**ДАТА УТВЕРЖДЕНИЯ**

**I ПАСПОРТ ПРОЕКТА**

|  |  |
| --- | --- |
| Наименование дополнительной профессиональной программы профессиональной переподготовки | Цифровое моделирование и суперкомпьютерные технологии |
| Наименование проекта | Замена персонажей в литературных произведениях с помощью больших языковых моделей |
| Шифр проекта (команды) | МАИ.2024.ДПППП-3.ЗАМПЕР |
| Заказчик проекта | Кафедра 806 «Вычислительная математика и программирование» |
| Руководитель темы от МАИ | Дмитрий Валерьевич Сошников, доцент кафедры 806 |
| Консультант | Дмитрий Валерьевич Сошников, доцент кафедры 806 МАИ |
| Рецензент темы | Дмитрий Валерьевич Сошников, доцент кафедры 806 МАИ |
| Целевая аудитория результата проекта (кто потребитель результата проекта) | Дети, для удовлетворения своих фантазий, путем добавления себя в качестве главного героя в любимые литературные произведения |
| Длительность проекта (даты начала и окончания) | 92 календарных дня (01.03.2024 – 31.05.2024) |
| Название команды | Медузы |
| Роли в проекте: | ФИО |
| Team Lead | Мамонтов Егор Олегович |
| Back-end developer | Андрюшин Лев Дмитриевич |
| Тестировщик | Андриянов Эрик Александрович |
| Back-end developer | Муратов Артём Алексеевич |
| Prompt-engineer | Комбаров Владислав Александрович |
| Составление и компоновка итогового отчёта | Маслова Анастасия Алексеевна |
| Prompt-engineer | Кудинов Денис Викторович |
| Front-end developer | Модин-Глазков Богдан Арсеньевич |

**Ссылки на ресурсы проекта**

|  |  |
| --- | --- |
| Ссылка на GitHub | https://github.com/TeoPlow/IT\_project\_jellyfish |
| Ссылка на доску в Trello | https://trello.com/b/vOBv7KlT/задачи |
| Сcылка на MIRO | https://miro.com/app/board/uXjVNGj7Ckg=/ |
| Ссылка на Google Collab | https://colab.research.google.com/drive/1SYFEXkvkhxT2c6WWP-x-xuxh1rvv-Wi6?usp=sharing |

**II ОПИСАНИЕ ПРОЕКТА**

|  |  |
| --- | --- |
| Образ результата | Сервис, позволяющий заменить персонажа в литературном произведении на другого персонажа, без сильного изменения сюжета |
| Цель проекта | Создать сервис, который будет помогать детям с большей вовлечённостью усваивать литературные произведения, и написать диплом к 31 мая 2024 года |
| Задачи проекта |  |
| 1 | Подготовительные работы.  Результат – готовое рабочее пространство |
| 2 | Координация работы команды.  Результат – налаженная коммуникация в команде |
| 3 | Создание системного проектного документа.  Результат – план разработки и паспорт проекта |
| 4 | Исследовательская работа по текстовым моделям.  Результат – понять, какая текстовая модель больше подходит под решение задачи |
| 5 | Разработка и тестирование программы для решения задачи.  Результат – готовый сервис |
| 6 | Оформление исследовательской работы.  Результат – готовая дипломная работы. |
| Ограничения и допущения, которые имеют или могут оказать существенное влияние на результат проекта | Ограничения на использование некоторых текстовых моделей из-за санкций.  Трудность в нахождении литературных произведений, так как они защищены от пиратства.  Невозможность нынешних текстовых моделей обработать сразу всё произведение.  Большое количество запросов к текстовым моделям стоит денег. |
| Необходимые ресурсы для выполнения проекта (компетенции исполнителей, материальные ресурсы и др.) | 1. Финансовые требования:    * Бюджет для закупки токенов   1600 руб.   1. Технологический стек:    * Программирование: Python.    * Фреймворки и библиотеки: пока не известны;    * Базы данных: литературные произведения. 2. Инфраструктура для применения технологий:    * Персональный компьютер. |
| Риски проекта | Риск отчисления участников проекта из высшего учебного заведения, что не позволит закончить проект. |

