Data- und AI-Governance-Rahmenwerk für Zero Trust Architektur (ZTA) mit Künstlicher Intelligenz (KI) in der Industrie 4.0

Version, Stand und Genehmigung

Dieses Dokument stellt die Version 2.0 des Data- und AI-Governance-Rahmenwerks dar, das am xx. xxx 2026 erstellt wurde. Es wurde von der Geschäftsführung, dem Chief Information Security Officer (CISO) und dem Datenschutzbeauftragten genehmigt, um eine maximale Auditierbarkeit zu gewährleisten. Die Genehmigung erfolgte durch Unterschriften, die in der gedruckten Version dieses Dokuments vorliegen, und dient als Nachweis für die formelle Annahme aller hierin festgelegten Regelungen.

Rahmenbedingungen und Geltungsbereich

Dieses Governance-Rahmenwerk gilt für alle Systeme der Künstlichen Intelligenz und Komponenten der Zero Trust Architektur, die in der Organisation entwickelt, beschafft, betrieben oder eingesetzt werden. Es umfasst insbesondere Anwendungen in operativen Technologien und Umgebungen der Industrie 4.0, wie etwa in der Produktion, der Lieferkette, der prädiktiven Wartung und der Qualitätskontrolle mit Unterstützung durch Künstliche Intelligenz. Die rechtliche und normative Grundlage dieses Rahmenwerks basiert auf der EU AI Act in der Fassung der Verordnung (EU) 2024/1689, insbesondere den Artikeln 9 bis 15 für Systeme mit hohem Risiko gemäß Annex III, die Use-Cases in kritischer Infrastruktur, Beschäftigung und Produktqualitätssicherung betreffen. Darüber hinaus orientiert es sich an der ISO/IEC 42001:2023 für das Artificial Intelligence Management System, an der ISO/IEC 27001:2022 für das Information Security Management System sowie an der IEC 62443 für die Sicherheit in operativen Technologien. Es integriert die Prinzipien der NIST SP 800-207 für Zero Trust Architecture und der DoD Zero Trust Reference Architecture. Der Geltungsbereich ist auf Systeme mit Relevanz für Künstliche Intelligenz oder Zero Trust Architektur beschränkt, wobei Systeme mit niedrigem Risiko vereinfachte Regelungen unterliegen, die in separaten Anhängen dieses Dokuments detailliert beschrieben werden.

Ziele und Grundsätze

Die primären Ziele dieses Rahmenwerks bestehen darin, Risiken für die Sicherheit, die Gesundheit, die Grundrechte und die Resilienz in operativen Technologien zu minimieren. Es soll die Nachvollziehbarkeit, die Auditierbarkeit und die Resilienz der Systeme gewährleisten, indem alle regulatorischen Anforderungen der EU AI Act und der ISO 42001 vollständig eingehalten werden. Darüber hinaus fördert es die verantwortungsvolle Nutzung von Künstlicher Intelligenz in dynamischen Umgebungen der Zero Trust Architektur. Die Grundsätze orientieren sich am Alignment mit den Prinzipien der NIST Zero Trust Architecture und der ISO 42001 Annex A. Dazu gehört das Prinzip, niemals blind zu vertrauen, sondern immer zu verifizieren, was durch dynamische und kontextbasierte Authentifizierung umgesetzt wird. Es wird ein Angriff vorausgesetzt, was eine kontinuierliche Überwachung und die Vergabe minimaler Privilegien erfordert. Die Provenienz von Daten und Entscheidungen muss über den gesamten Lebenszyklus hinweg gesichert sein. Eine menschliche Aufsicht ist bei Entscheidungen mit hohem Risiko obligatorisch. Der kontinuierliche Verbesserungsprozess folgt dem Plan-Do-Check-Act-Zyklus, der in allen Prozessen dieses Rahmenwerks verankert ist.

