



К 55-летию Геологического музея им. А. А. Чернова The 55th anniversary of the A. A. Chernov Geological Museum

УДК 550:004.043

DOI: 10.19110/geov.2023.5.3

Разработка базы данных фондов Геологического музея им. А. А. Чернова

К. С. Попвасев

Институт геологии ФИЦ Коми НЦ УрО РАН, Сыктывкар
kspopvasev@geo.komisc.ru

В современных музеях информационно-аналитические системы решают множество задач, связанных с музейной деятельностью. Применение общепринятых программных комплексов, разработанных для музеев, подведомственных Министерству культуры РФ, в естественно-научных музеях ограничено. Зачастую такие музеи разрабатывают собственные базы данных и информационные системы.

Данная работа посвящена разработке базы данных фондов Геологического музея им. А. А. Чернова. Изложены основные результаты исследования — проектирование даталогической модели. Сформулированы основные вопросы, требующие в дальнейшем тщательной проработки.

Ключевые слова: базы данных, даталогическое проектирование, естественно-научные музеи.

Development of the database of the collections of the Geological museum named after A. A. Chernov

K. S. Popvasev

Institute of Geology FRC Komi SC UB RAS, Syktyvkar

In modern museums, information and analytical systems complete many tasks related to museum activities. The application of generally accepted software systems, developed for subordinate museums of the Ministry of Culture of the Russian Federation, is limited in natural science museums. Often, natural science museums develop their own databases and information systems.

This article describes the development of a database of collections of the Geological Museum named after A. A. Chernov. We present main results of study — the design of a datalogical model.

Keywords: databases, datalogical design, natural science museums.

Введение

Информационные технологии широко используются во всех сферах музейной деятельности (фондовой, научной, экспозиционной, просветительской). Развитие и совершенствование систем накопления, систематизации, хранения и управления информацией является одним из приоритетных направлений во всех музеях.

За 55 лет существования Геологического музея им. А. А. Чернова в фондах накопилось огромное количество музейных предметов — более 220 тыс. единиц хранения. В музейном собрании выделяются: основной фонд, состоящий из выставочных и монографических коллекций, фонд рабочих коллекций и обменный фонд. Согласно основным направлениям геологической науки, коллекции, входящие в состав вышеуказанных фондов, разделены на палеонтологические, минералогические, литологические, петрографические, а также коллекции полезных ископаемых и руд. Дополнительно в музее выделяются историко-архивный фонд, включающий в себя архивы ученых-геологов (карты, схемы, фотографии, медали, личные вещи) и библиотечный фонд.

Учет и систематизация фондов музея ведётся на бумажных носителях, которые с течением времени подвергаются процессам естественного старения, износа по причине неправильного хранения, что, в свою очередь, порождает ряд негативных последствий, вплоть до потери информации.

За последние десятилетия в музее проведен большой объем работы по переводу сопроводительной документации в электронный формат. В частности, на 90 % выполнена работа по переводу содержания паспортов коллекций основного фонда, продолжаются работы по созданию электронных каталогов историко-архивного и библиотечного фондов. Несмотря на проделанную работу, вопросы накопления, обработки и хранения разнотипных по структуре данных остаются актуальными.

В связи с этим перед нами стоит задача — аккумулировать и систематизировать накопленные данные в единой базе с логичной многоуровневой структурой,

Для цитирования: Попвасев К. С. Разработка базы данных фондов Геологического музея им. А. А. Чернова // Вестник геонаук. 2023. 5(341). С. 25–29. DOI: 10.19110/geov.2023.5.3

For citation: Popvasev K. S. Development of the database of the collections of the geological museum named after A. A. Chernov. Vestnik of Geosciences, 2023, 5(341), pp. 25–29, doi: 10.19110/geov.2023.5.3

с возможностью оперативной корректировки и средствами быстрого поиска информации.

Для решения подобного рода задач существуют базы данных, обеспечивающие сбор, хранение, обработку, поиск и выдачу информации, необходимую в процессе работы. На отечественном рынке музейного программного обеспечения основными информационными системами являются общепринятые для сферы культуры АС «Музей» ГИВЦ Минкультуры России и КАМИС (ООО «КАМИС»). Эти программные средства охватывают широкий спектр музейной деятельности: формирование музейных баз данных; ведение приема, учета и движения музейных предметов; оформление учетно-хранительской документации; подготовку паспортов коллекций, инвентарных карточек, различных списков и каталогов; формирование электронных интерактивных публикаций и т. д. Однако их применение в геологических музеях ограничено в связи с тем, что геологические музейные коллекции имеют ряд специфических особенностей, отличающих их от художественных или исторических. К тому же приобретение и обслуживание таких систем требует значительных затрат.

