

ADHYA DAGAR

San Francisco, CA ♦ adhyadagar5@gmail.com ♦ [linkedin.com/in/adhya-dagar](https://www.linkedin.com/in/adhya-dagar) ♦ www.adhyadagar.me

Designing feedback loops that transform research into likable, usable, and reliable products and models, building the foundation of the agent-driven economy.

EDUCATION

Master of Science, Artificial Intelligence and Innovation May 2022 - May 2024
Carnegie Mellon University *Pittsburgh, PA*
Coursework: LLMs, AI Product & Service Design, ML in Production, Advanced NLP, Deep Learning

Bachelor of Technology, Computer Science Engineering July 2016 - May 2020
Vellore Institute of Technology *India*

EXPERIENCE

Product Manager, Model Evaluations, Adobe Firefly September 2025 - Present, *San Francisco, CA*

- **Leading Firefly Enterprise Evaluations** - Built human and automated evaluation systems for image, video, and audio generation models across video reframe, vectorization, image and video editing, and custom LoRAs models. Partnered with creative professionals and enterprise customers to define production-quality standards, translating expectations into measurable rubrics determining model readiness and launch decisions.
- **Driving cultural transformation toward evaluation-driven product development** – training PMs, standardizing eval study templates, launching enablement playbooks and embedded quality gates across Adobe’s generative AI workflows.
- **Scaling Agentic Evaluation & Improvement** – Built evaluation frameworks and test harnesses to assess and improve agentic capabilities across multimodal workflows.

Product Manager (Research & Evaluations), Pareto.AI March 2024 - August 2025, *San Francisco, CA*

- **Designed and deployed agent evaluation infrastructure** creating programmatically verifiable RL environments across 15+ professional domains (e.g., Data Science, Healthcare Operations, Finance) and 10+ OSS tools (e.g., OpenEMR, SuiteCRM, Metabase) to evaluate multi-step agent workflows, tool use, and complex reasoning, leading 20-member team to establish new standards for real-world agent performance assessment.
- **Designed a comprehensive legal eval framework** from proof of concept to deployment, alongside legal domain experts, strengthening frontier model legal reasoning across contract review, regulatory analysis, and case law.
- **Built an automated data quality engine** with a comprehensive evaluation taxonomy including LLM judges (instruction following, complexity), statistical checks (preference distribution, semantic similarity), and fraud detection reducing manual QA from 100% to 20%
- **Drove new frontier lab partnerships** by researching emerging evaluation needs, scoping pilots, and designing 10+ tailored proposals that converted into long-term collaborations and revenue growth for Pareto.

Product Manager, BitClass (Live cohort-based learning app) Jan 2022 - May 2022, *Bengaluru, IN*

- **Drove 170% user acquisition growth** from 78K to 210K monthly registrations through quick-checkout system that streamlined purchase flows, while launching short-form video content feature repurposed from recorded live sessions that boosted user retention by 60% across 1.3M user base
- **Optimized user experience through data-driven redesign** using A/B testing and conversion funnel analysis to improve course discovery, home screen navigation, and checkout flows, delivering measurable increases in user engagement and purchase conversion rates

PROJECTS

CAQA: Collaborative QA Generation: Developed three-agent framework using iterative optimization for low-resource medical QA synthesis, achieving 2x improvement on RQUGE metrics and passing Turing test validation.

ProjectArena: AI Management Tasks: Built evaluation framework for autonomous agents in project management workflows using Dockerized environments (Plane, Rocket.Chat), achieving 16.7% success rate across 30 multi-step instruction sets.

AWARDS & FELLOWSHIPS

Felicis AI Fellow: Part of Felicis Ventures’ 2024 AI Fellows cohort. Winner of Fellows Hackathon Prize; co-led winning team developing LLMs for gene discovery with mentorship from founders of Weights & Biases, Langchain, and Kaggle.

Corporate Startup Lab (CMU): Led AI strategy research on Fortune 500 AI adoption through CXO interviews and 200+ professional surveys. Published case studies on Shell-C3.ai's predictive maintenance of machinery using ML and architected solutions for Honda on their indoor navigation project.