

SHUBHAM JOSHI

AI Engineer | San Francisco, CA | (202) 440-8445 | shubhamjoshi@gmail.com

[Tacet, live product](#) | [AgentDock demo](#) | [Live model](#) | [Portfolio](#) | github.com/shubhamjoshi | [linkedin.com/in/shubham-joshi1](https://www.linkedin.com/in/shubham-joshi1)

SUMMARY

AI engineer who builds intelligent applications on top of foundation models, ships them to real users, then makes them reliable enough to keep running. I am now building Tacet, my own company, which turns an expert's judgment into tested rules that AI agents can run. I design the agent behaviour, write the backend, build the interface, fine tune the model when the model is the problem, and write the evaluation that decides whether it works. In the last year: 508 commits and 92,448 lines across 11 public repositories, 8 of them live, and a fine tuned transformer on Hugging Face. I build the whole thing, then I try to break it, and I report what broke.

SKILLS

Agentic AI, LLMs and Emerging Technologies: Agentic AI systems and agent architectures; Large Language Models and foundation models (Anthropic Claude, OpenAI GPT and Codex, Google Gemini); designing agent behaviour end to end: prompt engineering, reasoning and planning, tool and function calling, multi agent orchestration, Model Context Protocol servers and clients, retrieval and context engineering, structured output, multi provider routing and cost control; developing with Cursor, Claude Code and Codex, and a working familiarity with the frontier of applied AI

Evaluation, Guardrails and AI Safety: Evaluation harnesses and benchmark design, prespecified protocols with frozen fixtures, scenario and regression testing, adversarial and red team testing including prompt injection defense, deny by default permissions, human in the loop approval gates, secure credential handling, audit trails, confidence gating and abstention, honest reporting of negative results, responsible AI

Machine Learning and Deep Learning: PyTorch, Transformers and Hugging Face, transformer fine tuning and public model deployment (ModernBERT), scikit-learn, LightGBM, XGBoost, leakage safe splits, calibration, drift monitoring, SHAP and LIME interpretability, error analysis, taxonomy and label design, model selection under cost and latency budgets

Applied Computer Science and Software Engineering: Applied computer science and programming: Python, TypeScript, JavaScript, SQL; Next.js, React, Node, FastAPI, PostgreSQL, Prisma, REST APIs, system design, streaming over server sent events, durable job queues, idempotency, crash recovery, concurrency control, observability, performance profiling, cost tracking, automated testing, Docker, Git, CI/CD, services deployed and operated in production on Railway, Chrome extension and Tauri packaging; comfortable embedded in an unfamiliar codebase taking on new tasks

Collaboration and Communication: Cross-functional by habit: I collaborate with engineers, designers and non technical stakeholders, partner directly with clients to identify problems worth solving and drive integrated solutions with measurable impact, contribute reusable tooling, support colleagues, and bring strong communication skills to documentation, presenting findings, teaching, mentorship and continuous improvement

EXPERIENCE

Founder, Tacet - San Francisco, CA

Oct 2026 - Present

- Founded Tacet and built it solo from a HackNation and ElevenLabs hackathon prototype into a live multi tenant product. It watches an expert work, predicts each decision before they make it, asks one question only when it cannot explain what they did, and compiles the answer into an executable rule in the expert's own words that AI agents can call over API or MCP and new hires are coached against.
- Made the learning provable rather than claimed: a plain Python engine with Bayesian thresholds where language models only translate, sealed tests whose predictions are SHA-256 committed before any human label, and a compile gate that rejects a rule unless it quotes the expert verbatim and reproduces their past decisions. On a sample ledger, one answered question took held back accuracy from 96 to 100 percent across 44 transactions.
- Shipped it in nine days: 234 commits, 42,000 lines and 481 tests, including cross tenant checks generated from every route and ten real browser journeys against live models, plus key based auth, personal data scrubbing and a daily model spend cap.

Analytics Consultant, Johns Hopkins University, Client: Alliance of Community Health Plans - Washington, DC

Jul 2025 - Jul 2026

- Led technical execution for a six person team on a national client engagement, building the Python and SQL pipeline that reconciled four inconsistent federal sources into a 138,263 row dataset, catching a twelvefold enrollment overcount before it reached a model, and presenting in person to leadership, who said it would change where the organization spent its resources.

Co-Founder and Vice President, Business Analytics and AI Organization, Johns Hopkins University - Washington, DC

Jan 2026 - Present

- Co-founded and lead a 100 member organization alongside my degree, delivering a 3 hour hands on workshop on agents and modern AI development to 35 registrants and publishing the guide open source.

Founder, Digital Marketing and Web Studio, self-employed - Nagoya, Japan

Jan 2021 - May 2025

- Ran a studio single handed for four years for international clients in a third language, owning scoping, pricing, delivery and reporting.

PROJECTS

AgentDock, Agent Platform That Takes Real Actions

[Live app](#) | [Demo](#) | [Pitch deck](#) | [Code](#)

- Designed how the agent thinks and built all of it solo: model backed planning compiling a plain language goal into a validated multi step workflow, a reasoning and tool call layer executing over the real Model Context Protocol wire, and four first party connectors for Gmail, Calendar, Docs and Search. 225 commits and 35,539 lines in two months, with 66 test files, a PostgreSQL job queue with crash recovery and idempotent external actions.
- Solved what actually blocks agents from touching a real account, which is trust rather than capability: permissions resolve and clamp server side so prompt text can request capability but never grant it, every outbound action passes an approval gate, credentials are encrypted and budgets capped, and I wrote the adversarial suite proving prompt injection through tool output cannot fire an unapproved send.

Vibe Curator, Generative AI Product in Three Form Factors

[Live app](#) | [Code and releases](#)

- Built a real time creative environment across 104 commits and 19,913 lines where four providers land in one editable scene model: OpenAI for images, Anthropic for shaders, Gemini Veo for video, ElevenLabs for music. Model written shaders clear a static guard and a real compile pass before they render, and it ships as a web app, a Chrome extension and a signed macOS build.

OpsPilot and RxCheck, Fine Tuned and Evaluated AI

[Live model](#) | [Live app](#) | [Research report](#)

- Fine tuned and published a ModernBERT ticket router on 23,747 leakage safe examples, diagnosing instead of scaling: baselines put TF-IDF at 44 and ModernBERT at 52.3 percent, so the labels were the ceiling, and rebuilding ten labels into seven queues lifted accuracy to 73.1 percent, shipped with abstention that auto routes 62.3 percent at 85.5 percent accuracy.
- RxCheck reviews 152,416 records where the database decides and the model only explains, passing 26 of 26 scenarios even with every external service failing, and its research report publishes the results that made me look bad, including an official benchmark it passed only 22 of 30.

EDUCATION

Johns Hopkins University, Carey Business School - MS, Business Analytics and Artificial Intelligence, STEM designated, Dean's Scholarship Aug 2026

Nagoya University of Commerce and Business - BBA in Management, **GPA 3.63** (WES evaluated), Dean's Scholar honors, Monash exchange 2025

Coursework: Generative AI, Practical Machine Learning, Sequential Decision Making, Responsible AI, Statistical Analysis. **Certifications:** Anthropic Claude 101, Anthropic AI Fluency, Azure AI-900. **Languages:** English, Japanese, Nepali, Hindi.