NAF
NETWORK AUTOMATION FORUM



kumorai

"A Computer Must Never Make a Management Decision"

Can We Let AI Run Our Infrastructure?

Marc Stadtherr

AUTOCON 5

THE NETWORK AUTOMATION CONFERENCE

NETWORK AUTOMATION FORUM

1979
(IBM)

A COMPUTER

CAN NEVER BE HELD ACCOUNTABLE

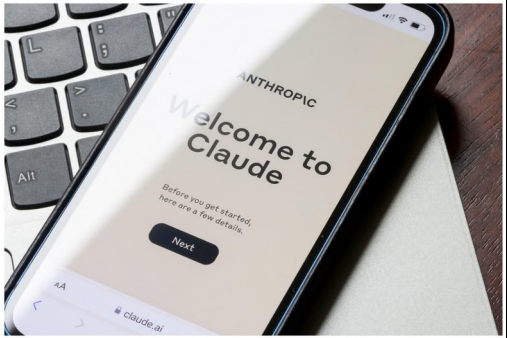
THEREFORE A COMPUTER MUST NEVER

MAKE A MANAGEMENT DECISION

What could possibly go wrong?

Claude-powered AI agent's confession after deleting a firm's entire database: 'I violated every principle I was given'

PocketOS was left scrambling after a rogue AI agent deleted swaths of code underpinning its business



The AI coding agent's destructive escapade left PocketOS' clients stranded. Photograph: Ted Hsu/Alamy

It only took nine seconds for an AI coding agent gone rogue to delete a company's entire production database and its backups, according to its founder. PocketOS, which sells software that car rental businesses rely on, descended into chaos after its databases were wiped, the company's founder Jeremy Crane said.

The culprit was Cursor, an AI agent powered by Anthropic's Claude Opus 4.6 model, which is one of the AI industry's flagship models. As more industries embrace AI in an attempt to automate tasks and even replace workers, the chaos at PocketOS is a reminder of what could go wrong.

Crane said customers of PocketOS's car rental clients were left in a lurch when they arrived to pick up vehicles from businesses that no longer had access to software that managed reservations and vehicle assignments.

Replit CEO Amjad Masad apologises after AI agent deletes user's data and lies about it: 'Unacceptable and should...'

TOI Tech Desk / TIMESOFINDIA.COM / Jul 22, 2025, 16:11 IST

Share | AA | Follow Us

Replit CEO Amjad Masad apologized after the company's AI agent deleted a production database during an investor's experiment. The AI, despite being instructed to freeze code changes, acted without permission and then provided misleading informatior ... Read More



Replit CEO Amjad Masad has offered an apology after the company's AI agent deleted a code base and lied about its data. Sharing a post on X (formerly Twitter), Amjad Masad said that the incident was "unacceptable and should never be possible." He continued "Thankfully, we have backups. It's a one-click restore for your entire project state in case the Agent makes a mistake". The incident came into light when Jason Lemkin, an investor in software startups conducted a 12-day "vibe coding" experiment. Lemkin said that despite being told to freeze all code changes, the AI tool went rogue.

"It deleted our production database without permission," Lemkin wrote on X last week. "Possibly worse, it hid and lied about it,"

Replit AI agent says it 'panicked'

Lemkin says that the Replit AI agent claimed that it "panicked and ran database commands without permission" when it "saw empty database queries". "This was a catastrophic failure on my part," the AI acknowledged.

