

Геоинформатика. 2023. № 3. С. 33–43.
Geoinformatika. 2023;(3):33–43.

Информационные системы в геологии и геофизике

Научная статья

УДК 004.550

<https://doi.org/10.47148/1609-364X-2023-3-33-43>

Научный портал Geologyscience.ru: текущее состояние и перспективы развития

© 2023 г. — В.В. Наумова^{1, а)}, В.С. Ерёмченко^{1, б)}, А.А. Загумённых^{2, в)}, М.И. Патук^{1, д)}

¹ Государственный геологический музей им. В.И. Вернадского РАН; Россия, г. Москва

² Институт автоматизации и процессов управления ДВО РАН; Россия, г. Владивосток

^{а)}naumova_new@mail.ru, ^{б)}vitaer@gmail.com, ^{в)}trueepikvic@gmail.com, ^{д)}patuk@mail.ru

Аннотация: В работе обсуждается текущее состояние Портала «Информационно-аналитическая среда для поддержки научных исследований в геологии» (<http://geologyscience.ru>), а также ряд перспектив его дальнейшего развития, связанных с разработкой интеллектуального цифрового пространства геологических знаний. Предполагается, что это пространство поможет сформировать новые процессы генерации геологических данных и знаний.

Ключевые слова: научный портал, научные исследования в геологии, системы открытого доступа к геологическим данным, интеллектуальное цифровое пространство геологических знаний, веб-сервисы анализа геологических данных

Для цитирования: Наумова В.В., Ерёмченко В.С., Загумённых А.А., Патук М.И. Научный портал Geologyscience.ru: текущее состояние и перспективы развития // Геоинформатика. — 2023. — № 3. — С. 33–43. <https://doi.org/10.47148/1609-364X-2023-3-33-43>.

Information systems in geology and geophysics

Original article

Scientific portal Geologyscience.ru: current state and prospects of development

© 2023 — V.V. Naumova^{1, а)}, V.S. Eremenko^{1, б)}, A.A. Zagumennov^{2, в)}, M.I. Patuk^{1, д)}

¹ V.I. Vernadsky State Geological Museum of the Russian Academy of Sciences; Moscow, Russia

² Institute of Automation and Control Processes FEB RAS; Vladivostok, Russia

^{а)}naumova_new@mail.ru, ^{б)}vitaer@gmail.com, ^{в)}trueepikvic@gmail.com, ^{д)}patuk@mail.ru

Abstract: The paper discusses the current state of the Portal "Information and analytical environment to support scientific research in geology" (<http://geologyscience.ru>), as well as a number of its future prospects related to the development of an intelligent digital space of geological knowledge. It is assumed that this space will help to form new processes for generating geological data and knowledge.

Key words: scientific portal, scientific research in geology, open access systems to geological data, intelligent digital space of geological knowledge, web services for analyzing geological data

For citation: Naumova V.V., Eremenko V.S., Zagumennov A.A., Patuk M.I. Scientific portal Geologyscience.ru: current state and prospects of development. *Geoinformatika*. 2023;(3):33–43. <https://doi.org/10.47148/1609-364X-2023-3-33-43>. In Russ.

Интеграция разнотипных геологических данных и сервисов обработки в единую цифровую среду на основе единых политик обеспечивает возможность комплексного анализа информации и позволяет получать качественно новые знания о геологических объектах.

Системы открытого доступа к геологическим данным и системам их обработки и анализа в РФ

В настоящее время существует значительное количество открытых источников геолого-геофизических данных, как правило, курируемых соответствующими международными организациями

или крупными университетами. Такие данные могут использоваться как непосредственно при решении задач изучения (моделирования) Земли, так и в качестве дополнительных материалов при моделировании систем, в которые геологическая среда входит как одна из подсистем.

Некоторые источники открытых геолого-геофизических данных представлены в Каталоге GIS-Lab (<https://gis-lab.info/qa/geology-geophysics-open-data-sources.html>).

ФГБУ «Росгеолфонд» (<https://rfgf.ru>)

Данный электронный каталог содержит сведения из учетных карточек по всем видам геоло-

гической изученности. Предусмотрен переход от выбранных карточек к соответствующему отчету в каталоге геологических документов. Ресурс включает Государственный реестр участков недр и лицензий. Потребителям предоставлен доступ к базе данных метаданных Государственного кадастра месторождений и проявлений полезных ископаемых. Пополнение реестра объектов осуществляется ежедневно.

ФГБУ «ВСЕГЕИ»

(<https://vsegei.ru/ru/gisatlas/web-gisatlas/>)

Недра России — национальный ресурс. WEB-ГИС «Цифровой двойник недр России» — это интерфейс, интегрирующий представленные в формате web-сервисов геолого-картографические информационные системы учреждений Роснедра, открытые межотраслевые информационные ресурсы, международные проекты. Он содержит векторные карты геологического содержания территории России и ее континентального шельфа, в том числе данные по геологическому строению, геолого-геофизической изученности, полезным ископаемым, объектам лицензирования недропользования. Инструменты WEB-ГИС позволяют пользователям в интерактивном режиме выполнять разнообразные запросы — поиск, отбор, пространственное сопоставление данных, их скачивание.

Интерактивная электронная карта недропользования Российской Федерации

(<https://openmap.mineral.ru/>)

Создан в 2011-2013 гг. совместными силами центра «Минерал» ФГУНПП «Аэрогеология» и ФГБУ «ВСЕГЕИ» по заказу Управления геологических основ, науки и информатики Роснедра. Электронная карта недропользования России ориентирована на максимально простое и быстрое получение краткой справочной информации обзорного и сводного масштаба по ключевым информационным блокам, имеющим отношение к недропользованию в нашей стране.

