана. Все же присутствие ограничений, описанных выше, регулярно создает проблемы

объединения и поддержания данных в актуальном состоянии. В частности, возникают случаи дублирования названий, если один таксон присутствует в разных коллекциях и его записи содержат ошибки или минимальные отличия в написании, вплоть до случайного удвоения знака пробела. В результате процесс сведения таблиц оказывается достаточно трудоемким. Решение конфликтов в данных требует многократного привлечения кураторов коллекций.

Проблема отсутствия учета особей (отдельных растений)

В текущем состоянии данные коллекционно-го фонда в основном представлены таксономическими таблицами. При этом фактическими элементами коллекции являются отдельные растения. Данная особенность продиктована использованием единой таблицы для учета всех данных. В таком варианте учет репрезентативных особей потребовал бы дублирования таксономических записей, что затруднило бы работу с таблицей.

Следствие – отсутствие фиксированной процедуры для учета таксонов, представленных несколькими особями, особенно если речь идет о растениях, полученных в разное время из разных источников.

Еще одним проблемным аспектом такого подхода является отсутствие учета местоположения особей в коллекциях. Фактически этот недостаток должен нивелироваться должной маркировкой. Но в ряде случаев этикетирование растений оказывается неполным, например, для однолетних, мелких растений или тех таксонов, которые представлены большим количеством особей. В результате задача поиска фактического представителя коллекции становится доступной только ее куратору, но очень затруднена для прочих сотрудников.

Принципы решения выявленных проблем

Проведя детальный анализ выявленных недостатков текущих подходов к ведению мета-данных коллекций, были предложены следующие решения.

1. *Использование для хранения данных полноценной реляционной базы данных.* Основным недостатком текущей схемы можно назвать отсутствие консистентности единой структуры данных и их дублирование. Очевидным решением является использование реляционной базы данных с отдельными таблицами для обособленных блоков информации. При этом следует раз-

работать и реализовать структуру связей между таблицами и систему ограничений на вносимую информацию. Тем самым сама структура базы данных должна запрещать внесение дублирующихся записей, поддерживать структуру взаимоотношений и принадлежности одних записей к другим. Например, запись об особи должна обязательно ссылаться на запись об одном и только об одном таксоне.

2. *Многопользовательское веб-приложение с централизованным хранилищем.* Набор метаданных коллекций предполагается хранить в виде единой базы, размещенной на сервере. При этом пользователи будут иметь неограниченный доступ к системе (с удаленных компьютеров и мобильных устройств, подключенных к интернету). Для удобства кураторов коллекций должен быть обеспечен быстрый доступ к выборкам таксонов и особей, что заменит необходимость ведения отдельных таблиц для каждой коллекции.

Решаемая проблема – вариации в структуре данных, дублирование данных, разрозненность данных.

3. *Система автоматизации и полуавтоматической проверки правильности написания названий таксонов.* Разрабатываемая система должна обеспечить полуавтоматическую проверку точности внесения таксономических названий. Для этого предложено использовать таксономическую основу, содержащую общепринятые написания таксонов и их статус (синоним или базионим). В качестве основы использована ежегодная выгрузка из базы данных проекта Catalogue of Life. Данный проект представляет собой объединение ряда таксономических баз данных, курируемых авторитетными исследовательскими организациями, такими как Ботанический сад Миссури, Королевский ботанический сад КЕУ, подразделение биоразнообразия Академии наук Китая, Смитсоновский институт и прочие (всего 176 наборов данных). База данных проекта включает таксономические названия живых организмов (по возможности с указанием библиографической ссылки на первоописание). Названия разделены на базионимы (принятые названия) и синонимы. Решение об отнесении названия к одной из категорий основывается на мнении экспертов, курирующих соответствующую базу [15]. Очевидно, что принятая в проекте таксономическая система (как и любая другая, претендующая на

глобальный охват) не может быть идеальной и верной во всех отношениях. Все же, база данных Catalogue of Life является крупнейшей открытой таксономической базой, названия в которой проходят экспертную проверку.

Решаемая проблема – минимизация опечаток и накапливаемых ошибок. Минимизация неопределенности в использовании синонимов научных названий.

4. *Четкое разделение таксономического названия и особи как двух взаимосвязанных, но отдельных блоков информации.* При этом под особью подразумевается фактическое растение, для которого могут быть отмечены данные о дате его поступления, высадки, гибели; его принадлежность к подразделению Донецкого ботанического сада; месторасположение на территории и прочее. В базе данных запись об особи должна иметь реляционную связь с записью таксономического названия по схеме «много к одному». То есть, одна особь может ссылаться только на одну таксономическую запись, но на одну таксономическую запись могут ссылаться несколько особей. Таксономические данные и записи об особях должны храниться в разных таблицах базы данных. В таком варианте изменение таксономической записи (внесение правок) не повлияет на данные особи. Схожим образом, при изменении таксономической принадлежности особи (например, при переопределении) изменяется только ссылка на таксон, но не сами записи таксона или особи.

Решаемая проблема – отсутствие учета отдельных растений.

5. *Модуль ГИС.* В системе должна быть представлена наглядная схема территории Донецкого ботанического сада с детализацией, достаточной для рассмотрения отдельных экспозиций. Для каждой особи растений могут сохраняться координаты местоположения на схеме. При этом модуль ГИС должен реализовать полноценную поддержку пространственных запросов. В частности, пользователь должен иметь возможность выбрать только те особи или таксоны, которые представлены на некотором участке.

Решаемая проблема – привязка данных об особях к их физическому местоположению; быстрый поиск образцов коллекции без необходимости детальных знаний о группе растений [8].

6. *Модуль фотографий.* Должна быть обеспечена возможность сохранения и удобного ото-

бражения фотографического материала для каждой таксономической единицы. Помимо сугубо эстетической роли, этот модуль должен обеспечивать упрощение идентификации образцов коллекции.

Программная реализация

Информационная система Донецкого ботанического сада была запланирована как веб-приложение (интернет-сайт) с богатым пользовательским интерфейсом, обеспечивающим внесение, первичную обработку, фильтрацию и презентацию таксономической и сопутствующей информации о коллекциях.

На первых этапах работы рассматривалась возможность использования Систем управления содержимым (CMS) по типу Wordpress, Joomla, Drupal и прочие. Применение CMS обеспечивает упрощенное, высокоуровневое создание и управление контентом сайтов. Все же этот подход был отклонен, так как разрабатываемая Информационная система должна реализовывать ряд специфических функций, которые не представлены в классических CMS и их дополнениях. Кроме того, самые распространенные системы управле-

ния содержимым основаны на языке программирования PHP, который плохо знаком сотрудникам Донецкого ботанического сада.

Нами было принято решение применить Каркас веб-приложений (Web application framework) Django 3.1 с использованием языка программирования Python 3.8. Несмотря на то, что данный подход намного более трудоемкий, чем применение CMS, он обеспечивает полный низкоуровневый контроль над всеми аспектами хранения, перемещения, преобразования и презентации информации.

