

EGE BULUT

USKUDAR / ISTANBUL

+90 507 526 2527 ✉ egebulut.iletisim@outlook.com [in linkedin.com/in/EgeBULUT](https://www.linkedin.com/in/EgeBULUT)

github.com/Ege-BULUT github.com/simplyegebulut

Who Am I?

Results driven **AI Solution Architect and R&D Team Lead** specializing in the end-to-end design and deployment of **enterprise grade AI ecosystems** and **SaaS microservices**. Proven expertise in architecting **privacy first RAG solutions**, **scalable LLM integrations**, and complex **multi-agent orchestrations** that automate critical workflows and drive significant operational efficiency. A strategic technical lead adept at bridging the gap between advanced **Gen AI research** and production-ready, **data driven software architectures**.

Experience

AI Solution Architect

.QIT (qit.codes)

May 2026 – Present

Remote

- Conducted contract-based **AI consulting** and **product development** for **Randemix AI, Converter, and Orchestrator platform**, directly automating complex Turkish customs clearance workflows.
- Developed an **AI normalization service** combining **RAG** and **Fuzzy Search (RapidFuzz)** to fix noisy inputs, misspellings, and OCR errors, translating unstructured data into valid, EVRIM-compatible outputs.
- Integrated **Milvus and Chroma vector databases** to achieve high-accuracy **semantic matching** across customs regimes.
- Executed **QLoRA fine-tuning** on domain-specific customs datasets, successfully boosting matching accuracy from an **initial 13% to a production-grade 96%**.
- Eliminated manual classification workflows by building **automated document processing pipelines** for e-invoices, emails, and directory-based file inputs, **speeding up the process by 1200%**.
- Architected and deployed these solutions within secure, **on-premise** enterprise environments to strictly comply with **local data privacy** constraints.

AI & R&D Team Lead

Rail-Flow GmbH

Jun 2024 – May 2026

Remote

- **Led and mentored** a cross-functional R&D **AI team** of 6 engineers, coordinating product and operations to deliver **3+ enterprise-grade AI** production systems.
- Delivered **privacy-first, on-prem RAG solutions** utilizing **Vector DBs (Cassandra, Supabase, OpenSearch)** and **multi-modal pipelines (OCR, NLP, Image Processing)** to safely process **5k+ customer documents daily**.
- **Containerized and orchestrated core AI microservices** using **Docker** and **Kubernetes**, deploying **Linux-based CI/CD pipelines** on Oracle Cloud (OCI) to **improve release speed by 50%**.
- Engineered **scalable inference stacks** and **deployment pipelines** using **FastAPI** and **Pydantic**, managing hybrid (cloud and on-prem) environments with **OpenAI, Meta, Google, xAI, and Ollama**.
- Architected a **unified AI orchestration environment** leveraging **LangChain** and **LangGraph** to build complex, reliable **multi-agent chain** flows tailored for high-volume logistics customers.
- Developed custom **MCPs (Model Context Protocol)** to connect **LLM agents** directly with 5+ core company products, streamlining data flow and **accelerating cross-app automated workflows by 120%**.

Senior AI Engineer

Simply Deliver

Jan 2023 – Jun 2024

Hybrid

- Engineered **3+ custom internal AI tools** to eliminate dependencies on third-party vendors, **reducing** corporate operational software **licensing costs by 80%**.
- Designed and architected scalable, cloud-based RAG systems on Oracle Cloud and Azure, leveraging **LangChain** and **vLLM** to achieve a **10x improvement in LLM query response times**.
- Implemented production-grade AI workflows using both **cloud APIs and local models (Ollama, LMStudio)**, cutting **API token spend by 78% through smart caching** and local fallback routing.
- **Optimized data extraction processes** using Streamlit and PDFMiner.six, increasing data extraction **accuracy to 95%+** across unstructured multi-page logistics documents.

QA Engineer (Temporary)

Simply Deliver

Oct 2022 – Jan 2023

Remote

- Carried hybrid responsibilities across **QA and R&D** during a critical team transition, maintaining **100% on-time delivery** for both sprint-based testing and feature rollouts.
- Developed and improved a new staging/testing environment for the Automatic Order Entry (AOE) team, improved testing speeds **200%**.
- Executed **40+ end-to-end (E2E) testing scenarios**, including **Sanity Checks, Smoke Tests, and ACP/APEX LCT framework tests** to ensure zero-defect deployments.
- Managed QA workflows using **Jira** and **Xray**, creating 15+ structured test plans that **reduced production bugs by 20%**.