**III КОМАНДА ПРОЕКТА**

|  |  |  |  |
| --- | --- | --- | --- |
| **ФИО** | **Роль** | **Компетенция** | **Задача проекта** |
| Мамонтов Егор Олегович | Team Leader | С++, Linux, Python, Excel, MIRO, Trello, Git | 1. Подготовительные работы:    * Собрать команду    * Создать рабочее пространство в Miro, Trello, Google Colab    * Создать место для коммуникации команды в Telegram, Discord.    * Создать GitHub репозиторий 2. Координация работы команды:    * Назначение задач участникам команды    * Контроль за выполнением поставленных задач    * Обеспечение коммуникации между командой и консультантом 3. Создание системного проектного документа:    * Координация и руководство процессом создания документа    * Придумать план разработки проекта    * Выбор лучшего инструмента для разработки проекта    * Создание паспорта проекта    * Создание плана разработки в MIRO и Trello 4. Оформление исследовательской работы:    * Поиск теоретических материалов по своей части    * Конспектирование всех действий по проекту |
| Андрюшин Лев Дмитриевич | старший Backend-developer,  заместитель Team Leader | С, С++, Figma, Excel, Linux, Python | 1. Подготовительные работы:    * Помощь TeamLeader’у в организации работ 2. Координация работы команды:    * Помощь TeamLeader’у в координации команды 3. Создание системного проектного документа:    * Выбор лучшего инструмента для разработки проекта 4. Back-End разработка основной программы:    * Определиться с окончательным видением программы    * Создать функцию приёма и обработки текста    * Создать функцию приёма и обработки задания с учётом промптов    * Создать код кластеризации текста на маленькие части    * Создать код объединения кластеров и заданий, чтобы получить готовый запрос для текстовой модели    * Подключить API текстовой модели к программе    * Реализовать возможность запроса и получения ответа от API    * Реализовать последовательную отправку кластеров к API и сохранению ответов в единый файл    * Реализовать вывод отредактированного произведения 5. Объединение и тестирование всех частей сервиса:    * Объединить интерфейс и основную программу в единый сервис 6. Оформление исследовательской работы:    * Поиск теоретических материалов по своей части    * Конспектирование всех действий по проекту |
| Андриянов Эрик Вячеславович | Тестировщик (QA) | С, С++, Python, DevOps | 1. Объединение и тестирование всех частей сервиса:    * Выбрать лучший инструмент для тестирования    * Протестировать основную программу    * Объединить интерфейс и основную программу в единый сервис    * Протестировать сервис целиком    * Указать на существующие ошибки разработчикам 2. Оформление исследовательской работы:    * Поиск теоретических материалов по своей части    * Конспектирование всех действий по проекту |
| Муратов Артём Алексеевич | Back-end developer | С++, С, Back-end разработка, CSS, DevOps, Machine Learning, HTML, Git, Figma, Docker, Django, Pandas, PyTorch, Python, SQL, Работа с XML, Scikit-Learn, Seaborn, СУБД Oracle, Linux, Matplotlib, Excel, PowerPoint | 1. Back-End разработка основной программы:    * Определиться с окончательным видением программы    * Создать функцию приёма и обработки текста    * Создать функцию приёма и обработки задания с учётом промптов    * Создать код кластеризации текста на маленькие части    * Создать код объединения кластеров и заданий, чтобы получить готовый запрос для текстовой модели    * Подключить API текстовой модели к программе    * Реализовать возможность запроса и получения ответа от API    * Реализовать последовательную отправку кластеров к API и сохранению ответов в единый файл    * Реализовать вывод отредактированного произведения 2. Объединение и тестирование всех частей сервиса:    * Объединить интерфейс и основную программу в единый сервис 3. Оформление исследовательской работы:    * Поиск теоретических материалов по своей части    * Конспектирование всех действий по проекту |
| Комбаров Владислав Александрович | Prompt-engineer | С, С++, Excel, Git, Linux, NumPy, Python, PowerPoint, Математическая статистика | 1. Исследовательская работа по текстовым моделям:    * Поиск потенциальных текстовых моделей    * Создание Dataset`а из текстовых произведений    * Сбор информации о текстовых моделях    * Поиск оптимальных промптов для решения задачи    * Тестирование различных текстовых моделей    * Создать оптимальный набор промптов для решения задачи    * Решить, какая текстовая модель лучше подходит для решения задачи 2. Оформление исследовательской работы:    * Поиск теоретических материалов по своей части    * Конспектирование всех действий по проекту |
| Маслова Анастасия Алексеевна | Составление и компоновка итогового отчёта | С++, Python, Математическая статистика | 1. Оформление исследовательской работы:    * Поиск теоретических материалов по своей части    * Поиск примера дипломной работы    * Конспектирование всех действий по проекту    * Соединение всей информации, собранной в ходе исследований    * Общее оформление всей работы |
| Кудинов Денис Викторович | Prompt-engineer | С++, С, Python, HTML | 1. Исследовательская работа по текстовым моделям:    * Поиск потенциальных текстовых моделей    * Создание Dataset`а из текстовых произведений    * Сбор информации о текстовых моделях    * Поиск оптимальных промптов для решения задачи    * Тестирование различных текстовых моделей    * Создать оптимальный набор промптов для решения задачи    * Решить, какая текстовая модель лучше подходит для решения задачи 2. Оформление исследовательской работы:    * Поиск теоретических материалов по своей части    * Конспектирование всех действий по проекту |
| Модин-Глазков Богдан Арсеньевич | Front-end developer | Back-end разработка, Data Science, Docker, Excel, Go, Jupyter, Machine Learning, Matplotlib, NumPy, Scikit-Learn, Математическая статистика, Seaborn, Python, Pandas, Linux, Git, С, С++ | 1. Front-end разработка интерфейса сервиса:    * Определиться с окончательным видением интерфейса    * Создать интерфейс программы 2. Объединение и тестирование всех частей сервиса:    * Объединить интерфейс и основную программу в единый сервис 3. Оформление исследовательской работы:    * Поиск теоретических материалов по своей части    * Конспектирование всех действий по проекту |