Rahmenbedingungen und Geltungsbereich

Dieser generische Anforderungskatalog gilt für alle Komponenten einer Zero Trust Architektur, die mit KI-Assistenz in industriellen Umgebungen der Industrie 4.0 eingesetzt werden. Er umfasst insbesondere Anwendungen in der Produktion, der Lieferkette, der prädiktiven Wartung, der Qualitätskontrolle sowie in sicherheitskritischen OT-Prozessen. Die normative Grundlage dieses Katalogs basiert auf der NIST SP 800-207 (Zero Trust Architecture), der EU AI Act (Verordnung (EU) 2024/1689, insbesondere Artikel 9 bis 15 für High-Risk-Systeme), der ISO/IEC 42001:2023 (Annex A Controls), der IEC 62443-3-3 (System Security Requirements) sowie der ISO/IEC 27001:2022. Der Katalog adressiert die identifizierten Mängel der vorherigen Version, indem er den Umfang erheblich erweitert, eine vollständige Mapping-Tabellen einführt, eine Risikobewertung integriert, Lifecycle-Aspekte abdeckt und Evidenz-Anforderungen explizit definiert.

Klassifizierung und Risikobewertung

Jede Anforderung wird einer Risikoklasse zugeordnet (basierend auf EU AI Act und IEC 62443 Security Levels SL 1–4). Die folgende Tabelle gibt einen Überblick über die Zuordnung.

Risikoklasse / Security Level	Beschreibung der Klasse	Typische Anwendung in Industrie 4.0	Priorität der Umsetzung
SL 1 / Minimal Risk	Schutz gegen unbeabsichtigte oder zufällige Bedrohungen	Reine Monitoring-Systeme ohne Steuerung	Mittel
SL 2 / Moderate Risk	Schutz gegen beabsichtigte Angriffe mit einfachen Mitteln	Standard-Qualitätskontrolle mit KI	Hoch
SL 3 / High Risk	Schutz gegen gezielte Angriffe mit erheblichen Ressourcen	Prädiktive Wartung mit Sicherheitsrelevanz	Sehr hoch
SL 4 / Very High Risk	Schutz gegen staatlich unterstützte Angriffe	Kritische Infrastruktur-Komponenten	Kritisch

Hinweise zur Vermeidung von Überklassifizierung

Im Kontext des EU AI Act (Verordnung (EU) 2024/1689) besteht die Pflicht zur Self-Assessment von AI-Systemen, ob sie in die High-Risk-Kategorie fallen (Annex III). Allerdings hat sich durch die hohen Strafen (bis zu 7% des globalen Jahresumsatzes für Verstöße) eine Tendenz zu "Over-Compliance" entwickelt: Unternehmen klassifizieren Systeme vorsichtshalber als High-Risk, um Risiken zu vermeiden. Studien zeigen, dass bis zu 40% der Klassifizierungen unklar sind, was zu Überklassifizierungen in 18-58% der Fälle führt (z. B. appliedAI-Institute Report 2023/2024). Dies bindet unnötig Ressourcen, behindert Innovation und verwässert den Fokus auf echte High-Risk-Systeme, wie die EU in Guidelines kritisiert (z. B. Pressemitteilung EC IP/25/2718: "Vermeidung unnötiger Belastungen durch korrekte Klassifizierung").

Im Gegensatz dazu betont IEC 62443 eine kosten-nutzen-orientierte Klassifizierung (SL 1-4), die Überklassifizierung vermeidet, da höhere Levels explizite Kostensteigerungen bedeuten. Empfehlung: Führen Sie eine dokumentierte Assessment durch, nutzen Sie EU-Beispiele für Non-High-Risk (Annex III-Ausnahmen: enge prozedurale Tasks, Unterstützung menschlicher Entscheidungen). Bei Unsicherheiten: Konsultieren Sie den EU AI Office oder externe Experten, um echte Risiken zu priorisieren und "Panik-Klassifizierungen" zu verhindern.

