Identity-Centric Security for AI Systems **IAM as Cybersecurity Infrastructure**

Reframing identity and access management from an administrative capability into the primary control plane for securing AI-driven enterprises.

By **Arunkumar Muthuramalingam**

Cyber Security and Infrastructure Operations manager at General Reinsurance

The Expanding Identity Attack Surface

The rapid enterprise adoption of LLMs, AI copilots, autonomous agents, APIs, and cloud-native automation has fundamentally redrawn the threat landscape. Traditional perimeter-based security models were designed for a world of human users and static network boundaries a world that no longer exists.

LLMs in Production

Language models embedded in workflows introduce novel trust and authorization challenges at every inference boundary.

Autonomous AI Agents

Agents that act, decide, and call external services operate with delegated authority that must be tightly scoped.

Machine-to-Machine APIs

Orchestration frameworks, plugins, and service meshes create dense webs of inter-service trust that demand continuous verification.

Cloud-Native Automation

Serverless functions, pipelines, and workload identities operate at scale with credentials that are rarely audited in real time.

Why the Perimeter Is No Longer the Boundary

Modern AI systems rely on complex, dynamic interactions between users, services, models, plugins, orchestration frameworks, and machine identities. In this environment, no single network boundary can enforce consistent security policy. **Identity has become the only control plane that travels with the workload.**

The Old Model

- Trusted internal network
- Static firewall rules
- Human-centric access policies
- Periodic credential audits

The New Reality

- Zero implicit trust, any location
- Dynamic, context-aware authorization
- Non-human identities dominate workloads
- Continuous, real-time verification required

Session Agenda

What We'll Cover Today

01

Identity as Control Plane

Reframing IAM from IT administration to foundational cybersecurity infrastructure for AI ecosystems.

02

Least Privilege at AI Scale

Granular, context-aware, and time-bound access controls for human and non-human actors.

03

Continuous Verification

Eliminating lateral movement and privilege escalation across AI pipelines, APIs, and cloud workloads.

04

Non-Human Identity Dominance

Governance strategies for AI agents, automation services, and machine-to-machine workloads.