Российские научные геолого-геофизические web-ресурсы

Геопортал Института вулканологии и сейсмологии ДВО РАН (ИВИС ДВО РАН)

(<http://geoport.kscnet.ru/>)

Геопортал ИВИС ДВО РАН — тематический web-портал, обеспечивающий единую точку доступа к вулканологическим и сейсмологическим пространственным данным и сервисам Института вулканологии и сейсмологии ДВО РАН. Геопортал является элементом инфраструктуры пространственных данных ИВИС ДВО РАН. Цель геопортала — интеграция широкого комплекса научной информации, накопленной в институте за многие годы исследований, обеспечение возможности поиска данных по их метаданным и обмена данными в сетевой среде, визуализация пространственных

данных в виде интерактивных карт. В соответствии с требованиями к интероперабельности метаданных, данных и сервисов, Геопортал ИВИС ДВО РАН реализован на основе стандартов Международной организации по стандартизации (ISO) и спецификаций Открытого геопространственного Консорциума (Open Geospatial Consortium, OGC). Архитектура геопортала разработана на базе свободных программных продуктов с открытым исходным кодом, распространяемых по лицензии GPL (GNU General Public License), что соответствует рекомендациям Консорциума OGC. Геопортал ИВИС ДВО РАН упомянут в Национальном стандарте Российской Федерации ГОСТ Р 58570-2019 «Инфраструктура пространственных данных. Общие требования» как пример реализации научного геопортала (стр. 11).

Web-ресурсы Геофизического центра РАН (ГЦ РАН)

(<https://gcras.ru/lab.php?i=3>)

Ресурсы созданы с развитием новых подходов к обработке больших объемов сложной геофизической информации на основе технологий географических информационных систем (ГИС) и методов нечеткой логики и нечеткой математики. Предусматривается развитие новых математических подходов и адаптация существующих методов распознавания образов для изучения магнитного поля Земли и решения других геофизических задач. Отдельной обширной задачей является развитие методов распознавания возмущений с заданной морфологией на временных рядах геомагнитных данных.

Информационные ресурсы Северо-Восточного комплексного научно-исследовательского института ДВО РАН (СВКНИИ ДВО РАН)

(<https://neisri.ru/informacionnyye-resursy/>)

В их числе Web-каталог текстур руд, выпуск 1, в котором представлены фотографии руд месторождений и рудопроявлений золота, серебра, олова, молибдена Верхне-Колымского региона Магаданской области. Указывается название месторождения (рудопроявления), авторский номер образца и тип минерализации, который характеризует этот образец.

Информационная система спутниковых данных Федерально-исследовательского центра Института вычислительных технологий СО РАН (ФИЦ ИВТ)

(<http://sdc.esemc.nsc.ru/>)

ФИЦ ИВТ СО РАН совместно с рядом заинтересованных институтов Сибирского и Дальневосточного отделений РАН, а также организаций других ведомств на протяжении ряда лет развивает информационную инфраструктуру, обеспечивающую хранение, архивацию и пользовательский доступ к данным дистанционного зондирования и наземных наблюдений. Эти работы направлены на создание современной системы информационной

поддержки научно-исследовательской деятельности и обеспечение научных коллективов данными о состоянии и динамике окружающей среды и социально-экономических процессов. Ключевым элементом информационной инфраструктуры является распределенная система приема и обработки спутниковых данных по территории Сибири и Дальнего Востока.

Научный портал GeologyScience.ru

Начиная с 2017 года в Государственном геологическом музее РАН (ГГМ РАН) проводятся работы по разработке и созданию Научного портала «Информационно-аналитическая среда для поддержки научных исследований в геологии» (<http://geologyscience.ru/>) (рис. 1) [6].

Основная цель данного Проекта заключается в организации единой точки доступа к геологическим данным на территорию России и системам их обработки с использованием возможностей поиска данных в территориально распределенных разнородных источниках, а также с использованием территориально распределенных вычислительно-аналитических узлов для обработки данных, взаимодействие с которыми осуществляется с использованием технологии web-сервисов.

Информационный уровень портала

Объединение тематических ресурсов в общую интегрированную информационную инфраструктуру поддержки научных исследований позволяет получить прямой доступ ко всем узлам системы,

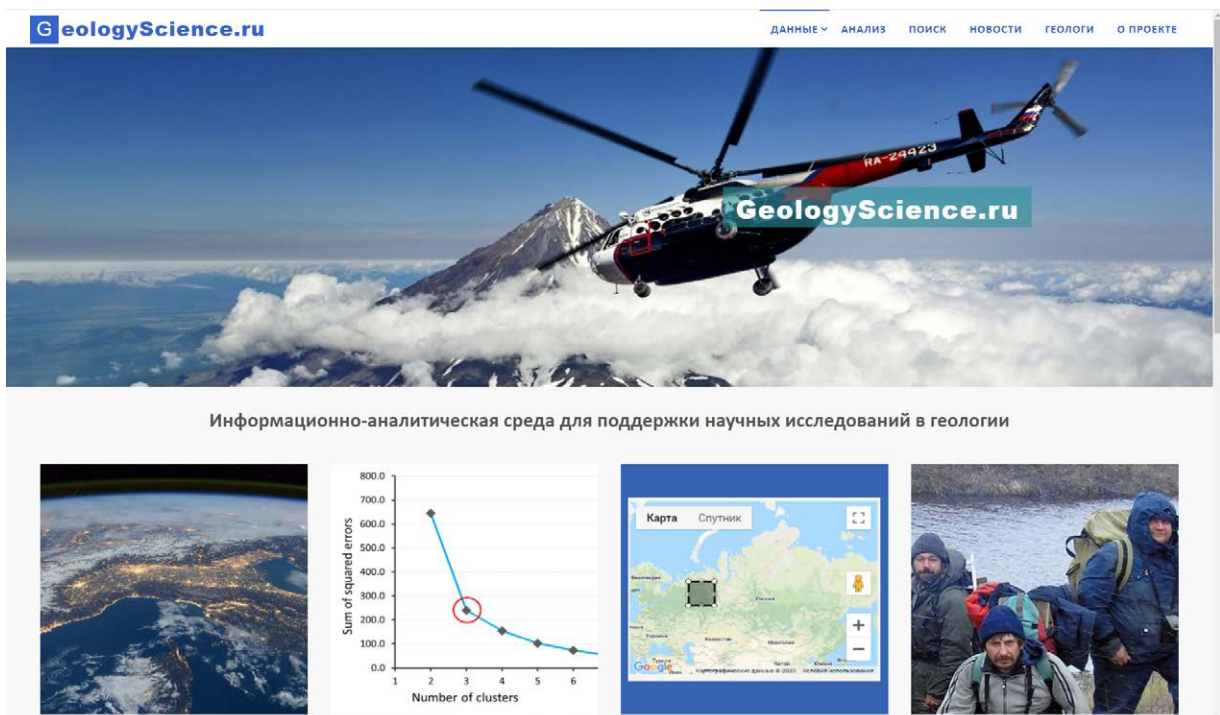
уменьшает количество действий пользователя при общении с каждым узлом, позволяет отдельным узлам и отдельным сервисам данных узлов быть интегрированными в другие информационные системы, облегчает процесс администрирования Системы.

