

ВЕДЕНИЕ БАЗ ДАННЫХ БИОЛОГИЧЕСКИХ КОЛЛЕКЦИИ В ИС SPECIFY

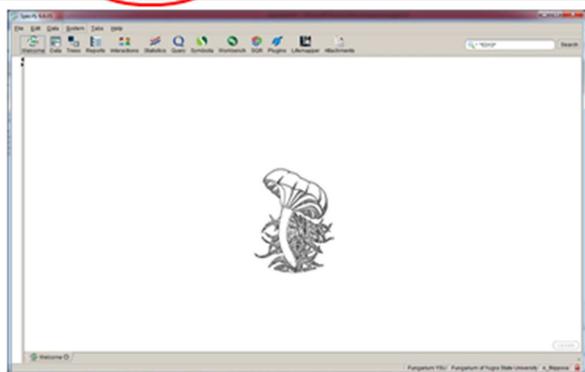
Филиппова Н.В.

Фунгарий Югорского университета

4500 образцов грибов , 3 типовых образца



На подобном семинаре в прошлом году мы с Вами говорили о том, что нам нужно освоить ИС для ведения наших биологических коллекций в международном стандарте, в том числе совместимом с GBIF. Мы решили остановиться на ИС Specify и мне было дано задание опробовать эту систему на примере нашей коллекции – Фунгария ЮГУ. В течение года я обучалась работе с этой системой, настраивала ее для нужд нашей коллекции и импортировала в систему данные (до этого база данных Фунгария велась в субд Access). Затем Дмитрий Карпов установил систему на сервере ЮГУ, так что в настоящее время коллекция может быть доступна для просмотра и редактирования в онлайн-режиме.



В целом ИС удобна и полностью отвечает нуждам ведения коллекции, относительно проста в освоении и мы планируем продолжать ее использование. Мы также можем использовать ее для накопления данных о разных коллекциях в ХМАО для последующего экспорта в GBIF. Например, для отдельных гербариев ООПТ в базе могут быть отдельные коллекции. Затем я могу импортировать в них все данные из таблиц в формате .csv (поля, которые нужно будет там заполнить, определим заранее). По желанию, сотрудник гербария может вести свой раздел самостоятельно, или высылать периодические дополнения в базу данных в форме тех же таблиц. Кроме того, ваша коллекция будет доступна для просмотра и поиска онлайн.

Теперь подробнее о том, что из себя представляет система и как мы с ней работаем на примере Фунгария ЮГУ.



Веб сайт информационной системы:

<http://specifyx.specifysoftware.org>

Адрес базы данных Фунгария:

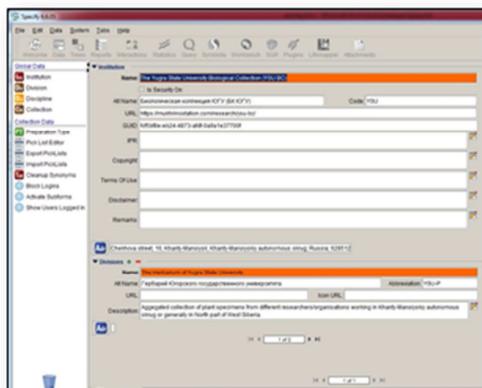
<https://fungariumysu.org/database-structure>



ИС Specify разрабатывается в институте Биоразнообразия Университета Канзас (США) группой программистов и биологов. На развитие программы был выигран грант NSF, однако система является бесплатной и построена на блоках свободного открытого кода. Таким образом достаточно гибка и может быть доработана специалистами для разных задач. Система используется во всем мире, достаточно распространена и надежна в плане ее развития в будущем. Specify основана на стандартах DarwinCore и совместима с другими системами работающими в рамках этого стандарта, в том числе GBIF. Поэтому в случае необходимости перевода данных в другую систему это может быть выполнено.

