

Zero Trust as Foundational Infrastructure for Enterprise AI Enablement

Author: Sven D. Olensky

Publisher: Agency Collapse Publishing

Date: November 2025

License: Creative Commons Attribution 4.0 International, CC BY 4.0

Copyright © 2025 Agency Collapse Publishing

Zenodo Metadata Block

Title:

Zero Trust as Foundational Infrastructure for Enterprise AI Enablement

Author:

Sven D. Olensky

Publisher:

Agency Collapse Publishing

Publication Type:

Report / White Paper

Version:

1.0 RELEASE

Publication Date:

2025-11-16

Description:

Zero Trust as Foundational Infrastructure for Enterprise AI Enablement presents Zero Trust as an operational force multiplier in the AI era. Grounded in enterprise security architecture, it reframes security from risk containment to strategic acceleration. Drawing from CISA's Zero Trust Maturity Model, NIST SP 800-207, the CSA AI Controls Matrix, and SABSA's design principles, it argues that foundational Zero Trust maturity offers a stable and practical substrate for responsible AI adoption. By operationalizing Zero Trust as both a control system and a capability accelerator, enterprises can achieve safe velocity. They gain the ability to innovate confidently, integrate AI broadly, and convert governance into sustainable growth.

Keywords:

Zero Trust, AI Security, Enterprise Architecture, Cybersecurity, AI Governance, Data Protection, Digital Trust, Safe Velocity, CISA, CSA, NIST SP 800-207, SABSA

DOI: 10.5281/zenodo.17625784

License:

Creative Commons Attribution 4.0 International (CC BY 4.0)

Language:

English

Upload Type:

Publication → Report

Communities:

Cloud Security Alliance (CSA)
AI Safety and Security Research

Related Identifiers (optional):

- CISA Zero Trust Maturity Model (v2.0, 2024)
- NIST SP 800-207: Zero Trust Architecture, NIST AI Risk Management Framework
- Cloud Security Alliance: AI Controls Matrix v1.0
- SABSA: Enterprise Security Architecture Framework

Contents

Abstract.....	4
From Guardrails to Growth	6
Zero Trust Baseline for AI Readiness.....	7
What is Zero Trust?	7
Minimum Viable Zero Trust Baseline.....	7
Why this Matters.....	8
Ad-hoc – No Zero Trust Posture	9
Zero Trust Foundational: The Minimum Viable Zero Trust Baseline.....	10
Zero Trust Foundational: Controls to Implement.....	11
Zero Trust Foundational: Evidence to Provide	12
Zero Trust Optimized: Evolved and Advanced Zero Trust Posture	13
Zero Trust Foundational: Mapping to existing frameworks and standards.....	14
From Architecture to Operation: Rolling Out Foundational Controls	15
Conditional Access Policies and MFA Rollout	15
Common pitfalls to avoid	16
Vulnerability Remediation	16
Common pitfalls to avoid	17
Data Classification and Data Loss Prevention	17
Common pitfalls to avoid	18
Cross-cutting Guidance Across All Control Families	18
Definition of Done	18
Governance and Communication	18
Telemetry.....	18
Safety Valves	19
Operational Realism	19
AI Enablement Patterns – Agentic Systems.....	19
Agentic Use Case: Customer Support Triage Agent	20
Governance at Speed: Turning Security into Strategic Acceleration	21
Conclusion: Safe Velocity.....	22
Acknowledgments	23
References	23

Abstract

Zero Trust, grounded in enterprise security architecture, is the operational force multiplier of the AI era. As both a business enabler and an architectural discipline, it lays the groundwork for secure, large-scale AI adoption: turning what was once risk containment into strategic acceleration.

AI integration exposes the limits of compliance-based security. Organizations accustomed to satisfying auditors now face an ecosystem where models consume sensitive data, learn from shared environments, and act with agency across distributed systems. Each generative or predictive capability introduces not only opportunity but irrevocable trust dependencies. Without verifiable control, enterprises risk losing ownership of their data, their logic, and their accountability.

This paper argues from practical experience that foundational Zero Trust maturity provides the most stable and scalable foundation for responsible AI enablement. It is not the only possible path, but it is a proven one that minimizes complexity and prevents avoidable risks. Drawing from open frameworks, including **CISA's Zero Trust Maturity Model**, **NIST SP 800-207**, the **Cloud Security Alliance AI Controls Matrix**, and **SABSA's business-driven design principles**, it proposes a unified model linking architectural trust, technical enforcement, and business agility.

By operationalizing Zero Trust as both control system and capability accelerator, enterprises can unlock safe velocity: the ability to innovate confidently, integrate AI broadly, and convert governance into growth.

Introduction

While many architectural approaches can support AI adoption, Zero Trust has proven to be the most coherent and operationally grounded foundation. It creates the preconditions that AI systems depend on without requiring organizations to rebuild their entire security model. In the age of AI, these qualities define the line between responsible innovation and uncontrolled exposure. The organizations that succeed with AI will be those that treat Zero Trust not as an overlay or compliance framework but as core infrastructure - the fabric through which every identity, data flow, and model interaction must pass.

This paper does not present Zero Trust as a certification or compliance regime. It treats Zero Trust as a practical architectural posture that consistently produces the operational conditions that modern AI systems require. The position taken here is not that Zero Trust replaces other established frameworks. It is that Zero Trust aligns naturally with them and offers a direct and practical way to operationalize their intentions

Artificial intelligence magnifies the consequences of weak trust boundaries. Models consume sensitive data, generate outputs that can be acted on autonomously, and integrate across systems that were never designed for mutual verification. Each model represents a new locus of decision-making and a new surface for compromise. Traditional, compliance-driven controls are insufficient in this environment because they measure configuration, not behavior. Zero Trust closes this gap by enforcing real-time decisions at every interface - human or machine - and by ensuring that every action is explicitly verified, context-aware, and auditable.

At its core, Zero Trust is a business enabler. When implemented correctly, it allows enterprises to innovate at speed without sacrificing assurance. The *business enablement effect* of Zero Trust can be defined as the measurable increase in controlled innovation rate without an increase in risk exposure. When applied to AI adoption, this means faster deployment of new capabilities, reduced incident rates, and shorter recovery times when anomalies occur. In practice, it is the difference between organizations that can scale AI confidently and those that are trapped by manual oversight and reactive governance.

The concept of a **force multiplier** also becomes tangible in this context. Every automated decision, policy evaluation, or identity verification executed by the system represents additional assurance achieved without additional headcount or process friction. When governance and telemetry are integrated directly into the technology stack, security ceases to be an external constraint and becomes part of the system's natural motion. This is how Zero Trust translates from architectural intent into operational advantage.

However, effective adoption does not begin with aspirational maturity models or abstract patterns. It begins with a **minimum viable baseline**: the smallest coherent set of controls that delivers verifiable protection across identity, device, network, application, and data layers. This baseline creates the conditions for AI to operate safely. It is measurable, repeatable, and achievable within a defined timeframe - typically two to three quarters of disciplined rollout - if approached methodically.

Zero Trust readiness, therefore, is not defined by the number of technologies deployed but by the quality of control implementation and the ability to observe and adjust those controls in real time. The journey starts with visibility. Controls are introduced in monitoring mode, tuned based on real data, and then gradually shifted into enforcement. Once stability is confirmed, automation takes over to sustain compliance and reduce drift. This phased approach replaces one-time hardening with a living feedback loop - a control system that learns from its own telemetry.