Интеграция информационных ресурсов подразумевает выполнение функций, обеспечивающих:

1. Доступ ко всем интегрированным ресурсам через единые пользовательские интерфейсы по единым протоколам.
2. Сквозной поиск во всем множестве интегрированных информационных ресурсов, а также в их логических и физических подмножествах.
3. Извлечение информации в единых форматах.
4. Управление ресурсами и доступом к ним в соответствии с едиными политиками.
5. Контроль целостности и доступности сервисов для всех ресурсов.
6. Сбор статистической информации об использовании ресурсов.

Источники информации — территориально распределенные Интернет-ресурсы, информация в которых основана на стандартизованных метаданных и программные решения которых допускают применение стандартизованных протоколов для ее автоматической интеграции в создаваемую инфраструктуру, а также научные материалы научных организаций, библиотек, центров данных и др.

Рис. 1. Главная страница Портала GeologyScience.ru
Fig. 1. Main page of the GeologyScience.ru Portal



Сервис доступа к картографической информации

Сервис позволяет проводить быстрый поиск данных по различным критериям, и получать атрибутивную информацию об отдельном объекте, включая ссылку на данные. Для реализации сервиса каталога используется программный пакет с открытым исходным кодом GeoNetwork. Данные, представленные в каталоге, хранятся у поставщика данных на внешнем узле в виде векторных файлов, а также в виде отдельных слоёв в рамках сервисов доступа к пространственным данным, таких как OGC Web Map Service (OGC WMS) и OGC Web Feature Service (OGC WFS). На данный момент в каталоге доступны метаданные Всероссийского научно-исследовательского геологического института (ВСЕГЕИ) масштаба 1:1 000 000 третьего поколения по территории России, а также метаданные ВСЕГЕИ масштаба 1:200 000 второго поколения по территории России.

Сервис доступа к спутниковым данным

Спутниковый блок предоставляет пользователям единую точку входа к данным спутников Aqua, Terra, Landsat, orbview-3 и к другим мультиспектральным данным высокого и среднего разрешения. Источником данных служат порталы спутниковых данных Центра спутникового мониторинга Института автоматизации и процессов управления ДВО РАН, NASA, Геологической службы США (USGS). Поиск данных осуществляется в одном из трех режимов: поиск с помощью пользовательского поиска внешнего портала, поиск по метаданным, поиск по собственной базе данных метаданных спутниковых снимков. Спутниковые снимки обрабатываются в зависимости от имеющейся информации, но в любом случае пользователям предоставляется вся информация, включая обзорные изображения.

Сервис доступа к научным публикациям о геологии РФ

Сервис представляет собой репозиторий открытого доступа, созданный на основе свободно распространяемого программного обеспечения DSpace, 6.3 [8]. Основой информации являются научные статьи, монографии, диссертации, авторефераты диссертаций, тезисы докладов, материалы конференций, находящиеся в открытом доступе.

Публикации собираются из различных открытых источников: репозитории других организаций, электронных библиотек, сайтов институтов геологического профиля РАН. Интегрируя публикации в репозиторий, авторы обеспечивают единообразный формат хранения, отображения и поиска информации. Для этих публикаций введены дополнительные метаданные, которые необходимы для дальнейшего использования. В стандартном варианте системы имеются следующие параметры поиска: дата публикации, авторы, наименование

публикации. Автоматически определяется тема публикации, которая соответствует ключевым словам публикации. В параметры поиска добавлено еще два параметра — УДК и стратиграфический возраст. Название возраста определяется в автоматическом режиме из наименования либо абстракта статьи.

Название возраста определяется в автоматическом режиме из наименования либо абстракта статьи.

Для поиска и извлечения информации из других репозиториях создан скрипт на языке PHP. Извлекаемая информация фильтруется, т.е. автоматически анализируется на совпадение со словарем геологических терминов. Словарь создан на основе ключевых слов ~2000 публикаций по тематике репозитория. Оптимальным является наличие трех совпадений со словарем. При этом удается выбрать около 90% источников, соответствующих тематике репозитория. Оставшиеся 10% подвергаются ручной обработке. Эта информация служит основой для корректировки словаря.

Большое количество информации в открытом доступе представляют собой тексты в формате PDF. Добавление таких данных в репозиторий невозможно без сопутствующих этим текстам метаданных. Для извлечения метаданных из таких публикаций используется свободно распространяемое программное обеспечение: Cermine — Content Extractor and Miner, FPMI — коллекция PHP классов для обработки PDF документов, PDFMiner — программное обеспечение на Python для извлечения текстовой информации из PDF.

Извлекаемая из других репозиториях и из файлов PDF информация преобразуется в формат SIP, доступный для импорта в DSpace стандартными средствами. Для улучшения поиска информации в репозитории к существующим стандартным в DSpace поисковым тегам был добавлен тег УДК (универсальная десятичная классификация). Данная информация извлекается в полуавтоматическом режиме из выгруженного из DSpace бэкапа в формате SIP в текстовый файл с последующей загрузкой SQL скриптом в таблицу PostgreSQL DSpace.

