

Научная статья
УДК 551.468.1
EDN NVWXVU

Программный продукт «ФотоБерега Крыма»

М. П. Вецало *, Е. А. Годин, Е. А. Исаева, Л. К. Галковская

Морской гидрофизический институт РАН, Севастополь, Россия

* e-mail: mvetsalo@mhi-ras.ru

Аннотация

Описывается информационно-поисковая система «ФотоБерега Крыма», разработанная сотрудниками группы «Банк океанографических данных» на основе концепции программного продукта «ФотоБерега» для систематизации и каталогизации коллекции цифровых изображений берегов Крымского полуострова, а также обеспечения эффективной работы с этой коллекцией при проведении научных исследований. Программная система написана на языке программирования *Python*. Интерфейс приложения разработан с использованием пакета *tkinter*. Центральной частью системы является каталог метаданных объектов съемки, который построен на основе фасетной классификации и включает описательные фасеты «Дата и время», «Вид съемки» и специализированные фасеты «Географический регион», «Генетический тип побережья». Для формирования поисковой выдачи в программной системе применен метод расширенного булева поиска. Загрузка новых изображений и редактирование метаданных существующих элементов каталога выполняется в редакторе метаданных. Работа с геоинформационной частью базы метаданных осуществляется в редакторе геоданных. Созданный программный продукт имеет значительный потенциал для дальнейшего развития и после соответствующей настройки может быть использован для работы с изображениями берегов других регионов, а также для систематизации и классификации коллекций изображений в самых различных областях.

Ключевые слова: берега, изображения, визуализация, систематизация, каталогизация, фасетная классификация, информационно-поисковая система

Благодарности: работа выполнена в рамках государственного задания ФГБУН ФИЦ МГИ по теме № 0827-2020-0004 «Прибрежные исследования». Авторы выражают благодарность инициатору разработки данного программного продукта д. г. н. Ю. Н. Горячкину, а также к. х. н. В. В. Долотову и к. ф.-м. н. А. В. Багаеву за плодотворное обсуждение полученных результатов.

Для цитирования: Программный продукт «ФотоБерега Крыма» / М. П. Вецало [и др.] // Экологическая безопасность прибрежной и шельфовой зон моря. 2024. № 4. С. 131–140. EDN NVWXVU.

© Вецало М. П., Годин Е. А., Исаева Е. А., Галковская Л. К., 2024



Контент доступен по лицензии Creative Commons Attribution-Non Commercial 4.0 International (CC BY-NC 4.0)
This work is licensed under a Creative Commons Attribution-Non Commercial 4.0 International (CC BY-NC 4.0) License

Software PhotoCoasts of Crimea

M. P. Vetsalo *, E. A. Godin, E. A. Isaeva, L. K. Galkovskaya

Marine Hydrophysical Institute of RAS, Sevastopol, Russia

* e-mail: mvetsalo@mhi-ras.ru

Abstract

The article describes the information retrieval system PhotoCoasts of Crimea developed by the staff of the Oceanographic Data Bank group based on the concept of the software PhotoCoasts to systematize and catalogue the collection of digital images of the Crimean Peninsula coasts. The system also ensures effective work with this collection while conducting scientific research. The software system is written in the Python programming language. The application interface is developed using the tkinter package. The system core is a catalogue of meta-information on photosurvey objects. The catalogue is based on faceted classification and includes descriptive facets “Date and Time”, “Type of Photosurvey” and specialised facets “Geographic Region”, “Coast Genetic Type”. The method of extended Boolean retrieval was applied to form the query results in the software system. New images are uploaded and metadata of existing catalogue elements are edited in the metadata editor. Work with the geoinformation part of the metadata base is performed in the geodata editor. The developed software has a significant potential for further evolution and after appropriate adjustment can be used for work with coast images of other regions. It also allows systematisation and classification of image collections in various fields.

