

DOI 10.54596/2958-0048-2024-3-58-66

УДК 727.65

МРНТИ 34.29.35

ОЦИФРОВКА ГЕРБАРНОЙ КОЛЛЕКЦИИ СЕВЕРО-КАЗАХСТАНСКОГО УНИВЕРСИТЕТА ИМЕНИ МАНАША КОЗЫБАЕВА КАК РЕЗУЛЬТАТ ИНТЕГРАЦИИ ЦИФРОВЫХ ТЕХНОЛОГИЙ В ГЕРБАРНОЕ ДЕЛО

Романчук В.В.^{1*}, Абильтаева А.А.²

^{1*} НАО «Северо-Казахстанский университет имени Манаша Козыбаева»
Петропавловск, Казахстан

² ТОО «Академия физической культуры и массового спорта»
Астана, Казахстан

* Автор для корреспонденции: romanchuk221193@mail.ru

Аннотация

В наши дни все сферы человеческих знаний прочно взаимодействуют с цифровыми технологиями. Данные технологии открывают новые подходы к изучению и методы исследования в большинстве научных областей.

В статье показана интеграция цифровых технологий в гербарное дело. Описана актуальность оцифровки гербарных коллекций. Представлен алгоритм создания цифрового гербария и предложен способ публикации данных. Отмечены ограничивающие факторы и преимущества работы с электронным гербарием.

Ключевые слова: цифровые технологии, оцифровка, электронный гербарий, акроним, штрихкодирование, набор данных, Index Herbariorum, GBIF.

МАНАШ ҚОЗЫБАЕВ АТЫНДАҒЫ СОЛТҮСТІК ҚАЗАҚСТАН УНИВЕРСИТЕТІНІҢ ГЕРБАРИЯЛАР ЖИНАҒЫНЫҢ ЦИФРЛІК ТЕХНОЛОГИЯЛАРДЫ ШӨПШІЛІК КҮТІМІНЕ ИНТЕГРАЦИЯЛАУ НӘТИЖЕСІНДЕГІ

Романчук В.В.^{1*}, Әбілтаева А.А.²

^{1*} «Манаш Қозыбаев атындағы Солтүстік Қазақстан университеті» КеАҚ
Петропавл, Қазақстан

² «Дене шынықтыру және бұқаралық спорт академиясы» ЖШС
Астана, Қазақстан

*Хат-хабар үшін автор: romanchuk221193@mail.ru

Аңдатпа

Қазіргі уақытта адам білімінің барлық салалары цифрлық технологиялармен тығыз байланыста. Бұл технологиялар көптеген ғылыми салаларда зерттеу мен зерттеу әдістерінің жаңа тәсілдерін ашады.

Мақалада гербарий жұмысына цифрлық технологияларды біріктіру көрсетілген. Гербарий жинақтарын цифрландырудың өзектілігі сипатталған. Цифрлық гербарий құру алгоритмі ұсынылған және деректерді жариялау әдісі ұсынылған. Электрондық гербариймен жұмыс істеудің шектеуші факторлары мен артықшылықтары атап өтілді.

Кілт сөздер: цифрлық технологиялар, цифрландыру, электронды гербарий, аббревиатура, штрих-кодтау, деректер жинағы, Index Herbariorum, GBIF.

**DIGITIZATION OF HERBARIUM COLLECTION OF MANASH KOZYBAEV
NORTH-KAZAKHSTAN UNIVERSITY AS A RESULT OF DIGITAL TECHNOLOGY
INTEGRATION IN HERBAROLOGY**

Romanchuk V.V.^{1*}, Abiltaeva A.A.²

^{1*}*Manash Kozybayev North Kazakhstan University NPLC, Petropavlovsk, Kazakhstan*

²*Academy of Physical Culture and Mass Sports LLP, Astana, Kazakhstan*

**Corresponding author: romanchuk221193@mail.ru*

Abstract

Nowadays, all areas of human knowledge are firmly interacting with digital technologies. These technologies open up new approaches to study and research methods in most scientific fields.