Основываясь на заявленных требованиях к Информационной системе был подобран стек технологий, представленный в таблице 1.

Помимо основных технологий при разработке использовалось большое количество вспомогательных сторонних библиотек языков программирования Python и JavaScript – 90 и 46 соответственно.

Организация данных

Основой Информационной системы является реляционная база данных. Отдельные блоки информации представлены в разных таблицах, логи-

Таблица 1. Перечень основных технологий, использованных для разработки Информационной системы Донецкого ботанического сада

Наименование	Роль
Python 3.8 [17]	Основной язык программирования серверной части приложения
Django 3.1 [7]	Каркас веб-приложения. Реализация основной логики серверной части приложения
Celery 5.0 [4]	Менеджер распределенных задач – обеспечение выполнения асинхронных процедур обработки информации
PostgreSQL 12 [16]	Основная система управления базами данных (СУБД), реализующая парадигму реляционных баз данных
PostGIS 3.1.1 [13]	Расширение PostgreSQL, обеспечивающее поддержку геопространственных данных
Redis 6.0.5 [14]	СУБД с парадигмой базы данных «ключ-значение» для хранения промежуточных данных (кэш)
Nginx 1.19 [12]	Основной HTTP сервер / обратный прокси-сервер
Gunicorn 20.0 [9]	WSGI HTTP сервер, выполняющий роль сервера приложений Django
Tilestache [19]	Сервер картографических данных (растровые и векторные плитки – raster tiles, vector tiles)
Mapnik 3.0 [1]	Система отрисовки картографических данных (генерирование растровых плиток – raster tiles по данным таблиц)
Materialize.css 1.0.0 [10]	Основной CSS каркас (стиль пользовательской части)
Leaflet 1.7.1 [20]	Отображение картографических данных на стороне пользователя
D3.js 6.5.0 [11]	Отображение интерактивных графиков на стороне пользователя
WeasyPrint 52.2 [21]	Подготовка этикеток особей (система генерирования PDF файлов)

чески связанных между собой. Принципиальная схема базовых таблиц представлена на рисунке 1.

Центральным блоком информации является таблица под условным названием «Таксономическая единица» (далее ТЕ). В данном блоке сохраняются записи о всех научных названиях, которые присутствуют или присутствовали в составе коллекций. Каждая запись ТЕ имеет свой уникальный идентификационный номер (далее ID), на который могут ссылаться записи других таблиц.

Блок «Таксономическая единица» логически ссылается на две таблицы: «Высшие таксоны» и «Данные Catalogue of Life». Обе эти таблицы являются производными архива базы данных проекта Catalogue of Life. Первая содержит перечень всех принятых названий таксонов на уровнях от царства до семейства, каждому из которых присвоен свой ID. Каждая таксономическая единица должна ссылаться на ID семейства, что позволяет

определить принадлежность ТЕ к высшим таксонам. Таблица «Данные Catalogue of Life» выполняет роль таксономической основы и служит для полуавтоматической проверки точности вносимых имен растений. Во время добавления в базу новой ТЕ вносимые данные сверяются с записями этой таблицы. В случае если найдено полное совпадение, запись из таблицы «Данные Catalogue of Life» будет полностью скопирована в блок «Таксономическая единица», и вносимая ТЕ получит ссылку на соответствующий ID. Так как такой тип внесения новых таксонов заявлен в системе как предпочтительный, то в большинстве случаев будут сохраняться записи из таксономической основы, а не внесенные пользователем вручную. Данный аспект призван минимизировать шансы на внесение записей с опечатками и прочими неточностями. Кроме того, в случае изменения информации в проекте Catalogue of

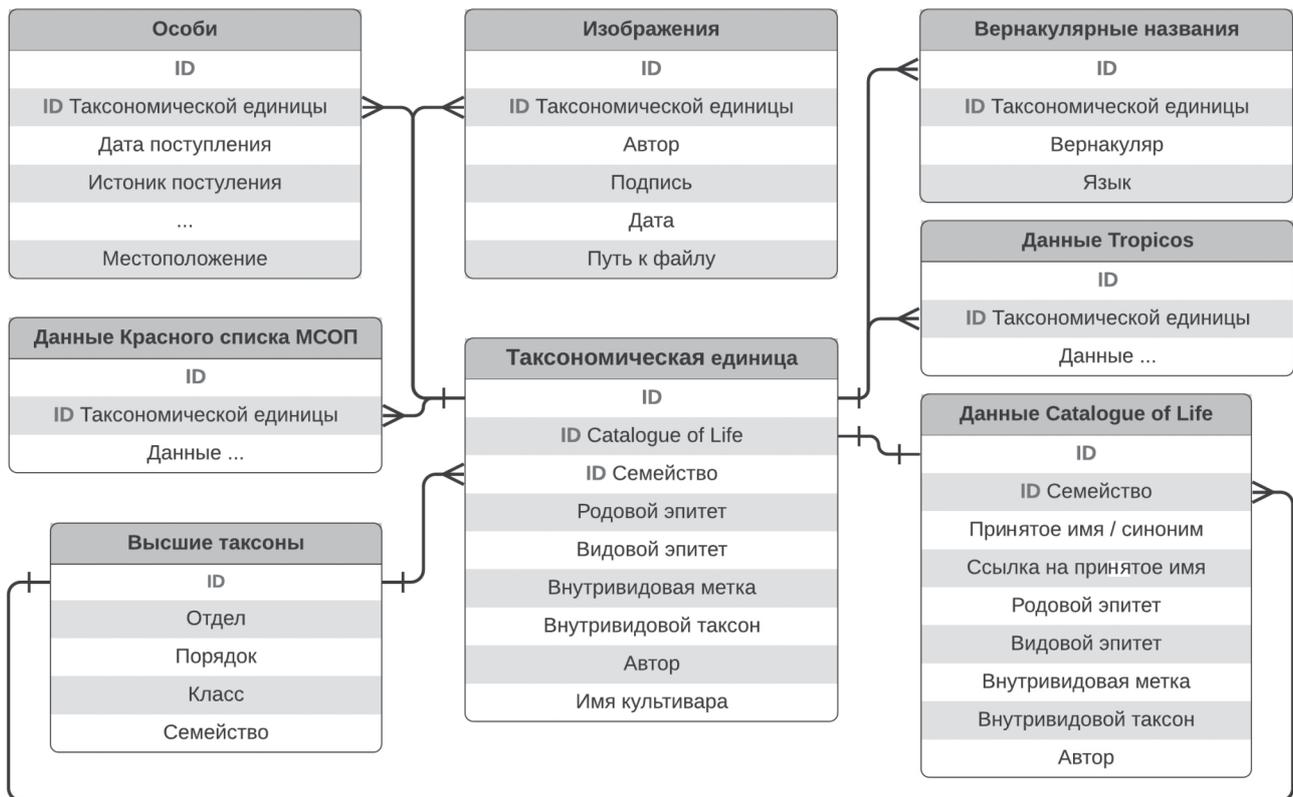


Рис. 1. Схема взаимосвязей между базовыми таблицами базы данных Информационной системы. Отдельные блоки представляют таблицы, полужирным на темном фоне выделены названия таблиц. Строки в блоках отображают названия колонок. Линиями представлены реляционные связи (взаимные ссылки между полями разных таблиц). *для краткости представлены не все таблицы и не все поля в таблицах

Fig. 1. The scheme of relationships between the base tables of the Information system database. Separate blocks represent tables, table names are highlighted in bold on a dark background. The lines in the blocks represent the column names. Lines represent relations (reciprocal links between fields in different tables). *for brevity, not all tables and not all fields in tables are presented

Life данные в Информационной системе могут быть автоматически актуализированы благодаря сохранению ссылок на ID.

