

Yusef Syed

yusefmsyed@gmail.com | github.com/YusefSyed | [Agent Eval](#) | [Product Portfolio](#)
Center Valley, PA and Toronto, ON | U.S.-Canadian dual citizen | No sponsorship required
Available full-time May - August 2027 in the U.S. or Canada

EDUCATION

University of Toronto

Toronto, ON

Faculty of Arts & Science - Mathematical & Physical Sciences; intended Mathematics and Statistics Expected May 2030

Relevant Coursework: Data Structures (dual enrollment), AP Computer Science A, AP Calculus BC, AP Physics C

AI EVALUATION & PRODUCT ENGINEERING

Agent Eval Mutation Lab | [GitHub](#) | *Python, SQLite, pytest, mypy*

Aug 2026

- Built a typed, resumable **Python 3.12+ evaluation engine** with workers isolated from oracle labels, transactional SQLite, content-addressed evidence, and canonical scheduling across **104 deterministic scorer tasks**.
- Preregistered and ran a **624-trial**, two-model/two-prompt/three-seed study; rejected the intervention after **5 of 6 frozen promotion gates failed**, released all outcomes, and did not retry invalid responses or retune post hoc.
- CI spans Python 3.12–3.14, enforces **81% branch coverage**, detects all **14 predeclared semantic mutations**, and reproduces **17 canonical artifacts byte-for-byte**.

Aesthetics AI | [App Store](#) | [Demo](#) | *React Native, Expo, Supabase, RevenueCat*

May 2026 - Present

- Shipped a subscription-enabled iOS product integrating RevenueCat with AI vision workflows; owned the React Native client, Supabase/PostgreSQL backend, and **8 Edge Functions**.
- Built schema-validated AI flows and a deterministic **36-scenario synthetic replay harness** covering representative production outputs, routing, and **14 safety cases**; protected evaluator integrity with checksum locks, offline network denial, and **4 mutation tests**.

Tiraz | [App Store](#) | [Demo](#) | *React Native, Expo, Supabase, Deno*

Jun 2026 - Present

- Released a privacy-first iOS wardrobe coach with AI outfit checks, closet recommendations, image cutouts, and subscriptions across React Native and a Supabase/Deno backend.
- Reproduced a historical private **150-image, 826-object** extraction evaluation (**F1 0.333**) and a separate 50-crop classifier result (**47/50; 0.94 macro accuracy**); preserved the failed product gate, with no private media or production-accuracy claim.

CallReclaim | [Demo](#) | [Case Study](#) | *Next.js, Supabase, Twilio, LLM APIs*

Jul 2026 - Present

- Designed fail-closed Twilio/LLM workflows with signed webhooks, STOP/HELP consent handling, deduplicated provider actions, owner-scoped access controls, and live-provider integrations disabled by default.
- Deployed the consented missed-call recovery pilot to Vercel after source-integrity, dependency, secret, smoke, and readiness checks; the system spans **40 API handlers** and **23 Supabase migrations**.

LEADERSHIP & COMMUNITY

Feeding the Valley | *Vice President*

2025 - 2026

- Led **60+ student volunteers**, raised **\$20K**, and helped prepare **1,700 meals** for local food banks.

FIRST Robotics Team 834 | *Fundraising Lead*

2022 - 2026

- Raised **\$5K+**, reached **2,000+ prospects**, and mentored **7+ members**; the team qualified twice for the FIRST World Championship.

TECHNICAL SKILLS

Languages: Python, TypeScript, JavaScript, SQL

Python & Testing: pytest, mypy, uv, SQLite, GitHub Actions

Product & Data: React, React Native, Expo, Next.js, Node.js, Supabase, PostgreSQL, Deno

APIs & Deployment: Structured LLM APIs, Twilio Voice/SMS, RevenueCat, EAS/TestFlight, Vercel