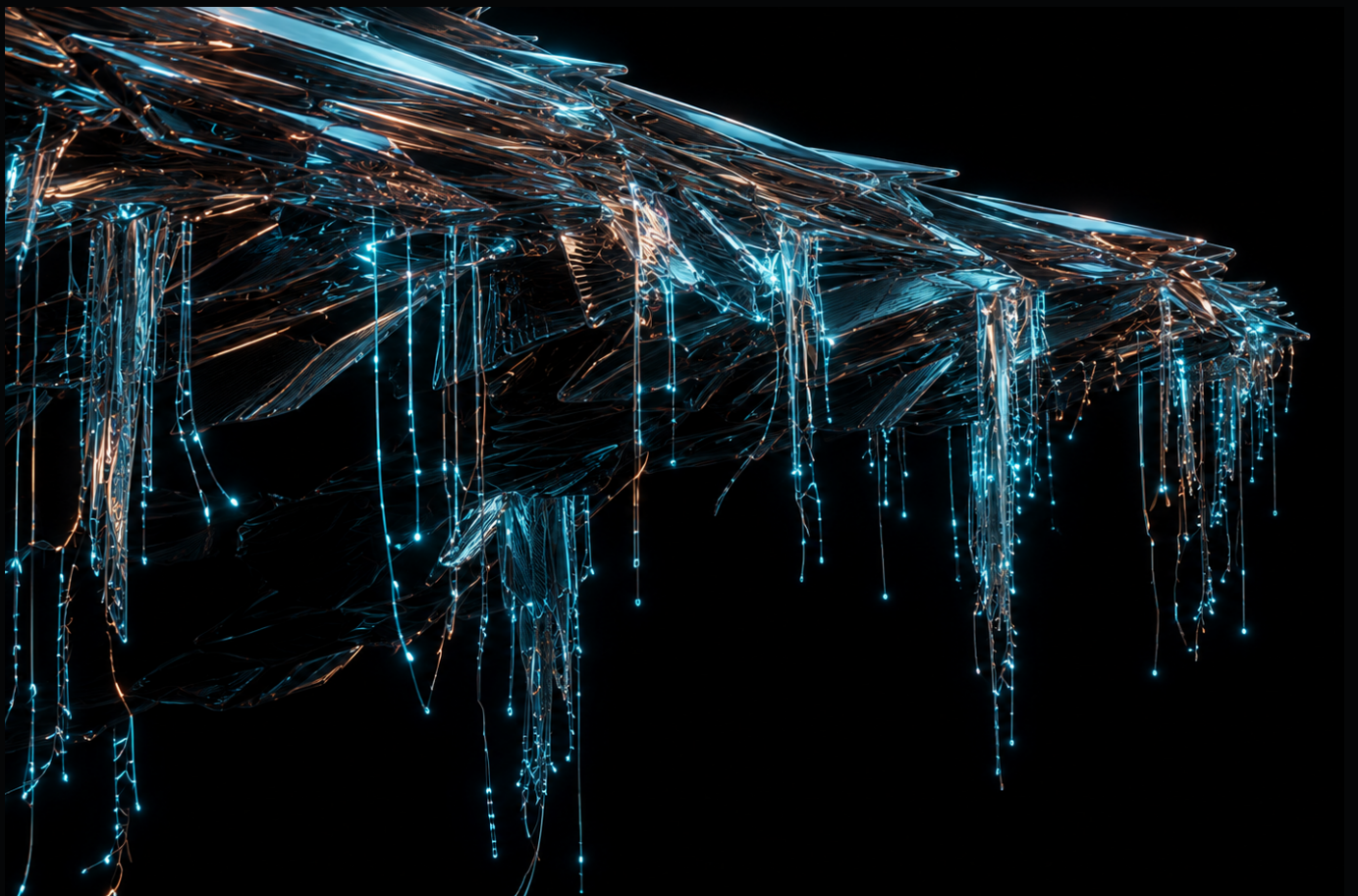


Evidence that AI agents are ready for the real world.

