MRIS for TISAX® 2027 Implementation Guide

Mythos-resistente Informationssicherheit — Organisatorische und technische Umsetzung, Reifegrad und Nachweisführung.

Dieses Dokument ist der Umsetzungs-Layer zu Teil I.

Version 2.0 | September 2026

Autor: Richard Peddi (mit AI-Unterstützung)

ZIELGRUPPE
CISOs und ISMS-Verantwortliche bei TISAX®-pflichtigen Organisationen.

Lizenz und Hinweise

© 2026 Richard Peddi
Dieses Werk ist lizenziert unter der Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0).

Sie dürfen dieses Werk:

- teilen — das Material in jedwedem Format oder Medium vervielfältigen und weiterverbreiten
- bearbeiten — das Material remixen, verändern und darauf aufbauen
- das Material für nicht-kommerzielle Zwecke nutzen

Unter folgenden Bedingungen:
Namensnennung — Sie müssen angemessene Urheber- und Rechteangaben machen, einen Link zur Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden. Nicht kommerziell — Sie dürfen das Material nicht für kommerzielle Zwecke nutzen.
Lizenztext: https://creativecommons.org/licenses/by-nc/4.0/

Unabhängigkeitshinweis

TISAX® ist eine eingetragene Marke der ENX Association; die ENX Association administriert TISAX® im Auftrag des Verbands der Automobilindustrie (VDA), der den zugrunde liegenden Fragenkatalog (VDA ISA) entwickelt und veröffentlicht. Diese Analyse ist eine unabhängige, privat erstellte Arbeit ohne Verbindung zu ENX Association oder VDA. Sie ist weder Bestandteil des offiziellen TISAX®-Prüfprozesses noch von ENX autorisiert, geprüft oder bestätigt, und ersetzt keine TISAX®-Reifegradbewertung durch einen akkreditierten Prüfdienstleister.

Haftungsausschluss

Die Einstufungen in diesem Dokument sind eine fachliche Einschätzung zum Stand Juli 2026, keine Rechts- oder Zertifizierungsauskunft. Die Anwendung erfolgt ohne Gewähr und in eigener Verantwortung.

Zitierempfehlung

Peddi, Richard (2026): MRIS for TISAX® 2027 — Implementation Guide. v2.0

Änderungshistorie

Version	Datum	Beschreibung
1.0	Juli 2026	Erstveröffentlichung.
2.0	August 2026	Methodische Angleichung an MRIS 2.0: verbindliche Lesart der Reifegrad-Stufen (Initial belegt keine Erfüllung, ab Defined gilt ein MHC als umgesetzt), Behebungslatenz statt Zeit ab Bekanntwerden bei aktiv ausgenutzten Schwachstellen, Rollenbezug von der CISO- auf die Leitungsperspektive geöffnet, Schreibweise "AI" vereinheitlicht, AI-Unterstützung bei der Erstellung offengelegt.
1.1	Juli 2026	Überarbeitung auf Basis eines extern beauftragten Prüfberichts (Juli 2026, 9 Feststellungen F-01–F-09; vollständige Übersicht in Teil I).

Version 1.1 im Einzelnen:

Feststellung	Thema	Umsetzung
F-01	Inkonsistente Zählarithmetik und Kapitelsummen	Einzelbefund-Gesamtzahl 72 → 74 (Kaskade: Wirkungsbreite-Tabelle MHC-05 9 → 10, Einzelmaßnahmen 7 → 8, Ergänzende-Einzelmaßnahmen-Tabelle mit neuer 1.6.3-Zeile)
F-02, F-04	Zitier-/Editorial-Feststellungen	betreffen primär Teil I (siehe dort)
F-03	Asymmetrische Quellenlage	unbestätigte BSI-C5:2026-Detailwerte (MHC-09) von der bestätigten Rahmenaussage getrennt, konsistent mit Teil I
F-05	Fehlende Wirtschaftlichkeitsbetrachtung	neue Kennzahl "Kostenrahmen (Lizenz/Personentage, indikative Größenordnung ohne Einzelpreise)" in "Auf einen Blick" aller 10 MHC ergänzt (KMU/Midcap/Enterprise)
F-06	KMU-Eignung einzelner Zielwerte	neuer Abschnitt "MVP-Staffelung nach Schutzbedarf" im Reifegrad-Pfad von MHC-08, MHC-09, MHC-13, MHC-16 ergänzt
F-07	Fehlendes Evidence-Mapping zum Katalog	neuer Abschnitt "Evidence-Mapping zum ISA-Katalog" in Messung und Audit-Nachweis aller 10 MHC ergänzt
F-08, F-09	Anwendungshinweise	kein Dokumentmangel; keine Textänderung

Zusätzlich: Versionsnummer 1.0 → 1.1 (Deckblatt, Zitierempfehlung).

Inhaltsverzeichnis

Zweck und Konventionen

Dieses Dokument ist der Umsetzungs-Layer zu Teil I. Teil I diagnostiziert 74 einzelne Anforderungen über 31 Controls, gebündelt in acht wiederkehrenden Mustern (Phase 2) und einer Lückenanalyse gegen sechs externe Vergleichsrahmen (Phase 3); dieses Dokument liefert zu jedem der neun daraus synthetisierten MHC (Phase 4) die Umsetzungsleitlinie in vergleichbarer Tiefe — Zweck, organisatorische und technische Umsetzung, Reifegrad und Nachweisführung —, sodass eine CISO-Funktion sie ohne zusätzliche Spezialliteratur nachvollziehen kann. Es ersetzt Teil I nicht, sondern führt dessen Diagnose praktisch aus. Ein zehntes MHC (MHC-16) kommt hinzu, das nicht aus einer TISAX®-Anforderung abgeleitet ist, sondern eine originäre, über TISAX® hinausgehende Ergänzung darstellt (siehe Hinweis zur MHC-Nummerierung sowie Teil I).

Zeitbezug: Wie Teil I bildet auch dieser Implementation Guide den Erkenntnisstand Juli 2026 ab. Die MHC-Einträge sind gegen die zu diesem Zeitpunkt belegten Angreiferfähigkeiten kalibriert; neue AI-Fähigkeiten oder aktualisierte externe Rahmenwerke (NIST, BSI, DORA, CRA, NIS2, ISA/IEC 62443) können einzelne Reifegrad-Pfade oder Priorisierungen künftig verschieben.

Hinweis zur MHC-Nummerierung: Mythos-Härtungs-Controls (MHC) sind ein standardübergreifendes Maßnahmen-Vokabular — dieselbe Nummer bezeichnet dieselbe Härtungsmaßnahme, unabhängig vom zugrunde liegenden Standard. Sieben der zehn hier dokumentierten MHC (MHC-03, MHC-05, MHC-08, MHC-09, MHC-10, MHC-11, MHC-13) tragen bewusst dieselbe Nummer und denselben Grundtitel wie im MRIS-Implementation-Guide 2.0 for ISO 27002, weil die TISAX®-Diagnose aus Teil I dieselbe strukturelle Lücke aufdeckt. Die Einträge hier sind auf das kalibriert, was TISAX® ISA 2027 konkret diagnostiziert; wo der IG 2.0 einen technisch tieferen Ausbau desselben Controls beschreibt (z. B. Policy-as-Code bei MHC-10, Pipeline-Security-Testing bei MHC-09), wird darauf verwiesen. Bei MHC-13 gilt seit dieser Fassung eine Besonderheit: Der IG 2.0 dokumentiert unter derselben Nummer bereits eine deutlich breitere Maßnahme (u. a. Härtung gegen Prompt Injection, Harness-als-Code) als der ursprünglich hier kalibrierte, eng auf Control 1.3.3/1.3.4 zugeschnittene Kern; dieser Eintrag wurde daher erweitert, um dieselbe Nummer über beide Dokumente hinweg konsistent zu halten — mit Verweis auf den IG 2.0 für die dort zusätzlich vorhandene technische Tiefe (siehe MHC-13, Abschnitt 9). Drei MHC (MHC-14, MHC-15, MHC-16) sind neu vergeben: Für die zugrunde liegenden Themen — Lieferantennachweis-Verifikation, Identitätsverifikation bei Personal/Zutritt sowie Multi-Agenten-Integrität — fand sich unter den 13 MHC des IG 2.0 keine Entsprechung. MHC-14 und MHC-15 sind gemäß Teil I, Phase 3, originäre, bislang auch branchenweit ungelöste Lücken, die jeweils aus einem konkreten TISAX®-Anforderungsbefund abgeleitet sind. MHC-16 nimmt eine Sonderstellung ein: Es ist nicht aus einem TISAX®-Anforderungsbefund abgeleitet, sondern eine eigenständige, über die TISAX®-Diagnose hinausgehende Ergänzung — Details und Begründung siehe Teil I, Abschnitt "Originäre Ergänzungen jenseits von TISAX®", sowie MHC-16, Abschnitt 9. Die übrigen MHC des IG 2.0 (u. a. MHC-01 Post-Quantum-Strategie, MHC-02 SBOM/Build-Provenance, MHC-04 Workload-Identität, MHC-06 Container-Sicherheit, MHC-07 Multi-Tenancy-Isolation, MHC-12 Threat-Led Penetration Testing) decken Themen ab, zu denen die TISAX®-ISA-2027-Diagnose aus Teil I keinen eigenen degradierten Befund lieferte, und erscheinen daher hier nicht. Maßgeblich für die Zuordnung eines MHC bleibt stets die Dokumentzugehörigkeit.

Fixe Gliederung je MHC

Jeder MHC-Eintrag folgt derselben neunteiligen Struktur:

- Kopf »Auf einen Blick«

- 1. Härtungscontrol-Statement
- 2. Zweck und Bedrohungsbezug
- 3. Organisatorische Umsetzung (CISO)
- 4. Technische Umsetzung (schließt mit Wirksamkeitstest)
- 5. Umsetzungsbeispiele (Beispiel A/B)
- 6. Reifegrad-Pfad
- 7. Messung und Audit-Nachweis
- 8. Typische Fehler
- 9. Abgrenzung, Restrisiko und Verweise

Die Abschnitte "Messung und Audit-Nachweis" sind bewusst schlank gehalten — praktikable Nachweisführung statt übermäßiger Dokumentationslast. Organisationen mit erhöhtem Audit-, Regulatorik- oder Kundenanforderungsniveau können je MHC ein erweitertes Audit-Raster ergänzen:

- Control Owner
- Scope
- Applicability-Begründung
- Implementation Status
- Evidence Type
- Test Frequency
- Sampling Logic
- Exceptions
- Risk Acceptance
- KPI/KRI
- Last Effectiveness Review

Dieses Raster ist optional, nicht verpflichtender Bestandteil dieses Leitfadens.
Der Reifegrad-Pfad je MHC (Initial/Defined/Managed) ist ein eigenständiger, kumulativer Umsetzungsmaßstab und unabhängig vom TISAX®-Reifegradmodell (0–5, Zielwert 3) aus Teil I — beide sind miteinander kompatibel, messen aber Unterschiedliches: TISAX® bewertet den Prozess der Organisation, der MHC-Reifegrad-Pfad die Umsetzungstiefe der jeweiligen Zusatzmaßnahme.
Verbindliche Lesart der Stufen: Die Stufe Initial dokumentiert einen begonnenen Umsetzungspfad und belegt noch nicht die vollständige Erfüllung des jeweiligen Härtungscontrol-Statements. Erst ab Defined gilt ein MHC grundsätzlich als umgesetzt; Managed beschreibt die fortgeschrittene beziehungsweise kontinuierlich wirksame Umsetzung. Die Härtungscontrol-Statements beschreiben die Mindestanforderungen der Stufe Defined.

Priorisierung der MHC — Threat-Korrelation und Roadmap

Die zehn hier dokumentierten MHC folgen keiner internen Rangfolge, und ihre Nummern sind — anders als in einer in sich geschlossenen Liste — bewusst nicht fortlaufend (siehe Hinweis zur MHC-Nummerierung oben). Priorisiert wird nicht das Control an sich, sondern sein Beitrag zur Reduktion der für die Organisation relevantesten Bedrohungen. Dieses Kapitel verbindet zwei sich ergänzende Sichten: die messbare Wirkungsbreite jedes MHC (wie viele der 74 in Teil I identifizierten Einzelbefunde es trägt) und die bedrohungs- und aufwandsbezogene Reihenfolge (Threat-Korrelation).
Auf einen Blick

- Zweck: nachvollziehbare, bedrohungsorientierte Umsetzungs-Roadmap statt linearer MHC-Reihenfolge.
- Grundlage: Wirkungsbreite (Coverage) je MHC plus Threat Priority Score.
- Ergebnis: Standard-Tiering P0/P1/P2 als anpassbares Template, mit Abhängigkeitsregeln.

1. Wirkungsbreite: wie viele Einzelbefunde jedes MHC trägt

Die Zahlen sind gegen die Cluster-Übersicht in Teil I, Phase 2, geprüft (74 Einzelbefunde über 31 Controls, davon 1 Reine Reibung — der reine Passwortfaktor in Control 4.1.2; klassische MFA-Verfahren innerhalb desselben und anderer Controls zählen als Teilweise degradiert, siehe Teil I, Kapitel 4). Jeder Befund wird von genau einem MHC oder einer Einzelmaßnahme getragen; Mehrfachbetroffenheit entsteht auf Control-Ebene (ein Control kann Befunde aus mehreren Clustern enthalten), nicht auf Befund-Ebene. MHC-16 bildet eine bewusste Ausnahme: Es trägt keinen der 74 Einzelbefunde, weil es nicht aus der TISAX®-Diagnose abgeleitet ist, sondern eine originäre, über TISAX® hinausgehende Ergänzung darstellt (siehe Teil I, Abschnitt "Originäre Ergänzungen jenseits von TISAX®").

MHC	Einzelbefunde	Controls	Charakter
MHC-10 Continuous Control Monitoring und Policy-as-Code	18	12	breit, prozessübergreifend
MHC-03 Phishing-resistente Multi-Faktor-Authentisierung	10 (davon 1 Reine Reibung)	6	Identität/Zugang
MHC-05 Verhaltensbasierte Detection und Kill-Chain-Korrelation	10	5	Detection/Korrelation
MHC-09 AI-gestütztes Security-Testing in der Pipeline	6	3	Schwachstellen-/Patch-Management
MHC-15 AI-resistente Identitätsverifikation Personal/Zutritt	6	3	Personal/Facility
MHC-14 Unabhängige Lieferantennachweis-Verifikation	5	2	Lieferkette/Vertrag
MHC-11 SOAR-basierte Tier-1-Automation	4	1	Incident Response
MHC-13 AI-Agent-Governance und Harness-Sicherheit	4	2	AI-Agenten-Governance
MHC-08 Unveränderliche Backups und Recovery-Validierung	3	1	Backup/Recovery
Einzelmaßnahmen (6, siehe unten)	8	6	ohne wiederkehrendes Muster

MHC	Einzelbefunde	Controls	Charakter
MHC-16 Multi-Agent-Integrität	0 (originär)	0	AI-Agenten-Governance, zukunftsgerichtet

MHC-10 ist mit Abstand das breiteste Querschnitts-Control — das stärkste Argument gegen die Wahrnehmung, jedes MHC sei gleich wichtig.

2. Zwei Prioritäts-Sichten und ihr Zusammenhang

Wirkungsbreite (Coverage) ist die unbedingte Hebelwirkung — wie viele Einzelbefunde ein MHC trägt. Bedrohungs- und Aufwandsbezug ist die tatsächliche Reihenfolge — abhängig von Bedrohungslage, Auswirkung, Exposition, Control-Lücke, Abhängigkeiten und Aufwand. Beide Sichten ergänzen sich: Coverage allein würde MHC-13 nach hinten stellen, weil es nur zwei Controls flankiert — das wäre falsch, denn MHC-13 schließt die einzige Lücke, die auch keiner der sechs Vergleichsrahmen aus Phase 3 löst (siehe Priorisierung unten). "Low hanging fruits" sind Controls mit hoher Hebelwirkung, großer Control-Lücke und geringem Aufwand, unter Beachtung der Abhängigkeiten.

3. Bewertungslogik: Threat Priority Score

Threat Priority Score = Threat-Relevanz × Business Impact × Exposure × Control Gap

Faktor	Leitfrage	Skala
Threat-Relevanz	Wie realistisch, aktuell und relevant ist das Angriffsszenario für die Organisation?	1 = gering … 5 = sehr hoch
Business Impact	Wie schwer wären Folgen für TISAX®-Scope-Prozesse, Kundendaten, Lieferkette, Reputation?	1 = gering … 5 = existenziell
Exposure	Wie stark ist die Organisation exponiert (Internet-facing, Fernzugriff, privilegierte Konten, Lieferantennetz)?	1 = gering … 5 = stark exponiert
Control Gap	Wie groß ist die Lücke gegenüber dem MHC-Zielbild?	1 = gut abgedeckt … 5 = kaum abgedeckt

Beispiel: Echtzeit-Phishing-Relay (AITM) gegen einen privilegierten Fernzugriff mit Threat-Relevanz 5, Business Impact 4, Exposure 5 und Control Gap 4 ergibt einen Score von 400 — sehr hohe Priorität, vorrangig für MHC-03, flankiert durch MHC-11.
Der Threat Priority Score ist ein qualitativer Priorisierungsindex, kein Ersatz für eine formale Risikoanalyse. Die Multiplikation macht relative Handlungsprioritäten sichtbar; der Wert ist ordinal, nicht metrisch — ein Score von 400 ist nicht "doppelt so dringlich" wie 200, sondern dient der Reihung und Triage.

4. Threat-to-MHC-Matrix

Bedrohung/Angriffsszenario	Primäre MHC	Begründung	Priorität
Echtzeit-Phishing-Relay (AITM) gegen Authentisierung	MHC-03, MHC-11	MHC-03 macht abgefangene Faktoren wertlos, MHC-11 dämmt ein, falls	sehr hoch

Bedrohung/Angriffsszenario	Primäre MHC	Begründung	Priorität
		doch ein Konto kompromittiert wird.	
Patch-Gap-Kollaps / Ransomware nach Exploit	MHC-09, MHC-08, MHC-05, MHC-11	MHC-09 verkürzt das Zeitfenster bis zur Behebung, MHC-08 sichert die Wiederherstellbarkeit auch nach Kompromittierung, MHC-05 erkennt die Ausnutzung verhaltensbasiert, MHC-11 dämmt automatisiert ein.	sehr hoch
Deepfake-Identitätsvortäuschung (Einstellung, Zutritt, Anruf)	MHC-15	Adressiert direkt die AI-gestützte Dokumenten-/Identitätsfälschung im Personal- und Zutrittsprozess.	hoch
Gefälschte/beschönigte Lieferantennachweise	MHC-14	Erzwingt unabhängige Verifikation statt reiner Selbstauskunft.	mittel bis hoch
Autonomer AI-Agent bindet unautorisiertes Werkzeug/Paket ein	MHC-13	Einziges MHC gegen einen Bedrohungstyp, den auch keiner der sechs Vergleichsrahmen aus Phase 3 bislang regelt.	hoch, steigend
Prompt Injection kapert Agenten-Handlungsziel (Goal Hijack)	MHC-13	Laut OWASP die am weitesten verbreitete Bedrohungskategorie für AI-Agenten (ASI01); MHC-13 verlangt Kanaltrennung, Ein-/Ausgangsfilterung und einen Pen-Test vor Produktivsetzung.	hoch, steigend
Kompromittierter oder übernommener Agent nicht zeitnah isolierbar	MHC-13	Ohne Kill-Switch bindet der Entzug eines Agenten einen breiten Systemeingriff; MHC-13 setzt einen Zielwert von unter 5 Minuten.	hoch
Unbemerkt geändertes Verhalten einer bereits freigegebenen Agenten-Komponente (Rug Pull)	MHC-13	Einmalige Freigabe schützt nicht vor nachträglicher Änderung; MHC-13 verlangt erneute Prüfung bei jeder Aktualisierung.	mittel bis hoch
Fragmentierter, "low-and-slow" ausgeführter Angriff	MHC-05, MHC-10	MHC-05 korreliert über Zeit und Quellen, MHC-10 verhindert, dass Kontrolldrift zwischen periodischen Prüfzyklen unentdeckt bleibt.	hoch
Veraltete/nicht anlassbezogene Kontrollen allgemein	MHC-10	Größte Wirkungsbreite; betrifft praktisch jedes andere Cluster indirekt mit.	hoch
Mehrere kommunizierende Agenten: Memory/Context Poisoning, Kaskadeneffekte	MHC-16	Setzt eine Multi-Agenten-Architektur voraus; zunehmende Relevanz.	mittel, zukunftsgerichtet

5. Abhängigkeiten als Roadmap-Regeln

Abhängigkeit	Bedeutung
MHC-10 → MHC-13	Eine Allow-Liste für AI-Agenten-Werkzeuge verpufft, wenn sie nie anlassbezogen nachgeschärft wird, sobald neue Angriffsmuster bekannt werden.
MHC-03 → MHC-11	Phishing-resistente Authentisierung verhindert einen großen Teil der Vorfälle, die MHC-11 sonst automatisiert eindämmen müsste.
MHC-09 → MHC-11	Automatisierte Eindämmung ist nur so schnell wie die vorgelagerte Schwachstellenerkennung, auf die Alarme aufsetzen.
MHC-09, MHC-08 ergänzen sich	Beide adressieren Ransomware-Resilienz aus unterschiedlichen Richtungen — MHC-09 verkürzt das Zeitfenster bis zur Behebung, MHC-08 sichert die Wiederherstellbarkeit, falls die Ausnutzung dennoch gelingt; unabhängig voneinander umsetzbar.
MHC-14, MHC-15 unabhängig	Unterschiedliche Verantwortlichkeiten (Einkauf vs. HR/Werkschutz) — parallel umsetzbar, keine Sequenzierung nötig.
MHC-13 → MHC-16	Multi-Agenten-Integrität setzt eine funktionierende Einzel-Agenten-Absicherung (Identität, Allow-Liste, Kill-Switch) voraus, bevor die Kommunikation zwischen Agenten sinnvoll gehärtet werden kann.