Таблица «Особь» содержит подробную информацию о фактических образцах, присутствующих в коллекции сейчас или в прошлом. Под образцом подразумевается единичное растение или пространственно ограниченная группа, например, участок произрастания травянистого растения. Каждая особь имеет уникальный идентификационный номер и ссылается на единственную запись «Таксономической единицы». Подразумевается, что на одну таксономическую единицу может ссылаться неограниченное количество особей. Кроме того, присутствует обязательная ссылка на административное подразделение Донецкого ботанического сада. Таким образом обеспечивается учет принадлежности особей к отделам или лабораториям. Именно особь является базовой единицей коллекции (ранее в этой роли использовалась таксономическая запись). Каждый куратор коллекции принимает самостоятельное решение о количестве вносимых особей. В рамках Информационной системы некоторый таксон считается представленным в коллекции, если к нему приписана хотя бы одна ныне живая «Особь». Для этого каждая особь содержит поле, отвечающее за состояние: живая/выпавшая. Следует отметить, что данный показатель является относительным. Так, например, однолетние растения считаются живыми, если сохраняется их популяция. Поводом для смены статуса может быть полное исчезновение живых особей и материала для размножения.

Таким образом, таблица «Таксономическая единица» является просто списком научных названий, который ничего не говорит о текущем объеме коллекции. Актуальное же состояние коллекции оценивается через учет ныне живых «Особей» и связанных с ними записей «Таксономических единиц». При таком подходе в случае гибели некоторой особи она просто перестает учитываться при оценке размера коллекции, но информация о ней не удаляется и может быть использована, например, для проведения исторического анализа. То есть, Информационная система в принципе не предусматривает удаления записей. Удобство работы с таким, казалось бы, избыточным набором данных обеспечивается выдачей отфильтрованных порций информации. Например,

для некоторого типа работы система выдает списки только «актуальных» таксонов (представленных живыми особями); в другом варианте может быть подготовлена выдача только таксонов или особей, присутствующих в коллекциях в некоторый период времени.

Для каждой особи могут быть внесены дополнительные сведения: дата поступления, источник поступления (город или место сбора из природы), материал поступления (семена, часть или целое растение). Кроме того, для особи может быть отмечено ее пространственное положение на схеме территории.

Для реализации системы фотокаталога реализована дополнительная таблица «Изображения». Фактически, вносимые изображения сохраняются на жестком диске в виде файлов. Данная же таблица содержит ссылки, адреса этих файлов и их привязку к записям о таксономических единицах. Кроме того, для каждого изображения сохраняются записи о дате фотосъемки, авторе изображения и расширенная подпись к изображению.

Остальные блоки представленной структуры ссылаются на таблицу «Таксономическая единица». Например, при внесении новой ТЕ система совершает автоматический запрос к базам проектов Красного списка МСОП и Tropicos. Если вносимое имя присутствует в этих базах, то полученные данные сохраняются в соответствующих таблицах. При этом каждая запись содержит ссылку на ID той таксономической единицы, для которой она была сделана.

Помимо указанных основных таблиц, информационная система подразумевает возможность хранения дополнительных данных, характеризующих таксономические единицы. На данный момент реализованы три таблицы: «Вернакулярные названия», «Экология», «Дополнительные сведения». Записи в каждой из этих таблиц ссылаются на «Таксономические единицы». Таким образом, для каждого таксона может быть указано одно или более вернакулярных, «народных» названий с указанием языка: Русский, Украинский или Английский. В таблице «Экология» может быть сохранена информация об отношении таксона к некоторой экологической группе растений, например: по отношению к увлажнению или освещению. Кроме того, структура таблицы позволяет сохранять данные о жизненной форме и феноритмотипе растения. Во всех случаях пользовате-

лю доступен выбор из заранее подготовленных категорий.

Таблица «Дополнительные сведения» предназначена для хранения неструктурированных данных. Она содержит всего 3 блока информации: категория записи, источник, собственно запись. В ней для любой таксономической единицы может быть сохранена произвольная информация на усмотрение пользователя. Например, подробное описание габитуса, ботаническое описание из некоторой флоры, примеры медицинского использования и так далее.

Информационная система содержит ряд вспомогательных таблиц. Например, данные о текущей административной структуре, таблица со списком зарегистрированных пользователей, таблицы для хранения адресов электронной почты и хэшей паролей. В отдельных специализированных таблицах хранятся пространственные данные, на основе которых формируется визуализация картографической подложки.

Графический интерфейс

Графический интерфейс пользователя реализован как классическое веб-приложение, доступ к которому обеспечивается с помощью интернет-браузера. По сути, для конечного пользователя Информационная система является обычным интернет-сайтом, доступным с любого устройства, обеспечивающего доступ в интернет.

Приложение построено по многостраничной схеме. Передача данных и взаимодействие с ними реализованы с применением протокола HTTP.

Для построения интерфейса использована схема с разделением HTML (структура и содержание страницы), CSS (списки каскадных стилей – визуальное оформление страницы) и JavaScript (вспомогательные функции, реализующие динамическое взаимодействие с элементами страницы в браузере). HTML страницы генерируются серверной частью приложения на основе параметров запросов. CSS и JavaScript компоненты созданы и скомпилированы заранее и представляют собой статические файлы.

Визуальное оформление интерфейса основано на CSS фреймворке Materialize.css. Общий стиль соответствует рекомендациям Material Design, подготовленным специалистами Google. Большое внимание при разработке уделено возможности и удобству работы с сайтом на мобильных устройствах, поэтому приложение полно-

стью отвечает требованиям парадигмы адаптивного веб-дизайна (Responsive Design) – размеры, отображение и компоновка структурных элементов страниц изменяются в зависимости от размера экрана используемого устройства.

Страницы приложения можно разделить на несколько групп, отвечающих разным процедурам работы с Информационной системой. Первую группу составляют формы для внесения и верификации новых данных.

