

Арсен Кенесбаев

Data Engineer | Python & SQL

- +7 707 446 6793
- Казахстан, Алматы
- arsen801777@gmail.com
- Telegram



Live-портфолио



linkedin

О себе

Data Engineer со специализацией на SQL, Python и автоматизации. Имею опыт проектирования RAG-экосистем, разработки нормализованных банковских БД и построения ETL-пайплайнов. Владею инструментами контейнеризации (Docker) и оркестрации (n8n). Автор интерактивного портфолио, демонстрирующего навыки full-cycle разработки данных.

Навыки

Core Data Engineering: SQL, Python, PostgreSQL, Oracle Database, ETL/ELT, PL/SQL
Data Processing & AI: Pandas, NumPy, Qdrant, PGVector, RAG-системы, Scikit-learn
Инфраструктура и Автоматизация: n8n, Docker, Linux, Zabbix, Grafana, Git/GitHub
Инструменты и Web: Flask, FastAPI, Waitress, Ngrok, Selenium, HTML/CSS/JavaScript

Проекты

RAG-экосистема с интеграцией в браузер Tech Stack: Python, Qdrant (Vector DB), Docker, n8n, JS, Google Drive API, JSON/MIME. <ul style="list-style-type: none">Разработал браузерное расширение для внедрения RAG-контекста в Google AI Studio; реализовал сложную валидацию MIME-типов и парсинг JSON, чтобы API распознавал внедренные данные как нативные файлы.Внедрил рекурсивную суммаризацию (chain summarization) и триггер «Thought Cleaning» для удаления внутренних логов (CoT), что снизило потребление токенов на ~70% и предотвратило переполнение контекстного окна.Построил локальный пайплайн данных: команды в Telegram запускают загрузку данных из Google Drive в Docker-контейнер с Qdrant через кастомный скрипт автоматизации.	09/2025 - 11/2025
Веб-сайт системы рекомендаций фильмов Tech Stack: Flask, TMDb API, Scikit-learn, TF-IDF, Cosine Similarity, Matplotlib, Waitress, Ngrok <ul style="list-style-type: none">Создал систему рекомендаций фильмов на основе контентной фильтрации (TF-IDF + косинусное сходство) и ассистента в стиле Акинатора.Обработал датасет из 5,000+ фильмов TMDb, внедрил адаптивную фильтрацию и интегрировал командное портфолио.Развернул приложение через Waitress + Ngrok, реализовав аутентификацию пользователей и переключение тем (Dark/Light mode).	11/2024 - 12/2024
База данных онлайн-банкинга (Oracle APEX) Tech Stack: Oracle APEX, SQL, Java, Miro, Canva <ul style="list-style-type: none">Спроектировал базу данных (в стиле Kaspi) с нормализованными схемами для клиентов, карт, кредитов и транзакций.Написал 30+ оптимизированных запросов (joins, подзапросы, оконные функции), добавил триггеры и индексы для ускорения работы.Сгенерировал тестовые данные с помощью Mockaroo и Java; визуализировал структуру и презентовал проект в Miro и Canva.	10/2023 - 12/2023

Образование

Сулейман Демирель Университет(SDU) Бакалавр, Компьютерные науки	Каскелен, Казахстан 09/2022 - 05/2026
Сертификаты: <ul style="list-style-type: none">Google Data Analytics(ссылка)PostgreSQL(Michigan University)<ul style="list-style-type: none">HCIA Datacom - HuaweiSelenium на Python - StepikPython Backend - Skylearn(TechOrda)	Языки: <ul style="list-style-type: none">Казахский (родной)Русский (свободное владение)Английский (Advanced, ielts 7.5)

Опыт работы

KazNetCom DevOps-инженер (стажировка) <ul style="list-style-type: none">Развернул систему видеонаблюдения Frigate в Docker, перенес обработку видео на GPU (NVIDIA CUDA) для максимальной эффективности сервера.Создал решение для мониторинга на стеке Zabbix/Grafana с real-time дашбордами для контроля состояния ЦОД, задержек и доступности сервисов.Интегрировал Zabbix с Telegram для создания автоматизированной системы оповещений о сбоях и восстановлении работы базовых станций.Разработал кастомную GIS-платформу с уникальной интеграцией draw.io, позволяющей инженерам редактировать схемы сети прямо на интерактивной карте.	Алматы, Казахстан 08/2025 - 9/2025
Агентство РК по регулированию и развитию финансового рынка Инженер по автоматизации и данным (стажировка) <ul style="list-style-type: none">Спроектировал и внедрил две end-to-end системы на n8n: конвейер генерации отчетов (GLPI/Bitrix24) и RAG-систему для AI-ассистента.Разработал полный цикл RAG: от сбора данных (веб-скрейпинг) до векторизации (Embeddings) и хранения в PostgreSQL/PGVector.Реализовал сложную интеграцию с GLPI и Bitrix24, включая реверс-инжиниринг API, обогащение данных и извлечение информации из .docx с помощью LLM (Ollama).Обеспечил отказоустойчивость асинхронных процессов, внедрив механизм внешних блокировок (distributed locks) в PostgreSQL для решения проблемы состояний гонки (race conditions).Спроектировал PULL-архитектуру для обхода сетевых ограничений, блокировавших стандартные PUSH-вебхуки.	Алматы, Казахстан 06/2025 - 08/2025