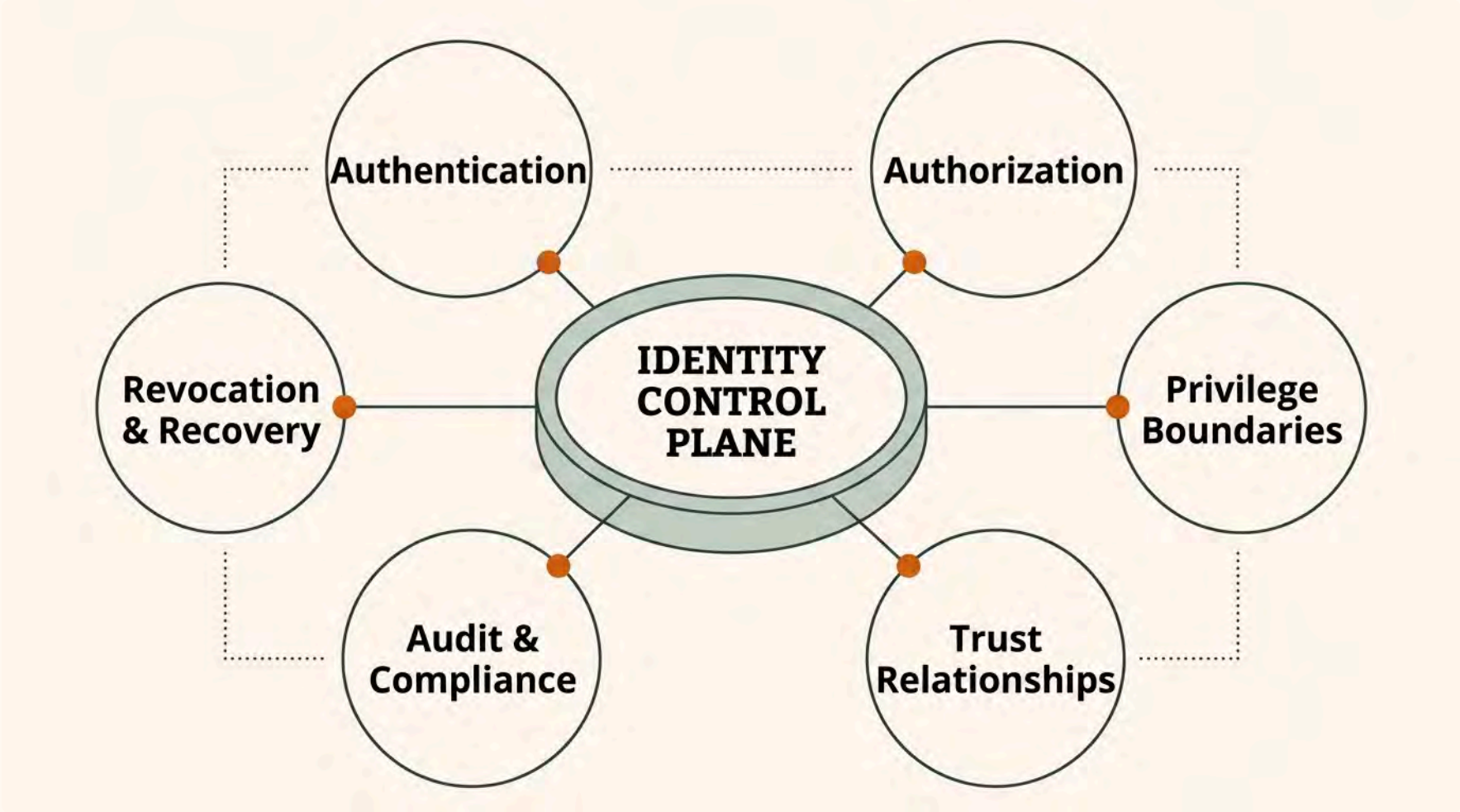
05

Resilience Through Governance

Rapid containment, revocation, and recovery when AI systems are compromised.

Identity as the Primary Control Plane

Rather than treating IAM as a supporting administrative function, modern AI security architecture demands that identity be elevated to the role of **structural security infrastructure** governing authentication, authorization, privilege boundaries, and trust relationships across every distributed component.



Least Privilege: Reducing Attack Surface at AI Scale

Identity-driven least privilege is the most direct mechanism for shrinking the blast radius of any compromise in an AI system. Granularity, context-awareness, and time-bounding are the three axes that define modern least-privilege architecture.

Granular Access Controls

Permissions scoped to the exact resource, action, and data classification required never broader. Applied uniformly to human users, service accounts, AI agents, and automation workloads.

