**НЕГОСУДАРСТВЕННОЕ ОБРАЗОВАТЕЛЬНОЕ ЧАСТНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ**

**«МОСКОВСКИЙ ФИНАНСОВО-ПРОМЫШЛЕННЫЙ УНИВЕРСИТЕТ «СИНЕРГИЯ»**

|  |  |  |
| --- | --- | --- |
| **Факультет/Институт** |  | Информационные системы и технологии |
|  |  | (наименование факультета/ Института) |
| **Направление/специальность** |  | Разработка, сопровождение и обеспечение безопасности информационных систем |
| **подготовки:** |  | (код и наименование направления /специальности подготовки) |
| **Форма обучения:** |  | очная |
|  |  | (очная, очно-заочная, заочная) |
|  |  |  |

**Лабораторный практикум № 1**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **по дисциплине** |  | **Инструментальные средства информационных систем** | | |
|  | | |  | (наименование дисциплины) |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Обучающийся** |  | **Шипилов Евгений Александрович** |  |  |
|  |  | (ФИО) |  | (подпись) |
| **Группа** |  | **ВБИо-206рсоб** |  |
|  |  |  |  |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Преподаватель** |  | **Сибирев Иван Валерьевич** |  |  |
|  |  | (ФИО) |  | (подпись) |

**Москва 2025 г.**

**Отчёт по лабораторной работе №3**

**Задание 1**

Моделирование Информационной системы сервиса автоматизированной стенограммы ООО “**EchoText”** (сервис автоматизированной стенограммы)

**Задание 2**

Компания ООО “**EchoText”** специализируется на преобразовании аудио- и видеоконтента в текстовый формат с использованием искусственного интеллекта (AI) и обработки естественного языка (NLP). Отчёт формирует основу для разработки ключевых компонентов компании, включая управление, технологии и стратегию развития.

**Общие сведения о компании.**

**Сфера деятельности:**

**Основной продукт:** SaaS-платформа для автоматической и ручной транскрипции.

**Дополнительные услуги:** Анализ тональности речи, суммаризация текста, интеграция с CRM и медиаплатформами.

**Миссия:**

**«Превращать речь в ценность через точность, скорость и безопасность цифровых решений».**

**Общий штат сотрудников: 22 человека**

**Способы ведения бизнеса (Анализ рынка и стратегия):**

**-Основные конкуренты**

|  |  |
| --- | --- |
| **Компания** | **Преимущества перед** ООО “**EchoText”** |
| Otter.ai | Поддержка редких языков, ручная проверка. |
| Rev.com | Более низкая цена для малого бизнеса. |

**Конкурентная стратегия:**

**-Фокус на B2B-нишах: Юридические и медицинские транскрипции.**

**-Гибридная модель: AI + человеческий контроль для точности 99.9%.**

**Цели компании:**

**На 1 год**

1. **Запуск MVP с поддержкой 3 языков.**
2. **Привлечение 100 корпоративных клиентов.**
3. **Внедрение CI/CD для ускорения релизов.**
4. **Сертификация ISO 27001 (безопасность данных).**
5. **Партнёрство с 2 университетами для тестирования продукта.**

**На 3 года**

1. **Покрытие 90% рынка юридической транскрипции в ЕС.**
2. **Разработка мобильного приложения с оффлайн-режимом.**
3. **Снижение стоимости обработки на 40% за счёт оптимизации моделей.**
4. **Интеграция с Microsoft Teams и Slack.**
5. **Выход на оборот $5 млн.**

**На 5 лет**

1. **Запуск real-time стенограммы для live-ивентов.**
2. **Внедрение blockchain для аудита транскриптов.**
3. **Достижение нулевого углеродного следа.**
4. **Покрытие 50 языков, включая редкие диалекты.**
5. **Статус «единорога» (оценка $1 млрд).**

