

Ali A. Malik

Toronto, ON • hello@aliahmedmalik.ca • linkedin.com/in/aliahmedmalik • aliahmedmalik.ca

Summary

Senior product manager for AI and AI-assisted platforms, with 6 years taking products from 0 to 1: NLP analytics used by 700,000 people, LLM tooling that cut delivery 40%, and the evaluation layer frontier-AI teams use to trust their training data. At Sapien, owns strategy, roadmap, and success metrics for Proof of Quality, a rubric-graded layer where LLM judges and human experts score model and agent output, and has carried it **from first spec to live enterprise design partners**. Builds the first version in code, runs discovery and pilots directly with customers, and keeps engineering, design, and go-to-market on one plan.

Experience

Sapien

Senior Product Manager

Toronto, ON

October 2025 – Present

Human-data and evaluation infrastructure for frontier AI: expert annotation, RLHF, and rubric-based quality attestation.

- **Own Sapien's core product, Proof of Quality:** roadmap, evaluation framework, and success metrics for a rubric-graded layer where LLM judges and human experts score training data and agent outputs. Shaped V1 scope and wrote the launch specs.
- Ran **120+ customer discovery and go-to-market conversations** as a forward-deployed PM. Those conversations produced the V1 requirements, 3 pilot proposals for security firms including CertiK and Sherlock, and a **converted design partner** whose CEO chose Sapien over extending his existing tooling after a live demo.
- Proved the product with a live enterprise partner by re-evaluating its AI security auditor against 5 senior human auditors: **21 of 40 findings flagged as false positives in 77 minutes**, an auditable verdict on 98%, and an approved Phase 2 partnership.
- Designed and pressure-tested the pricing model in customer conversations, landing on **pay-for-the-proof pricing and unit economics** (\$1,000 per project, per-event fees, \$500 per report) so customers pay for verified quality instead of review hours.
- **Resolved 5 launch-blocking defects in the V1 spec** before engineering kickoff, including a quality metric defined two ways and AI review panels copying each other's votes; wrote the rubric guardrails engineering builds against.
- Built Sapien's **internal knowledge brain**: a RAG index over all company documents, a Graphify knowledge graph, and a custom agent that answers from it, so planning, strategy, and team updates draw on one accurate source of truth.
- **Ships code alongside engineers**: 6 pull requests in the production monorepo, including a 7-layer release gate (UI regression, accessibility, LLM-persona evaluation) that guards every PR. Prototyped the rater marketplace MVP solo.

Tempo (Y Combinator 2023)

AI Product Manager

Toronto, ON

December 2024 – October 2025

Design-to-code AI tooling that turns Figma files into production applications.

- Shipped **6 production web applications** on Tempo's in-house LLM code-generation and agent platform, cutting delivery 40%.
- Ran 20 concurrent client builds with 8 developers and 6 designers, **shipping 1 to 3 features weekly**; owned backlogs, sprint plans, and client communication, and integrated Stripe, BigCommerce, Odoo, and Sanity into modular build pipelines.

Surf (acquired by Datacy)

Senior Product Manager

Toronto, ON

January 2023 – December 2024

NLP and behavioral analytics SaaS; 700,000 users across consumer, business, and marketplace products.

- Led 2 agile teams across 3 products, owning the roadmap and KPIs for 34 sprints and **shipping 70 features**.
- **Grew active users 35%** through campaign experiments driven by behavioral analytics and predictive models.
- Managed enterprise accounts for Amazon, Electronic Arts, Logitech, and Riot Games at **\$1.5 million average deal size**.

Trufan (acquired by Surf)

Product Manager

Toronto, ON

September 2020 – December 2022

Social data and audience intelligence startup.

- Built and launched **2 consumer platforms serving 500,000 users** with a 6-person team, on NLP over social and web data.
- Delivered a live dashboard for **120,000 browser extension users** and data marketplace integrations with Snowflake and LiveRamp.

Education

London School of Economics and Political Science

London, United Kingdom

Bachelor of Science, Business

Skills

AI Evaluation: Evaluation frameworks, rubric and golden-set design, LLM-as-judge with human escalation, inter-rater reliability, RLHF.

AI Systems: Model selection, LLM and agent workflows, prompt engineering, RAG, MCP, guardrails, human oversight, Claude Code, OpenAI.

Product: Use-case selection, 0-to-1, roadmaps and OKRs, specifications, success metrics, pricing and unit economics, build-vs-buy.

Leadership: Executive and cross-functional alignment, design-partner pilots, customer discovery, forward-deployed GTM, Agile, Jira.

Technical: SQL, Snowflake, TypeScript, React, Next.js, Node.js, Playwright, Storybook, GitHub Actions, Supabase, Netlify, Vercel.

Domain: Web3 security (Solidity audits, staking, commit-reveal), computer vision annotation QA, eCommerce (Stripe, Odoo).

Speaking: Keynote on training and evaluating AI models, **AI Futurist Conference**, Toronto, July 2026.

Independent: Founder, Kleta (digital agency); building **AGNCY** and **Ustadh**, two SaaS products. Case studies: aliahmedmalik.ca/projects.