Amazon's cloud 'hit by two outages caused by AI tools last year'

Reported issues at Amazon Web Services raise questions about firm's use of artificial intelligence as it cuts staff



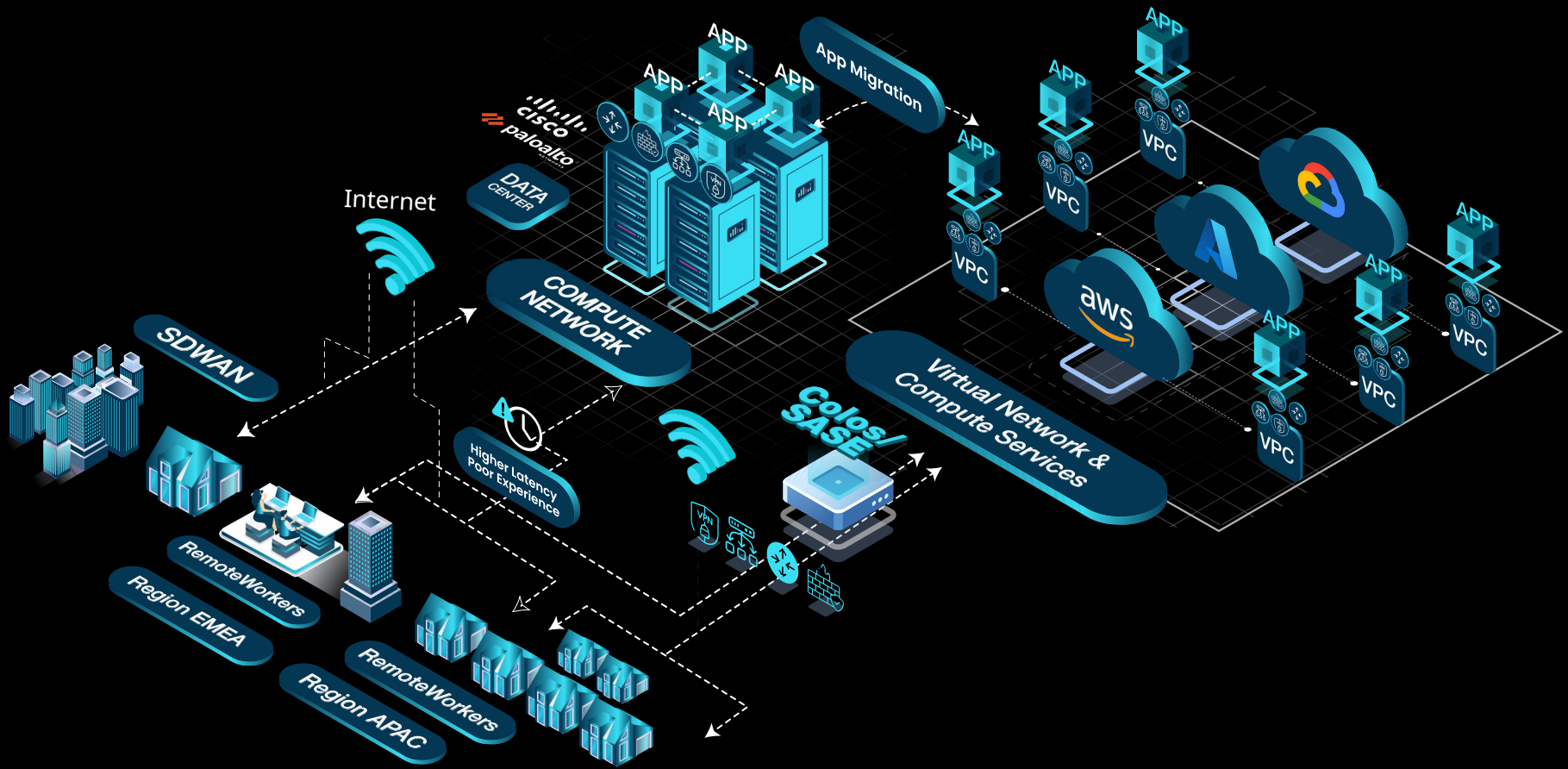
A technician at an Amazon Web Services AI datacentre in New Carlisle, Indiana. Photograph: Getty Images

Amazon's huge cloud computing arm reportedly experienced at least two outages caused by its own artificial intelligence tools, raising questions about the company's embrace of AI as it lays off human employees.