Научно-исследовательские работы по проектированию базы данных фондов Геологического музея им. А. А. Чернова проводились С. И. Плосковой (Плоскова, 1997, 1998). По результатам была построена информационная модель музея на основе реляционных баз данных, а также опробована программа создания и использования базы данных на уровне отдельной коллекции. Программа была реализована на языке Clipper в среде операционной системы MS-DOS. Однако в связи выходом из строя запоминающего устройства (HDD) программа и ее исходный код были утрачены. Тем не менее часть наработок С. И. Плосковой сохранилась.

В настоящее время развитие и распространение веб-технологий, повышение доступности интернета совершили настоящую революцию. В интернете представлены бесплатные инструменты и технологии разработки программного обеспечения, подробная литература, документальная поддержка. Эти факторы существенно снижают затраты на создание специализированных информационных систем и баз данных и позволяют разрабатывать уникальные, «заточенные» под конкретные цели программные продукты.

В силу данных обстоятельств для хранения, систематизации и оперативного представления сведений, необходимых для дальнейшего использования в научно-исследовательской, экспозиционной и просветительской деятельности, в Геологическом музее им. А. А. Чернова разрабатывается база данных фондов музея.

Результаты

Процесс проектирования базы данных состоит из нескольких этапов, отражающих переход от частного словесного описания информационной структуры предметной области к описанию объектов в терминах модели (Голицина и др., 2005). Ниже перечислены основные этапы проектирования баз данных (Карпова, 2001):

- 1) сбор сведений и системный анализ предметной области;
- 2) инфологическое проектирование;

3) выбор системы управления базами данных (СУБД);

4) даталогическое проектирование;

5) физическое проектирование.

Первым и важнейшим этапом при проектировании базы данных является описание проблемно-предметной области. В нашем случае такой областью являются естественно-научные музеи геологической тематики. Геологические музеи представляют собой центры накопления знаний о Земле как о геологическом объекте. В большинстве случаев геологические объекты в силу своих характеристик несовместимы с пространством музея и не могут быть размещены в нем целиком. Они представлены только коллекциями — своеобразными моделями реально существующего (или существовавшего) объекта. В состав коллекции, в свою очередь, входят геологические образцы, тесно связанные с самим природным телом. Разностороннее изучение образца позволяет накапливать множественную информацию о нем. Появление новых научных гипотез, методов исследования, детализации и корректировки номенклатур, развитие приборной базы позволяют исследователю неоднократно изучать один и тот же образец, получая новые данные. Из этого следует, что образцы обладают переменным информационным полем. Часто бывает так, что один-единственный предмет, не имеющий самостоятельной музейной ценности, при включении в коллекцию приобретает совершенно другие свойства. К примеру, образец гранита, имеющий данные о привязке, возрасте, химическом и минеральном составе и т. д. — это совсем иное, чем гранит без точной привязки. Именно уточняющие факты придают научную ценность геологическому музейному предмету, не имеющему непосредственной привлекательности. Таким образом, разрабатываемая база данных должна позволять включать образец во все тематические выборки по заложенным признакам, независимо от физического места хранения.

Следующий этап разработки базы данных — проектирование инфологической модели. Под инфологической моделью данных понимается описание предметной области в терминах и единицах, не поддерживаемых СУБД. При разработке базы данных фондов музея за основу взята модель «сущность – связь», разработанная Питером Ченом (Chen, 1976). В этой модели поддерживаются три типа моделируемых объектов: «сущность», «атрибут» и «связь». Атрибут — некоторое свойство, характеристика сущности. Под сущностью понимаются «единица хранения» и «коллекция». В качестве перечня свойств конкретной сущности могут быть взяты общие и специфические признаки объектов — носителей информации. Одноименные сущности, для которых определены совпадающие множества свойств, объединяются во множество сущностей (например, множество образцов, множество коллекций). Между множествами сущностей могут действовать различные связи. Конкретный музейный образец всегда связан с коллекцией, а коллекция составляет часть фонда. Связи, как и сущности, могут обладать свойствами. Существуют различные типы бинарных связей, т. е. связей, определенных на двух множествах сущностей. Если их рассматривать не относительно множества связей как такового, а относительно одного из множеств сущностей, на котором определено данное мно-



жество связей, то различаются четыре типа множеств связей: „1:1” (один к одному), „1:N” (один к многим), „N:1” (многие к одному), „N:N” (многие к многим). Наиболее распространенными типами связей при построении модели являются „1:N” и „N:1”. Связи этих типов используются для представления отношений типа «состоит из», «включает», «составлен», а также естественной иерархии объектов. После определения сущностей и их свойств выявляются связи, свойства и типы связей, строится диаграмма «сущность – связь». На диаграмме отображены в виде прямоугольников множества сущностей, множество связей отображается в виде ромба (рис. 1).

