

УДК 004.85

РАЗРАБОТКА ЦИФРОВОЙ ИНТЕЛЛЕКТУАЛЬНОЙ СФЕРЫ ГЕОЛОГИЧЕСКИХ ЗНАНИЙ

В. В. Наумова^[0000-0002-3001-1638]

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

naumova_new@mail.ru

Аннотация

На основе проведенного анализа цифровых технологий будущего, а также современных решений для поддержки научных исследований в геологии коллектива авторов Государственного геологического музея им. В. И. Вернадского РАН разработана и предложена архитектура цифровой интеллектуальной сферы геологических знаний. Основная концепция работы заключается в создании интегрированной цифровой экосистемы, где геологические коллекции, научные данные и экспертиза естественно-научных музеев становятся частью единого, доступного и интерактивного пространства. Это не просто оцифровка экспонатов, а построение связной среды для науки, образования и популяризации.

Ключевые слова: *Единое цифровое пространство геологических знаний, цифровая геологическая среда будущего, геологические коллекции, научные данные, экспертиза естественно-научных музеев, цифровые инструменты, ИИ-ассистенты для геологии.*

ВВЕДЕНИЕ

В работах [1, 2] рассмотрен ряд перспективных цифровых технологий, которые способны радикально изменить развитие сразу многих областей и в настоящее время уже применяются или имеют перспективы применяться в геологических исследованиях. Это:

- искусственный интеллект (ИИ);
- технологии беспроводной связи;
- системы распределенного реестра;

- большие данные и квантовые вычисления;
- виртуальная и дополненная реальности (VR и AR);
- цифровые инструменты и виртуальные ассистенты.

Отметим, что развитие цифровых технологий будущего не только определяет решение частных геологических задач, но и позволяет находить принципиально новые концептуальные решения для построения цифровых информационных экосистем и платформ для геологических исследований.

Основная концепция таких работ заключается в создании интегрированной цифровой экосистемы, где научные данные различных типов (например, полевые, экспериментальные, лабораторные и др.), геологические коллекции и экспертизы естественно-научных музеев становятся частью единого, доступного и интерактивного пространства. Это построение связной среды для науки, образования и популяризации.

Интеграция научных данных и геологических коллекций музеев открывает огромные возможности для анализа геологических объектов, создавая мощный синергетический эффект.

Технологии, обеспечивающие интеграцию

Онтологии и *Linked Open Data*: единые стандарты описания (например, Darwin Core для палеонтологии) – позволяют «связать» данные из разных источников.

Искусственный интеллект и *машинное обучение* используются для автоматической классификации образцов по изображениям и выявления скрытых закономерностей в больших массивах разрозненных данных.

Блокчейн применяется для фиксации происхождения образца и всей истории его исследований, что критически важно для научной достоверности.

Интеграция превращает музеи из хранилищ редкостей в активные узлы глобальной исследовательской сети. Она позволяет не просто изучать отдельные образцы, но и анализировать целые геологические системы в пространстве и времени, отвечая на фундаментальные вопросы о прошлом и будущем Земли. Это качественно новый уровень геологических исследований, стирающий границы между институтами, музеями и дисциплинами.

В настоящее время идет активное формирование подобных цифровых экосистем, в которых научные данные и музейные коллекции и становятся частью единого, доступного и интерактивного пространства.

Единой целостной платформы «всё в одном» для интеграции геологических научных данных и музейных коллекций, их обработки и комплексного анализа пока не существует, но есть набор взаимосвязанных технологий, стандартов и инициатив, которые вместе образуют такую среду.

Конкретные примеры и реализуемые проекты

1. Specify 7 (с расширением GeoSpecify) – специализированное решение для геологии. Это система управления коллекциями (CMS) с открытым исходным кодом, предназначенная для геологических музеев, университетов и исследовательских групп, которым нужна глубокая работа со стратиграфией и возрастами. Это, пожалуй, самый яркий пример платформы, которая целенаправленно развивается для нужд геологии. В 2024 г. при поддержке Швейцарской академии наук создано расширение GeoSpecify, которое добавило критически важные функции:

- *работа с разными типами объектов (MOTs)*: в одной коллекции могут сосуществовать минералы, горные породы, метеориты и керны. Для каждого типа имеются свои форма ввода данных и система нумерации, что решает проблему «универсальных» форм, перегруженных ненужными полями;
- *гибкие системы классификации*: в биологии – одно древо видов, а в геологии у минералов, горных пород и метеоритов – совершенно разные классификации и справочники. GeoSpecify позволяет использовать несколько систем именования внутри одной базы данных;
- *составные объекты (COGs)*: очень важно для геологии, например, когда имеется образец породы с несколькими вкраплениями минералов. Система позволяет создать запись для всего образца и отдельные связанные записи для каждого минерала внутри него;
- *стратиграфия и абсолютный возраст*: платформа поддерживает детальное описание стратиграфии (лито-, био-, магнитостратиграфия) и, что важно,

позволяет вносить данные об абсолютном возрасте (с учетом метода датирования, погрешности и диапазона), а не только об относительном.

2. Arctos – мощная экосистема для интеграции наук. Это масштабируемая облачная CMS с открытым исходным кодом, предназначенная для крупных исследовательских консорциумов и музеев, объединяющих биологические, геологические и культурные коллекции. Arctos является примером «зрелой» платформы, которая уже десятилетия успешно работает с гетерогенными данными.

- *Универсальность:* в Arctos можно управлять не только геологическими образцами (минералы, палеонтология), но и генетическими ресурсами, полевыми дневниками, архивными документами, живыми организмами и даже археологическими артефактами.

- *Концепция «расширенного образца» (Digital Extended Specimen):* это ключевая философия Arctos. Платформа связывает сам образец (камень/окаменелость) со всеми ассоциированными данными: полевыми заметками, фотографиями, генетическими последовательностями, публикациями. Создается «паутина знаний», а не просто каталог.

- *Сообщество:* Arctos развивается большим сообществом профессионалов, что гарантирует его устойчивость и соответствие современным стандартам (включая принципы FAIR для научных данных).

3. MuseumPlus + GIS. Это коммерческая система управления коллекциями (разработчиком является zetcom). Предназначена для музеев, где важен мощный пространственный анализ (археология, геология, антропология). Эта платформа интересна своим подходом к геоданным.

В отличие от других систем MuseumPlus имеет полноценный ГИС-модуль (Add-on): вы можете не просто ввести GPS-координаты в текстовое поле, но и работать с интерактивной картой, рисовать полигоны (ареалы) и отмечать точные места находок.

Визуальный анализ позволяет искать и анализировать объекты на карте. Например, можно выделить область на карте и получить все геологические образцы, найденные в этом районе. Это очень востребовано для анализа распространения минералов или источников сырья в древности.

4. DiSSCo (Distributed System of Scientific Collections). Это крупнейшая европейская инициатива по объединению естественно-научных коллекций. В ее рамках разрабатываются цифровые расширения, включая сервисы для аналитики на основе искусственного интеллекта (ИИ) (например, автоматическое распознавание видов по изображениям).

5. Федеральная платформа «Научное наследие России». В ее геологическом блоке заложены принципы интеграции данных из музеев РАН, университетов и научных институтов с аналитическими инструментами.

В последние годы в Государственном геологическом музее им. В. И. Вернадского (ГГМ) РАН велась разработка двух Порталов: для поддержки научных геологических исследований (<https://geologyscience.ru>) и Открытых данных ГГМ РАН (<http://data.sgm.ru>). Первый осуществляет распределенную интеграцию разных типов геологических данных по территории РФ, второй – интегрирует музейные образцы ГГМ РАН и мировых геологических музеев (с территории РФ). Портал <https://geologyscience.ru> содержит сервисы для аналитики, в том числе на основе ИИ.

Сотрудниками Научного отдела ГГМ РАН впервые в мировой практике естественно-научных музеев был осуществлен комплексный переход от теоретических концепций к действующим программным прототипам в пяти ключевых направлениях на стыке геологии и искусственного интеллекта.

1. Разработана основная концепция работ, которая заключается в разработке и построении интегрированной цифровой экосистемы, где научные данные, геологические коллекции и экспертные заключения естественно-научных музеев становятся частью единого, доступного и интерактивного пространства. Это не просто оцифровка экспонатов, а построение связной среды для науки, образования и популяризации. Интеграция научных данных и геологических коллекций музеев открывает огромные возможности для анализа геологических объектов, создавая мощный синергетический эффект.

В последнее время на основе анализа существующих мировых решений, современных ИТ-технологий и наших предыдущих работ выполнено проектиро-

вание Единого цифрового пространства геологических знаний, в которое интегрируются научные данные, геологические коллекции и экспертиза естественно-научных музеев.