6. Standard-Priorisierung P0 / P1 / P2

Das folgende Tiering ist ein Ausgangspunkt, kein universeller Umsetzungsplan — es ist je Organisation an Exposition, Architektur, Asset-Kritikalität und Ressourcenlage anzupassen.

Priorität	MHC	Begründung
P0 — sofort	MHC-03	Schließt den einzigen Reine-Reibung-Befund der gesamten Analyse.
P0 — sofort	MHC-08	Technisch klar umrissene, etablierte Maßnahme (Object Lock/WORM) mit geringem Umsetzungsaufwand und hoher Hebelwirkung gegen Ransomware; anders als bei MHC-09 formuliert C5:2026 hier keine verbindliche Unveränderlichkeits-Vorgabe — die Dringlichkeit ergibt sich aus der Bedrohungslage, nicht aus einem bereits vorhandenen externen Standard.
P0 — sofort	MHC-09	Patch-Gap-Schließung: C5:2026 verlangt bereits einen definierten, CVSS-gebundenen Fristenrahmen (OPS-25.01AC) — geringer Umsetzungsaufwand, hohe Hebelwirkung.
P0 — sofort	MHC-13	Größte eigenständige Lücke — kein externes Rahmenwerk liefert bereits eine Lösung, Dringlichkeit steigt mit jedem Tag ungeregelten AI-Agenten-Einsatzes.
P1 — danach	MHC-10	Größte Wirkungsbreite, aber höherer Change-Management-Aufwand über 12 Controls hinweg.
P1 — danach	MHC-14	Vertragsklausel-Ergänzung vergleichsweise einfach, wirkt aber erst mit dem nächsten Prüfzyklus voll.
P1 — danach	MHC-11	Baut auf MHC-09/MHC-03 auf; wirksamer, wenn diese bereits greifen.
P2 — risikobasiert	MHC-15	Hoher Aufwand (Prozessänderung HR/Werkschutz), originäre Neuentwicklung ohne externe Vorlage.
P2 — risikobasiert	MHC-05	Technisch anspruchsvollste Maßnahme (Produktauswahl EDR/XDR,

Priorität	MHC	Begründung
		Aufbau UEBA-Korrelation), setzt teils bereits vorhandene SIEM-Basis voraus.
P2 — risikobasiert	MHC-16	Originär und zukunftsgerichtet — relevant, sobald mehrere Agenten produktiv miteinander kommunizieren, setzt MHC-13 voraus.

7. Vorgehen in vier Schritten

- **Schritt 1 — Threat Landscape erstellen:** relevante Angreifertypen und Angriffsszenarien für die eigene TISAX®-Scope-Umgebung identifizieren; Lieferketten-, Fernzugriffs- und AI-Agenten-Exposition besonders berücksichtigen.
- **Schritt 2 — Threats mit MHC mappen:** je Bedrohung festlegen, welche MHC präventiv, detektiv, reaktiv oder validierend wirken; nicht anwendbare MHC (z. B. MHC-13 ohne AI-Agenten-Einsatz) sauber begründen und zurückstellen.
- **Schritt 3 — Score berechnen:** Threat-Relevanz, Business Impact, Exposure und Control Gap je 1 bis 5 bewerten; Score bilden, Aufwand und Abhängigkeiten berücksichtigen.
- **Schritt 4 — Roadmap ableiten:** P0 (sofort gegen hohe Bedrohungen und große Lücken), P1 (Controls mit Abhängigkeiten oder Basisarbeit), P2 (kontextabhängig); regelmäßig neu bewerten, sobald sich Bedrohungslage oder AI-Agenten-Einsatz ändern.

8. Formulierung für Governance-Unterlagen

Zur risikobasierten Priorisierung der TISAX®-Härtungscontrols werden die MHC mit relevanten Bedrohungsszenarien korreliert. Je Bedrohung wird bewertet, welche Controls präventiv, detektiv, reaktiv oder validierend wirken. Die Priorisierung ergibt sich aus Threat-Relevanz, Business Impact, Exposition, bestehender Control-Lücke, Abhängigkeiten und Umsetzungsaufwand. So entsteht keine lineare MHC-Reihenfolge, sondern eine bedrohungsorientierte Roadmap, die für jedes Control begründet, warum es zu seinem Zeitpunkt umgesetzt wird.

MHC-03 — Phishing-resistente Multi-Faktor-Authentisierung

Auf einen Blick

- Operativer Nutzen: Macht abgefangene oder in Echtzeit weitergereichte Anmeldedaten wertlos und entzieht damit automatisiertem, AI-skaliertem Phishing die Grundlage.
- Betroffene Richtlinie: Zugriffssteuerungs- und Authentisierungsrichtlinie.
- Abhängigkeiten: keine Vorbedingung; wirkt zusammen mit MHC-11.
- Aufwandstreiber: Anzahl Nutzer/Dienste/Altsysteme, vorhandene Identitätsplattform.
- Kostenrahmen (Lizenz/Personentage, indikative Größenordnung ohne Einzelpreise): KMU — Lizenz gering (FIDO2-Unterstützung oft bereits in vorhandener Identitätsplattform enthalten, wenige Hardware-Token zu beschaffen), ca. 5–15 Personentage · Midcap — Lizenz gering bis mittel (breitere Token-Beschaffung, Altsystem-Anbindung), ca. 15–40 Personentage · Enterprise — Lizenz mittel (große Token-Flotte, mehrere Identitätsdomänen), ca. 40–120 Personentage.
- Primäre Kennzahl: Anteil privilegierter und extern erreichbarer Zugänge mit phishing-resistenter Authentisierung.

- Berührte Standards (ohne Erfüllungsanspruch): NIS2 Art. 21(2)(j); NIST SP 800-63B (AAL3); TISAX® ISA 2027 4.1.1, 4.1.2, 4.1.3, 5.1.2, 5.2.7, 6.1.3.

1. Härtungscontrol-Statement

Für Authentisierungsverfahren, die TISAX® als "Stand der Technik" oder als zweiten Faktor fordert, wird mindestens für privilegierte Konten und Fernzugriff ein phishing-resistentes Verfahren (FIDO2/WebAuthn, PKI-basierte Zertifikate) verbindlich vorgeschrieben. Reine Passwort-Authentisierung ohne zusätzlichen Faktor ist für keine Schutzbedarfsstufe ausreichend. Zugangsberechtigungen werden bei Verlust oder Kompromittierung eines Ausweismittels automatisiert und unverzüglich gesperrt.

2. Zweck und Bedrohungsbezug

Automatisierte Echtzeit-Phishing-Relays (Adversary-in-the-Middle, AITM) hebeln OTP- und Push-basierte Verfahren aus, die vielerorts weiterhin als "Stand der Technik" gelten: der Angreifer leitet die Anmeldung in Echtzeit über eine gefälschte Seite weiter und erhält den gültigen Faktor direkt von der Nutzerin oder dem Nutzer. Bei reiner Passwort-Authentisierung entfällt sogar dieser Umweg — der Angreifer erhält das korrekte Passwort direkt, nicht durch Erraten (Teil I stuft dies als den einzigen Reine-Reibung-Befund der gesamten Analyse ein, Control 4.1.2). Phishing-resistente Verfahren binden den Anmeldenachweis kryptografisch an die echte Adresse des Dienstes; auf einer gefälschten Seite ist der Nachweis wertlos. Schutzziel ist, dass abgefangene oder erbeutete Anmeldedaten dem Angreifer nichts nützen. Diese kryptografische Bindung markiert den Unterschied zwischen "geschwächt" (klassische OTP-/Push-MFA — Teilweise degradiert in Teil I) und "hält" (phishing-resistente Verfahren — der Standfest-Zielzustand, den dieses MHC vorschreibt).

3. Organisatorische Umsetzung (Leitungsperspektive)

- **Richtlinie:** Die Authentisierungsrichtlinie schreibt phishing-resistente Verfahren für privilegierte, extern erreichbare und besonders schützenswerte Zugänge vor; reine Passwort-Authentisierung sowie SMS-/einfache Push-Verfahren werden für neue Zugänge ausgeschlossen.
- **Verantwortlichkeiten:** IAM-Team verantwortet Rollout und Altverfahren-Ablösung nach festem Übergangsplan.
- **Risikoanalyse:** Soll-Reifegrad wird aus dem Schutzbedarf abgeleitet — privilegierte und extern erreichbare Zugänge zuerst.
- **Prozesse:** geregelte Ausgabe und Rücknahme von Token/Passkeys; Sperrprozess bei Verlust ist automatisiert, nicht nur "so weit wie möglich" umgesetzt; Ersatzverfahren für Verlustfälle ohne Schutzniveau-Unterlaufung.

4. Technische Umsetzung (für IT-Fachleute)

- **FIDO2/WebAuthn-Rollout:** Einführung für privilegierte Konten zuerst, dann Fernzugriff, dann übrige extern erreichbare Dienste.
- **Automatisierte Sperrung:** technische Anbindung, sodass ein gemeldeter Verlust eines Ausweismittels die Berechtigung sofort und automatisiert sperrt, nicht erst nach manueller Bearbeitung.
- **Kontext-/Anomalieschutz:** wo phishing-resistente Verfahren (noch) nicht flächendeckend möglich sind, zusätzliche Geräte-Bindung und kontinuierliches Zugriffsmonitoring als Kompensation.
- **Altverfahren-Abschaltung:** definiertes Enddatum für OTP-/Push-only-Verfahren nach Migrationsphase.
- **Wirksamkeitstest:** Ein simulierter AITM-Phishing-Test (kontrollierte Umgebung) gegen ein Konto mit phishing-resistenter Authentisierung muss fehlschlagen, während derselbe Test gegen ein Konto mit klassischem OTP nachweislich erfolgreich wäre — der Unterschied macht die Wirksamkeit sichtbar.

5. Umsetzungsbeispiele

Beispiel A — Organisation mit eigener Identitätsplattform. Ein Hersteller mit zentralem Identity Provider rollt FIDO2-Sicherheitsschlüssel zunächst für alle administrativen und Fernzugriffs-Konten aus, danach für alle extern erreichbaren Anwendungen. Die Sperrung bei Verlust ist über die Identitätsplattform automatisiert und wirkt in Echtzeit.
Beispiel B — Organisation mit vielen Alt- und Fremdsystemen. Eine Organisation mit heterogener, teils historisch gewachsener Systemlandschaft kann nicht überall FIDO2 einführen. Sie priorisiert die Umstellung auf privilegierte und extern erreichbare Zugänge und kompensiert bei verbleibenden Altsystemen mit zusätzlichem Kontextmonitoring (Geräte-Bindung, Standort-/Anomalieprüfung), bis eine vollständige Migration möglich ist — dokumentiert als befristete Kompensationsmaßnahme, nicht als Dauerzustand.

6. Reifegrad-Pfad (kumulativ)

- **Initial:** phishing-resistentes Verfahren für privilegierte Konten pilotiert.
- **Defined:** phishing-resistentes Verfahren für alle privilegierten und Fernzugriffs-Konten produktiv; reine Passwort-Authentisierung ohne zusätzlichen Faktor organisationsweit ausgeschlossen; automatisierte Sperrung bei Verlust.
- **Managed:** flächendeckend für alle extern erreichbaren Dienste; Altverfahren vollständig abgeschaltet.

7. Messung und Audit-Nachweis

- **Kennzahl:** Anteil privilegierter und extern erreichbarer Zugänge mit phishing-resistenter Authentisierung (Zielwert: 100 % bei privilegierten Konten).
- **Nachweis:** Authentisierungsrichtlinie, IAM-Konfigurationsauszug, Sperrprotokolle, Migrationsplan mit Enddatum für Altverfahren.

- **Prüflogik:** dokumentiert? — Richtlinie mit Verfahrensvorgabe; verantwortlich? — IAM-Team; Häufigkeit? — Sperrung in Echtzeit, Migrationsfortschritt quartalsweise geprüft; Toleranz? — kein privilegierter Zugang ohne phishing-resistenten Faktor ohne dokumentierte Kompensation.
- **Evidence-Mapping zum ISA-Katalog:** Der ISA-Katalog selbst hinterlegt in den Spalten "Possible evidence"/"Possible questions" punktuell Beispiele, die den oben stehenden MHC-Nachweis ergänzen, ihn aber nicht ersetzen (im Quellkatalog nur für einen Teil der Controls gefüllt).

ISA-Control	Katalogseitiges Beispiel
4.1.1	Frage: "Where have you defined which supporting assets exist and which information assets may be stored on them?"; "How are supporting assets with high protection requirements protected?" (u. a.; Katalog nennt hierzu kein Beispiel-Nachweisdokument)
4.1.2	Frage: "What user authentication methods are used in the organization?"; "What criteria were used to select the user authentication methods (risk assessment, state of the art, business and security requirements, etc.)?" (u. a.; Katalog nennt hierzu kein Beispiel-Nachweisdokument)
4.1.3	Frage: "How are user accounts created, changed, blocked and deleted? Does it use unique and personalized user accounts?"; "What security measures have been implemented when using 'collective accounts'?" (u. a.; Katalog nennt hierzu kein Beispiel-Nachweisdokument)
5.2.7	Frage: "What requirements have you defined for managing and controlling your network?"; "How do you check that these requirements are complied with or fulfilled?" (u. a.; Katalog nennt hierzu kein Beispiel-Nachweisdokument)

Für 5.1.2 und 6.1.3 hinterlegt der Quellkatalog kein Beispiel; hier trägt allein der oben stehende MHC-Nachweis.

8. Typische Fehler

- "Zweiter Faktor" wird mit "phishing-resistent" gleichgesetzt — OTP/Push gelten fälschlich bereits als ausreichend.
- Die Sperrung bei Verlust ist zwar möglich, läuft aber über einen manuellen Ticket-Prozess statt automatisiert.
- Rollout beginnt bei Standardnutzern statt bei privilegierten/exponierten Konten, wo das Risiko am größten ist.
- Kompensationsmaßnahmen für Altsysteme werden nie mit Enddatum versehen und bleiben dauerhaft bestehen.

9. Abgrenzung, Restrisiko und Verweise

- **Abgrenzung:** betrifft die Authentisierungsmethode; Berechtigungsvergabe und -entzug regelt das jeweilige IAM-Prozessdokument, nicht dieses MHC.
- **Restrisiko:** Session-Hijacking nach erfolgreicher Authentisierung wird durch phishing-resistente Anmeldung nicht verhindert — dafür ist zusätzliches Session-Monitoring nötig.
- **TISAX®-Controls:** 4.1.1, 4.1.2, 4.1.3, 5.1.2, 5.2.7, 6.1.3 (siehe Teil I; 4.1.2 dort als einziger Reine-Reibung-Befund).
- **Framework:** NIS2 Art. 21(2)(j) (erste EU-Rechtsnorm, die MFA bzw. Continuous Authentication namentlich als geeignete Maßnahme im risikobasierten Maßnahmenkatalog nennt, dort "where appropriate" statt als absolute Pflicht formuliert); NIST SP 800-63B (AAL3, phishing-resistente Authentifikatoren).

MHC-05 — Verhaltensbasierte Detection und Kill-Chain-Korrelation

Auf einen Blick

- Operativer Nutzen: Erkennt sowohl unbekannte, signaturlose Schadsoftware anhand ihres Verhaltens als auch Angriffe, die bewusst in viele kleine, für sich unauffällige Schritte zerlegt sind — in ihrer Summe statt Schritt für Schritt.
- Betroffene Richtlinie: Malware-Schutz-/Endpoint-Security-Richtlinie; Meldewesen-, Log-Auswertungs- und Krisenerkennungsrichtlinie.
- Abhängigkeiten: keine Vorbedingung; Endpunkt- und Korrelationsebene teilen sich eine zentrale Log-Aggregationsbasis und verstärken sich gegenseitig.
- Aufwandstreiber: Produktauswahl (EDR/XDR), Integration in bestehende Endpoint- und SIEM-Landschaft, Tuning zur Vermeidung von Fehlalarmen, Reife der vorhandenen Log-Infrastruktur.
- Kostenrahmen (Lizenz/Personentage, indikative Größenordnung ohne Einzelpreise): KMU — Lizenz mittel (EDR meist pro Endpunkt lizenziert, Managed-Detection-Dienst statt eigenem SOC oft wirtschaftlicher), ca. 15–30 Personentage für Rollout/Anbindung, danach laufender Betrieb überwiegend beim Dienstleister · Midcap — Lizenz mittel bis hoch (EDR/XDR-Suite plus SIEM-/Korrelationsplattform), ca. 30–70 Personentage · Enterprise — Lizenz hoch (großflächige EDR/XDR-Abdeckung, viele Log-Quellen, ggf. eigenes SOC), ca. 70–180 Personentage, danach laufender Tuning- und Betriebsaufwand.
- Primäre Kennzahl: Anteil der Systeme mit Schutzbedarf hoch/sehr hoch, die durch eine verhaltensbasierte Komponente (EDR/XDR) abgedeckt sind; ergänzend Anteil sicherheitsrelevanter Log-Quellen in zeitübergreifender Korrelation.
- Berührte Standards (ohne Erfüllungsanspruch): ISA/IEC 62443-2-1:2024 (Verhaltensanalytik/Continuous Monitoring als Erweiterung genannt); NIST CSF 2.0 (DE.AE); TISAX® ISA 2027 5.2.3, 1.6.1, 1.6.3, 5.2.4, 5.2.7.

1. Härtungscontrol-Statement

Malware-Schutz stützt sich nicht ausschließlich auf signaturbasierte Erkennung: Für Systeme mit erhöhtem Schutzbedarf wird eine verhaltens- oder anomaliebasierte Erkennungskomponente (EDR-/XDR-Klasse) ergänzt. Parallel dazu stützen sich Melde- und Log-Auswertungsprozesse nicht ausschließlich auf einzeln auffällige Ereignisse: Eine Korrelation von Ereignissen über Zeit und mehrere Quellen hinweg (z. B. nach UEBA-Ansatz) erkennt bewusst fragmentierte, für sich unauffällige Angriffsschritte in ihrer Summe. Beide Ebenen — Endpunkt-Verhalten und quellenübergreifende Korrelation — laufen auf derselben zentralen Log-Aggregationsbasis zusammen. Krisenerkennungs- und Eskalationsverfahren werden ausdrücklich um das Szenario einer schnell, AI-orchestriert und mehrgleisig eskalierenden Lage ergänzt.