Во вторую группу входят страницы профайлов – индивидуальные страницы, отображающее всю внесенную информацию о таксономических единицах и особях. Профайлы являются основным способом добавления новых данных, связанных с таксонами или особями, или правки ранее внесенной информации.

Третью группу представляют страницы для осуществления поиска по ранее внесенным данным. Реализован «Быстрый поиск» по слову или фразе. Результатом поиска является список всех таксономических единиц, в названиях которых (научных или вернакулярных) содержится фраза или часть фразы из поискового запроса (поиск проводится с учетом синонимичных названий). «Расширенный поиск» позволяет получить список таксонов или особей, отвечающих комбинации критериев, например, по принадлежности к административному подразделению, дате внесения, статусу таксона в международном Красном списке МСОП и так далее. «Пространственный поиск» позволяет получить выборку таксонов, особи которых расположены на одном или нескольких участках территории Донецкого ботанического сада.

Отдельную группу составляют страницы вывода отчетов. Основным типом отображения результатов поиска является табличное представление таксонов или особей. В таком виде подаются: общий список таксонов коллекционного фонда, результаты расширенного поиска, таблица отчета о динамике поступления и выпада образцов. Программно таблицы основаны на JavaScript библиотеке Handsontable, которая позволяет создавать интерактивные представления с функциями, отчасти сходными с табличными редакторами типа Excel. Например, во всех таблицах реализована функция динамической сортировки и фильтрации по колонкам. Кроме табличного вывода присутствуют специфические страницы отчетов,

резюмирующие данные в графической (графики и гистограммы) и картографической форме. В большинстве случаев выводимые данные доступны для скачивания в виде отдельных файлов.

Процедура ввода данных

Одной из ключевых задач Информационной системы является повышение надежности вносимых данных. Прежде всего речь идет о минимизации возможных опечаток в научных названиях. Кроме того, поставлена цель снизить вероятность внесения синонимичных названий как отдельных таксонов. Данные задачи реализуются в процессе ввода новых данных путем последовательных проверок относительно ранее внесенных названий и базы данных Catalogue of Life.

Логически внесение новых данных можно выразить следующей последовательностью: ввод таксономической информации, ввод вернакулярных названий, указание принадлежности к экологическим группам, внесение произвольных дополнительных данных, добавление изображений, внесение данных об имеющихся или ранее имевшихся особях. Процедура ввода реализована с помощью классических веб-форм с применением текстовых полей, выпадающих списков для выбора фиксированных вариантов, чекбоксов для указания бинарных (да/нет) характеристик.

Блок-схема действий пользователя при вводе данных новой таксономической единицы представлена на рисунке 2.

При переходе к соответствующей странице сайта пользователю предлагается заполнить форму с таксономическими данными. По умолчанию, данные сверяются с базой Catalogue of Life, что отмечено положением соответствующего переключателя. В этом случае требуется указать родовое, видовое название и, если необходимо, внутривидовой эпитет и название культивара. Опционально может быть указан автор научного названия. После заполнения всех необходимых полей данные отправляются в виде запроса на сервер. Там информация подвергается ряду проверок.

1. Проверка уникальности. Если внесенные таксономические данные совпадают с уже имеющимся в базе таксономом, то пользователю будет отправлено соответствующее сообщение с указанием ссылки на профайл этого таксона. Таким образом система исключает возможность внесения в базу двух одинаковых записей.

2. Проверка на синонимичность. Если вносимые данные соответствуют синониму ранее сохраненного в системе таксона, то пользователю будет направлено соответствующее пояснение и

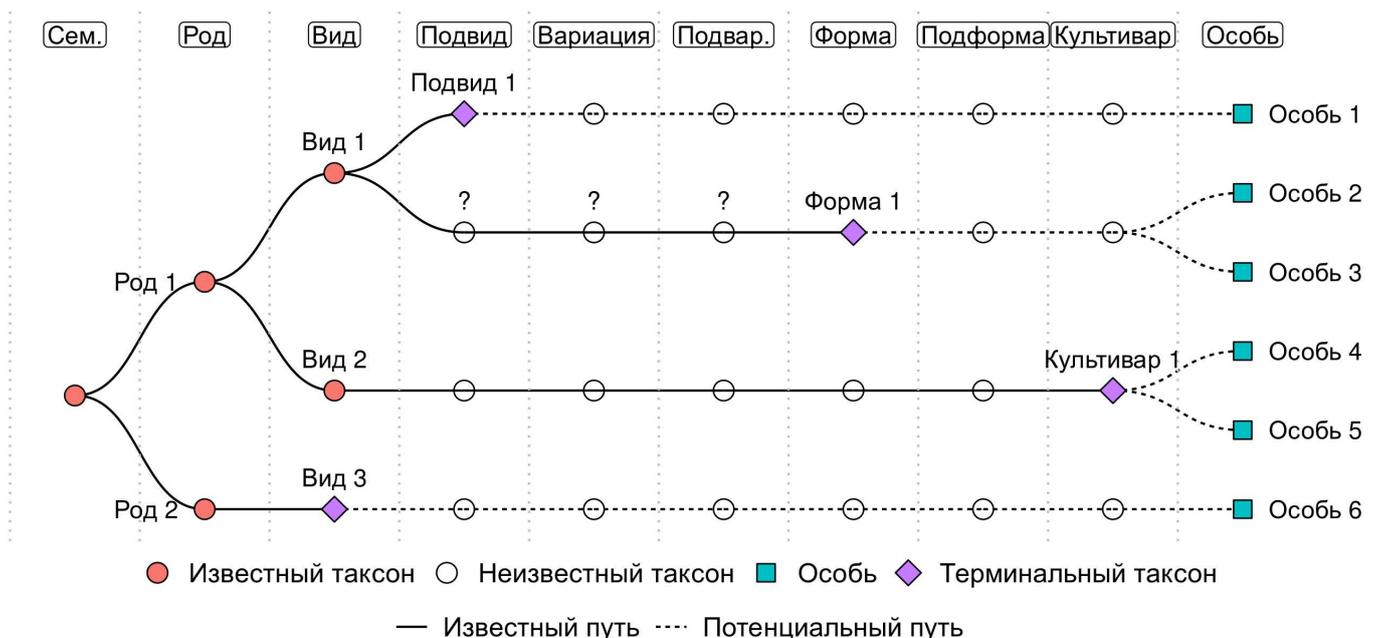


Рис. 2. Блок-схема процедуры ввода нового таксона в Информационную систему Донецкого ботанического сада
Fig. 2. Flowchart of the procedure for entering a new taxon into the Information system of the Donetsk Botanical Garden

ссылка на профайл имеющегося названия (базионима). Проверка производится согласно базе синонимов Catalogue of Life. Если таксономическая позиция пользователя не совпадает с принятой в Catalogue of Life и он предпочитает внести синонимичное название как отдельный таксон, предусмотрено отключение сверки с Catalogue of Life. В этом случае данные будут внесены без проверки. Кроме того, указание о синонимичности будет изъято из проверочной базы, то есть внесенное название в дальнейшем будет рассматриваться системой как базионим.