Recent threat data underscores why this discipline matters. Reports such as IBM's *X-Force Threat Intelligence Index* and Verizon's *Data Breach Investigations Report* show that credential misuse, supply-chain compromise, and data exfiltration remain the top enterprise breach vectors. These are precisely the risks mitigated by the core Zero Trust pillars: identity assurance, segmentation, and continuous verification. As AI expands the scale and autonomy of digital operations, these controls become not only security imperatives but business requirements for sustained trust.

The sections that follow define this minimum viable Zero Trust baseline, explain how it underpins AI readiness, and demonstrate how its staged rollout transforms architectural intent into daily operational reality. By translating principle into measurable control, Zero Trust becomes the true infrastructure of the AI era - the system of verification through which innovation can safely accelerate.

From Guardrails to Growth

Security controls can increase agility when designed into the system's normal operating path. Enterprise security architecture literature frames controls as mechanisms that enable movement by bounding consequences. The Zero Trust model advanced this position by replacing implicit trust with explicit verification and continuous evaluation. For AI, this means identity-bound access for people and machine actors, segmented compute and data paths, and policy-enforced behavior for training and inference. Properly implemented, controls do not slow delivery; they shorten decision cycles by making accept-or-deny outcomes immediate and auditable.

The implication for AI programs is direct. A service account does not receive broad, static rights. It receives a short-lived scope tied to a purpose and a context signal. A model does not pull raw data by convention. It requests through a broker that enforces classification and redaction rules. This architecture replaces review queues with real-time approval by policy, turning guardrails into growth mechanisms.

Zero Trust Baseline for AI Readiness

A credible AI foundation requires consistent maturity across pillars and supporting capabilities. The architectural lineage of this approach traces to the SABSA framework's emphasis on traceability between business drivers and technical controls. Its contextual-to-component logic remains relevant as a means of aligning risk and value. The same principle of continuous alignment, assurance, and adaptation underlies the phased rollout model presented here: visibility first, enforcement second, automation last. Beyond that reference, the implementation described in this paper stands on its own as an operationally grounded evolution of Zero Trust.

The relationship between Zero Trust and the NIST AI Risk Management Framework becomes clear when viewed through operational expectations. AI RMF describes the work required to map risks, measure behavior, manage controls, and govern AI systems. Each function presumes identity integrity, authoritative inventories, provenance, enforceable access, and continuous verification. Without these prerequisites, AI RMF becomes a policy description without a path to execution. Zero Trust supplies the operational substrate that AI RMF depends on. Identity provides a stable reference. Segmentation provides containment. Telemetry provides visibility. Verification provides integrity. Together these elements make AI RMF actionable rather than aspirational. This relationship is not hierarchical. Zero Trust does not supersede AI RMF, and AI RMF does not override Zero Trust. Each operates at a different layer, with Zero Trust defining the environment and AI RMF governing the lifecycle and risk posture of the AI systems within it.

What is Zero Trust?

Zero Trust is a security model that continuously verifies every access request to **ANY resource** and **asset**, based on dynamic context, assuming no implicit trust and emphasizing least privilege, segmentation, and breach containment. The default stance is to **Assume Breach**.

Minimum Viable Zero Trust Baseline

The minimum viable Zero Trust baseline represents the smallest enforceable set of controls that delivers measurable risk reduction while enabling AI systems to operate safely at enterprise scale. It focuses on five pillars - identity, device, network, application, and data - implemented with observability and governance as cross-cutting functions. At minimum, a viable baseline includes the following capabilities: enforced MFA for privileged identities, verified device compliance for interactive sessions, segmentation between training and inference environments, API gateways enforcing identity propagation, and DLP controls applied to sensitive data classifications.

Observability and governance function as binding layers across all pillars because accountability and telemetry determine whether the controls behave as intended.

Implementing these baseline controls establishes the conditions required for safe innovation. Each capability can be introduced incrementally, starting in monitoring mode, verified through telemetry, and advanced to enforcement once confidence is established. The following sections

describe how this staged rollout progresses across control families and how governance mechanisms sustain assurance once enforcement is active.

Together, these capabilities define the operational minimum for Zero Trust enablement. They form a measurable baseline that can be deployed incrementally without architectural redesign and provide the prerequisite assurance required for advanced AI adoption. In this context, automation refers to the point where policy enforcement and remediation occur without manual intervention.

In this context, *posture* refers to an internal assertion of alignment with the baseline controls described in this paper, not an external compliance certification.

Why this Matters

Conservative modeling from 2024 industry reports **places the average cost of a breach between \$8MM-10MM per event**, with credential misuse and lateral movement as primary vectors.

Each missing control multiplies the blast radius of every failure. This effect appears not only in the technical impact of incidents but also in insurance disputes, where carriers increasingly deny claims when controls documented on paper are not present in practice.

Zero Trust foundational controls do not eliminate breaches, but they sharply limit cost propagation and claim denials.

This is the financial definition of containment.

Ad-hoc – No Zero Trust Posture

Domain	Capability Name	Level 1 – Legacy / Ad Hoc
DETECT	Detect Anomalies and Events	Ad hoc alerts; minimal tuning; high false positives.
DETECT	Monitor Systems and Assets	Partial log collection; limited retention; blind spots.
DETECT	Manage Detection & Readiness Protocols	No SOC playbooks; alert handling varies by analyst.
GOVERNANCE	Establish Organizational Direction	No formal InfoSec charter; priorities shift with incidents; unclear authority.
GOVERNANCE	Define Risk Management Strategy	Risk treated informally; no risk appetite or criteria.
GOVERNANCE	Manage InfoSec Resources	People/budget reactive; tooling sprawl; unclear capacity.
GOVERNANCE	Define InfoSec Policy	Policies incomplete/outdated; inconsistent enforcement.
GOVERNANCE	Provide InfoSec Oversight	Ad hoc reviews; minimal leadership reporting.
GOVERNANCE	Manage Supply Chain Security Risks	Vendor risk unmanaged or one-off questionnaires.
IDENTITY	Catalog and Manage Security Assets	Unknown assets; manual spreadsheets; shadow IT.
IDENTITY	Assess and Evaluate Risks	No standard method; sporadic assessments.
IDENTITY	Optimize Risk Identification Protocols	KRIs undefined; manual signal collection.
LEGAL	Support Legal Discovery	eDiscovery handled case-by-case; data not preserved.
LEGAL	Investigative Support	Forensics ad hoc; evidence handling inconsistent.
LEGAL	Regulatory & Contractual Readiness	Unclear obligations; reactive responses.
PROTECT	Manage Identities and Access	Shared accounts; weak auth; inconsistent offboarding.
PROTECT	Train and Educate Personnel	One-off or annual CBT only; low engagement.
PROTECT	Protect and Secure Data	Unclassified data; inconsistent encryption; broad access.
PROTECT	Implement Protective Measures	Baseline configs inconsistent; patching informal.
PROTECT	Maintain Systems and Infrastructure	Manual maintenance; undocumented dependencies.
PROTECT	Deploy Protective Technologies	Tools deployed inconsistently; coverage gaps.
RECOVER	Plan and Execute Recovery	Backups unreliable; recovery untested.
RECOVER	Optimize Recovery Protocols	Improvements opportunistic; little measurement.
RECOVER	Manage Recovery Communication	Stakeholders uninformed; milestones unclear.
RESPOND	Response Planning & Management	No formal IR plan; unclear roles; ad hoc approvals.
RESPOND	Manage Security Event Communication	Uncoordinated communications; stakeholders surprised.
RESPOND	Assess Event Scope & Impact	Root cause unclear; scoping slow.
RESPOND	Mitigate and Contain Event Impact	Manual containment; inconsistent eradication.
RESPOND	Optimize Response Protocols (Improvements)	Post-incident reviews rare; lessons not tracked.

Ad-Hoc: Most Organizations are at this Level or below.