2. Сравнительный анализ современных архитектур позволил нам выделить модель DeepSeek R1, относящуюся к классу систем с расширенными возможностями логического вывода. На базе выбранной модели разработан веб-сервис [3–5], предоставляющий уникальный функционал, способный выполнять сравнительный анализ до 5 научных статей стандартного объема; поддержку мультязычных источников (ввод текстов на английском, китайском, русском и других языках); формирование структурированных отчетов на русском языке с выделением ключевых тезисов, противоречий и паттернов. Предложенный подход был протестирован для сравнительного анализа геологических публикаций.

3. Создан прототип цифрового ассистента геолога-исследователя [6–8], разработанного для поддержки научных исследований в геологии. Система позволяет структурировать процесс исследования, обеспечивая пользователя интеллектуальным интерфейсом взаимодействия с нейронной сетью, а также доступом к научной информации, геологическим данным и цифровым инструментам обработки и анализа. Дальнейшее развитие проекта связано с расширением протоколов взаимодействия между агентскими системами, что позволит увеличить функциональность и автономность цифрового ассистента, а также возможность непрерывной адаптации используемых ИИ-моделей к актуальным научным задачам.

4. Исследования современного ландшафта технологий и подходов работы со спутниковыми данными, в том числе форматов хранения данных и моделей ИИ, выявили предпосылки для создания распределенной системы работы со спутниковыми данными для решения на естественном языке научных и прикладных задач в геологии. В результате была предложена концептуальная схема такой системы, интегрирующая рассмотренные технологии, а также разработан ее прототип, реализующий базовый функционал поиска и визуализации спутниковых данных от облачных провайдеров.

5. В результате работ, выполненных ранее, успешно решена главная задача – осуществлен переход от теоретической концепции и дизайн-проекта к работающему программному прототипу образовательной Метавселенной «История Земли». Реализована функциональная альфа-версия интерактивной среды на базе движка Unity, размещенная в открытом онлайн-доступе (формат WebGL). Внедрена и апробирована оригинальная система навигации, основанная на визуальной концепции Роберта Хейзена, позволяющая пользователю интуитивно выбирать геологическую эпоху для изучения. Создана и интегрирована тестовая интерактивная сцена «Красная Земля», подтверждающая возможность построения иммерсивных пространств, моделирующих прошлое планеты. Сформирована техническая и архитектурная основа для дальнейшего развития проекта: добавления новых эпох, интеграции базы знаний портала «История Земли», внедрения ИИ-агентов в качестве гидов и образовательных игровых механик.

Основная цель развития этих работ состоит в разработке и адаптации методов и технологий для создания Интегрированной цифровой научно-образовательной платформы «Интеллектуальная сфера геологических знаний», т. е. комплексной системы, которая собирает информацию о геологии Земли, изучает, анализирует, выявляет закономерности и генерирует новое понимание на основе интеграции коллекций, данных и экспертизы с помощью искусственного интеллекта.

Основная задача заключается в интеграции успешно апробированных ранее разрозненных инструментов в единую сквозную цифровую экосистему: платформу, в которой этапы научного цикла (от анализа литературы и спутниковых данных к выдвижению гипотез и работе с коллекциями, до визуализации и популяризации знаний) поддерживаются сквозными ИИ-сервисами.

На основании анализа существующих мировых исследований и технологий, а также ранее полученных в ГГМ РАН решений сформулированы ключевые направления работ, в которых необходимо проводить исследования:

- 1) от интеллектуальной экосистемы к цифровой интеллектуальной сфере геологических знаний;
- 2) от цифрового каталога к интеллектуальной системе знаний;

- 3) от обработки текстов к генерации научных гипотез;
- 4) от анализа спутниковых данных к автономному прогнозированию;
- 5) от Метавселенной для популяризации к инструменту визуальной науки.

Сформулированы следующие конкретные задачи для разработки Единой цифровой интеллектуальной сферы геологических знаний.

1. От прототипов к промышленной платформе

Для Единого цифрового пространства:

- создание ядра интеллектуальной экосистемы, включающего механизмы управления ИИ-агентами и потоками данных, интеграции мультимодальных ИИ-моделей, генерации сценариев работы и интерфейсов пользователя под задачи научных исследований в геологии;

- создание открытых API для интеграции цифровых каталогов других музеев и научных организаций с семантическим ядром на базе LLM (DeepSeek R1).

