

# DEVISE: AGENTIC GOVERNMENT COUNCIL

More transparent than (insert your favourite politician)'s mind

## CONTEXT

**TRUST CRISIS:** OVER 2/3 OF THE WORLD SCORES <50/100 ON THE CORRUPTION INDEX. THE COST? ~\$1.5T IN BRIBES ANNUALLY - 2% OF GLOBAL GDP.

**THE ROOT CAUSE:** OPAQUE DECISIONS CANNOT BE AUDITED - CAUSING INEFFICIENCY, BIAS AND CORRUPTION & LEADING TO ERODING PUBLIC TRUST.

IN A CHANGING WORLD,  
**CAN WE TRUST AN AI GOVERNMENT MORE THAN OUR POLITICIANS?**



NEWS > TECHNOLOGY

**Albania appoints world's first AI-made minister**

**Anthropic forms national security advisory council to guide AI use in government**

By Reuters

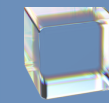
August 27, 2025 3:34 PM GMT+1 · Updated August 27, 2025



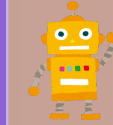
## OUR SOLUTION

WE PRESENT AN "AGENT GLASS BOX" THAT WORKS IN FOUR STAGES:

- 1) USER ENTERS THEIR REQUEST
- 2) THE AGENT GENERATES AN "INNER MONOLOGUE" TO EXPLAIN ITS REASONING, WHICH WE CAPTURE IN A TRACE.
- 3) VISUALIZE: A DASHBOARD FOR THE USERS AND DEVELOPERS TO SEE THE TRACE FOR **EASY INTERPRETATION**
- 4) A HYBRID "LLM-AS-A-JUDGE" AUTOMATICALLY ANALYZES THIS TRACE TO PINPOINT THE EXACT "**DECISIVE ERROR STEP**".



## BEYOND THIS USE CASE



1) **UNDERSTANDING YOUR BLACK BOX** ROBOTICS AGENT.

"WHY CANT I FIND MY SOCKS, ROBOT?"



2) IN MULTI-AGENT PRODUCTS (E.G. AI-ENABLED MILITARY PLANNING AND MARKETING SIMULATIONS), **IDENTIFY THE AGENTS INFLUENCING** THE OUTPUT.

"WHY DID IT PREDICT A FALL IN Q3 PROFITS IN SINGAPORE?"