Ad-Hoc is the default for many organizations nowadays. Lack of Zero Trust MVP posture also means partially significant gaps and security risks that can translate into actual financial liabilities due to insufficient risk mitigation and management.

Legacy environments rely on implicit trust, perimeter filtering, and periodic audits. Controls are ad hoc, ownership is unclear, and verification is retrospective rather than continuous. The minimum viable Zero Trust baseline replaces this with explicit verification at every interface: identities are authenticated with strong assurance, device posture is validated at session start, data is labeled and governed, and all decisions are logged. The shift is measurable: from assumed safety at the edge to demonstrated control in real time.

Zero Trust Foundational: The Minimum Viable Zero Trust Baseline

When Zero Trust Foundational is fully implemented then the baseline control families form a coherent operating minimum:

- enforceable identity assurance,
- verified device and workload integrity,
- segmented environments,
- policy-aware APIs, and
- data protection with lineage.

This is not an aspirational maturity tier; it is the smallest enforceable set of controls that enables AI systems to run safely at enterprise scale.

Domain	Capability Name	Level 2 – Foundational Zero Trust
DETECT	Detect Anomalies and Events	Use-cases mapped to risks; tuning lifecycle; detections across identity/device/network/app/data.
DETECT	Monitor Systems and Assets	Central log ingestion with required retention; monitoring SLAs; coverage for privileged activity.
DETECT	Manage Detection & Readiness Protocols	SOC runbooks, tiering, SLA/MTR; purple-team exercises; shift handoffs standardized.
GOVERNANCE	Establish Organizational Direction	Approved InfoSec charter, mission, roadmap; roles/RACI defined; ZT principles (verify explicitly, least privilege, segment) embedded in strategy.
GOVERNANCE	Define Risk Management Strategy	Documented risk appetite/tolerance; recurring assessments; baseline ZT risk criteria (identity, device, workload, data, network).
GOVERNANCE	Manage InfoSec Resources	Workforce plan; budget tied to risk reduction; central tool inventory; shared services identified (SOC/IAM).
GOVERNANCE	Define InfoSec Policy	Policy suite current, approved; mapped to ZT (identity verification, segmentation, least privilege, telemetry).
GOVERNANCE	Provide InfoSec Oversight	Formal oversight bodies; standard KPIs/KRIs; ZT posture reviews across identity/device/network/data.
GOVERNANCE	Manage Supply Chain Security Risks	Tiered assessments; minimum ZT controls in contracts (MFA, logging, segmentation); continuous monitoring for critical suppliers.
IDENTITY	Catalog and Manage Security Assets	Authoritative inventories (devices, apps, identities, data, vendors) with owners.
IDENTITY	Assess and Evaluate Risks	Defined methodology; periodic assessments; threat modeling for key systems.
IDENTITY	Optimize Risk Identification Protocols	KRIs across identity/device/workload/data; standardized intake; escalation thresholds.
LEGAL	Support Legal Discovery	Legal hold process; preservation/collection procedures; counsel engaged; privacy-by-design.
LEGAL	Investigative Support	Forensic readiness plan; imaging procedures; secure storage; role-based access.
LEGAL	Regulatory & Contractual Readiness	Mapped obligations (HIPAA/PCI/GDPR/GLBA, contracts); standard clauses for ZT controls in MSAs.
PROTECT	Manage Identities and Access	Central IAM; MFA enforced; RBAC/SoD; periodic reviews; device posture checks for sensitive access.
PROTECT	Train and Educate Personnel	Role-based training (incl. privileged users); phishing simulations; developer secure coding; ZT awareness.
PROTECT	Protect and Secure Data	Data classification; DLP; encryption at rest/in transit; key mgmt; least privilege on data stores.
PROTECT	Implement Protective Measures	Hardened baselines (CIS Benchmarks); patch SLAs; vuln mgmt; change control.
PROTECT	Maintain Systems and Infrastructure	Standard maintenance windows; dependency maps; backup/restore validated for critical systems.
PROTECT	Deploy Protective Technologies	Standard stack (EDR, NGFW, email/web security) with coverage targets; central management.
RECOVER	Plan and Execute Recovery	RTO/RPO defined; tiered apps; regular restore tests; offline/immutable backups.
RECOVER	Optimize Recovery Protocols	Post-recovery reviews; prioritized backlog; dependency hardening.
RECOVER	Manage Recovery Communication	Recovery comms plan; status templates; external customer scripts.
RESPOND	Response Planning & Management	IR plan approved/tested; roles/RACI; legal/PR engaged; ZT containment patterns defined.
RESPOND	Manage Security Event Communication	Comms matrix, templates, notification thresholds; planned cadence.
RESPOND	Assess Event Scope & Impact	Standard analysis procedures; identity/device/data scoping; forensics support.
RESPOND	Mitigate and Contain Event Impact	Standard options (isolate device, disable token, block IP); eradication procedures; patch/IOC deployment.
RESPOND	Optimize Response Protocols (Improvements)	Blameless postmortems; actions tracked to closure; content/playbook updates.

Foundational Zero Trust: Overview over recommended baseline.

Zero Trust Foundational: Controls to Implement

Technical controls are observable and enforceable by default, with telemetry, exception tracking, and drift detection integrated into daily operations. Below shows an acceptable list of technical controls that can be put in place, and once completed, a Zero Trust Foundational posture can be claimed. Note that each control represents a significant effort on an organizational level and requires Leadership commitment and alignment with priorities.