Для цифрового ассистента:

Развитие в полноценную мультиагентную научную среду, где специализированные ИИ-агенты (по минералогии, палеонтологии, геодинамике), использующие общую базу знаний и данные коллекций, могут интегрироваться для решения комплексных задач.

Для системы работы со спутниковыми данными:

Реализация полного цикла – от запроса на естественном языке до автоматической обработки снимков предобученными нейросетевыми моделями и генерации аналитической справки.

2. Создание синергетического контура «знание – образование»

- «История Земли» становится не только образовательным, но и научным инструментом. Ученые могут загружать в симулятор свои модели (палеорекострукции, модели рудообразования) и визуализировать их в интерактивном 4D-пространстве (3D + время), проводить виртуальные эксперименты. Это новое качество интерпретации данных и междисциплинарной коллаборации.

- Прямая интеграция Метавселенной «История Земли» с базой знаний и инструментами анализа. Пример сценария: пользователь (ученый или студент) через ассистента находит в литературе гипотезу о климате девона, запрашивает

визуализацию и погружается в соответствующую интерактивную сцену в Метавселенной, «населенную» реконструированной флорой и фауной на основе данных коллекций ГГМ РАН.

ПРОЕКТИРОВАНИЕ ИНТЕГРИРОВАННОЙ ЦИФРОВОЙ НАУЧНО-ОБРАЗОВАТЕЛЬНОЙ ПЛАТФОРМЫ «ИНТЕЛЛЕКТУАЛЬНАЯ СФЕРА ГЕОЛОГИЧЕСКИХ ЗНАНИЙ»

На этапе проектирования этой цифровой платформы рассмотрены концептуальная постановка, а также архитектура подобной платформы, которая может быть применима не только для геологических систем, но и для систем в других тематических направлениях наук о Земле.

Далее приведен список базовых элементов архитектуры предлагаемой платформы:

1. Система накопленных знаний и теорий

- Ядро: фундаментальные геологические дисциплины (минералогия, петрология, стратиграфия, тектоника и т. д.).
- Теоретическая база: основополагающие концепции (тектоника плит, актуализм, цикличность геологических процессов).
- Модели: цифровые, концептуальные и физические модели месторождений, бассейнов, геодинамических процессов.

2. Методология и инструменты познания

- Полевые и лабораторные методы: от геологического молотка и картографии до масс-спектрометрии и микроскопии.
- Вычислительные и цифровые технологии: ГИС (географические информационные системы), 3D-моделирование, big data анализ, машинное обучение для интерпретации данных.
- Методы интерпретации: умение «читать» геологическую историю по разрезам, керну, геофизическим данным.

3. Динамика и развитие

- Самообогащение: новые данные (например, с дистанционного зондирования или глубокого бурения) постоянно проверяют и развивают теории.

- Конвергенция с другими науками: тесная связь с биологией (палеонтология), химией, физикой, экологией, информатикой.
- Эволюция парадигм: смена глобальных концепций (например, переход от фиксизма к мобилизму – тектонике плит).

Таким образом, мы разрабатываем интеллектуальную сферу геологических знаний живой, развивающаяся системой, которая включает:

- содержание (накопленные факты, теории, модели);
- методы (способы получения и проверки новых знаний);
- сообщество (носителей и передатчиков этих знаний);
- цель (понимание строения и истории Земли для решения практических задач человечества).

Итак, представлен цикл «наблюдение → интерпретация → теория → прогноз → новое наблюдение», поддерживаемый профессиональным сообществом и технологическим прогрессом. Это не просто «книжная» наука, а целая инфраструктура для генерации, хранения, применения и трансляции критически важных знаний о нашей планете.

Цифровая интеллектуальная сфера геологических знаний (ЦИСГЗ) становится новой парадигмой и технологической инфраструктуры, которая трансформирует традиционную геологию. Ее можно определить следующим образом.

ЦИСГЗ – это целостная, динамическая и интерактивная среда, в которой все элементы геологических знаний (данные, модели, методы, экспертиза) существуют, создаются, интегрируются, анализируются и передаются в цифровой форме, обеспечивая качественно новый уровень понимания недр и принятия решений. Это «цифровой двойник» геологической реальности и геологического мышления.

Перечислим ключевые компоненты и характеристики ЦИСГЗ.

