

Использование мультимодальных эмбедингов как эффективного способа обработки климатической информации в рамках целевых задач Климатической доктрины Российской Федерации

О. Ю. ГАВЕНКО^{1,2,*}, Н. А. ШАШОК²

¹Институт динамики систем и теории управления им. В.М. Матросова СО РАН, 664033, Иркутск, Россия

²Федеральный исследовательский центр информационных и вычислительных технологий, 630090, Новосибирск, Россия

*Контактный автор: Гавенко Ольга Юрьевна, e-mail: olga.yu.gavenko@mail.ru

Поступила 20 мая 2025 г., доработана 20 июня 2025 г., принята в печать 01 июля 2025 г.

Использование передовых технологий позволяет обеспечить наукоемкую поддержку принятия решений широкого круга проблем, связанных с антропогенными и техногенными факторами. Автоматизация такой поддержки требует проведения фундаментальных и прикладных научных исследований, направленных на разработку аналитических и управляющих систем, способных корректно обрабатывать и интерпретировать климатические данные. Эффективным способом обработки климатической информации является использование мультимодальных эмбедингов, объединяющих информацию из разнородных источников. Разработка информационной системы, в которой применяются мультимодальные климатические данные, соответствует стратегическим целям Российской Федерации в области экологического мониторинга и климатической политики РФ.

Ключевые слова: поддержка принятия решений, разработка аналитических систем, климатическая информация, мультимодальные эмбединги.

Цитирование: Гавенко О.Ю., Шашок Н.А. Использование мультимодальных эмбедингов как эффективного способа обработки климатической информации в рамках целевых задач Климатической доктрины Российской Федерации. Вычислительные технологии. 2025; 30(4):133–144. DOI:10.25743/ICT.2025.30.4.012.

Введение

Климатическая доктрина Российской Федерации, утвержденная Указом Президента РФ от 26 октября 2023 г., является системой целей, принципов, задач и механизмов реализации государственной климатической политики РФ — круга вопросов, связанных с изменениями климата и последствиями этих изменений [1]. Основа Доктрины — фундаментальные и прикладные научные знания в области климата и смежных областей, при этом научное обоснование Доктрины включает в себя признание влияния человека на климат. Таким образом, антропогенная и техногенная составляющие процессов

изменения климата признаны объектами научных исследований климатических изменений. В документе подчеркивается определяющая роль политического процесса, что обоснованно для государственных документов такого уровня, определено, что проблемы “не могут быть решены посредством только научных методов” [1], однако признано, что факторы изменений климата, такие как глобальное потепление (в том числе связанное с выбросами парниковых газов), климатические аномалии, экстремальные погодные явления, их неизбежное влияние на жизнь и их угроза устойчивому развитию и благополучию, предопределяют необходимость развития климатического мониторинга и проведения всесторонних научных исследований в области климата. В Доктрине подчеркивается необходимость учета изменения климата, что определено одним из ключевых долговременных факторов безопасности РФ. Подход к этому вопросу, как указано в Доктрине, должен быть основан “на всестороннем научном исследовании экологических, экономических и социальных факторов” [1]. Ключевой фактор изменений климата — антропогенные выбросы парниковых газов — в рамках обновленной климатической Доктрины определен как основная проблема, решение которой состоит в переходе к низкоуглеродной экономике. Этот процесс в настоящее время наблюдается в различных областях человеческой деятельности, и наукоемкие технологии в поддержке принятия решения широкого круга соответствующих задач приобретают приоритетное значение.

Научная обоснованность принимаемых мер в области климата — один из основных принципов климатической политики РФ, что подразумевает, в том числе, использование самых передовых технологий. В соответствии с концепцией Доктрины на передний план выходит наукоемкая поддержка принятия решений, связанных именно с антропо- и техногенными факторами, при этом приоритетом исследований в области изменений климата определено снижение уровня неопределенности оценок изменений и их последствий. В Доктрине подчеркивается, что климатическая политика подлежит регулярной и своевременной корректировке с учетом новых знаний, а также технологического развития, что, естественно, подразумевает наукоемкие технологии, в том числе информационные: первым пунктом основных задач обозначено “развитие информационной и научной основы климатической политики, включая усиление научно-технического и технологического потенциала РФ, для обеспечения полноты и достоверности информации о состоянии климата, антропогенном и ином воздействии на климат, его происхождении и будущем изменении и последствиях такого изменения” [1].