**Организационная структура компании** ООО “**EchoText”**

**Уровни иерархии:**

**Уровень 1: Стратегическое руководство.**

**Уровень 2: Руководители направлений.**

**Уровень 3: Исполнители.**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Должность** | **Подразделение** | **Уровень** | **Кол-во штатных единиц** | **Основные обязанности** |
| Генеральный директор | Руководство | 1 | 1 | Стратегическое управление, координация отделов, развитие бизнеса. |
| Руководитель R&D | R&D (AI/NLP) | 2 | 1 | Управление разработкой AI-моделей, интеграция NLP, оптимизация алгоритмов. |
| Data Scientist | R&D (AI/NLP) | 3 | 4 | Обучение и тестирование моделей распознавания речи, работа с датасетами. |
| ML Engineer | R&D (AI/NLP) | 3 | 2 | Развертывание моделей в production, настройка ML-пайплайнов. |
| Руководитель IT-инфраструктуры | Техническая поддержка | 2 | 1 | Управление облачной инфраструктурой, безопасность, DevOps. |
| DevOps Engineer | Техническая поддержка | 3 | 2 | Настройка CI/CD, мониторинг серверов, автоматизация развертывания. |
| Backend Developer | Техническая поддержка | 3 | 1 | Разработка API, интеграция с внешними сервисами (Zoom, Teams). |
| Руководитель лингвистического отдела | Лингвистический отдел | 2 | 1 | Контроль качества транскрипций, управление командой редакторов. |
| Лингвист-редактор | Лингвистический отдел | 3 | 4 | Ручная проверка и коррекция текстов, локализация контента. |
| Руководитель отдела продаж | Маркетинг и продажи | 2 | 1 | Продвижение продукта, заключение B2B-контрактов, аналитика рынка. |
| Менеджер по продажам | Маркетинг и продажи | 3 | 2 | Работа с клиентами, проведение демо-сессий, поддержка партнеров. |
| Финансовый директор | Административный отдел | 2 | 1 | Управление бюджетом, финансовое планирование, отчетность. |
| Юрист | Административный отдел | 3 | 1 | Юридическое сопровождение, GDPR-комплаенс, договоры с клиентами. |

**Схема:Изображение выглядит как текст, диаграмма, План, снимок экрана

Контент, сгенерированный ИИ, может содержать ошибки.**

**Техническая архитектура компании** ООО “**EchoText”**

**Пояснения:**

1. **Гибридная обработка данных:**

**- Аудиофайлы обрабатываются AI-движком, а результаты проверяются лингвистами через внутренний интерфейс (не включен в таблицу).**

