

ANSHAY AGARWAL

AI R&D Engineer | Founder | anshayagr@gmail.com | +91-8800472674 | [LinkedIn](#) | [GitHub](#)

R&D-FOCUSED AI ENGINEER

Principal AI Architect with 10+ years specializing in Low-Latency Autonomous Agents and Real-Time Signal Processing. Ex-Qualcomm/Nvidia Lead Engineer who scaled XR performance by 96%. Expert in deploying production-grade RAG pipelines and Voice-First AI interfaces.

Stack: Chat & Voice AI Agents, RAG Pipelines, WebSockets (Real-time Audio)

EXPERIENCE

STEALTH AI VENTURE | Remote | Lead AI Systems Architect | AI, Python | Oct 2025 - Jan 2026

- **Engineered a <500ms latency Voice-to-Action pipeline** by orchestrating raw PCM streams via WebSockets, bypassing standard API bottlenecks.
- **Built a fault-tolerant State Machine** for multi-turn conversations, integrating Gemini 2.5 with custom Voice Activity Detection (VAD) locks to prevent hallucination during silence.
- **Deployed scalable containerized microservices** handling concurrent audio streams with 99.9% uptime.

SCHOOLBOOK.AI | Co-Founder & Lead AI Architect | AI, Python | Jun 2025 - Sep 2025

- **Architected a Multimodal RAG Pipeline:** Built a high-performance retrieval system integrating Vector Databases (Embeddings) with Gemini Pro to process unstructured text (NCERT) into structured, context-aware outputs.
- **Engineered Autonomous Agentic Workflows:** Orchestrated multi-step reasoning chains using LangGraph and Python, enabling the AI to autonomously execute tool-use behaviors and dynamic planning.
- **Full-Stack GenAI Deployment:** Designed the end-to-end system orchestration, managing backend services and frontend integration for a low-latency interactive AI response loop.

QUALCOMM | Senior Lead Software Engineer | AI, Python, C++ | Apr 2023 - Mar 2025

- **Optimized XR Rendering Pipeline:** Reduced GPU Flicker correction latency by **96% (35ms → 1.3ms)**, directly enabling 90Hz refresh rates for Premium Tier Snapdragon chipsets.
- Spearheaded the design of novel HDR flows for Premium Tier chips, redefining image quality benchmarks for upcoming hardware.

GREyscaleAI | Software Engineer | C++, Python, OpenCV | Feb 2021 - Mar 2023

- **Architected Multi-Modal Inspection Pipelines:** Synchronized high-throughput data streams from heterogeneous sensors (**X-Ray + RGB**), enabling real-time automated defect recognition.
- Designed a real-time embedded UI orchestrating communication between hardware sensors and user inputs.
- Led a complete codebase refactoring initiative that streamlined build processes.

NVIDIA | System Software Engineer | C++, Python | Aug 2017 - Jul 2018

- **Imaging Stack Optimization:** Enhanced the Tegra Camera imaging stack, resolving critical **Lens Shading** and **Bayer Demosaicing** artifacts for high-precision computer vision applications

DRDO | Scientist | C++, MATLAB, OpenCV | Aug 2013 - Jul 2015

- Engineered solutions for lens shading correction in Thermal Imagers (LREO/MREO), critical for night vision.
- Built a custom Raspberry Pi-based remote controller for Thermal Imagers, eliminating external procurement costs.

PROJECTS

Calox - AI-Native Personal Analytics Agent | *Flutter, Hive, Gemini*

- **Architected an Offline-First AI Platform:** Integrated a local NoSQL database (Hive) with cloud-based LLM inference for privacy-centric analytics.
- **Engineered an Autonomous Analyst:** Developed an agent to ingest unstructured logs and generate trend visualizations

Ai-Powered Youtube Video Summarizer | React Native, LangChain, Vector DB

- Developed a cross-platform mobile app (**Expo/React Native**) for podcast and video summarization using **LLMs and RAG**.
- Integrated vector databases for semantic search and managed full-stack deployment for real-time inference.

SKILLS

Languages: Python, C++, JavaScript

AI/ML: Agents (LangGraph), LLMs (Gemini, LangChain), RAG, Deep Learning (FastAI, TensorFlow), Computer Vision (OpenCV), ASR (Whisper)

Systems: Git, Linux, Embedded Systems, MQTT, Real-time Audio Processing

ACHIEVEMENTS

- **Top Ranks:** GATE (AIR 423, 99.9%), IIT-JEE (Top 1%), GRE (321/340 - 170/170 Quant)

EDUCATION

IIT Delhi | M.Tech, Computer Technology | 2015 - 2017

- **Dissertation:** "Thermal Video Stabilization" for night vision systems.
- **Project:** Brain-Computer Interface (BCI) using EEG motor imagery signals for cursor control.

IIT Mandi | B.Tech, Electrical Engineering | 2009 - 2013

- **Dissertation:** Automobile Collision Prevention System
- **Leadership:** Founding Member of Robotics Section; Built IIT Mandi's first robot.

SPECIALIZATIONS

- **Wharton School:** Entrepreneurship Specialization (2021-2022)
- **DeepLearning.AI:** TensorFlow Professional Cert (2020), Deep Learning Specialization (2019)