

# **Yahya Obeid**

Santa Clara, California, US | +1 (408) 591-2747 | [yobeid2002@gmail.com](mailto:yobeid2002@gmail.com)  
[github.com/yahyaobeid/](https://github.com/yahyaobeid/) | <https://www.linkedin.com/in/yahyaobeid/>

## **PROFESSIONAL SUMMARY**

Software Engineer with 3+ years of experience in backend and AI-driven systems, specializing in Python, Django, JavaScript, and scalable API development. Built production-ready tools at Oracle and PerfAI, improving system performance, data flow efficiency, and API security. Experienced in relational databases, backend frameworks(Django,Flask), and modern UI frameworks (React/Angular). Passionate about event-driven architectures (Kafka) and leveraging AI productivity tools to accelerate development.

## **WORK EXPERIENCE**

### **AI Software Engineer, Oracle (Contract)**

**JUN 2025 – PRESENT**

- Architected and developed the backend API for AuditAIQ, an internal AI compliance platform that ingests and compares policy documents against operational artifacts (Confluence pages, Jira exports, and data files such as CSV, JSON, and XLSX) to detect policy violations. Designed distributed task pipelines using Django REST Framework and Celery, transforming manual, multi-day audit reviews into automated analyses completed in under an hour.
- Designed and implemented a full multi-model extraction pipeline integrating Llama Maverick and Grok LLMs into AuditAIQ's backend for large-scale document analysis. Developed intelligent chunking and preprocessing logic to handle long text inputs, prompt-based structured extraction templates, and schema validation to generate normalized JSON outputs. Optimized inference accuracy and latency, reducing processing failures by 70% across high-volume audit datasets.
- Deployed RabbitMQ-based task queues for parallel LLM inference and background data normalization, improving throughput for large-scale audit datasets.

### **Software Engineer, PerfAI**

**AUG 2023 – MAY 2025**

- Developed a full-stack automated API testing platform using Next.js and NestJS, enabling customers to generate and execute large-scale functional and regression tests across internal APIs. Designed validation workflows for schema consistency, response integrity, and latency tracking, improving data reliability and boosting user engagement by 15%.
- Built APIPrivacy, a NestJS-based microservice that detects and reports sensitive data exposures (PII, PCI, HIPAA) across internal APIs. Designed real-time scanning pipelines and schema-based validation to flag non-compliant responses, generating automated alerts and remediation reports that strengthened data security and compliance posture across production environments.
- Trained and deployed endpoint classification models on Vertex AI using TensorFlow and Scikit-Learn to detect APIs handling regulated data (PII, PCI, HIPAA) and governance operations (e.g., sorting, access control). Integrated the models into PerfAI's core platform to automate endpoint categorization, improving detection accuracy and reducing manual audit workload for engineering and compliance teams.

### **Software Engineer Intern, Z2Data**

**OCT 2020 - OCT 2021**

- Led the development of automated, responsive email templates powering internal and client-facing communications at Z2Data. Built modular layouts and dynamic placeholders for personalized content using HTML, CSS, and JavaScript, and integrated with internal REST APIs and SendGrid for scalable deployment. Streamlined communication workflows, reducing manual update effort and accelerating campaign turnaround time.
- Contributed to the enhancement and maintenance of Z2Data's production website built with ASP.NET MVC and SQL Server, resolving application and UI issues that affected data visualization and client dashboards. Refactored legacy code and optimized queries, improving page load performance by 30%.
- Pioneered the development of a **RESTful API** using **Python**, **Django REST Framework**, and **PostgreSQL** to support Z2Data's help page, enabling faster and more accurate part searches. Optimized database queries and serializer logic to improve search latency and data handling efficiency, significantly enhancing user experience.

## SKILLS

**Languages:** Python, JavaScript, Typescript, Java, SQL, HTML, CSS

**Technologies:** Django, RabbitMQ, Celery, NextJS, Rancher, Node.js, React, Flask, FastAPI, NestJS, TensorFlow,

Scikit-learn, Vertex AI, LangChain, OpenAI APIs, Pydantic, Pandas, NumPy, LangGraph, Kafka, Postgres, SQLite

**Cloud & Devops:** Docker, Kubernetes, GCP, AWS (EC2, S3, Lambda), REST APIs, CI/CD (GitHub Actions), NGINX,

Redis

## EDUCATION

**Bachelors, Computer Science, San Francisco State University, 2023, 3.7**

## PROJECTS

### <https://triphelix.com>, AI Agents for Travel Assistance

- Architected and built Triphelix, an AI-driven travel assistant that generates context-aware itineraries using curated system prompts and LangGraph for agent orchestration. Implemented real-time itinerary updates and low-latency streaming via WebSockets and Next.js, backed by a Django API for seamless data management.
- Architecting the next phase of Triphelix's AI agent ecosystem to perform autonomous searches for flights, hotels, and car rentals, leveraging multi-agent coordination and real-time data APIs to identify the cheapest, location-optimized options. Integrating geospatial filtering and interest-based embeddings to recommend destinations and accommodations that match user preferences and travel patterns.
- Hosted Triphelix on an AWS EC2 server running Ubuntu, orchestrating deployment through Docker and docker-compose for seamless container management. Configured an NGINX reverse proxy for routing between backend and frontend services, implemented SSL/TLS certificates for secure HTTPS traffic, and optimized server configuration for production scalability and uptime.