1. Информационная система для обработки и анализа мультимодальных климатических данных

Согласно Доктрине, научная обоснованность климатической политики должна базироваться на пополнении знаний о климате. Этот процесс, в свою очередь, связан с систематическими наблюдениями, мониторингом, обработкой данных для получения адекватной оценки проблем и, соответственно, принятием решений, а также с фундаментальными и прикладными научными исследованиями для разработки необходимых анализирующих и управляющих систем. Упреждающая адаптация к изменениям климата — приоритет климатической политики РФ, что не может быть реализовано без наукоемких технологий, позволяющих проектировать и реализовывать прогрессивные информационные системы, способные хранить, обрабатывать и использовать данные для

поддержки принятия решений в задачах, связанных с климатическими изменениями и упреждающими адаптациями. Научное и информационное обеспечение исследований, связанных с мерами по адаптации и смягчению антропогенного влияния на климат, — одно из основных направлений реализации климатической политики. Как указано в Доктрине, основная задача научного обеспечения — предоставление “достоверной научной информации для принятия соответствующих решений” [1]. Подчеркивается, что принципиальное значение для научного обеспечения климатической политики имеют: точность данных для объективной оценки климатических и антропогенных процессов и разработка инструмента в виде информационно-аналитических систем, позволяющих хранить, обрабатывать и анализировать получаемые данные, что, в свою очередь, дает поддержку принятия решений в стабильных и чрезвычайных ситуациях, связанных с антропогенными и техногенными процессами, а также в разработке стратегий адаптации и смягчения антропогенного воздействия на климат.

Таким образом, современные системы обработки информации, включающие с себя самые передовые технологические решения и инструменты, представляются эффективным инструментом для реализации задач, изложенных в Климатической доктрине РФ.

2. Проектирование информационной системы для обработки климатических данных: общие положения

Информационная система, обрабатывающая и анализирующая климатические мультимодальные данные, предназначена для работы с информацией, получаемой как от внешних источников (мониторинговых систем и глобальной сети), так и из внутренних хранилищ данных. Многообразие форматов климатических данных (текстовый, числовой, графический, видео- и аудиоформаты, географические данные, данные мониторинга) и необходимость обработки не только имеющихся объемов данных, но и поступающих в непрерывном режиме, — два фактора, обосновывающих необходимость разработки обязательного элемента системы: инструмента, позволяющего адекватно обрабатывать все виды данных для их дальнейшего использования в системах вопрос – ответ и в системах поддержки принятия решений. В аспекте обработки и анализа разнородных динамических и статических данных разработка и реализация современного инструмента для хранения климатической информации в векторном формате становится оптимальным решением для информационной системы требуемого типа.

Проектирование информационной системы, способной обрабатывать и анализировать климатические данные, представляет собой комплексную задачу, связанную с мониторингом природных и техногенных процессов, а также с оценкой антропогенных рисков. Такая система должна быть способна интегрировать данные из разнородных источников, включая спутниковые снимки, данные датчиков, текстовые отчеты и научные публикации, обеспечивая их семантическую интерпретацию и контекстуальный анализ. Это позволит не только повысить точность прогнозов и оценок, но и создать инструмент для оперативного реагирования на изменения климатической обстановки, что соответствует стратегическим интересам Российской Федерации в области экологической безопасности и устойчивого развития.

Следует отметить, что существующие в обозначенной области системы и наборы данных преимущественно ориентированы на обработку текстовых документов. Однако значительная часть климатических данных, несущих критически важный контекст,

представлена в форматах, выходящих за рамки текстовой информации. К ним относятся данные мониторинга окружающей среды, спутниковые снимки, видеозаписи и пр. В качестве примеров таких данных можно привести экологические словари [2], информацию, предоставляемую Федеральной службой по гидрометеорологии и мониторингу окружающей среды (включая карты, новостные сводки и табличные данные) [3], данные Climate Forecast System Reanalysis (CFSR) [4], а также научные статьи с открытым доступом через такие платформы, как КиберЛенинка [5] и др.