Context-Aware Authorization

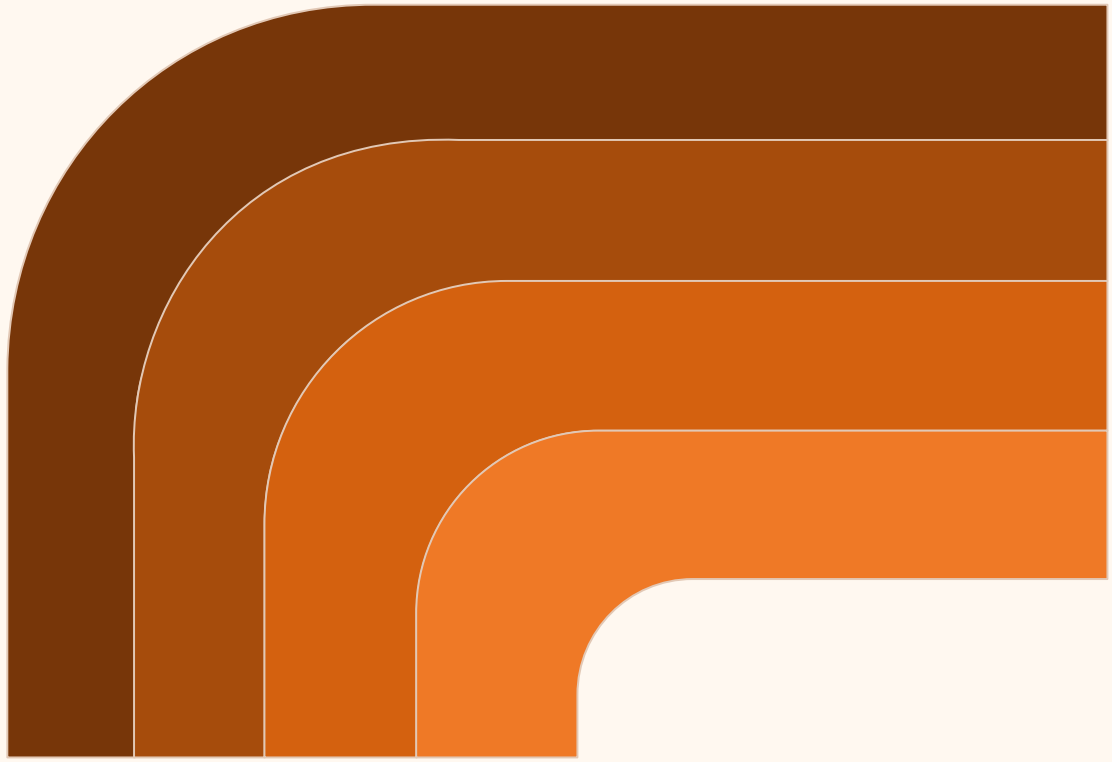
Access decisions informed by real-time signals: device posture, network location, time of day, behavioral baselines, and request risk score. Static role assignments are insufficient for dynamic AI pipelines.

Time-Bound Credentials

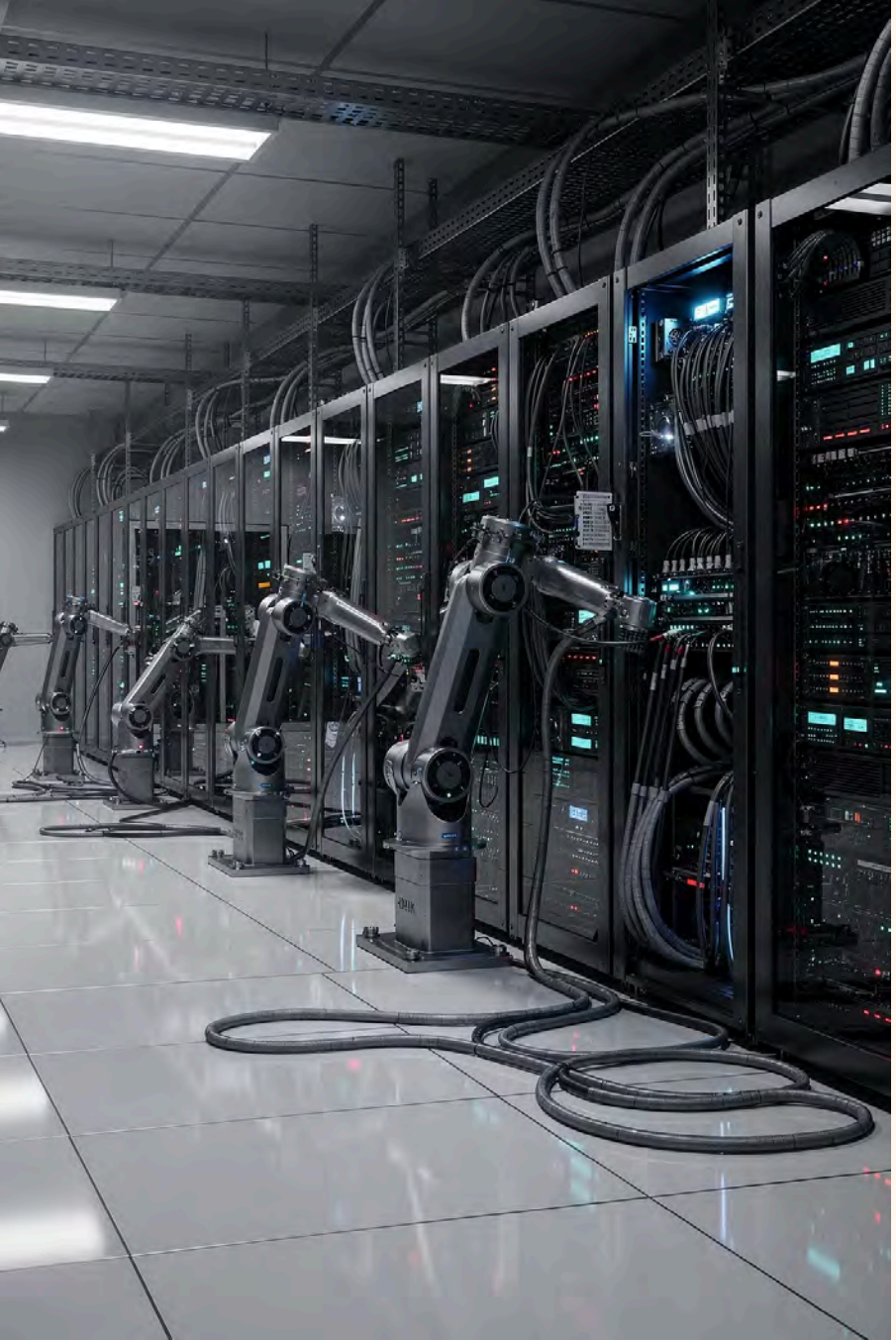
Ephemeral tokens and just-in-time privilege elevation replace long-lived credentials. Access is granted for the minimum duration required and automatically revoked upon session or task completion.