Keywords: coasts, images, visualisation, systematisation, cataloguing, faceted classification, information retrieval system

Acknowledgments: The work was performed under state assignment of MHI RAS no. 0827-2020-0004 “Coastal studies”. The authors are grateful to DrSci (Geogr.) Yu. N. Goryachkin, who initiated the development of the software, as well as to PhD (Chem.) V. V. Dolotov and PhD (Phys.-Math.) A. V. Bagaev for fruitful discussion of the results.

For citation: Vetsalo, M.P., Godin, E.A., Isaeva, E.A. and Galkovskaya, L.K., 2024. Software PhotoCoasts of Crimea. *Ecological Safety of Coastal and Shelf Zones of Sea*, (4), pp. 131–140.

Введение

Изображения являются наиболее длительно существующим способом фиксации состояния природных объектов, пройдя путь от рисунка до фото- и киносъемки с высокой степенью разрешения. При условии определения места и времени получения изображений с достаточной точностью их сопоставление позволяет проанализировать происходящие изменения состояния тех или иных объектов под воздействием природных и антропогенных факторов. Вместе с тем значительная часть изображений до настоящего времени не оцифрована, хранится в разрозненных наборах, должным образом не систематизирована и не каталогизирована. Такое положение дел может привести к утрате многих материалов и не обеспечивает их доступность для широкого круга заинтересованных пользователей. Сказанное полностью справедливо и для изображений морских берегов, в том числе Крымского побережья. Сложившаяся ситуация может быть исправлена путем организации работ

по переводу изображений в цифровую форму и создания специального программного обеспечения (ПО) для работы с ними.

В настоящее время существует значительное количество программных продуктов для создания каталогов цифровых изображений, но, как правило, они недостаточно эффективны для решения задач научных исследований. Анализ информации из открытых источников^{1), 2), 3)} и личный опыт использования показали, что в большинстве существующих программных продуктов каталог строится на основе иерархии файловой системы компьютера пользователя (*Adobe Bridge*, *ACDSee Photo Studio* и др.) или собственной иерархической структуры (папки в *Adobe Lightroom* и *Corel AfterShot Pro*, коллекции в *darktable*, альбомы *digiKam* и др.) со всеми присущими таким классификациям недостатками. В частности, жесткая структура иерархии затрудняет включение новых уровней деления. Классификация становится чрезвычайно громоздкой и сложной в использовании при большом количестве уровней и недостаточно информативной при их малом количестве. Такие продукты также ориентированы в первую очередь на работу с метаданными изображения (параметры съемки, геолокация и т. п.), но не обладают достаточно развитыми возможностями определения характеристик объекта съемки, необходимыми для реализации исследовательских задач. Следует отметить, что значительная часть существующего программного обеспечения является коммерческой и имеет закрытый исходный код, не позволяющий модифицировать продукт в соответствии с требованиями пользователя. Дополнительным препятствием для приобретения и использования существующего коммерческого ПО каталогизации цифровых изображений в данный момент является санкционная политика ряда государств в отношении Российской Федерации.

Принимая во внимание описанные выше недостатки существующего программного обеспечения, сотрудники отдела гидрофизики шельфа и группы «Банк океанографических данных» ФГБУН ФИЦ МГИ разработали концепцию программного продукта «ФотоБерега», включающую общие подходы к созданию ПО для визуализации, систематизации и каталогизации цифровых изображений при проведении научных исследований:

- обеспечение максимальной независимости от внешних факторов, использование только свободно распространяемых компонентов с открытым исходным кодом;
- работа с каталогом (в том числе его расширяемость) с учетом специфики применения изображений при решении научных задач;
- поддержка базы геoinформации;
- поддержка работы с метаданными цифровых изображений (время съемки, геопозиционирование);
- массовый импорт изображений.

¹⁾ Сажко Д. 10 приложений для организации фотоколлекции // Лайфхакер : сайт. 24.12.2018. URL: <https://lifehacker.ru/kak-organizovat-kollekciyu-fotografij/> (дата обращения 25.11.2021).

²⁾ Обзор Windows-приложений для наведения порядка в фотоархивах // Хабр : сайт. 12.06.2014. URL: <https://habr.com/ru/post/226123/> (дата обращения 25.11.2021).