1. Единое цифровое пространство данных (*Digital Data Fabric*):

- интеграция разрозненных данных: объединение в единые платформы данных дистанционного зондирования (ДЗЗ), геофизики, геохимии, керна, скважин, полевых наблюдений;

- семантические онтологии и стандарты: системы понятий и связей (например, GeoSciML), которые, например, позволяют машинам «понимать», что «гранит» – это «интрузивная магматическая порода», а не просто слово;

- большие геоданные (GeoBigData): обработка петабайтов сейсмических данных, спутниковых снимков, результатов моделирования.

2. Сквозные цифровые рабочие потоки (*Digital Workflows*):

- автоматизация рутинных операций;
- воспроизводимость исследований: любой анализ или модель можно повторить, проверить и развить, так как все этапы будут зафиксированы цифровым следом;

- коллаборативные платформы: возможность для распределенных команд геологов, геофизиков и инженеров работать в одной цифровой среде с одними данными в реальном времени.

3. Цифровые модели и симуляция как ядро знаний:

- 3D/4D геологические модели: не просто карты, а объемные динамические модели геологических объектов;

- цифровые двойники (*Digital Twins*) объектов недр: виртуальные копии геологических объектов, непрерывно обновляемые новыми данными и позволяющие проводить «что... если» сценарии;

- предиктивное моделирование и машинное обучение: алгоритмы, предсказывающие вероятность обнаружения рудных тел, свойства коллектора в межскважинном пространстве, риск обрушения пород;

- искусственный интеллект как «коллега-геоаналитик»;

- генерация гипотез: ИИ анализирует огромные массивы данных и находит скрытые закономерности, предлагая геологу новые идеи для интерпретации;

- усиление экспертизы: не замена геолога, а инструмент, расширяющий его когнитивные способности;

- открытость, доступность и инфраструктура;

- национальные/корпоративные геопорталы: единые точки доступа к данным (например, геологические службы, предоставляющие данные в виде веб-сервисов);

- облачные платформы (Cloud Computing): мощности для хранения и расчетов (например, суррогатное моделирование фильтрационных задач);
- принципы FAIR: данные должны быть Findable (обнаруживаемые), Accessible (доступные), Interoperable (интероперабельные), Reusable (повторно используемые).

ЦИСГЗ – это экосистема, где данные превращаются в информацию, информация – в цифровые модели и симуляции, а те – в прогнозное знание и оптимальные решения, замыкая цикл через сбор новых данных. Это новый способ геологического мышления и деятельности, возможный только в цифровую эпоху. Главная цель построения ЦИСГЗ состоит в том, чтобы сделать недра «прозрачными», понятными и управляемыми в цифровом пространстве до и вместо физического вмешательства.

Предполагается, что ЦИСГЗ будет иметь сложную, многослойную архитектуру. На рис. 1 представлены возможные ключевые компоненты.



Рис. 1. Ключевые слои цифровой интеллектуальной сферы геологических знаний.

Опишем основные слои ЦИСГЗ.

1. Инфраструктура и стандарты

- технологическая платформа – облачная (IaaS, PaaS) или гибридная инфраструктура для хранения и вычислений;

- стандарты данных: геологические онтологии и словари (GeoSciML, RESQML, PRODML, EarthChem);
- метаданные: протоколы описания происхождения, качества и методов получения данных;
- правила доступа, использования и коммерциализации цифровых активов.

2. Единое цифровое пространство данных (Data Space/Fabric)

- унифицированные хранилища и базы данных:
 - реляционные БД: каталоги образцов, скважин, месторождений;
 - NoSQL БД и Data Lakes: для неструктурированных и слабоструктурированных данных (сейсмические кубы, 3D-модели, спутниковые снимки, PDF-публикации);
 - системы управления пространственными данными: платформы для работы с геопривязанной информацией;
- системы сбора данных (IoT).

3. Ядро знаний – модели и симуляции

- платформы для построения и хранения моделей:
 - статические (3D) геологические модели (литология, структура, стратиграфия);
 - динамические (4D) гидродинамические и фильтрационные модели;
 - геомеханические и геохимические модели;
- цифровые двойники конкретных объектов (месторождение, карьер, участок недр);
- вычислительные модули и симуляторы: пакеты для математического моделирования процессов.

4. Интеллектуальный слой (AI/ML & Analytics Layer)

- библиотеки алгоритмов и инструментов:
 - машинное обучение / глубокое обучение: для распознавания образов (сейсмические фации, минералы на снимках керна), прогнозирования свойств, кластеризации;
 - аналитика больших данных (Big Data Analytics): для выявления пространственно-временных закономерностей;

- платформы для геостатистики и продвинутой визуализации;
- цифровые помощники и чат-боты: системы, способные на естественном языке отвечать на запросы типа «покажи все скважины, вскрывшие пласт Ю2, с содержанием золота выше 1 г/т».