3. Если вносимое название не является дубликатом или синонимом ранее внесенных таксонов, то система попытается найти совпадающее название в базе Catalogue of Life среди базионимов и синонимов. Основная цель этой проверки – предотвратить возможные опечатки или неверное указание автора и семейства. Система использует полнотекстовый поиск, что позволяет найти частичные совпадения. Если найдено полное совпадение с таксоном из Catalogue of Life, то система копирует в основную таблицу все основные сведения, например, перечень высших таксонов, и переведет пользователя на страницы ввода дополнительных данных. Если во вносимом названии допущены ошибки написания, то процедура вернет список таксонов с наиболее похожими названиями. В таком случае пользователю предлагаются варианты замены вносимого названия. Кроме того, если вносимое название является синонимом, то пользователю будет предложено заменить его на базионим. Как и в предыдущем пункте, пользователь не обязан использовать предложенные замены. Если есть основания считать, что вносимое название является правомерным, то в форме может быть убран флажок с переключателя «Сверять с Catalogue of Life». После этого любое название, внесенное пользователем, будет добавлено в базу (за исключением случаев полного дублирования).

Предложенная процедура сверки с Catalogue of Life обеспечивает повышение точности и единообразия вносимых данных. Все же эта функция может быть отключена пользователем, что важно в случаях если таксон отсутствует в базе Catalogue of Life или его статус вызывает обоснованные сомнения у пользователя, например, отнесение к рангу синонимов.

При внесении данных со всеми проверками система автоматически создаст ссылку на семейство таксона; при отключенных проверках семейство должно быть указано пользователем вручную. Для обеспечения консистентности данных вся таксономия от царства до семейства полностью повторяет данные Catalogue of Life без возможности ее изменения или дополнения пользователем.

После принятия таксономического названия (при успешном выполнении проверок или при их игнорировании) данные перемещаются во временное хранилище, а пользователь перенаправляется на следующие страницы ввода. Последующие формы для ввода вернакуляров, добавления изображений и указания экологических характеристик являются необязательными. Эти данные в последующем доступны для редактирования и добавления со страницы профайла таксономической единицы.

После принятия таксономического названия система совершает ряд автоматических действий. Например, производятся запросы к базам данных Красного списка МСОП и проекта Tropicos. Если указанный таксон присутствует в этих базах, запросы возвращают дополнительные сведения, которые сохраняются в соответствующие таблицы. Среди этих данных могут быть: хромосомные числа с библиографическими ссылками, данные о присутствии таксонов в национальных флорах, статус и категория опасности исчезновения согласно Красному списку МСОП. Если база данных МСОП содержит перечень стран, в которых распространен таксон, система генерирует и сохраняет карту распространения.

По завершении процедуры ввода данных новой таксономической единицы система производит окончательную проверку целостности данных и формирует постоянные записи в соответствующих таблицах. Кроме того, производится обработка фотографического материала, в частности, проводится их оптимизация и создание уменьшенных версий – иконок.

После этого таксономическая единица считается внесенной в базу данных. Пользователю предлагается внести данные об особях это таксона. Как отмечалось выше, единицами учета являются именно особи. Поэтому таксон без живых особей не считается присутствующим в коллекции и не учитывается при подготовке отчетов.

Геоинформационная система

Задача учета пространственного местоположения особей решается с применением подходов геоинформационных систем. Структура реализованной ГИС состоит из трех условных компонентов: пространственные данные, система преобразования данных в визуальное представление (цифровые изображения), система отображения данных на стороне пользователя (как элемент веб-страницы).

Пространственные данные представлены координатами местоположения отдельных особей и объектами, формирующими пространственную схему территории Донецкого ботанического сада. При этом используются географические координаты в проекции EPSG:3395 (Web Mercator).

Для построения схемы территории использовано 4 базовых слоя (таблицы с пространственным типом данных). Первые три слоя – это площадные объекты 3-х иерархических уровней детализации. Первый уровень – участки территории и комплексы построек. Второй уровень – отдельные элементы зданий (например, отдельные теплицы) и более мелкие участки территории. Третий уровень – элементы внутри тепличного комплекса, отдельные экспозиции. Четвертый слой содержит данные о линейных объектах – дороги, тротуары, тропинки и ограждения. Следует отметить, что эти данные представлены в векторной форме (совокупность координат узлов и соотношений между узлами). В таком представлении данные могут быть обновлены или отредактированы с помощью распространенных ГИС-редакторов, например Qgis.

Данные пространственной схемы используются для динамического генерирования карты-подложки. Для этого применяется классическая веб-технология – плиточная веб-карта (tiled web map). Такой подход используется в большинстве картографических продуктов, например Google Maps, Яндекс.Карты и прочие. В рамках этой технологии на стороне пользователя карта формируется из отдельных изображений (плиток) квадратной формы. Для просмотра доступны разные масштабы или уровни детализации (zoom). Технически процесс отображения карты можно описать следующим алгоритмом: пользователь открывает окно с картой; программное обеспечение страницы рассчитывает номера «плиток», которые нужны для формирования карты на дан-

ном масштабе; номера плиток передаются серверу; сервер возвращает цифровые изображения (плитки); программное обеспечение располагает плитки в нужном порядке, чтобы сформировать карту. При перемещении или масштабировании области видимости, программное обеспечение рассчитывает номера новых плиток, которые вновь запрашиваются у сервера, после чего карта обновляется.

В разрабатываемой Информационной системе программным обеспечением для отображения плиточных карт является библиотека leaflet.js. Процесс формирования цифровых изображений из векторных данных реализован на основе программного продукта Mapnik. Визуальное оформление карт производится на основе правил, специально разработанных для проекта. За процедуру обмена данными между сервером и пользовательским приложением отвечает картографический сервер Tilestache.

Данные о местоположении особей передаются браузеру в цифровом виде и отображаются поверх карты-подложки в виде маркеров различной формы и цвета.

Реализованная ГИС позволила добиться поставленной задачи хранения и управления данными о пространственном положении объектов коллекционного фонда. Кроме того, реализованы процедуры пространственного поиска. Пользователь может сделать запрос и получить список-выборку таксонов или особей, расположенных в пределах одного или нескольких участков на карте. При этом данные являются полностью динамическими и независимыми. Если организация территории в будущем изменится, принадлежность особей к участкам будет автоматически переопределена.

Дополнительные функции

Примененная модульная архитектура и решение разрабатывать систему полностью своими силами предоставляют неограниченные возможности для ее дальнейшего развития.