³⁾ Шляхтина С. Каталогизация и хранение цифровых снимков // КомпьютерПресс : сайт. Ноябрь 2004. URL: <https://compress.ru/article.aspx?id=12397> (дата обращения: 25.11.2021).

В статье описаны результаты работ по реализации этой концепции при разработке специализированной информационно-поисковой системы «ФотоБерега Крыма» [1, 2], оптимизированной для решения практической задачи – систематизации и классификации коллекции цифровых изображений берегов Крымского п-ова, собранных в отделе гидрофизики шельфа МГИ за более чем полуторавековой период, и эффективной работы с ними.

Подходы и методы

Центральной частью системы является каталог метаинформации объектов съемки. Каталог строится на основе фасетной классификации, которая обладает большей семантической силой, чем иерархическая [3, 4]. На первом этапе формирования каталога определяется набор понятий (терминов), необходимых для описания элемента каталога (изображения). Затем термины семантически либо по иному принципу группируются в классы (фасеты). Классификация элементов каталога, в свою очередь, не задается заранее, а строится путем выборки элементов из фасетов и формирования из них линейной цепочки, называемой фасетной формулой. Место каждого фасета в фасетной формуле является строго фиксированным. При этом для решения задачи информационного поиска по каталогу достаточно определить порядок преобразования информационных потребностей пользователя в фасетную формулу (поисковый запрос) и рассматривать полученную на основе этой формулы классификацию как поисковую выдачу.

Для формирования поисковой выдачи в программной системе применен метод расширенного булева поиска [5], при котором результат определяется логическим выражением, формируемым на основе запроса пользователя. Поисковый запрос (фасетная формула) преобразуется в логическое выражение, которое применяется к метаданным каждого элемента каталога. С целью повышения скорости формирования поисковой выдачи применяется обратное индексирование базы метаданных изображений [5, 6] по отдельным фасетам. Ускорение работы приложения при использовании индекса достигается путем значительного (в общем случае) уменьшения количества участвующих в операциях элементов каталога и перехода к использованию операций над множествами вместо выполнения логических операций (булева поиска) для всех изображений, присутствующих в базе программы.

Реализация

Для реализации программной системы был выбран язык программирования *Python*, имеющий как развитую стандартную библиотеку, так и достаточное количество свободно распространяемых сторонних библиотек с открытым исходным кодом.

Каталог системы «ФотоБерега Крыма» включает в себя описательные фасеты «Дата и время» (*D*), «Вид съемки» (*T*) и специализированные – «Географический регион» (*R*), «Генетический тип побережья» (*G*). Фасет «Дата и время», в свою очередь, разбит на три субфасета: «Дата начала периода» (*D_B*), «Дата окончания периода» (*D_E*) и «Сезонность съемки» (*S*). Результирующая фасетная формула для применяемой классификации имеет вид

$$\langle D_B : D_E : S : R : T : G \rangle.$$

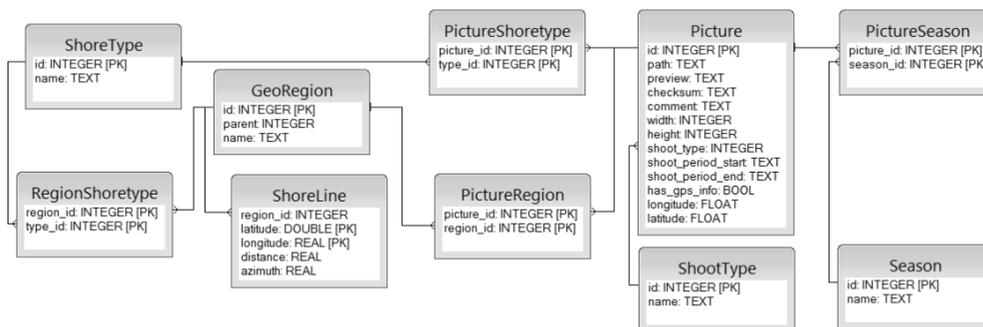
Часть фасетов содержит заданное при разработке приложения конечное множество понятий, определяемое семантикой соответствующего признака или спецификой разрабатываемой системы. Так, в качестве описаний для сезонности съемки (субфасет «Сезонность съемки») используется общепринятое деление годового цикла на календарные времена года {«зима», «весна», «лето», «осень»}, а состав фасета «Генетический тип побережья» определяется геоморфологией Крымского п-ова и соответствует монографии [7].