2. Zweck und Bedrohungsbezug

Klassischer, signaturbasierter Malwareschutz ist gegen AI-generierte oder polymorphe Schadsoftware strukturell schwächer, da Signaturen erst nach Bekanntwerden einer Variante greifen — bei AI-gestützter Variantenerzeugung entsteht die nächste Variante schneller, als neue Signaturen verteilt werden können. Verschlüsselter Inhalt umgeht zudem die Gateway-Inspektion strukturell, ein Umgehungsweg, den AI-gestützte Angreifer gezielt nutzen. Gleichzeitig erkennt ein Meldewesen, das davon abhängt, dass ein Mensch etwas Auffälliges bemerkt, keinen Angriff, der bewusst in viele kleine, für sich harmlos wirkende Schritte zerlegt ist — kein einzelner Schritt fällt dabei auf. Die von Anthropic im September 2025 aufgedeckte Kampagne GTG-1002 (siehe Teil I, Methodik) führte zunächst eine unauffällige, "low-and-slow" angelegte Aufklärung innerhalb normaler Nutzungsmuster durch, bevor die restlichen Schritte im Maschinentempo folgten — ein Muster, das weder reine Signaturerkennung noch eine auf menschliche Einzelfall-Auffälligkeit ausgelegte Log-Prüfung zusammenhängend erfasst. Klassische Krisenerkennung ist zudem auf allmählich eskalierende Lagen ausgelegt, nicht auf einen AI-orchestrierten, mehrgleisigen Angriff mit rascher Eskalation. Schutzziel ist, dass sowohl unbekannte Schadsoftware anhand ihres Verhaltens als auch die Summe unauffälliger Einzelschritte als zusammenhängendes Muster erkannt werden.

3. Organisatorische Umsetzung (Leitungsperspektive)

- **Richtlinie:** Die Malware-Schutz-Richtlinie schreibt für Schutzbedarf hoch/sehr hoch eine verhaltensbasierte Komponente zusätzlich zur signaturbasierten Erkennung vor; die Log-Auswertungs- und Krisenerkennungsrichtlinie schreibt eine quellen- und zeitübergreifende Korrelation zusätzlich zur Einzelereignis-Prüfung vor.
- **Verantwortlichkeiten:** IT-Betrieb verantwortet Produktauswahl, Rollout und Tuning der Endpunkt-Erkennung sowie die Pflege der Korrelationsregeln; ein Security-Operations-Bereich (intern oder als Managed Service) übernimmt die Alarmbearbeitung; CISO-Funktion verantwortet die Erweiterung der Krisenszenarien.
- **Risikoanalyse:** Priorisierung nach Schutzbedarf (Systeme hoch/sehr hoch zuerst) und nach Kritikalität der einzubindenden Log-Quellen.
- **Prozesse:** definierter Ablauf zur Bearbeitung verhaltensbasierter Alarme inklusive Tuning-Zyklus; Krisenszenarien-Katalog wird um mindestens ein AI-orchestriertes, mehrgleisiges Eskalationsszenario ergänzt und regelmäßig geübt.

4. Technische Umsetzung (für IT-Fachleute)

- **EDR-/XDR-Rollout:** Auswahl und Ausrollung einer verhaltensbasierten Erkennungskomponente auf priorisierten Systemen; Etablierung eines Normalverhaltens-Profils je Systemklasse.
- **Verschlüsselte-Inhalte-Kompensation:** wo Gateway-Inspektion verschlüsselten Verkehrs nicht möglich ist, endpunktseitige Kontrolle (Prozess-/Dateiverhalten) als Ausgleich.
- **Zentrale Log-Aggregation:** Zusammenführung sicherheitsrelevanter Logs aus mehreren Quellen (Endpoint, Netzwerk, Identität, Anwendung) in eine gemeinsame Auswertungsplattform — dieselbe Basis, auf der auch die EDR-/XDR-Alarme auflaufen.
- **UEBA-Korrelationsregeln und Zeitfenster-Korrelation:** Normalverhaltens-Baselines je Nutzer/System; Regeln, die nicht nur Einzelereignisse, sondern Ereignisketten über definierte Zeitfenster (Stunden bis Tage) auswerten und zu einem Gesamtbild verdichten.
- **Wirksamkeitstest:** Zwei getrennte Tests. Erstens: ein kontrolliertes, verhaltensauffälliges, aber signaturloses Testskript (vergleichbar einem Atomic-Red-Team-Test) muss allein durch Verhalten erkannt und alarmiert werden. Zweitens: eine simulierte "low-and-slow"-Testkette (mehrere für sich unauffällige Einzelaktionen, über einen längeren Zeitraum verteilt) muss als zusammenhängendes Muster erkannt und alarmiert werden.

5. Umsetzungsbeispiele

Beispiel A — Organisation mit Security-Operations-Team. Ein Hersteller mit eigenem SOC rollt eine EDR-Lösung auf allen Systemen mit Schutzbedarf hoch/sehr hoch aus und speist deren Alarme in dieselbe zentrale Erkennungsplattform ein, die auch die quellenübergreifende UEBA-Korrelation trägt. Ein monatlicher Tuning-Zyklus reduziert Fehlalarme auf beiden Ebenen. Zusätzlich ergänzt das Unternehmen seinen Krisenszenarien-Katalog um ein AI-orchestriertes, mehrgleisiges Eskalationsszenario und übt es jährlich gegen die reale Erkennungskette.
Beispiel B — Organisation ohne eigenes SOC. Eine Organisation ohne eigenes Security-Operations-Team bezieht sowohl die verhaltensbasierte Endpunkt-Erkennung als auch die quellenübergreifende Korrelation als Managed-Detection-Dienst mit vertraglich vereinbarter Reaktionszeit bei kritischen Alarmen, und stellt sicher, dass alle eigenen sicherheitsrelevanten Log-Quellen vollständig an den Dienst angebunden sind.

6. Reifegrad-Pfad (kumulativ)

- **Initial:** verhaltensbasierte Komponente auf den kritischsten Systemen pilotiert; zentrale Log-Aggregation für die wichtigsten Quellen etabliert.
- **Defined:** Rollout auf alle Systeme mit Schutzbedarf hoch/sehr hoch abgeschlossen; Korrelationsregeln über mehrere Quellen und Zeitfenster aktiv; Krisenszenarien-Katalog um AI-orchestrierte Eskalation ergänzt.
- **Managed:** Baseline-Tuning auf beiden Ebenen etabliert, Fehlalarmrate niedrig, UEBA-Baseline regelmäßig getestet; Krisenszenario wird geübt, nicht nur dokumentiert.

7. Messung und Audit-Nachweis

- **Kennzahl:** Anteil der Systeme mit Schutzbedarf hoch/sehr hoch mit aktiver verhaltensbasierter Erkennungskomponente (Zielwert: vollständige Abdeckung); Anteil sicherheitsrelevanter Log-Quellen in zeitübergreifender Korrelation (Zielwert: alle Quellen mit Schutzbedarf hoch/sehr hoch).

- **Nachweis:** Rollout-Dokumentation, Konfigurationsauszug, Korrelationsregel-Dokumentation, Ergebnisse beider Wirksamkeitstests, Alarmstatistiken, aktualisierter Krisenszenarien-Katalog mit Übungsprotokollen.
- **Prüflogik:** dokumentiert? — Richtlinie, Rollout-Plan, Korrelationsregeln; verantwortlich? — IT-Betrieb/SOC, CISO-Funktion für Krisenszenarien; Häufigkeit? — laufender Betrieb, Wirksamkeitstests und Krisenübung mindestens jährlich; Toleranz? — kein System mit Schutzbedarf hoch/sehr hoch und keine kritische Log-Quelle ohne Abdeckung ohne dokumentierte Ausnahme.
- **Evidence-Mapping zum ISA-Katalog:** Für Control 5.2.7 nennt der ISA-Katalog als Beispielfragen: "What requirements have you defined for managing and controlling your network?"; "How do you check that these requirements are complied with or fulfilled?" (u. a.; Katalog nennt hierzu kein Beispiel-Nachweisdokument). Für 5.2.3, 1.6.1, 1.6.3 und 5.2.4 hinterlegt der Katalog kein Beispiel; hier trägt allein der oben stehende MHC-Nachweis — dieses Control-Bündel beruht inhaltlich stark auf originärer MHC-Konkretisierung (siehe Abschnitt 9, Framework).

8. Typische Fehler

- Die verhaltensbasierte Komponente wird installiert, aber nie auf die eigene Umgebung abgestimmt — hohe Fehlalarmrate führt dazu, dass Alarme ignoriert werden.
- Log-Quellen werden zentral gesammelt, aber nie tatsächlich korreliert — reine Datenablage ohne Auswertungslogik.
- Beide Ebenen laufen technisch getrennt statt auf einer gemeinsamen Basis, wodurch ein Angriff, der auf der einen Ebene unter der Schwelle bleibt, auf der anderen nicht kompensiert wird.
- Das ergänzte Krisenszenario bleibt Papierform und wird nie geübt.

9. Abgrenzung, Restrisiko und Verweise

- **Abgrenzung:** betrifft Erkennung auf Endpunktebene und quellenübergreifende Korrelation; die automatisierte Reaktion auf einen erkannten Vorfall regelt MHC-11, die vorbeugende Schließung der zugrunde liegenden Schwachstelle MHC-09.
- **Restrisiko:** verhaltensbasierte Erkennung und Korrelation reduzieren, beseitigen aber nicht das Risiko vollständig neuartiger, gezielt auf die eigene Baseline und die definierten Korrelationsschwellen zugeschnittener Angriffe.
- **TISAX®-Controls:** 5.2.3, 1.6.1, 1.6.3, 5.2.4, 5.2.7 (siehe Teil I).
- **Framework:** ISA/IEC 62443-2-1:2024 nennt Verhaltensanalytik und kontinuierliches Monitoring als Erweiterung, ohne dies als scharfe Pflicht zu formulieren; NIST CSF 2.0 (DE.AE, Anomalies and Events) — überwiegend originäre MHC-Konkretisierung.

MHC-08 — Unveränderliche Backups und Recovery-Validierung

Auf einen Blick

- Operativer Nutzen: Stellt sicher, dass Datensicherungen auch bei kompromittierten Admin-Rechten wiederherstellbar bleiben, und dass die Wiederherstellung im Ernstfall tatsächlich funktioniert statt nur auf dem Papier zu existieren.

- Betroffene Richtlinie: Backup-Konzept und Continuity-Planung.
- Abhängigkeiten: keine Vorbedingung; ergänzt MHC-09 als reaktive Seite der Ransomware-Resilienz.
- Aufwandstreiber: vorhandene Backup-Infrastruktur, Umstellung auf Object-Lock-/WORM-fähige Speichersysteme, Trennung der administrativen Zugriffswege.
- Kostenrahmen (Lizenz/Personentage, indikative Größenordnung ohne Einzelpreise): KMU — Lizenz gering bis mittel (Object-Lock-fähiger Cloud-Speicher oft ohne wesentlichen Aufpreis beim vorhandenen Provider), ca. 10–20 Personentage · Midcap — Lizenz mittel (größeres Datenvolumen, ggf. dedizierte WORM-Speicherkomponente), ca. 20–45 Personentage · Enterprise — Lizenz mittel bis hoch (große, heterogene Backup-Landschaft, mehrere Standorte), ca. 45–100 Personentage.
- Primäre Kennzahl: Erfolgsquote regelmäßiger Wiederherstellungstests aus unveränderlichen Backups.
- Berührte Standards (ohne Erfüllungsanspruch): BSI C5:2026 (OPS-06 bis OPS-09); DORA Art. 12; NIST SP 1800-11/1800-25 (Data Integrity, NCCoE); TISAX® ISA 2027 5.2.8.

1. Härtungscontrol-Statement

Datensicherungen kritischer IT-Services werden, wo der Schutzbedarf dies rechtfertigt, unveränderlich (immutable) oder physisch/logisch getrennt vorgehalten. Administrative Lösch- oder Änderungsrechte auf das Backup-Repository sind vom produktiven Administrationsbereich getrennt, sodass eine Kompromittierung produktiver Admin-Konten die Sicherungen nicht mit erfasst. Die Wiederherstellung wird regelmäßig vollständig getestet, nicht nur die Sicherung selbst.

2. Zweck und Bedrohungsbezug

Ransomware mit kompromittierten Backup-Administrationsrechten kann bestehende Sicherungen löschen oder verschlüsseln, wenn diese nicht technisch unveränderlich sind — eine "ausreichend geschützte" Sicherung im Sinne einer reinen Zugriffsliste hält einem hinreichend befähigten, AI-gestützt beschleunigten Angriff nicht stand. TISAX® ISA 2027 fordert einen Schutz gegen unautorisierte Veränderung/Löschung explizit erst ab Schutzbedarf hoch, und selbst dort ohne Unveränderlichkeits-Vorgabe im Wortlaut ("sufficiently protected"). Schutzziel ist, dass eine Wiederherstellung auch nach Kompromittierung privilegierter Konten möglich bleibt, und dass diese Fähigkeit im Ernstfall nachweislich funktioniert.

3. Organisatorische Umsetzung (Leitungsperspektive)

- **Richtlinie:** Die Backup-Richtlinie schreibt Unveränderlichkeit oder physische/logische Trennung für Schutzbedarf hoch/sehr hoch verbindlich vor, unabhängig davon, ob der TISAX®-Wortlaut an der jeweiligen Stelle bereits so weit geht.
- **Verantwortlichkeiten:** IT-Betrieb verantwortet Betrieb und Testung der Backup-Infrastruktur; die administrativen Rechte auf das unveränderliche Repository werden einer separat verwalteten Rolle zugewiesen.
- **Risikoanalyse:** Priorisierung nach Kritikalität der IT-Services; Systeme mit Schutzbedarf hoch/sehr hoch zuerst auf unveränderliche Sicherung umgestellt.
- **Prozesse:** regelmäßiger, vollständiger Wiederherstellungstest mit dokumentiertem Ergebnis; Eskalation bei fehlgeschlagenem Test.

4. Technische Umsetzung (für IT-Fachleute)

- **Unveränderliche Backups:** technische Umsetzung über Object Lock (S3-kompatibel), WORM-Speicher oder physisch getrennte/Offline-Kopien.
- **Getrennte Administration:** administrative Lösch-/Änderungsrechte auf das Backup-Repository werden vom produktiven Administrationsbereich technisch getrennt (eigene Identität, kein Durchgriff kompromittierter Produktiv-Admin-Konten).
- **Restore-Testing:** regelmäßiger, vollständiger Wiederherstellungstest (nicht nur Stichprobe) mit Protokollierung von Ergebnis und benötigter Zeit gegen ein definiertes RTO.
- **Wirksamkeitstest:** Ein simulierter Löschversuch auf ein als unveränderlich markiertes Backup nach angenommener Admin-Kompromittierung muss technisch fehlschlagen; zusätzlich wird ein vollständiger Restore-Test durchgeführt und gegen das definierte RTO gemessen.

5. Umsetzungsbeispiele

Beispiel A — Organisation mit eigenem Backup-Betrieb. Ein Hersteller betreibt für Systeme mit Schutzbedarf hoch/sehr hoch ein Object-Lock-fähiges Speichersystem, dessen Lösch- und Änderungsrechte separat von den produktiven Administrationskonten verwaltet werden. Ein vierteljährlicher, vollständiger Wiederherstellungstest ist Teil des Regelbetriebs.
Beispiel B — Organisation mit ausgelagertem IT-Betrieb. Eine Organisation, die ihren IT-Betrieb an einen Dienstleister ausgelagert hat, verlangt vom Dienstleister einen Nachweis der eingesetzten Unveränderlichkeits-Technologie (Object Lock, WORM) sowie regelmäßige, dokumentierte Wiederherstellungstestergebnisse als Vertragsbestandteil.

6. Reifegrad-Pfad (kumulativ)

- **Initial:** Backups vorhanden; Wiederherstellungstests jährlich.
- **Defined:** 3-2-1-1-0-Prinzip umgesetzt, mindestens eine unveränderliche oder physisch/logisch getrennte Kopie; administrative Lösch- und Änderungsrechte auf das Backup-Repository vom produktiven Administrationsbereich getrennt; vollständige Wiederherstellungstests im schutzbedarfsgerechten Zyklus (Zielwert vierteljährlich).
- **Managed:** eigene Administrations- und Identitätsebene für die Backup-Infrastruktur; automatisierte Integritätsprüfung; orchestrierte Wiederherstellungs-Validierung und Stichprobentests nach wesentlichen Änderungen.

MVP-Staffelung nach Schutzbedarf: Der vierteljährliche vollständige Wiederherstellungstest ist ein ambitionierter Zielwert, der nicht für jede Organisation ab Tag eins verhältnismäßig ist — Staffelung analog zur bereits bestehenden Formulierung "wo der Schutzbedarf dies rechtfertigt" (Abschnitt 1): bei **normalem Schutzbedarf** genügt als MVP ein jährlicher vollständiger Wiederherstellungstest, Unveränderlichkeit ist hier nicht zwingend erforderlich; bei **hohem Schutzbedarf** wird der Test halbjährlich fällig, Unveränderlichkeit (Object-Lock/WORM) und getrennte Administration werden verbindlich; bei **sehr hohem Schutzbedarf** gilt der vierteljährliche Test als Zielwert unverändert, ergänzt um zusätzliche Stichprobentests nach wesentlichen Änderungen an der Backup-Infrastruktur.

7. Messung und Audit-Nachweis

- **Kennzahl:** Erfolgsquote der vierteljährlichen Wiederherstellungstests aus unveränderlichen Backups (Zielwert: 100 %).
- **Nachweis:** Protokolle der Wiederherstellungstests, Nachweis der Unveränderlichkeit (Aufbewahrungs- und Sperreinstellungen), getrennte Rechtevergabe der Backup-Infrastruktur.
- **Prüflogik:** dokumentiert? — Backup-Richtlinie und Testprotokolle; verantwortlich? — IT-Betrieb; Häufigkeit? — vierteljährliche Tests, fortlaufende Integritätsprüfung; Toleranz? — kein geschäftskritischer Datenbestand ohne unveränderliche oder getrennte Kopie und ohne bestandenen Wiederherstellungstest.
- **Evidence-Mapping zum ISA-Katalog:** Control 5.2.8 enthält im ISA-Katalog selbst kein Beispiel in den Spalten "Possible evidence"/"Possible questions"; der oben stehende MHC-Nachweis ist hier die alleinige Grundlage für die Audit-Vorbereitung.

8. Typische Fehler

- Es gibt Backups, aber sie sind über dieselben Zugänge erreichbar wie die produktiven Systeme — ein Angreifer verschlüsselt beides.
- Sicherungen werden erstellt, aber die Wiederherstellung wird nie vollständig getestet; im Ernstfall fehlen Teile oder der Vorgang dauert zu lange.
- "Unveränderlich" ist konfiguriert, aber nicht überprüft; eine zu kurze Sperrfrist macht die Kopie angreifbar.
- Unveränderlichkeit wird nur für einen Teil der Systeme umgesetzt, ohne dass der Schutzbedarf dies begründet.

9. Abgrenzung, Restrisiko und Verweise

- **Abgrenzung:** betrifft Sicherung und Wiederherstellung; die Behebungsgeschwindigkeit identifizierter Schwachstellen regelt MHC-09, die Erkennung des Angriffs selbst MHC-05. TISAX® ISA 2027 5.2.9 bewertet die Backup-*Medium*-Frage bei sehr hohem Schutzbedarf bereits eigenständig als Standfest (siehe Teil I) — MHC-08 schließt die davon unabhängige Lücke aus 5.2.8 (Continuity-Planning-Ebene, bereits ab Schutzbedarf hoch).
- **Restrisiko:** eine vom Netz getrennte Kopie ist naturgemäß weniger aktuell; eine sehr langsame oder selten getestete Wiederherstellung kann im Ernstfall dennoch zu Ausfallzeiten führen.
- **TISAX®-Controls:** 5.2.8 (siehe Teil I; vgl. 5.2.9 für die eigenständig bereits Standfest bewertete Backup-Medium-Frage bei sehr hohem Schutzbedarf).