5. Интерфейсы и сервисы

- веб-порталы и геосервисы: удобные интерфейсы для доступа к данным, картам и моделям через браузер;
- сервисы Application Programming Interface (API): позволяют программам автоматически «общаться» друг с другом и запрашивать данные/расчеты;
- специализированное профессиональное ПО: цифровые рабочие места (Digital Workspaces) геолога, геофизика, инженера-нефтяника, интегрированные в общую систему;
- системы поддержки принятия решений (DSS): панели индикаторов (Dashboards), которые агрегируют ключевые показатели и прогнозы для менеджеров.

6. Системное управление

- управление жизненным циклом данных (Data Governance): политики, роли и процессы, обеспечивающие качество, безопасность и актуальность данных;
- системы кибербезопасности: защита критически важной цифровой интеллектуальной собственности;
- учебные модули и симуляторы: для обучения новых специалистов работе внутри ЦИСГЗ.

Таким образом, ЦИСГЗ – это не просто один программный продукт, а взаимосвязанный комплекс технологий, стандартов, людей и процессов, образующих целостную экосистему для геологической деятельности нового типа.

В заключение рассмотрим современное развитие интеллектуальных платформ будущего, которые генерируют новые геологические знания. Интеграция подобной платформы в ЦИСГЗ является следующим шагом развития экосистемы.

ИНТЕЛЛЕКТУАЛЬНЫЕ ПЛАТФОРМЫ, КОТОРЫЕ ГЕНЕРИРУЮТ НОВОЕ ЗНАНИЕ НА СТЫКЕ ГЕОЛОГИЧЕСКИХ ДАННЫХ, МУЗЕЙНЫХ КОЛЛЕКЦИЙ И ЭКСПЕРТНЫХ ОЦЕНОК

Такие системы, как правило, строят на архитектуре графов Знаний (Knowledge Graphs) и используют различные техники ИИ, чтобы находить скрытые связи и закономерности. Перечислим их.

1. KnowWhereGraph – междисциплинарный граф знаний о пространстве и месте. Это масштабный гео-граф знаний, который объединяет данные о природных и социальных системах в единую, предварительно интегрированную структуру. Он связывает данные по ключевым измерениям – пространству, месту и времени, что делает данные FAIR-совместимыми и готовыми для ИИ. Платформа позволяет задавать сложные междисциплинарные вопросы, которые были бы невозможны при работе с изолированными базами данных. Например, можно искать корреляции между геологическим составом породы (из музейной коллекции), историей землепользования (из спутниковых данных) и климатическими изменениями (из архивов наблюдений). Хотя платформа охватывает и другие сферы, ее ядро является идеальным решением для интеграции «сухих» геологических данных с контекстной информацией, что критически важно для понимания закономерностей.

2. StratiGraph – ИИ и графы знаний для стратиграфии. Этот проект нацелен на сбор данных, обработку, анализ и генерацию нового понимания.

Система StratiGraph создана для работы с последовательностью напластований, что напрямую связано с музейными коллекциями (образцы пород и окаменелости привязаны к конкретным слоям).

Для сбора данных: используются ИИ-модели, которые работают прямо на полевых устройствах. Они помогают распознавать изображения, классифицировать стратиграфические единицы и даже проводить первичный анализ артефактов *на месте находки*, что решает проблему чистоты данных.

Создается стратиграфический граф знаний. Он динамически связывает пространственные данные (3D-модели раскопов), временные последовательности и интерпретационные слои (мнения экспертов).

С помощью ИИ автоматизируется создание аннотированных библиографий и проводится подготовка данных к публикации, что радикально сокращает время от открытия до научного вывода.

3. GeoGPT + LightRAG – специализированные языковые модели для построения графа знаний для геологического прогнозирования на основе текстовых данных (научных статей, отчетов и др.). Используется модель GeoGPT [9], которая обучена на специализированной геологической литературе. Эксперименты показали, что она распознает названия минералов и месторождений на 17–28% лучше, чем общие модели типа DeepSeek-V3.