Generischer Anforderungskatalog

Der folgende Katalog ist thematisch gruppiert und enthält für jede Anforderung eine eindeutige ID, eine vollständige Beschreibung, die Konformitätsprüfung, den Bewertungsstatus, Referenzen zu den Kernstandards sowie die geforderte Evidenz. Alle Felder sind in vollständigen Sätzen formuliert, um juristische Lesbarkeit und Auditierbarkeit zu gewährleisten.

Anforderungs-ID	Beschreibung der Anforderung	Konformitätsprüfung / Nachweismethode	Bewertungsstatus	Referenzen zu Standards	Geforderte Evidenz
ZTA-01	Dynamische Richtlinien-Durchsetzung muss in Echtzeit erfolgen, wobei jede Zugriffsentscheidung anhand aktueller Kontextdaten (Identität, Gerät, Verhalten, OT-Prozesszustand) getroffen wird.	Die Policy Engine muss alle Zugriffsanfragen in Echtzeit evaluieren und protokollieren; Testszenarien mit simulierten Kontextänderungen müssen durchgeführt werden.	Offen	NIST SP 800-207 Tenet 3 & 6, IEC 62443 SR 2.1, EU AI Act Art. 15	Audit-Logs der Policy Engine, Testprotokolle, Konfigurationsscreenshots

ZTA-02	Identitäts- und Zugriffsmanagement muss kontinuierlich verifizieren, dass nur stark authentifizierte und autorisierte Entitäten Zugriff erhalten, inklusive Geräte- und Service-Identitäten.	Multi-Factor-Authentifizierung und Device-Posture-Checks müssen für alle Zugriffe implementiert sein; Least-Privilege-Policies müssen durchgesetzt werden.	Offen	NIST SP 800-207 Tenet 1 & 4, IEC 62443 SR 1.1–1.7, ISO 42001 A.3.2	IAM-Konfiguration, Auth-Logs, Least-Privilege-Matrix
ZTA-03	Datenprovenienz und Integrität müssen über den gesamten AI- und ZTA-Lifecycle nachweisbar sein, inklusive Herkunft, Veränderung und Verwendung der Daten.	Alle Daten müssen mit kryptografischen Signaturen versehen und in unveränderbaren Logs gespeichert werden; Bias- und Qualitätschecks müssen dokumentiert sein.	Offen	EU AI Act Art. 10, ISO 42001 A.7.1–A.7.5, IEC 62443 SR 3.1–3.9	Provenienz-Chain-Dokumentation, Hash- und Signatur-Reports
ZTA-04	Vollständige Auditierbarkeit und Nachvollziehbarkeit aller Entscheidungen (ZTA-Policy & KI-Inferenz) muss gewährleistet sein, inklusive Explainability für KI-Entscheidungen.	Audit-Trails müssen zeitgestempelt, unveränderbar und suchbar sein; Model Cards und Explainability-Methoden müssen für alle KI-Modelle vorliegen.	Offen	EU AI Act Art. 13, ISO 42001 A.8.1–A.8.5, NIST SP 800-207 Tenet 7	Vollständige Audit-Logs, Model Cards, Explainability-Reports
ZTA-05	Menschliche Aufsicht muss bei High-Risk-Entscheidungen obligatorisch implementiert sein, mit klar definierten Eskalationsregeln und Triggers.	Human-in-the-Loop / on-the-Loop-Mechanismen müssen für kritische Entscheidungen vorhanden sein; Eskalation bei Confidence < 85 % oder Drift > 10 %.	Offen	EU AI Act Art. 14, ISO 42001 A.9.1–A.9.3	Oversight-Protokolle, Eskalationsregeln-Dokument, Testfälle
ZTA-06	Micro-Segmentation und Restricted	Netzwerke müssen in Zonen und Conduits	Offen	IEC 62443 SR 5.1–5.7, NIST SP	Netzwerkdiagramm (Zone/Conduit),