Обработка климатических данных с целью извлечения аналитически значимой информации и поддержки принятия решений, соответствующих потребностям пользователей, представляет собой активно развивающееся направление научных и прикладных исследований как в Российской Федерации, так и в международном научном сообществе. В контексте российской повестки разработка интеллектуальных информационных систем, ориентированных на обработку климатической информации, охватывает широкий спектр междисциплинарных областей, включая экологию, экономику, геоинформатику и управление природными ресурсами.

В частности, можно рассматривать следующие задачи и проекты:

- проектирование системы поддержки принятия решений для классификации экологического состояния окружающей среды [6];
- построение автоматизированной информационно-аналитической системы, позволяющей региональным инвестиционным и экологическим аналитикам вырабатывать рациональные управленческие решения в области оценки эколого-экономической привлекательности угледобывающей отрасли региона [7];
- создание систем мониторинга, предназначенных для прогнозирования и предотвращения загрязнения окружающей среды при размещении промышленных объектов и технических систем [8].

В качестве примера зарубежных интеллектуальных платформ и модели искусственного интеллекта (ИИ) можно привести такие проекты, как Climate X Spectra [9], анализирующий риски для недвижимых активов через интеграцию климатических сценариев и физического моделирования, MIT Climate Models [10], моделирующие различные поведения климата, системы раннего предупреждения [11], а также модели искусственного интеллекта по типу ClimateGPT [12].

Однако перечисленные решения демонстрируют, что применение больших языковых моделей и нейролингвистических методов остается ограниченным. Существующие реализации преимущественно используют классические алгоритмы машинного обучения и экспертные системы. Отдельные языковые модели, такие как ClimateGPT, применяются исключительно без привязки к платформам, а также без использования какой-либо уточняющей информации в дальнейшем, что, соответственно, требует переобучения либо предоставления данных со стороны пользователя. В медиапространстве отмечается растущий интерес к интеграции LLM общего домена, включая модели типа ChatGPT [13], в задачи экологического анализа, поскольку использование этих моделей как в общем домене, так и в узконаправленных, демонстрирует их потенциал для работы с экологическими данными и в задачах поддержки принятия решений по вопросам экологии. Для использования моделей общего домена в узконаправленных задачах также необходимо предоставлять дополнительную информацию, о которой пользователь может не знать. Одним из решений такой задачи может быть разработка ИИ-агента с использованием RAG [14].

3. Архитектурные особенности проектирования информационной системы для обработки мультимодальных климатических данных

Проектирование информационных систем для обработки мультимодальных климатических данных, в частности при разработке интеллектуальных систем на основе ИИ-агентов, сопряжено с архитектурными вызовами, обусловленными спецификой мультимодальности и использования разных типов данных. Такие системы должны эффективно интегрировать, хранить, обрабатывать, анализировать и использовать разнородные данные: спутниковые снимки, картографические данные, показания наземных сенсоров, выходные файлы численных моделей, результаты полевых наблюдений и др. В силу отсутствия фиксированности множества типов данных ключевыми особенностями проектирования подобной системы становятся следующие положения:

- обеспечение масштабируемости для работы с большими объемами данных и вероятно высокой скоростью поступления информации;
- обеспечение гибкости для поддержки разнообразных форматов и структур;
- реализация механизмов предобработки данных, позволяющих согласовывать разные форматы;
- создание пользовательского интерфейса для анализа и визуализации.

Необходимо также учитывать, что данные могут изначально храниться в несвязанных друг с другом хранилищах. Следовательно, возникает задача эффективного поиска информации, зачастую близкой друг к другу больше по смысловому параметру, чем по семантическим параметрам.

3.1. Мультимодальные эмбединги: векторное представление данных как эффективный способ обработки мультимодальной информации

В исследованиях, касающихся технологий обработки информации, одной из ключевых проблем становится эффективное использование мультимодальных эмбедингов — векторных представлений данных, которые объединяют информацию из разнородных источников (текст, изображения, аудио, видео и т. д.), а также интеграция эмбедингов, сгенерированных специализированными моделями векторизации, адаптированными под конкретные форматы данных. Эмбединги, созданные с помощью этих моделей, хотя и демонстрируют высокую точность применения в алгоритмах семантического поиска в рамках своих доменов, зачастую несовместимы между собой, поскольку либо образуют несовместные векторные пространства, либо используют несовместные семантические характеристики, что затрудняет создание универсальных информационных систем, способных обрабатывать гибридные данные. В свою очередь, модели, генерирующие мультимодальные эмбединги, несмотря на объединение нескольких типов данных в рамках одного векторного пространства, ограничены для добавления новых типов данных либо новых источников данных, а также могут испытывать недостаточную глубину анализа характеристик.