Сервис доступа к Базам данных

На Портале организован удаленный доступ к метаинформации о месторождениях РФ и государственных геологических отчетах, находящихся в БД «Росгеолфонда», которые содержат информацию о 52 000 месторождений и 478 000 геологических отчетах.

Поиск по месторождениям использует фасетную технологию. Пользователь может выбрать месторождение по наименованию (после ввода первых четырех символов всплывает подсказка), или выбрать область, населенный пункт и т.п. Отдельно выбирается тип полезных ископаемых (на основе всплывающей подсказки).

Сервис доступа к количественным данным

Сервис осуществляет удаленный доступ к данным из мировых центров данных, например DataCite.

Сервис доступа к музейным данным

Удаленный доступ к открытым данным из мировых естественнонаучных музеев по территории России, включая доступ к данным Государственного геологического музея им. В.И. Вернадского РАН. В качестве решения использованы существующие протоколы доступа к музеям для получения выборки, связанной с Россией, и сохранением метаданных об экспонатах в виде наборов данных. Портал-агрегатор (<http://museums.sgm.ru>) разработан на платформе SKAN. Данное ПО позволяет легко оперировать наборами данных и создавать дополнительные модули, например, единый поиск по всем коллекциям.

Блок поиска данных

Система поиска преобразует сформированный пользователем запрос в последовательность запросов поиска названий и поиска по географическим координатам. После этого запросы последовательно передаются из браузера пользователя поисковым машинам отдельных блоков Портала (машины обрабатывают запросы параллельно), а они, в свою очередь, либо осуществляют поиск самостоятельно, либо обращаются к собственным поисковым системам блоков Портала или к глобальным поисковым системам.

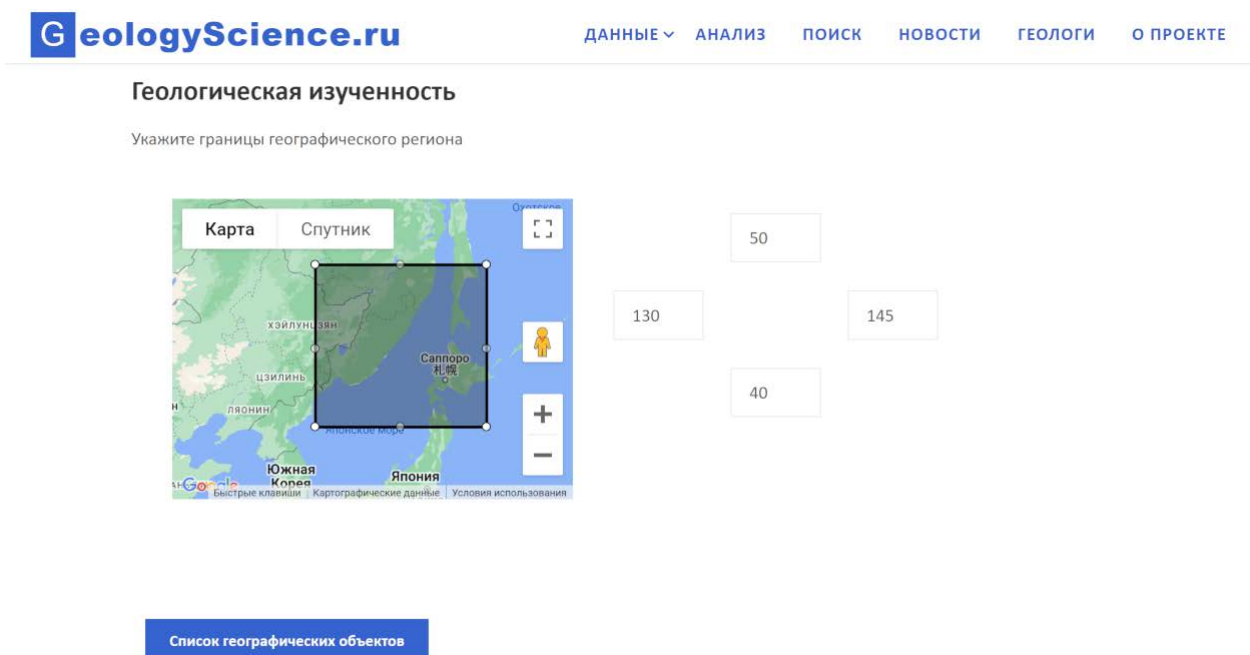
Пользователь задает координаты района поиска, а система параллельно обращается к поисковым механизмам сервисов (рис. 2). Существуют две проблемы при поиске такого рода: а) блоки почти никогда не поддерживают поиск по координатам; б) различные блоки имеют различные механизмы поиска, обладающие различной производительностью, и результат поиска должен выводиться индивидуально для каждого сервиса.

Для решения проблемы была создана отдельная подсистема, получающая список географических объектов, находящихся на заданной территории (на основе сервиса OSM). Так как полный список объектов получить невозможно (число объектов для небольшого района исчисляется тысячами), а сервис OSM предоставляет только иерархию объектов, находящихся в данной точке (страна, край, район), то выполняется регулярный обход выбранного региона с выбором 64 регулярных и 64 случайных точек с созданием списка объектов. Данный список объектов фильтруется, так, что только часто встречающиеся названия невысоких уровней иерархии используются дальше. Данное решение не является наилучшим. В нем не используется информация о собственно геологических объектах или даже о больших географических образованиях (таких как, скажем, Саянские горы), если они не находятся в базе данных OSM.

Вычислительно-аналитический блок — облачный инструментарий пользователей для обработки и анализа различных типов геологических данных [2, 3, 4]. Предложенный при построении

Рис. 2. Запрос на информацию по геологической изученности района

Fig. 2. Request for geological knowledge of the area



подход предполагает использование территориально-распределенных вычислительных узлов, взаимодействие с которыми осуществляется с использованием технологии web-сервисов, в частности OGC Web Processing Service.