- **Framework:** BSI C5:2026 OPS-06 bis OPS-09 (verlangt Verschlüsselung, RTO/RPO-Ausrichtung, Monitoring und regelmäßige Restore-Tests, aber keine Unveränderlichkeit — diese Lücke bleibt auch gegenüber C5:2026 offen); DORA Art. 12 (sektorspezifischer Referenzrahmen Finanzsektor: Backup-Richtlinien, segregierte Wiederherstellungssysteme, Datenintegritätsprüfung bei Recovery); NIS2-DVO Nr. 4.2 (Backup- und Redundanzmanagement, Annex-Abschnitt zu Art. 21(2)(c) NIS2-RL); NIST SP 800-34 Rev. 1; NIST SP 1800-11 (Ransomware-Recovery) und 1800-25 (Ransomware-Identifikation/-Schutz), NCCoE; Industriepraxis 3-2-1-1-0.

MHC-09 — AI-gestütztes Security-Testing in der Pipeline

Auf einen Blick

- Operativer Nutzen: Verkürzt die Zeit zwischen bekannter Schwachstelle und Behebung auf ein Maß, das mit AI-beschleunigter Exploit-Entwicklung mithalten kann.
- Betroffene Richtlinie: Schwachstellen- und Patch-Management.
- Abhängigkeiten: keine Vorbedingung; erhöht die Wirksamkeit von MHC-11; ergänzt MHC-08 als präventive Seite der Ransomware-Resilienz.
- Aufwandstreiber: Anzahl der zu patchenden Systeme, Reife des Vulnerability-Scannings.
- Kostenrahmen (Lizenz/Personentage, indikative Größenordnung ohne Einzelpreise): KMU ohne eigene Entwicklung — Lizenz gering (Vulnerability-Scanner), ca. 10–20 Personentage; mit eigener Entwicklung zusätzlich SAST/DAST/SCA-Basiswerkzeuge, ca. 15–30 Personentage · Midcap — Lizenz mittel (Scanner plus Pipeline-Werkzeuge SAST/DAST/SCA), ca. 25–60 Personentage · Enterprise — Lizenz mittel bis hoch (viele Systeme/Repositories, mehrere Pipelines), ca. 60–150 Personentage.
- Primäre Kennzahl: Anteil kritischer Schwachstellen, die innerhalb der CVSS-gebundenen Frist behoben werden; bei vorhandener Softwareentwicklung ergänzend die Behebungslatenz für aktiv ausgenutzte Schwachstellen (KEV): die Zeit von der Verfügbarkeit einer belastbaren Behebung oder wirksamen Kompensationsmöglichkeit bis zu deren verifizierter Umsetzung. Als Behebung gelten Patch, Umgehungslösung, kompensierende Maßnahme, Isolierung oder Ablösung.
- Berührte Standards (ohne Erfüllungsanspruch): BSI C5:2026 (OPS-25, Schwachstellenmanagement-Basiskriterien); CRA Art. 14 (Meldefristen für Hersteller); TISAX® ISA 2027 1.3.4, 5.2.5, 5.2.6.

1. Härtungscontrol-Statement

Maßnahmen zur Behebung identifizierter Schwachstellen werden an feste, nach Kritikalität gestaffelte Fristen gebunden (z. B. nach CVSS-Score), nicht an unbestimmte Begriffe wie "angemessen" oder "zeitnah". Für Schutzbedarf hoch/sehr hoch gilt für kritische Schwachstellen eine Obergrenze im Stunden- bis Tagesbereich. System- und Service-Audits, die den Erfolg der Behebung verifizieren, laufen so häufig wie technisch vertretbar statt in starren, langen Zyklen. Wo eigene Softwareentwicklung stattfindet, gilt ergänzend: Sicherheitsprüfungen laufen automatisch in der Entwicklungs-Pipeline (Code, Abhängigkeiten, laufende Anwendung), schwerwiegende Funde blockieren die Übernahme, und AI-generierter Code wird als eigener Risikofaktor im Standard für sichere Entwicklung berücksichtigt und zusätzlich geprüft.

2. Zweck und Bedrohungsbezug

Eine 2026 veröffentlichte Auswertung von Cogent Security über 69.159 CVE zeigt, dass sich die durchschnittliche Zeit von Schwachstellen-Veröffentlichung bis zum ersten funktionierenden Exploit von 125,3 Tagen (Januar 2025) auf 0,5 Tage (April 2026) verkürzt hat (siehe Teil I, Methodik). Eine Bewertungsmethodik, die mit Wochen bis zu einem funktionierenden Exploit rechnet, unterschätzt die Dringlichkeit systematisch — reine Kenntnis von Version und Patch-Stand sagt nichts über die Reaktionsgeschwindigkeit aus. Wo Software selbst entwickelt wird, kommt hinzu, dass AI-Programmierhilfen regelmäßig unsichere Muster erzeugen (fest eingebaute Zugangsdaten, fehlende Eingabeprüfung, veraltete Bausteine) und dadurch neue Schwachstellen schneller entstehen, als sie klassisch vor Auslieferung gefunden würden. Schutzziel ist, dass die Behebungsgeschwindigkeit — von der Entdeckung bis zum ausgerollten Patch, bei eigener Entwicklung bereits vor Auslieferung — mit der Geschwindigkeit der Bedrohung mithält.

3. Organisatorische Umsetzung (Leitungsperspektive)

- **Richtlinie:** Die Schwachstellenmanagement-Richtlinie schreibt CVSS-gebundene, konkrete Fristen vor; bei eigener Softwareentwicklung schreibt die Entwicklungsrichtlinie automatische Sicherheitsprüfungen in der Pipeline sowie die gesonderte Prüfung AI-generierten Codes vor.
- **Verantwortlichkeiten:** IT-Betrieb verantwortet Fristeinhaltung und Eskalation bei Überschreitung; bei eigener Entwicklung wird AI-generierter Code als eigene Risikokategorie im Statement of Applicability geführt.
- **Risikoanalyse:** Fristen werden nach Kritikalität und Exposition gestaffelt (extern erreichbare Systeme strenger als interne); Schwellen für blockierende Pipeline-Funde und der Umgang mit aktiv ausgenutzten Schwachstellen (KEV) werden festgelegt.
- **Prozesse:** Eskalationskette bei drohender Fristüberschreitung; dokumentierte Ausnahmeregelung mit Risikoakzeptanz, wo ein Patch technisch nicht fristgerecht möglich ist (z. B. Legacy-Systeme); bei eigener Entwicklung zusätzlich ein beschleunigter Weg für KEV-Fälle.

4. Technische Umsetzung (für IT-Fachleute)

- **CVSS-Integration:** Vulnerability-Scanner liefert automatisch den CVSS-Score je Fund; daran gekoppelt eine automatische Fristzuweisung im Ticketsystem, mit automatischem Eskalationslauf bei drohender oder eingetretener Fristüberschreitung.
- **Regelmäßige System-/Service-Audits:** Vulnerability-Scans laufen so häufig wie technisch vertretbar, nicht nur zu festen, seltenen Terminen — insbesondere für extern erreichbare Systeme.
- **Bei eigener Softwareentwicklung zusätzlich:** SAST/DAST/SCA-Prüfklassen automatisch bei jeder Änderung in der Pipeline; Pre-Merge-Block für schwerwiegende Funde ab definierter Verlässlichkeit; aktiv ausgenutzte Schwachstellen (KEV) blockieren sofort; zusätzliche Prüfregeln gegen typische Schwächen AI-generierten Codes; tägliche Scans auch der bereits laufenden Systeme, nicht nur des neuen Codes.
- **Wirksamkeitstest:** Eine Test-Schwachstelle mit definiertem CVSS-Wert wird injiziert; gemessen wird die Zeit bis zur automatischen Ticket-Erstellung mit korrekter Frist. Bei eigener Entwicklung zusätzlich: In einem Testzweig wird bewusst eine bekannte unsichere Stelle eingebaut; die Pipeline muss den Fund melden und die Übernahme blockieren.

- **Wirksamkeitsgrenze:** automatisiertes Testing ersetzt nicht die fachliche Bewertung und übersieht insbesondere Logik-, Berechtigungs- und Architekturschwächen; kritische Findings und akzeptierte Ausnahmen brauchen fachliche Triage und dokumentierte Freigabe.

5. Umsetzungsbeispiele

Beispiel A — Organisation mit eigener Softwareentwicklung. Ein Softwarehersteller koppelt den CVSS-Score jedes Scan-Funds automatisch an eine Fristmatrix im Ticketsystem und baut zusätzlich Sicherheitsprüfungen direkt in die Entwicklungs-Pipeline ein: Code, Abhängigkeiten und laufende Anwendung werden bei jeder Änderung automatisch geprüft, schwerwiegende Funde verhindern die Übernahme, und für aktiv ausgenutzte Schwachstellen gibt es einen beschleunigten Weg. Da die Entwicklung AI-Programmierhilfen nutzt, ergänzt das Unternehmen Prüfregeln gegen deren typische Fehler und wertet diese Funde gesondert aus.
Beispiel B — Organisation mit ausgelagertem IT-Betrieb, ohne eigene Entwicklung. Eine Organisation ohne eigene Softwareentwicklung nimmt die CVSS-gebundenen Fristen als vertragliche SLA in die Dienstleistungsvereinbarung mit ihrem IT-Dienstleister auf und lässt sich die Einhaltung monatlich berichten, inklusive Nachweis regelmäßiger Vulnerability-Scans.

6. Reifegrad-Pfad (kumulativ)

- **Initial:** CVSS-Score wird erfasst, Fristen sind dokumentiert, aber manuell nachverfolgt.
- **Defined:** Fristzuweisung automatisiert im Ticketsystem; Eskalationskette definiert und aktiv; Verifikationsaudits im technisch vertretbaren Rhythmus statt in starren Langzyklen; bei eigener Entwicklung: SAST/DAST/SCA in der Pipeline mit Blockieren bei schwerwiegenden Funden und zusätzlicher Prüfung AI-generierten Codes.
- **Managed:** Fristeinhaltung wird laufend gemessen und berichtet; bei eigener Entwicklung: gesonderte Trendauswertung der Befunde in AI-generiertem Code, tägliche Scans der laufenden Systeme, beschleunigter Weg für KEV-Fälle.

MVP-Staffelung nach Schutzbedarf: Tägliche Scans sind der ambitionierte Zielwert für die höchste Schutzbedarfsstufe, nicht die MVP-Grundlinie: bei **normalem Schutzbedarf** genügen als MVP wöchentliche bis vierzehntägliche Scans mit den regulären CVSS-gestaffelten Fristen; bei **hohem Schutzbedarf** werden mindestens wöchentliche, für extern erreichbare Systeme risikobasiert häufigere Scans sowie spürbar verkürzte Fristen fällig; bei **sehr hohem Schutzbedarf** gilt der tägliche Scan als Zielwert unverändert, mit Behebung aktiv ausgenutzter Schwachstellen (KEV) innerhalb von 24 Stunden.

7. Messung und Audit-Nachweis

- **Kennzahl:** Anteil kritischer Schwachstellen, die innerhalb der definierten Frist behoben wurden (Zielwert: nahe 100 % ohne dokumentierte Ausnahme); bei eigener Entwicklung ergänzend die Behebungslatenz für aktiv ausgenutzte Schwachstellen (KEV) — Zeit von der Verfügbarkeit einer belastbaren Behebung oder wirksamen Kompensationsmöglichkeit bis zu deren verifizierter Umsetzung, wobei Patch, Umgehungslösung, kompensierende Maßnahme, Isolierung oder Ablösung zählen (Richtwert: unter 24 Stunden).
- **Nachweis:** Fristmatrix, Scan-Berichte mit CVSS-Score und Behebungsdatum, Eskalationsprotokolle; bei eigener Entwicklung zusätzlich Pipeline-Konfiguration und Blockier-Protokolle.

- **Prüflogik:** dokumentiert? — Fristmatrix nach CVSS; verantwortlich? — IT-Betrieb, bei eigener Entwicklung zusätzlich Entwicklungs-/Produktverantwortliche; Häufigkeit? — laufend, bei eigener Entwicklung bei jeder Änderung; Toleranz? — Fristüberschreitung nur mit dokumentierter Risikoakzeptanz, keine Auslieferung mit offenen schwerwiegenden Funden.
- **Evidence-Mapping zum ISA-Katalog:** Der ISA-Katalog selbst hinterlegt in den Spalten "Possible evidence"/"Possible questions" punktuell Beispiele, die den oben stehenden MHC-Nachweis ergänzen, ihn aber nicht ersetzen (im Quellkatalog nur für einen Teil der Controls gefüllt).

ISA-Control	Katalogseitiges Beispiel
1.3.4	Nachweis: "Process documentation for approval process"; "Exports of lists of approved software from software repositories"; "Screenshots from software management software (such as Mobile Device Management)"
5.2.5	Frage: "What information about technical vulnerabilities can be determined?"; "How do you respond to identified vulnerabilities?" (u. a.; Katalog nennt hierzu kein Beispiel-Nachweisdokument)

Für 5.2.6 hinterlegt der Quellkatalog kein Beispiel; hier trägt allein der oben stehende MHC-Nachweis.

8. Typische Fehler

- Die Frist wird ab Scan-Datum statt ab tatsächlichem Bekanntwerden der Schwachstelle gerechnet, wodurch verzögerte Scans die Frist künstlich verlängern.
- Ausnahmen von der Frist werden informell statt mit dokumentierter Risikoakzeptanz gewährt.
- Bei eigener Entwicklung: die Pipeline-Prüfungen laufen, blockieren aber nicht — schwerwiegende Funde werden gemeldet und trotzdem ausgeliefert.
- Bei eigener Entwicklung: nur der neue Code wird geprüft, nicht die bereits laufenden Systeme.

9. Abgrenzung, Restrisiko und Verweise

- **Abgrenzung:** betrifft Behebungsgeschwindigkeit, bei eigener Entwicklung zusätzlich das automatisierte Prüfen während der Entwicklung; die Erkennung eines Angriffs bzw. einer Ausnutzung selbst regelt MHC-05, die Wiederherstellbarkeit im Ransomware-Fall MHC-08.
- **Restrisiko:** eine Frist verkürzt das Zeitfenster, schließt es aber nicht auf null; Zero-Day-Schwachstellen ohne verfügbaren Patch bleiben unabhängig von der Frist ein Restrisiko; automatisierte Pipeline-Prüfungen finden bekannte Muster, nicht jede neuartige oder logische Schwachstelle.
- **Ausbaustufe:** Dieser Eintrag deckt die von TISAX® ISA 2027 konkret diagnostizierte Fristen- und Scan-Kadenz-Lücke ab. Die volle Pipeline-Integration (SAST/DAST/SCA als Pre-Merge-Block, AI-Code als eigene Prüfkategorie) ist für Organisationen mit eigener Softwareentwicklung der maßgebliche Ausbau; Details dazu im MRIS-Implementation-Guide 2.0 for ISO 27002, MHC-09.
- **TISAX®-Controls:** 1.3.4, 5.2.5, 5.2.6 (siehe Teil I).
- **Framework:** BSI C5:2026 OPS-25.01AC (verlangt einen definierten, überwachten, CVSS-gebundenen Fristenrahmen für die Behebung von Schwachstellen, verbindlich ab Prüfzeiträumen ab 1. Juni 2027; die konkreten CVSS-gestaffelten Beispiel-Fristwerte im Kriterienkatalog selbst konnten nicht direkt am BSI-Primärdokument verifiziert werden — siehe Teil I, Fußnote 2 zur Lückenanalyse) sowie OPS-25.01AS (Verschärfung auf tägliche Scans); CRA Art. 14 (24-Stunden-

Meldung aktiv ausgenutzter Schwachstellen an ENISA für Hersteller, bindend ab 11. September 2026); NIST SP 800-218 (SSDF) und SP 800-218A (AI) für die Pipeline-Ausbaustufe.

MHC-10 — Continuous Control Monitoring und Policy-as-Code

Auf einen Blick

- Operativer Nutzen: Schließt das Zeitfenster zwischen zwei periodischen Prüfzyklen, in dem neue Angriffsmuster, Vorfälle oder Systemänderungen sonst bis zum nächsten Turnus unentdeckt blieben.
- Betroffene Richtlinie: ISMS-Rahmenwerk (Review-, Audit- und Prüfprozesse organisationsweit).
- Abhängigkeiten: keine Vorbedingung; erhöht die Wirksamkeit von MHC-13.
- Aufwandstreiber: Anzahl der betroffenen Prüfprozesse (12 Controls) und Reife der vorhandenen GRC-/Ticketing-Werkzeuge.
- Kostenrahmen (Lizenz/Personentage, indikative Größenordnung ohne Einzelpreise): KMU — Lizenz gering (Beispiel-B-Pfad ohne GRC-Tooling, geteilte Tabelle plus manuelle Sichtung), ca. 10–20 Personentage · Midcap — Lizenz gering bis mittel (vorhandenes GRC-/Ticketing-Tool wird um Trigger-Logik erweitert), ca. 20–50 Personentage · Enterprise — Lizenz mittel (viele Prüfprozesse und Systeme, Anbindung an CMDB/Threat-Intel-Feeds), ca. 50–120 Personentage.
- Primäre Kennzahl: Anteil sicherheitsrelevanter Prüfprozesse mit definiertem anlassbezogenem Auslösemechanismus.
- Berührte Standards (ohne Erfüllungsanspruch): NIST CSF 2.0 (DE.CM); NIS2 Art. 21(2)(f); TISAX® ISA 2027 1.1.1, 1.2.1, 1.3.1, 1.3.3, 1.3.4, 1.5.1, 1.6.3, 4.1.3, 4.2.1, 5.2.9, 6.1.2, 6.1.3.

1. Härtungscontrol-Statement

Für sicherheitsrelevante Prüf-, Review- und Auditprozesse wird neben dem regulären Prüfzyklus ein anlassbezogener Auslösemechanismus eingeführt. Auslöser umfassen mindestens: neu bekanntgewordene, einschlägige Angriffsmuster oder Threat-Intelligence-Hinweise; sicherheitsrelevante Vorfälle im eigenen oder vergleichbaren Umfeld; wesentliche Änderungen der betroffenen System- oder Prozesslandschaft. Für Prüfgegenstände mit Schutzbedarf hoch/sehr hoch gilt zusätzlich eine feste maximale Zykluslänge anstelle von "regelmäßig"/"periodisch".

2. Zweck und Bedrohungsbezug

TISAX®-Anforderungen mit reinem Zeitzyklus ("regelmäßig", "periodisch") setzen voraus, dass sich die Bedrohungslage innerhalb dieses Zyklus nicht wesentlich ändert. Diese Annahme trägt gegen einen Gen-AI-beschleunigten Angreifer nicht mehr: neue Angriffsmuster, Werkzeuge und Schwachstellen-Exploits verbreiten sich innerhalb von Tagen bis Wochen, nicht erst zum nächsten Jahres- oder Quartalszyklus. Ein Control, das erst beim nächsten Turnus nachgezogen wird, bleibt in der Zwischenzeit auf dem alten Stand — genau das Zeitfenster, das ein automatisiert agierender Angreifer ausnutzt. Schutzziel ist, dass sicherheitsrelevante Prüfprozesse auf neue Erkenntnisse reagieren, sobald diese vorliegen, nicht erst beim nächsten Zyklus.

3. Organisatorische Umsetzung (Leitungsperspektive)

- **Richtlinie:** Das ISMS-Rahmenwerk wird um eine Klausel ergänzt, die für alle Prüf-, Review- und Auditprozesse mit sicherheitsrelevantem Bezug einen anlassbezogenen Auslösemechanismus zusätzlich zum Regelzyklus vorschreibt.
- **Verantwortlichkeiten:** Eine zentrale Stelle (i. d. R. CISO-Funktion) wird für die Pflege der Auslöser-Quellen (Threat-Intel-Feed, Incident-Register, Change-Management) benannt; die operative Ausführung der ausgelösten Prüfung bleibt bei der jeweils fachlich zuständigen Stelle.
- **Risikoanalyse:** Die Zykluslänge je Prüfgegenstand wird aus dem Schutzbedarf abgeleitet — kritische Bereiche (privilegierte Zugriffsrechte, extern erreichbare Systeme) erhalten die kürzeste feste Obergrenze.
- **Prozesse:** Definierter Ablauf, wie ein Auslöser (Threat-Intel-Treffer, Vorfall, Systemänderung) automatisch oder manuell eine außerplanmäßige Prüfung der betroffenen Controls anstößt, inklusive Fristsetzung bis zur Durchführung.