Continuous Identity Verification Across AI Pipelines

Static authentication at session initiation is insufficient when AI agents autonomously traverse APIs, vector databases, orchestration layers, and cloud workloads. **Every hop is a new trust boundary that must be independently verified.**



Continuous verification directly addresses the two most dangerous post-compromise behaviors in AI environments: **lateral movement** across interconnected services, and **privilege escalation** through misconfigured trust relationships or stolen tokens. Real-time behavioral monitoring and anomaly detection must be embedded at every layer of the AI stack not bolted on at the perimeter.



Chapter 3

The Non-Human Identity Problem

> In modern enterprise AI environments, non-human identities AI agents, automation services, and machine-to-machine workloads now significantly outnumber human users. Most organizations cannot enumerate them.

Non-Human Identities: Scale, Risk, and Governance

Service accounts, API keys, OAuth tokens, workload identities, and AI agent credentials represent the largest and fastest-growing identity population in enterprise environments and historically the least governed. This gap is now a primary attack vector.

Why Non-Human Identities Are High Risk

- Long-lived credentials rarely rotated
- Overly permissive scopes granted at creation
- No MFA or behavioral baseline
- Orphaned accounts persist after decommission
- Difficult to correlate across distributed systems

Governance Requirements

- Comprehensive discovery and inventory
- Automated credential rotation policies
- Scoped, short-lived tokens per workload
- Behavioral anomaly detection for agents
- Lifecycle management tied to service lifecycle

AI Agent Identity: Trust Delegation at Scale

Autonomous AI agents are unique identity principals they act with delegated authority, call external APIs, access sensitive data stores, and spawn sub-agents. Each of these operations must be governed by a clearly defined identity with scoped, auditable permissions.

Agent Authentication

Every agent must authenticate with a verifiable, non-shareable identity. Shared credentials between agent instances create untraceable accountability gaps.

Scoped Authorization

Agent permissions must be scoped to the specific task context. An agent summarizing documents should never hold write access to production databases.