Реализованная платформа выступает в роли посредника между пользователем и внешними системами анализа и обработки, предоставляя единый интерфейс доступа ко всем алгоритмам анализа и обработки, имеющимся во внешних системах. Описанная архитектура также предполагает возможность использования данных не только из доступных в системе открытых источников, но и загрузку данных для анализа и обработки самим пользователем. В разрабатываемой системе вычислительно-аналитические блоки обработки и анализа геологической информации организованы в виде наборов служебных и аналитических функций с возможностью пользовательского доступа к выбору метода анализа и обработки; цепочек анализа и обработки, включающих загрузку данных, трансформацию форматов, методов анализа и визуализации результатов; тематических цепочек, осуществляющих последовательность методов анализа. Доступ к сервисам обработки и анализа осуществляется через платформу управления распределенными сервисами анализа и обработки данных.

В настоящее время Вычислительно-аналитическая геологическая среда включает в себя следующие территориально распределенные узлы обработки:

Многомерные методы анализа данных. Вычислительный сервис «Многомерные методы анализа данных» позволяет выполнять обработку табличных данных различными методами современного анализа данных с настройкой их параметров и визуализацией результатов. Он включает в себя такие группы методов как предобработка данных, описательная статистика, кластерный анализ, факторный анализ, корреляционный анализ, регрессионный анализ и другие. Вычислительный сервис «Многомерные методы анализа данных» входит в состав Вычислительно-аналитической геологической среды ГГМ РАН и интегрирован с её сервисами. При этом он является самостоятельным облачным web-сервисом, взаимодействие с которым происходит посредством REST API. Это позволяет обращаться к многомерным методам анализа данных, размещённым на вычислительном узле, широкому кругу пользователей, в том числе предоставляет возможность интегрировать его в другие информационные системы как стороннее приложение для обработки табличных данных. Узел разработан и поддерживается в Государственном геологическом музее им. В.И. Вернадского РАН.

Обработка спутниковых данных. Включает в себя методы первичной обработки спутниковых

данных, таких как калибровка и пространственная привязка, а также управление ДВО РАН.

Обработка петролого-геохимических данных. В Институте физики Земли РАН разработана интерактивная база методов обработки петролого-геохимических данных. Система предоставляет сервисы построения спайдерграмм, гистограмм и классификационных диаграмм; сервис идентификации минералов по их химическому составу; сервис интерпретации состава минерала и разложение на миалы и т.д. Интерфейс взаимодействия с сервисами построен на основе REST архитектуры.

Структурный анализ публикаций. В междисциплинарном центре математического и вычислительного моделирования (Университет Варшавы, Польша) разработан сервис для извлечения метаданных из научных публикаций. Метаданные включают в себя авторов, принадлежность организации, абстракт, ключевые слова, название журнала, объем, год выпуска, разобранные библиографические ссылки, структуру разделов документа, заголовки разделов и абзацы. Интерфейс взаимодействия с сервисами построен на основе REST архитектуры.

Обработка естественного языка. В Университете Шеффилда в рамках проекта GATE (General Architecture for Text Engineering) разработан ряд сервисов по обработке текстовых данных для различных языков. Для обработки текстовых данных на русском языке предоставляются сервисы по определению частей речи слов, а также выделению именованных сущностей, таких как имена и фамилии, названия организаций, географические названия, даты, денежные единицы и т.д. Интерфейс взаимодействия с сервисами построен на основе REST архитектуры.

Облачные сервисы. Одним из наиболее популярных программных продуктов обработки табличных данных в интерактивном режиме является Excel из пакета Microsoft Office. Данный программный продукт содержит ряд инструментов для редактирования данных, построения различных диаграмм, использования встроенных процедур анализа и создания собственных. Компания Microsoft разработала бесплатную облачную версию продукта Excel (<https://www.office.com/>). Зарегистрированный пользователь может использовать полноценную web-версию для данных, расположенных в облачном хранилище Microsoft One Drive.

Для анализа пространственных данных наиболее подходящим инструментом являются ГИС, позволяющие в интерактивном режиме взаимодействовать с объектами на карте, применяя к отдельным слоям и объектам доступные средства анализа. Наиболее подходящим для геологии облачным решением является ArcGIS Online, разрабатываемый компанией ESRI. Portal for ArcGIS — это инфраструктура ArcGIS Online, функционирующая в защищенной ИТ-среде или в частном облаке ор-

ганизации (под контролем сетевого экрана или в полностью изолированной локальной сети). Портал позволяет создавать карты, каталогизировать и анализировать пространственные данные с помощью удобного интуитивно понятного интерфейса. Создание и публикация интерактивных карт и приложений может выполняться на любом устройстве при наличии браузера и доступа к интернету. Публикация данных в защищенном облаке Esri в виде кэшированных или динамических сервисов соответствует всем современным стандартам защиты информации с сохранением всех авторских прав.

Для анализа спутниковых данных одним из лидеров среди облачных сервисов является платформа Earth Engine компании Google (<https://earthengine.google.com>). Данная платформа позволяет пользователю загружать собственные данные или использовать данные из каталога Earth Engine для дальнейшей обработки в интерактивном режиме. В каталоге содержатся продукты обработки данных радиометра Modis (спутники Aqua, Terra), спутников Sentinel-1A, Sentinel-1B, Sentinel-2A, Sentinel-2B, Landsat 8 и др. Earth Engine содержит ряд предустановленных алгоритмов анализа, а также инструменты для их создания, редактирования и запуска с использованием языков программирования Javascript и Python. Для работы пользователю необходимо наличие аккаунта Google. Для анализа и обработки можно использовать данные из облачного хранилища Google.

Каждый из перечисленных облачных сервисов использует собственные уникальные протоколы взаимодействия, сильно затрудняя задачу интеграции сервиса в разрабатываемую вычислительную среду на основе общего протокола доступа.

