

Uğur Özcan

Senior Data Engineer | Tech Lead | AI & Backend Specialist

Mugla/TURKEY
+90 (534) 364 4085
ugurozcan108@gmail.com
<https://uozcan12.github.io>

Professional Summary

Strategic and results-driven **Senior Data Engineer & Technical Lead** with 7+ years of experience in architecting scalable data ecosystems and AI-driven solutions. Proven expertise in **optimizing Big Data pipelines (Spark)**, achieving up to **58% performance gains**, and leading cross-functional teams in high-stakes environments like **Huawei**. Specialized in **LLM fine-tuning (SFT)**, advanced **Prompt Engineering**, and complex **Data Governance** across global infrastructures. A former **Founding Engineer** with a unique blend of deep technical mastery in **3D simulation** and a strong track record of mentoring engineering teams to deliver mission-critical software.

Work Experience & Internships

- | | | |
|-------------------|---|-------------------------|
| 07/2025 - Present | Huawei
<i>Senior Data Engineer / Tech Lead</i> | Remote, Turkey |
| | <ul style="list-style-type: none">Technical Leadership: Responsible for the technical direction of the team, enforcing best practices, and overseeing codebase review processes to ensure high code quality and consistency.Performance Optimization: Architected improvements for TPP Spark and File System codebases, achieving a 58.9% reduction in execution time (from 908s to 373s) through parallel processing.Data Governance: Engineered end-to-end ETL pipelines for MENA projects, integrating data from ElasticSearch, MongoDB, and SAP into Agent Databases.Project Management: Successfully delivered 10+ high-priority crawling projects and developed an internal Kanban System (DTWF) for workload tracking.Multilingual Support: Managed TPP/PP configurations and dataset acquisition across 7 languages including Arabic, Russian, Portuguese, Spanish, French, Turkish, and Swahili.System Development: Developed a Django-based service for the CV (Computer Vision) team and delivered a full-stack mini-service (FE-BE integration) for the CCTV Environment project.Training: Provided professional Django Training sessions to multiple cross-functional teams within the SADC department. | |
| 01/2023 - 07/2025 | Huawei
<i>Data Engineer</i> | Remote, Turkey |
| | <ul style="list-style-type: none">LLM Project: Directed SFT data preparation (Summarization, Brainstorming) and performed advanced Prompt Engineering for GPT-3.5 and GPT-4.Data Engineering: Spearheaded large-scale image/text dataset acquisition by managing VMs and Docker containers for the Petal Search ecosystem.Internal Tools: Developed 'Common Crawler v1.0' and 'Text Preprocessing Tool', significantly increasing team-wide ETL and data analysis efficiency.NLP Initiatives: Orchestrated textual data aggregation and categorization to support natural language processing and flower data template initiatives.System Maintenance: Managed automated crawling pipelines, addressed critical bug fixes, and integrated new content providers for Search services.DevOps & Quality: Implemented codebase review processes and managed team repositories to ensure high code quality and seamless version control. | |
| 11/2022 - 01/2023 | Quick and Quality Coding Soft
<i>Software Engineer</i> | Remote, Turkey |
| | <ul style="list-style-type: none">NLP Project: Developed an automated analysis system for medical notes using Python, OpenAI API, and HuggingFace. | |
| 06/2022 - 11/2022 | Osoci
<i>Backend Developer</i> | Istanbul, Turkey |

- Infrastructure: Architected backend services and database structures using **NestJS, Docker Compose, and MongoDB**.
 - Media Services: Researched and integrated real-time stream services utilizing **Janus WebRTC** and Ant Media Server.
 - DevOps: Established CI/CD pipelines via **Jenkins** and implemented system monitoring with **Grafana**.
- 06/2019 - 06/2022 **Fited** **İzmir, Turkey**
Founding Engineer & Software Engineer
- 3D Simulation: Developed a web-based Scoliosis Brace design platform using **ThreeJS** and complex **Mathematical Equations**.
 - Cloud Architecture: Managed full-stack infrastructure on **Google Cloud Platform (GCP)** using App Engine, Compute Engine, and Firebase.
 - R&D: Conducted reverse engineering on **Rhinoceros7/Grasshopper** using NURBS Python and rhino3dm libraries for 3D modeling automation.
 - Workflow: Built automated testing suites with **Selenium** and developed backend interfaces using **Flask**.
- 10/2017 - 06/2018 **IZTECH - CSIGB Research Group** **İzmir, Turkey**
Msc. Student, Software Engineer
- Computer Vision: Developed deep learning models for Automated Latent Fingerprint Recognition (TUBITAK 1001) using **PyTorch and TensorFlow**.
 - Biometrics Research: Conducted advanced research on Face, Iris, and **Gait Recognition** algorithms using OpenCV.
 - Data Analysis: Generated and analyzed synthetic biometric databases for AFIS (Automated Fingerprint Identification System).
- 06/2016 - 09/2016 **Radboud University** **Nijmegen, Netherlands**
Erasmus Internship Programme
- **Big Data Analysis:** Processed and analyzed over 1GB of social media (Twitter) data focusing on complex topics like public health and sociology.
 - **Machine Learning:** Developed and deployed Classification and Clustering algorithms on the university's high-performance servers.
 - **Backend Development:** Engineered RESTful APIs using **Flask API** and **Eve API** to serve project data.
 - **Database Management:** Designed and managed the project's data storage layers using **MongoDB**.
 - **Full-Stack Contribution:** Developed the Relevancer Project dashboard and implemented core content features.
 - Project Links: relevancer.science.ru.nl/about | bitbucket.org/hurrial/relevancer
- 06/2015 - 08/2015 **Computer Engineering Department** **Muğla, Turkey**
Internship
- **Cybersecurity & Privacy:** Researched database encryption and security protocols using **CryptDB** on Ubuntu environments.
 - **Software Development:** Designed and implemented Graphical User Interfaces (GUI) using **NetBeans** and Java.

Skills, Qualifications & Interests

Languages	Python, JavaScript (Node.js), TypeScript, Java, SQL, R, Haskell
AI & Data	LLM (GPT-4), NLP, PyTorch, TensorFlow, Keras, Pandas, NumPy, Scikit-Learn, Computer Vision
Backend & Cloud	Docker, Kubernetes, Jenkins, GCP (Google Cloud), MongoDB, PostgreSQL, Redis, Flask, FastAPI, NestJS, Django
Web & 3D	ThreeJS, React/Vue.js familiarity, HTML5/CSS3, WebGL
Tools & OS	Linux/Ubuntu, Git, LaTeX, Grafana, Selenium, Rhino7/Grasshopper