Состав фасета «Географический регион» и субфасетов «Дата начала периода» и «Дата окончания периода» не фиксирован в коде приложения и допускает модификацию пользователем. В частности, в системе предусмотрена возможность работы с данными о регионе съемки, которые являются основой для базы геоинформации.

Для хранения каталога и данных геоинформационной части программной системы использована встраиваемая *SQL*-ориентированная свободно распространяемая СУБД с открытым исходным кодом *SQLite3*⁴⁾. Основные метаданные фотографии и идентификаторы ее связей с элементами фасетов каталога содержатся в таблице *Picture*. При отображении структуры каталога для каждого фасета и субфасета «Сезонность съемки» была выделена отдельная таблица (рис. 1). Данные субфасетов «Дата начала периода» и «Дата окончания периода» не существуют без соответствующей фотографии и реализованы как атрибуты изображения в таблице *Picture*. В общем случае с целью получения нормализованной базы данных связь между конкретной фотографией и фасетами каталога строится одним из следующих способов:

- для фасетов с кратностью связи 1 : 1 или 1 : *N* связь организуется путем включения соответствующего идентификатора в таблицу *Picture*;
- для фасетов с кратностью связи *M* : *N* создается дополнительная таблица.

Отображение иерархического фасета «Географический регион» на реляционную структуру базы данных осуществляется с помощью списка смежности по таблице *GeoRegion* [8, 9].



Р и с . 1 . Схема базы данных программной системы «ФотоБерега Крыма»
 F i g . 1 . The scheme of the database for the software system PhotoCoasts of Crimea

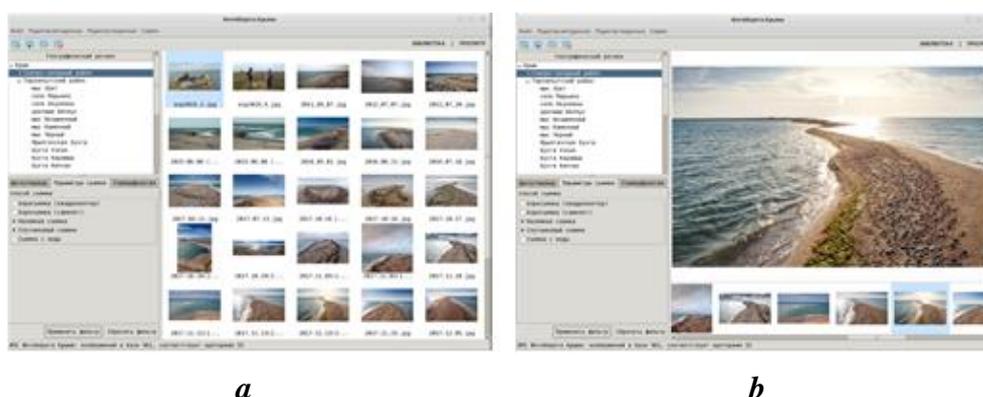
⁴⁾ SQLite : сайт. 2000 — . URL: <https://sqlite.org/index.html> (дата обращения 02.12.2024).

Для хранения исходных данных программная система использует иерархию в файловой системе жесткого диска. Каждой фотографии соответствует отдельная директория, содержащая оригинал цифрового изображения и миниатюру.

Интерфейс приложения разработан с использованием пакета *tkinter*⁵⁾ и состоит из главного окна, редактора метаданных и редактора геоданных (рис. 2).

