

# AI RED TEAM

## PROBLEM

Unvalidated LLM outputs are high-risk

- Hallucinations erode trust
- Compliance breaches are costly
- Resources wasted on rework



## SOLUTION

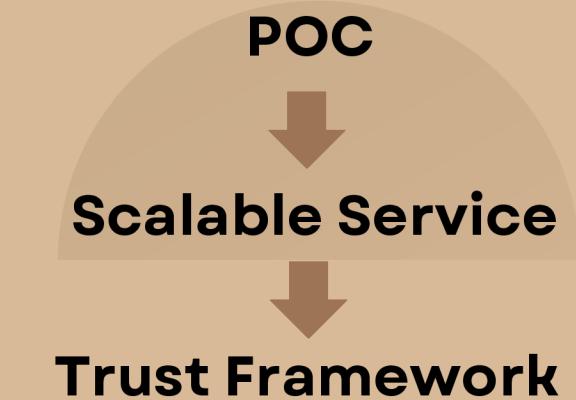
SAFETY SEAL for LLM

- Validation before delivery
  - Every prompt and response is checked for accuracy, compliance, and relevance BEFORE it reaching the user
- Compliance
  - Accuracy, compliance enforced
- TRUST
  - Trust built into workflows

# VALUE SAFETY & TRUSTED AI



# FUTURE AI VALIDATION FRAMEWORK



## Framework Design

- Start small, Scale fast
- Standardize Guardrails
- Make trust an Enterprise-wide norm

# Architecture

