

# Waleed Ahmed

*Applied AI Engineer*

waleedahmed.ca  
hello@waleedahmed.ca  
linkedin.com/in/waleed-ahmed  
github.com/TRAP33ZOID  
Toronto, ON, Canada  
437-450-9195

## RELEVANT EXPERIENCE

**Founder** — *ZoidAI* — September 2025 – Present

- Architected and deployed a production multi-tenant SaaS platform serving dental clinics with AI voice agents, implementing **PostgreSQL** with **Row-Level Security** policies to enforce complete tenant isolation and prevent cross-clinic data leakage
- Built end-to-end **RAG pipeline** with **pgvector** for semantic search, handling document ingestion through **LangChain** text splitters, embedding generation, and retrieval to power both voice and text-based AI agents with domain-specific knowledge
- Engineered dual knowledge base synchronization system that maintains consistency between **VAPI** phone infrastructure and internal vector database, processing **hundreds of documents** across multiple tenants with real-time updates
- Implemented **Stripe** webhook handlers and subscription management system with usage-based billing, integrating payment lifecycle events with tenant provisioning and credit allocation workflows
- Deployed to production on **Vercel** with serverless architecture, managing API routes, edge middleware for authentication, and real-time **WebSocket** connections for voice agent testing

**AI Consultant** — *Zoid Consulting* — May 2024 – Present

- Consulted directly with small business owners to identify operational pain points, architected custom AI automation solutions, and delivered production voice agents and chatbots tailored to specific business workflows
- Built production-grade AI voice agents using **VAPI**, **Bland AI**, and **Twilio**, integrating with client scheduling systems, CRMs, and phone infrastructure to automate appointment booking and customer inquiries
- Designed and deployed domain-specific **RAG knowledge bases** for each client, processing business documentation, FAQs, and policies to ensure agents provided accurate, context-aware responses
- Owned full customer lifecycle from initial outreach and technical discovery through solution architecture, deployment, and ongoing iteration based on real-world usage feedback
- Managed infrastructure costs, client billing, and profitability while maintaining **multiple concurrent deployments** across different business verticals including healthcare and professional services

## TECHNICAL PROJECTS

**Enterprise RAG Evaluation Pipeline** — *Python, LangChain, Qdrant, FastAPI*

- Built production-ready RAG system with **Qdrant** vector database, implementing hybrid search combining semantic similarity and keyword matching to improve retrieval accuracy by **30%**
- Developed comprehensive evaluation framework measuring retrieval precision, answer faithfulness, and context relevance using **RAGAS** metrics and synthetic test dataset generation
- Implemented chunking strategy optimization pipeline, testing multiple approaches (recursive, semantic, fixed-size) and embedding models (**OpenAI**, **Cohere**) to maximize retrieval quality
- Exposed system via **FastAPI** endpoints with request validation, rate limiting, and monitoring, deployed with **Docker** for consistent runtime across development and production environments

## PROFESSIONAL SUMMARY

Applied AI engineer with hands-on experience building and deploying production agentic systems for enterprise customers. Built multi-tenant SaaS platform from scratch with RAG pipelines, vector databases, and real-time voice agent infrastructure. Direct consulting experience architecting custom AI solutions for small businesses, handling technical discovery, deployment, and iteration. Strong foundation in modern AI stack including agent orchestration, fine-tuning workflows, and production deployment practices. Comfortable working directly with customers to translate business requirements into technical solutions. Passionate about shipping reliable AI systems that solve real problems.

## SKILLS

### Languages & Frameworks

Python, TypeScript, JavaScript, React, Next.js, FastAPI, Flask, Node.js, SQL

### Vector Databases & Search

pgvector, Qdrant, Pinecone, Weaviate, Semantic Search, Hybrid Search, Embedding Models

### Databases & Infrastructure

PostgreSQL, Supabase, Redis, Row-Level Security, Database Design, Query Optimization

### DevOps & Deployment

Docker, Docker Compose, Vercel, AWS (S3, Lambda), Git, GitHub Actions, CI/CD, Serverless Architecture

### Enterprise Integration

REST APIs, WebSockets, Webhooks, OAuth, Multi-tenancy, Stripe Integration, Twilio, VAPI

### Tools & Platforms

VAPI.ai, Bland AI, Tableau, Jupyter, VS Code, Linux, RAGAS, LangSmith

## EDUCATION

**Bachelor of Computer Engineering** — *Toronto Metropolitan University* — May 2026  
**Diploma of Data Science** — *Lighthouse Labs* — Oct 2023