



Общество с ограниченной ответственностью «Скайори»
шоссе Космонавтов., д.111И к1, помещ. 77, Пермь, 614066
ОГРН 1195958032181, ИНН 5905061802, КПП 590501001

Корпоративный ИИ Эксперт «Skyori AI Expert»

Описание функциональных характеристик

Листов 5

Пермь, 2026

писание и назначение системы

Корпоративный ИИ Эксперт «Skyori AI Expert» (далее – ИИ Эксперт, система) является функциональным решением, предназначенным для автоматизации процессов анализа документов и предоставления информационно-аналитических материалов пользователям. Система обеспечивает оперативный поиск релевантной информации в массиве документов по заданным тематическим вопросам, автоматизацию подготовки аналитических материалов с указанием использованных источников, а также диалоговый режим работы для уточнения запросов и поддержания контекста беседы.

ИИ Эксперт реализован в виде веб-приложения.

писание функциональных характеристик

адачи, которые решает система

ИИ Эксперт предназначен для решения следующих задач:

- сбор, обработка и систематизация текстовых документов;
- обеспечение семантического поиска релевантных фрагментов документов с использованием технологий векторного поиска и больших языковых моделей;
- автоматизация подготовки аналитических материалов, содержащих обобщенную информацию по загруженным документам с указанием источников;
- обеспечение диалогового режима взаимодействия с системой для уточнения запросов и поддержания контекста беседы;
- снижение временных затрат сотрудников на рутинные операции по поиску и первичному анализу текстовой информации.

Основные пользователи ИИ Эксперта:

- сотрудники, выполняющие поиск информации в большом массиве информации;
- аналитики, формирующие обзоры на основе большого количества разнородных источников информации.

ункции, которые выполняет система

ИИ Эксперт обеспечивает поддержку и выполнение следующих функций:

- загрузка и хранение текстовых документов в объектном хранилище с организацией иерархической структуры индексов;
- автоматическая индексация загруженных документов: извлечение текста, разбиение на фрагменты, векторизация с помощью модели эмбедингов и сохранение в векторном хранилище для последующего семантического поиска;
- семантический поиск релевантных фрагментов документов по запросу пользователя на основе векторных представлений (эмбедингов);

- генерация структурированных ответов на запросы пользователей с использованием больших языковых моделей на основе найденных релевантных фрагментов документов и контекста диалога;
- отображение блока источников с указанием использованных документов и коэффициента релевантности каждого источника, с возможностью просмотра выдержки из текста;
- диалоговый режим работы с поддержанием контекста беседы (передача предыдущих запросов и ответов на вход нового запроса для уточнения и детализации);
- одновременный запрос к нескольким индексам с компоновкой результатов в единый ответ;
- управление чатами (создание, переименование, удаление, добавление в избранное, поиск по списку чатов);
- управление источниками поиска (добавление, удаление документов, участвующих в генерации ответа);
- экспорт результатов запросов и перечня использованных источников в текстовом формате;
- настройка параметров генерации ответов: выбор языковой модели, количества источников, температуры и других параметров;
- управление промптами – создание и редактирование системных инструкций для настройки поведения языковой модели при генерации ответов;
- управление индексами: создание, редактирование, удаление индексов и вершин иерархии, запуск индексации, мониторинг статусов индексации.

Возможности ИИ Эксперта позволяют разграничить доступ пользователей к функциям в зависимости от их роли.

Ролевой моделью ИИ Эксперта определены автоматизированные рабочие места следующих пользователей: пользователь, администратор.

Пользователю предоставляются следующие возможности:

- создание новых чатов и отправка запросов;
- выбор индексов и настроек чата для поиска;
- просмотр истории сообщений в формате «вопрос-ответ» с форматированными ответами и блоком источников;
- работа с чатами: переименование, удаление, добавление в избранное, поиск по списку чатов;
- работа с источниками: добавление, удаление источников поиска;
- экспорт чата и источников в текстовом формате;

Администратор обладает всеми правами пользователя, а также дополнительно имеет следующие возможности:

- управление индексами (создание, редактирование, удаление индексов и вершин иерархии);
- запуск индексации всех индексов или отдельного индекса;
- управление настройками чатов: создание и редактирование промптов, настройка параметров генерации ответов (языковая модель, количество источников, количество фрагментов, размер контекста, температура);

Входной информацией являются:

- тексты документов, загружаемые в объектное хранилище;
- запросы пользователей, вводимые через веб-интерфейс в диалоговом режиме;
- настройки чатов: выбор индексов, параметры генерации;
- системные промпты и инструкции, настраиваемые администратором;

Выходной информацией являются:

- сгенерированные ответы ИИ Эксперта на запросы пользователей с указанием использованных источников (документов);
- выдержки из текстов документов, отображаемые при клике на источник;
- файлы экспорта чатов и источников в текстовом формате;
- метаданные индексов, документов и фрагментов, хранящиеся в базе данных.

требования к ресурсам для работы системы

ИИ Эксперт организован на основе клиент-серверной архитектуры с использованием доступа через веб-приложение.

Рекомендуемая конфигурация технических средств:

- Сервер ИИ Эксперта: два 4-ядерных процессора, ОЗУ не менее 16 Гб, не менее 20 Гб дискового пространства.

Веб-приложение реализовано с использованием следующих технологий:

- Технологии: .NET Core, Docker, LLM, OpenSearch, клиент-сервер, микросервисы.
- Язык программирования: C#, Python, TypeScript, HTML, CSS
- Операционная система: Ubuntu, Debian, CentOS, AstraLinux, РедОС, ALT Linux.
- СУБД: PostgreSQL.
- Объектное хранилище: MinIO.
- Векторное хранилище: OpenSearch.
- Языковая модель: подключение к LLM-провайдерам осуществляется через стандартные API-интерфейсы (REST API /

Рекомендуемая конфигурация рабочих станций пользователей, достаточная для работы ИИ Эксперта:

- процессор с тактовой частотой не ниже 2 ГГц на системной плате, обеспечивающей оптимальное взаимодействие остального комплекса технических средств;
- оперативная память объемом не менее 4 Гб;
- монитор, обеспечивающий отображение не менее 24 бит цветов в разрешении не менее 1024×768;
- жесткий диск с объемом свободного пространства не менее 20 Гб;
- периферийные устройства ввода (клавиатура, манипулятор мышь).

Рекомендации к установленным на рабочих станциях браузерам:

- Яндекс Браузер (версия не ниже 24);
- Google Chrome (версия не ниже 131).