Пользователь определяет требуемую совокупность характеристик изображения, формируя поисковый запрос с помощью панели поиска, расположенной в левой части главного окна приложения. Поисковая панель содержит отдельный раздел для каждого фасета каталога и позволяет составлять запрос интуитивно понятным способом без использования специализированных языков запросов. Результат поиска отображается в рабочей области главного окна и визуализируется в двух основных режимах. Режим библиотеки позволяет оценить объем поисковой выдачи и получить общее представление о результатах поиска. В свою очередь, режим просмотра обеспечивает возможность детального рассмотрения каждого снимка и связанной с ним метainформации. Так как цифровые изображения, получаемые с использованием современных средств съемки, могут иметь высокое разрешение, которое существенно превышает разрешение компьютерного монитора, в режиме просмотра реализована функция масштабирования снимка, в том числе автоматического масштабирования под размер окна приложения.

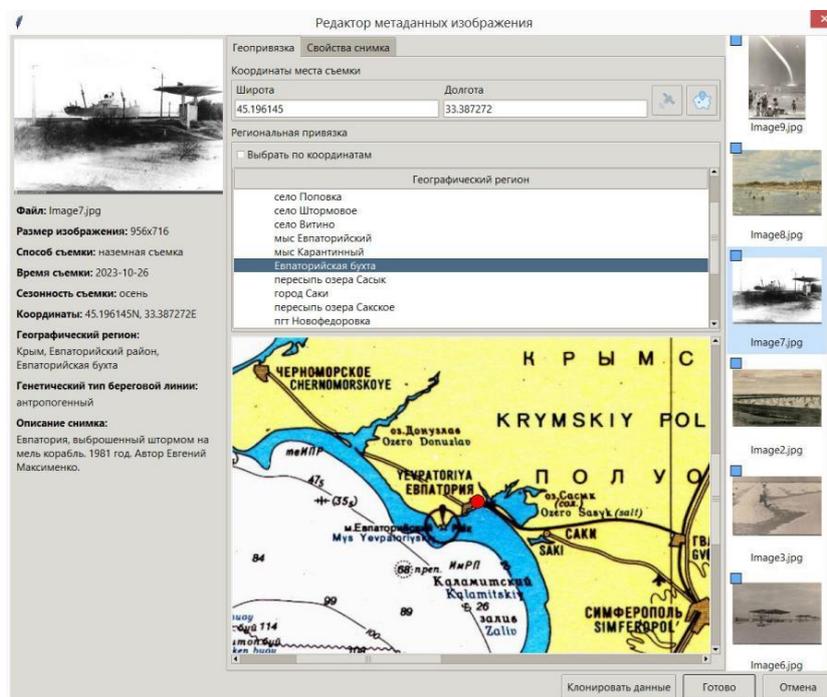
Загрузка новых изображений и редактирование метаданных существующих элементов каталога осуществляются в редакторе метаданных (рис. 3). С его помощью можно изменить координаты места съемки, ее географический регион, дату, сезон, способ съемки, генетический тип побережья,



Р и с . 2 . Пользовательский интерфейс главного окна приложения, режим «БИБЛИОТЕКА» (*a*) и режим «ПРОСМОТР» (*b*)

Fig. 2. The user interface of the application main window: LIBRARY mode (*a*) and VIEW mode (*b*)

⁵⁾ Python Software Foundation : Graphical User Interfaces with Tk : сайт. 2001 – .
URL: <https://docs.python.org/3/library/tk.html> (дата обращения 02.12.2024).