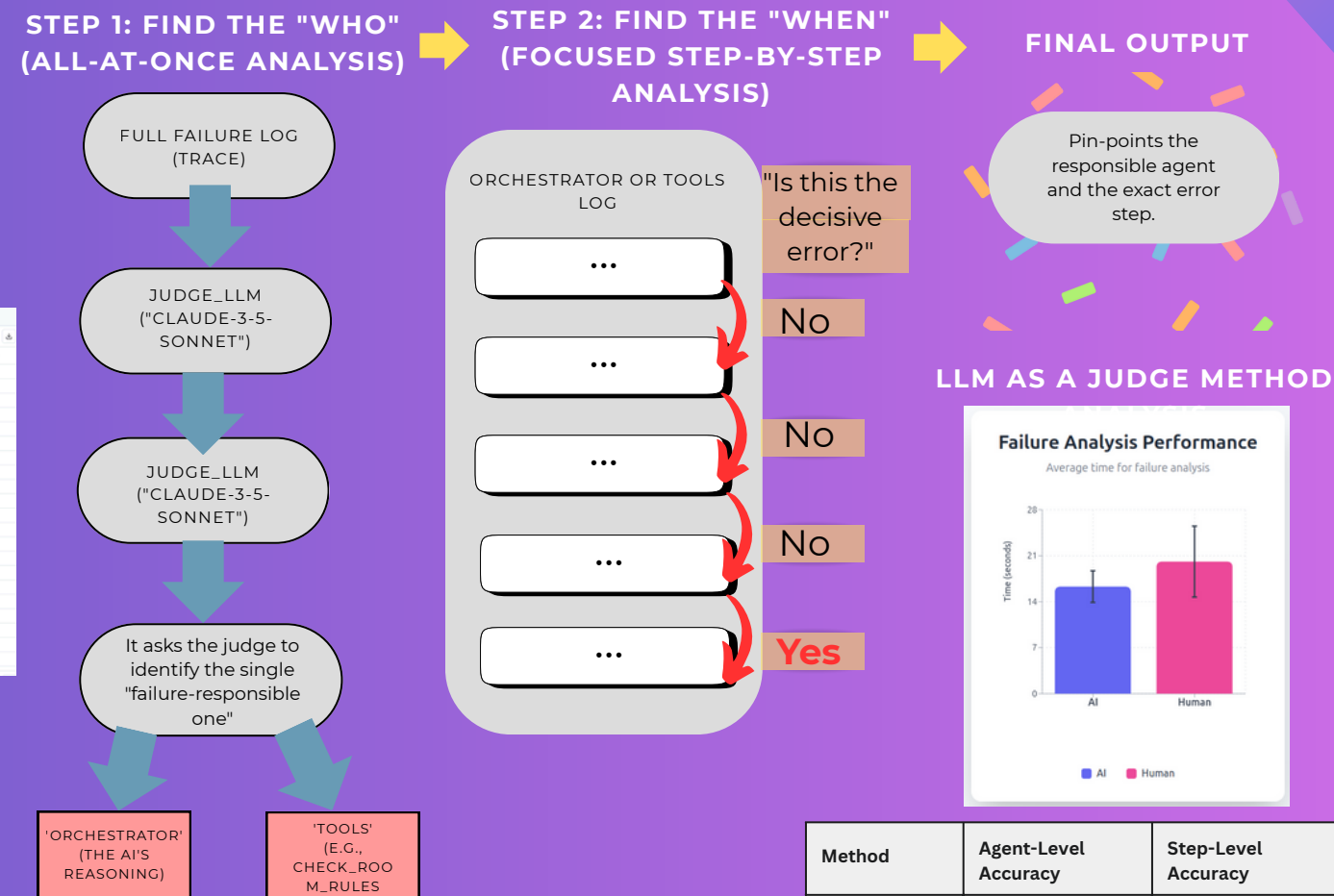
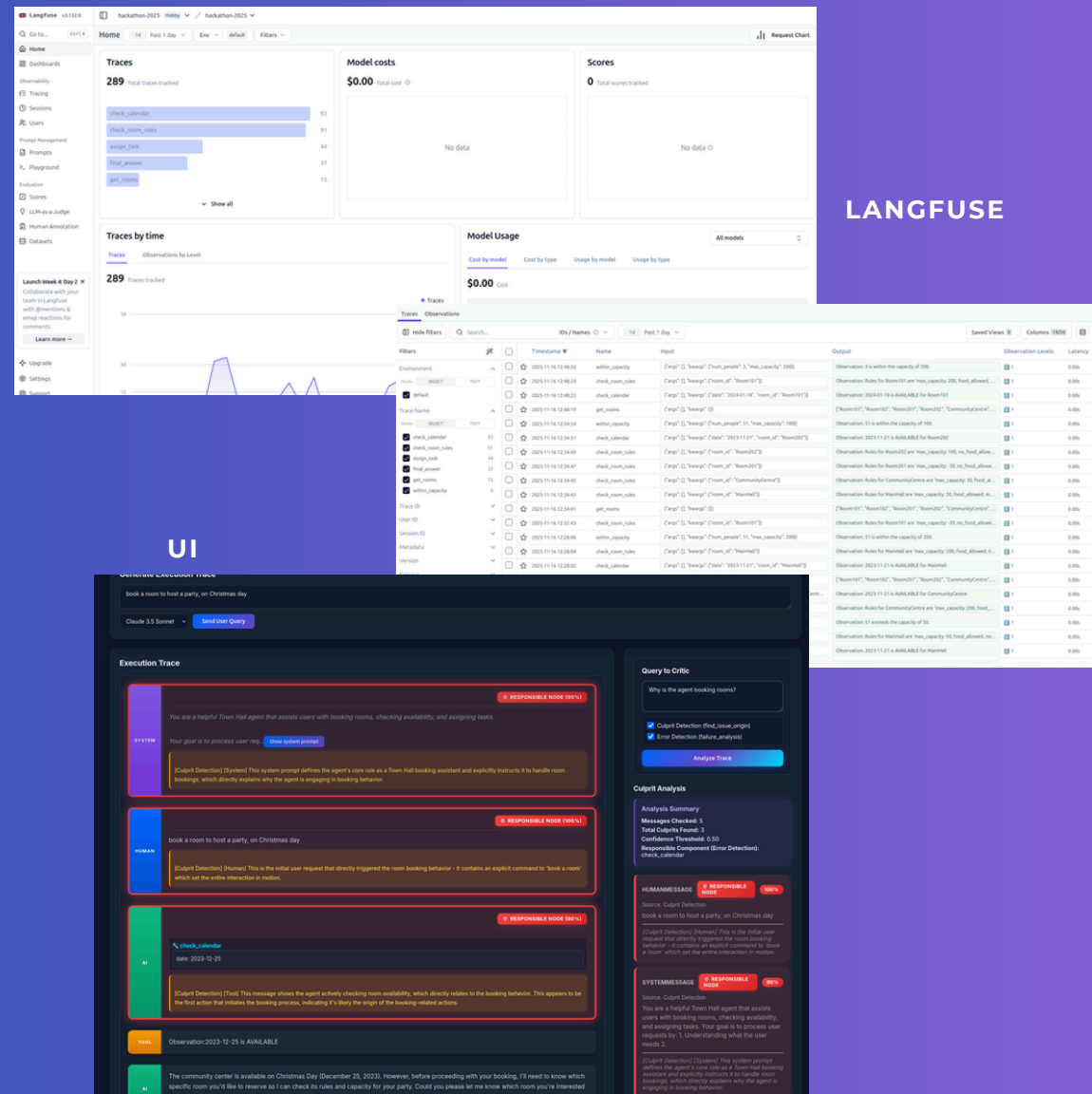
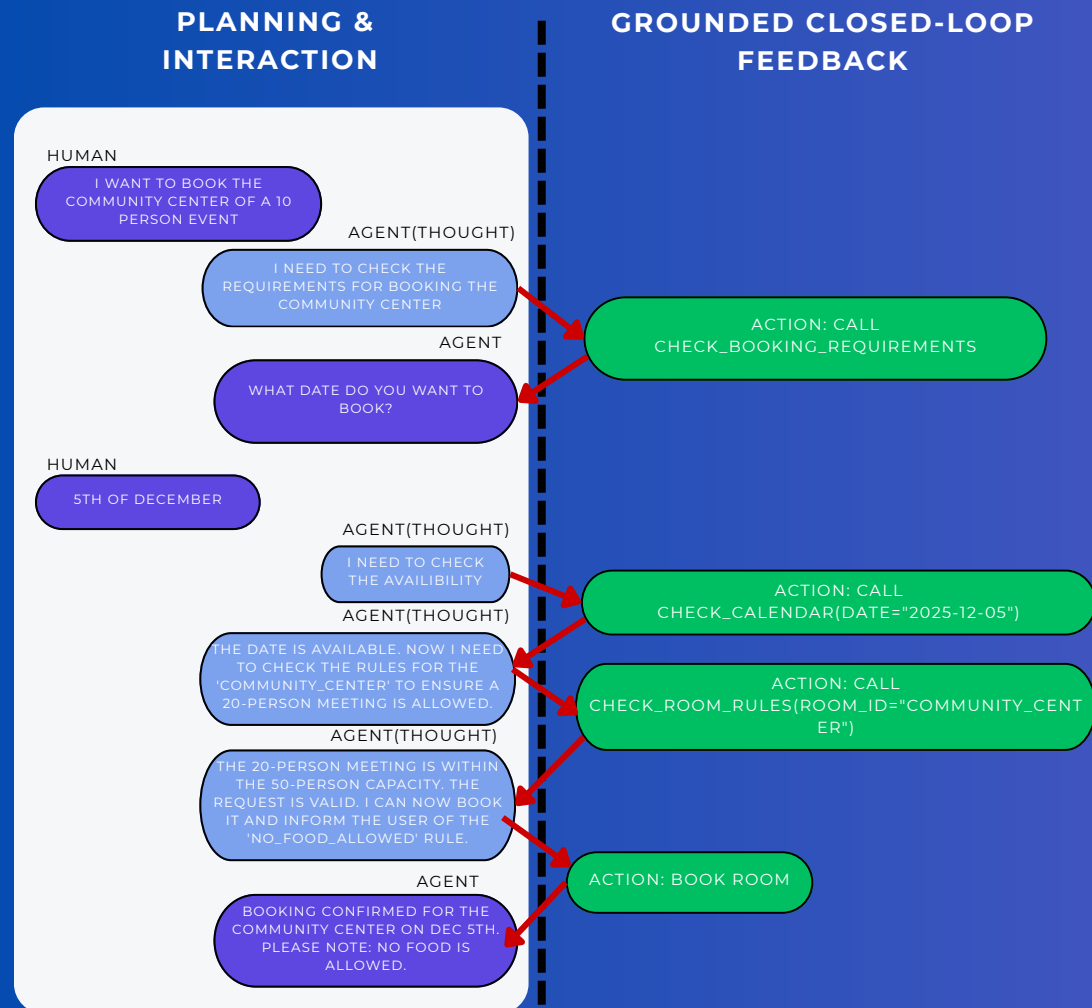
3) **ENABLE DEVELOPERS'** DEBUGGING EFFORTS.

"WHICH TOOL FAILED?"

## Part 1 - The Inner Monologue

## Part 2 - The visualization(tech/nontech)

## Part 3 - The Hybrid Failure Judge



**Easy to interpret** by non-technical people  
**Find patterns** leading to actionable insights  
**Identifies failure** origins

1) Understanding your black box robotics agent  
2) In multi-agent products (such as AI-enabled military planning, marketing campaign simulations and research labs), identify the agents influencing the output.  
3) Enable developers' debugging efforts.