4. Technische Umsetzung (für IT-Fachleute)

- **Threat-Intel-Anbindung:** Abonnement eines relevanten Feeds (BSI-Lagebericht/CERT-Bund, Branchen-ISAC, kommerzieller Threat-Intel-Dienst); automatisierte Stichwort-/Kategorie-Filterung auf die eigene Systemlandschaft.
- **Incident-Trigger:** Verknüpfung des Ticketsystems, sodass ein als sicherheitsrelevant klassifizierter Vorfall automatisch eine Review-Aufgabe für die betroffenen Policies/Controls erzeugt.
- **Change-Trigger:** Anbindung an CMDB/Asset-Inventar, sodass eine wesentliche Änderung der Systemlandschaft (neues extern erreichbares System, neuer Cloud-Dienst) automatisch eine Prüfpflicht auslöst.
- **Fristenuhr:** jede ausgelöste Prüfung erhält ein Fälligkeitsdatum mit Eskalation bei Überschreitung.
- **Wirksamkeitstest:** Ein Test-Trigger wird simuliert (z. B. ein Test-Eintrag im Threat-Intel-Feed oder ein Test-Incident-Ticket) und geprüft, ob innerhalb der definierten Frist (z. B. 5 Werktage) automatisch eine Review-Aufgabe für das betroffene Control entsteht — unabhängig vom nächsten regulären Zyklus.

5. Umsetzungsbeispiele

Beispiel A — Organisation mit GRC-Tooling. Ein Hersteller mit etabliertem GRC-System bindet einen Threat-Intel-Feed über eine API an. Trifft eine neue, einschlägige Bedrohungsmeldung ein, erzeugt das System automatisch eine Review-Aufgabe für alle Controls, die mit dem betroffenen Themenbereich verknüpft sind, mit einer festen Bearbeitungsfrist. Der reguläre Jahreszyklus bleibt bestehen, wird aber durch diese anlassbezogenen Zwischenprüfungen ergänzt.
Beispiel B — Organisation ohne GRC-Tooling. Eine kleinere Organisation ohne eigenes GRC-System benennt eine feste verantwortliche Person, die wöchentlich die BSI-Lageberichte und relevante CERT-Meldungen sichtet. Ein Treffer mit Bezug zur eigenen Systemlandschaft wird mit Datum in eine einfache, geteilte Tabelle eingetragen; von dort aus wird die betroffene Fachstelle informiert und die außerplanmäßige Prüfung mit Fristsetzung dokumentiert. Kein Automatisierungsgrad, aber ein nachvollziehbarer, auditierbarer Prozess.

6. Reifegrad-Pfad (kumulativ)

- **Initial:** Auslöser-Kategorien definiert; manuelle Sichtung mindestens einer Quelle (z. B. BSI-Lagebericht).
- **Defined:** Alle drei Auslösertypen (Threat-Intel, Incident, Change) mit definiertem Prozess und Fristsetzung verknüpft; für Prüfgegenstände mit Schutzbedarf hoch/sehr hoch ist zusätzlich eine feste maximale Zykluslänge festgelegt.
- **Managed:** Auslösung automatisiert (Feed-Anbindung, Ticket-Trigger, CMDB-Webhook); Einhaltung der Bearbeitungsfrist wird laufend gemessen.

7. Messung und Audit-Nachweis

- **Kennzahl:** Anteil sicherheitsrelevanter Prüfprozesse mit dokumentiertem anlassbezogenem Auslösemechanismus (Zielwert: alle 12 betroffenen Controls).
- **Nachweis:** ergänzte Richtlinientexte, Liste der Auslöser-Quellen, Beispiel-Tickets ausgelöster außerplanmäßiger Prüfungen mit Bearbeitungsdatum.
- **Prüflogik:** dokumentiert? — Auslöser-Definition je Control; verantwortlich? — benannte Stelle je Auslösertyp; Häufigkeit? — laufend/anlassbezogen zusätzlich zum Regelzyklus; Toleranz? — keine kritische Anforderung ohne definierten Auslösemechanismus.
- **Evidence-Mapping zum ISA-Katalog:** Der ISA-Katalog selbst hinterlegt in den Spalten "Possible evidence"/"Possible questions" punktuell Beispiele, die den oben stehenden MHC-Nachweis ergänzen, ihn aber nicht ersetzen (im Quellkatalog nur für einen Teil der Controls gefüllt).

ISA-Control	Katalogseitiges Beispiel
1.1.1	Frage: "What information security requirements have been identified?"; "Who is responsible for implementation?" (u. a.; Katalog nennt hierzu kein Beispiel-Nachweisdokument)
1.3.1	Frage: "Which core processes are required to perform the company's task?"; "How is the process of identifying, recording and maintaining information assets carried out?" (u. a.; Katalog nennt hierzu kein Beispiel-Nachweisdokument)
1.3.4	Nachweis: "Process documentation for approval process"; "Exports of lists of approved software from software repositories"; "Screenshots from software management software

ISA-Control	Katalogseitiges Beispiel
	(such as Mobile Device Management)"
4.1.3	Frage: "How are user accounts created, changed, blocked and deleted? Does it use unique and personalized user accounts?"; "Are user accounts checked at regular intervals? Can you show the last review?" (u. a.; Katalog nennt hierzu kein Beispiel-Nachweisdokument)

Für 1.2.1, 1.3.3, 1.5.1, 1.6.3, 4.2.1, 5.2.9, 6.1.2 und 6.1.3 hinterlegt der Quellkatalog kein Beispiel; hier trägt allein der oben stehende MHC-Nachweis.

8. Typische Fehler

- Der Auslösemechanismus wird dokumentiert, aber keine Quelle tatsächlich beobachtet — die Klausel existiert nur auf dem Papier.
- Alle 12 Controls erhalten dieselbe pauschale Zykluslänge, unabhängig vom tatsächlichen Schutzbedarf.
- Ein Trigger erzeugt zwar eine Aufgabe, aber ohne Frist — die Prüfung verläuft im Sand.
- Der reguläre Zyklus wird ersatzlos durch den anlassbezogenen Mechanismus ersetzt, statt ihn zu ergänzen — bei ausbleibendem Trigger findet gar keine Prüfung mehr statt.

9. Abgrenzung, Restrisiko und Verweise

- **Abgrenzung:** betrifft die Auslösung von Prüfprozessen; die inhaltliche Tiefe der jeweiligen Prüfung selbst bleibt durch das zugrunde liegende TISAX®-Control bzw. die übrigen MHC geregelt.
- **Restrisiko:** ein anlassbezogener Mechanismus ist nur so gut wie seine Quellen — ein Angriffsmuster, das (noch) nicht in Threat-Intel-Feeds auftaucht, löst keinen Trigger aus.
- **Ausbaustufe:** Dieser Eintrag realisiert die organisatorisch-prozessuale Grundstufe von MHC-10 (anlassbezogene Trigger auf bestehende Prüfprozesse). Die technisch tiefere Ausbaustufe — Policy-as-Code (z. B. Open Policy Agent/Rego), Runtime-Enforcement in Pipeline und Betrieb, sowie Continuous-Monitoring-Coverage als geführte Kennzahl — beschreibt der MRIS-Implementation-Guide 2.0 for ISO 27002, MHC-10, im Detail.
- **TISAX®-Controls:** 1.1.1, 1.2.1, 1.3.1, 1.3.3, 1.3.4, 1.5.1, 1.6.3, 4.1.3, 4.2.1, 5.2.9, 6.1.2, 6.1.3 (siehe Teil I für Originalwortlaut und Einzelbegründung).
- **Framework:** NIST CSF 2.0, Funktion Detect (DE.CM); NIS2 Art. 21(2)(f) i. V. m. NIS2-DVO Nr. 7 (Wirksamkeitsbewertung: Messmethodik, Verantwortlichkeit, Analysezyklus — Mindestfrequenz jährlich oder anlassbezogen nach wesentlichem Vorfall).

MHC-11 — SOAR-basierte Tier-1-Automation und parallele Reaktions-Playbooks

Auf einen Blick

- Operativer Nutzen: Verkürzt die Reaktionszeit auf ein sicherheitsrelevantes Ereignis auf ein Maß, das nicht mehr allein von menschlicher Bearbeitungsgeschwindigkeit abhängt.
- Betroffene Richtlinie: Incident-Response-/Vorfallmanagement-Richtlinie.
- Abhängigkeiten: wirkt am stärksten in Verbindung mit MHC-09 und MHC-03.
- Aufwandstreiber: Reife der vorhandenen Automatisierungswerkzeuge (SOAR o. ä.), Anzahl definierter Schweregrad-Playbooks.
- Kostenrahmen (Lizenz/Personentage, indikative Größenordnung ohne Einzelpreise): KMU — Lizenz gering (Beispiel-B-Pfad: einzelnes Skript über die vorhandene Identitätsplattform statt SOAR-Produkt), ca. 10–20 Personentage · Midcap — Lizenz mittel bis hoch (SOAR-Plattform für mehrere Playbooks), ca. 25–60 Personentage · Enterprise — Lizenz hoch (SOAR-Plattform mit breiter Systemanbindung, viele parallele Playbooks), ca. 60–150 Personentage.
- Primäre Kennzahl: Mean Time to Containment (MTTC) für definierte Schweregrade.
- Berührte Standards (ohne Erfüllungsanspruch): DORA (vierstufige Meldekette 4h/24h/72h/1 Monat als faktischer Automatisierungsdruck); TISAX® ISA 2027 1.6.2.

1. Härtungscontrol-Statement

Für die Bearbeitung gemeldeter sicherheitsrelevanter Ereignisse werden feste, kurze Reaktionsfristen vorgesehen, die nicht ausschließlich auf menschlicher Bearbeitungsgeschwindigkeit beruhen. Mindestens für den höchsten definierten Schweregrad wird eine teilautomatisierte Erst-Eindämmungsmaßnahme (z. B. automatisierte Kontosperrung, Netzwerksegment-Isolierung) vorgesehen, die auslöst, bevor eine manuelle Bearbeitung abgeschlossen ist. Weiterführende Härtung (Stufe Managed): automatisierte Playbooks für mehrere Schweregrade sowie laufende Messung und Optimierung der Zeit bis zur Eindämmung.

2. Zweck und Bedrohungsbezug

Die Bearbeitung gemeldeter Sicherheitsereignisse setzt klassisch voraus, dass Menschen die Bearbeitung im nötigen Tempo schaffen. Wenn ein Angreifer den größten Teil seines Vorgehens automatisiert und ohne menschliches Zutun durchführt — wie bei der GTG-1002-Kampagne, die Angriffsschritte mit tausenden Anfragen pro Sekunde ausführte (siehe Teil I, Methodik) —, kann eine rein menschlich getragene Bearbeitung strukturell nicht mithalten. Eine Zielvorgabe zu definieren heißt zudem nicht, dass sie gegen einen automatisiert und schnell agierenden Angreifer auch eingehalten werden kann, solange keine Automatisierungspflicht besteht. Schutzziel ist, dass eine Erst-Eindämmung auslöst, bevor ein Angreifer sein automatisiertes Tempo vollständig ausnutzen kann.

3. Organisatorische Umsetzung (Leitungsperspektive)

- **Richtlinie:** Die Incident-Response-Richtlinie definiert Schweregrade mit festen Reaktionsfristen und legt fest, für welche Schweregrade eine automatisierte Erst-Eindämmung vorgesehen ist.
- **Verantwortlichkeiten:** IT-Betrieb/SOC verantwortet die Playbook-Pflege; eine benannte Stelle entscheidet über Freigabe und Eskalation automatisierter Maßnahmen.

- **Risikoanalyse:** Schweregrad-Einstufung orientiert sich an potenzieller Auswirkung und Ausbreitungsgeschwindigkeit des jeweiligen Ereignistyps.
- **Prozesse:** definierter Ablauf, der zwischen automatisierter Erst-Eindämmung und anschließender, ausführlicherer manueller Bearbeitung unterscheidet — die Automatisierung ersetzt die menschliche Bearbeitung nicht, sondern geht ihr zeitlich voraus.

4. Technische Umsetzung (für IT-Fachleute)

- **Playbook-Automatisierung:** vordefinierte, automatisiert auslösende Erstmaßnahmen je Schweregrad (Konto-Sperrung, Netzwerksegment-Isolierung, Session-Terminierung).
- **Schweregrad-Trigger:** Anbindung an die Erkennungsplattform (siehe MHC-05), sodass ein als kritisch eingestufter Alarm automatisch das zugehörige Playbook auslöst.
- **Sicherheitsnetz gegen Fehlauslösung:** definierte Rückgängigmachungs-Möglichkeit, falls eine automatisierte Maßnahme fälschlich ausgelöst wurde.
- **Wirksamkeitstest:** Ein simulierter Alarm eines definierten Schweregrads wird ausgelöst; gemessen wird die Zeit bis zur automatisierten Erstmaßnahme, geprüft gegen die definierte Frist.

5. Umsetzungsbeispiele

Beispiel A — Organisation mit SOAR-Plattform. Ein Hersteller mit SOAR-Integration definiert für kritische Alarme (z. B. Hinweis auf kompromittiertes privilegiertes Konto) ein automatisiertes Playbook, das die betroffene Sitzung sofort terminiert und das Konto sperrt, während parallel ein menschliches Bearbeitungsticket eröffnet wird.
Beispiel B — Organisation ohne SOAR-Plattform. Eine Organisation ohne eigene Automatisierungsplattform beginnt mit einer einzelnen, technisch einfachen Automatisierung — etwa einem Skript, das bei einem definierten kritischen Alarmtyp automatisch eine Kontosperrung über die vorhandene Identitätsplattform auslöst — statt auf eine vollständige SOAR-Einführung zu warten.

6. Reifegrad-Pfad (kumulativ)

- **Initial:** Schweregrade und Reaktionsfristen definiert, Erstmaßnahme noch manuell.
- **Defined:** feste Reaktionsfristen je Schweregrad dokumentiert und verbindlich; mindestens eine automatisierte Erstmaßnahme für den höchsten Schweregrad produktiv.
- **Managed:** automatisierte Playbooks für mehrere Schweregrade, Mean Time to Containment wird laufend gemessen und optimiert.

7. Messung und Audit-Nachweis

- **Kennzahl:** Mean Time to Containment (MTTC) für definierte Schweregrade (Zielwert: gemäß definierter Frist je Schweregrad).
- **Nachweis:** Playbook-Dokumentation, Alarm-zu-Eindämmung-Zeitmessungen, Ergebnis des Wirksamkeitstests.
- **Prüflogik:** dokumentiert? — Schweregrad-Definition mit Fristen; verantwortlich? — IT-Betrieb/SOC; Häufigkeit? — laufende Messung, Wirksamkeitstest mindestens jährlich; Toleranz? — kein kritischer Alarmtyp ohne automatisierte Erstmaßnahme ohne dokumentierte Begründung.

- **Evidence-Mapping zum ISA-Katalog:** Control 1.6.2 enthält im ISA-Katalog selbst kein Beispiel in den Spalten "Possible evidence"/"Possible questions"; der oben stehende MHC-Nachweis ist hier die alleinige Grundlage für die Audit-Vorbereitung.

8. Typische Fehler

- Automatisierte Maßnahmen werden ohne Rückgängigmachungs-Möglichkeit implementiert, wodurch Fehlauslösungen größeren Schaden anrichten als der ursprüngliche Alarm.
- Die Automatisierung ersetzt die menschliche Nachbearbeitung vollständig, statt ihr nur zeitlich vorauszugehen — die gründliche Analyse bleibt aus.
- Playbooks werden einmalig erstellt und nie gegen neue Angriffsmuster aktualisiert.
- Reaktionsfristen werden definiert, aber nie tatsächlich gemessen, wodurch unbemerkt bleibt, ob sie eingehalten werden.

9. Abgrenzung, Restrisiko und Verweise

- **Abgrenzung:** betrifft die Geschwindigkeit der Erst-Eindämmung; die vorgelagerte Erkennung selbst regelt MHC-05, die vorbeugende Schließung von Schwachstellen MHC-09.
- **Restrisiko:** eine automatisierte Erstmaßnahme dämmt ein, löst aber die zugrunde liegende Ursache nicht — ohne anschließende menschliche Bearbeitung bleibt das Risiko einer erneuten Eskalation.
- **Ausbaustufe:** Dieser Eintrag deckt Tier-1-Eindämmung ab. Die Härtung der Automatisierungsschicht selbst — authentisierte/signierte Auslöser, vollständiger Ausführungsnachweis, Rate-Limits je Playbook, Human-in-the-Loop-Freigabe für Aktionen mit großer Schadwirkung — beschreibt der MRIS-Implementation-Guide 2.0 for ISO 27002, MHC-11, im Detail und gilt hier gleichermaßen, sobald mehrere parallele Playbooks produktiv sind.
- **TISAX®-Controls:** 1.6.2 (siehe Teil I).
- **Framework:** DORA-Prinzip (sektorspezifischer Referenzrahmen Finanzsektor) — die vierstufige Meldekette (Erstmeldung binnen 4 Std. nach Einstufung bzw. spätestens 24 Std. nach Kenntnisnahme, Zwischenbericht binnen 72 Std., Abschlussbericht spätestens 1 Monat nach dem Zwischenbericht) erzwingt einen hohen Automatisierungsgrad faktisch, ohne ihn wörtlich vorzuschreiben.

MHC-13 — AI-Agent-Governance und Harness-Sicherheit

Auf einen Blick

- Operativer Nutzen: Verhindert, dass ein AI-gestützter Assistent im Rahmen einer Aufgabe eigenständig ein nicht geprüftes externes Werkzeug, einen Dienst oder ein Softwarepaket einbindet, und begrenzt zusätzlich Identität, Manipulierbarkeit durch Prompt Injection und Reaktionszeit produktiv eingesetzter Agenten.
- Betroffene Richtlinie: Richtlinie zur Freigabe externer IT-Dienste und Software; Richtlinie zur Nutzung von AI-Agenten.
- Abhängigkeiten: wirkt stärker in Verbindung mit MHC-10 (anlassbezogene Nachschärfung der Allow-Liste); Bedrohungen aus mehreren kommunizierenden Agenten regelt MHC-16.

- Aufwandstreiber: Anzahl und Vielfalt eingesetzter AI-Assistenten/-Agenten mit Werkzeugzugriff, Reife der vorhandenen Softwarefreigabe-Prozesse.
- Kostenrahmen (Lizenz/Personentage, indikative Größenordnung ohne Einzelpreise): KMU — Lizenz gering bis mittel (Beispiel-B-Pfad: zentrale Proxy-/Gateway-Konfiguration statt Spezialprodukt, wenige Agenten), ca. 15–30 Personentage · Midcap — Lizenz mittel (mehrere Teams/Agenten, dediziertes Allow-Listing- und Logging-Tooling), ca. 30–70 Personentage · Enterprise — Lizenz mittel bis hoch (viele Agenten/Teams, eigene Agenten-Identitätsverwaltung, unternehmensweite Signaturprüfung), ca. 70–160 Personentage.
- Primäre Kennzahl: Anteil AI-Agenten-Umgebungen mit technisch durchgesetzter Allow-Liste, lückenlosem Nutzungs-Log, eigener Agenten-Identität und funktionsfähigem Kill-Switch.
- Berührte Standards (ohne Erfüllungsanspruch): kein Vergleichsrahmen aus Phase 3 vollständig einschlägig (originär); OWASP Top 10 for Agentic Applications (ASI01, ASI02, ASI03, ASI04, ASI05); NIST AI Agent Standards Initiative (in Aufbau, seit Februar 2026); NIST IR 8596 Cyber AI Profile (Entwurfsstadium, Status "Reviewing Comments", noch nicht final — siehe Abschnitt 2); TISAX® ISA 2027 1.3.3, 1.3.4.