Governance Meta Layer	Pillar/Capability	Configuration/Feature Required	Standard/Control/Reference
DETECT / RESPOND	Visibility	Logs are aggregated from identity, device, and network systems into a central SIEM.	CISA ZTMM Visibility; NIST SP 800-53 AU-6; CIS Control 8.2
DETECT / RESPOND	Visibility	Security teams perform regular manual log reviews and alert triage.	CISA ZTMM Visibility; CIS Control 8.3
DETECT / RESPOND	Visibility	Dashboards are established to visualize overall security posture.	Microsoft Sentinel; CISA ZTMM Visibility
DETECT / RESPOND	Vulnerability Management	A centralized vulnerability scanning platform is deployed.	CISA BOD 22-01; NIST SP 800-40; CIS Control 7.1
DETECT / RESPOND	Vulnerability Management	Critical and high vulnerabilities are remediated within defined service-level targets (typically 7–30 days).	NIST SP 800-40; CISA BOD 22-01; CIS Control 7.3
DETECT / RESPOND	Vulnerability Management	Scan results are reviewed and tracked through the SIEM or equivalent dashboard.	CISA ZTMM; NIST SP 800-53 RA-5
GOVERNANCE / LEGAL	Governance	A formal Zero Trust strategy document is created and approved.	CISA ZTMM Governance; NIST SP 800-53 PL-2
GOVERNANCE / LEGAL	Governance	Ownership for each Zero Trust pillar is assigned and maintained.	CISA ZTMM Governance; Microsoft Zero Trust Strategy Guide
GOVERNANCE / LEGAL	Governance	Security awareness training includes Zero Trust principles and practices.	CISA ZTMM Governance; NIST SP 800-53 AT-2
GOVERNANCE / LEGAL	Governance	A single policy or standard operating procedure reflects the organization's Zero Trust architecture.	CISA ZTMM Governance; NIST SP 800-207
PROTECT	Applications	Single sign-on is configured for all cloud-based applications.	Microsoft Zero Trust Applications; CIS Control 6.2
PROTECT	Applications	Conditional Access is enforced for sensitive or privileged applications.	Microsoft CA; CISA ZTMM Applications; NIST SP 800-53 AC-6
PROTECT	Applications	Privileged Identity Management (PIM) is used to control administrative role elevation.	Microsoft Entra PIM; CISA ZTMM Applications; CIS Control 4.3
PROTECT	Applications	Manual access reviews are performed for sensitive application roles.	CISA ZTMM Applications; CIS Control 6.3
PROTECT	Data	A data classification scheme is defined, approved, and documented.	CISA ZTMM Data; CIS Control 3.4; NIST SP 800-60
PROTECT	Data	Data loss prevention (DLP) is enforced in at least one communication channel.	Microsoft DLP; CISA ZTMM Data; CIS Control 3.5
PROTECT	Data	Audit logging is enabled across all data repositories.	Microsoft Purview; NIST SP 800-53 AU-2, AU-12; CISA ZTMM Data
PROTECT	Data	Manual access reviews are conducted for data sets containing sensitive information.	CISA ZTMM Data; CIS Control 3.8
PROTECT	Devices	Device registration (Azure AD Join or Azure AD Registered) is enabled for all managed devices.	CISA ZTMM Devices; Microsoft Zero Trust Devices; CIS Control 1.1
PROTECT	Devices	Device compliance policies are enforced, including encryption, antivirus, and OS patch levels.	Microsoft Intune; CISA ZTMM Devices; CIS Control 4.1, 4.4
PROTECT	Devices	Device compliance status is used as a condition in Conditional Access policies.	Microsoft Conditional Access; CISA ZTMM Devices
PROTECT	Devices	A partial device inventory is maintained within a mobile device management platform.	CISA ZTMM Devices; CIS Control 1.2
PROTECT	Identity	Multi-factor authentication is enforced for all users through Conditional Access.	CISA ZTMM v2.0 Identity; NIST SP 800-63B (AAL2); CIS Control 6.3
PROTECT	Identity	Legacy authentication protocols (IMAP, POP3, SMTP Basic) are disabled.	Microsoft Security Baseline; CISA ZTMM; CIS Control 6.7
PROTECT	Identity	Conditional Access policies evaluate user risk, device compliance, and application sensitivity.	Microsoft CA; CISA ZTMM Identity; NIST SP 800-207
PROTECT	Identity	Single sign-on integration is established with Microsoft Entra ID for core enterprise applications.	CISA ZTMM; Microsoft Zero Trust Identity; CIS Control 6.1
PROTECT	Network	Internal network segmentation is implemented (for example, VLANs or network security groups).	CISA ZTMM Network; NIST SP 800-207; CIS Control 13.1
PROTECT	Network	A Secure Web Gateway or proxy is used for outbound traffic inspection.	CISA ZTMM Network; Microsoft Defender for Endpoint; CIS Control 13.10
PROTECT	Network	Location-based Conditional Access policies restrict access from untrusted or high-risk regions.	Microsoft CA Named Locations; CISA ZTMM Network
RECOVER	Automation	At least one alert or workflow is automated (for example, user onboarding or alert notifications).	CISA ZTMM Automation; Microsoft Sentinel Playbooks
RECOVER	Automation	A documented manual playbook or standard operating procedure exists for incident response.	CISA ZTMM Automation; CIS Control 17.3

Foundational Zero Trust: Technical controls.

Zero Trust Foundational: Evidence to Provide

Concrete evidence can be supplied to ensure the proper posture is achieved. Some examples of artifacts that can be gathered and provided can be seen in the next table below.

Domain	Capability Name	Foundational Zero Trust - Evidence Examples
DETECT	Detect Anomalies and Events	Use-case catalog; tuning tickets; UEBA models; FP/FN trend reports.
DETECT	Monitor Systems and Assets	Ingestion map; retention policies; coverage dashboards; attestation reports.
DETECT	Manage Detection & Readiness Protocols	Runbooks; SOAR workflows; exercise reports; MTTR metrics; QA checklists.
GOVERNANCE	Establish Organizational Direction	Approved charter, strategy deck, RACI, OKRs, governance calendar, board/steering minutes.
GOVERNANCE	Define Risk Management Strategy	Risk policy; risk register with owners; KRI dashboards; signed risk acceptances; quarterly risk reports.
GOVERNANCE	Manage InfoSec Resources	Headcount plan; skills matrix; budget vs outcome map; service catalog; vendor map; renewal calendar.
GOVERNANCE	Define InfoSec Policy	Policy library; approval logs; exception register; control-to-policy mappings.
GOVERNANCE	Provide InfoSec Oversight	Steering minutes; KPI packs; closed audit findings; supplier scorecards.
GOVERNANCE	Manage Supply Chain Security Risks	TPRM policy; assessment records; contractual clauses; access reviews; SBOMs; offboarding logs.
IDENTITY	Catalog and Manage Security Assets	CMDB/asset registry; discovery reports; ownership matrix; posture scores.
IDENTITY	Assess and Evaluate Risks	Assessment reports; risk heatmaps; threat models; ASM dashboards; test reports.
IDENTITY	Optimize Risk Identification Protocols	KRI catalog; intake SOPs; correlation rules; simulations.
LEGAL	Support Legal Discovery	Legal hold notices; collection logs; chain-of-custody; counsel opinions; retention proofs.
LEGAL	Investigative Support	Forensic SOPs; toolkit & images; chain-of-custody; lab accreditation; case timelines.
LEGAL	Regulatory & Contractual Readiness	Obligations register; clause library; audit reports; customer trust portal.
PROTECT	Manage Identities and Access	MFA logs; SSO configs; access review certs; PAM/JIT records; device posture policies.
PROTECT	Train and Educate Personnel	Training plan; completion/assessment scores; simulation metrics; curriculum; cert records.
PROTECT	Protect and Secure Data	Classification policy; DLP alerts; key rotation logs; data access reviews; lineage reports.
PROTECT	Implement Protective Measures	Baseline standards; patch compliance; vuln trends; change tickets; drift reports.
PROTECT	Maintain Systems and Infrastructure	Runbooks; CMDB dependencies; maintenance calendar; SLO reports; backup verification logs.
PROTECT	Deploy Protective Technologies	Coverage reports; policy configs; purple-team results; architecture diagrams.
RECOVER	Plan and Execute Recovery	BCP/DR plans; restore test results; immutable backup logs; RTO/RPO reports.
RECOVER	Optimize Recovery Protocols	Review reports; backlog items; exercise/chaos logs; resilience KPI dashboards.
RECOVER	Manage Recovery Communication	Comms plan; status logs; customer notices; regulator correspondence.
RESPOND	Response Planning & Management	IR plan; test/exercise reports; approval matrix; containment playbooks.
RESPOND	Manage Security Event Communication	Comms plan; stakeholder lists; notification logs; post-incident comms review.
RESPOND	Assess Event Scope & Impact	Triage SOP; case notes; chain-of-custody; scoping timelines.
RESPOND	Mitigate and Contain Event Impact	Containment playbooks; SOAR actions; IOC deployment logs; success/rollback rates.
RESPOND	Optimize Response Protocols (Improvements)	AARs; action tracker; metrics pack; updated runbooks.

Foundational Zero Trust: Evidence that can illustrate successful implementation.

Zero Trust Optimized: Evolved and Advanced Zero Trust Posture

The table below describes the highest level of Zero Trust Posture: **Optimized Zero Trust**. Implemented Optimized Zero Trust requires **significant** investment and support by the entire organizational management chain as well as the employees. Many of the controls are usually out of reach for most organizations, due to cost implications and operational impact.

