

I ПАСПОРТ ПРОЕКТА

Наименование проекта	Система автоматического анализа электронного документа (PDF)
Шифр проекта (команды)	MAI's Hive
Заказчик проекта	ПАО «МТС»
Руководитель темы от МАИ	Булакина Мария Борисовна
Целевая аудитория результата проекта (кто потребитель результата проекта)	Бухгалтерия, секретари
Длительность проекта (даты начала и окончания)	11.11.2024 - 25.05.2025
РОЛИ В ПРОЕКТЕ:	ФИО
Team lead	Мишин Сергей Алексеевич
UX/UI-дизайнер	Барменков Артемий Сергеевич
Технический писатель	Бондаренко Виктория Арсеновна
Frontend-разработчик	Борзова Дарья Максимовна
Backend-разработчик	Иванченко Макар Дмитриевич
Аналитик	Казец Полина Олеговна
Frontend-разработчик	Савинов Никита Олегович
Data-инженер	Сергеев Владимир Андреевич
QA-инженер	Федоров Ярослав Артемович
Backend-разработчик	Шведов Александр Иванович
Дата создания первой версии паспорта проекта	15.12.2024

Ссылки на ресурсы проекта

Ссылка на GitHub	https://github.com/oryce/mai-ck
------------------	---

II ОПИСАНИЕ ПРОЕКТА

Образ результата	<i>IT-продукт</i>
Цель проекта	<i>Разработка интерактивной системы анализа электронных документов</i>
Задачи проекта	
1	<i>Провести анализ существующих решений</i>
2	<i>Определить функциональные возможности, составить техническое задание</i>
3	<i>Разработать архитектуру системы</i>
4	<i>Разработать backend-часть системы</i>
5	<i>Разработать frontend-часть системы</i>
6	<i>Протестировать систему на различных данных</i>
7	<i>Оформить техническую документацию и отчет</i>
Результат проекта (конкретный итоговый результат проекта)	<i>Интерактивная система анализа электронных документов, обеспечивающая автоматизацию процессов и повышение эффективности работы персонала</i>
Ограничения и допущения, которые имеют или могут оказать существенное влияние на результат проекта	<p><i>Ограничения вычислительных мощностей для обработки больших объемов данных в реальном времени.</i></p> <p><i>Совместимость форматов: Сложности в поддержке редких или устаревших форматов документов.</i></p> <p><i>Точность OCR: Ограниченная точность распознавания текста в отсканированных или низкокачественных изображениях.</i></p> <p><i>Ограничение по времени выполнения проекта.</i></p> <p><i>Допущение о наличии стабильного интернет-соединения для использования приложения.</i></p>
Необходимые ресурсы для выполнения проекта (компетенции исполнителей, материальные ресурсы и др.)	<p><i>Серверные мощности.</i></p> <p><i>Программное обеспечение:</i></p> <ol style="list-style-type: none"> <i>1. Среды разработки.</i> <i>2. Библиотеки и фреймворки.</i> <i>3. Системы управления проектами.</i> <i>4. Средства тестирования.</i>

	<p>5. Системы контроля версий.</p> <p>6. Большая языковая модель (LLM) с открытым исходным кодом.</p> <p>Временные ресурсы.</p>
Риски проекта	<p>Высокие объемы данных.</p> <p>Ограниченная поддержка некоторых форматов (например, устаревших PDF или редких типов документов).</p> <p>Задержки в разработке из-за технических проблем.</p>