Итак, можно выделить две главные проблемы использования как мультимодальных эмбедингов, так и разных эмбедингов совместно, сгенерированных различными моделями:

1. Зависимость от структуры входных данных и данных обучения. Добавление нового типа данных требует перепроектирования архитектуры или трудоемкого дообучения модели, что нарушает масштабируемость.

2. **Согласованность.** Даже при использовании общих метрик близости векторы из разных модальностей могут занимать несогласованные области в векторном пространстве, снижая точность семантического поиска или классификации.

3.2. Использование модульной архитектуры в задаче проектирования систем для семантического поиска по разным типам данных

Для систем, в которых необходимо динамически подстраиваться под новые типы данных без полного переобучения (а это именно системы обработки климатических данных как специфического вида информации), обеспечивать семантическую согласованность эмбедингов, а также сохранять баланс между универсальностью и специализацией для конкретных задач, авторами статьи предлагается модульная архитектура, в которой каждый отдельный источник данных может обрабатываться модулем поиска, независимым от других модулей, выполняющих такую же функцию. Подобная архитектура предполагает использование некоторого общего модуля, являющегося связующим звеном модулей поиска и входной точкой пользовательских запросов к поиску данных в источниках. Такой подход требует использования общего контракта API между модулями, а также некоторых описывающих модули поиска метаданных, позволяющих связующему модулю обращаться к модулям поиска по указанному API. Более того, подобный подход позволяет расширять систему, не используя новую модель для генерации эмбедингов: достаточно разработать новый модуль со своей моделью и включить его в систему. В некотором роде описываемый подход близок к микросервисным архитектурам, преимуществами которых, как правило [15], считаются:

- внедрение независимых обновлений частей, не связанных друг с другом;
- параллельная разработка несколькими командами (в случае предлагаемой модульной архитектуры каждый поисковый модуль может иметь своего разработчика);
- технологическое разнообразие, т. е. использование различных инструментов, языков и библиотек при условии соблюдения общего контракта API.

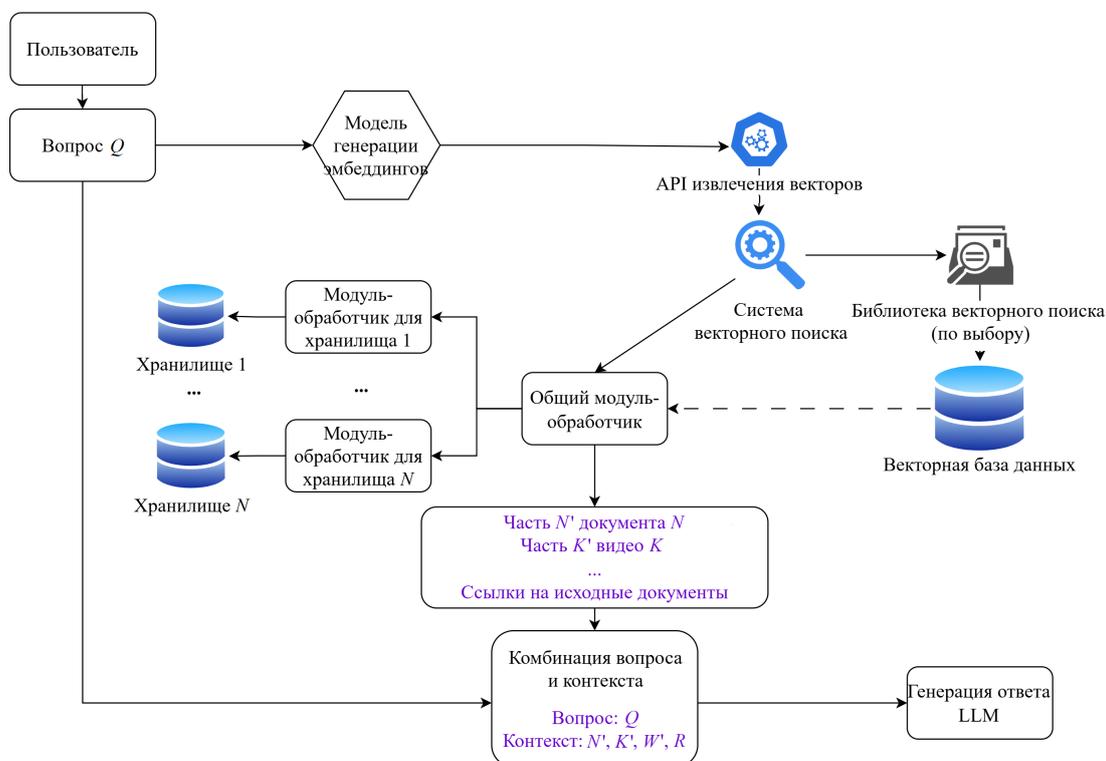
Представляется, что контракт API модулей не должен являться чрезмерно усложненным для облегчения разработки и достаточной может быть следующая функциональность.