Audit & Traceability

Every action taken by an AI agent must be attributable to a specific identity and logged with sufficient fidelity to support incident investigation and compliance reporting.

Centralized Identity Governance: Resilience by Design

Distributed, siloed identity management creates irreconcilable blind spots. When an AI system is compromised, the ability to **contain, revoke, and recover** at speed depends entirely on whether identity is centrally governed or fragmented across teams and tools.

1

Detect

Centralized telemetry enables rapid identification of anomalous identity behavior across all AI workloads.

2

Contain

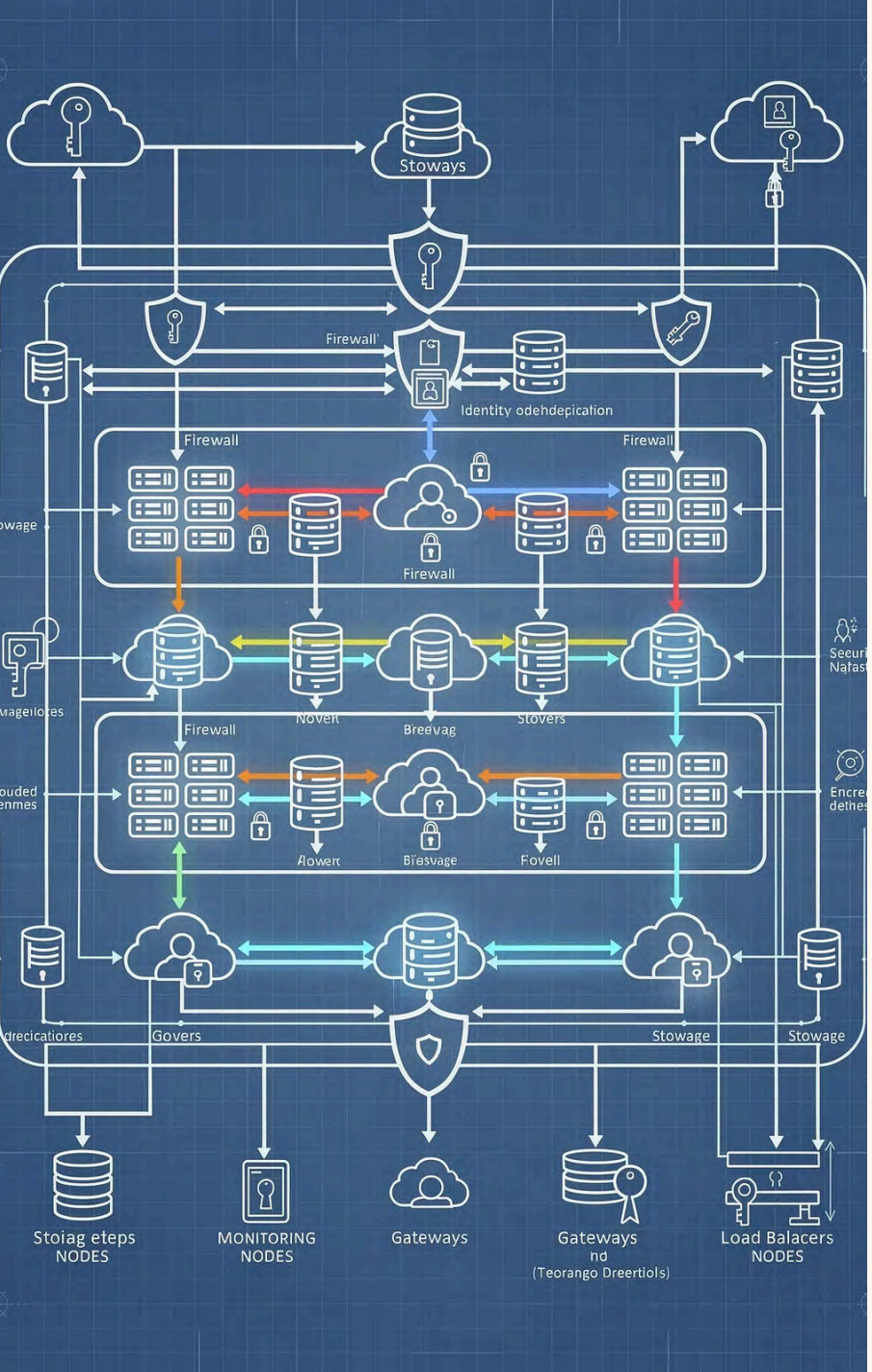
A single governance plane allows immediate revocation of compromised credentials across all downstream systems simultaneously.