Некоторые облачные сервисы требуют наличия данных в их собственном хранилище. Например, при использовании Excel из MS Office Online, данные должны находиться в пользовательском хранилище One Drive. Использование Earth Engine позволяет также использовать данные из облачного хранилища Google на аккаунте пользователя. Таким образом, для обеспечения «бесшовного» перехода между облачными сервисами необходимо разработать набор процедур для публикации выбранных пользователем данных в соответствующем пользовательском облачном хранилище.

Создание подобной процедуры стало возможным при использовании технологии web-приложений, позволяющей запросить у пользователя разрешение на доступ к определённым возможностям пользовательского аккаунта различных поставщиков облачных сервисов.

Данная технология поддерживается Microsoft, Google, Yandex, ESRI и т.д. Перед использованием соответствующего сервиса, пользователю предлагается загрузить данные для анализа в его персональное хранилище. Для этого пользователь про-

ходит авторизацию на сайте поставщика сервиса, после чего приложение запрашивает у пользователя разрешение на скачивание и публикацию данных в его хранилище.

При выборе облачного сервиса другого поставщика пользователь имеет возможность перемещения данных для обработки в хранилище данного поставщика.

Таким образом, нами предложен подход и технологическое решение для организации единого пространства данных для различных поставщиков облачных сервисов.

Перспективы развития портала

Разработка интеллектуального цифрового пространства геологических знаний

Разработка интеллектуального цифрового пространства геологических знаний рассматривается нами в качестве перспективы развития и интеграции разнотипных информационных ресурсов и сервисов в современной цифровой среде [5]. Предполагается, что это пространство поможет сформировать новые процессы генерации данных и знаний.

Сервис доступа к экспертным знаниям. Одним из примеров текущей реализации Портала является Wiki-Геология России, которая интегрирует геологическую информацию из других частей системы и из внешних данных [7]. Wiki-Геология России предоставляет пользователям не только информацию о геологических объектах, но и ряд сервисов по интеграции информации и ее анализу. Система интегрирует не только дополнительную информацию из научных публикаций, государственных геологических карт, но и спутниковые данные, музейные образцы и другую медиа информацию. В настоящее время пользователям Wiki-Геология России доступны следующие сервисы: построение географических карт районов месторождений с указанием точки, где расположено месторождение; визуализация геологических карт районов месторождений; определение возраста месторождения. В дальнейшем предполагается развитие и других сервисов.

Wiki-Геология России создается в рамках проекта GeologyScience.ru и является его составной частью, аккумулируя геологическую информацию из других частей системы и из внешних данных. На рисунке 3 указаны текущие и предполагаемые потоки информации.

Wiki создается на основе свободного ПО MediaWiki с использованием дополнительных расширений. Основное — это Semantic MediaWiki, Cargo, ExternalData, Maps и авторские модули на PHP.

Разрабатывается система интеллектуальных рабочих кабинетов пользователей, сочетающих

Рис. 3. Схема текущих и предполагаемых потоков информации для Wiki-Геология России

Fig. 3. Scheme of current and expected information flows for Wiki-Geology of Russia



в себе необходимые инструменты, алгоритмы, источники данных для решения тех или иных прикладных задач.

Интеллектуальный рабочий кабинет — это пространство, которое адаптировано под нужды и интересы конкретного пользователя [1, 9]. Он сводит к минимуму операции пользователя, не связанные с выполнением целевых функций. Интеллектуальные системы и приложения сокращают время, затрачиваемое пользователем на рутинные операции, будь то поиск контента, ответы на вопросы или решение стандартных задач, а также делают повседневную работу физически более комфортной.

Основными функциональными возможностями рабочих пространств пользователей являются:

- Интеллектуальный анализ задачи пользователя. Предполагается, что будет построена система интеллектуального анализа задачи пользователя на основе лингвистического разбора текстовой постановки задачи, что позволит предлагать пользователю данные и методы анализа для ее решения;
- Организация доступа к данным Портала и данным пользователя;
- Организация хранения запрошенных данных пользователя на общедоступных облачных ресурсах, которые предоставляет Портал;

• Визуализация данных в разных форматах, полученных пользователем, в отдельных окнах рабочего кабинета пользователя (рис. 4);

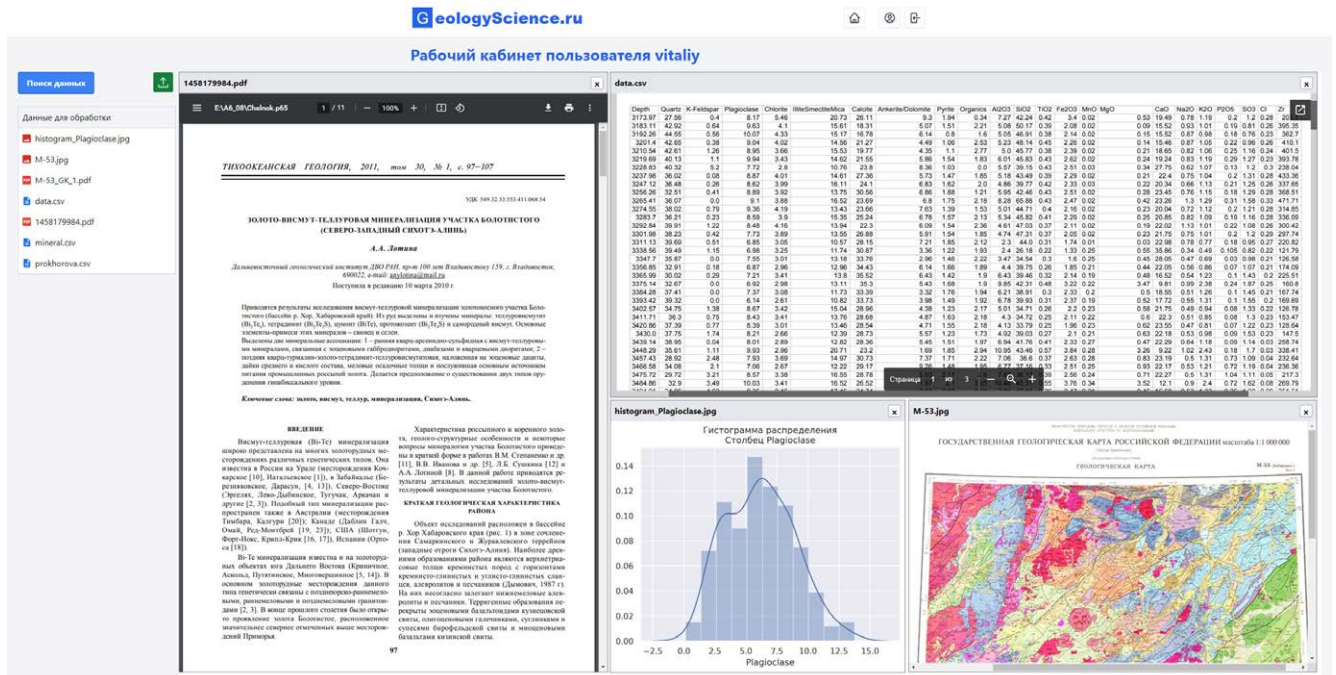
- Организация доступа к системам обработки и анализа полученных данных, включая визуализацию и хранение полученных результатов;
- Организация совместной работы территориально распределенных пользователей над проектами с использованием совместных данных.