Например, одна из реализованных дополнительных задач – создание стандартизированных этикеток через генерирование PDF файлов, основанных на заранее подготовленных шаблонах. Этикетки формируются для особей (функция доступна со страницы профайла особи). Система автоматически заполняет поля с имеющимися данными и подстраивает размер записей для

оптимального визуального расположения. Длинные названия (например, таксоны с указанием вида, внутривидового таксона и культивара) располагаются в 2 или 3 строки. Аналогично, если надпись не помещается в пределах одной строки, происходит уменьшение размера шрифта. Важным элементом этикеток является QR-код, в котором зашифрован интернет-адрес профайла соответствующей особи. Таким образом пользователь с мобильным устройством, оснащенным камерой, может быстро обратиться к соответствующей странице. Предполагается, что такая процедура облегчит контроль за актуальностью данных, например при перемещении особей, что потребует изменить координаты местоположения в профайле.

Одно из запланированных направлений работы – создание системы электронного гербария Донецкого ботанического сада. Предполагается, что в нем будет задействована уже имеющаяся система контроля таксономических данных. Основная задача будет заключаться в разработке оптимальной схемы хранения и доставки пользователю сканированных версий гербарных листов в высоком разрешении. Данная задача может быть решена с применением уже отработанной схемы оптимизации фотографий таксономических единиц.

Еще одно реализуемое направление – система мониторинга параметров микроклимата. В рамках проекта отдельный модуль Информационной системы будет выступать в роли «хаба» для приема, хранения и обработки данных, поступающих с датчиков. Кроме того, запланированы функции для краткосрочного прогноза температурного режима в оранжерейном комплексе.

Организация количественного учета коллекционного фонда

В ходе предварительного обсуждения выяснилось, что задача формирования количественной сводки по базе, несмотря на кажущуюся простоту, содержит ряд противоречий. Первая целевая статистика такого отчета должна отображать общий объем коллекции – это количество уникальных названий растений. Одновременно есть необходимость подсчета представленности коллекции по отдельным таксономическим уровням (семейство, род, вид и так далее). Несогласованность мнений по поводу принципа формирования такого отчета возникает при попытке сопос-

тавить общий объем и представленность таксономических уровней. Например, на практике при описании коллекций сложилась традиция формировать записи вида: «коллекция содержит N таксонов, включая K видов, L подвидов, ... и M культиваров». При этом возникает проблема отсутствия аддитивности в количествах разных таксономических групп. То есть $N \neq K + L + \dots + M$. Источник проблемы – совмещение концептуально разных показателей: общий объем коллекции (количество уникальных имен, которые могут принадлежать разным уровням иерархии) и количество уникальных таксонов в отдельных иерархических уровнях. При этом нет фактических оснований рассматривать объем родительского уровня как сумму объемов дочерних уровней.

Исходя из этого дополнительной задачей Информационной системы Донецкого ботанического сада стала разработка и внедрение унифицированной системы количественного учета коллекций.

При разработке системы исходили из следующих фактов.

1. Культивар является таксоном внутривидового уровня.

Обоснование – Международный кодекс номенклатуры культивируемых растений (МКНКР). Преамбула пункт 1:

«...Цель данного кодекса предоставить стабильный метод наименования таксонов культивируемых растений...». Здесь и далее в МКНКР культивар называется таксоном.

2. Выделение группы растений в качестве культивара не отрицает их принадлежность к более высоким таксономическим уровням. То есть растения, представляющие культивар некоторого вида, также являются представителями этого вида.

Обоснование:

а. Международный кодекс номенклатуры водорослей, грибов и растений (МКН), статья 28, пункт 3:

«Ничто не препятствует использованию в отношении культивируемых организмов имен, опубликованных в соответствии с требованиями данного кодекса».

б. Наследуемость элементов подмножества в математике.

Для описания релевантности групп организмов современная таксономия использует древовидную иерархическую структуру. Любой таксон является формальным названием для некото-

рого множества организмов. Более низкие таксоны являются подмножествами более высоких таксонов. Например, вид является подмножеством множества организмов, составляющих некоторый род. Также древовидная структура подразумевает, что ни у одного таксона не может быть более одного родительского таксона, например вид не может быть подмножеством сразу двух родов. Кроме того, древовидная структура запрещает существование «висящих» узлов – то есть не допускается, чтобы некоторый таксон (кроме корневого) не имел родительского таксона. Из этого следует, что если в коллекции присутствует организм, являющийся элементом множества организмов, формирующих некоторый таксон, то эта же особь является элементом всех надмножеств организмов, то есть является элементом более высоких таксонов. Например, если в коллекции есть хотя бы одна особь некоторого подвида, то можно утверждать, что в коллекции есть представитель вида, рода, семейства и так далее, которые являются родительскими для указанного подвида.

С учетом вышеуказанных замечаний и положений было решено разграничивать общий объем коллекции и ее таксономический состав по уровням как не пересекающиеся метрики. Под объемом коллекции подразумевается количество уникальных таксономических имен (любого уровня), к которым относятся живые особи коллекции. Таксономический состав считается независимо как количества уникальных названий в отдельных уровнях иерархии.

Описанные принципы представлены в таблице, отображающей некоторую гипотетическую коллекцию (табл. 2) и соответствующий ей дре-

вовидный граф таксономической структуры (рис. 3).

В рассматриваемом наборе данных объем коллекции (уникальные названия) соответствует количеству терминальных узлов графа (рис. 3). Размер коллекции составляет 4 названия (учетные единицы), относящиеся к разным таксономическим уровням от вида до культивара.

Графическое представление также снимает разночтения относительно состава коллекции по отдельным таксономическим уровням. В коллекции есть представители 2-х родов, 3-х видов, 1-го подвида, 1-й формы и 1-го культивара.

Таким образом предлагается следующая процедура подсчета:

А. Для определения объема коллекции (количества учетных единиц/ таксонов):

1. Выделить уникальные строки таблицы с учетом столбцов от рода до культивара.
2. Подсчитать количество уникальных строк.

Пример для таблицы 1:

Уникальные строки:

- 1) Род_1 Вид_1 Подвид_1
- 2) Род_1 Вид_1 Форма_1
- 3) Род_1 Вид_2 Культивар_1
- 4) Род_2 Вид_3

То есть, таблица содержит 4 уникальных строки. Отсюда объем коллекции составляет 4 единицы.

В. Для подсчета представленности отдельных таксономических уровней:

1. Выделить строки таблицы, содержащих не пустые записи для целевого таксономического уровня.
2. Отбросить все столбцы правее целевого таксономического уровня.