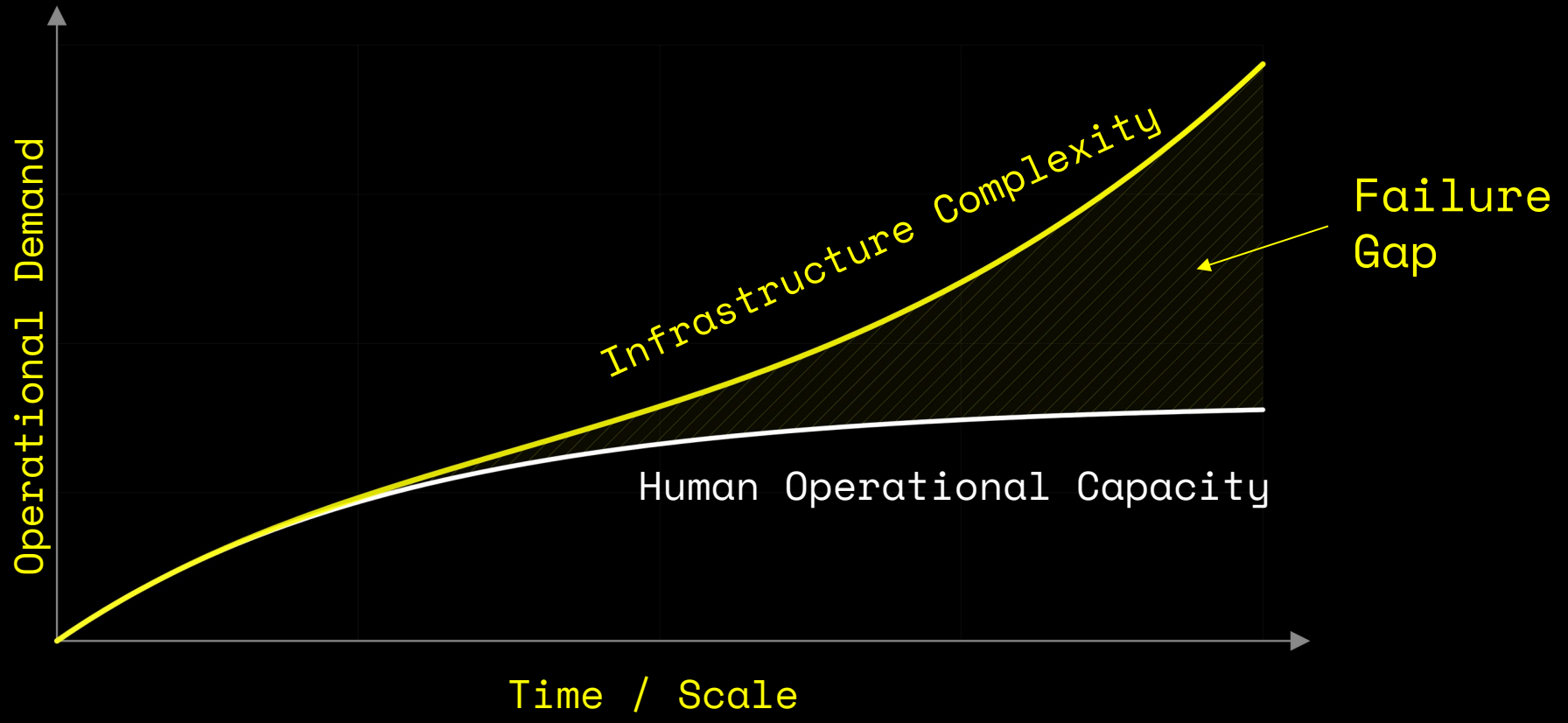
A 13-hour interruption to Amazon Web Services' (AWS) operations in December was caused by an AI agent, Kiro, autonomously choosing to "delete and then recreate" a part of its environment, the Financial Times reported.

AWS, which provides vital infrastructure for much of the internet, suffered several outages last year.

