

ПРОГРАММА ДЛЯ ЭВМ «СИНОД»

(вопросно-ответная система на основе базы знаний)

Описание функциональных характеристик

*экземпляра программного обеспечения,
предоставленного для проведения экспертной проверки*

1. Назначение программного обеспечения

Программа для ЭВМ «Синод» (далее — Программа) является вопросно-ответной системой на основе персональной базы знаний с применением технологии RAG (Retrieval-Augmented Generation — генерация с дополнением поиском).

Назначение Программы — предоставить пользователю инструмент для накопления структурированных знаний в виде статей и получения точных, контекстуально обоснованных ответов на вопросы, сформированных на основе содержимого собственной базы знаний пользователя с применением большой языковой модели.

Программа предназначена для использования физическими лицами, организациями, исследователями, аналитиками и иными пользователями, которым требуется эффективный инструмент работы с накопленной текстовой информацией.

2. Задачи, решаемые Программой

Программа «Синод» решает следующие задачи:

Таблица 2.1 — Задачи Программы

№	Задача	Описание
1	Формирование персональной базы знаний	Пользователь создаёт, редактирует и удаляет статьи, содержащие структурированную текстовую информацию по произвольным темам.
2	Поиск релевантной информации	По запросу пользователя Программа выполняет поиск статей из базы знаний, наиболее релевантных заданному вопросу.
3	Генерация ответа на вопрос	Программа формирует развёрнутый ответ на вопрос пользователя на основе найденных статей с применением языковой модели.
4	Разграничение доступа к данным	Каждый пользователь работает только со своей базой знаний; доступ к данным других пользователей исключён.
5	Обеспечение истории взаимодействий	Программа сохраняет историю вопросов и ответов пользователя для последующего просмотра.

3. Функциональные характеристики

3.1. Перечень функций

Программа реализует следующие функции, сгруппированные по функциональным областям:

3.1.1. Управление учетными записями пользователей

- Регистрация нового пользователя (ввод логина и пароля, создание учётной записи).
- Аутентификация пользователя (вход в систему по логину и паролю).
- Завершение сеанса работы (выход из системы).
- Управление профилем пользователя (изменение параметров учётной записи).

3.1.2. Управление базой знаний

- Создание статьи базы знаний с указанием заголовка, краткого описания и основного текста (контекста).
- Просмотр списка статей базы знаний пользователя.
- Просмотр содержимого отдельной статьи.
- Редактирование статьи (изменение заголовка, описания, контекста).
- Удаление статьи из базы знаний.
- Поиск статей по ключевым словам (фильтрация списка).

3.1.3. Вопросно-ответный интерфейс (Q&A)

- Приём текстового вопроса от пользователя.
- Автоматический семантический поиск релевантных статей базы знаний по смыслу запроса (на основе векторных представлений / эмбедингов), а не только по совпадению ключевых слов (RAG-механизм).
- Формирование контекстуального промпта: вопрос пользователя + найденные фрагменты статей.
- Передача промпта языковой модели и получение сгенерированного ответа.
- Отображение ответа пользователю с указанием использованных источников (статей).
- Сохранение пары «вопрос — ответ» в истории взаимодействий.

3.1.4. История взаимодействий

- Просмотр истории ранее заданных вопросов и полученных ответов.
- Повторный просмотр отдельных записей истории.

3.1.5. REST API

- Предоставление программного интерфейса (REST API) для взаимодействия клиентской части с серверной.
- Поддержка операций CRUD для статей базы знаний через API.
- Поддержка операции генерации ответа через API.
- Аутентификация запросов к API посредством JWT-токенов, выдаваемых сервером аутентификации Keycloak (OAuth 2.0 / OpenID Connect).

3.2. Сводная таблица функций

Таблица 3.1 — Функции Программы и их описание

№	Функц. область	Функция	Результат выполнения
1	Учётные записи	Регистрация пользователя	Создана учётная запись; пользователь может войти в систему
2	Учётные записи	Аутентификация (вход)	Пользователю выдан токен сессии; открыт доступ к личному кабинету
3	База знаний	Создание статьи	Статья сохранена в базе данных и доступна для поиска
4	База знаний	Редактирование статьи	Обновлённые данные статьи сохранены
5	База знаний	Удаление статьи	Статья удалена из базы данных и исключена из поиска
6	Q&A	Генерация ответа на вопрос	Пользователь получил текстовый ответ, сформированный на основе его базы знаний
7	История	Просмотр истории	Пользователь видит список сохранённых пар вопрос–ответ
8	API	CRUD-операции через REST API	Внешние системы могут управлять базой знаний программно