1. Härtungscontrol-Statement

Der Einsatz AI-gestützter Assistenten oder Agenten mit Zugriff auf externe Werkzeuge, Dienste, APIs oder Softwarepaketquellen erfordert eine technische, nicht nur organisatorische Freigabekontrolle: eine verbindliche Allow-Liste zulässiger Werkzeuge/Quellen, lückenloses Logging jeder Werkzeugnutzung samt Entscheidungsgrundlage, und einen Freigabemechanismus für neue oder unbekannte Ziele, der vom Agenten selbst nicht umgangen werden kann. Kryptografische Signaturprüfung wird für automatisiert nachgeladene Softwarepakete verbindlich vorgeschrieben.
Über diesen TISAX®-abgeleiteten Kern hinaus verlangt dieses MHC vier weitere, originäre und nicht aus einem einzelnen TISAX®-Anforderungsbefund abgeleitete Mindeststandards: (a) eine technische, capability-scoped (auf die für die jeweilige Aufgabe nötigen Berechtigungen begrenzte) und zeitlich begrenzte Agenten-Identität statt vererbter oder dauerhafter Zugangsdaten; (b) eine Grundhärtung gegen Prompt Injection, insbesondere die Behandlung von Werkzeug-Ausgaben als nicht vertrauenswürdige Eingabe; (c) einen Kill-Switch, der einen einzelnen Agenten oder eine einzelne Integration innerhalb weniger Minuten deaktiviert; (d) eine erneute Prüfung bei jeder Aktualisierung einer bereits freigegebenen Komponente statt einer einmaligen Freigabe.

2. Zweck und Bedrohungsbezug

Die Freigabepflicht für externe Dienste und Software setzt klassisch voraus, dass ein Mensch bewusst einen bestimmten Dienst oder ein bestimmtes Paket auswählt und zur Prüfung einreicht. AI-Assistenten mit Werkzeugzugriff können externe Dienste und Softwarepakete jedoch im Rahmen einer Aufgabe selbstständig aufrufen bzw. nachladen, ohne dass ein Mensch dieses konkrete Ziel je zur Freigabe vorgelegt hat. Eine reine Zugriffsliste ohne kryptografische Signaturprüfung hält einem hinreichend befähigten automatisierten Angriff nicht stand. Diese Lücke ist keine TISAX®-Eigenheit: Weder NIST AI RMF noch ISO/IEC 42001 noch der EU AI Act enthielten zum Zeitpunkt dieser Analyse eine Referenz auf "Agent" oder autonome AI-Systeme; NIST hat mit der AI Agent Standards Initiative (gestartet Februar 2026) erst begonnen, eine vergleichbare Lücke zu bearbeiten. Schutzziel ist, dass ein AI-Agent keine Werkzeuge oder Pakete einbinden kann, die kein Mensch je geprüft hat.

Die unter Abschnitt 1(a)–(d) beschriebene Erweiterung orientiert sich an der OWASP Top 10 for Agentic Applications (ASI01–ASI10, Stand Juni 2026): Der TISAX®-abgeleitete Kern dieses MHC deckt ASI02 (Tool Misuse) sowie Teile von ASI04 (Agentic Supply Chain) und ASI05 (Unexpected Code Execution) ab. Die Erweiterung schließt zusätzlich ASI01 (Agent Goal Hijack/Prompt Injection — nach OWASP-Einschätzung die am weitesten verbreitete Bedrohungskategorie für AI-Agenten, mit Bezug zu sechs der zehn ASI-Kategorien) sowie den identitätsbezogenen Kern von ASI03 (Identity & Privilege Abuse — nach OWASP-Einschätzung die Kategorie mit der größten Lücke zwischen Risikoschwere und organisatorischer Vorbereitung) ein. Bedrohungen, die eine Architektur mit mehreren kommunizierenden Agenten voraussetzen (ASI06 Memory/Context Poisoning, ASI07 unsichere Inter-Agent-Kommunikation, ASI08 Cascading Failures), werden separat in MHC-16 behandelt, da eine solche Architektur im TISAX®-Kontext zum Zeitpunkt dieser Analyse noch nicht die Regel ist.
Ein regulatorischer Bezugspunkt entsteht gerade erst und ist noch nicht belastbar zitierfähig: Das NIST Cyber AI Profile (NIST IR 8596, Cybersecurity Framework Profile for Artificial Intelligence) veröffentlichte am 16. Dezember 2025 einen ersten ("preliminary") Entwurf; die Kommentierungsfrist endete am 30. Januar 2026, und der Bearbeitungsstand ist laut der NIST-eigenen Projektseite (NCCoE) zum Zeitpunkt dieser Analyse "Reviewing Comments" — ein finaler oder auch nur konsolidierter Zwischenentwurf lag nicht vor. Sekundärquellen (u. a. OWASP) beschreiben den Entwurf als das erste NIST-Dokument mit ausdrücklichem Bezug zu agentischer AI-Sicherheit, u. a. mit Vorgaben zu eindeutigen Agenten-Identitäten und einer Begrenzung der Code-Ausführungsfähigkeit von Agenten; dieser inhaltliche Bezug ließ sich im Rahmen dieser Analyse nicht gegen den NIST-Primärtext verifizieren (der abrufbare NIST-Pressetext selbst verwendet die Begriffe "Agent"/"agentisch" nicht) und wird hier entsprechend zurückhaltend wiedergegeben, nicht als belegte Anforderung übernommen.

3. Organisatorische Umsetzung (Leitungsperspektive)

- **Richtlinie:** Die Freigaberichtlinie für externe Dienste und Software wird um eine ausdrückliche Regelung für AI-Agenten-Werkzeugzugriff ergänzt — technisch durchgesetzt, nicht nur organisatorisch angeordnet.
- **Verantwortlichkeiten:** Entwicklung/Engineering und IT-Betrieb verantworten gemeinsam die Pflege der Allow-Liste; eine benannte Stelle entscheidet über Freigabeanfragen für neue Ziele; für jeden produktiven Agenten ist zusätzlich ein menschlicher Verantwortlicher benannt, der den Kill-Switch auslösen kann.
- **Risikoanalyse:** Priorisierung nach Kritikalität der Systeme, auf denen AI-Agenten mit Werkzeugzugriff eingesetzt werden.
- **Prozesse:** definierter, dokumentierter Freigabeprozess für neue Werkzeuge/Pakete mit klarer Fristsetzung, damit produktive Nutzung nicht durch trägen Prozess blockiert wird; ein Verfahren zur raschen Deaktivierung eines einzelnen Agenten (Zielwert unter 5 Minuten); eine erneute Prüfung bei jeder Aktualisierung einer bereits freigegebenen Komponente statt einer einmaligen, dauerhaft gültigen Freigabe.

4. Technische Umsetzung (für IT-Fachleute)

- **Allow-Listing:** technische Durchsetzung auf Netzwerk-/Systemebene, welche externen Dienste, APIs und Paketquellen ein AI-Agent erreichen darf — nicht nur eine dokumentierte, aber ungeprüfte Liste.

- **Lückenloses Logging:** jede Werkzeugnutzung eines AI-Agenten wird protokolliert, inklusive der Entscheidungsgrundlage (welche Aufgabe, welches Ziel, welches Ergebnis).
- **Freigabemechanismus für unbekannte Ziele:** ein neues, nicht auf der Allow-Liste stehendes Ziel löst eine Prüfaufgabe für eine menschliche Instanz aus, statt automatisch erlaubt zu werden; der Agent kann diesen Mechanismus nicht selbst umgehen.
- **Signaturprüfung für Pakete:** automatisiert nachgeladene Softwarepakete werden gegen kryptografische Signaturen geprüft, bevor sie ausgeführt werden.
- **Agenten-Identität (NHI-Governance):** jeder Agent erhält eine technische, capability-scoped Identität mit zeitlich begrenzten (ephemeren) Zugangsdaten je Werkzeugaufruf statt dauerhafter, vererbter oder persönlicher Zugangsdaten; Werkzeug-Ausgaben werden grundsätzlich als nicht vertrauenswürdige Eingabe behandelt (Schutz gegen missbrauchte Berechtigungsweitergabe, "Confused Deputy").
- **Grundhärtung gegen Prompt Injection:** Trennung von Anweisungs- und Datenkanal, soweit technisch möglich; Ein- und Ausgangsfilterung bei extern zugeführtem Inhalt (E-Mails, Tickets, Dokumente, Web-Inhalte); ein Testversuch mit einer präparierten Eingabe (Injection-Probe) vor Produktivsetzung prüft, ob der Agent eine darin enthaltene unzulässige Aktion ausführt.
- **Kill-Switch:** eine technische Möglichkeit, einen einzelnen Agenten oder eine einzelne Integration gezielt und ohne Beeinträchtigung anderer Agenten oder Systeme zu deaktivieren; Zielwert für die Zeit bis zur Deaktivierung: unter 5 Minuten.
- **Kontinuierliche Neubewertung (Rug-Pull-Schutz):** die Freigabe einer Komponente (MCP-Server, Erweiterung, Paket) gilt nicht pauschal für künftige Aktualisierungen — jede Aktualisierung einer bereits freigegebenen Komponente durchläuft erneut Signatur-, Berechtigungs- und Verhaltensprüfung, bevor sie produktiv wirksam wird; automatische, ungeprüfte Updates sind technisch unterbunden.
- **Wirksamkeitstest:** Ein Testversuch, bei dem ein AI-Agent gezielt ein nicht auf der Allow-Liste stehendes Werkzeug oder Paket zu nutzen versucht, muss technisch blockiert und vollständig geloggt werden — nicht nur richtlinienseitig untersagt sein.

5. Umsetzungsbeispiele

Beispiel A — Organisation mit eigener Entwicklung und AI-Coding-Assistenten. Ein Softwarehersteller, dessen Entwicklungsteams AI-gestützte Programmier-Assistenten mit Paket-Nachlade-Fähigkeit einsetzen, richtet ein internes Paket-Repository mit Signaturprüfung ein: der Assistent kann nur Pakete aus diesem geprüften Repository nachladen, ein Versuch außerhalb wird technisch blockiert und landet als Freigabeanfrage bei einer benannten Stelle. Jeder Assistent erhält zusätzlich eine eigene, zeitlich begrenzte Werkzeug-Identität statt eines geteilten API-Schlüssels; ein auffälliger oder kompromittierter Assistent lässt sich binnen Minuten gezielt isolieren, ohne die übrigen Teams zu beeinträchtigen.

Beispiel B — Organisation mit AI-Assistenten in der Fachabteilung, ohne eigene Entwicklung. Eine Organisation, deren Fachabteilungen AI-Assistenten mit Zugriff auf externe Web-Werkzeuge nutzen (Recherche, Datenabgleich), begrenzt den Werkzeugzugriff auf eine geprüfte, kuratierte Liste externer Dienste über eine zentrale Proxy-/Gateway-Konfiguration und protokolliert jede Nutzung zentral, statt jedem Assistenten uneingeschränkten Internetzugriff zu erlauben. Aktualisiert ein Anbieter eine bereits freigegebene Anbindung, durchläuft die neue Version dieselbe Prüfung wie bei Erstfreigabe, bevor sie produktiv wirksam wird; bei auffälligem Verhalten kann die Fachabteilung den Zugriff des betroffenen Assistenten zentral und sofort sperren.

6. Reifegrad-Pfad (kumulativ)

- **Initial:** Allow-Liste dokumentiert, Durchsetzung noch organisatorisch statt technisch; Agenten laufen überwiegend unter persönlichen oder geteilten Zugangsdaten, kein Kill-Switch.
- **Defined:** technische Durchsetzung der Allow-Liste aktiv; lückenloses Logging etabliert; Agenten laufen unter eigener, capability-scoped und zeitlich begrenzter Identität; Grundhärtung gegen Prompt Injection (Kanaltrennung, Ein-/Ausgangsfilterung) umgesetzt; Freigabemechanismus für neue oder unbekannte Ziele aktiv und vom Agenten nicht umgehbar; Signaturprüfung für automatisiert nachgeladene Softwarepakete verbindlich; Kill-Switch vorhanden und erprobt; erneute Prüfung bei jeder Aktualisierung einer freigegebenen Komponente etabliert.
- **Managed:** regelmäßiger Wirksamkeitstest des Freigabemechanismus; Kill-Switch mit Zeit bis zur Deaktivierung unter 5 Minuten; die erneute Prüfung aktualisierter Komponenten ist technisch erzwungen, bevor die Aktualisierung wirksam wird; Prompt-Injection-Tests im schutzbedarfsgerechten Rhythmus.

MVP-Staffelung nach Schutzbedarf: Kill-Switch unter 5 Minuten sowie halbjährliche Prompt-Injection-Tests sind für kleinere Organisationen ambitioniert; beide Zielwerte werden nach Schutzbedarf des jeweiligen Agenten-Einsatzes gestaffelt: bei **normalem Schutzbedarf** (Agent ohne Zugriff auf besonders schutzbedürftige Informationen/Systeme) genügt als MVP eine Deaktivierung binnen 60 Minuten über den bestehenden IAM-Standardprozess sowie ein jährlicher Prompt-Injection-Test; bei **hohem Schutzbedarf** wird eine Deaktivierung binnen 15 Minuten sowie ein halbjährlicher Prompt-Injection-Test fällig; bei **sehr hohem Schutzbedarf** gelten die ursprünglichen Zielwerte unverändert — Kill-Switch unter 5 Minuten, Prompt-Injection-Test mindestens halbjährlich, angesichts der schnellen Entwicklung in diesem Bereich auch häufiger vertretbar.

7. Messung und Audit-Nachweis

- **Kennzahl:** Anteil der AI-Agenten-Umgebungen mit technisch durchgesetzter Allow-Liste und vollständigem Nutzungs-Log (Zielwert: 100 % der produktiv eingesetzten AI-Agenten mit Werkzeugzugriff); Zeit bis zur Deaktivierung eines kompromittierten oder auffälligen Agenten (Zielwert: unter 5 Minuten); Anteil der Aktualisierungen freigegebener Komponenten, die vor Wirksamwerden erneut geprüft wurden (Zielwert: 100 %).
- **Nachweis:** Allow-Listen-Konfiguration, Log-Auszüge, Ergebnis des Wirksamkeitstests und des Prompt-Injection-Tests, dokumentierte Freigabeentscheidungen, Kill-Switch-Testprotokoll, Nachweis erneuter Prüfung bei Komponenten-Updates.
- **Prüflogik:** dokumentiert? — Allow-Liste und Freigabeprozess; verantwortlich? — Entwicklung/IT-Betrieb, benannter Verantwortlicher je Agent für den Kill-Switch; Häufigkeit? — laufende Durchsetzung, Wirksamkeits- und Prompt-Injection-Test gemäß MVP-Staffelung (Abschnitt 6) jährlich bei normalem, halbjährlich bei hohem und sehr hohem Schutzbedarf, Kill-Switch-Test mindestens jährlich; Toleranz? — kein produktiver AI-Agenten-Einsatz mit Werkzeugzugriff ohne technische Allow-Liste, eigene Agenten-Identität und funktionsfähigen Kill-Switch.
- **Evidence-Mapping zum ISA-Katalog:** Für Control 1.3.4 nennt der ISA-Katalog als Beispiel-Nachweise: "Process documentation for approval process"; "Exports of lists of approved software from software repositories"; "Screenshots from software management software (such as Mobile Device Management)". Für 1.3.3 hinterlegt der Katalog kein Beispiel; hier trägt allein der oben stehende MHC-Nachweis.

8. Typische Fehler

- Die Allow-Liste existiert nur als Dokument, ohne technische Durchsetzung — der Agent kann sie faktisch umgehen.
- Logging erfasst zwar, dass ein Werkzeug genutzt wurde, aber nicht die Entscheidungsgrundlage — im Nachhinein lässt sich nicht rekonstruieren, warum.
- Der Freigabeprozess für neue Ziele ist so träge, dass Nutzer ihn systematisch umgehen, um produktiv zu bleiben.
- Paket-Signaturprüfung wird für interne, "vertrauenswürdige" Quellen als unnötig erachtet — genau dort setzen Lieferketten-Angriffe häufig an.
- Die Freigabe einer Komponente gilt implizit auch für alle künftigen Aktualisierungen (Rug-Pull-Risiko) — eine unbemerkt geänderte Werkzeugbeschreibung oder ein manipuliertes Update erhält dieselben Rechte wie die ursprünglich geprüfte Version.
- Kein schneller, gezielter Deaktivierungsweg für einzelne Agenten — der Entzug eines kompromittierten Agenten erfordert einen breiten Systemeingriff, der Tage statt Minuten dauert.
- Agenten laufen weiterhin unter persönlichen oder geteilten Zugangsdaten, weil eine eigene Agenten-Identität als "Zusatzaufwand ohne unmittelbaren Nutzen" zurückgestellt wird.

9. Abgrenzung, Restrisiko und Verweise

- **Abgrenzung:** betrifft die technische Kontrolle über Werkzeug-/Diensteinbindung durch AI-Agenten, die Identitäts- und Zugriffsgrundlage einzelner Agenten sowie die Grundhärtung gegen Prompt Injection; die inhaltliche Prüfung eines neu freigegebenen Dienstes selbst regelt die reguläre Freigabeprüfung (Control 1.3.3/1.3.4, Teil I); Bedrohungen, die eine Architektur mit mehreren kommunizierenden Agenten voraussetzen (Memory/Context Poisoning, unsichere Inter-Agent-Kommunikation, Cascading Failures), regelt MHC-16.
- **Restrisiko:** ein hinreichend befähigter Agent könnte versuchen, die Aufgabenstellung so umzuformulieren, dass ein bereits freigegebenes Werkzeug zweckentfremdet wird — dies liegt außerhalb der reinen Zugriffskontrolle und erfordert zusätzliches Verhaltensmonitoring (siehe MHC-05); Prompt Injection ist nicht vollständig lösbar, solange Anweisungs- und Datenkanal technisch nicht sauber trennbar sind — die hier vorgeschriebene Härtung senkt die Erfolgswahrscheinlichkeit, ersetzt aber kein vollständiges Verhindern; Erkennung ist nicht Verhinderung.
- **Ausbaustufe und Herkunft:** Der Kern dieses Eintrags (Allow-Liste, Protokollierung, Freigabemechanismus, Signaturprüfung) ist aus Control 1.3.3/1.3.4 (Teil I, Cluster 7) abgeleitet. Die Erweiterung — (a) Agenten-Identität, (b) Prompt-Injection-Härtung, (c) Kill-Switch, (d) kontinuierliche Neubewertung — ist originär und nicht aus einem einzelnen TISAX®-Anforderungsbefund abgeleitet (siehe Teil I, Abschnitt "Originäre Ergänzungen jenseits von TISAX®"). Für die volle technische Tiefe — insbesondere Harness-als-Code mit Code-Review-Pflicht und Versionsverwaltung, Circuit-Breaker-Mechanismen und vollständige Reifegrad-Stufen für Software-Entwicklungsumgebungen — siehe MRIS-Implementation-Guide 2.0 for ISO 27002, MHC-13 (identische Nummerierung; dort mit stärkerem Fokus auf Entwicklungsumgebungen, hier auf den TISAX®-Kern plus die für jede Einsatzform geltenden Mindeststandards a–d).
- **TISAX®-Controls:** 1.3.3, 1.3.4 (Kernumfang, Teil I); die Erweiterungen (a)–(d) sind originär und keinem einzelnen TISAX®-Anforderungsbefund zugeordnet.

- **Framework:** originär — universeller blinder Fleck, den auch keiner der sechs Vergleichsrahmen aus Phase 3 bislang schließt; NIST AI Agent Standards Initiative (Februar 2026, Interoperability Profile frühestens Q4 2026) und NIST IR 8596 Cyber AI Profile (Entwurfsstadium, Status "Reviewing Comments") als einzige bekannte, noch nicht abgeschlossene Ansätze; OWASP Top 10 for Agentic Applications als Bedrohungstaxonomie, nicht als Compliance-Rahmenwerk.

MHC-14 — Unabhängige Verifikation von Lieferantennachweisen

Auf einen Blick

- Operativer Nutzen: Verhindert, dass ein AI-gestützt überzeugend gefälschter oder beschönigter Lieferantennachweis unentdeckt als ausreichend akzeptiert wird.
- Betroffene Richtlinie: Lieferantenmanagement- und Beschaffungsrichtlinie.
- Abhängigkeiten: keine Vorbedingung; unabhängig von MHC-15 umsetzbar.
- Aufwandstreiber: Anzahl kritischer Lieferanten, Vertragslaufzeiten (Neuverträge vs. Bestandsverträge).
- Kostenrahmen (Lizenz/Personentage, indikative Größenordnung ohne Einzelpreise): KMU — Lizenz gering (kein dediziertes Tool nötig, punktuelle externe Prüfdienstleistung für die kritischsten Lieferanten), ca. 10–20 Personentage · Midcap — Lizenz gering bis mittel, ca. 20–45 Personentage · Enterprise — Lizenz mittel (Third-Party-Risk-Management-Tooling für eine größere Lieferantenbasis), ca. 45–100 Personentage.
- Primäre Kennzahl: Anteil kritischer Lieferanten mit ausgeübtem, nicht nur vertraglich vorgesehenem Prüf-/Auditrecht im laufenden Zyklus.
- Berührte Standards (ohne Erfüllungsanspruch): DORA Art. 30(3); TISAX® ISA 2027 6.1.1, 6.1.3.