The article shows the integration of digital technologies in herbarium science. The relevance of digitization of herbarium collections is described. An algorithm for creating a digital herbarium is presented and a method for publishing the data is proposed. The limiting factors and advantages of working with electronic herbarium are noted.

Keywords: digital technologies, digitization, electronic herbarium, acronym, barcoding, data set, Index Herbariorum, GBIF.

Введение

В Северо-Казахстанском университете имени Манаша Козыбаева (далее СКУ) хранится большая гербарная коллекция, насчитывающая порядка 8 000 экземпляров засушенных растений. Особая ценность коллекции заключается в том, что в ней имеется значительное количество исторического гербария, датированного 1960-1970 годами.

Гербарный фонд университета начал формироваться в 1960 году по инициативе Мельницкой Альбины Константиновны, преподавателя систематики растений [1]. С тех пор фонд ежегодно пополнялся сотнями новых гербарных образцов. Имеющаяся коллекция — это результат кропотливого и методичного труда не одного поколения студентов и преподавателей. На основании собранного материала было написано множество научных работ: статей, дипломных работ, диссертаций, методических пособий. Данный гербарий используется как средство обучения в ходе преподавания цикла ботанических дисциплин так как является единственным в своём роде натуральным наглядным объектом, способным доподлинно передать морфологическое строение растений.

В наши дни цифровые технологии находятся на пике своего развития, повсеместно наблюдается их интеграция в другие отрасли науки. Обладая огромным потенциалом, цифровые технологии открывают новые возможности в изучении различных сфер человеческих знаний.

Феномен цифровых технологий предусматривает использование новых подходов в работе с гербарием. В частности, появление высокотехнологичного оборудования позволило проводить молекулярно-генетический анализ гербария, что в свою очередь поставляет информацию для работы систематиков растений. Как известно, классификация и филогения растений была существенно пересмотрена за последние несколько десятилетий главным образом за счёт молекулярных исследований.

Другим актуальным направлением в работе с гербарием стала его оцифровка, данный метод является не только прогрессивным, но и необходимым с практической точки зрения. Суть метода заключается в переводе этикеточных данных и изображений гербарных образцов в цифровую форму.

Гербарий является хрупким объектом и требует особых условий хранения. Кроме того, зачастую при ботанических исследованиях возникает сложность с доступом к интересующим образцам, особенно при международных исследованиях, когда отсутствует физическая возможность посетить учреждение и изучить гербарные образцы. Актуальность метода оцифровки гербария заключается в том, что посредством цифровых копий появилась возможность сохранить коллекционные фонды гербария для потомков и открыть к ним общий доступ.

Материалы и методы исследования

Метод систематического литературного обзора положен в основу теоретической части исследования, с помощью данного метода было изучено состояние вопроса создания электронного гербария в мире и в Казахстане. Посредством метода был составлен алгоритм работы [2].

При помощи методов инвентаризации, отбора и сравнительно-морфологического анализа был определён объём работы, составлен таксономический список представителей видов растений подкласса Ламииды (*Lamiidae*), гербарии которых будут подлежать оцифровке в рамках исследования [3].

Методы штрихкодирования, стандартизации, верификации, «точка-радиус» могут быть объединены по общим понятием – методы оцифровки. Метод штрихкодирования выражался посредством маркировки гербарных листов разработанными штриховыми кодами Code-128. Метод стандартизации заключался в работе с набором данных Data с использованием стандарта Darwin Core (DwC). Метод верификации обеспечил проверку набора данных для устранения ошибок перед их публикацией на платформе org. Метод «точка-радиус» был использован для осуществления геопривязки гербарных образцов к месту их сбора [4].

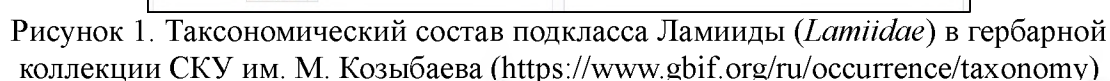
Результаты исследования

В Северо-Казахстанском университете им. М. Козыбаева работы по оцифровке гербария были начаты в 2024 году. Первым был оцифрован гербарий растений подкласса Ламииды (*Lamiidae*). В общей сложности в цифровой формат переведены этикеточные данные 400 гербарных образцов.

