

Keegan Choudhury

AI Product Manager | Agent Workflows, LLM Evaluation & Developer Tools

Washington, DC Metro | Open to relocation: California, including the SF Bay Area

keeganarko@gmail.com | keegan-studio.com | linkedin.com/in/keeganchou

Product manager, 3+ years, two on Microsoft AI: Excel Formula Repair (public web launch), an LLM support assistant used daily by support engineers, and a co-authored KDD workshop LLM-evaluation benchmark.

EXPERIENCE

Capital One — Product Manager | August 2025–Present

- Won leadership approval for a two-week go-live delay after shadow testing exposed request-coverage gaps and old-versus-new response mismatches; led the Venture X card-portfolio migration to a platform serving 4.3M requests/month with no customer-facing outages.
- Leading discovery for 13-engineer Collections & Recoveries monitoring (SMS, email, mobile, web): shared alerts, remediation guidance, rollback workflows, and business, operations, engineering, and product views.
- Defined a reusable card-data API strategy for a planned TSYS replacement; validated test-card issuance.

Microsoft — Product Manager | July 2023–August 2025

- Partnered with user researchers to surface formula fixes that were syntactically valid but missed user intent; helped shape Formula Repair in the Copilot pane and the grid's right-click menu; contributed demos and leadership readouts for its public Excel for the web launch.
- Co-authored FoRepBench (KDD 2025 workshop): scenario design, eval criteria, and writing for a 618-example benchmark of four LLMs; GPT-4.1 execution match: 80% on synthetic vs. 41% on forum-sourced cases.
- Owned discovery and tailored demos for 10+ agent use cases; onboarded 3+ internal teams to agent-framework prototypes (persistent memory, reusable context, human-in-the-loop oversight), which they used to start prototyping conversational agents for information retrieval, teaching, and clinic intake.
- Led a five-person team building an LLM support assistant used daily by support engineers: graph-based documentation retrieval, LLM troubleshooting, and escalation to an engineer with relevant context.
- Co-built hand-graded evals of completion and cited documents and diagrams, verified by support engineers: ~95% of test cases passed without needing an engineer.
- Helped define its production escalation rules, weighing cost, case context, elapsed time, and troubleshooting attempts; modeled \$1.07M in annual productivity value for the business case from estimated hours saved.
- Co-inventor, published patent application US20250378320A1: agent-guided completion of structured artifacts.

SELECTED AI PROJECTS

Assembly — Creator | 2026–Present

- Solo-built Assembly with AI coding agents: a beta app whose coordinator delegates across Claude, Codex, and Grok agents on users' own subscriptions (TypeScript, MCP; 3,300+ tests).
- Blocked context compaction until each agent's conversation is saved and hash-verified (Claude, Codex).
- Built agent messaging with 8 delivery states that never overwrites a person's unsent draft.
- Added visual agent-activity views from tester feedback; testers built their own finance, email, and fitness agent integrations, and all five early testers returned without prompting.

EARLIER EXPERIENCE

Tesla — Technical Program Manager Intern | September–December 2022

- Automated vehicle-door part validation, raising line throughput ~23% (SQL, Postman, KPI dashboard).

Microsoft — Program Manager Intern, HoloLens 2 | May–August 2022

EDUCATION

Texas A&M University | B.S. Computer Science | Summa Cum Laude | GPA 3.94 | May 2023

SKILLS

AI: LLM evaluation (execution checks, LLM-as-a-judge), agent orchestration, graph-based retrieval

Technical: Python, SQL, PostgreSQL, TypeScript, React, Electron, MCP, Semantic Kernel

Product: User research, shadow testing and launch readiness, business cases, API strategy, roadmaps