Система имеет несколько вариантов для установки: базовая Specify 6 установка на локальном компьютере; Specify 7 браузерная версия требующая установки на сервере; возможно использование облачного хранения SpCloud и удаленного менеджмента коллекции через Specify 7; есть также версия для iPad SpecifyInsight – в основном просмотр, поиск и демонстрация коллекции на планшете.

В системе создано 4 уровня: UGRABIO > ЮГУ > Фунгарий ЮГУ & Гербарий ЮГУ

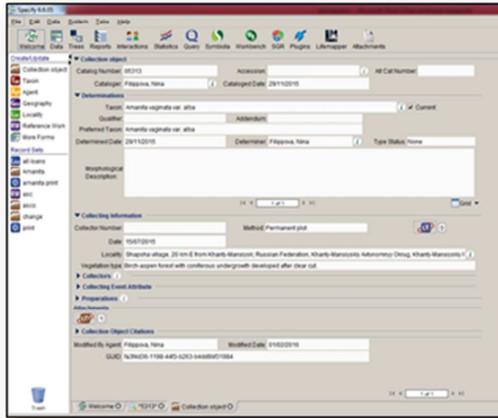


В настоящее время заполнение и менеджмент коллекции Фунгария ЮГУ ведется на локальном компьютере. Параллельно на сервере ЮГУ установлены Specify 6 и Specify 7, которые регулярно обновляются (раз в 2-3 месяца). В будущем мы настроим возможность непосредственной работы с базой на сервере ЮГУ.

Базу мы организовали изначально так, чтобы в ней можно было разместить многие коллекции Югры. Самый верхний уровень назван UGRABIO. Следующий уровень – ЮГУ, внутри которого сейчас два подразделения (Фунгарий и Гербарий) со своими таксономическими деревьями, формами ввода данных и пр. Оба подразделения имеют в настоящее время по одной коллекции. Поэтому, если другие коллекции будут импортировать свои данные в базу, им также будет создан определенный уровень и собственная коллекция.

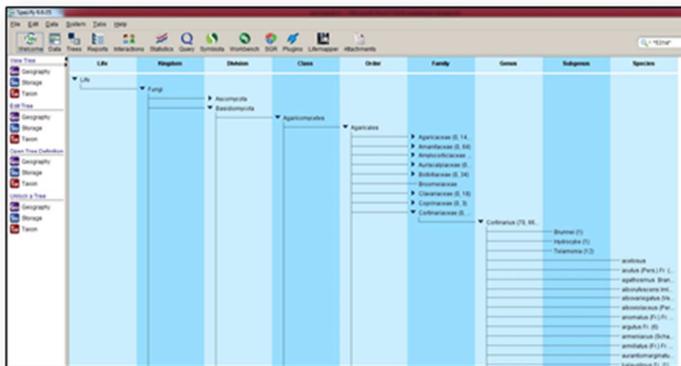
В основе Specify лежит подготовленная система большого числа связанных таблиц, которые в общем соответствуют потребностям баз данных для хранения информации разного рода коллекций. Однако в каждом случае можно и нужно делать настройку необходимых таблиц и полей для каждого конкретного случая, создание формата полей и форм ввода. Это делается с помощью оконного меню, а также редактированием кода xml (для создания форм). Эту работу может выполнять как пользователь, так и менеджер (который настроит один раз эти параметры для коллекции и пользователь в это вникать не будет). В нашей базе данных были использованы стандартные таблицы и поля, наиболее часто применяющиеся в ботанических коллекциях. Выпадающие списки, форма ввода были изменены для нужд коллекции грибов. К сожалению, мы не пользовались опытом других фунгариев на этом этапе и в будущем будем более подробно общаться на эту тему с коллегами и улучшать какие-то моменты.