С применением указанных языковых моделей генерируются знания:

- GeoGPT автоматически извлекает сущности (минералы, породы, месторождения, разломы) и связи между ними из миллионов текстов;
- на основе этого строится граф знаний геологоразведки;
- далее используется механизм LightRAG, который позволяет модели отвечать на сложные вопросы, получая релевантный контекст из этого графа, а не просто «фантазируя» на основе общих данных.

Идеально подходит для задач, где нужно на основе гигантского объема опубликованных знаний выявить новые перспективные зоны для поиска полезных ископаемых.

ЗАКЛЮЧЕНИЕ

На основании анализа существующих мировых исследований и технологий, а также решений, ранее полученных в ГГМ РАН, разрабатывается Единая сквозная цифровая экосистема: платформа, в которой этапы научного цикла (от анализа литературы и спутниковых данных к выдвижению гипотез и работе с коллекциями до визуализации и популяризации знаний) поддерживаются сквозными ИИ-сервисами.

Основная концепция работы заключается в создании интегрированной цифровой экосистемы, где геологические коллекции, научные данные и экспертиза естественно-научных музеев становятся частью единого, доступного и интерактивного пространства. Это не просто оцифровка экспонатов, а построение связной среды для науки, образования и популяризации.

На этапе проектирования предложена концептуальная постановка и разработана архитектура платформы, которая может быть применима не только

для геологических систем, но и для систем в других тематических направлениях наук о Земле. Также обсуждены направления автоматического получения геологических знаний и разработки цифровых геологических баз знаний для использования в ЦИСГЗ.

Благодарности

Работа выполняется в рамках Государственной темы Государственного геологического музея им. В. И. Вернадского РАН (FMMR-2025-0001), рег. № 1022061000107-4-1.5.1 «Естественнонаучные музеи в едином цифровом пространстве геологических знаний».

СПИСОК ЛИТЕРАТУРЫ

1. Наумова В.В., Патук М.И., Еременко А.С., Загуменнов А.А., Еременко В.С. Цифровые технологии будущего для научных исследований в геологии // Электронные библиотеки. 2025. Том 28, № 4. С. 788–805.

<https://doi.org/10.26907/1562-5419-2025-28-4-788-805>

2. Naumova V.V., Eremenko A.S., Zagumenov A.A., Eremenko V.S. Digital Technologies of the Future for Scientific Research in Geology // Automatic Documentation and Mathematical Linguistics. 2025. Vol. 59, Suppl. 4. P. S298–S303.

<https://doi.org/10.3103/S0005105525701195>

3. Mikhail I. Patuk, Vera V. Naumova. Using Semantic Search to Select and Rank Geological Publications // Automatic Documentation and Mathematical Linguistics. 2024. Vol. 58, Suppl. 5. P. S292–S296.

<https://doi.org/10.3103/S0005105525700372>

4. Патук М.И., Наумова В.В. Анализ текстов геологических публикаций с использованием больших языковых моделей // Электронные библиотеки. 2025. Т. 28, № 4. С. 806–821. <https://doi.org/10.26907/1562-5419-2025-28-4-806-821>

5. Patuk M.I., Naumova V.V. Comparative Analysis of Geological Texts Using Large Language Model // Automatic Documentation and Mathematical Linguistics. 2025. Vol. 59, Suppl. 4. P. S304–S309. <https://doi.org/10.3103/S0005105525701201>

6. Еременко В.С., Наумова В.В. Виртуальный ассистент геолога-исследователя // Электронные библиотеки. 2025. Т. 28, № 4. С. 781–787.

<https://doi.org/10.26907/1562-5419-2025-28-4-781-787>

7. Eremenko V.S., Naumova V.V. Digital Assistant for Geologist-Researchers// Automatic Documentation and Mathematical Linguistics. 2025. Vol. 59, Suppl. 4. P. 781–787. <https://doi.org/10.3103/S0005105525701183>

8. Naumova V.V., Zagumennov A.A., Eremenko V.S. Digital Tools and Virtual Assistants to Support Scientific Research in Geology // Automatic Documentation and Mathematical Linguistics. 2024. Vol. 58, Suppl. 5. P. S287–S291. <https://doi.org/10.3103/S0005105525700360>

9. Yifan Zhang, Cheng Wei, Zhengting He, Wenhao Yu. GeoGPT: An assistant for understanding and processing geospatial tasks// International Journal of Applied Earth Observation and Geoinformation. 2024. Vol. 131. Art. 103976. <https://doi.org/10.1080/15230406.2024.2404868>