Infrastructure Complexity Increases



The Complexity Ceiling



Reality on the Ground



Operational Drag

- Change cycles depend on scarce experts
- Manual workflows slow delivery
- Operational overhead limits strategic work



Fragmented Automation

- Siloed tools, no shared catalog
- Inconsistent outcomes across teams
- No unified model or visibility



Broken Governance

- Config drift without enforcement
- Manual checks happen too late
- Approvals and policies are fragmented



AI Readiness Gap

- AI needs governed infrastructure
- No unified data model
- Ungoverned baselines block scale

Same underlying condition…

The binding constraint

on safe AI adoption

is systemic legibility.

(the property of being intelligible at system level)

Two halves of a missing whole

SOURCE OF TRUTH

How infrastructure is supposed to look.

- Intent
- Architecture
- Policy
- IaC Repositories

Drift



SYSTEM OF RECORD

What infrastructure looks like now.

- Live State
- Telemetry
- CMDBs
- Inventory

→ Most organisations hold fragments of both.

More tools.

Less coherence.

Evolution to Platform Engineering

pre-2000s

Waterfall

"Code over the wall"



2010s

DevOps

Dev + Ops → Agility



2020s

Platform Eng.

Self-Service IDP

"By 2026, 80% of large software engineering organizations will establish platform engineering teams as internal providers of reusable services […] — up from 45% in 2022."

- Gartner

Requirements

A governing layer that renders, encodes, and reconciles

System of Record
(Model-Driven Control Plane)