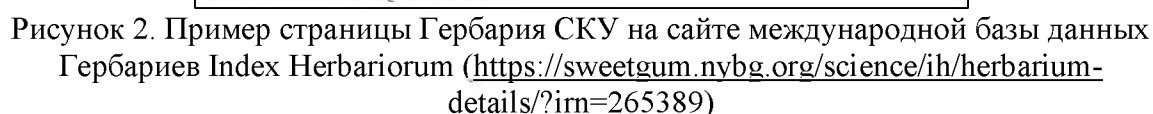
Оцифровка гербария осуществлялась по следующему алгоритму:

1. подготовка гербария, который будет подлежать оцифровке;
2. получение Гербарием СҚУ международного акронима. Регистрация Гербария в азе данных Гербариев – Index Herbariorum (IH);
3. регистрация Гербария СҚУ в Глобальном фонде сохранения биоразнообразия –
4. штрихкодирование гербарных образцов;
5. подготовка набора данных Occurrence Data для публикации в GBIF.org;
6. верификация набора данных посредством GBIF Data Validator;
7. авторизация в программном обеспечении Integrated Publishing Toolkit (IPT);
8. публикация набора данных на сайте GBIF.org.

В гербарном фонде СҚУ им. М. Козыбаева подкласс Ламииды (*Lamiidae*) представлен 3 порядками, 8 семействами, 33 родами и 47 видами (Рис. 1).



Следующим этапом было получение Гербарием SKU международного индекса – акронима, присваиваемого международной базой данных Гербариев Index Herbariorum (IH), основанной в 1935 году. На сегодняшний день в IH зарегистрировано более 3000 действующих Гербариев со всего мира. По результатам рассмотрения заявки Гербарию SKU был присвоен акроним – SQU и создана страница на сайте NYBG STEERE HERBARIUM (<https://sweetgum.nybg.org/science/ih/>) [5] (Рис. 2).



В качестве способа предоставления оцифрованных данных была выбрана международная глобальная платформа для публикации данных о биоразнообразии – GBIF (Global Biodiversity Information Facility). Данная платформа даёт возможность учреждениям и исследователям по всему миру делиться данными о биоразнообразии Земли. GBIF это не просто интернет-ресурс, это глобальный информационный фонд биоразнообразия, созданный в 2001 году. Секретариат фонда располагается в Копенгагене, работа осуществляется посредством региональных узлов. Для облегчения публикации данных GBIF имеет собственное программное обеспечение Integrated Publishing Toolkit (IPT), стандарт данных Darwin Core (DwC), предназначенный для унификации публикуемых данных, и четыре типа наборов данных (Data sets) [6].

На этапе штрихкодирования гербарных образцов предназначенный для оцифровки гербарий был снабжён штриховыми кодами, содержащими буквенную (SQU) и цифровую (00001-00400) последовательность в формате Code-128 (Рис. 3).



Рисунок 3. Гербарный образец, снабжённый штриховым кодом

Для публикации данных гербарных коллекций используется набор данных GBIF – Data. Этот тип данных составляет основу всех данных публикуемых в GBIF. Помимо метаданных набор содержит информацию о дате и месте сбора гербарных образцов, что в свою очередь позволяет осуществлять геопривязку находок, тем самым создавая картирование ареалов произрастания видов.

Набор данных создаётся в таблице Microsoft Excel, столбцы именуются в соответствии со стандартом Darwin Core. В данной работе таблица с набором данных Data содержит 37 столбцов (Рис. 4). На данном этапе также была осуществлена геопривязка гербарных образцов к местам их сбора.

