

WALEED SHOAIB

Islamabad, Pakistan · Pakistani National

waleedshoaib20@gmail.com · +92 332 577 6168

linkedin.com/in/waleedshoaibai · github.com/waleedshoaib2 · waleedshoaibai.substack.com

Summary

CTO and AI Engineer with **4+ years** building and deploying **multi-modal AI systems** in production across **healthcare, financial services, and computer vision**. Lead a team shipping 3 enterprise products serving clients across the **Middle East and USA**. Core expertise in **LLMs, agentic workflows, RAG pipelines**, vector/graph databases, and cloud-native deployment on **GCP and AWS**. IEEE-published researcher focused on AI systems that improve with use, not just at demo.

Experience

Aeyron

Lewes, DE, USA & Islamabad, PK

Chief Technology Officer

Apr 2026 – Present

- Own **architecture, technical roadmap, and end-to-end delivery** for a team shipping 3 production AI products (healthcare documentation, bilingual document intelligence, financial analysis) to enterprise clients across the Middle East and USA.
- Architected and deployed **YOLO-based computer vision systems** for real-time detection, managing the full lifecycle: data annotation strategy, model training, optimization, and low-latency inference.
- Designed and scaled **RAG architectures** combining vector search with structured and unstructured data, integrated into **agentic workflows** for intelligent automation, validation, and decision-making.
- Led **model fine-tuning and domain adaptation** across vision models and LLMs, improving accuracy for specialized enterprise applications in healthcare and finance.
- Established **monitoring and observability** (Prometheus, Grafana), reducing inference latency and ensuring robust production deployment.

Senior AI Engineer

May 2025 – Apr 2026

- Architected end-to-end **LLM-based agentic systems** using LlamaIndex and ReAct, enabling autonomous financial analysis and decision-making.
- Built a **production-grade forex intelligence pipeline** ingesting market and news data into ClickHouse, generating real-time summaries and trend signals with **sub-500ms latency**.
- Designed scalable **ETL and aggregation pipelines** for macroeconomic indicators, supporting continuous updates and low-latency querying.
- Implemented **multi-agent collaboration** (ingestion, reasoning, validation), improving analysis accuracy and throughput by **35%**.
- Integrated **RAG pipelines** with Milvus and FinBERT for contextual retrieval and sentiment-driven insights.

Catalic AI

Remote

AI Engineer

Mar 2025 – May 2025

- Designed **ETL pipelines** for large-scale **FOMC datasets**, reducing end-to-end processing time by **10%**.
- Built a **research automation agent** (LlamaIndex + ReAct) with long-term semantic memory and Milvus search, deployed on GCP via Terraform and ArgoCD, increasing research efficiency by **15%**.

Objex

Montreal, Canada

Software Engineer

Mar 2024 – Feb 2025

- Developed a **cloud-native, microservices-based platform** on GCP; integrated **RAG components** using Firestore, improving search and recommendation by **40%**.
- Optimized **GraphQL APIs**, reducing query latency by **25%**; built event-driven pipelines using Google Cloud Pub/Sub.

Eastlancer

Islamabad, PK

Software Engineer

Mar 2022 – Feb 2024

- Built **scalable backend APIs** (Node.js, Express); integrated AI-assisted services and third-party APIs; improved performance through caching and load balancing.

Selected Projects

Receiptly AI: Bilingual **English/Arabic** receipt and PDF extraction using **Google Gemini + PaddleOCR PP-StructureV3**; custom JSON schemas, GPU deployment. Serving Middle East enterprise clients.

Cornerstone AI: Three-stage **RAG** pipeline that learns clinician writing styles and generates patient-specific **DOCX** reports; async services on LanceDB, Dockerized to GCP with CI/CD.

GeoAsset: Multi-sensor geospatial asset detection: satellite (**YOLO-OBB + SAHI**), street (**YOLO-World** open-vocab), and LiDAR; fuses OSM data, exports to GeoJSON/PostGIS.

DataLadder: Retail-shelf panorama stitching via **EfficientLoFTR** matching and multiband blending; async FastAPI/S3 service, **~48× faster** exposure compensation.

SteriSmart: Blender/FreeCAD pipeline generating synthetic multi-view datasets of orthopedic surgical instruments for CV model training.

Publications

W. Shoaib, B. Ismail, and M. Salik, “Point-Cloud-Native L-Band Augmentation of SWOT KaRIn: Recovering Canopy-Shadowed Dark Water with NISAR,” *IEEE Geoscience and Remote Sensing Letters*, 2026. [DOI]

- First combined use of **NISAR L-band SAR** and **SWOT** pixel clouds; recovers dark-water pixels missed under canopy across four wetland regimes.
- **+11.3 pp** recall lift on Congo, **+17.9 pp** on Amazon; disjoint 95% bootstrap CIs vs. baseline.

W. Shoaib, “Why Production AI Fails on Day 90 — and What Evaluator-First Architecture Fixes,” *Whitepaper*, 2026.

- Proposes an **evaluator-first** architecture where evaluation agents are built before generation agents, enabling production AI systems that **self-correct and improve** post-deployment instead of degrading.

B. Ismail and **W. Shoaib**, “Semantic-Conditioned Spectral Fusion for Label-Efficient Multispectral Adaptation of Frozen Remote-Sensing Vision–Language Models,” *IEEE Geoscience and Remote Sensing Letters*, 2026. (*Under Review*) [DOI]

- Parameter-efficient module (~0.1M params) adapting frozen RGB remote-sensing VLMs to multispectral imagery via FiLM-conditioned spectral fusion.
- **+4.6% overall accuracy** on EuroSAT-MS; **+1.7% mAP** on BigEarthNet-v2, concentrated in the low-label regime.

Technical Skills

AI & LLMs	Python, LangChain, LlamaIndex, ReAct, Reflexion, RAG, DeepEval, HALU Eval
Computer Vision	YOLO, Object Detection, Real-time Inference, Fine-tuning, Domain Adaptation, OpenCV
NLP & ML	spaCy, FinBERT, BioBERT, Transformers, NER, Semantic Search, Embeddings, Vector Similarity
Data & Retrieval	Milvus, Neo4j, Memgraph, ClickHouse, Redis, PostgreSQL, PGVector, Firebase
Cloud & Infra	GCP, AWS, Kubernetes, Terraform, Docker, CI/CD, ArgoCD, Prometheus, Grafana
Backend	FastAPI, REST, GraphQL, Express, Nest JS

Education

COMSATS University Islamabad

Wah Cantt, Punjab, Pakistan

Bachelor of Science in Computer Science — CGPA: 3.33 / 4.0

Certifications

Neo4j Certified Professional (Nov 2025) · **Docker Foundations Professional** · **Microservices Foundations Professional**