

JULIAN MERCER

AI PRODUCT MANAGER

San Francisco, CA | (415) 555-0148 | julian.mercer@email.com | linkedin.com/in/julianmercerc

PROFESSIONAL SUMMARY

AI Product Manager with 8+ years of experience launching language-model and machine-learning products at OpenAI, Microsoft, and Anthropic. Connects user workflows, evaluation design, safety, latency, and cost to responsible adoption and durable business value.

PROFESSIONAL EXPERIENCE

Senior AI Product Manager OpenAI, San Francisco, CA

Jul 2022 - Present

- Owned strategy and launch for an enterprise language-model workflow adopted by 42,000 teams, reaching \$86M annualized revenue with governed rollout controls.
- Defined task-specific evaluations with Research and customers across quality, safety, latency, and cost, improving successful completion 14% at scale.
- Prioritized model, retrieval, and interface improvements from 18,000 failure traces, reducing unsupported-answer rate 31% across subsequent roadmap reviews.
- Launched admin controls and audit logs with Security and Legal, shortening regulated-enterprise review cycles 26% and resolving recurring deployment objections.
- Built usage and unit-economics dashboards by task and model, reducing inference cost per successful workflow 22% while preserving quality thresholds.

AI Product Manager Microsoft, Redmond, WA

May 2019 - Jun 2022

- Led machine-learning assistance in Microsoft 365 used by 3.8M monthly users, increasing accepted suggestions 19%.
- Designed human-in-the-loop fallback and confidence UX, reducing harmful automation escalations 37% and giving users a clear recovery path.
- Established slice-based model monitoring across language, tenant size, and task complexity, detecting four launch-blocking regressions for the next planning cycle.
- Partnered with Engineering on latency budgets and caching, improving p95 response time from 4.8 to 2.6 seconds.

Product Manager Anthropic, San Francisco, CA

Aug 2016 - Apr 2019

- Translated early customer research into requirements for grounded document analysis across legal and support workflows without weakening customer guardrails.
- Built evaluation rubrics and reviewer calibration that increased inter-rater agreement from 0.58 to 0.76.
- Ran controlled pilots with 24 design partners, improving weekly retained usage 33% over two iterations.
- Documented model limitations and escalation paths, improving pilot trust scores by 12 points and making support ownership explicit before launch.

CORE SKILLS

AI product strategy | LLM evaluation | Machine learning | RAG | Prompt design | Human-in-the-loop | Responsible AI | Model monitoring | Experimentation | Product analytics | APIs | Latency | Unit economics | Enterprise security | Roadmapping

EDUCATION

B.S. in Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA

CERTIFICATIONS

DeepLearning.AI Generative AI for Product Leaders | Microsoft Azure AI Engineer Associate | Pragmatic Institute Foundations