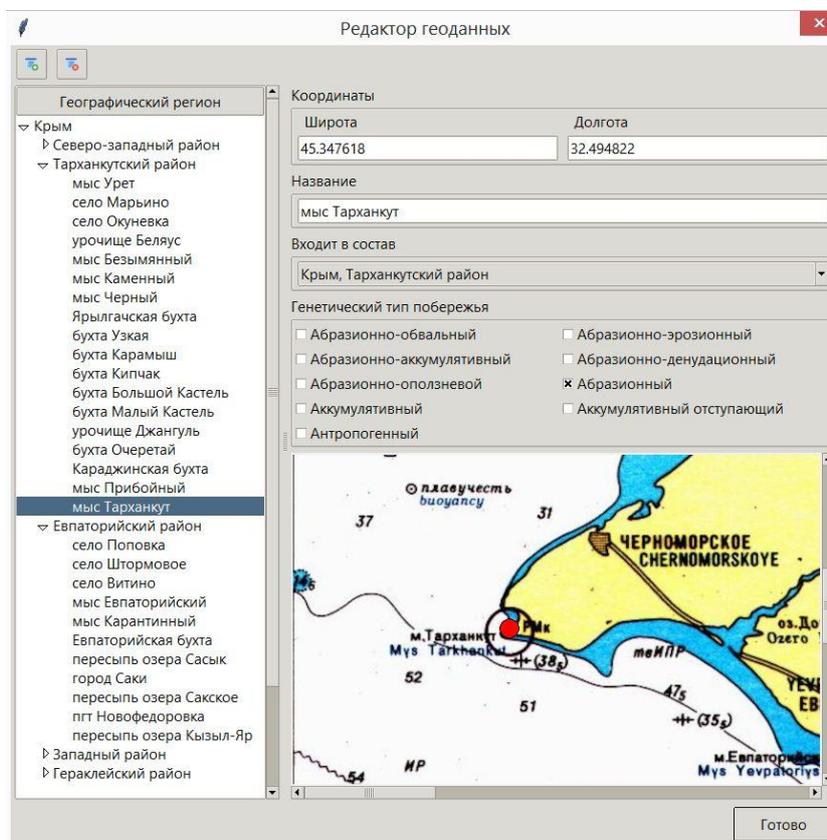
Р и с . 3 . Окно редактора метаданных изображений

Fig. 3. The Image metadata editor window

а также внести коррективы в описание снимка. Режим массовой загрузки позволяет пользователю быстро добавить в базу изображения со сходными метаданными, например, результаты экспедиционной съемки на конкретной местности.

При заполнении метаданных изображения в разрабатываемой программной системе была реализована возможность чтения данных из заголовков *Exif* загружаемых файлов (в том числе координат места съемки по данным *GPS* и даты съемки) и внедрена функциональность, позволяющая по координатам места съемки определить регион и генетический тип побережья. В качестве региона съемки выбирается ближайший из известных приложению. В частности, в базе ИПС «ФотоБерега Крыма» присутствует информация о 140 регионах побережья из монографии [7]. Для поиска региона, ближайшего к точке съемки, в программе применяется индексация геоданных береговой линии на равномерной сетке в полярной системе координат [10] с центром в точке 45.5° с. ш., 34.0° в. д.

Работа с геоинформационной частью базы метаданных осуществляется в редакторе геоданных (рис. 4), с помощью которого можно создавать, удалять и корректировать сведения о регионах съемки, в том числе координаты, наименование и генетический тип побережья.



Р и с . 4 . Окно редактора геоданных
 F i g . 4 . The Geodata editor window

Заклучение

На базе концепции программного продукта «ФотоБерега» была реализована специализированная информационно-поисковая система «ФотоБерега Крыма», оптимизированная для решения практической задачи – систематизации, классификации и работы с коллекцией цифровых изображений берегов Крымского п-ова при проведении научных исследований. В статье описана структура каталога изображений и способ хранения метаданных объекта съемки. Изложен порядок работы с системой при выполнении поиска информации и загрузке новых изображений. Ключевой особенностью разработанного программного продукта является наличие модуля формирования базы геоинформации о крымском побережье. Использование геоинформации совместно с возможностью чтения метаданных цифрового изображения и массовой загрузкой изображений существенно облегчает внесение информации в каталог системы.

Разработанная информационно-поисковая система «ФотоБерега Крыма» имеет значительный потенциал для дальнейшего развития. Ее функциональные возможности можно расширить и адаптировать для работы с изображениями

берегов других регионов. Продукт является универсальным и после соответствующей настройки может найти применение для систематизации, классификации и работы с коллекциями цифровых фотографий в самых разных областях науки.