Domain	Capability Name	Level 3 – Optimized / Service-Ready Zero Trust
DETECT	Detect Anomalies and Events	UEBA/behavioral analytics; detections adapt to trust scores; threat intel & deception integrated.
DETECT	Monitor Systems and Assets	Cross-telemetry correlation (identity, device, network, cloud, SaaS); multi-tenant monitoring; SOC2/ISO attestations.
DETECT	Manage Detection & Readiness Protocols	SOAR automation; continual readiness drills; QA gates on content; customer runbooks for service delivery.
GOVERNANCE	Establish Organizational Direction	Strategy cascades to portfolio/program OKRs; ZT operating model measured quarterly across BUs and suppliers; external assurance of governance.
GOVERNANCE	Define Risk Management Strategy	Continuous risk scoring tied to business impact; automated policy adjustment; external regulatory alignment across portfolio companies.
GOVERNANCE	Manage InfoSec Resources	Service catalog with cost models; capacity & coverage SLOs; chargeback/showback; onboarding playbooks for portfolio companies.
GOVERNANCE	Define InfoSec Policy	Policy-as-code/guardrails; exception lifecycle with expiry & business owner accountability.
GOVERNANCE	Provide InfoSec Oversight	Independent assurance/audits; exec risk dashboards; supplier/portfolio oversight integrated.
GOVERNANCE	Manage Supply Chain Security Risks	Adaptive 3rd-party access; SBOM/SCRM integrated; high-risk suppliers attested; automated offboarding.
IDENTITY	Catalog and Manage Security Assets	Real-time discovery; asset posture feeds policy; unmanaged assets quarantined.
IDENTITY	Assess and Evaluate Risks	Continuous risk analytics; attack surface mgmt; red/purple teaming tunes controls.
IDENTITY	Optimize Risk Identification Protocols	Automated/correlated risk signals; policy adapts to KRI movement; impact simulations.
LEGAL	Support Legal Discovery	Repeatable eDiscovery service; tooling integrated with data governance; cross-tenant holds with audit trails.
LEGAL	Investigative Support	Accredited lab/process; timeline & artifact correlation; tenant-aware investigations.
LEGAL	Regulatory & Contractual Readiness	Proactive attestations (SOC 2 Type II/ISO 27001); regulated data service-readiness; evidence portals for customers.
PROTECT	Manage Identities and Access	Adaptive access (risk/behavior/device); JIT/JEA; continuous verification; strong federation; PAM integrated.
PROTECT	Train and Educate Personnel	Personalized, metrics-driven learning; just-in-time nudges; certification programs; supplier/user onboarding packages.
PROTECT	Protect and Secure Data	Attribute/label-based access; tokenization; privacy engineering; continuous lineage & usage analytics.
PROTECT	Implement Protective Measures	Immutable images; golden pipelines with policy-as-code; risk-based patch orchestration & auto-remediation.
PROTECT	Maintain Systems and Infrastructure	Self-healing infra; SRE practices; SLOs & error budgets; infra compliance attested for regulated tenants.
PROTECT	Deploy Protective Technologies	Control efficacy measured vs ATT&CK; policy orchestration; multi-tenant-ready for service delivery.
RECOVER	Plan and Execute Recovery	Automated recovery with attestation; rejoin requires re-verification (identity/device/workload); cross-tenant recovery playbooks.
RECOVER	Optimize Recovery Protocols	Game-day exercises; chaos testing in lower envs; resilience metrics in exec scorecards.
RECOVER	Manage Recovery Communication	Joint comms with partners/regulators; real-time status portals; SLA-aligned updates.
RESPOND	Response Planning & Management	Orchestrated response with automated policy actions; enterprise playbooks; external SLAs supported.
RESPOND	Manage Security Event Communication	Regulatory & timely breach notifications; joint comms with partners; predefined external SLA language.
RESPOND	Assess Event Scope & Impact	Automated scoping with identity/asset graphing; blast-radius estimation; repeatable for service customers.
RESPOND	Mitigate and Contain Event Impact	Automated kill-switch policies; least-privilege restoration; coordinated supplier/tenant actions.
RESPOND	Optimize Response Protocols (Improvements)	Metrics-driven improvements (MTTD/MTTR/dwell); customer feedback loop; control efficacy informs roadmap.

Optimized Zero Trust: Advanced Deployments and Compliance Posture.

Zero Trust Foundational: Mapping to existing frameworks and standards

All the technical controls mentioned above are solidly grounded in existing frameworks and standards, as can be seen in the next table. This alignment draws directly from **NIST CSF 2.0**, the **CISA Zero Trust Maturity Model**, and the **CSA Cloud Controls Matrix**, which together provide the scaffolding for the baseline controls described here.

AI RMF is not mapped in this table because it governs the AI lifecycle rather than operational controls. It depends on the foundational Zero Trust capabilities established earlier in this document.

Domain	Capability Name	Key References
DETECT	Detect Anomalies and Events	NIST CSF DE.AE; ISO 27035; MITRE D3FEND
DETECT	Monitor Systems and Assets	NIST CSF DE.CM; ISO 27001 A.5/A.8; CIS 8
DETECT	Manage Detection & Readiness Protocols	NIST CSF DE.DP; NIST 800-61r2; ISO 27035
GOVERNANCE	Establish Organizational Direction	NIST CSF ID.GV; ISO 27014; COBIT EDM/APO; CIS v8 IG2/IG3
GOVERNANCE	Define Risk Management Strategy	NIST CSF ID.RM; ISO 31000; COSO ERM; NIST 800-207 (signals)
GOVERNANCE	Manage InfoSec Resources	COBIT APO07/APO12; ISO 27001 A.5 & A.6; NIST CSF GOV
GOVERNANCE	Define InfoSec Policy	ISO 27001/27002 A.5; NIST CSF ID.GV; CIS v8
GOVERNANCE	Provide InfoSec Oversight	COBIT EDM; ISO 27014; NIST CSF GOV; SOC 2
GOVERNANCE	Manage Supply Chain Security Risks	NIST CSF ID.SC; NIST 800-161r1; ISO 27036; CIS Control 15
IDENTITY	Catalog and Manage Security Assets	NIST CSF ID.AM; CIS 1-2; ISO 27002 A.5/A.8
IDENTITY	Assess and Evaluate Risks	NIST CSF ID.RA; ISO 31000; MITRE ATT&CK; NIST 800-30
IDENTITY	Optimize Risk Identification Protocols	NIST CSF ID.RM; COSO ERM; NIST 800-207 (signals)
LEGAL	Support Legal Discovery	EDRM; ISO 27050; ISO 27701
LEGAL	Investigative Support	ISO 27037/27041; NIST 800-101; SWGDE
LEGAL	Regulatory & Contractual Readiness	SOC 2; ISO 27001/27018; PCI DSS; HIPAA; GDPR
PROTECT	Manage Identities and Access	NIST CSF PR.AC; NIST 800-207; CIS 5-6; ISO 27001 A.5/A.8/A.9
PROTECT	Train and Educate Personnel	NIST CSF PR.AT; ISO 27002 A.6.3; NIST SSDF
PROTECT	Protect and Secure Data	NIST CSF PR.DS; ISO 27018/27701; CIS 3; NIST 800-57
PROTECT	Implement Protective Measures	NIST CSF PR.IP; CIS Benchmarks; NIST 800-40; ITIL Change
PROTECT	Maintain Systems and Infrastructure	NIST CSF PR.MA; SRE; ISO 20000; ISO 27001 A.8
PROTECT	Deploy Protective Technologies	NIST CSF PR.PT; MITRE ATT&CK; CIS Controls
RECOVER	Plan and Execute Recovery	NIST CSF RC.RP; ISO 22301; NIST 800-184
RECOVER	Optimize Recovery Protocols	NIST CSF RC.IM; ISO 22301; SRE
RECOVER	Manage Recovery Communication	NIST CSF RC.CO; ISO 22301; RS.CO
RESPOND	Response Planning & Management	NIST CSF RS.RP; ISO 27035; NIST 800-61r2
RESPOND	Manage Security Event Communication	NIST CSF RS.CO; ISO 27035; applicable laws/regulations
RESPOND	Assess Event Scope & Impact	NIST CSF RS.AN; NIST 800-61r2; MITRE ATT&CK
RESPOND	Mitigate and Contain Event Impact	NIST CSF RS.MI; ISO 27035; NIST 800-61r2
RESPOND	Optimize Response Protocols (Improvements)	NIST CSF RS.IM; ISO 27035; SRE

Foundational Zero Trust: Standard and Framework Mapping.