**- Для критичных задач (медицинские/юридические транскрипции) используется ручная постобработка.**

1. **Масштабируемость:**

**- Kubernetes позволяет автоматически добавлять ресурсы при пиковых нагрузках (например, обработка 10k файлов в час).**

1. **Безопасность:**

**- Данные пользователей шифруются на всех этапах (E2E).**

**- Регулярные аудиты соответствия стандартам GDPR и HIPAA.**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Уровень архитектуры** | **Компонент** | **Технологии/Инструменты** | **Описание** | **Назначение** |
| **Инфраструктура** | Облачные сервисы | AWS EC2, S3, Lambda; Google Cloud Storage | Виртуальные машины для обработки данных, хранилище для аудио/текстов. | Масштабируемость, отказоустойчивость, снижение затрат на физическое оборудование. |
|  | ML-кластеры | Kubernetes, Docker, NVIDIA DGX | Кластеры для обучения и инференса моделей распознавания речи. | Параллельная обработка больших объемов данных, ускорение обучения моделей. |
|  | Серверы предобработки | FFmpeg, SoX (Sound eXchange) | Конвертация аудио/видео в стандартизированные форматы (WAV, MP3). | Устранение шумов, нормализация громкости, разделение каналов. |
| **Обработка данных** | AI-движок | OpenAI Whisper, PyTorch, Hugging Face | Нейросети для транскрипции, суммаризации, анализа тональности. | Преобразование речи в текст с контекстной обработкой. |
|  | Конвейер данных | Apache Kafka, Airflow | Оркестрация потоков данных между этапами обработки. | Управление очередями задач, мониторинг задержек. |
| **Хранение данных** | Основные базы данных | PostgreSQL (метаданные), MongoDB (логи) | Учетные записи пользователей, история транзакций, аудио-логи. | Быстрый доступ к структурированным и неструктурированным данным. |
|  | Кэширование | Redis, Memcached | Временное хранение промежуточных результатов транскрипции. | Снижение нагрузки на серверы, ускорение отклика. |
| **Интеграции** | API-шлюз | FastAPI, GraphQL, Swagger | RESTful API для интеграции с внешними системами. | Подключение клиентов к платформе через единую точку входа. |
|  | Партнерские интеграции | Zoom API, Microsoft Teams SDK, Slack Webhooks | Автоматическая загрузка записей с платформ для расшифровки. | Упрощение workflow для корпоративных клиентов. |
| **Клиентский интерфейс** | Веб-приложение | React.js, WebAssembly, Web Audio API | Интерфейс для загрузки файлов, редактирования текста, экспорта результатов. | Кроссплатформенный доступ, работа с аудио в браузере. |
|  | Мобильное приложение | Flutter, TensorFlow Lite | Оффлайн-транскрипция, синхронизация с облаком. | Расшифровка записей без интернета (базовые функции). |
| **Безопасность** | Шифрование | AES-256, TLS 1.3, Let’s Encrypt | Защита данных при передаче и хранении. | Предотвращение утечек, соответствие GDPR и HIPAA. |
|  | Аутентификация | OAuth 2.0, JWT, Auth0 | Управление доступом пользователей и ролями (RBAC). | Защита от несанкционированного доступа. |
|  | Резервирование | AWS Glacier, RAID-массивы | Архивация данных и резервное копирование. | Восстановление информации при сбоях. |
| **Мониторинг и аналитика** | Логирование | ELK-стек (Elasticsearch, Logstash, Kibana) | Сбор и визуализация логов серверов и приложений. | Выявление аномалий, анализ производительности. |
|  | Производительность | Prometheus, Grafana | Мониторинг загрузки CPU/GPU, времени отклика API. | Оптимизация ресурсов, предотвращение downtime. |
| **DevOps** | CI/CD | GitLab CI/CD, ArgoCD | Автоматизация тестирования, сборки и деплоя. | Ускорение вывода новых версий продукта. |