Таблица 2. Гипотетический пример таксономических записей

№	Род	Вид	Подвид	Вариация	Подвариация	Форма	Подформа	Культивар	Особь
1	Род 1	Вид 1	Подвид 1						Особь 1
2	Род 1	Вид 1				Форма 1			Особь 2
3	Род 1	Вид 1				Форма 1			Особь 3
4	Род 1	Вид 2						Культ. 1	Особь 4
5	Род 1	Вид 2						Культ. 1	Особь 5
6	Род 2	Вид 3							Особь 6

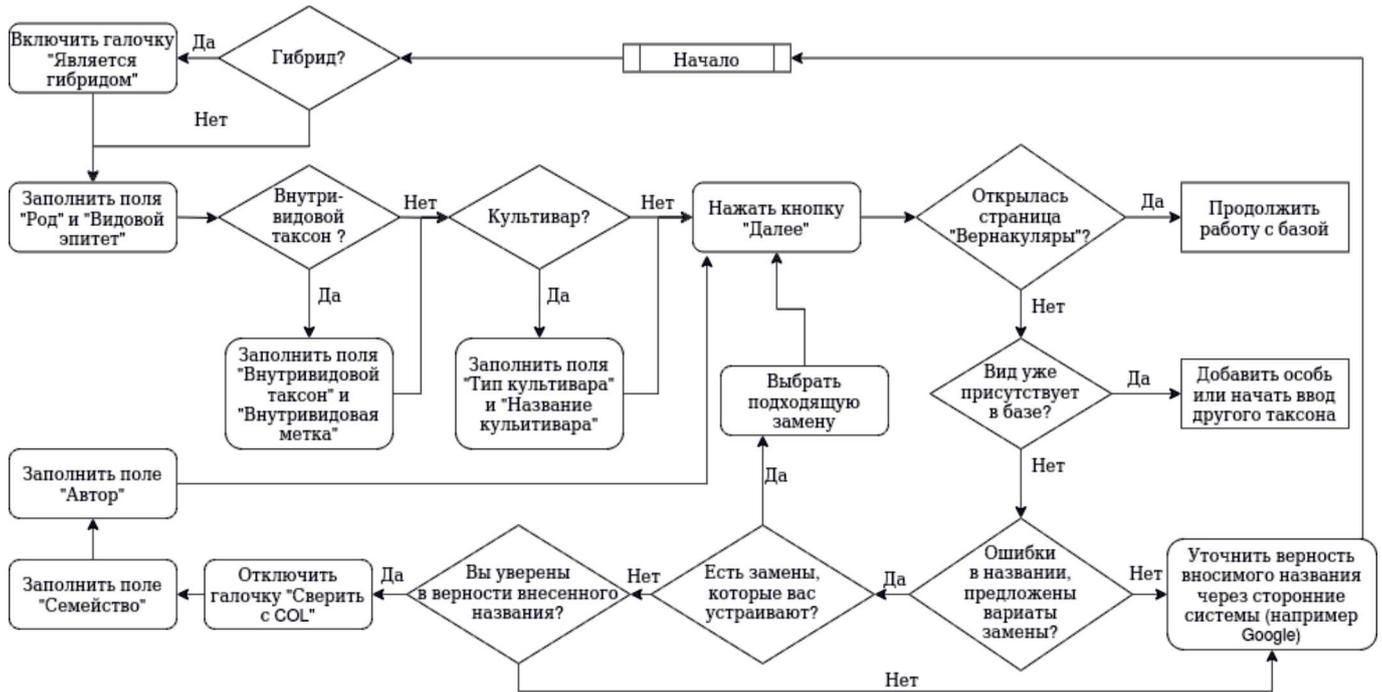


Рис. 3. Граф таксономической иерархии, соответствующий записям в таблице 2
 Fig. 3. Taxonomic hierarchy graph corresponding to the entries in table 2

Например, при оценке представленности уровня «Форма» следует оставить только строки под номером 2 и 3. И отбросить столбцы, следующие после столбца «Форма», то есть «Подформа», «Культивар» и «Особь».

3. В оставшейся отфильтрованной таблице подсчитать количество уникальных рядков.

Например, в таблице 1, если целевой таксономический уровень – «Форма», после выполнение пунктов 1 и 2 остается только одна уникальная строка:

1) Род_1 Вид_1 Форма_1

То есть, в коллекции присутствуют таксоны, относящиеся к единственной записи таксономического уровня «Форма».

Полный текст отчета по размеру коллекции может иметь следующий вид: «Коллекция содержит 4 уникальных названия (таксона), относящиеся к 2-м родам, 3-м видам, 1-му подвиду, 1-й форме и 1-му культивару».

Сравнение с аналогами

Для оценки функционального наполнения созданной Информационной системы имеет смысл сравнить ее с аналогами, представленными на рынке. Наиболее близким по своим целям и функциям коммерческим решением является IrisBG. По заявлениям разработчиков, этой сис-

темой пользуются более 180 ботанических садов, арборетумов и частных коллекций по всему миру. В таблице 3 представлено сравнение функций IrisBG и Информационной системы Донецкого ботанического сада.

Как следует из сравнения, единственное преимущество IrisBG перед разработанной Информационной системой Донецкого ботанического сада – это присутствие функциональности интернет-магазина. Следует отметить, что данная функция не выделялась как необходимая для нашей системы. Отсутствие менеджмента гербарного материала является временным и запланировано для будущей разработки.

Главным преимуществом Информационной системы перед IrisBG является использование таксономической основы. Эта функция была одним из ключевых требований при разработке системы. Также Информационная система Донецкого ботанического сада включает полноценное ГИС-решение для учета пространственного местоположения особей, тогда как альтернатива использует сторонние решения на основе провайдеров по типу Google Maps или наборы статических карт (изображений). Как результат, наша система предоставляет более гибкое представление территории ботанического сада, а также обес-

Таблица 3. Сравнение функций коммерческой IrisBG и Информационной системы Донецкого ботанического сада

Элементы	IrisBG	Информационная система Донецкого ботанического сада
Картографический сервис	Да (ограниченно)	Да
Гербарные данные	Да	Нет (в планах)
Управление магазином	Да	Нет (не планируется)
WEB интерфейс коллекций	Да	Да
WEB интерфейс магазина	Да	Нет (не планируется)
Многоязычная версия	Да	Нет (не планируется)
Использование таксономической основы	Нет	Да
Расширенные отчеты	Да	Да
Пространственный поиск	Нет	Да
Полнотекстовый поиск	Да	Да
Сверка с Красным списком МСОП	Нет	Да
Сверка с базой Tropicos	Нет	Да
Расширяемость	Да (ограниченно)	Да
Неограниченное количество пользователей	Да	Да
Установка облачной базы данных	Да	Да
Хостинг облачной базы данных	Да	Да
Установка WEB интерфейса	Да	Да
Хостинг WEB интерфейса	Да	Да
Стоимость	1 764 276,8 руб.	своими силами
Годовое обслуживание	244 228,4 руб.	своими силами

печивает функции пространственного поиска, который отсутствует в IrisBG.

Учитывая высокую стоимость коммерческого решения (1 764 276 рублей на момент написания статьи) и отсутствие в ней критически важных функций, решение о собственной разработке системы выглядит вполне оправданным.