После выбора модели базы данных проводится выбор СУБД. При выборе учитывались прежде всего возможности работы системы с большими объемами данных, функциональные возможности, наличие подробной литературы и возможность интеграции с другими программными продуктами. Под заданные кри-

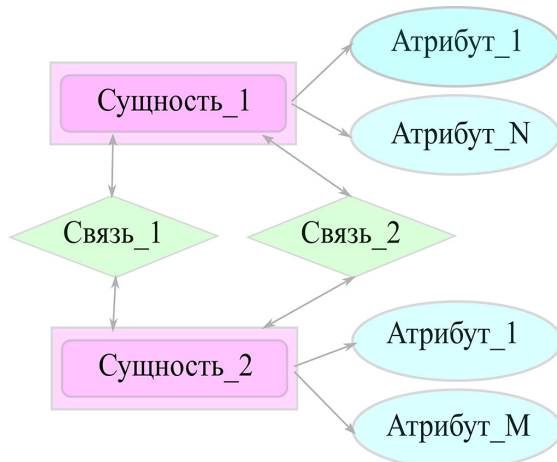


Рис. 1. Пример модели «сущность – связь» Питера Чена

Fig. 1. Example of the «entity – relationship» model by Peter Chen

терии из наиболее популярных и развитых подходит СУБД MySQL. Данная СУБД обладает следующими преимуществами при её выборе (Филиппов, 2007): является реляционной моделью данных; свободно распространяется; адаптирована для веб-приложений; имеет хорошую базу технической документации, литературы и практических примеров; содержит множество плагинов; хорошо интегрируется в другое программное обеспечение; имеет низкие требования к ресурсам аппаратного обеспечения.

После того как определено множество переменных, описывающих проблемно-предметную область, функциональные зависимости между ними и возможные ключи для получения оптимальной структуры базы данных в единицах, допустимых выбранной СУБД, строится даталогическая модель в специализированном веб-приложении (<https://dbdesign.online/>).

Модель (рис. 2) включает в себя 18 таблиц, из которых 6 («Фонд», «Печатная работа», «Издательство», «Раздел геологической науки», «Авторы», «Место хранения») описывают библиотечный фонд, 16 – основной фонд, фонд рабочих коллекций и обменный фонд. Четырьмя таблицами описывается историко-архивный фонд («Фонд», «Коллекции», «Место хранения»,

«Автор»). Согласно разработанной даталогической модели можно установить, какие атрибуты характеризуют определенную сущность. На данном этапе проектирования количество сущностей немного меньше, чем количество связей. Однако на этапе физического проектирования в целях уменьшения громоздкости число отношений возрастет в связи с их разбиением. Для лучшего восприятия связи между основными таблицами выделены цветом: желтым – с таблицей «Авторы», зеленым – с таблицей «Привязка», красным – с таблицей «Коллекции», синим – с таблицей «Место хранения». Помимо перечисленных, в базе данных присутствуют таблицы: «Тип коллекции», «Биостратиграфический возраст», «Абсолютный возраст», «Свита». Таким образом, даталогическая модель охватывает всю совокупность образцов музейного собрания.

Завершающий этап – физическое проектирование базы данных. На данном этапе решается задача грамотного размещения объектов базы данных в пространстве памяти, строятся индексы, определяется целесообразность использования хеширования и кластеризации (Карпова, 2001). Физическое проектирование баз данных состоит из нескольких этапов:

1. Создание таблиц с учетом особенностей выбранной СУБД. На данном этапе реализуются связи, представления, проводится индексирование.

2. Реализация бизнес-правил в средствах выбранной СУБД при помощи создания хранимых процедур и триггеров.

3. Определение нагрузки на элементы базы данных и проведение оценки пропускной способности и времени отклика на запросы. При необходимости проводится оптимизация физической модели базы данных.

4. Обеспечение безопасности информации. Определяются пользовательские группы, производится разграничение прав доступа.

5. Осуществление постоянного мониторинга базы данных и СУБД. При необходимости физическая модель базы данных модернизируется.