III КОМАНДА ПРОЕКТА

ФИО	Роль	Компетенция	Задача проекта
Мишин Сергей Алексеевич	<i>Team lead</i>	<ol style="list-style-type: none"> 1. Управление разработкой с использованием Agile/Scrum 2. Архитектурное проектирование 3. OCR, NLP 	<ol style="list-style-type: none"> 1. Организация процесса разработки и распределения задач 2. Обеспечение технического качества системы и ее архитектуры 3. Управление коммуникацией внутри команды и с заинтересованными сторонами 4. Контроль выполнения задач и соблюдения сроков
Барменков Артемий Сергеевич	<i>UX/UI-дизайнер</i>	<ol style="list-style-type: none"> 1. Figma, Adobe XD 2. Опыт проектирования интерфейсов для сложных систем 3. Знание принципов адаптивного и интерактивного дизайна 	<ol style="list-style-type: none"> 1. Разработка интуитивно понятного интерфейса для взаимодействия с системой 2. Обеспечение удобного пользовательского опыта (UX) с учетом различных сценариев 3. Создание современного визуального стиль, включая цветовую схему и иконки
Бондаренко Виктория Арсеновна	<i>Технический писатель</i>	<ol style="list-style-type: none"> 1. Знание ГОСТ 7.32-2017 2. Владение текстовыми редакторами: Microsoft Word и LaTeX 3. Умение структурировать и форматировать документы в соответствии со стандартами 	<ol style="list-style-type: none"> 1. Создание спецификаций, технических описаний, отчетов об испытаниях, руководств по установке и настройке 2. Регулярное обновление документации в соответствии с изменениями в продукте 3. Приведение документации к единому стилю и формату
Борзова Дарья Максимовна	<i>Frontend-разработчик</i>	<ol style="list-style-type: none"> 1. Владение JavaScript, React 2. Умение работать с REST API и WebSocket 3. Знание HTML, CSS и адаптивной верстки 	<ol style="list-style-type: none"> 1. Разработка пользовательского интерфейса на основе макетов 2. Реализация загрузки и отображения документов 3. Интеграция с Backend API для обработки данных 4. Тестирование и отладка интерфейса

Иванченко Макар Дмитриевич	Backend-разработчик	<ol style="list-style-type: none"> 1. Boost, gRPC, REST API 2. Работа с базами данных (SQL, NoSQL) и многопоточность 	<ol style="list-style-type: none"> 1. Разработка модулей обработки текста (токенизация, ключевые слова) 2. Интеграция OCR-инструментов для распознавания текста 3. Оптимизация производительности системы
Казец Полина Олеговна	Аналитик	<ol style="list-style-type: none"> 1. Построение моделей процессов (BPMN, UML) 2. Анализ данных и работа с SQL 	<ol style="list-style-type: none"> 1. Сбор требований от пользователей и заказчиков, их формализация и документирование 2. Проектирование функционала системы, описание сценариев использования 3. Анализ входных данных, определение методов их обработки и требований к качеству 4. Взаимодействие с командой разработки и тестирования для реализации и проверки требований
Савинов Никита Олегович	Frontend-разработчик	<ol style="list-style-type: none"> 1. Владение JavaScript, React 2. Умение работать с REST API и WebSocket 3. Знание HTML, CSS и адаптивной верстки 	<ol style="list-style-type: none"> 1. Создание интерактивных элементов и визуализации данных 2. Оптимизация интерфейса для работы с большими документами 3. Реализация мультиязычного интерфейса 4. Улучшение производительности и скорости рендеринга
Сергеев Владимир Андреевич	Data-инженер	<ol style="list-style-type: none"> 1. ETL/ELT 2. Работа с реляционными и NoSQL базами данных (PostgreSQL, MongoDB) 3. Разработка и настройка распределенных систем обработки данных (Apache Spark, Kafka) 	<ol style="list-style-type: none"> 1. Разработка ETL-процессов для обработки данных из электронных документов 2. Настройка хранилища данных для хранения и быстрого доступа к документам 3. Интеграция данных с модулями анализа текста и OCR 4. Обеспечение безопасности, мониторинга и отказоустойчивости системы обработки данных