Основная форма для ввода информации об образце:



Таким образом может выглядеть основная форма ввода информации по образцу: стандартные поля о дате, времени, номере коллекции. Дополнительно открывается раздел о локалитете (связь с географическим деревом); под-форма об определении образца (связь с таксономическим деревом); под-форма об экологических условиях произрастания (субстрат с выпадающим списком и в свободной форме, растение – хозяин); под-форма о единицах хранения, связанных с образцом; и под-форма о связанных иллюстрациях и пр. файлах.

Таксономическое дерево грибов:



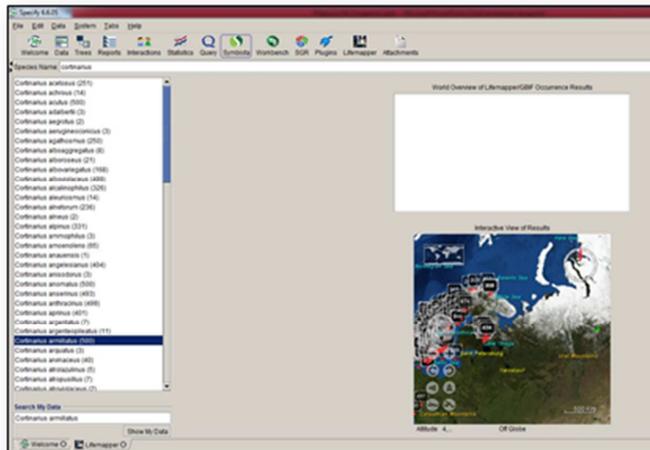
Таксономические и географические деревья являются отдельным блоком в структуре Specify. Структура географического дерева есть по умолчанию вплоть до районов. Разработчики системы рекомендуют пользоваться также таксономическими деревьями, загруженными для каждой дисциплины из проекта Tree of Life. Однако в настоящее время мы пользуемся вручную созданным деревом и планируем доработку этой проблемы в ближайшее время.

Менеджмент коллекции:
заемы, отправка в дар, обмен, печать
ЭТИКЕТОК И ПР.



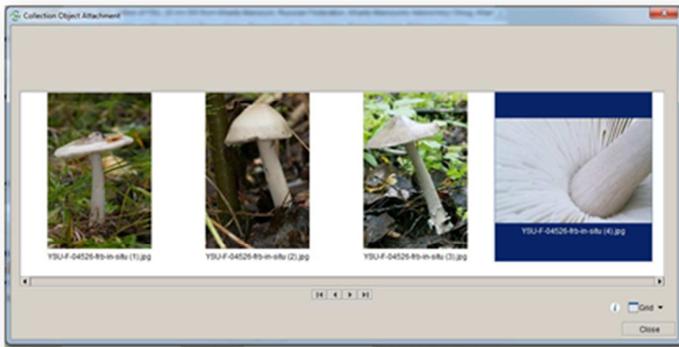
Отдельный блок функционала системы посвящен отслеживанию операций с образцами: такие как отправка в заем, обмен или дар. Здесь возможно создание перечня отправляемых образцов и их последующее отслеживание в базе, так что будет видно их текущее местонахождение. С другой стороны, сохраняется статистика позволяющая подводить итоги подобных операций. Например, в настоящее время в базе числится 18 заемов, из них 8 открытых (т.е. образцы еще не возвращены в коллекцию). Также можно сформировать формат этикеток для разных коллекций, который будет сохранен в системе. Затем этикетки можно распечатать через этот шаблон, используя созданную подборку образцов.

Specify + GIS:
GoogleEarth, WorldWind, LifeMapper



Кратко остановлюсь на картировании точек находок с помощью нескольких GIS модулей. Они позволяют визуализировать не только точки находок в базе (на рисунке помечены зеленым), но точки распространения загружаемые из GBIF (красным). Другие функции GIS модулей я еще не изучала подробно, но это может сделать Илья Филиппов поскольку он в этой области хорошо работает.