Выводы

В ходе проведенной работы проанализированы текущие подходы к формированию таксономических списков и прочих наборов метаданных коллекций растений в Донецком ботаническом саду. Установлен ряд проблемных аспектов, в частности дублирование баз данных, накопление

ошибок в научных названиях, отсутствие непосредственного учета особей растений.

Для устранения указанных недостатков решено разработать информационную систему, реализующую функции унификации и полуавтоматической проверки таксономических данных. Оптимальным форматом реализации системы признано веб-приложение, обеспечивающее доступ к централизованному хранилищу с любого устройства, поддерживающего работу в сети интернет. Веб-приложение разработано на основе фреймворка Django с применением сторонних библиотек языка Python.

Обеспечение полуавтоматической проверки таксономических данных реализовано с исполь-

зованием таксономической основы на базе материалов проекта Catalogue of Life.

Задача учета пространственного положения отдельных особей решена путем реализации полноценной ГИС с динамическим генерированием собственной картографической подложки. Это же решение обеспечило возможность пространственного поиска особей.

Конечным результатом работы стала фактическая разработка интернет-приложения информационной системы и ее развертывание на сервере Донецкого ботанического сада. К настоящему моменту система запущена в эксплуатацию.

Сравнение реализованной системы с коммерческими аналогами продемонстрировало ее преимущества как в плане функционального наполнения, так и с позиции финансовых затрат на создание и поддержание.

1. *Artem Pavlenko*, 2017, mapnik [Электронный ресурс]. URL: <https://mapnik.org/> (дата обращения 10.03.2021).
2. *Bennett B.C.*, *Balick M.J.* Does the name really matter? The importance of botanical nomenclature and plant taxonomy in biomedical research // *Journal of Ethnopharmacology*. 2014. Vol. 152, N 3. P. 387–392.
3. *Berendsohn W.G.* A taxonomic information model for botanical databases: the IOPI Model // *Taxon*. 1997. Vol. 46, N 2. P. 283–309.
4. *Celery 5.0.5 documentation*, 2021 [Электронный ресурс]. URL: <https://docs.celeryproject.org/> (дата обращения 10.03.2021).
5. *Dalcin E.C.*, *Estevão da Silva L.A.*, *Cabanilla C.C.* et al. Data Quality Assessment at the Rio de Janeiro Botanical Garden Herbarium Database and Considerations for Data Quality Improvement // 8th International Conference on Ecological Informatics (Brasilia, 3–7 December 2012). 2012. P. 3–7.
6. *Davis F.W.*, *Stoms D.M.*, *Estes J.E.*, *Scepan J.*, *Scott J.M.* An information systems approach to the preservation of biological diversity // *International Journal of Geographical Information Science*. 1990. Vol. 4, N 1. P. 55–78.
7. *Django Software Foundation*, 2021, Django. [Электронный ресурс]. URL: <https://www.djangoproject.com/> (дата обращения 10.03.2021).
8. *Dobesova Z.* Geographic information systems for botanical garden – Steps of design and realization // 12th International Multidisciplinary Scientific GeoConference and EXPO – Modern Management of Mine Producing, Geology and Environmental Protection, SGEM 2012. Vol. 3. P. 377–384.
9. *Gunicorn*, 2021 [Электронный ресурс]. URL: <https://gunicorn.org/> (дата обращения 10.03.2021).
10. *Materialize*, 2021, materialize [Электронный ресурс]. URL: <https://materializecss.com/> (дата обращения 10.03.2021).
11. *Mike Bostock*, 2020, d3js [Электронный ресурс]. URL: <https://d3js.org/> (дата обращения 10.03.2021).
12. *Nginx*, 2021 [Электронный ресурс]. URL: <https://nginx.org/> (дата обращения 10.03.2021).
13. *PostGIS Project Steering Committee*, 2021, PostGIS [Электронный ресурс]. URL: <https://postgis.net/> (дата обращения 10.03.2021).
14. *Redis*, 2021 [Электронный ресурс]. URL: <https://redis.io/> (дата обращения 10.03.2021).
15. *Roskov Y.*, *Ower G.*, *Orrell T.* et. al., eds. (2019). *Species 2000 & ITIS Catalogue of Life, 2019 Annual Checklist*. Species 2000: Naturalis, Leiden, the Netherlands. ISSN 2405-884X. [Электронный ресурс]. URL: www.catalogue-of-life.org/annual-checklist/2019 (дата обращения 10.03.2021).
16. *The PostgreSQL Global Development Group*, 2021, PostgreSQL 12.6 [Электронный ресурс]. URL: <https://www.postgresql.org/docs/12/> (дата обращения 10.03.2021).
17. *The Python Language Reference*, 2021 [Электронный ресурс]. URL: <https://docs.python.org/3.8/reference/> (дата обращения 10.03.2021).
18. *Thiers B.M.*, *Tulig M.C.*, *Watson K.A.* Digitization of The New York Botanical Garden Herbarium // *Brittonia*. 2016. Vol. 68, N 3. P. 324–333.
19. *TileStache*, 2021 [Электронный ресурс]. URL: <http://tilestache.org/> (дата обращения 10.03.2021).
20. *Vladimir Agafonkin*, 2021, leafletjs [Электронный ресурс]. URL: <https://leafletjs.com/> (дата обращения 10.03.2021).
21. *WeasyPrint*, 2021 [Электронный ресурс]. URL: <https://weasyprint.org/> (дата обращения 10.03.2021).
22. *Wickham H.* Tidy Data // *Journal of Statistical Software*. 2014. Vol. 59, N 10. P. 1–23.

Поступила в редакцию: 11.03.2021

UDC 004(62+65)+582

THE INFORMATION SYSTEM OF DONETSK BOTANICAL GARDEN DEVELOPMENT

I.I. Strelnikov, S.A. Prykhodko, A.Z. Glukhov, A.V. Nikolaeva

Public Institution «Donetsk Botanical Garden»

The article is devoted to the aspects of the development of the Donetsk Botanical Garden information system, the main purpose of which is the management of collections taxonomic lists and the living plants registration. At the reconnaissance stage, the current approaches to maintaining collections metadata bases were analyzed. A number of potential limitations and disadvantages such as duplication of data, accumulation of errors, lack of a stable approach to the registration of plants individuals were identified. The proposed model of the information system is designed to eliminate the indicated shortcomings and unify the procedures for handling data for all departments of the botanical garden. The system is implemented as a web application using modern technologies. The task of increasing the reliability and accuracy of taxonomic data was solved using the taxonomic framework. An important feature of the system is the use of GIS technologies to take into account the spatial position of individuals.

Key words: Information system, database, Donetsk Botanical Garden, Python, Catalogue of Life

Citation: Strelnikov I.I., Prykhodko S.A., Glukhov A.Z., Nikolaeva A.V. The Information system of Donetsk Botanical Garden development // *Industrial Botany*. 2021. Vol. 21, N 1. P. 36–52.
