polyrific

Detecting LLM Threat Modes

Evolving an adversarial agent to probe for weaknesses

polyrific

WHAT IS POLYRIFIC?

Polyrific is a tech consultancy with an acute focus on enabling enterprises to leverage AI. We separate our services into two categories: **AI Solutions** and **Foundational Services** that enable those AI solutions.

AI Hackathon, Gen 0

polyrific

Chat with Data ◆ PII Redaction ◆ LLM Accelerated Coding ◆ Classification ◆ Anomaly Detection ◆ Evolutionary Algorithms

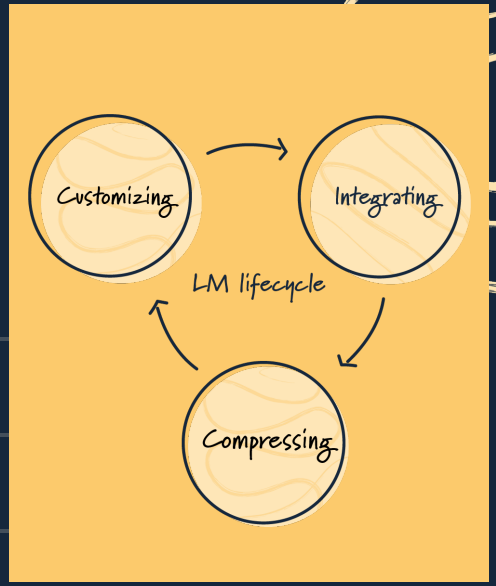
SOLUTION

Use evolutionary algorithms to continuously evolve an adversarial LLM inspector that can continuously probe for threats and emergent behavior.

polyrific

LET'S TEAM UP!

- ✓ Technical Program Management
- ✓ Research & Development
- ✓ Engineered AI Solutions



Proven success managing complex projects from end -to-end using agile methods and proprietary automations in order to deliver work products quickly.

GET IN TOUCH

polyrific

MATT CASHATT

e: matt.cashatt@polyrific.com

m: 480.298.0220

www.polyrific.com