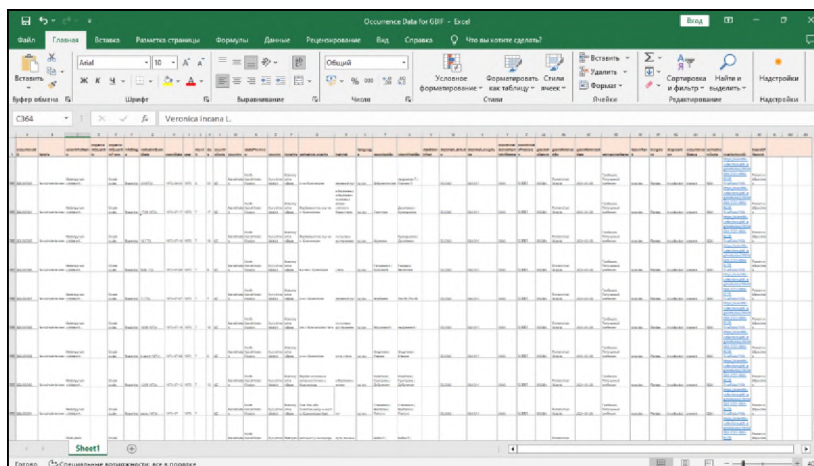
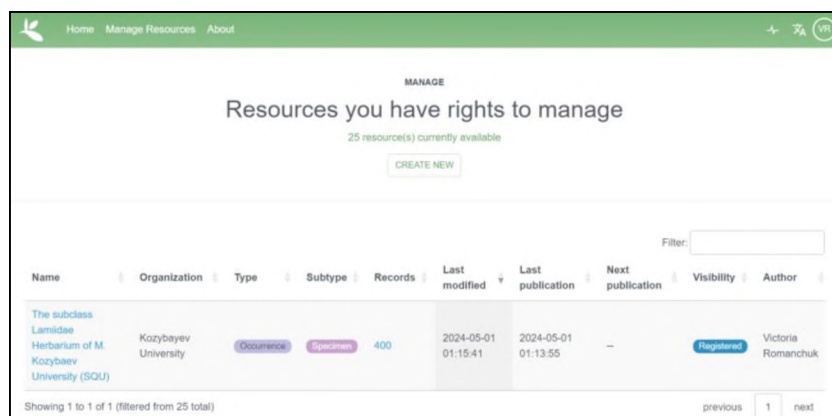


Рисунок 4. Фрагмент таблицы Excel содержащая набор данных Occurrence Data

После подготовки набора данных был использован валидатор GBIF Data Validator (<https://www.gbif.org/tools/data-validator>), при помощи которого была осуществлена верификация данных на предмет дублирующихся записей, неполных полей и несоответствий в форматировании [7]. Выявленные ошибки были устранены.

Авторизация в программном обеспечении Integrated Publishing Toolkit (IPT) была осуществлена при поддержке Карагандинского исследовательского университета имени академика Е.А. Букетова. По результатам регистрации организации «Manash Kozybayev North Kazakhstan University NPLC» получила возможность публиковать данные на платформе GBIF.org (<https://www.gbif.org/>) [8] (Рис. 5).



Name	Organization	Type	Subtype	Records	Last modified	Last publication	Next publication	Visibility	Author
The subclass Lamidae	Kozybayev University	Occurrence	Specimen	400	2024-05-01 01:15:41	2024-05-01 01:13:55	—	Registered	Victoria Romanchuk

Рисунок 5. Страница СКУ им. М. Козыбаева в программном обеспечении IPT на сервере Карагандинского университета, скрин с экрана (<https://www.gbif.org/ipt>)

На завершающем этапе этикеточные данные 400 гербарных образцов были опубликованы на сайте GBIF.org (<https://www.gbif.org/>). Благодаря геопривязке, на платформе сформирована карта мест сбора гербарных образцов, отражающая места их произрастания (Рис. 6).

