

Alex Anthony

San Francisco, CA | [Website](#) | [LinkedIn](#)

Staff AI Product Manager — 10+ years building AI agents, conversational search, LLM infrastructure, evals, and enterprise data platforms. Former AI researcher (Cycorp + Rutgers cognitive science / natural language semantics). I ship AI products with speed and quality.

SELECTED AI PRODUCT HIGHLIGHTS

AI contract data extraction (Ironclad)

Led re-architecture to adopt an LLM-first backend → ~80% F1 loss reduction; >90% online accept rate across hundreds of extraction types; scaled throughput from 10s to 10ks of documents per hour

Conversational Search (Ironclad)

Launched agentic search combining metadata, lexical, neural search, and LLM coordination for question-answering across full contract repositories and conversational record search in table views; replaced legacy keyword-only search.

Evals Platform (Ironclad)

Built evals operating model from scratch — reference-based, rubric judging, experiment tracking, launch-readiness → 80% internal adoption in 2 months.

LLM Infrastructure Platform (Indeed)

Prompt Toolkit and Evals Service from 0 → 1 → 1,000+ users; scaled GenAI Gateway to billions of requests and trillions of tokens per quarter.

NL→SQL Analytics Agent (Indeed)

Co-created a skunkworks analytics NL-to-SQL app “Spock” for answering natural language data questions; reached hundreds of users and spun out into a dedicated team.

EXPERIENCE

Ironclad — San Francisco, CA

Staff Product Manager, AI

Mar 2026–current

Senior Product Manager, AI

Aug 2025–Feb 2026

- Product lead for AI extraction agents, Conversational Search, evals platform, and human data. Partner with engineering, design, and GTM to turn ML capabilities into high quality enterprise product experiences.

Conversational Search & AI Agents

- Launched Ironclad's conversational search to GA — an agentic in-product AI assistant enabling QA across full contract repositories and conversational record search in table views.
- Architecture combines metadata, lexical, neural search, and LLM coordination to replace legacy keyword-only search.

Evals & AI Quality

- Established Ironclad's AI evals operating model end-to-end: diagnosed evaluation and human data as bottlenecks, aligned leadership on investment, and championed & delivered a platform evals + observability solution.
- Achieved 80% internal adoption in 2 months; created company-wide enablement program (training, talks, office hours) that helped 6+ feature teams implement offline evals and hill-climb on AI quality.
- The platform supports reference-based evals, rubric evals, experiment tracking, and launch-readiness practices.

AI Extraction & Human Data

- Led complete re-architecture of AI data extraction to an agentic, foundation model-powered backend — ~80% reduction in F1 loss, 90% online accept rate across 100+ extraction types.
- Built Human Data function from zero: hired and managed 4 Legal SMEs, secured six-figure quarterly budget, signed external labeling services, delivered 10K+ labeled documents and 100K+ labels.

Indeed — Austin, TX

Senior Technical Product Manager, LLM Infrastructure 2024–2025

- Led 0→1 development of Prompt Toolkit and Evaluation Service — an end-to-end platform for prompt engineering and LLM evaluations — driving adoption from zero to 1,000+ users in two quarters.
- Scaled GenAI Gateway to support hundreds of internal clients at billions of requests and trillions of tokens per quarter.
- Defined and executed build-vs-buy strategy, evaluating 20+ and implementing 5 third-party solutions to accelerate AI infrastructure maturity.

Senior Technical Product Manager, Data & Analytics 2023–2024

- Led migration of 1,000+ analytics ETLs from BigQuery to AWS (S3, Glue, Athena), saving >\$3M.

- Co-captained cross-organizational Data Mesh transformation and co-created an internal NL→SQL analytics copilot that retrieved high-quality prior queries, adapted them to user intent, and turned distributed data knowledge into reusable analytics workflows; reached hundreds of users and was funded/spun out into a dedicated team.

Product Manager, Data Engineering & Orchestration **2021–2022**

- Launched and scaled Chorus, a dependency-driven data orchestration platform supporting 10+ teams and 1,000+ dataset ETLs.
- Delivered high-impact analytical datasets supporting 7M+ queries from ~5,000 employees and 1,000+ product experiments.

Bloomberg – Remote

Product Manager, Enterprise Data License **2020–2021**

- PM for >\$1B ARR Data License product; direct report to Chief Data Officer.
- Designed and built a data discovery application; led internal adoption and sales efficacy campaign.

Cycorp – Austin, TX

Member of the Technical Staff **2016–2020**

- Led 0→1 creation of Constellation, productizing Cyc's semantic AI data integration → contributed to >\$4M new revenue.
- Solutions architecture lead for \$10M+ Fortune 100 partnership building AI-enabled product recommenders — owned end-to-end from pitch through delivery.
- Lead developer on AI application for detecting aggressive accounting estimates, deployed to hundreds of analysts at a leading global investment bank.

EDUCATION

Rutgers University – PhD (Philosophy) [ABD] **2010–2016**

- Researcher, instructor, Natural language semantics, logic, philosophy of language.
- Publication: ["Experience, evaluation, and faultless disagreement."](#) Inquiry (2016).
- Primary instructor for undergraduate lectures and seminars: Formal Logic, Philosophy of Cognitive Science, Philosophy of Literature.

Wesleyan University – BA (Philosophy) **2006–2010**

- High Honors
- Phi Beta Kappa
- Wise Prize for best thesis: "Dispositional predicates in context."