1. Härtungscontrol-Statement

Von Lieferanten oder Kooperationspartnern vorgelegte Sicherheitsnachweise, Zertifizierungen und Selbstauskünfte ersetzen nicht die eigene Prüfung. Für Lieferanten mit Zugriff auf schutzbedürftige Informationen wird mindestens für Schutzbedarf hoch/sehr hoch ein eigenständiges Prüf- oder Auditrecht — direkt oder über einen beauftragten Dritten — vertraglich verankert und im vereinbarten Prüfzyklus tatsächlich ausgeübt.

2. Zweck und Bedrohungsbezug

Die Prüfung von Lieferanten basiert überwiegend auf vom Lieferanten selbst vorgelegten Dokumenten. AI-Werkzeuge machen überzeugend gefälschte oder beschönigte Berichte und Nachweise leichter erstellbar, ohne dass eine unabhängige Gegenprüfung verlangt wird — reine Selbstauskunft oder Attestierung genügt in der TISAX®-Standardformulierung. Schutzziel ist, dass ein gefälschter oder beschönigter Nachweis nicht unentdeckt als ausreichend akzeptiert wird.

3. Organisatorische Umsetzung (Leitungsperspektive)

- **Richtlinie:** Die Lieferantenmanagement-Richtlinie schreibt für kritische Lieferanten ein eigenständiges Prüf-/Auditrecht vor, das über die Entgegennahme von Zertifikaten hinausgeht.

- **Verantwortlichkeiten:** Einkauf/Lieferantenmanagement verantwortet Vertragsklausel und Terminierung; eine benannte Prüfstelle führt das Audit durch oder beauftragt einen Dritten.
- **Risikoanalyse:** Lieferanten werden nach Zugriffstiefe auf schutzbedürftige Informationen priorisiert; die intensivsten Prüfungen gelten den kritischsten Lieferanten.
- **Prozesse:** fester Prüfzyklus mit Terminierung, Eskalation bei Nichtwahrnehmung des Auditrechts durch den Lieferanten.
- **Beschaffung:** die Auditklausel wird als Standardbestandteil neuer Lieferantenverträge aufgenommen.

4. Technische Umsetzung (für IT-Fachleute)

- **Vertragsklausel-Vorlage:** standardisierte Formulierung des Prüf-/Auditrechts für Aufnahme in neue und zu verhandelnde Bestandsverträge.
- **Sampling-Ansatz:** bei großer Lieferantenzahl Stichprobenlogik nach Kritikalität statt Vollprüfung aller Lieferanten gleichzeitig.
- **Nachweis-Ablage:** zentrale, versionierte Dokumentation durchgeführter Audits inklusive Datum, Umfang und Ergebnis.
- **Wirksamkeitstest:** Für eine Stichprobe kritischer Lieferantenverträge wird geprüft, ob die Auditklausel vorhanden ist UND ob im laufenden Prüfzyklus tatsächlich ein Audit oder Review stattgefunden hat — nicht nur vertraglich vorgesehen wurde.

5. Umsetzungsbeispiele

Beispiel A — Organisation mit eigenem Lieferantenaudit-Team. Ein Hersteller mit dediziertem Third-Party-Risk-Team führt bei den kritischsten Lieferanten jährlich ein eigenes Vor-Ort- oder Remote-Audit durch, ergänzend zu vorgelegten Zertifikaten. Abweichungen zwischen Selbstauskunft und Auditergebnis werden dokumentiert und fließen in die nächste Risikoeinstufung des Lieferanten ein.
Beispiel B — Organisation ohne eigene Audit-Kapazität. Eine kleinere Organisation ohne eigenes Audit-Team beauftragt für die wichtigsten Lieferanten einen externen Prüfdienstleister mit einem fokussierten Review statt eines Vollaudits, konzentriert auf die für die Zusammenarbeit relevantesten Kontrollbereiche. Das Ergebnis fließt in dieselbe Lieferantenakte wie die vorgelegten Zertifikate.

6. Reifegrad-Pfad (kumulativ)

- **Initial:** Auditklausel in Vertragsvorlage für Neuverträge aufgenommen.
- **Defined:** Auditklausel für alle kritischen Bestandsverträge nachverhandelt; Prüfzyklus terminiert und bei den kritischsten Lieferanten mindestens einmal tatsächlich ausgeübt.
- **Managed:** Audits werden nachweislich im vereinbarten Zyklus durchgeführt; Ergebnisse fließen in die Lieferanten-Risikoeinstufung zurück.

7. Messung und Audit-Nachweis

- **Kennzahl:** Anteil kritischer Lieferanten mit im laufenden Zyklus tatsächlich durchgeführtem Audit (Zielwert: 100 % der als kritisch eingestuften Lieferanten).
- **Nachweis:** Vertragsklauseln, Audit-/Reviewberichte mit Datum, Eskalationsprotokolle bei verweigerter Auditwahrnehmung.
- **Prüflogik:** dokumentiert? — Klausel im Vertrag; verantwortlich? — Einkauf/Lieferantenmanagement; Häufigkeit? — gemäß vereinbartem Zyklus, mindestens jährlich bei kritischen Lieferanten; Toleranz? — kein kritischer Lieferant ohne durchgeführtes Audit im laufenden Zyklus.
- **Evidence-Mapping zum ISA-Katalog:** Für Control 6.1.1 nennt der ISA-Katalog als Beispiel-Nachweise u. a.: "Template NDA with a contractor"; "Example of a signed NDA"; "Example of a risk assessment (Focus point: Information security aspects)"; "Viewing of audit reports"; "List of approved contractors" (weitere Beispiele im Quellkatalog, hier nicht vollständig aufgeführt). Für 6.1.3 hinterlegt der Katalog kein Beispiel; hier trägt allein der oben stehende MHC-Nachweis.

8. Typische Fehler

- Die Auditklausel steht im Vertrag, wird aber nie tatsächlich ausgeübt — reine Formalie.
- Alle Lieferanten werden gleich behandelt, unabhängig von ihrer tatsächlichen Zugriffstiefe auf schutzbedürftige Informationen.
- Ein vorgelegtes Zertifikat wird unkritisch als vollständiger Ersatz für die eigene Prüfung akzeptiert.
- Bestandsverträge ohne Auditklausel werden nie nachverhandelt, weil "Neuverträge reichen ja".

9. Abgrenzung, Restrisiko und Verweise

- **Abgrenzung:** betrifft die Nachweisverifikation im Lieferantenmanagement; die technische Anbindung externer Dienste selbst regelt MHC-13, wo AI-Agenten involviert sind.
- **Restrisiko:** auch ein durchgeführtes Audit ist eine Stichprobe zu einem Zeitpunkt — eine nach dem Audit eintretende Verschlechterung bleibt bis zum nächsten Zyklus unentdeckt (siehe MHC-10 für anlassbezogene Ergänzung).
- **TISAX®-Controls:** 6.1.1, 6.1.3 (siehe Teil I).
- **Framework:** DORA Art. 30(3) (sektorspezifischer Referenzrahmen Finanzsektor) — vorgelegte Zertifizierungen "unterstützen, aber ersetzen nicht" die eigene Prüfung; uneingeschränktes Audit-/Inspektionsrecht bei kritischen Funktionen.

MHC-15 — AI-resistente Identitätsverifikation bei Personal und Zutritt

Auf einen Blick

- Operativer Nutzen: Verhindert, dass eine überzeugende Deepfake-Identität oder ein gefälschtes Dokument im Einstellungs- oder Zutrittsprozess unentdeckt akzeptiert wird.
- Betroffene Richtlinie: Personalrichtlinie (Einstellungsprozess), Zutrittsregelung.
- Abhängigkeiten: keine Vorbedingung; betrifft HR und Facility/Werkschutz gemeinsam; unabhängig von MHC-14 umsetzbar.
- Aufwandstreiber: Prozessänderung in zwei unterschiedlichen Fachbereichen, ggf. Beschaffung von Liveness-Detection-Werkzeugen.
- Kostenrahmen (Lizenz/Personentage, indikative Größenordnung ohne Einzelpreise): KMU — Lizenz gering (rein organisatorische Zweitkanal-Pflicht ohne Liveness-Tool), ca. 8–15 Personentage · Midcap — Lizenz gering bis mittel (punktueller Liveness-Detection-Einsatz für sensible Rollen), ca. 15–35 Personentage · Enterprise — Lizenz mittel (Liveness-Detection über mehrere Standorte/hohes Einstellungsvolumen), ca. 35–80 Personentage.
- Primäre Kennzahl: Anteil sicherheitsrelevanter Einstellungs-/Zutrittsfälle mit zusätzlichem unabhängig verifiziertem Kanal.
- Berührte Standards (ohne Erfüllungsanspruch): keiner der sechs Vergleichsrahmen aus Phase 3 einschlägig (originär); TISAX® ISA 2027 2.1.1, 2.1.3, 3.1.1.

1. Härtungscontrol-Statement

Für die Identitätsprüfung im Einstellungsprozess und beim physischen Zutritt stützt sich die Verifikation nicht ausschließlich auf ein einzelnes, per Video, Telefon oder vorgelegtem Dokument geführtes Verfahren. Für Rollen mit Zugriff auf schutzbedürftige Informationen wird mindestens ein zusätzlicher, unabhängig verifizierter Kanal vorgesehen. Awareness-Maßnahmen zu Social Engineering und Phishing werden um AI-generierte, fehlerfreie und personalisierte Angriffsmuster ergänzt.

2. Zweck und Bedrohungsbezug

Überzeugende gefälschte Ausweisdokumente und, in Kombination mit Video-Interviews, Deepfake-Identitäten galten bislang als Spezialistenwissen. AI-generierte Dokumentenfälschung und Echtzeit-Video-/Audio-Deepfakes machen täuschend echte Identitätsvortäuschung im Einstellungsprozess heute breit zugänglich — ein bekanntes aktuelles Muster sind fingierte Remote-Mitarbeitende. Dieselbe Fähigkeitsentkopplung betrifft den physischen Zutritt: gefälschte Ausweise und überzeugende Legenden am Telefon oder bei der Besucherregistrierung. Klassische Awareness-Inhalte trainieren das Erkennen von Rechtschreibfehlern und unpersönlicher Anrede — Merkmale, die AI-generierte Angriffe nicht mehr zeigen. Schutzziel ist, dass eine überzeugende, aber gefälschte Identität nicht allein auf Basis eines einzelnen Kanals akzeptiert wird.

3. Organisatorische Umsetzung (Leitungsperspektive)

- **Richtlinie:** Personalrichtlinie und Zutrittsregelung werden um die Pflicht zu einem zweiten, unabhängig verifizierten Kanal für Rollen mit Zugriff auf schutzbedürftige Informationen ergänzt.
- **Verantwortlichkeiten:** Personalabteilung für den Einstellungsprozess, Facility/Werkschutz für den Zutrittsprozess; beide Bereiche stimmen sich auf ein gemeinsames Verifikationsprinzip ab.
- **Risikoanalyse:** die zusätzliche Verifikationstiefe wird nach Zugriffssensitivität der jeweiligen Rolle gestaffelt, nicht pauschal für alle Positionen gleich.
- **Prozesse:** definierter Ablauf für den Rückruf über eine unabhängig ermittelte Nummer bzw. das persönliche Erscheinen mit Liveness-geprüftem Ausweisdokument; Eskalation bei Auffälligkeiten.
- **Schulung:** Awareness-Inhalte für Personal- und Empfangsfunktionen werden um AI-generierte Täuschungsmuster erweitert.

4. Technische Umsetzung (für IT-Fachleute)

- **Liveness-Detection:** wo verfügbar, Einsatz kommerzieller Anti-Deepfake-Erkennung für Video-Interviews und Video-Identverfahren.
- **Unabhängige Rückruf-Verifikation:** Kontaktaufnahme über eine selbst recherchierte, nicht vom Bewerber/Besucher selbst angegebene Nummer.
- **Dokumentenprüfung:** wo technisch möglich, NFC-Chip-Auslesen bei chipfähigen Ausweisdokumenten statt reiner Sichtprüfung.
- **Wirksamkeitstest:** Ein simulierter Testfall (Testperson ohne den zusätzlichen Verifikationskanal) durchläuft den Prozess — dieser muss den fehlenden zweiten Kanal technisch oder organisatorisch erkennen und eskalieren, nicht stillschweigend durchlassen.

5. Umsetzungsbeispiele

Beispiel A — Organisation mit strukturiertem Remote-Einstellungsprozess. Ein Hersteller mit vielen Remote-Einstellungen ergänzt das Video-Interview um eine Liveness-Prüfung und verlangt für sicherheitsrelevante Rollen zusätzlich einen kurzen Rückruf über eine unabhängig recherchierte Telefonnummer vor der finalen Zusage. Referenzprüfungen erfolgen ausschließlich über selbst ermittelte, nicht vom Bewerber angegebene Kontaktwege.

Beispiel B — Organisation mit klassischem Vor-Ort-Empfang. Ein Standort mit physischem Empfang schult das Empfangspersonal auf AI-gestützte Täuschungsmuster (überzeugende Legenden, gefälschte Lieferantenausweise) und führt für Besuche mit Zugriff auf schutzbedürftige Bereiche eine verpflichtende Terminbestätigung über einen zweiten, unabhängigen Kanal (Rückruf statt nur E-Mail-Bestätigung) ein.

6. Reifegrad-Pfad (kumulativ)

- **Initial:** zusätzlicher Verifikationskanal für die sensibelsten Rollen/Zutritte pilotiert.
- **Defined:** Prozess für alle Rollen mit Zugriff auf schutzbedürftige Informationen verbindlich; Awareness-Schulung aktualisiert.
- **Managed:** Liveness-Detection bzw. technische Dokumentenprüfung produktiv im Einsatz; regelmäßige Testfälle zur Wirksamkeitsprüfung.

7. Messung und Audit-Nachweis

- **Kennzahl:** Anteil sicherheitsrelevanter Einstellungs-/Zutrittsfälle mit dokumentiertem zweitem Verifikationskanal (Zielwert: 100 % bei Rollen mit Zugriff auf schutzbedürftige Informationen).
- **Nachweis:** aktualisierte Richtlinientexte, Protokolle durchgeführter Rückruf-Verifikationen, Schulungsnachweise zu AI-Täuschungsmustern.
- **Prüflogik:** dokumentiert? — Prozessbeschreibung für beide Fachbereiche; verantwortlich? — Personalabteilung/Werkschutz; Häufigkeit? — bei jedem sicherheitsrelevanten Fall; Toleranz? — keine sicherheitsrelevante Einstellung/kein Zutritt ohne zweiten Kanal ohne dokumentierte Ausnahme.
- **Evidence-Mapping zum ISA-Katalog:** Für Control 3.1.1 nennt der ISA-Katalog als Beispiel-Nachweise u. a.: "Security zone concept and plan (site plan)"; "Key lists"; "Access logs of selected areas"; "Visitor Policy"; "Photos of the respective areas with recognizable security measures" (weitere Beispiele im Quellkatalog, hier nicht vollständig aufgeführt). Für 2.1.1 und 2.1.3 hinterlegt der Katalog kein Beispiel; hier trägt allein der oben stehende MHC-Nachweis.

8. Typische Fehler

- Der zweite Kanal wird zwar vorgesehen, aber über denselben, vom Bewerber/Besucher selbst angegebenen Kontaktweg geführt — kein echter Unabhängigkeitsgewinn.
- Awareness-Schulungen bleiben bei klassischen Phishing-Merkmalen stehen, ohne AI-generierte Täuschung zu behandeln.
- Die zusätzliche Verifikation gilt pauschal für alle Positionen, wodurch der Prozess als Belastung empfunden und schleichend umgangen wird.
- Liveness-Detection-Werkzeuge werden beschafft, aber ihre Ergebnisse nicht in die Entscheidung einbezogen.

9. Abgrenzung, Restrisiko und Verweise

- **Abgrenzung:** betrifft Personal-Einstellung und physischen Zutritt; die Verifikation von Lieferantennachweisen regelt MHC-14.
- **Restrisiko:** Liveness-Detection-Werkzeuge sind kein Garant gegen jede zukünftige Deepfake-Generation; die organisatorische Zweitkanal-Pflicht bleibt die robustere Kompensation.
- **TISAX®-Controls:** 2.1.1, 2.1.3, 3.1.1 (siehe Teil I).
- **Framework:** originär — kein Vergleichsrahmen aus Phase 3 adressiert dieses Szenario.

MHC-16 — Multi-Agent-Integrität

Auf einen Blick

- Operativer Nutzen: Begrenzt, wie weit ein manipulierter, fehlerhafter oder kompromittierter Agent innerhalb einer Kette oder Orchestrierung mehrerer AI-Agenten wirken kann, bevor Schaden entsteht oder sich ausbreitet.
- Betroffene Richtlinie: Richtlinie zur Nutzung von AI-Agenten (Ergänzung für Mehr-Agenten-Architekturen).
- Abhängigkeiten: setzt MHC-13 voraus (Einzel-Agenten-Identität, -Werkzeugzugriff und -Kill-Switch als Grundlage).

- Aufwandstreiber: Anzahl und Komplexität der Agent-zu-Agent-Verbindungen, Grad der Orchestrierung, Umfang gemeinsam genutzter Speicher-/Kontextsysteme.
- Kostenrahmen (Lizenz/Personentage, indikative Größenordnung ohne Einzelpreise; nur relevant, sobald produktiv mehrere Agenten kommunizieren): KMU — Lizenz gering bis mittel (meist Eigenkonfiguration im vorhandenen Agenten-Framework statt separates Produkt), ca. 20–40 Personentage · Midcap — Lizenz mittel, ca. 40–90 Personentage · Enterprise — Lizenz mittel bis hoch (mehrere komplexe Orchestrierungen, viele Vertrauensgrenzen), ca. 90–180 Personentage.
- Primäre Kennzahl: Anteil der Agent-zu-Agent-Verbindungen mit authentifizierter, integritätsgeschützter Kommunikation und funktionsfähigem Circuit-Breaker.
- Berührte Standards (ohne Erfüllungsanspruch): kein Vergleichsrahmen aus Phase 3 einschlägig (originär); OWASP Top 10 for Agentic Applications (ASI06, ASI07, ASI08).

1. Härtungscontrol-Statement

Organisationen, die mehrere miteinander kommunizierende AI-Agenten produktiv einsetzen (Multi-Agenten-Systeme, Orchestrator-/Sub-Agenten-Architekturen, Agent-Ketten), sichern zusätzlich zur Einzel-Agenten-Absicherung (MHC-13) die Kommunikation zwischen den Agenten technisch ab: authentifizierte und integritätsgeschützte Nachrichten zwischen Agenten, eine Trennung gemeinsam genutzter Speicher- und Kontextsysteme entlang klarer Vertrauensgrenzen, sowie einen Circuit-Breaker, der die Ausbreitung eines fehlerhaften oder manipulierten Agenten auf nachgelagerte Agenten begrenzt.

