Техническое задание

Хакатон: ИИ-агент для научных исследований и подготовки статей

1. Общая информация

Название проекта: ИИ-агент для научных исследований и подготовки статей

Заказчик: Жезказганский университет имени О.А. Байконурова Разработчик: Проведение Хакатона. 7-8 августа 2025 г. ОФЛАЙН

Дата составления Т3: 25.07.2025

2. Цель проекта

Создать прототип ИИ-сервиса, который помогает преподавателям, магистрантам и студентам Жезказганского университета автоматизировать процессы поиска литературы, анализа научных текстов, написания аннотаций и оформления источников для написания научных публикаций, их оформления и идеи для будущих исследований, гипотезы, возможные методологии, помощь в оформлении заявок на Конкурсы проводимых НЦГНТЭ и др.

3. Функциональные требования

- 1. Обработка входных данных. ИИ-агент должен уметь принимать и обрабатывать следующие типы входных данных: текстовые запросы, загруженные статьи (PDF, DOCX, TXT), ключевые слова. Необходимо реализовать парсинг текста и выделение метаданных (название, авторы, аннотация).
- 2. Поиск научной литературы. ИИ-агент должен использовать внешние API (CrossRef, Semantic Scholar и др.) для поиска релевантной литературы по ключевым словам, темам или тезисам. Результаты должны отображать название, DOI, источник и краткую аннотацию.
- 3. Генерация аннотаций и рефератов. На основе загруженной статьи или текста ИИ должен генерировать краткий обзор (резюме), выделяя цель исследования, методы, результаты и выводы.
- 4. Автоматический перевод. Поддержка перевода научного текста между казахским, русским и английским языками. Предусмотреть использование сторонних сервисов (Google Translate API, HuggingFace models) или локальных моделей (MarianMT, KazTrans).
- 5. Генерация и оформление библиографии. ИИ-агент должен автоматически создавать список литературы по введённым DOI или источникам, оформленный по выбранному стилю (ГОСТ, APA, MLA).
- 6. Проверка на плагиат (семантическая). На основе семантического сравнения текстов ИИ должен выявлять схожесть с другими научными текстами. Можно использовать алгоритмы сравнения по эмбеддингам или внешние API.
- 7. Структурный анализ научной статьи. ИИ-агент должен распознавать и обозначать ключевые компоненты статьи: введение, цели, гипотезы, методы, результаты, выводы. Возможна визуализация структуры.
- 8. Интерфейс взаимодействия. Пользовательский интерфейс должен обеспечивать простое взаимодействие с агентом: загрузка файлов, ввод запросов, отображение результатов в удобной форме. Варианты: веб-интерфейс (Streamlit/Flask) или Telegram-бот.
- 9. Языковая локализация. Интерфейс и функции ИИ должны поддерживать переключение языков: казахский, русский, английский.
- 10. Модуль предложений тем и гипотез. ИИ должен на основе заданной области знаний или загруженной статьи предлагать идеи для будущих исследований, гипотезы, возможные методологии и направления для написания научной статьи. Создаёт краткие обзоры по теме и/или предлагает темы для исследований.
- 11. Модуль помощник в оформлении заявок на Конкурсы проводимых НЦГНТЭ https://www.ncste.kz/ru/competition-2.

Наименование процессов и полезные функции

Процесс (задача)	Полезные функции ИИ-агента
Поиск научной информации	Автоматический подбор релевантных статей из
	международных баз
Анализ научных публикаций	Реферирование, выделение ключевых идей, построение
	структуры
Создание и редактирование научного	Генерация аннотаций, введений, выводов, на основе
текста	введённых данных
Перевод научного контента	Поддержка казахского, русского и английского языков
Проверка источников и цитирование	Автоматическое оформление по ГОСТ, APA, MLA
Проверка оригинальности	Семантическая проверка на заимствования
Помощь в формировании гипотез и	Предложения идей, подбор теоретических моделей и
обзоров	источников
Помощник в оформлении заявок на	Оформление заявок по требованиям Конкурсов НЦГНТЭ и
Конкурсы	др. https://www.ncste.kz/ru/competition-2

4. Языки интерфейса

- Казахский
- Русский
- Английский

5. Технические требования

- Доступность 24/7, время отклика ≤ 3 сек;
- Адаптивный дизайн (мобильные и десктопные устройства);
- Совместимость с LMS университета, eGov и ENIS;
- База знаний редактируемая из админ-панели;
- Интеграция с Telegram, WhatsApp, сайт университета.

6. Архитектура и технологии

- Использование LLM (например, GPT, KazNLP);
- Web-бэкенд: Python (FastAPI/Django) или Node.js;
- Фронтенд: React/HTML5;
- База данных: PostgreSQL/MySQL;
- Интеграция с внешними АРІ (по согласованию);
- Панель администратора для загрузки программ, тестов, мероприятий.

7. Этапы реализации

✓ Проведение Хакатона 7-8 августа 2025 г. ОФЛАЙН

8. Безопасность и соответствие требованиям

- Соответствие Закону РК "О персональных данных";
- Шифрование данных;
- Поддержка анонимного режима.

9. Критерии приёмки

- Точность ответов $\geq 90\%$ по тестовому набору;
- Полная поддержка казахского, русского и английского языков;
- Корректная работа на всех платформах;
- Интеграция с мессенджерами и сайтом и LMS системой вуза.

10. Контакты и ответственные

- Заказчик: Жезказганский университет имени О.А. Байконурова
- Исполнитель: Проведение Хакатона
- Ответственный за контент: проректор по HuCP/PR-отдел
- Ответственный за тех. часть: ІТ-отдел университета и подрядчик.