3

Recover

Pre-staged identity playbooks accelerate credential re-issuance and privilege re-scoping without manual coordination delays.

Organizations with centralized identity governance contain breaches significantly faster than those with fragmented identity silos the difference between minutes and days of lateral exposure.



Infrastructure-Level Impacts: The Three Pillars

Reduced Attack Surface

Least-privilege access controls granular, context-aware, and time-bound minimize the permissions available to any compromised identity, dramatically limiting what an attacker can reach.

Limited Lateral Movement

Continuous verification at every service boundary prevents compromised credentials from being used to traverse AI pipelines, APIs, vector databases, and cloud workloads undetected.

Faster Incident Recovery

Centralized governance enables rapid containment through immediate, system-wide credential revocation and pre-built recovery playbooks that eliminate coordination delays under pressure.

Architectural Principles: Building Identity-First AI Systems

Implementing identity-centric security requires architectural decisions made at system design time not retrofitted after deployment. The following principles provide a strategic framework for LLM-powered systems built to be scalable, resilient, and trustworthy.

1

Identity by Design

Define identity boundaries, trust relationships, and permission scopes before building AI pipelines not after observing production failures.

2

Zero Implicit Trust

Every component human or machine must authenticate and authorize independently, regardless of network location or prior session state.

3

Continuous Enforcement

Replace point-in-time access decisions with continuous, signal-driven authorization that adapts to runtime context and behavioral signals.

4

Unified Governance

Consolidate human, service, and AI agent identities under a single governance framework with consistent policy, audit, and lifecycle management.

Common Failure Patterns to Avoid

Most identity-related compromises in AI environments are not novel attacks they exploit predictable governance gaps. Recognizing these patterns is the first step to eliminating them.

1

Overprivileged Service Accounts

AI workloads granted broad permissions at provisioning time "for convenience" never scoped down as requirements crystallized. A single compromised token grants blast-radius access across the entire platform.

2

Credential Sprawl

API keys, OAuth tokens, and agent credentials embedded in code, configuration files, or environment variables ungoverned, unrotated, and invisible to the identity governance program.

3

Implicit Agent Trust

Orchestration frameworks that pass identity context between components without re-authentication, allowing a compromised sub-agent to inherit the trust level of the parent orchestrator.

4

Siloed Identity Operations

Human identity managed by IT, service accounts owned by DevOps, and AI agents ungoverned by either creating irreconcilable blind spots that slow detection and containment.

Key Takeaways: Identity Is the Security Boundary

As enterprises scale LLM-powered systems, the security architecture must evolve to match. The network perimeter cannot follow workloads across clouds, agents, and APIs **but identity can.**

Elevate IAM to Infrastructure Status

Identity governance is not an IT function it is the foundational control plane for AI security. Treat it with the same engineering rigor as networking or compute.

Govern Non-Human Identities First

AI agents and automation services represent your largest and least-governed identity population. Inventory, scope, and continuously monitor them before they become your largest liability.

Architect for Continuous Verification

Build AI pipelines that re-authenticate and re-authorize at every trust boundary. Assume any credential can be compromised at any time, and design accordingly.

Organizations that implement identity-centric architecture before scaling AI systems dramatically outperform those that attempt to retrofit governance after incidents occur.

Thank You!