	Data Flow müssen implementiert sein, um Lateral Movement in OT- und IT-Netzwerken zu verhindern.	segmentiert sein; Datenflüsse dürfen nur explizit erlaubte Pfade nutzen.		800-207 Tenet 5	Firewall- und Segmentation-Rules
ZTA-07	Kontinuierliches Monitoring und Incident Response müssen für Anomalien, Drift und Sicherheitsereignisse eingerichtet sein, inklusive automatisierter Alerts.	SIEM-Integration mit KI-basierter Anomalie-Erkennung muss vorhanden sein; Incident-Response-Plan muss getestet werden.	Offen	NIST CSF Detect/Respond, ISO 42001 A.10.1–A.10.3, IEC 62443 SR 6.1	Monitoring-Dashboards, Incident-Reports, Testprotokolle
ZTA-08	Lifecycle-Management für AI- und ZTA-Komponenten muss alle Phasen (Design, Entwicklung, Deployment, Monitoring, Decommissioning) abdecken.	Ein vollständiger AI Lifecycle Prozess muss dokumentiert und mit Risiko-Assessments verknüpft sein.	Offen	ISO 42001 A.6.1–A.6.2.8, EU AI Act Art. 9	Lifecycle-Diagramm, Phasen-Dokumentation, Decommissioning-Plan
ZTA-09	Alle Kommunikationen müssen unabhängig von der Netzwerkposition vollständig gesichert werden, einschließlich Verschlüsselung in Transit und End-to-End-Sicherung für OT-Datenströme und KI-Inferenz-Daten.	Alle Verbindungen müssen TLS 1.3 oder höher nutzen; OT-spezifische Protokolle müssen durch sichere Gateways oder Wrappers geschützt werden; Penetrationstests auf unverschlüsselte Kommunikation müssen regelmäßig durchgeführt werden.	Offen	NIST SP 800-207 Tenet 2 (All communication secured regardless of location), IEC 62443 SR 3.1–3.9 (System Integrity), EU AI Act Art. 15 (Cybersecurity)	TLS-Konfigurationsberichte, OT-Protokoll-Analyse, Penetrationstest-Reports, Verschlüsselungs-Matrix
ZTA-10	Zugriffe müssen sitzungsbezogen und mit Just-in-Time / Just-Enough-Access	Policy Engine muss session-basierte Tokens mit kurzer Lebensdauer	Offen	NIST SP 800-207 Tenet 3 (Per-session access),	Session-Log-Analyse, Token-Lebensdauer-Konfiguration,

	gewährt werden, wobei Zugriffe automatisch bei Session-Ende oder Kontextänderung widerrufen werden.	ausstellen; automatische Revocation bei Anomalien oder Zeitüberschreitung muss implementiert sein.		NIST Tenet 4 (Least privilege), IEC 62443 SR 2.1 (Least Privilege)	Revocation-Testprotokolle
ZTA-11	Kontinuierliche Bewertung der Sicherheitslage (Continuous Posture Assessment) muss für alle Entitäten (User, Device, AI-Modell, OT-Gerät) durchgeführt werden, inklusive Device Health Checks und Behavioral Analytics.	Echtzeit-Monitoring von Device-Compliance, User-Verhalten und Modell-Drift; KI-gestützte Anomalie-Erkennung muss integriert sein.	Offen	NIST SP 800-207 Tenet 6 (Continuous verification), ISO 42001 A.10 (Continuous Improvement), EU AI Act Art. 15 (Robustness)	Device-Posture-Reports, UEBA-Dashboards, Drift-Detection-Logs
ZTA-12	Automatisierung und Orchestrierung von Security-Maßnahmen muss implementiert sein, um Policies dynamisch anzupassen, Incidents zu isolieren und Response-Prozesse zu automatisieren.	SOAR-Integration mit ZTA-Komponenten; automatisierte Quarantäne bei erkannten Bedrohungen; KI-gestützte Policy-Optimierung muss vorhanden sein.	Offen	NIST SP 800-207 (Automation in ZTA), IEC 62443 SR 6.1 (Timely Response), Cloud Security Alliance ZTA AI-Integration	Automation-Workflow-Diagramme, SOAR-Konfiguration, Incident-Automatisierungs-Tests
ZTA-13	Schutz der Daten als Kernressource muss durch Klassifizierung, Verschlüsselung at-rest, Tokenisierung und Data Loss Prevention gewährleistet sein, insbesondere für Trainingsdaten und Inferenz-	Datenklassifizierungs-Schema muss existieren; sensible OT- und KI-Daten müssen verschlüsselt gespeichert werden; DLP-Regeln müssen greifen.	Offen	NIST SP 800-207 Tenet 1 (All data sources as resources), EU AI Act Art. 10 (Data Quality & Governance), IEC 62443 SR 3.9 (Data Confidentiality)	Datenklassifizierungs-Matrix, Encryption-at-Rest-Reports, DLP-Alert-Logs