Приложения: 4000 макро- и около 3000 микро-фотографий, 70 фотографий местообитаний, около 100 графических рисунков, около 350 таблиц измерений и 20 pdf статей



Связанные с образцом файлы (фотографии, pdf и пр.) также можно хранить в системе. В базе фунгария загружено около 4000 макро- и около 3000 микро-фотографий, 70 фотографий местообитаний, около 100 графических рисунков микроструктур, а также 350 таблиц с измерениями микроструктур и около 20 pdf файлов публикаций. Такое разнообразие файлов было бы очень сложно хранить без использования базы данных и в этом смысле она очень помогает все систематизировать.

Наконец, чтобы импортировать в базу ваши данные нужно сохранить их в форме таблицы формата csv, при этом импортируемый файл проходит через разметку и последовательный ввод в базу в результате чего обеспечивается безошибочный импорт.

Как говорилось выше, есть специальный функционал для экспорта базы в другие системы с помощью разметки по шаблону DarwinCore. Также будет осуществляться экспорт базы в GBIF. Для этого будет установлено ПО (IPT) которое будет периодически обновлять данные наших коллекций, представленные в GBIF.

- Каждый может установить Specify на локальном компьютере и начать им пользоваться
- Мы можем также вести объединенную базу данных для основных коллекций Югры на сервере ЮГУ: **основная цель** такой базы будет
 - размещение ваших данных онлайн
 - публикация данных в GBIF
- **Обозримой целью** (в течение года) может быть создание 4-5 коллекций внутри Specify (по 100 образцов)

На этом часть моего доклада, где я хотела познакомить вас с функциями Specify завершена. Особенности работы в системе мы можем изучать индивидуально. Я бы хотела пригласить вас к совместной работе, которая может быть построена следующим образом.

С одной стороны, каждый может установить Specify на локальном компьютере и начать им пользоваться. Я со своей стороны могу помочь в настройке, но нужно понимать что на обучение уйдет какое-то время потому что система достаточно сложная.

Другой вариант – мы можем вести объединенную базу данных для основных коллекций Югры на сервере ЮГУ. При этом отдельные исследователи могут передавать данные в форме иксель-таблиц, которые будут импортироваться в систему время от времени. Основная цель такой базы будет 1) размещение ваших данных онлайн; 2) публикация данных в GBIF.

Обозримой целью (в течение года) может быть создание 4-5 коллекций внутри Specify куда будут импортированы данные Ваших коллекций. После этого мы можем подключить нашу систему к GBIF и работать в том же режиме в будущем.

Совместная работа над созданием БД коллекций ЮГРЫ:

- Назначить ответственных за передачу данных о коллекциях на платформу Specify
- Подготовить разделы в БД UGRABIO соответствующие коллекциям, собрать метаданные о коллекциях
- Подготовить шаблоны для импорта в Specify (определить поля и формат данных) для каждой коллекции
- Импортировать данные в каждую коллекцию (для начала – не более 100 образцов от каждой коллекции)
- Параллельно разрабатывать политику сбора данных и их передачи в GBIF
- Подключить базу к IPT и сделать первый экспорт в GBIF

Можно нарисовать такой план:

1. Назначить ответственных за передачу данных о коллекциях на платформу Specify (Звягина Е.А. – грибы, Васина А.Л. – сосудистые растения, Лапшина Е.Д. – мхи, Бабушкин Е. – моллюски, Акопян Э. – насекомые и т.д.)
2. Подготовить разделы в БД UGRABIO соответствующие коллекциям, собрать метаданные о коллекциях в отдельной таблице (Филиппова Н.В. совместно с ответственными).
3. Подготовить шаблоны для импорта в Specify (определить поля и формат данных) для каждой коллекции (Филиппова Н.В. совместно с ответственными).
4. Импортировать данные в каждую коллекцию (для первого шага – не более 100 образцов от каждой коллекции) (каждый из ответственных).
5. Параллельно разрабатывать политику сбора данных и их передачи в GBIF.
6. Подключить базу к IPT и сделать первый экспорт.