

CHRIS McNOSKY

AI EVALUATION, AGENT RELIABILITY & TECHNICAL PROGRAM MANAGEMENT

Dallas-Fort Worth, TX | 817-360-0687 | cmcnosky@gmail.com

linkedin.com/in/chris-mcnosky | github.com/cmcnosky | cmcnosky.github.io/why-hire

PROFESSIONAL SUMMARY

AI evaluation and technical program leader who turns ambiguous requirements into testable acceptance criteria, adversarial evaluations, reproducible evidence, and clear release decisions. Directs AI-assisted technical work while owning scope, architecture, quality gates, and final judgment. Earlier experience spans purchasing, vendor documentation, regulated operations, and deadline-driven coordination.

SELECTED IMPACT

- **Submitted a difficult Tokio macro change as a non-draft upstream PR.** Scoped and directed the AI-assisted Rust contribution for issue #3974 after two earlier contributor attempts closed. Derived the design contract from those failure modes; full upstream CI is green and the PR awaits maintainer review.
- **Held an AI-produced change out of integration despite 1,336 passing tests.** In my evaluation lab, a second, separately prompted adversarial review found two paths where acceptance logic could approve spec-violating behavior; correction and retest were required before integration.
- **Improved evaluator reliability while preserving the failure record.** When Stinger mislabeled a correct refusal, preserved the original evidence, specified the correct classification, and required regression coverage for six independently worded refusals before accepting the correction.

PROFESSIONAL EXPERIENCE

Founder / AI Evaluation & Governance - NoFuckery AI | 2026-Present

- Translate ambiguous objectives into specifications, acceptance criteria, failure modes, and clear proceed, correct, hold, or stop decisions.
- Direct AI-assisted implementation and separate review across Python and Rust systems; inspect diffs, logs, tests, reports, and source material before approval.
- Build and publish evidence-led systems and analyses, including Stinger and the Tokio contribution; preserve contradictory evidence and require correction plus regression testing.

Founder / Sole Operator - High Pie | October 2022-Present

- Own supplier selection, sourcing, production, pricing, inventory, regulated-product compliance documentation, fulfillment, and customer issue resolution for a hemp-product business.

Purchasing Coordinator - Lennar | May 2018-October 2022

- Reviewed bids, contracts, budgets, and vendor records; reconciled cost and documentation discrepancies before field execution.
- Coordinated vendors, trades, and internal stakeholders against construction deadlines across cost, quality, schedule, documentation, and compliance.

CORE SKILLS

AI evaluation and agent reliability: adversarial testing; evaluation design; acceptance criteria; rubric and scorecard design; completion-claim verification; failure classification; correction and regression gates; agent red teaming

Technical program management and quality operations: requirements definition; test strategy; root-cause analysis; release readiness; risk triage; cross-functional coordination; vendor and stakeholder management

Evidence, governance, and tools: evidence review; fail-closed controls; technical research and writing; Python; Git, GitHub, and CI; contracts, budgets, and vendor records; Excel, Power Query, and Power Pivot

EDUCATION

University of North Texas - B.A., Sociology, Economics, and History

Brooklyn Law School - Coursework toward J.D., 2014-2017; degree not conferred