СПИСОК ЛИТЕРАТУРЫ

1. Информационная поддержка исследований прибрежных зон Черного и Азовского морей / Е. А. Годин [и др.] // *Натурные и теоретические исследования – в практику берегопользования : сборник материалов, представленных на Всероссийскую конференцию с международным участием: XXIX Береговая конференция: Калининград, 2022. С. 330–333.*
2. *Вецало М. П., Годин Е. А. Разработка программной системы базы данных фотографических изображений берегов Крыма // Моря России : вызовы отечественной науки. Тезисы докладов Всероссийской научной конференции. Севастополь, 26–30 сентября 2022 г. Севастополь : ФГБУН ФИЦ МГИ, 2022. С. 287–289. EDN KZNERH.*
3. *Черный А. И. Введение в теорию информационного поиска. Москва : Наука, 1975. 238 с.*
4. *Ранганатан Ш. Р. Классификация двоеточием. Основная классификация. Москва : ГПНТБ СССР, 1970. 422 с.*
5. *Маннинг К. Д., Прабхакар Р., Шютце Х. Введение в информационный поиск. Москва : Вильямс, 2011. 520 с.*
6. *Witten I. H., Moffat A., Bell T. C. Managing gigabytes: compressing and indexing documents and images. Second edition. Morgan Kaufmann, 1999. 519 p.*
7. *Горячкин Ю. Н., Долотов В. В. Морские берега Крыма. Севастополь : Колорит, 2019. 256 с.*
8. *Богданов Д. В. Оптимальный способ хранения и обработки древовидных структур в базах данных // Программные продукты и системы. 2009. № 1. С. 140–142. EDN NOJDWN.*
9. *Тарасов С. В., Бураков В. В. Способы реляционного моделирования иерархических структур данных // Информационно-управляющие системы. 2013. № 6. С. 58–66. EDN RPREYR.*
10. *Bentley J. L., Friedman J. H. Data structures for range searching // ACM Computing Surveys. 1979. Vol. 11, iss. 4. P. 397–409. <https://doi.org/10.1145/356789.356797>*

Поступила 14.03.2024 г.; одобрена после рецензирования 7.06.2024 г.; принята к публикации 18.09.2024 г.; опубликована 20.12.2024 г.

Об авторах:

Вецало Максим Петрович, ведущий инженер-программист, Морской гидрофизический институт РАН (299011, Россия, г. Севастополь, ул. Капитанская, д. 2), **SPIN-код: 4199-5264, ORCID ID: 0000-0002-3543-2124, Scopus ID: 57222028338, mvetsalo@mhi-ras.ru**

Годин Евгений Александрович, научный сотрудник, Морской гидрофизический институт РАН (299011, Россия, г. Севастополь, ул. Капитанская, д. 2), **SPIN-код: 9561-8338, ORCID ID: 0000-0002-6469-1379, Scopus ID: 56950615200, ResearcherID: AEP-0342-2022, godin_ea@mhi-ras.ru**

Исаева Елена Андреевна, ведущий инженер-программист, Морской гидрофизический институт РАН (299011, Россия, г. Севастополь, ул. Капитанская, д. 2), **SPIN-код: 5366-7440, ORCID ID: 0000-0002-1860-0026, Scopus ID: 57191413519, isaeva-ea@mhi-ras.ru**

Галковская Людмила Константиновна – ведущий инженер-программист, Морской гидрофизический институт РАН (299011, Россия, г. Севастополь, ул. Капитанская, д. 2), **SPIN-код: 9205-0925, galkovskaya@gmail.com**

Заявленный вклад авторов:

Вецало Максим Петрович – разработка алгоритма и программная реализация программного продукта, подготовка текста статьи и иллюстративного материала

Годин Евгений Александрович – постановка проблемы, определение состава и структуры системы, подготовка текста статьи

Исаева Елена Андреевна – подготовка массива изображений, тестирование программного продукта, подготовка текста статьи и иллюстративного материала

Галковская Людмила Константиновна – подготовка массива изображений, тестирование программного продукта

Все авторы прочитали и одобрили окончательный вариант рукописи.