From Architecture to Operation: Rolling Out Foundational Controls

Zero Trust becomes real when controls are introduced in staged, observable steps. The practical sequence is monitor first, enforce second, automate last. The following rollouts represent a minimum viable baseline that organizations can execute within roughly 120 days while maintaining service stability. Each control family starts in report-only mode, adds governance for exceptions, then shifts into scoped enforcement and finally broad enforcement with automation. These rollouts are the practical expression of the minimum viable Zero Trust baseline. They show that enablement is achieved by disciplined phasing, not overnight switches.

Monitoring builds trust in controls. Scoped enforcement contains risk while learning from production behavior. Automation closes the loop and turns governance into an operating condition. This is the path from architecture to operation that teams can run starting tomorrow.

Conditional Access Policies and MFA Rollout

Phase	Description	Policy	MONTH								
			1	2	3	4	5	6	7	8	9
0	Preparation	Create Policies and Apply in Report-only mode									
0	Governance	document exceptions									
0	Preparation	Enable Logging and Monitoring for Policies									
0	Preparation	Modify/Finetune Policies before enforcement									
1	Immediate Risk Blocking	Block Legacy Authentication									
1	Immediate Risk Blocking	Require MFA for All Admins									
2	Risk-Adaptive Controls	Require MFA for High-Risk Users									
2	Device Compliance Enforcement	Block Access from Unmanaged Devices									
3	Risk-Adaptive Controls	Require MFA Outside Trusted Locations									
4	Geo Restriction Enhancements	Block Access from Unsupported Countries/Regions									

LEGEND

Phase 0: Preparation
Phase 1: Immediate Risk Blocking
Phase 2: Risk-Adaptive Controls
Phase 3: Risk-Adaptive Controls
Phase 4: Geo Restriction Enhancements

Phase 0 Preparation and Governance form the foundation for this rollout. Policies are created and placed in report only mode so the environment can be observed without impact. Logging captures behavior patterns, false positives, false negatives, and anomalies. Policy refinement happens during this period. All exceptions are added to the risk register with clear ownership and expiration dates. This phase establishes visibility and accountability before enforcement begins.

Phase 1 Immediate Risk Blocking removes the most dangerous access paths. Legacy authentication is eliminated because it bypasses MFA. All administrative accounts are required to use MFA. Privileged access becomes the first hardened boundary, which prevents compromise at the highest consequence layer.

Phase 2 Risk Adaptive Controls and Device Compliance ties trust decisions to context. Users flagged as high risk by identity protection signals must complete MFA. Device posture becomes mandatory. Unmanaged devices are denied access. Identity and device integrity now shape the trust decision together.

Phase 3 Contextual Friction adds location-based verification. This phase comes before geographic restrictions because contextual checks surface issues with lower user impact, which allows corrections before tighter boundaries are applied. Access attempts from outside-trusted locations require additional authentication. This reduces misuse tied to unexpected sign in locations without affecting normal daily activity.

Phase 4 Geo Restriction Enhancements limits access to approved countries and regions. This reduces the exposure surface tied to high-risk geographies and finalizes the tightening of access control across the environment.

Common pitfalls to avoid

- Enforcing unmanaged device blocks before validating break-glass access.
- Enabling geo restrictions without confirming emergency operations access for executives and admins on travel.
- Failing to publish a short user-facing FAQ and/or communication before each step change.

Vulnerability Remediation

Phase	Description	Policy	MONTH								
			1	2	3	4	5	6	7	8	9
0	Preparation	Align VM Policies									
0	Preparation	Asset Identification									
0	Preparation	Confirm Findings on Assets									
0	Governance	Document Exceptions									
1	Remediate	Zero-Day (actively exploited)									
2	Remediate	Critical / Prioritized High									
2	Remediate	High									
3	Remediate	Medium / Low									

LEGEND

Phase 0: Preparation
Phase 1: Immediate Risk Blocking
Phase 2: Risk-Adaptive Controls
Phase 3: Risk-Adaptive Controls

Phase 0 Preparation and Governance align vulnerability management with the broader security baseline. All assets in scope are identified and mapped to owners. Findings are validated to reduce noise. Exceptions are recorded in the risk register with justification and expiration. This creates a mature signal before remediation pressure begins. The same classification scheme and DLP controls apply to data entering or leaving AI services, including retrieval systems and internal model platforms.

Phase 1 Zero Day Handling triggers immediate isolation for systems affected by actively exploited vulnerabilities. Patching and testing proceed in parallel to reduce exposure time. This is the highest urgency condition, and the workflow reflects that urgency.

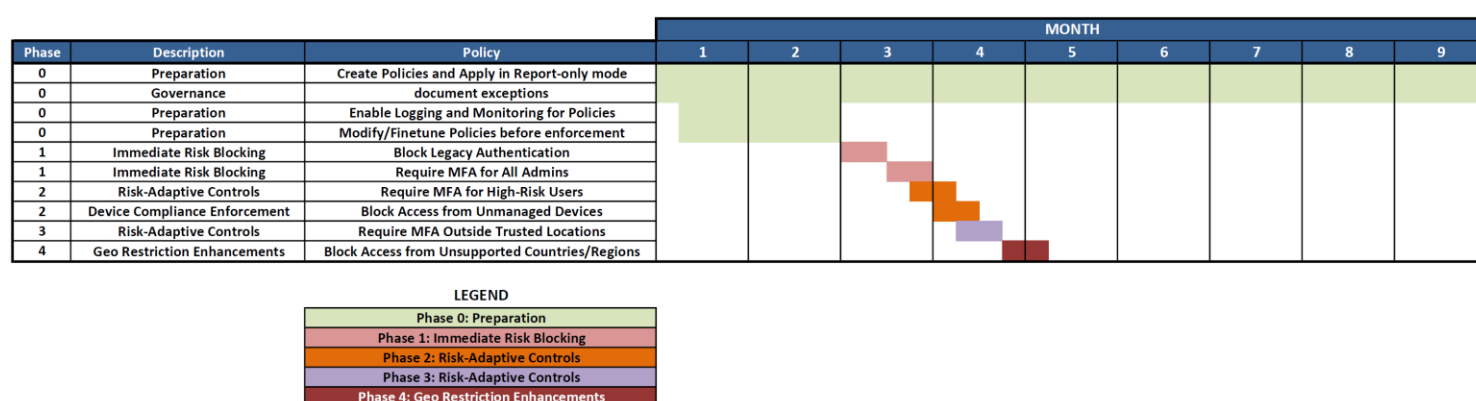
Phase 2 Critical and High Severity Remediation enforces strict timelines for high impact issues. Critical and high findings older than thirty or sixty days appear in leadership dashboards so accountability cannot drift. Remediation is expected and directly tied to named system owners.

Phase 3 Medium and Low Severity Governance manages the long tail of the backlog. In this model, findings older than 180 days become exceptions rather than ignored items, which serves as a policy target rather than a statement of current industry reality. This prevents slow moving exposures from settling into operational inertia.