**Системная архитектура компании** ООО “**EchoText”**

**Пояснения:**

**Взаимосвязи уровней:**

**- Технологическая платформа обеспечивает ресурсы для работы сервисов.**

**- Сервисы используют модули для реализации бизнес-логики.**

**- Интерфейсы предоставляют точки взаимодействия с внешними системами и пользователями.**

**Специфика VerbaFlow:**

**- Гибридная обработка: AI + ручная проверка (модуль контроля качества).**

**- Поддержка live-трансляций через WebSocket.**

**- Фокус на B2B-интеграции (Zoom, Teams).**

|  |  |  |  |
| --- | --- | --- | --- |
| **Уровень** | **Компонент** | **Технологии/Инструменты** | **Описание** |
| **1. Технологическая платформа** | Облачная инфраструктура | AWS, Google Cloud | Базовый слой для хранения, вычислений и масштабирования. |
|  | ML-инфраструктура | Kubernetes, Docker, NVIDIA GPU | Оркестрация контейнеров, обучение и инференс AI-моделей. |
|  | Обработка аудио/видео | FFmpeg, SoX, Web Audio API | Конвертация, нормализация и предобработка медиафайлов. |
|  | Базы данных и кэши | PostgreSQL, MongoDB, Redis | Хранение метаданных, аудио-логов, кэширование промежуточных результатов. |
| **2. Сервисы** | AI-сервис | OpenAI Whisper, PyTorch, Hugging Face | Распознавание речи, суммаризация текста, анализ тональности. |
|  | Сервис управления данными | Apache Kafka, Airflow | Оркестрация потоков данных, управление очередями задач. |
|  | Сервис интеграций | FastAPI, GraphQL, Zoom API | RESTful API для внешних систем, интеграция с платформами (Zoom, Teams). |
|  | Сервис безопасности | AES-256, OAuth 2.0, Auth0 | Шифрование данных, аутентификация, RBAC. |
|  | Клиентский сервис | React.js, Flutter, WebAssembly | Веб- и мобильные интерфейсы для загрузки файлов и работы с транскриптами. |
|  | DevOps-сервис | GitLab CI/CD, Prometheus, Grafana | Автоматизация сборки, деплоя и мониторинга. |
| **3. Модули** | Модуль транскрипции | Whisper (базовая модель) + дообученные слои | Преобразование речи в текст с поддержкой 20+ языков. |
|  | Модуль контроля качества | Внутренний интерфейс для лингвистов | Ручная проверка транскрипций, исправление ошибок AI. |
|  | Модуль аналитики | Elasticsearch, Kibana | Визуализация метрик (точность, скорость обработки, пользовательская активность). |
|  | Модуль уведомлений | Webhooks, SMTP | Оповещение пользователей о завершении задач, email-рассылки. |
|  | Модуль резервного копирования | AWS S3 Glacier, RAID | Архивация данных, восстановление при сбоях. |
| **4. Интерфейсы** | REST API | Swagger, Postman | Внешнее взаимодействие: загрузка файлов, получение транскриптов. |
|  | WebSocket | Socket.IO | Стриминговая обработка аудио в реальном времени для live-трансляций. |
|  | Интеграционные интерфейсы | Zoom SDK, Microsoft Teams API | Прямая загрузка записей конференций из платформ. |
|  | Веб-интерфейс | React.js + Redux | Управление профилем, настройками, просмотр истории транскрипций. |
|  | Мобильный интерфейс | Flutter | Оффлайн-режим, синхронизация с облаком. |
|  | Админ-панель | Django Admin | Управление пользователями, аудит безопасности, аналитика. |

**Схема:**

**Изображение выглядит как текст, диаграмма, Шрифт, линия

Контент, сгенерированный ИИ, может содержать ошибки.**

**Бизнес-архитектура в представлении ментальной карты**

**Изображение выглядит как текст, снимок экрана, диаграмма, круг

Контент, сгенерированный ИИ, может содержать ошибки.**

**Связи между критически важными бизнес-процессами компании** ООО “**EchoText”**

|  |  |  |  |
| --- | --- | --- | --- |
| **Процесс 1** | **Процесс 2** | **Тип связи** | **Описание взаимодействия** |
| **Разработка AI-моделей** | **Обработка данных** | Технологическая зависимость | AI-модели используются для автоматической транскрипции аудио/видео. |
| **Обработка данных** | **Клиентский опыт** | Влияние на качество | Точность транскрипции влияет на удовлетворенность клиентов. |
| **Продажи и маркетинг** | **Интеграции** | Расширение возможностей | Интеграция с Zoom/Teams повышает ценность продукта для B2-сегмента. |
| **Интеграции** | **Клиентский опыт** | Удобство использования | Прямая загрузка записей из платформ упрощает workflow клиентов. |
| **Безопасность данных** | **Обработка данных** | Обеспечение соответствия | Шифрование данных гарантирует конфиденциальность при обработке. |
| **Масштабирование** | **Разработка AI-моделей** | Ресурсная зависимость | Оптимизация облачных затрат позволяет выделять больше ресурсов на R&D. |
| **Продажи и маркетинг** | **Клиентский опыт** | Обратная связь | Данные NPS-опросов помогают корректировать маркетинговые стратегии. |
| **Безопасность данных** | **Интеграции** | Риск-менеджмент | Сертификация GDPR необходима для работы с европейскими клиентами через API. |
| **Масштабирование** | **Безопасность данных** | Устойчивость | Внедрение «зеленых» технологий снижает риски, связанные с регуляторными требованиями. |
| **Разработка AI-моделей** | **Масштабирование** | Инновации | Локализация моделей под новые языки открывает доступ к рынкам Азии и Европы. |
| **Клиентский опыт** | **Продажи и маркетинг** | Рост лояльности | Улучшение CX увеличивает LTV клиентов и снижает CAC. |
| **Обработка данных** | **Масштабирование** | Оптимизация | Автоскейлинг Kubernetes позволяет обрабатывать больше запросов без потери скорости. |