**IV ЗАДАЧИ ПРОЕКТА (ОЦЕНКА ПО ВРЕМЕНИ)**

|  |  |  |
| --- | --- | --- |
| **Задача** | **Подзадачи** | **Время на выполнение (чел.-час)** |
| 1. Подготовительные работы | * 1. Собрать команду   2. Создать рабочее пространство в Miro, Trello, Google Colab;   3. Создать место для коммуникации команды в Telegram, Discord;   4. Создать GitHub репозиторий;   5. Помощь Team Leader’у в организации работ; | 10 часов |
| 2. Координация работы команды | * 1. Назначение предварительных задач участникам команды;   2. Контроль за выполнением поставленных задач;   3. Обеспечение коммуникации между командой и консультантом;   4. Помощь Team Leader’у в координации команды; | 10 часов |
| 3. Создание системного проектного документа | 3.1 Координация и руководство процессом создания документа;  3.2 Придумать план разработки проекта;  3.3 Создание паспорта проекта;  3.4 Выбор лучшего инструмента для разработки проекта;  3.5 Создание плана разработки в MIRO и Trello; | 10 часов |
| 4. Исследовательская работа по текстовым моделям | 4.1 Поиск потенциальных текстовых моделей;  4.2 Создание Dataset`а из текстовых произведений;  4.3 Сбор информации о текстовых моделях;  4.4 Поиск оптимальных промптов для решения задачи;  4.5 Тестирование различных текстовых моделей;  4.6 Создать оптимальный набор промптов для решения задачи;  4.7 Решить, какая текстовая модель лучше подходит для решения задачи; | 100 часов |
| 5. Front-end разработка интерфейса сервиса | * 1. Определиться с окончательным видением интерфейса;   2. Создать интерфейс программы; | 80 часов |
| 6. Back-End разработка основной программы | 6.1 Определиться с окончательным видением программы;  6.2 Создать функцию приёма и обработки текста;  6.3 Создать функцию приёма и обработки задания с учётом промптов;  6.4 Создать код кластеризации текста на маленькие части;  6.5 Создать код объединения кластеров и заданий, чтобы получить готовый запрос для текстовой модели;  6.6 Подключить API текстовой модели к программе;  6.7 Реализовать возможность запроса и получения ответа от API;  6.8 Реализовать последовательную отправку кластеров к API и сохранению ответов в единый файл;  6.9 Реализовать вывод отредактированного произведения; | 100 часов |
| 7. Объединение и тестирование всех частей сервиса | 7.1 Выбрать лучший инструмент для тестирования;  7.2 Протестировать основную программу;  7.3 Объединить интерфейс и основную программу в единый сервис;  7.4 Протестировать сервис целиком;  7.5 Указать на существующие ошибки разработчикам; | 50 часов |
| 8. Оформление исследовательской работы | 8.1 Поиск теоретических материалов по своей части   * 1. Поиск примера дипломной работы   8.3 Конспектирование всех действий по проекту  8.4 Соединение всей информации, собранной в ходе исследований  8.5 Общее оформление всей работы | 50 часов |
| **ИТОГО ПЛАНИРУЕМОЕ ВРЕМЯ НА ПРОЕКТ:** |  | **410 часов** |