Common pitfalls to avoid

- Treating scanner output as ground truth without validation.
- Allowing silent ownership gaps; every asset requires a named owner with escalation path.
- Closing tickets without confirming patch effectiveness by rescanning.

Data Classification and Data Loss Prevention



Phase 1 Taxonomy Definition finalizes the label model and ensures each label has a clear definition and protection logic. This creates the identity of the data and sets the semantic boundary for all downstream controls.

Phase 2 Passive Labeling Enablement turns on labeling across e.g. Exchange, SharePoint, OneDrive, and Teams without forcing any user action. The environment begins to reveal how sensitive content moves in real life, which is often different from architectural assumptions.

Phase 3 Stabilization and Audit Observation uses *audit only* mode to gather signal quality. The program observes exposure routes, accidental misclassifications, false positives, and usage patterns. That observation informs the tuning that follows.

Phase 4 Mandatory Tagging Adoption requires labeling during creation or upload in areas that already show stable behavior. Classification becomes a routine part of daily work, and the identity layer shifts from theoretical to operational.

Phase 5 Awareness through Policy Tips introduces real time prompts without blocking. Users receive cues based on label behavior which conditions the environment for upcoming enforcement. This reduces friction and improves user accuracy.

Phase 6 Scoped Enforcement for High Sensitivity Labels applies justify or block controls only to the labels with the highest risk profile. Other labels remain in audit or awareness mode. This delivers meaningful protection without destabilizing the workforce.

Phase 7 Broad Enforcement Expansion widens enforcement once the organization demonstrates stable labeling habits and mature exception handling. Routing and response processes are predictable by this point, which allows wider coverage.

Phase 8 Collaboration Channel Coverage extends DLP to Teams messages and chat-based file sharing. Because the earlier stages have normalized the behavior model, this final phase lands with minimal friction and completes the control surface.

Common pitfalls to avoid

- Skipping taxonomy confirmation, which leads to mislabeling downstream.
- Moving to block without a training period using policy tips.
- Enabling Teams DLP without testing guest sharing scenarios.

Cross-cutting Guidance Across All Control Families

Definition of Done

A rollout reaches a true **Definition of Done** only when the control behaves as a stable part of normal operations. This means the report only signal has settled across multiple monitoring cycles with no unexpected surges or unexplained gaps. Exceptions trend downward because owners close them rather than renew them. User visible errors fall to a level that does not disrupt daily work. Most routine remediations are handled automatically through the platform rather than through manual intervention. When these conditions are met, the control has crossed from project work into operational reality. It no longer depends on extraordinary effort or constant surveillance, and it begins to behave as part of the enterprise trust fabric.

Governance and Communication

Governance and proper communication remain the spine that holds each rollout together. Every change benefits from a simple explanation that tells people what is shifting, why it matters, and where support can be found. Clear communication reduces friction and lowers the noise that often surrounds security transitions. Exception management follows the same principle. Each exception is tracked with an owner and an expiration date so that drift does not become normalized. Expiration dates are treated as real work items rather than reminders or administrative cleanup. This approach keeps the organization focused on closure instead of accumulation.

Telemetry

Telemetry creates both the confidence and the restraint needed for safe enforcement. A consistent view of policy behavior, enforcement latency, and automated coverage helps teams understand where the system is stable and where it is brittle. Drift detection and restoration time are observed closely because they signal whether the environment is ready to absorb change. These measurements form a continuous feedback loop that keeps all three rollouts grounded in evidence rather than assumptions.

Safety Valves

Fallback is always an option. Safety valves protect the organization when enforcement moves from observation to real impact. Issues happen. Misconfigurations occur even in the strictest Change Management environments, which is why every control requires a clearly defined rollback plan. A controlled break glass process provides privileged access under strict time limits. Canary groups and ring style deployments place early enforcement on a small and representative cohort so unexpected issues can surface before they reach the wider population. **These mechanisms create room for correction without exposing the enterprise to unnecessary risk.**

Operational Realism

Operational realism matters throughout these programs. Many organizations begin without unified telemetry, consistent asset attribution, or fully reliable inventories for identities, devices, or data.

These dependencies shape the sequence of implementation. Monitoring must come before enforcement. Verification of authoritative sources must come before automation.

Once signals are stable and coverage is confirmed, enforcement becomes both safe and predictable.

AI Enablement Patterns – Agentic Systems

While this paper centers on implementing and operationalizing the minimum viable Zero Trust baseline, future research will expand these foundations into applied reference designs. Those forthcoming models will illustrate how foundational controls manifest within AI system architectures and data pipelines, translating policy intent into repeatable engineering blueprints.

The rise of **agentic systems** changes the nature of enterprise AI adoption. Basic model invocation is simple to govern because inputs and outputs move in predictable channels. Agents move differently. They chain decisions, call tools, execute code, and route work across identity boundaries and data tiers. They behave more like distributed decision engines than applications.

Zero Trust becomes the only reliable way to maintain clear boundaries when the execution path is dynamic. Identity must follow the agent through every tool and action. Segmentation must apply inside the agent runtime, not only at the perimeter. Provenance must capture the full sequence of decisions.

Frameworks such as MCP illustrate this shift by formalizing tool access and delegation. Without a Zero Trust governed runtime, these capabilities become **untraceable execution surfaces**. With a Zero Trust governed runtime, agent autonomy becomes **bounded, auditable, and aligned** with enterprise controls.

Different agent frameworks use different runtime patterns, but the underlying security requirement remains the same because once actions are delegated across boundaries, Zero Trust verification becomes the governing force.

For example, these designs will formalize recurring structures such as: **segmented model lifecycles**, where training, testing, and deployment occur in discrete trust zones with governed data

preparation and restricted compute access; **continuous model attestation**, in which models are signed at build time and verified at release through integrated MLOps approval gates; **telemetry-linked access**, where inference APIs validate short-lived tokens and correlate all calls for audit; **policy-enforced data lineage**, ensuring every dataset and derivative carries origin, transformation, and sensitivity tags enforced by retrieval brokers and vector store controls; and **shared-responsibility maps** that define explicit operational duties across architecture, security engineering, platform, and data-science functions.

These constructs will serve as the next layer of architectural guidance, building on the control-level foundations established here.

Agentic Use Case: Customer Support Triage Agent

Customer support triage is one of the clearest examples of where agentic systems generate immediate and measurable return on investment. The workflows are repetitive, high volume, and tied to structured knowledge and historical case data. An agent can read an entire ticket, ingest the customer's account history, search documentation, query internal systems, and produce a draft resolution that a human can review in seconds instead of minutes. Even partial automation reduces handling time, improves consistency, and lowers operational cost at scale. This use case also exposes the architectural impact of agentic systems. The agent is no longer generating text. It is invoking tools, reading sensitive records, crossing identity boundaries, and taking actions inside critical systems. The promise of efficiency sits beside the need for identity lineage, tool segmentation, and runtime provenance. This example captures both the economic upside and the structural risks of agentic AI in a way that every enterprise immediately understands.

When enterprises begin testing agentic systems in customer support workflows, the default implementation pattern tends to collapse identity at the point of execution. A human agent authenticates to the support console, presses an assist button, and the request is handed to the model through a shared service account. The agent runtime then calls internal tools such as CRM systems, order platforms, knowledge bases, and account records using the same shared identity. In the logs, the workflow becomes a sequence of service calls with no link back to the human who initiated them. There is no lineage, no ability to reconstruct which decision triggered which action, and no boundary that prevents privilege creep. The appearance of automation hides the fact that this pattern breaks the basic assumptions of accountability the moment the agent touches a sensitive system.