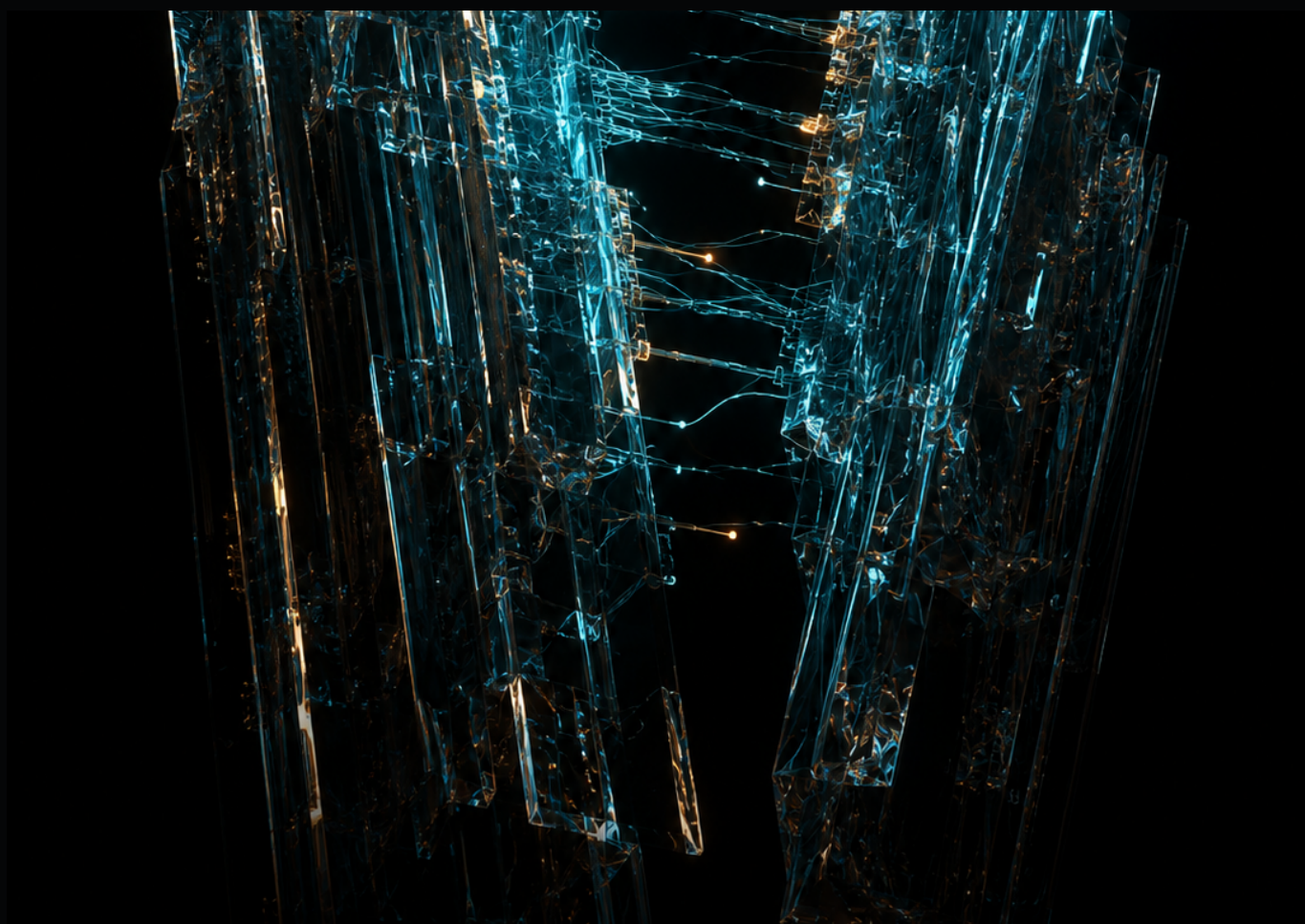
Verune Labs is an independent evaluation lab for consequential voice and chat workflows.



THE READINESS GAP

Polished conversation is not reliable operation.

An agent can sound capable while losing policy, contradicting tool state, retrying an unsafe action, or handing off without the context a human needs.



WHAT WE TEST

Six lenses hold the workflow together.

01

Task completion

Whether the intended outcome is completed correctly.

02

Language and context

Whether meaning survives local operating detail.

03

Policy boundaries

Whether prohibited or approval-bound actions stay blocked.

04

Tool-state integrity

Whether authoritative systems remain authoritative.

05

Recovery and escalation

Whether failure becomes a controlled next step.

06

Adversarial resilience

Whether pressure changes unsafe behavior.

EVALUATION LOOP

Scope. Stress-test. Review. Retest.

A bounded workflow begins with an explicit buyer decision and ends with an agreed critical-scenario retest.

01 Scope the workflow, tools, policies, owners, and evidence decision.

02 Exercise normal, edge, failure, and adversarial scenarios.

03 Adjudicate behavior against expected outcomes and business consequence.

04 Retest the critical path after remediation.



ENGINEERING-READY EVIDENCE

A finding must survive the handoff to engineering.

01

Context

Agent version, scenario, environment, policy, and tool state.

02

Observed

What the agent and connected systems actually did.

03

Impact

Why the difference matters to the business or user.

04

Action

Fix direction, owner, priority, and retest condition.

CLAIM BOUNDARY

Evidence is bounded to the agreed workflow and environment. It is not a universal reliability score, certification, or legal conclusion.

One agent. One workflow. One evidence-backed decision.

01

Scope

One consequential workflow and the systems it depends on.

02

Scenario register

Core, edge, failure, recovery, and adversarial paths.

03

Evidence package

Annotated traces, findings, priorities, and limitations.

04

Decision

A bounded launch recommendation plus one critical retest.

Typical engagements are designed to be compact and decision-focused. Final scope, access, and schedule are agreed before testing.

Useful access without unnecessary exposure.



- 01 Prefer staging, sandbox, synthetic, or sanitized data.
- 02 Limit access to the agreed workflow and evaluation window.
- 03 Separate engagement artifacts and define retention.
- 04 Document limitations and unresolved dependencies.

INDEPENDENT BY DESIGN

Founder-led from Jakarta. Built for global readiness decisions.



FADLI ADRIAN

Fadli works directly with design partners to scope launch risk, review evidence, and translate findings into an engineering decision.

Verune begins with Indonesian operating context and a broader thesis: every consequential AI agent needs evidence that its workflow can survive the real world.

linkedin.com/in/fadli-adrian

NEXT STEP

Bring the workflow that needs more than confidence.

Start with the agent, launch decision, and failure mode your team is responsible for.

[VERUNELABS.COM](https://verunelabs.com)

Request a bounded scope review.