EVELOPMENT OF A DIGITAL INTELLIGENT GEOLOGICAL KNOWLEDGE SPHERE

Vera V. Naumova ^[0000-0002-3001-1638]

State Geological Museum named after Vladimir Vernadsky of RAS, Moscow, Russia
naumova_new@mail.ru

Abstract

Based on the analysis of future digital technologies and modern solutions for supporting scientific research in geology, the authors of the State Geological Museum named after Vladimir Vernadsky of RAS team have developed and proposed an architecture for a digital intelligent framework of geological knowledge. The main concept of the work is to create an integrated digital ecosystem where geological collections, scientific data, and the expertise of natural science museums become part of a single, accessible, and interactive space. This is not just about digitizing exhibits, but about building a connected environment for science, education, and outreach.

Keywords: *unified digital space for geological knowledge, digital geological environment for the future, geological collections, scientific data, expertise from natural science museums, digital tools, AI assistants for geology.*

REFERENCES

1. *Naumova V.V., Patuk M.I., Eremenko A.S., Zagumennov A.A., Erjomenko V.S.* Cifrovye tehnologii budushhego dlja nauchnyh issledovanij v geologii // Russian Digital Libraries Journal. 2025. Vol. 28, No. 4. P. 788–805 [In Russ.]
<https://doi.org/10.26907/1562-5419-2025-28-4-788-805>
2. *Naumova V.V., Eremenko A.S., Zagumennov A.A., Eremenko V.S.* Digital Technologies of the Future for Scientific Research in Geology // Automatic Documentation and Mathematical Linguistics. 2025. Vol. 59, Suppl. 4. P. S298–S303.
<https://doi.org/10.3103/S0005105525701195>
3. *Mikhail I. Patuk, Vera V. Naumova.* Using Semantic Search to Select and Rank Geological Publications // Automatic Documentation and Mathematical Linguistics. 2024. Vol. 58, Suppl. 5. P. S292–S296.
<https://doi.org/10.3103/S0005105525700372>
4. *Patuk M.I., Naumova V.V.* Analiz tekstov geologicheskikh publikacij s ispol'zovaniem bol'shih jazykovyh modelej // Russian Digital Libraries Journal. 2025. Vol. 28, No. 4. P. 806–821 [In Russ.]
<https://doi.org/10.26907/1562-5419-2025-28-4-806-821>
5. *Patuk M.I., Naumova V.V.* Comparative Analysis of Geological Texts Using Large Language Model // Automatic Documentation and Mathematical Linguistics. 2025. Vol. 59, Suppl. 4. P. S304–S309.
doi: <https://doi.org/10.3103/S0005105525701201>
6. *Eremenko V.S., Naumova V.V.* Virtual'nyj assistent geologa-issledovatelja // Russian Digital Libraries Journal. 2025. Vol. 28, No. 4. P. 781–787 [In Russ.].
<https://doi.org/10.26907/1562-5419-2025-28-4-781-787>
7. *Eremenko V.S., Naumova V.V.* Digital Assistant for Geologist-Researchers // Automatic Documentation and Mathematical Linguistics. 2025. Vol. 59, Suppl. 4. P. 781–787. <https://doi.org/10.3103/S0005105525701183>
8. *Naumova V.V., Zagumennov A.A., Eremenko V.S.* Digital Tools and Virtual Assistants to Support Scientific Research in Geology // Automatic Documentation and Mathematical Linguistics. 2024. Vol. 58, Suppl. 5. P. S287–S291.
<https://doi.org/10.3103/S0005105525700360>

9. Yifan Zhang, Cheng Wei, Zhengting He, Wenhao Yu. GeoGPT: An assistant for understanding and processing geospatial tasks// International Journal of Applied Earth Observation and Geoinformation. 2024. Vol. 131. Art. 103976. <https://doi.org/10.1080/15230406.2024.2404868>

СВЕДЕНИЯ ОБ АВТОРЕ



НАУМОВА Вера Викторовна – д. г.-м. н., г. н. с.,
зав. Научным отделом Государственного геологического
музея им. В. И. Вернадского РАН, Москва, Россия.

Vera V. NAUMOVA – Prof., head of SGM scientific de-
partment, *State Geological Museum named after Vladimir Ver-*
nadsky of RAS, Moscow, Russia.

email: naumova_new@mail.ru

ORCID: 0000-0002-3001-1638

Материал поступил в редакцию 28 апреля 2026 года