	Outputs in KI-Systemen.				
ZTA-14	Sichtbarkeit und Analytics müssen umfassend implementiert sein, um alle Zugriffe, Anomalien und Kontextdaten zu sammeln, zu analysieren und für Threat Hunting sowie Compliance-Reporting zu nutzen.	Zentrale SIEM- oder Analytics-Plattform mit KI-Unterstützung; vollständige Log-Sammlung aus allen ZTA-Komponenten und OT-Systemen.	Offen	NIST SP 800-207 Tenet 7 (Collect & Analyze Data), ISO 42001 A.8 (Transparency), IEC 62443 SR 6.1 (Monitoring)	SIEM-Dashboard-Screenshots, Log-Retention-Policy, Analytics-Reports
ZTA-15	Resilienz gegenüber Ausfällen und Angriffen muss durch Redundanz, Failover-Mechanismen und Backup/Restore-Prozesse für kritische ZTA- und KI-Komponenten gewährleistet sein, ohne OT-Verfügbarkeit zu gefährden.	Hochverfügbarkeits-Architektur für Policy Engine und KI-Modelle; regelmäßige Disaster-Recovery-Tests; OT-spezifische Non-Disruptive-Recovery.	Offen	IEC 62443 SR 7.1–7.8 (Resource Availability), NIST CSF Recover, EU AI Act Art. 15 (Robustness & Accuracy)	HA-Konfigurationsdiagramme, DR-Testprotokolle, Failover-Simulation-Results

Mapping zu Kernstandards

Die Anforderungen ZTA-09 bis ZTA-15 sind vollständig auf die folgenden Standards gemappt. Die detaillierte Statement-of-Applicability-Tabelle (inkl. aller 38 ISO 42001 Controls und IEC 62443 SRs) befindet sich weiterhin im Anhang.

- NIST SP 800-207: Tenet 2 (Secured Communication), Tenet 3 (Per-Session), Tenet 6 (Continuous), Tenet 7 (Data Collection), Automation-Orchestrierung
- EU AI Act: Art. 10 (Data & Quality), Art. 13 (Transparency), Art. 14 (Oversight), Art. 15 (Accuracy, Robustness, Cybersecurity)
- ISO 42001 Annex A: A.7 (Data), A.8 (Transparency/Explainability), A.10 (Incident & Improvement)
- IEC 62443-3-3: SR 3 (Integrity), SR 5 (Restricted Data Flow), SR 6 (Timely Response & Monitoring), SR 7 (Resource Availability)

Anhang: Vollständiges Statement of Applicability und Changelog

Im Anhang dieses Dokuments befindet sich die vollständige Mapping-Tabelle zu allen relevanten Controls (38 aus ISO 42001, SRs aus IEC 62443-3-3). Der Changelog dokumentiert: In Version 3.0 (Erweiterung) wurden ZTA-09 bis ZTA-15 hinzugefügt, um weitere NIST-Tenets, OT-Resilienz-Anforderungen (IEC 62443), AI-

spezifische Robustheit (EU AI Act) und Automatisierungs-Aspekte abzudecken. Zusätzlich wurde der Abschnitt "Klassifizierung und Risikobewertung" um Hinweise zur Überklassifizierung erweitert.

Unterschrift / Genehmigung