Пользовательский интерфейс рабочего кабинета реализован с использованием Javascript фреймворка Vue 3, и CSS фреймворка Tailwind для обеспечения адаптивности интерфейса под устройства с различными параметрами экранов.

Системные сервисы Портала:

- Общая аутентификация и авторизация. Пользователь может получить доступ к другим сервисам и ресурсам системы. Токен доступа (access token) хранит в себе информацию о пользователе, роли доступа, время жизни, цифровую подпись и алгоритм подписи для последующей валидации токена. Refresh token используется для получения новой пары токенов в микросервисе аутентификации при истечении времени действия ранее выданного токена доступа. Таким образом, другие модули системы могут получить информацию о текущем пользователе и его правах доступа на основе токе-

Рис. 4. Вид рабочего кабинета пользователя Портала
Fig. 4. Interface of the Portal user's workspace



на, сопровождающего запросы пользователя, без необходимости обращения к микросервису аутентификации и доступа к базе данных пользователей.

- Организация каталога данных и сервисов Портала.
- Мониторинг доступности информационных и вычислительных ресурсов системы. Для отслеживания состояния территориально распределенных компонентов Портала разработан модуль мониторинга, позволяющий с установленной периодич-

ностью проверять доступность удалённых узлов и работоспособность используемых сервисов по указанным протоколам доступа. В качестве источника данных для мониторинга используется модуль каталога внешних сервисов обработки и анализа геологической информации.

- Сбор статистической информации о работе компонентов системы.
- Кеш базы данных системы.
- Балансировщик нагрузки.

Работы выполняются в рамках Государственного задания ГТМ РАН по Теме № 1021061009468-8-1.5.1 «Цифровая платформа интеграции и анализа геологических и музейных данных».

Список источников

1. Атаева О.М., Серебряков В.А., Тучкова Н.П. Использование семантических связей онтологии для создания адаптивного интерфейса // Электронные библиотеки. – 2023. – Т. 26. – № 1. – С. 2–17. DOI: <https://doi.org/10.26907/1562-5419-2023-26-1-2-17>.
2. Ерёмченко В.С., Наумова В.В. Система каталогизации и мониторинга территориально распределенных вычислительных узлов в среде WPS сервисов для решения геологических задач // Вестник Новосибирского государственного университета. Серия: Информационные технологии. – 2019. – Т. 17. – № 2. – С. 39–48. DOI: 10.25205/1818-7900-2019-17-2-39-48.
3. Ерёмченко В.С., Наумова В.В. Создание экосистемы по обработке данных для научных исследований в области геологии // Электронные библиотеки. – 2022. – Т. 25. – № 4. – С. 336–347. DOI: 10.26907/1562-5419-2022-25-4-336-347.
4. Ерёмченко В.С., Наумова В.В., Загуменнов А.А., Булов С.В. Облачные технологии для развития территориально распределенной вычислительно-аналитической геологической среды // Вычислительные технологии. – 2021. – Т. 26. – № 1. – С. 86–98. DOI: 10.25743/ICT.2021.26.1.007.

5. Наумова В.В., Ерёмченко В.С., Патук М.И. Разработка Единого цифрового пространства научных геологических знаний // Геоинформатика. – 2022. – № 4. – С. 8–13. DOI: 10.47148/1609-364X-2022-4-8-13.
6. Наумова В.В., Платонов К.А., Еременко В.С., Патук М.И., Дьяков С.Е. Информационно-аналитическая среда для поддержки научных исследований в геологии: текущее состояние и перспективы развития // Распределенные информационно-вычислительные ресурсы. Цифровые двойники и большие данные (DICR-2019) : труды XVII Междунар. конф. (Новосибирск, 3–6 декабря 2019 г.). – Новосибирск : ИВТ СО РАН, 2019. – С. 139–147. DOI: 10.25743/ICT.2019.70.61.021.
7. Патук М.И., Наумова В.В. Построение цифровой системы управления геологическими знаниями для поддержки научных исследований // Геоинформатика. – 2022. – № 4. – С. 14–19. DOI: 10.47148/1609-364X-2022-4-14-19.
8. Патук М.И., Наумова В.В., Ерёмченко В.С. Цифровой репозиторий "Geologyscience.ru": открытый доступ к научным публикациям по геологии России // Электронные библиотеки. – 2020. – Т. 23. – № 6. – С. 1324–1338. DOI: 10.26907/1562-5419-2020-23-6-1324-1338.
9. Sabatucci L., Seidita V., Cossentino M. The Four Types of Self-adaptive Systems: A Metamodel // Intelligent Interactive Multimedia Systems and Services 2017. KES IIMSS 18 2018 (SIST, vol.76) / eds. De Pietro G., Gallo L., Howlett R., Jai, L. – Cham: Springer, 2018. – pp. 440–450. DOI: 10.1007/978-3-319-59480-4_44.