In a Zero Trust aligned model, the AAA chain looks **different**. The support agent still initiates the workflow, but their identity is carried into the agent session as a signed and verifiable context attribute. The agent has a distinct workload identity. Each tool the agent may invoke has its own identity with narrow permissions. When the agent calls a tool, the call includes the agent identity and a constrained representation of the human's authority. The tool verifies that the agent is authorized to call it and that the human is permitted to perform the requested action. Every step becomes a real authorization event rather than an internal assumption. This produces full provenance of the execution path. A reviewer can trace which human initiated the session, which agent identity acted, which tools were invoked, what resources were touched, and what each step attempted to do.

This is what Zero Trust means in an agentic environment. The agent becomes **an identity-bearing workload** rather than a feature **hidden inside an application**. Actions are evaluated. Tool boundaries remain real. The organization keeps the ability to audit and govern autonomous workflows as they span multiple systems. Without this structure, agentic automation becomes a tangle of unaccountable service accounts and opaque decision chains. With it, enterprises can adopt high value agentic use cases without abandoning the operational foundations that keep the environment safe.

Later work will expand these foundations into detailed patterns for agent execution, platform governance, and runtime assurance.

Governance at Speed: Turning Security into Strategic Acceleration

Governance slows delivery when it operates outside daily workflows. In a Zero Trust program, governance is codified as policy, enforced by pipelines and platforms, and evidenced by telemetry.

- **Policy as code.** Infrastructure and application templates embed access, segmentation, and logging defaults. Noncompliant resources are blocked at creation or auto-remediated.
- **Runtime assurance.** Policy decision points evaluate identity, device, data class, and anomaly signals. Enforcement occurs in-line. Exceptions are explicit, time-bound, and logged.
- **Product decisions.** Shadow AI attempts indicate unmet needs. By analyzing blocked actions and exceptions, teams can prioritize sanctioned capabilities with correct controls.
- **Process improvement.** Repeated denials or drift in a data pipeline may point to classification gaps or brittle sharing models. Adjusting tags and scopes reduces friction without increasing exposure.
- **Partner trust.** Demonstrable lineage, access logs, and model integrity checks enable new data collaborations and customer assurances. Security becomes a feature of the service.

Compliance establishes minimum thresholds. Zero Trust architectures generate assurance data that can inform product decisions, process improvement, and partner trust.

In this mode, the organization treats security telemetry as a design input. It reallocates effort from manual oversight to system improvements that raise both assurance and productivity. That is the force multiplier effect in practice.

Conclusion: Safe Velocity

The argument in this paper **does not depend on Zero Trust being the only valid architectural strategy**. It depends on Zero Trust being a practical and repeatable way to create the conditions under which AI adoption becomes safe, observable, and scalable. Organizations that achieve these conditions through other models can reach similar readiness.

Zero Trust, applied as enterprise architecture, enables AI with control and transparency. It converts verification into a default behavior, makes identity and data stewardship explicit, and turns governance into an inline function. The result is **safe velocity**, defined as a controlled innovation rate under stable or improving risk metrics.

Safe velocity is observable. Indicators include reduced mean time to deploy AI services, stable or improving incident frequency per release, decreasing manual exception counts, and shrinking policy drift restoration times. These metrics translate Zero Trust from architectural philosophy into measurable operational performance.

For example, it can be expressed through the following indicators:

- **Deployment Frequency:** the number of AI service releases or policy changes per quarter.
- **Incident Rate per Release:** security or compliance incidents per deployment; must remain flat or trend downward.
- **Policy Drift MTTR:** median time to detect and correct configuration drift; should decrease over time.
- **Exception Closure Rate:** percentage of exceptions closed on schedule.

When deployment frequency rises and the other metrics remain stable or improve across consecutive release cycles, the organization is operating within a healthy pattern of safe velocity. The exact thresholds vary by environment, but the pattern remains consistent.

Enterprises that adopt this model realize two compounding benefits. First, AI delivery becomes faster because approvals are executed by policy and evidence is produced automatically. Second, risk is bounded because trust is recalculated at each step and drift is corrected quickly. Security architecture, long seen as overhead, becomes business infrastructure and a consistent business enablement mechanism.

That is the agenda for AI at scale.

Acknowledgments

The author thanks the open research and standards community, including contributors to the CISA Zero Trust Maturity Model, NIST Zero Trust publications, the Cloud Security Alliance AI Controls Matrix, and the SABSA community, whose public frameworks underpin this analysis.

Special thanks to **John Kindervag**, creator of the Zero Trust Model, for redefining how trust boundaries are understood in enterprise systems.

Deepest respect and gratitude to the late **John Sherwood**, Chief Architect of SABSA and co-author of *Enterprise Security Architecture: A Business-Driven Approach*, whose articulation of **security as a business enabler** and his enduring “**brakes on a car**” metaphor deeply inspired me more than twenty years ago.

His legacy remains foundational to every discussion linking security, architecture, and trust.

References

CISA. (2023). *Zero Trust Maturity Model (Version 2.0)*. Cybersecurity and Infrastructure Security Agency. <https://www.cisa.gov/zero-trust-maturity-model>

CISA. (2023). *Secure by Design*. Cybersecurity and Infrastructure Security Agency. <https://www.cisa.gov/securebydesign>

Cloud Security Alliance. (2025). *AI Controls Matrix (AICM v1)*. Cloud Security Alliance. <https://cloudsecurityalliance.org/artifacts/ai-controls-matrix>

ENISA. (2023). *Multilayer framework for good cybersecurity practices for AI*. European Union Agency for Cybersecurity. <https://www.enisa.europa.eu/publications/multilayer-framework-for-good-cybersecurity-practices-for-ai>

IBM Institute for Business Value. (2025). *IBM X-Force 2025 Threat Intelligence Index*. IBM. <https://www.ibm.com/reports/threat-intelligence>

ISO/IEC. (2022). *ISO/IEC 27001:2022 Information security, cybersecurity and privacy protection - Information security management systems - Requirements*. International Organization for Standardization. <https://www.iso.org/standard/27001>

Kindervag, J. (2010). *No More Chewy Centers: Introducing the Zero Trust Model of Information Security*. Forrester Research.

Microsoft. (2024). *Security for AI: A practical roadmap for CISOs*. Microsoft Security Blog. <https://www.microsoft.com/en-us/security/blog/2024/11/13/security-for-ai-a-practical-roadmap-for-cisos>

Microsoft. (2025). *How the Secure Future Initiative brings Zero Trust to life*. Microsoft Security Blog. <https://www.microsoft.com/en-us/security/blog/2025/05/15/how-the-microsoft-secure-future-initiative-brings-zero-trust-to-life>

NIST. (2020). *SP 800-207: Zero Trust Architecture*. National Institute of Standards and Technology. <https://doi.org/10.6028/NIST.SP.800-207>

NIST. (2023). *SP 800-207A: A Zero Trust Architecture Model for Access Control in Cloud-Native Applications in Multi-Location Environments*. National Institute of Standards and Technology. <https://csrc.nist.gov/pubs/sp/800/207/a/final>

NIST. (2024). *The NIST Cybersecurity Framework (CSF) 2.0*. National Institute of Standards and Technology. <https://doi.org/10.6028/NIST.CSWP.29>

Sherwood, J., Clark, A., & Lynas, D. (2005). *Enterprise Security Architecture: A Business-Driven Approach*. CRC Press.

The Center for Internet Security. (2024). *CIS Critical Security Controls v8.1*. Center for Internet Security. <https://www.cisecurity.org/controls/v8-1>

Verizon. (2024). *2024 Data Breach Investigations Report*. Verizon. <https://www.verizon.com/business/resources/reports/2024-dbir-data-breach-investigations-report.pdf>