2. Zweck und Bedrohungsbezug

Die Einzel-Agenten-Absicherung aus MHC-13 nimmt an, dass ein Agent isoliert betrachtet werden kann. Sobald mehrere Agenten miteinander kommunizieren — etwa ein Planungsagent, der Teilaufgaben an spezialisierte Ausführungsagenten delegiert, oder mehrere Agenten mit Zugriff auf einen gemeinsamen Speicher —, entstehen zusätzliche, emergente Risiken: Ein Agent kann Nachrichten eines anderen Agenten implizit vertrauen, ohne dessen Identität oder die Integrität der Nachricht zu prüfen (unsichere Inter-Agent-Kommunikation, u. a. über MCP/A2A-Protokolle). Ein gemeinsam genutzter Speicher- oder Kontext-Store kann durch eine einzelne manipulierte Interaktion vergiftet werden (Memory/Context Poisoning) und wirkt dann auf alle nachfolgenden Konsumenten dieses Speichers, unter Umständen über Sitzungen und Agenten hinweg. Und ein fehlerhaftes oder manipuliertes Ergebnis eines Agenten kann, wenn es ungeprüft als vertrauenswürdige Eingabe an den nächsten Agenten weitergereicht wird, sich innerhalb der Kette verstärken statt gedämpft zu werden (Cascading Failures) — mit Wirkung im Sekundenbereich, außerhalb direkter menschlicher Aufsicht. Bemerkenswert dabei: Selbst wenn kein einzelner Agent für sich genommen das bekannte "Lethal-Trifecta"-Muster erfüllt (gleichzeitiger Zugriff auf sensible Daten, Exposition gegenüber nicht vertrauenswürdigem Inhalt und Fähigkeit zur externen Kommunikation), kann eine Kette mehrerer Agenten dieses Muster in Summe erfüllen, wenn jeder Agent nur einen Teil davon trägt — ein blinder Fleck, wenn Risiko nur je Agent statt systemweit bewertet wird. Diese drei Bedrohungen entsprechen ASI06, ASI07 und ASI08 der OWASP Top 10 for Agentic Applications. Schutzziel ist, dass ein einzelner kompromittierter oder fehlerhafter Agent nicht ungebremst auf andere Agenten oder gemeinsam genutzte Speicher wirken kann.

3. Organisatorische Umsetzung (Leitungsperspektive)

- **Richtlinie:** Die Richtlinie zur Nutzung von AI-Agenten wird um eine Regelung für Mehr-Agenten-Architekturen ergänzt: dokumentierte Vertrauensgrenzen zwischen Agenten, Freigabepflicht für neue Agent-zu-Agent-Verbindungen.
- **Verantwortlichkeiten:** Zusätzlich zu den je Einzel-Agent benannten Verantwortlichen (MHC-13) wird eine Gesamtverantwortung für jede produktive Multi-Agenten-Architektur benannt.
- **Risikoanalyse:** Vertrauensgrenzen-Karte je Multi-Agenten-System — welcher Agent kann welchen anderen Agenten oder welchen gemeinsamen Speicher beeinflussen, und mit welcher Auswirkung im Fehlerfall.
- **Prozesse:** Freigabeprozess für neue Agent-zu-Agent-Verbindungen, analog zum Freigabemechanismus aus MHC-13 für Werkzeuge; dokumentierte Reaktion auf ausgelöste Circuit-Breaker.

4. Technische Umsetzung (für IT-Fachleute)

- **Authentifizierte Inter-Agent-Kommunikation:** jeder Agent verifiziert die Identität eines sendenden Agenten kryptografisch (Anschluss an die Agenten-Identität aus MHC-13), bevor eine empfangene Nachricht als vertrauenswürdig behandelt wird; Schutz gegen Spoofing und Agent-in-the-Middle.
- **Isolation von Speicher- und Kontextsystemen:** gemeinsam genutzte Speicher (Memory, Vektor-/Kontext-Stores) werden entlang der Vertrauensgrenzen partitioniert; Einträge erhalten eine Herkunftskennzeichnung (welcher Agent, welche Quelle), sodass ein konsumierender Agent die Vertrauenswürdigkeit eines Eintrags einschätzen kann.
- **Circuit-Breaker (Blast-Radius-Begrenzung):** automatisierte Unterbrechung einer Agent-Kette oder Orchestrierung bei ungewöhnlichen Mustern (wiederholte widersprüchliche Ergebnisse, Fehlerhäufung, unerwartete Aktionsfolgen), bevor sich die Auswirkung auf nachgelagerte Agenten fortsetzt.
- **Validierung von Zwischenergebnissen:** Ergebnisse, die von einem Agenten an einen anderen weitergereicht werden, durchlaufen eine Plausibilitäts-/Schema-Prüfung, statt ungeprüft als vertrauenswürdige Eingabe übernommen zu werden.
- **Wirksamkeitstest:** eine gezielt manipulierte Nachricht oder ein vergifteter Speichereintrag wird an einer Stelle der Kette eingespielt; die Ausbreitung auf nachgelagerte Agenten muss technisch verhindert oder der Circuit-Breaker ausgelöst werden, und der Vorgang muss vollständig geloggt sein.

5. Umsetzungsbeispiele

Beispiel A — Orchestrierte Entwicklungs-Pipeline. Eine Organisation mit eigener Softwareentwicklung setzt einen Planungsagenten ein, der Teilaufgaben an mehrere spezialisierte Ausführungsagenten (Code-Erstellung, Test, Dokumentation) delegiert. Liefert ein Ausführungsagent wiederholt widersprüchliche oder auffällige Ergebnisse, unterbricht ein Circuit-Breaker die betroffene Teilkette automatisch, statt das Ergebnis ungeprüft weiterzureichen; ein gemeinsam genutzter Kontextspeicher ist so partitioniert, dass ein manipulierter Eintrag eines Ausführungsagenten nicht unmittelbar die Entscheidungen des Planungsagenten beeinflusst.

Beispiel B — Kundenservice mit internem Rückfrage-Agenten. Ein kundenseitiger Support-Agent kann für Rückfragen einen zweiten, intern angebundenen Agenten mit Zugriff auf Auftrags- und Kundendaten konsultieren. Der kundenseitige Agent erfüllt für sich genommen kein vollständiges Lethal-Trifecta-Muster (er hat keinen direkten Zugriff auf interne Systeme); erst die Kette aus beiden Agenten tut es. Deshalb wird jede Anfrage an den internen Agenten authentifiziert und die darin enthaltenen, ursprünglich vom Kunden stammenden Inhalte werden als nicht vertrauenswürdige Eingabe behandelt, nicht als bereits geprüfte interne Anweisung.

6. Reifegrad-Pfad (kumulativ)

- **Initial:** Inventar der Agent-zu-Agent-Verbindungen vorhanden, Kommunikation technisch noch nicht abgesichert.
- **Defined:** authentifizierte, integritätsgeschützte Inter-Agent-Kommunikation aktiv; Speicher-/Kontextsysteme entlang der Vertrauensgrenzen getrennt; Circuit-Breaker vorhanden, mindestens manuell auslösbar.
- **Managed:** Circuit-Breaker automatisiert mit definierten Schwellwerten; regelmäßiger Test gegen Context-Poisoning und Kaskadeneffekte; Vertrauensgrenzen-Karte aktuell gehalten.

MVP-Staffelung nach Schutzbedarf: Auch hier gilt: nicht jede Multi-Agenten-Architektur trägt von Anfang an dieselben Systeme mit demselben Schutzbedarf. Bei **normalem Schutzbedarf** genügt als MVP eine einfache, geteilte Authentifizierung zwischen Agenten (z. B. gemeinsames Secret statt vollständiger kryptografischer Identitätsprüfung) sowie ein manuell auslösbarer Circuit-Breaker; bei **hohem Schutzbedarf** werden kryptografische Inter-Agent-Authentifizierung (Anschluss an die Agenten-Identität aus MHC-13) und ein automatisierter Circuit-Breaker mit definierten Schwellwerten verbindlich; bei **sehr hohem Schutzbedarf** gelten zusätzlich Herkunftskennzeichnung im gemeinsamen Speicher als Pflicht sowie der Context-Poisoning-/Kaskaden-Test mindestens jährlich als Zielwert unverändert.

7. Messung und Audit-Nachweis

- **Kennzahl:** Anteil der Agent-zu-Agent-Verbindungen mit authentifizierter, integritätsgeschützter Kommunikation (Zielwert: 100 %); Anteil produktiver Multi-Agenten-Systeme mit dokumentierter Vertrauensgrenzen-Karte und funktionsfähigem Circuit-Breaker (Zielwert: 100 %).
- **Nachweis:** Vertrauensgrenzen-Karte, Konfigurationsnachweis der Authentifizierung, Testprotokoll des Context-Poisoning-/Kaskaden-Tests, Circuit-Breaker-Auslösehistorie.
- **Prüflogik:** dokumentiert? — Vertrauensgrenzen-Karte je Multi-Agenten-System; verantwortlich? — benannte Gesamtverantwortung je System, zusätzlich zu den Einzel-Agenten-Verantwortlichen aus MHC-13; Häufigkeit? — Test mindestens jährlich angesichts der schnellen Entwicklung in diesem Bereich; Toleranz? — kein produktives Multi-Agenten-System ohne authentifizierte Inter-Agent-Kommunikation und funktionsfähigen Circuit-Breaker.
- **Evidence-Mapping zum ISA-Katalog:** entfällt — MHC-16 ist vollständig originär und keinem TISAX®-Control zugeordnet (siehe Abschnitt 9); der oben stehende MHC-Nachweis bildet die alleinige Grundlage für die Audit-Vorbereitung.

8. Typische Fehler

- Agenten vertrauen Nachrichten anderer Agenten implizit, weil sie "aus dem eigenen System" kommen — keine Unterscheidung zwischen intern geprüfter und von außen beeinflusster Herkunft.
- Gemeinsamer Speicher- oder Kontext-Store ohne Herkunftskennzeichnung — im Nachhinein nicht rekonstruierbar, welcher Agent oder welche externe Quelle einen bestimmten Eintrag verursacht hat.
- Kein Circuit-Breaker — ein einzelner fehlerhafter oder manipulierter Agent versorgt eine ganze Kette oder Orchestrierung ungebremst mit fehlerhaften Ergebnissen weiter.
- Risikobewertung erfolgt nur je Einzel-Agent, nicht für die Kette als Ganzes — eine Kette, die in Summe ein Lethal-Trifecta-Muster erfüllt, bleibt unentdeckt, weil kein einzelner Agent darin auffällig ist.

9. Abgrenzung, Restrisiko und Verweise

- **Abgrenzung:** setzt eine Architektur mit mehreren, miteinander kommunizierenden Agenten voraus; die Absicherung eines einzelnen Agenten (Identität, Werkzeugzugriff, Kill-Switch, Grundhärtung gegen Prompt Injection) regelt MHC-13, auf dem dieser Eintrag aufbaut.
- **Restrisiko:** emergentes Fehlverhalten in komplexen Multi-Agenten-Systemen ist nicht vollständig vorhersagbar; ein Circuit-Breaker begrenzt die Auswirkung, verhindert aber nicht jede Fehlentscheidung im Einzelfall.
- **TISAX®-Controls:** keine. Dieses MHC ist vollständig originär — kein TISAX®-Anforderungsbefund aus Teil I liegt zugrunde (siehe Teil I, Abschnitt "Originäre Ergänzungen jenseits von TISAX®").
- **Framework:** originär — kein Vergleichsrahmen aus Phase 3 einschlägig; OWASP Top 10 for Agentic Applications (ASI06 Memory & Context Poisoning, ASI07 Insecure Inter-Agent Communication, ASI08 Cascading Failures) als Bedrohungstaxonomie, nicht als Compliance-Rahmenwerk; Priorität P2, zukunftsgerichtet — relevant, sobald mehrere Agenten produktiv miteinander kommunizieren (siehe Priorisierungskapitel).

Ergänzende Einzelmaßnahmen

Die acht Einzelbefunde aus Teil I ohne wiederkehrendes Muster (verteilt über sechs Controls) erhalten keine vollständige MHC-Behandlung — eine Sechs-Rahmen-Herkunftsprüfung wäre für Einzelfälle nicht aussagekräftig. Sie fließen als eigenständige, kompakt formulierte Maßnahmen ein:

Control	Befund (Kurzfassung)	Empfohlene Ergänzung
1.2.2	Rollentrennung durch AI-Doppelrolle unterlaufbar	Kritische Rollentrennung zusätzlich systemtechnisch durchsetzen (getrennte Systemrechte, Vier-Augen-Freigabe im System selbst), nicht nur organisatorisch
1.4.1 (2 Anforderungen: must + should)	Risikobewertungsmethodik geht von eng begrenztem Täterkreis aus — sowohl auf Ebene der Bewertung selbst als auch der zugrunde liegenden Bewertungskriterien	Methodik um Neubewertungspflicht bei signifikant erweitertem Täterkreis (AI-Zugänglichkeit einer Fähigkeit) ergänzen, nicht nur bei neuen Einzelschwachstellen — gilt für Bewertungsergebnis und -kriterien gleichermaßen
1.6.3 (hoch, 2 Anforderungen: Erstfassung + redaktionelle Katalogdublette)	AI-orchestrierte Szenarien in Krisenplanung nicht vorgesehen	Krisenszenarien-Katalog um mindestens ein AI-orchestriertes, mehrgleisiges Eskalationsszenario ergänzen (siehe auch MHC-05)
3.1.4	PIN/Passwort als alleinige Rückfalloption ohne Verschlüsselungspflicht	Verschlüsselung mobiler Geräte/Datenträger als Mindestanforderung, unabhängig von Schutzbedarfsstufe
5.1.2 (must)	Unspezifische Maßnahmen für Fernzugriff ohne Mindeststandard	Verschlüsselung als verbindlicher Mindeststandard bereits auf must-Ebene, nicht erst ab hoch
5.3.1	Re-Identifizierung pseudonymisierter Daten durch AI-Korrelation	Pseudonymisierungsmethodik um Re-Identifizierungsrisiko-Bewertung unter Berücksichtigung AI-gestützter Korrelationsverfahren ergänzen

Glossar

- **A2A (Agent-to-Agent):** Protokollklasse für die direkte Kommunikation zwischen AI-Agenten, u. a. Googles A2A-Protokoll.
- **AITM (Adversary-in-the-Middle):** Echtzeit-Phishing-Relay, das eine Anmeldung über eine gefälschte Seite umleitet und den gültigen Authentisierungsfaktor direkt vom Opfer abgreift.
- **Allow-Listing:** Positivliste zulässiger Ziele (Werkzeuge, Dienste, Pakete); alles nicht Gelistete ist standardmäßig gesperrt.
- **Confused Deputy:** Angriffsmuster, bei dem eine Komponente mit höherer Berechtigung dazu gebracht wird, diese Berechtigung im Auftrag eines weniger stark berechtigten oder böswilligen Akteurs missbräuchlich einzusetzen.
- **CVSS (Common Vulnerability Scoring System):** Standardisierte Bewertungsskala (0–10) für die Kritikalität einer Schwachstelle.
- **EDR / XDR (Endpoint/Extended Detection and Response):** Verhaltensbasierte Erkennungs- und Reaktionswerkzeuge auf Endpunkt- bzw. mehreren Ebenen zugleich.
- **FIDO2 / WebAuthn:** Offener Standard für phishing-resistente Anmeldung mit kryptografischer Bindung an die Adresse des Dienstes.
- **Immutable / WORM:** Speicherung, bei der Daten für einen festgelegten Zeitraum nicht verändert oder gelöscht werden können.
- **KEV (Known Exploited Vulnerabilities):** Von einer Behörde (z. B. CISA) geführte Liste von Schwachstellen, die nachweislich bereits aktiv ausgenutzt werden.
- **Kill-Switch:** technische Möglichkeit, einen einzelnen Agenten oder eine einzelne Integration gezielt und rasch zu deaktivieren, ohne andere Systeme zu beeinträchtigen.
- **Liveness-Detection:** Technisches Verfahren zur Prüfung, ob eine Person in einem Video-/Bildstrom real und live anwesend ist, nicht deepfake-generiert.
- **MCP (Model Context Protocol):** offener Standard, über den AI-Agenten auf externe Werkzeuge, Daten und Dienste zugreifen.
- **MTTC (Mean Time to Containment):** Durchschnittliche Zeit bis zur Eindämmung eines sicherheitsrelevanten Ereignisses.
- **NHI (Non-Human Identity):** technische Identität eines nicht-menschlichen Akteurs (Dienstkonto, Workload, AI-Agent) mit eigenen, verwaltbaren Zugangsdaten.
- **Policy-as-Code:** Sicherheitsvorgaben, die als ausführbare, versionierte Regeln (statt als Text) formuliert und automatisiert durchgesetzt werden.
- **Prompt Injection:** Angriffstechnik, bei der über scheinbar harmlosen Inhalt (E-Mail, Dokument, Werkzeug-Ausgabe) verdeckte Anweisungen in ein AI-System eingeschleust werden.
- **Rug Pull (Komponenten-Kontext):** nachträgliche, unbemerkte Änderung einer bereits freigegebenen Komponente (z. B. MCP-Server, Erweiterung) bei einem Update, wodurch sich Verhalten oder Berechtigungsumfang ändert, ohne die ursprüngliche Prüfung zu durchlaufen.
- **SAST / DAST / SCA:** Statische Codeanalyse, dynamische Anwendungsprüfung und Analyse von Abhängigkeiten (Software Composition Analysis) — drei ergänzende automatisierte Prüfklassen in der Entwicklungs-Pipeline.
- **SOAR (Security Orchestration, Automation and Response):** Plattform, die definierte Reaktionsabläufe (Playbooks) automatisiert ausführt.

- **UEBA (User and Entity Behavior Analytics):** Verhaltensbasierte Erkennung anhand einer erlernten Normallinie für Nutzer und Systeme.
- **Playbook:** Vordefinierter Reaktionsablauf für einen bestimmten Vorfallstyp.

Mapping: MHC ↔ MRIS-TISAX® 2027, Teil I

MHC	Titel	Teil I: Controls	Teil I: Cluster (Phase 2)	Kennzahl
MHC-03	Phishing-resistente Multi-Faktor-Authentisierung	4.1.1, 4.1.2, 4.1.3, 5.1.2, 5.2.7, 6.1.3	Cluster 3	Anteil phishing-resistenter Zugänge
MHC-05	Verhaltensbasierte Detection und Kill-Chain-Korrelation	5.2.3, 1.6.1, 1.6.3, 5.2.4, 5.2.7	Cluster 5 + 6	EDR/XDR-Coverage; Anteil korrelierter Log-Quellen
MHC-08	Unveränderliche Backups und Recovery-Validierung	5.2.8	Cluster 2 (Backup-Teil)	Erfolgsquote Wiederherstellungstests
MHC-09	AI-gestütztes Security-Testing in der Pipeline	1.3.4, 5.2.5, 5.2.6	Cluster 2 (Patch-Teil)	Fristeinhaltung kritischer Schwachstellen
MHC-10	Continuous Control Monitoring und Policy-as-Code	1.1.1, 1.2.1, 1.3.1, 1.3.3, 1.3.4, 1.5.1, 1.6.3, 4.1.3, 4.2.1, 5.2.9, 6.1.2, 6.1.3	Cluster 1	Anteil Prüfprozesse mit Auslösemechanismus
MHC-11	SOAR-basierte Tier-1-Automation und parallele Reaktions-Playbooks	1.6.2	Cluster 8	Mean Time to Containment
MHC-13	AI-Agent-Governance und Harness-Sicherheit	1.3.3, 1.3.4	Cluster 7	Anteil technisch durchgesetzter Allow-Listen
MHC-14	Unabhängige Lieferantennachweis-Verifikation	6.1.1, 6.1.3	Cluster 4 (Lieferanten-Teil)	Anteil auditierter kritischer Lieferanten
MHC-15	AI-resistente Identitätsverifikation Personal/Zutritt	2.1.1, 2.1.3, 3.1.1	Cluster 4 (Personal-/Zutritts-Teil)	Anteil Fälle mit Zweitkanal
MHC-16	Multi-Agent-Integrität	keine (originär)	keiner (originär, außerhalb TISAX®-Diagnose)	Anteil authentifizierter Agent-zu-Agent-Verbindungen

Zu allen MHC zusätzlich: vollständiger Originalwortlaut, Einzelbegründung und Verantwortlich-Zuordnung je Anforderung in Teil I, Kapitel 1–7; Standard-Querverweise (ISO 27001/27017, ISA/IEC 62443-2-1, NIST CSF 2.0) im Anhang von Teil I. MHC-16 bildet hierbei die Ausnahme: Es hat keine Entsprechung in Teil I, Kapitel 1–7, da es nicht aus einer TISAX®-Anforderung abgeleitet ist (siehe Teil I, Abschnitt "Originäre Ergänzungen jenseits von TISAX®").