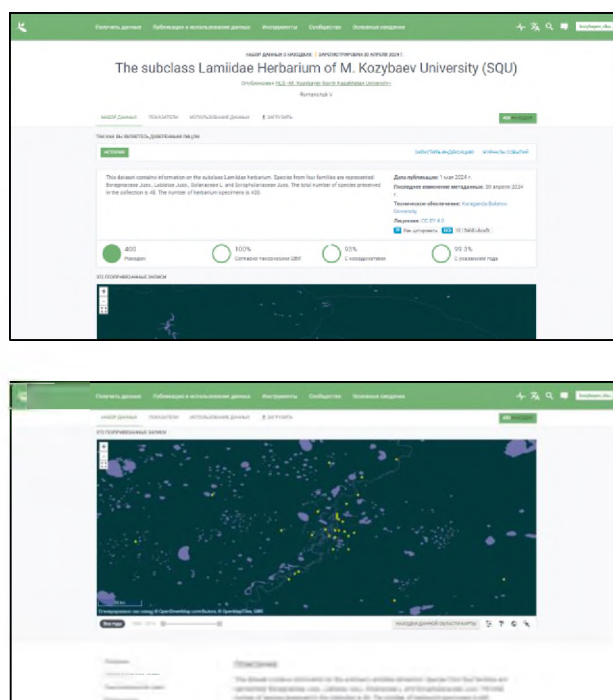


Рисунок 6. Набор данных о находках с геопривязанными записями оцифрованного гербария СКУ на сайте GBIF.org, скрин с экрана (<https://www.gbif.org/>)

Обсуждение и заключение

Планируется осуществить оцифровку всего гербария, хранящегося в коллекционном фонде Северо-Казахстанского университета им. М. Козыбаева. Ограничением исследования выступают следующие факторы: отсутствие штата сотрудников Гербария, отсутствие оборудования и финансирования.

В университете отсутствуют сотрудники Гербария, работой с гербарным фондом занимаются отдельные преподаватели и обучающиеся. В Гербарии отсутствует оборудование необходимое для осуществления полноценной работы по созданию электронного гербария, в частности специальные сканеры либо мощные цифровые камеры с сопутствующим оборудованием для создания высокоточных сканов или фотографий гербарных образцов.

Для успешной работы по оцифровке гербарной коллекции СКУ необходима коллаборация с IT специалистами, следует создать собственный сервер для работы с гербарием и установить программное обеспечение Integrated Publishing Toolkit (IPT) для самостоятельной публикации данных на платформе GBIF.org. При содействии программистов станет возможным завершение работы по сканированию штриховых кодов гербариев.

Платформа GBIF.org обладает множеством преимуществ для работы с данными. Во-первых, данная платформа является международной, общепринятой и хорошо себя зарекомендовавшей. GBIF успешно работает более 20 лет и предоставляет возможность тысячам учреждений делиться имеющимися данными и миллионам пользователей получать к ним доступ.

В контексте гербария загруженные данные представлены на сайте исчерпывающе. Помимо полной информации о публикующей организации пользователь имеет возможность ознакомиться с полным списком оцифрованного гербария. Каждый образец

снабжён детальным описанием находки, содержащим информацию о видовой принадлежности образца, месте и времени сбора, сборщиках, вплоть до даты определения вида и его «народных» названиях.

На основе загруженных данных на сайте формируется иерархическая таксономия их набора, карта с геопривязанными записями о находках, графические показатели и галерея изображений находок. Основная информация доступна уже при оцифровке этикеточных данных, галерея изображений доступна в случае сканирования или фотографирования гербариев и включения этой информации в набор данных Occurrence Data. На сайте доступны множество фильтров по всем имеющимся в наборе данных показателям. Вся информация доступна для скачивания.

Оцифровка гербария является крайне актуальным направлением. Цифровой гербарий позволяет сохранить ботанические коллекции от ветшания в ходе использования и обезопасить потерю данных в случае пожаров и других бедствий. Публикация оцифрованных данных открывает к гербарным коллекциям всеобщий доступ, позволяя исследователям со всего мира использовать эти данные в своих исследованиях, способствуя тем самым изучению и сохранению биоразнообразия Земли.