Федоров Ярослав Артемович	QA-инженер	<ol style="list-style-type: none"> 1. Selenium, Postman 2. JMeter, LoadRunner 3. OWASP, Burp Suite 	<ol style="list-style-type: none"> 1. Проверка корректности работы системы: загрузка документов, анализ текста, поиск данных 2. Измерение производительности системы при обработке больших документов и высокой одновременной нагрузке 3. Проверка удобства и корректности работы интерфейса на разных устройствах и экранах 4. Проверка защиты данных, уязвимостей и правильности управления доступом
Шведов Александр Иванович	Backend-разработчик	<ol style="list-style-type: none"> 1. Boost, gRPC, REST API 2. Работа с базами данных (SQL, NoSQL) и многопоточность 	<ol style="list-style-type: none"> 1. Проектирование серверной архитектуры системы 2. Создание и настройка REST API 3. Интеграция и оптимизация работы с базой данных 4. Обеспечение безопасности данных и API

IV ЗАДАЧИ ПРОЕКТА (ОЦЕНКА ПО ВРЕМЕНИ)

№	Название задачи	Подзадачи	Время (в часах)
1	Настройка проекта в среде разработки	<ol style="list-style-type: none"> 1. Установка и настройка среды разработки для фронтенда и бекенда (ReactJS, Python, PostgreSQL) 2. Настройка базового проекта на Python (FastAPI или Django) и фронтенда (ReactJS) 3. Интеграция с PostgreSQL 	10
2	Разработка бекенда	<ol style="list-style-type: none"> 1. Создание базовых классов и структуры проекта 2. Настройка подключения к базе данных PostgreSQL 3. Настройка REST API для взаимодействия с фронтендом 	25
3	Разработка системы авторизации пользователей	<ol style="list-style-type: none"> 1. Реализация регистрации пользователей 2. Реализация входа в систему 3. Реализация защиты маршрутов (JWT/Session) 	20
4	Разработка базы данных	<ol style="list-style-type: none"> 1. Определение структуры данных: результаты анализа, пользователи 2. Создание таблиц и связей 3. Оптимизация базы данных для работы с большими файлами. 	12
5	Микросервис: распознавание текста (OCR)	<ol style="list-style-type: none"> 1. Интеграция готового OCR-решения (например, Tesseract) 2. Разработка API для передачи документов и возврата текста 3. Оптимизация обработки для различных форматов документов 	30
6	Микросервис: анализ текста документа – Интеграция с LLaMa через API	<ol style="list-style-type: none"> 1. Интеграция с LLaMa через API 2. Создание промпта для определения типа документа (договор, накладная и т.д.) 3. Тестирование и отладка ответа модели 	25
7	Микросервис: анализ подписи и печати	<ol style="list-style-type: none"> 1. Разработка алгоритма для проверки наличия подписи и печати 2. Интеграция алгоритмов анализа изображений (например, OpenCV) 3. Настройка API для анализа загруженных документов 	40
8	Разработка frontend на ReactJS	<p>Создание компонентов для:</p> <ol style="list-style-type: none"> 1. Окна авторизации/регистрации 2. Загрузки и отображения документов 3. Просмотра результатов анализа 4. Личного кабинета пользователя 	50

9	Написание стилей	<ol style="list-style-type: none"> 1. Реализация пользовательского интерфейса (CSS/SCSS) 2. Создание адаптивного дизайна 	20
10	Интеграция бекенда и фронтенда	<ol style="list-style-type: none"> 1. Настройка взаимодействия через REST API 2. Тестирование работы основных функций (авторизация, загрузка документов, просмотр результатов) 	30
11	Разработка системы тестирования	<ol style="list-style-type: none"> 1. Написание тестов для бекенда (юнит-тесты и интеграционные тесты) 2. Написание тестов для фронтенда. 3. Тестирование микросервисов 	15
12	Документация	<ol style="list-style-type: none"> 1. Написание руководства для пользователя. 2. Описание API (Swagger/OpenAPI) 3. Подготовка презентации 	15
13	Развертывание приложения	<ol style="list-style-type: none"> 1. Настройка серверной инфраструктуры (Docker, CI/CD) 2. Размещение бекенда и фронтенда 3. Финальное тестирование в реальных условиях 	15
ИТОГО ПЛАНИРУЕМОЕ ВРЕМЯ НА ПРОЕКТ :			307