Artificial Intelligence Engineer

Simply Deliver

Jun 2022 – Nov 2022

Hybrid

- Developed and deployed initial production-ready **NLP** and **LLM**-powered features using **Python** and **LangChain**, contributing to a **25% increase** in user automation engagement.
- Built robust **data ingestion pipelines** that integrated with **Vector Databases (ChromaDB, Qdrant)**, improving **semantic search** query retrieval speeds by **80%** for internal datasets.
- **Fine-tuned open-source LLMs** (via **Ollama** and **Hugging Face**) for specific **classification** and text extraction tasks, achieving an **92% accuracy** rate on unstructured text data.
- **Containerize AI microservices** using **Docker**, reducing local environment setup time for new features from days to minutes.
- Integrated **API based solutions (OpenAI, Google, Ollama, HuggingFace, VLLM, LMStudio)** into existing software architectures, optimizing API call structures to manage token efficiency by 75% and halving the latency.

Software Developer Intern

Simply Deliver

Jan 2022 – May 2022

Hybrid

- Acquired deep technical proficiency in **Oracle APEX** and **PL/SQL**, developing **database-driven web applications**
- Successfully closed **80+** development and QA **tickets** during the internship.
- Engineered automated document processing workflows utilizing **Regex** to parse and **extract structured data from unstructured** logistics related **documents**.

Projects

Rail-Flow AI Engine | *Project Leader*

2024 – 2026

- Built AI engine reducing processing time from 4 hours to 1 minute (**240x improvement**).
- Scaled from solo project to 3-person team; delivered demo in 3 months.
- Implemented fully local AI with custom + Ollama models to ensure privacy.
- Contributed to **10% growth** in active client base.
- Tech Stack: **LangChain, LangGraph, LangFuse, PostgreSQL, Milvus, embedchain, CSS, HTML, Angular, Restful APIS, SaaS, MicroServices, CI/CD, Python, Agent Orchestration**

EKS-Smart EdTech Web Platform | *Lead Developer*

2021 – 2023

- Built an entire Education Technologies Platform that provides cutting-edge Educational AI Products
- Built web-pages that supports interactive test types (voice analysis, handwriting supports etc)
- Created web-pages displaying AI Tutors, live scoring boards, Student analysis report pages with interactive charts
- Tech Stack: **Angular, Nginx, PostgreSQL, Milvus, TypeScript, JavaScript, SCSS, HTML, Restful API**

SimplyCo PDF Interpreter & Dataset Tool | *Project Leader*

2022 – 2023

- Saved **€100K annually** and improved response time from 24s to 4s.
- Led end-to-end development from architecture to web dashboard, **CI/CD and DevOps**.
- Enabled **600% faster** dataset preparation for AI training.
- Tech Stack: **Python, Angular, OCR, PostgreSQL, TypeScript, JavaScript, CSS, HTML, Restful API,**

Relevant Coursework

- AI Team Leadership
- LangChain, LangFlow, LangGraph
- **Ollama**, Chains, Agents, Memory
- **RAG**, LLM, NLP, OCR, AOP
- **Generative AI**, ComfyUI, LORA
- SQLAlchemy, PostgreSQL, PLSQL
- MSSQL, MySQL, Oracle DB
- **MongoDB**, Firebase, **Vector** DBs
- **Cassandra**, Supabase, OpenSearch
- Local and Cloud AI Models
- **Agent Oriented Programming**
- FastAPI, Pydantic, Hugging Face
- Streamlit, Flask, Django, API
- Llama, GPT, Mistral, Deepseek
- SaaS, **RESTful**, Microservices

Technical Skills

Languages: Python, C#, Java, .NET, Typescript, JavaScript, SQL, Angular, NodeJS

Technologies: VectorDB, RAG, Linux, PostgreSQL, Windows & Ubuntu Servers, Docker, Kubernetes, LangChain

Education

Manisa Celal Bayar University

B.Sc. Computer Engineering

Sep 2018 – May 2023

Manisa