4. Входные данные

Входными данными для Программы являются данные, вводимые пользователем через веб-интерфейс или передаваемые через REST API:

Таблица 4.1 — Входные данные Программы

Наименование	Тип / формат	Описание
Логин пользователя	Строка, UTF-8	Уникальный идентификатор учётной записи
Пароль пользователя	Строка, UTF-8	Секретная фраза для аутентификации; хранится в хэшированном виде
Заголовок статьи	Строка, до 255 симв.	Краткое наименование статьи базы знаний
Описание статьи	Строка, UTF-8	Краткая аннотация содержания статьи
Контекст статьи	Текст, UTF-8	Основной текстовый контент статьи — используется при поиске и формировании ответа
Вопрос пользователя	Текст, UTF-8	Произвольный текстовый запрос, на который система должна сформировать ответ

5. Выходные данные

Программа формирует следующие выходные данные:

Таблица 5.1 — Выходные данные Программы

Наименование	Тип / формат	Описание
Токен аутентификации (JWT)	Строка, JSON	Выдаётся при входе в систему; используется для авторизации последующих запросов
Список статей базы знаний	JSON / HTML-страница	Перечень статей пользователя с заголовками и описаниями
Содержимое статьи	JSON / HTML-страница	Полный текст статьи (заголовок, описание, контекст)
Ответ на вопрос пользователя	Текст, UTF-8	Сгенерированный языковой моделью развёрнутый текстовый ответ с учётом контекста базы знаний
Источники ответа	Список ссылок / JSON	Перечень статей базы знаний, использованных при формировании ответа
История взаимодействий	JSON / HTML-страница	Список сохранённых пар «вопрос — ответ» с отметками времени

6. Режимы работы Программы

Программа функционирует в непрерывном режиме (24/7) как серверное веб-приложение и поддерживает одновременную работу нескольких пользователей.

Таблица 6.1 — Режимы работы Программы

Режим	Описание
Интерактивный (пользовательский)	Пользователь взаимодействует с Программой через веб-браузер: управляет статьями базы знаний и задаёт вопросы
Программный (API)	Внешние системы взаимодействуют с Программой через REST API: получение, создание, изменение, удаление статей, а также генерация ответов
Фоновый (автоматический)	Поиск релевантных статей и генерация ответа выполняются автоматически без участия пользователя после получения вопроса

7. Затрачиваемые ресурсы

7.1. Аппаратные ресурсы

Таблица 7.1 — Потребление аппаратных ресурсов в рабочем режиме

Компонент	Типовое потребление	Пиковое потребление
-----------	---------------------	---------------------

Процессор (CPU)	5–15% (4-ядерный CPU)	До 80% при генерации ответа LLM
Оперативная память (RAM)	4–6 ГБ (все компоненты)	До 12–16 ГБ при запуске LLM
Дисковое пространство	5–10 ГБ (БД + образы)	До 20 ГБ с весами LLM-модели
Сетевой трафик	Менее 1 Мбит/с (веб-трафик)	Зависит от числа пользователей

***Примечание.** Потребление ресурсов языковой моделью (LLM) является доминирующим фактором. При использовании квантизированной версии модели объём занимаемой памяти снижается до 4–6 ГБ.*

7.2. Временные характеристики

Таблица 7.2 — Временные характеристики выполнения функций

Операция	Типовое время	Примечание
Загрузка страницы веб-интерфейса	< 1 с	Локальная сеть
Создание / редактирование статьи	< 1,5 с	Запись в БД
Поиск релевантных статей	< 3 с	Запрос к PostgreSQL
Генерация ответа (LLM)	5–30 с	Зависит от мощности CPU/GPU и длины ответа
Аутентификация пользователя	< 3 с	Проверка JWT

8. Ограничения и условия применения

- Программа предназначена для работы в среде с поддержкой Docker и Docker Compose на Linux-сервере.
- Веб-интерфейс Программы рассчитан на использование через современные веб-браузеры (Chrome 110+, Firefox 110+ и совместимые) без необходимости установки дополнительного клиентского программного обеспечения.
- Программа не предназначена для обработки персональных данных в понимании Федерального закона от 27.07.2006 № 152-ФЗ без надлежащей дополнительной конфигурации.
- Качество и релевантность генерируемых ответов определяются полнотой и качеством наполнения базы знаний пользователя.
- Программа не осуществляет доступ к сети Интернет в процессе генерации ответов; языковая модель функционирует локально.