

How to Make AI Agents Enterprise-Ready

*We solved this **500 years ago.***

Philip Stanislaus

CTO, Causa Prima



CAUSA
PRIMA

WHAT WE BUILD

Agents that touch real money.



Read

>



Extract

>



Validate

>



Approve

>



Pay



CAUSA
PRIMA

ANY AGENT CAN BE COMPROMISED



**Prompt
injection**



Hallucination



**Confused
deputy**



Bugs

You can't tell from the output.



**CAUSA
PRIMA**

THE TEMPTING ANSWER

Monitor everything?

Slow. Expensive. *Misses what matters.*



CAUSA
PRIMA



It's an office building.

Trust the hallways. Lock the doors.



FOUR DOORS WORTH LOCKING



Front door

API auth



Job description

Agent scope



Locked rooms

Data + PII



Mailroom

External actions



CAUSA
PRIMA

BUT BOUNDARIES AREN'T ENOUGH



The confused deputy.

Don't break the vault — trick someone with a key.



CAUSA
PRIMA

THE 500-YEAR-OLD ANSWER

1494

Luca Pacioli: Double-entry bookkeeping.

Record it twice. Discrepancies surface themselves.



CAUSA
PRIMA

SAME PRINCIPLES, APPLIED TO AGENTS



**Verify the
source**



**Separate
duties**



**Hard-coded
checks**

LLMs inform. Rules enforce.



**CAUSA
PRIMA**

ZERO TRUST

Not *if*. **When.**

Assume any agent can be compromised — and build so it doesn't matter.

That's **zero-trust**.



Does it *matter*?

In a well-designed system — no.



CAUSA
PRIMA

HOW WE BUILD IT — WITH CLAUDE

CLAUDE.md is just the start.



Spec-driven dev



TDD · coverage



Code-review agents



Adversarial testers



Architecture reviews



Synthetic personas



CAUSA
PRIMA

Closed loop as the target.

Hundreds of systems. Always drifting.

Downtime · rate limits · API changes · no versioning



Tests run
continuously



Invariant
trips



Agent
diagnoses



Linear issue
+ PR



CAUSA
PRIMA

Drift breaks an invariant — and fixes itself before we notice.

We are hiring.

Building agents that touch real money?
Reach out — let's talk.

Philip Stanislaus · CTO, Causa Prima



**CAUSA
PRIMA**



Follow us

causaprima.ai