EKJOT SINGH

Email: ekjotmakhija@gmail.com LinkedIn: ekjot-singh-thefirst GitHub: ekjotsinghmakhija Portfolio: ekjot.me

PROFESSIONAL SUMMARY

Machine Learning Engineer and Systems Architect with expertise in building scalable AGI systems, distributed infrastructure, and high-performance ML pipelines. Proven track record in architecting deterministic reasoning frameworks and optimizing GPU utilization using PyTorch and CUDA. Adept at leading cross-functional teams and implementing Agile/Scrum methodologies to deliver open-source solutions like Tealbase and Metanthropic AI.

EDUCATION

Vellore Institute of Technology Bhopal, India
Bachelor of Technology - B.Tech, Computer Science *Sep 2024 – Sep 2028*

- **CGPA: 9.07 / 10.0**
- Specialized in Artificial Intelligence and Systems Engineering.
- Relevant Coursework: Distributed Systems, Advanced Algorithms, Operating Systems, Neural Networks.

EXPERIENCE

Metanthropic AI Remote
Founder & Lead Systems Architect *Nov 2025 – Present*

- Architecting an independent AGI research lab; pioneered a deterministic reasoning framework that reduces hallucination rates in LLMs by enforcing strict logic constraints during inference.
- Engineered high-throughput training pipelines using PyTorch and Custom CUDA kernels, optimizing GPU utilization for large-scale transformer models.
- Developed novel Intrinsic Safety scaling laws, bridging biological intuition with digital reasoning to create mathematically bounded AI systems.
- Spearheaded a distributed research team, establishing CI/CD workflows for reproducible ML experiments and model versioning.

Tealbase Remote
Founder & Principal Engineer *Aug 2025 – Present*

- Designed and implemented a high-performance, open-source BaaS (Backend-as-a-Service) mirroring Firebase's DX but built on the robustness of PostgreSQL and Go.
- Engineered a Change-Data-Capture (CDC) system to power realtime websocket subscriptions, handling thousands of concurrent connections with sub-100ms latency.
- Implemented an automated REST API generation engine that parses Postgres schemas to instantly provision typed endpoints, reducing boilerplate code by 90%.
- Architected enterprise-grade security middleware featuring JWT-based auth and row-level security (RLS) to ensure multi-tenant data isolation.

PROJECTS

The Digital Boardroom – *Python, FastAPI, Multi-Agent Systems* Dec 2025

- Engineered a multi-agent consensus engine orchestrating Blind Peer Reviews between GPT-4, Claude 3.5, and Gemini Pro to eliminate single-model stochasticity.
- Developed a Chairman synthesis module that aggregates divergent reasoning paths into a single high-fidelity output, improving reasoning accuracy on complex tasks.
- Optimized API orchestration latency using asynchronous task queues to handle parallel agent inferences.

Azhn – *TypeScript, Node.js, SVG* Aug 2025

- Developed a high-performance rendering engine that compiles HTML/CSS directly into scalable SVGs, eliminating the need for headless browser snapshots.
- Implemented server-side rendering optimizations for Open Graph image generation, reducing generation time by 40% compared to Puppeteer-based solutions.

Buttery-Auth – *Node.js, MongoDB, Docker* Jul 2025

- Engineered a containerized authentication microservice supporting OAuth 2.0, Magic Links, and secure session management.
- Implemented automated JWT rotation and secure cookie handling to mitigate XSS and CSRF attacks.

TECHNICAL SKILLS

Languages: Rust, Go, Python, C++, Java, TypeScript, SQL (PLpgSQL), Elixir
Core Competencies: Distributed Systems, Systems Architecture, High-Performance Computing, API Design
AI & ML: PyTorch, CUDA, TensorFlow, Transformers, Mechanistic Interpretability, Scikit-Learn
Infrastructure: Docker, Kubernetes, Redis, PostgreSQL, Nginx, Linux/Unix
Leadership & Soft Skills: Agile/Scrum, Team Leadership, Cross-functional Collaboration, Strategic Planning
Web Technologies: Next.js, React, Tailwind CSS, Framer Motion, WebSockets

PUBLICATIONS

Dataset Distillation for the Pre-Training Era (2025): Proposed Linear Gradient Matching to condense massive datasets, optimizing pre-training efficiency.

The Platonic Representation Hypothesis (2025): Research note analyzing the convergence of diverse foundation models towards shared latent representations.