Karim Jibai

Beirut, Lebanon | karim.jibai002@gmail.com | (+961) 76-414803
LinkedIn: https://www.linkedin.com/in/karim-jibai001/ | **GitHub**: https://github.com/GitDevKarim

PROFESSIONAL SUMMARY

Data & Backend Engineer with 3+ years of experience building scalable data pipelines, cloud architectures, and AI-powered systems. Proficient in Python, Go, SQL, and Azure, with hands-on expertise in RAG pipelines, vector databases, and real-time API development. Proven ability to deliver production-grade solutions across remote and on-site environments. Available across EU and US-compatible time zones (UTC+2/3).

TECHNICAL SKILLS

- **Languages**: Python, SQL, Golang (Go), Bash, Java, C, JavaScript/Typescript, HTML, CSS
- **Frameworks & Tools**: Git, Docker, Linux, Microsoft Azure, Google Cloud Platform (GCP), Microsoft Fabric, Redis, Cloudflare, Apache NiFi, Apache Spark, N8n, Flask, Django, FastAPI, Node.js, ExpressJS
- **AI & Data Intelligence**: scikit-learn, OpenAI API, Claude Code, LangChain, LlamaIndex, spaCy, NLTK, NumPy, Pandas, RAG systems, AI Agents, Prompt Engineering
- **Vector & Search Databases**: Chroma, FAISS, Pinecone, Azure Cognitive Search, MongoDB, Apache Solr
- **Databases**: Microsoft SQL Server, MySQL, Oracle DB, Azure SQL

PROFESSIONAL EXPERIENCE

AI Economic Pulse, Beirut, Lebanon **September 2024 – Present**

Data Engineer

- Built a hybrid RAG-based deduplication engine using OpenAI embeddings and vector databases (FAISS, Pinecone), improving semantic matching and ML feature quality.
- Architected and maintained end-to-end data pipelines using Azure Data Factory, Databricks, and ADLS, processing high-volume daily datasets at scale.
- Developed and deployed Python-based web scrapers and API integrations (including Google APIs) on Azure Functions for automated ingestion from multiple data sources.
- Designed data workflows to enhance dataset quality and contextual retrieval accuracy, improving downstream model performance.

Software Engineer (Backend & Cloud)

- Designed and implemented a modular backend architecture in Go, organizing APIs by domain (analytics, ML insights) for scalability and maintainability.
- Built and maintained 20+ production-grade REST APIs serving real-time analytics and pipeline outputs.
- Integrated backend services with Azure SQL and Data Lake to support high-throughput data access for analytics and downstream API consumption.
- Engineered cloud architecture using Azure Front Door, Cloudflare (DNS, CDN, security), and Redis caching, optimizing performance and routing, reducing average API response latency by ~40%.
- Implemented observability using OpenTelemetry and Application Insights and deployed containerized backend services using Docker on Azure.
- Engineered backend systems powering data-intensive analytics platforms, integrating ML outputs and large-scale data pipelines into low-latency APIs for real-time consumption.

Forestack, Remote, Contract **Early 2026**

Data Engineer

- Developed end-to-end data pipelines in Microsoft Fabric integrating CRM systems and external APIs, enabling automated ingestion and near real-time synchronization of business data.
- Designed and implemented transformation workflows to clean, structure, and standardize data for downstream analytics and reporting.
- Orchestrated data ingestion, transformation, and storage processes within a unified data platform, ensuring consistency and scalability.
- Automated pipeline operations and monitoring, improving reliability and reducing manual intervention in data workflows.

SOFTECH, Beirut, Lebanon **November 2022 – September 2024**
Data Engineer

- Designed and implemented database schemas and entity-relationship diagrams (ERDs) to establish a scalable data foundation.
- Developed data-migration and ingestion pipelines with Apache NiFi, moving legacy data to new systems and integrating continuous ingestion from REST APIs.
- Built a Flask-based monitoring service that tracked pipeline health through NiFi's API and issued real-time alerts on failure or delay.
- Optimized Python connectors linking SQL and NoSQL databases, improving integration reliability and accelerating ETL throughput.

Makerspace Engineering Solutions, Remote, Contract **2024**
Software Engineer [GitHub]

- Engineered scalable backend architecture supporting real-time scheduling and coordination of emergency-response meetings, integrating hardware components with Google Cloud services.
- Built webhook-driven automations using Twilio to process SMS triggers from hardware buttons and synchronize scheduling with Google Calendar API for instant meeting creation and notifications.

EDUCATION

Lebanese American University, Beirut, Lebanon
Bachelor of Science, Computer Science