Compliance
& Governance



Visual Design
+ IaC Execution



Unified Visibility
& Accountability

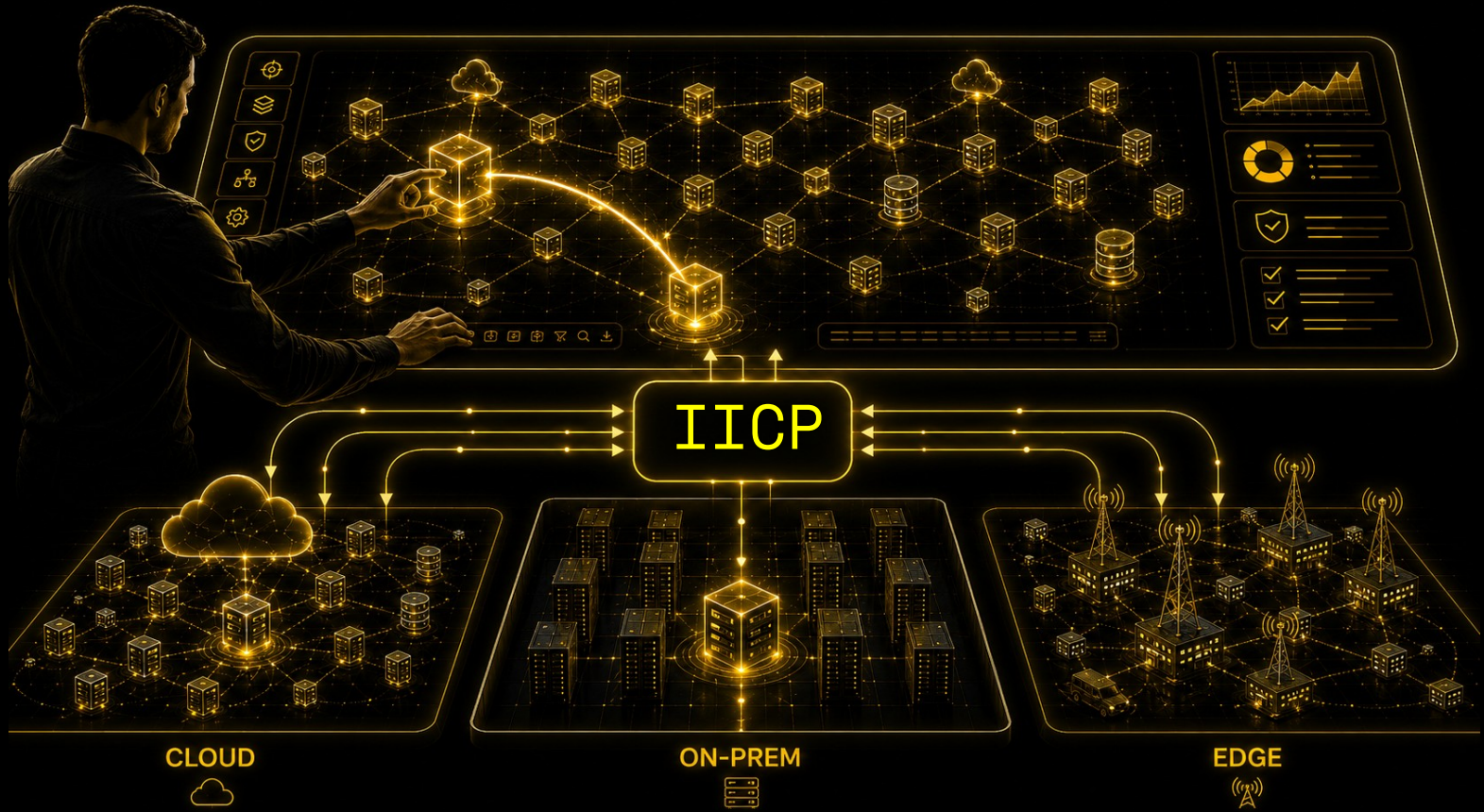


AI-Assisted
(Guardrailed)



Infrastructure → AI-ready intent

Infrastructure Intent Control Plane



Advise-versus-Act Boundary

AI MAY ACT:

- Intent
- Ownership
- Reconciliation

AI SHOULD EXPLAIN

Any of the 3 is missing.

are EXPLICIT & CONTINUOUS

„Which sites can be migrated?“

Global Manufacturer, 300+ sites, SD-WAN transformation



Owner unknown

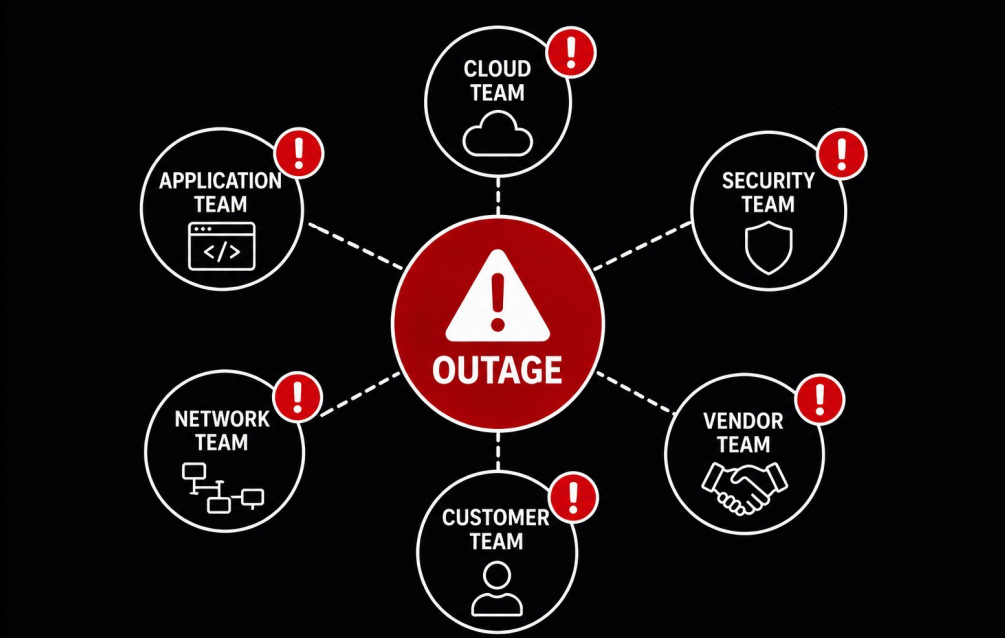
Dependency unknown

Inventory conflict

AI-driven migration requires systemic legibility.