В ходе проведения анализа проектируемых информационных полей был сделан вывод, что переход к физическому проектированию базы данных фондов Геологического музея им. А. А. Чернова невозможен из-за выявленных проблем. Во-первых, для предотвращения многочисленных ошибок в ходе заполнения базы данных требуется создать справочники названий минералов, горных пород и палеонтологических видов. Во-вторых, необходимо составить словарь местных стратиграфических шкал, увязанных с региональной и общей стратиграфической шкалами, а также подготовить материал на уровне свит с указанием следующих параметров: возраста, имя свиты, мощности, текстового блока литологического описания, указания фоссилий и их латинских названий. В-третьих, нужно подготовить словари административно-территориального, физико-географического деления региона. Реализация систем справочников в базе данных приведет к значительному сокращению объема за счет отсутствия дублирования информации от записи к записи.

И все же главной задачей при создании музейной базы данных фондов Геологического музея им. А. А. Чернова является не просто перенос учетной документа-

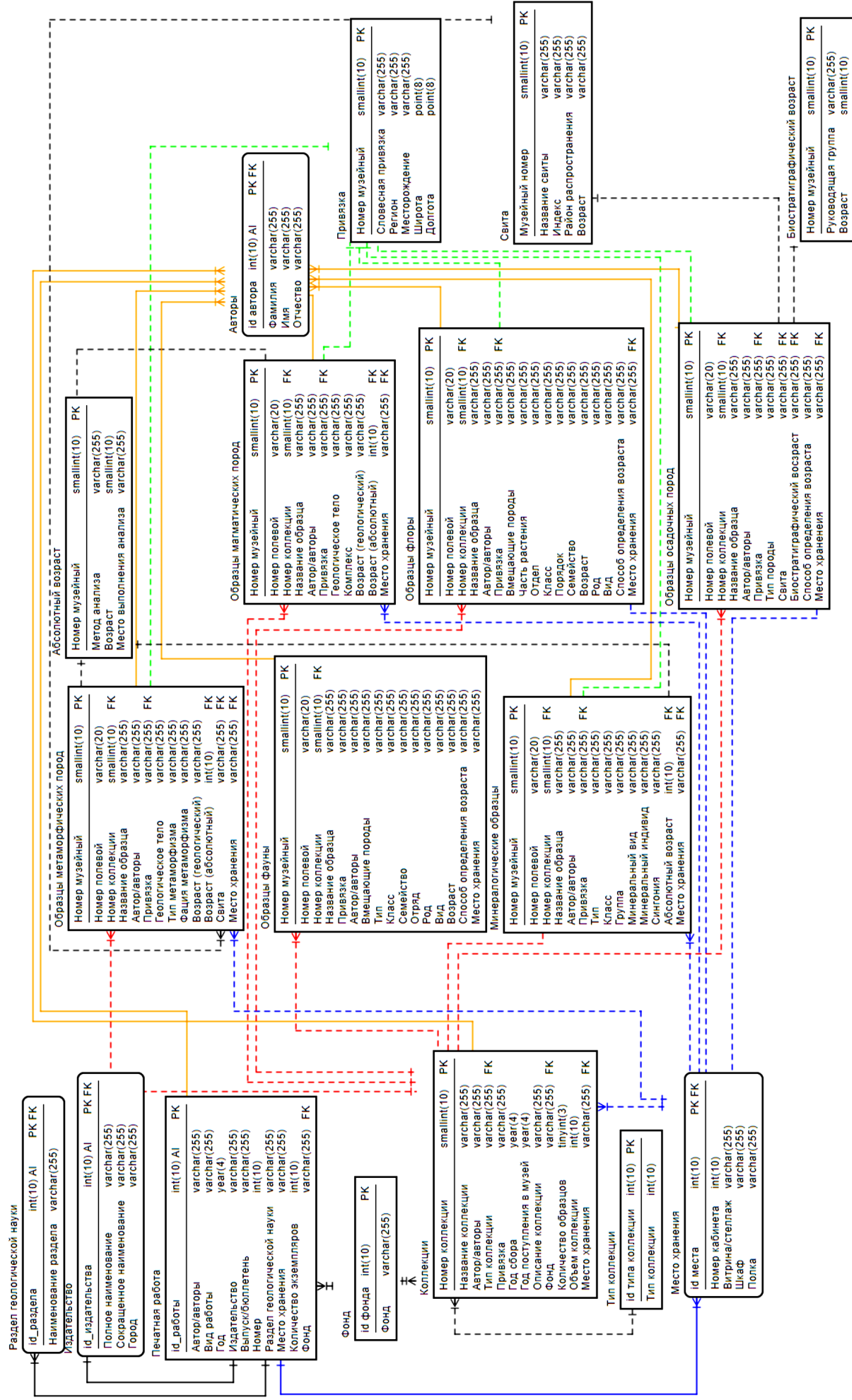


Рис. 2. Даталогическая модель данных фондов Геологического музея им. А. А. Чернова (пояснения в тексте)
 Fig. 2. Datological model of the collections of the Geological Museum named after A. A. Chernov (explanations in the text)



ции, а разработка удобной поисковой системы, позволяющей по ряду выбранных ключевых слов собрать информацию, необходимую для специалистов из разных областей, а не только геологов.