1. **Функциональность осуществления запроса.** В данном случае при обращении на поиск требуется передать не только текст запроса, но и количество ближайших запросов, информацию о том, насколько близки к запросу могут быть результаты, а также выполнить некоторую фильтрацию выходных данных. В качестве выходных данных должны быть не только обнаруженные при поиске части документов и данных, но и описание исходного источника данных и полной версии исходного документа, поскольку таким образом пользователь будет иметь возможность осуществить проверку вывода.
2. **Описание модуля.** Информационная функция, возвращающая описание самой себя, включая описание исходного источника данных.
3. **Описание своего “здоровья”.** В рамках данного метода осуществляется проверка, что сам модуль, его исходный источник, а также какие бы то ни были связанные хранилища работают штатно. При поломке любого элемента, с которым взаимодействует модуль, к модулю нет необходимости обращаться, но следует оповещать администратора системы.

Ключевой проблематикой предложенного подхода является высокая ресурсоемкость поддержки множества параллельно функционирующих моделей. Во-первых, даже при использовании хорошо оптимизированных моделей одновременная обработка запросов через несколько модулей с разными моделями требует значительных вычислительных мощностей, особенно в системах с определенными требованиями к задержкам. Во-вторых, могут возникать сложности с балансировкой нагрузки.

В качестве возможного направления улучшения архитектуры авторами статьи предлагается использование адаптивного подхода, предоставляющего пользователю возможности выбора источников или типов данных с помощью связующего модуля, т.е. динамический выбор источников данных на основе некоторого описания искомого хранилища, в качестве которого может выступать хранение метаописаний как в некотором хранилище связующего модуля, так и в конфигурационных файлах запускаемого приложения. В случае хранения метаописаний модулей поиска в специальном хранилище связующего модуля необходимо также иметь возможность добавлять метаописания в хранилища удобным способом с помощью API, доступного некоторой определенной группе пользователей. Второй вариант — хранение метаописаний в конфигурационных файлах — предполагает, что метаописания может добавлять только системный администратор системы, на которой запущено приложение, после чего может возникнуть необходимость перезапуска, как, например, происходит в контейнеризированных приложениях, управляемых с помощью систем оркестрации контейнеров.

В обоих случаях происходит большее расхождение в контрактах модуля-связки и модулей поиска: на вход функциональности осуществления запроса модуля-связки необходимо передавать массив значений, представляющий собой список используемых в запросе хранилищ или типов данных. При отсутствии данных поведение может быть определено как использующее все доступные модули поиска.



Процесс генерации ответа на запрос пользователя
The process of generating a reply to user request

Роль описываемых модулей при использовании системы схематически представлена на рисунке.

Таким образом, использование модульной архитектуры, как описано выше, привносит следующие преимущества в информационную систему, обрабатывающую большие массивы мультимодальных данных:

1. Масштабируемость. Добавление новых модулей не требует изменения основной архитектуры.
2. Гибкость. Пользователи могут динамически выбирать источники данных через связующий модуль.