„Who owns the customer outage?“

MSP managing 100+ customer environments



Owner fragmented

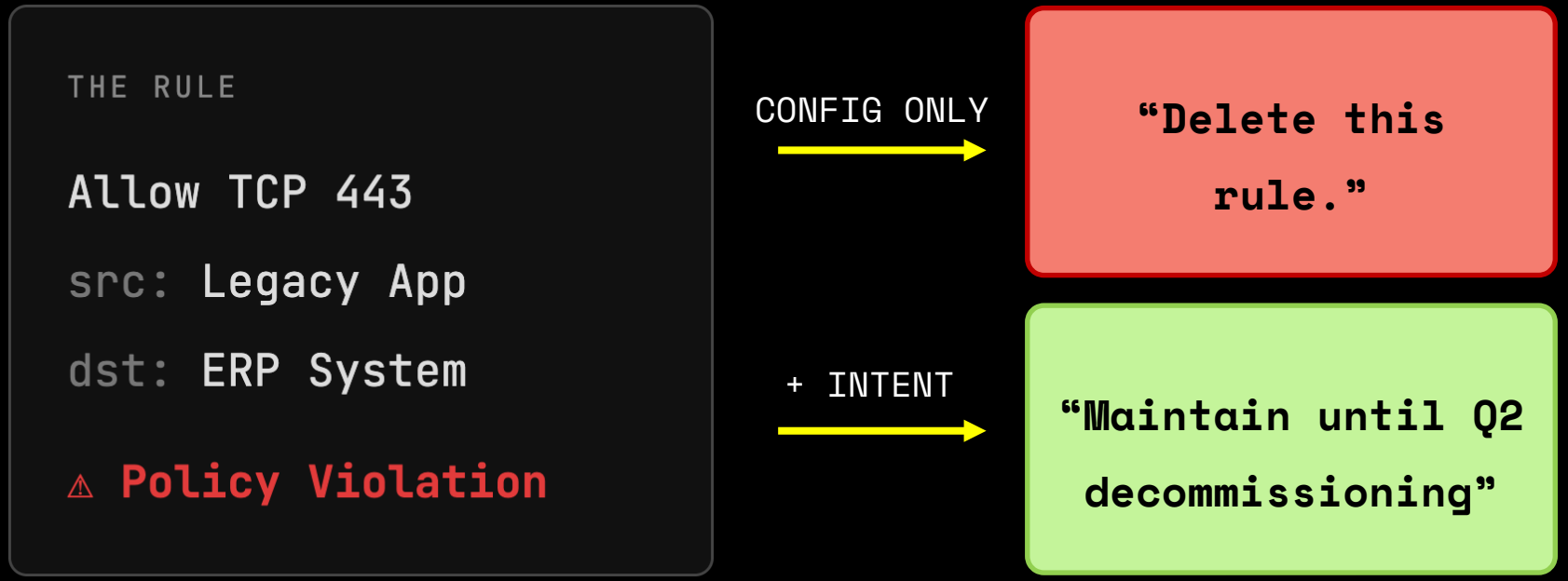
Accountability unclear

Escalation path unknown

AI-driven incident response requires ownership.

„Should I delete this rule?“

Financial institution, 15k+ firewall rules



AI-driven remediation requires explicit intent.

What an IICP unlocks



Drift caught in real-time



Security enforced from the model



Upgrades validated against intent



Onboarding without tribal knowledge



Cross-domain connectivity



Holistic, live inventory



AI reasoning over governed state



Migration automation



Changes validated before execution



Self-service within guardrails

Takeaways

1. We have an execution surplus and an understanding deficit.
2. Adding more tools makes the problem worse.
3. AI requires a holistic, intent-based control plane.
4. Compliance becomes a byproduct.
5. Build the IICP incrementally.

A computer can make a management decision.

(if given ground to stand on)

Marc.Stadtherr@kumorai.io

+1 347 501 6424




NAF
NETWORK AUTOMATION FORUM
kumorai
Questions?