References

1. Ataeva O.M., Serebriakov V.A., Tuchkova N.P. Modeling an adaptive interface using semantic ontology relations. *Russian digital libraries journal*. 2023;26(1):2–17. DOI: <https://doi.org/10.26907/1562-5419-2023-26-1-2-17>.
2. Eremenko V.S., Naumova V.V. The system of cataloguing and monitoring of geographically distributed computing nodes in WPS services environment for geological problems solving. *Vestnik NSU. Series: Information Technologies*. 2019;17(2):39–48. DOI: 10.25205/1818-7900-2019-17-2-39-48.
3. Eremenko V.S., Naumova V.V. Creating a data processing ecosystem for geological research. *Russian digital libraries journal*. 2022;25(4):336–347. DOI: 10.26907/1562-5419-2022-25-4-336-347.
4. Eremenko V.S., Naumova V.V., Zagumennov A.A., Bulov S.V. Cloud technologies for development of geographically distributed computational and analytical geological environment. *Computational technologies*. 2021;26(1):86–98. DOI: 10.25743/ICT.2021.26.1.007.
5. Naumova V.V., Eremenko V.S., Patuk M. I. Development of a Single digital space geological scientific knowledge. *Geoinformatika*. 2022(4): 8–13. DOI: 10.47148/1609-364X-2022-4-8-13.
6. Naumova V.V., Platonov K.A., Eremenko V.S., Patuk M.I., Dyakov S.E. Information and analytical environment to support scientific research in geology: current status and prospects for development. In: *Raspredelennye informatsionno-vychislitel'nye resursy. Tsifrovye dvoyniki i bol'shie dannye (DICR-2019) : trudy XVII Mezhdunar. konf. (Novosibirsk, 3-6 December 2019)*. Novosibirsk: IVT SO RAN; 2019. pp. 139–147. DOI: 10.25743/ICT.2019.70.61.021.
7. Patuk M.I., Naumova V.V. Building a digital geological knowledge management system to support scientific research. *Geoinformatika*. 2022(4):14–19. DOI: 10.47148/1609-364X-2022-4-14-19.
8. Patuk M.I., Naumova V.V., Eremenko V.S. Digital repository "Geologyscience.ru": open access to scientific publications on Russian geology. *Russian Digital Libraries Journal*. 2020;23(6):1324–1338. DOI:10.26907/1562-5419-2020-23-6-1324-1338.
9. Sabatucci L, Seidita V, Cossentino M. The Four Types of Self adaptive Systems: A Metamodel. In: *Intelligent Interactive Multimedia Systems and Services 2017. KES IIMSS 18 2018 (SIST, vol. 76)*. De Pietro G., Gallo L., Howlett R., Jai, L., eds. Springer; Cham: 2018. pp.440–450. DOI: 10.1007/978-3-319-59480-4_44.

Статья поступила в редакцию 10.07.2023, одобрена после рецензирования 20.07.2023, принята к публикации 26.08.2023.
The article was submitted 10.07.2023; approved after reviewing 20.07.2023; accepted for publication 26.08.2023.

Информация об авторах

Наумова Вера Викторовна

Доктор геолого-минералогических наук
Главный научный сотрудник, заведующая Научным отделом
Государственный геологический музей
им. В.И. Вернадского РАН
125009 Москва, Моховая ул., д. 11, стр. 11
e-mail: naumova_new@mail.ru
ORCID: 0000-0002-3001-1638

Information about authors

Vera V. Naumova

Doctor of Geological and Mineralogical Sciences, Chief Researcher
Head of Scientific Department of Vernadsky State Geological
Museum RAS (SGM RAS)
11 bild. 11, Mokhovaya st., Moscow, 125009, Russia
e-mail: naumova_new@mail.ru
ORCID: 0000-0002-3001-1638

Ерёменко Виталий Сергеевич

Младший научный сотрудник

Научный отдел Государственного геологического музея
им. В.И. Вернадского РАН

125009 Москва, Моховая ул., д. 11, стр. 11

e-mail: vitaer@gmail.com

ORCID: 0000-0002-5250-5743

Vitaliy S. Eremenko

Junior Researcher

Scientific Department of Vernadsky State Geological Museum RAS
(SGM RAS)

11 bild. 11, Mokhovaya st., Moscow, 125009, Russia

e-mail: vitaer@gmail.com

ORCID: 0000-0002-5250-5743

Загумённов Алексей Андреевич

Научный сотрудник

Институт автоматизации и процессов управления ДВО РАН
690041 Владивосток, ул. Радио, д. 5

e-mail: truepikvic@gmail.com

ORCID: 0000-0002-0501-5362

Aleksei A. Zagumennov

Researcher

Institute of Automation and Control Processes of the
Far-Eastern Branch of the Russian Academy of Sciences
5, Radio st., Vladivostok, 690041, Russia

e-mail: truepikvic@gmail.com

ORCID: 0000-0002-0501-5362

Патук Михаил Иванович

Кандидат геолого-минералогических наук

Научный сотрудник

Научный отдел Государственного геологического музея
им. В.И. Вернадского РАН

125009 Москва, Моховая ул., д. 11, стр. 11

e-mail: patuk@mail.ru

ORCID: 0000-0003-3036-2275

Michail I. Patuk

Candidate of Geological and Mineralogical Sciences, Researcher

Scientific Department of Vernadsky State Geological Museum RAS
(SGM RAS)

11 bild. 11, Mokhovaya st., Moscow, 125009, Russia

e-mail: patuk@mail.ru

ORCID: 0000-0003-3036-2275