Литература:

1. Галиев Ж.М., Калкаманова А.Б. Гербарный фонд кафедры общей биологии Северо-Казахстанского государственного университета имени М. Козыбаева: учебно-методическое пособие студентам по ботаническим дисциплинам. – Петропавловск: СҚУ им. М. Козыбаева, 2006.
2. Романчук В.В., Тлеубергенова Г.С. Таксономический состав подкласса Lamiidae в гербарной коллекции кафедры «Биология» СҚУ им. М. Козыбаева. Вестник Северо-Казахстанского Университета им. М. Козыбаева. 2023; (4 (60)): 24-34. <https://doi.org/10.54596/2958-0048-2023-4-24-34>
3. Романчук В.В., Мадиева А.Н., Тлеубергенова Г.С. Состояние вопроса создания электронного гербария в Казахстане. Вестник Северо-Казахстанского Университета им. М. Козыбаева. 2024; (1 (61)): 17-27. <https://doi.org/10.54596/2958-0048-2024-1-17-27>
4. Романчук В.В. Оцифровка и индексация гербариев растений подкласса Lamiidae гербарного фонда кафедры «Биология» СҚУ им. М. Козыбаева. Диссертация на соискание степени магистра естественных наук. – Петропавловск: СҚУ им. М. Козыбаева, 2024.
5. Index Herbariorum - The William & Lynda Steere Herbarium. URL: <https://sweetgum.nybg.org/science/ih/> (дата обращения: 07.09.2024).
6. GBIF. URL: <https://www.gbif.org/ru/> (дата обращения: 07.09.2024).
7. GBIF Data Validator. URL: <https://www.gbif.org/tools/data-validator> (дата обращения: 08.09.2024).
8. IPT. URL: <https://www.gbif.org/ru/ipt> (дата обращения: 08.09.2024).

References:

1. Galiev Zh.M., Kalkamanova A.B. Gerbarnyj fond kafedry obshchej biologii Severo-Kazahstanskogo gosudarstvennogo universiteta imeni M. Kozybaeva: uchebno-metodicheskoe posobie studentam po botanicheskim disciplinam. – Petropavlovsk: SKGU im. M. Kozybaeva, 2006.
2. Romanchuk V.V., Tleubergenova G.S. Taksonomicheskij sostav podklassa Lamiidae v gerbarnoj kollekcii kafedry «Biologiya» SKU im. M. Kozybaeva. Vestnik Severo-Kazahstanskogo Universiteta im. M. Kozybaeva. 2023; (4 (60)): 24-34. <https://doi.org/10.54596/2958-0048-2023-4-24-34>
3. Romanchuk V.V., Madieva A.N., Tleubergenova G.S. Sostoyanie voprosa sozdaniya elektronnoho gerbariya v Kazahstane. Vestnik Severo-Kazahstanskogo Universiteta im. M. Kozybaeva. 2024; (1 (61)): 17-27. <https://doi.org/10.54596/2958-0048-2024-1-17-27>
4. Romanchuk V.V. Ocifrovka i indeksaciya gerbariev rastenij podklassa Lamiidae gerbarnogo fonda kafedry «Biologiya» SKU im. M. Kozybaeva. Dissertaciya na soiskanie stepeni magistra estestvennyh nauk. – Petropavlovsk: SKU im. M. Kozybaeva, 2024.
5. Index Herbariorum - The William & Lynda Steere Herbarium. URL: <https://sweetgum.nybg.org/science/ih/> (data obrashcheniya: 07.09.2024).

6. GBIF. URL: <https://www.gbif.org/ru/> (data obrashcheniya: 07.09.2024).
7. GBIF Data Validator. URL: <https://www.gbif.org/tools/data-validator> (data obrashcheniya: 08.09.2024).
8. IPT. URL: <https://www.gbif.org/ru/ipt> (data obrashcheniya: 08.09.2024).

Information about the authors:

Romanchuk V.V. – corresponding author, Master of Science, researcher, Kozybayev University, Petropavlovsk, Kazakhstan; e-mail: romanchuk221193@mail.ru;

Abiltaeva A.A. – Master of Science, Senior lecturer of the Department of Sports Education and Coaching, Academy of Physical Culture and Mass Sports, Astana, Kazakhstan; e-mail: A_abiltayeva@apems.edu.kz.