

Emre Turan

Turangenesi.com • [LinkedIn](#) • [GitHub](#) • turangenesi@gmail.com

Summary

AI engineer (6+ yrs) building LLM agents, multi-agent systems, and production AI. Currently building Headroom, an agent-oversight firewall (LangGraph human-in-the-loop) and a paper showing that, past a point, more human review can make a system less safe. B.S. Statistics & Data Science, UCLA.

Skills

AI & Agentic Systems: Multi-Agent Orchestration, LangGraph, LangChain, MCP Integrations, Agentic Workflow Design, Tool Routing, Context Engineering, Agent Memory & State Management, Human-in-the-Loop System Design, LLM Evaluation & Observability, LLM Fine-tuning, Prompt Engineering

Engineering & Infrastructure: Python, C#, SQL, Testing (pytest), System Design, Event-Driven Architecture, FastAPI, REST APIs, Docker, CI/CD (GitHub Actions), AWS, GCP, Heroku, PostgreSQL, MongoDB, SQLite, Pinecone, PyTorch, Hugging Face Transformers

Work Experience

[AI Objectives Institute](#) | Lead AI Engineer

San Francisco • May 2024 - Jan 2026

- Built and deployed the open-source [Elicitation Bot](#) (sole engineer) — full-stack, multimodal WhatsApp deliberation agent (Twilio + GPT-4o + Whisper) for [Talk to the City's](#) national-scale public consultations, reaching thousands of students across 55 universities and 12+ international conferences (Search for Common Ground, Taiwan AI College Alliance, Notre Dame).
- Ran 4 months of weekly live sessions (10–20K messages/day) — sharded onboarding across 5 Twilio numbers to clear the per-number 1K-new-users/day cap (a day-one launch blocker), then lifted delivery success 96% → 98% by fixing Heroku rate-limit and cold-start failures and sizing horizontal/vertical scaling to the app's CPU/memory profile.
- Designed API-driven LLM evaluation pipelines for the '[Moral Learning](#)' project, automating multilingual survey generation and assessment of human values, translation quality, and alignment.

[Generative Alpha](#) | AI Developer & Data Engineer

San Francisco • Summer 2023

- Engineered an SEC API pipeline processing a decade of filings for US30 stocks, normalizing and transforming raw documents into structured training data for AI agents targeting quantitative trading and financial advisory on Wall Street.
- Fine-tuned GPT-4 on generated Q&A pairs derived from SEC filings to improve financial data comprehension and model performance.

[Exploind & Sentium](#) | Data Scientist & Quantitative Analyst

Los Angeles • 2021–2023

- Built and deployed full-stack data platforms (Python, PostgreSQL, TimescaleDB, AWS, Docker) with ETL pipelines, high-frequency data management, and scalable cloud infrastructure via Lambda and CloudFront.
- Fine-tuned GPT-4 with LangChain for NLP-based analysis pipelines, integrating human-in-the-loop techniques and automated insight generation.

Research & Projects

Headroom — agent-oversight firewall + calibration study • [Demo](#) • [GitHub](#) • [arXiv](#) — March 2026 – Present

- Showed, in a paper and replay-based eval, that agent oversight is an inverted-U: past a point, more human review makes a system *less* safe, because reviewers fatigue. Reframes oversight as a measurement problem — reviewer capacity and disagreement — not just classification.
- Framed the guard as selective classification under asymmetric cost — risk-coverage, AURC, and Neyman–Pearson on a 125-action adversarial set, against a measured Fleiss' $\kappa = 0.52$ reviewer-agreement floor.
- Engineered the runtime on a LangGraph state graph that pauses risky file / shell / git / deploy actions mid-task via `interrupt()` for resumable approval (layered rules → LLM risk scoring → human review), with an MCP server any agent can route tool calls through, append-only audit log, and live dashboard.

OperatorLock — server-enforced behavioral risk-control system for futures • [Demo](#) • [GitHub](#) — Jan – Mar 2026

- Built a system that hard-blocks impulsive trades at the API through nine non-overridable gates: post-exit cooldowns, per-bar tempo tokens, Renko zone gates, and a daily stop/limit. Most platforms only warn; this one won't let the order through.
- Architected an event-driven Flask + Firestore state machine, fed by six TradingView webhook streams and bridged to a live C# Rithmic (RAPI+) execution layer.

Education

University of California, Los Angeles (UCLA), CA

- B.S. Statistics and Data Science — graduated 2024