Предварительная спецификация модулей поиска представляет собой следующий документ в формате спецификации openAPI:

```

openapi: 3.0.0
info:
  title: API для модулей поиска
  version: 1.0.0
  description: API для взаимодействия связующего модуля с модулями поиска.
paths:
  /search:
    post:
      summary: Выполнение поискового запроса.
      requestBody:
        required: true
        content:
          application/json:
            schema:
              type: object
              properties:
                query:
                  type: string
                  description: Текст запроса.
                limit:
                  type: integer
                  description: Количество возвращаемых результатов.
                similarity_threshold:
                  type: number
                  format: float
                  description: Минимальный порог схожести результатов.
                filters:
                  type: object
                  description: Параметры фильтрации (например, указание порядка документов или избавление от определенных документов).
      responses:
        '200':
          description: Успешный ответ.
          content:
            application/json:
              schema:
                type: object
                properties:
                  results:
                    type: array
                    items:
                      type: object
                      properties:
                        snippet:
                          type: object
                          description: Найденная часть документа.
                        source:
                          type: string
                          description: Идентификатор источника данных.
                        document_url:
                          type: string
                          description: Ссылка на полный документ.
  /info:
    get:

```

```
summary: Получение метаописания модуля.
responses:
  '200':
    description: Успешный ответ.
    content:
      application/json:
        schema:
          type: object
          properties:
            name:
              type: string
              description: Название модуля.
            data_source:
              type: string
              description: Описание источника данных.
            version:
              type: string
              description: Версия модуля.

/health:
get:
summary: Проверка доступности модуля.
responses:
  '200':
    description: Модуль работает корректно.
    content:
      application/json:
        schema:
          type: object
          properties:
            status:
              type: string
              enum: [OK, ERROR]
            details:
              type: string
              description: Детали ошибки (если есть).
```

Заключение

Информационная система, обрабатывающая и анализирующая климатические мультимодальные данные, соответствует стратегическим целям Российской Федерации в области экологического мониторинга и климатической политики РФ, обозначенным в указах Президента РФ и в положениях Климатической доктрины Российской Федерации. Научная обоснованность принимаемых мер, самые передовые наукоемкие технологии обеспечивают оценку рисков и поддержку принятия решений по широкому кругу антропогенных и техногенных проблем. Реализация поставленных задач требует проведения фундаментальных и прикладных научных разработок аналитических и управляющих систем, способных эффективно обрабатывать и интерпретировать климатические данные, что, в свою очередь, связано с мониторингом природных и техногенных процессов и оценкой антропогенных рисков.

Специфические особенности обработки и анализа мультимодальных динамических и статических данных обосновывают необходимость разработки и реализации инструмента для хранения климатической информации в векторном формате, что становится оптимальным решением для описанной в работе информационной системы. Использование мультимодальных эмбедингов, объединяющих разнородные данные, а также интеграция эмбедингов, сгенерированных специализированными моделями векторизации, представляет, как показано в статье, эффективный подход к обработке климатических данных. Применение предложенной авторами статьи модульной архитектуры при проектировании информационной системы, обрабатывающей климатические данные, дает

ряд преимуществ, таких как масштабируемость, гибкость, использование механизмов предобработки данных, позволяющих согласовывать разные форматы, а также дает возможность повысить точность прогнозов и оценок и создать инструмент для оперативного реагирования на изменения климатической обстановки, что соответствует стратегическим интересам Российской Федерации в области экологической безопасности и устойчивого развития.