Выводы

В процессе исследования были получены следующие результаты.

Проведено описание Геологического музея им. А. А. Чернова как проблемно-предметной области.

Представлен системный анализ потоков геологической информации (путей движения внутри проблемной области исходных документов, содержащих геологическую информацию).

Составлена информационная модель проблемной области (описание в некоторых условных единицах информационного пространства проблемной области).

На сегодняшний день база данных геологических фондов музея находится на стадии перехода от даталогического проектирования к физическому. Ведутся работы по созданию справочников и словарей. Параллельно с этим в средствах выбранной СУБД проводятся работы по созданию базы данных, пользователей, пользовательских типов, кластеров и таблиц. Дальнейшая работа направлена на создание представлений, синонимов и последовательностей, разграничение прав доступа, создание процедур, функций и триггеров.

Работа выполнена в рамках темы НИР «Развитие минерально-сырьевого комплекса Тимано-Североуральско-Баренцевоморского региона на основе эффективного прогноза, геологического моделирования, геолого-экономической оценки ресурсного потенциала и новых технологий переработки полезных ископаемых» (122040600011-5; FUUU-2022-0059).

Автор выражает благодарность рецензентам за ценные советы и критические замечания.

Литература / References

Голицына О. Л., Максимов Н. В., Попов И. И. Базы данных: Учебное пособие. М.: ФОРУМ: ИНФРА-М, 2005. 352 с.
Golitsina O. L., Maksimov N. V., Popov I. I. *Bazy dannykh* (Databases): Textbook, Moscow: FORUM: INFRA-M, 2005, 352 p.

Карпова Т. С. Базы данных: модели, разработка, реализация. СПб.: Питер, 2001. 304 с.

Karpova T. S. *Bazy dannykh: modeli, razrabotka, realizatsiya* (Databases: models, development, implementation). St. Petersburg: Piter, 2001, 304 p.

Онлайн-инструмент для проектирования и моделирования схемы базы данных. URL: <https://dbdesign.online/> (дата обращения: 02.05.2022).

Online Database Schema Design and Modeling Tool URL: <https://dbdesign.online/> (date of access: 02.05.2022)

Плоскова С. И. Концепция информационной модели научного регионального геологического музея // Структура, вещество, история литосферы Тимано-Североуральского сегмента: Материалы VII науч. конф. Института геологии Коми НЦ УрО РАН. Сыктывкар: Геопринт, 1998. С. 126–128.

Ploskova S. I. *Kontseptsiya informatsionnoy modeli nauchnogo regional'nogo geologicheskogo muzeya* (The concept of the information model of the scientific regional geological museum). Structure of matter, history of the lithosphere of the Timan-North Ural segment: proceedings of the 7th scientific conference of the Institute of Geology of the Komi SC UB RAS. Syktyvkar: Geoprint, 1998, pp. 126–128.

Плоскова С. И. Информационная модель научного регионального геологического музея // Структура вещества, история литосферы Тимано-Североуральского сегмента: Материалы VIII науч. конф. Института геологии Коми НЦ УрО РАН. Сыктывкар: Геопринт, 1999. С. 133–135.

Ploskova S. I. *Informatsionnaya model' nauchnogo regional'nogo geologicheskogo muzeya* (Information model of the scientific regional geological museum). Structure of matter, history of the lithosphere of the Timan-North Ural segment: proceedings of the 8th scientific conference of the Institute of Geology of the Komi SC UB RAS. Syktyvkar: Geoprint, 1999, pp. 133–135.

Филиппов В. А. Электронные хранилища информации и WEB-технологии. М.: Отдельное издание, 2017. 80 с.

Filippov V. A. *Elektronnyye khranilishcha informatsii i WEB-tekhnologii* (Electronic storage of information and WEB technologies). Moscow. Otdelnoye izdaniye, 2017, 80 p.

Chen P. P.-S. The Entity-Relationship Model — Toward a Unified View of Data // ACM Transactions on Database Systems. March 1976. Vol. 1, № 1. P. 9–36.

Поступила в редакцию / Received 31.03.2023