Список литературы

- [1] Указ Президента Российской Федерации от 26.10.2023 г. № 812. Адрес доступа: <http://www.kremlin.ru/acts/bank/49910>. (Дата обращения 20.05.2025)
- [2] **Фадеев С.В.** Экологический словарь. Санкт-Петербург; 2011: 35.
- [3] ЕИП Росгидромета: открытые данные. Адрес доступа: <https://eip.meteo.ru/opendata>. (Дата обращения 20.05.2025)
- [4] Climate forecast system reanalysis (CFSR). Available at: <https://climatedataguide.ucar.edu/climate-data/climate-forecast-system-reanalysis-cfsr>. (Accessed May 20, 2025)
- [5] Научная электронная библиотека “КиберЛенинка”. Адрес доступа: <https://cyberleninka.ru>. (Дата обращения 20.05.2025)
- [6] **Корневский Н.А., Крупчатников Р.А., Башир А.С., Климов Н.С.** Интеллектуальная система поддержки принятия решений по контролю и управлению состоянием окружающей среды. Вестник Курской государственной сельскохозяйственной академии. 2014; (2):78–80.
- [7] **Медведев А.В., Прокопенко Е.В., Кисяков И.М.** Система поддержки принятия решений в оценке экономической эффективности угледобывающей отрасли с учетом экологических ограничений. Уголь. 2021; 12(1149):28–33.
- [8] **Гладилин Д.Е., Тарасова Е.В., Кирильчук И.О.** Принятие решений в экологических СППР. Молодежь и системная модернизация страны: сборник научных статей 5-й Международной научной конференции студентов и молодых ученых. 2020; (5):12–15.
- [9] Climate risk data analytics with spectra by climate X. Available at: <https://www.climate-x.com/spectra>. (Accessed June 27, 2025)
- [10] Climate models. Available at: <https://climate.mit.edu/explainers/climate-models>. (Accessed June 27, 2025)
- [11] Искусственный интеллект для снижения риска бедствий: возможности, проблемы и перспективы. Адрес доступа: <https://wmo.int/ru/media/magazine-article/iskusstvennyu-intellekt-dlya-snizheniya-riska-bedstviy-vozmozhnosti-problemy-i-perspektivy>. (Дата обращения 27.06.2025)
- [12] **Thulke D., Gao Y., Pelsler P., Brune R., Jalota R., Fok F., Ramos M., Wyk I., Nasir A., Goldstein H., Tragemann T., Nguyen K., Fowler A., Stanco A., Gabriel J., Taylor J., Moro D., Tsybalov E., Waal J., Matusov E., Yaghi M., Shihadah M., Ney H., Dugast C., Dotan J., Erasmus D.** ClimateGPT: towards AI synthesizing interdisciplinary research on climate change. Available at: <https://arxiv.org/pdf/2401.09646>. (Accessed June 27, 2025)
- [13] **Reyhani H.S., Pasandideh S.M., Johnson S.N.** Artificial intelligence in ecology: a commentary on a Chatbot’s perspective. The Bulletin of the Ecological Society of America. 2023; 104(4). Available at: <https://esajournals.onlinelibrary.wiley.com/doi/10.1002/bes2.2097>.

- [14] RAG, AI agents, and agentic RAG: an in-depth review and comparative analysis. Available at: <https://www.heaptrace.com/blog-posts/rag-ai-agents-and-agentic-rag-an-in-depth-review-and-comparative-analysis>. (Accessed July 22, 2025)
- [15] Harris C. Microservices vs. monolithic architecture. Available at: <https://www.atlassian.com/microservices/microservices-architecture/microservices-vs-monolith>. (Accessed June 27, 2025)

Вычислительные технологии, 2025, том 30, № 4, с. 133–144. © ФИЦ ИВТ, 2025
Computational Technologies, 2025, vol. 30, no. 4, pp. 133–144. © FRC ICT, 2025

ISSN 1560-7534
eISSN 2313-691X

INFORMATION TECHNOLOGIES

DOI:10.25743/ICT.2025.30.4.012

Multimodal embeddings as an effective way for processing climate information within the framework of the Climate Doctrine of the Russian Federation

O. YU. GAVENKO^{1,2,*}, N. A. SHASHOK²

¹Matrosov Institute for System Dynamics and Control Theory of SB RAS, 664033, Irkutsk, Russia

²Federal Research Center for Information and Computational Technologies, 630090, Novosibirsk, Russia

*Corresponding author: Olga Yu. Gavenko, e-mail: olga.yu.gavenko@mail.ru

Received May 20, 2025, revised June 20, 2025, accepted July 01, 2025.

Abstract

Application of advanced technologies allows knowledge-intensive supporting for the decision-making on a wide range of issues related to anthropogenic and technogenic factors. Automation of such support requires fundamental and applied scientific research aimed at developing analytical and control systems which are capable correctly process and interpret climate data. An effective way to process climate information is to use multimodal embeddings that combine information from diverse sources. The development of an information system that uses multimodal climate data corresponds to the strategic goals of the Russian Federation in the environmental monitoring and climate policy.

Keywords: decision-making support, development of analytical systems, climate data, multimodal embeddings.

Citation: Gavenko O.Yu., Shashok N.A. Multimodal embeddings as an effective way for processing climate information within the framework of the Climate Doctrine of the Russian Federation. Computational Technologies. 2025; 30(4):133–144. DOI:10.25743/ICT.2025.30.4.012. (In Russ.)

References

1. Ukaz Prezidenta Rossiyskoy Federatsii ot 26.10.2023 g. No. 812 [Decree of the President of the Russian Federation dated 10.26.2023 No. 812]. Available at: <http://www.kremlin.ru/acts/bank/49910>. (Accessed May 05, 2025) (In Russ.)
2. Fadeev S.V. Ekologicheskiy slovar' [Ecological dictionary]. Saint Petersburg; 2011: 35. (In Russ.)
3. EIP Rosgidrometa: otkrytye dannye [Unified information platform of Roshydromet: open data]. Available at: <https://eip.meteo.ru/opendata>. (Accessed May 20, 2025) (In Russ.)
4. Climate forecast system reanalysis (CFSR). Available at: <https://climatedataguide.ucar.edu/climate-data/climate-forecast-system-reanalysis-cfsr>. (Accessed May 20, 2025)
5. Nauchnaya elektronnyaya biblioteka "KiberLeninka" [Scientific electronic library "CyberLeninka"]. Available at: <https://cyberleninka.ru>. (Accessed May 20, 2025) (In Russ.)

6. **Korenevsky N.A., Krupchatnikov R.A., Bashir A.S., Klimov N.S.** Intelligent decision support system for monitoring and managing the state of the environment. *Bulletin of the Kursk State Agricultural Academy*. 2014; (2):78–80. (In Russ.)
7. **Medvedev A.V., Prokopenko E.V., Kisyakov I.M.** Decision support system for assessing the economic efficiency of the coal mining industry taking into account environmental constraints. *UGOL'*. 2021; 12(1149):28–33. (In Russ.)
8. **Gladilin D.E., Tarasova E.V., Kirilchuk I.O.** Decision-making in environmental decision support systems. *Youth and the Systemic Modernization of the Country: Collection of Scientific Articles of the 5th International Scientific Conference of Students and Young Scientists*. 2020; (5):12–15. (In Russ.)
9. Climate risk data analytics with spectra by climate X. Available at: <https://www.climate-x.com/spectra>. (Accessed June 27, 2025)
10. Climate models. Available at: <https://climate.mit.edu/explainers/climate-models>. (Accessed June 27, 2025)
11. *Iskusstvennyy intellekt dlya snizheniya riska bedstviy: vozmozhnosti, problemy i perspektivy* [Artificial intelligence for disaster risk reduction: opportunities, challenges and prospects]. Available at: <https://wmo.int/ru/media/magazine-article/iskusstvennyy-intellekt-dlya-snizheniya-riska-bedstviy-vozmozhnosti-problemy-i-perspektivy>. (Accessed June 27, 2025) (In Russ.)
12. **Thulke D., Gao Y., Pelsler P., Brune R., Jalota R., Fok F., Ramos M., Wyk I., Nasir A., Goldstein H., Tragemann T., Nguyen K., Fowler A., Stanco A., Gabriel J., Taylor J., Moro D., Tsybalov E., Waal J., Matusov E., Yaghi M., Shihadah M., Ney H., Dugast C., Dotan J., Erasmus D.** ClimateGPT: towards AI synthesizing interdisciplinary research on climate change. Available at: <https://arxiv.org/pdf/2401.09646>. (Accessed June 27, 2025)
13. **Reyhani H.S., Pasandideh S.M., Johnson S.N.** Artificial intelligence in ecology: a commentary on a Chatbot's perspective. *The Bulletin of the Ecological Society of America*. 2023; 104(4). Available at: <https://esajournals.onlinelibrary.wiley.com/doi/10.1002/bes2.2097>.
14. RAG, AI agents, and agentic RAG: an in-depth review and comparative analysis. Available at: <https://www.heaptrace.com/blog-posts/rag-ai-agents-and-agentic-rag-an-in-depth-review-and-comparative-analysis>. (Accessed July 22, 2025)
15. **Harris C.** Microservices vs. monolithic architecture. Available at: <https://www.atlassian.com/microservices/microservices-architecture/microservices-vs-monolith>. (Accessed June 27, 2025)