Sicherheitsanforderungen - SBK

- 1. Zielsetzung und Einordnung
- 2. Begriffe
- 3. Geltungsbereich4. Anforderungsklassen
- 5. Basisanforderungen
- 6. Anforderungen für Besonders schützenswerte Daten
- 7. Anforderungen für KRITIS
- 8. Mindestnachweise für die Sicherheitsabnahme
- 9. Ausnahmen
- 10. Abnahme
- 11. Quellenbasis

---

1. Zielsetzung und Einordnung

Diese Sicherheitsvorgabe beschreibt verbindliche Mindestanforderungen für den Aufbau und Betrieb eines AI Model Hub. Ziel ist es, für die eingesetzten GPU-Server, das Linux-Betriebssystem, die Kubernetes-Plattform, die GPU-Rechen-Pods, die Modellbereitstellung und die API-Anbindung aus einer Middleware heraus ein angemessenes und nachweisbares Sicherheitsniveau sicherzustellen.

Die Plattform darf erst produktiv angebunden und genutzt werden, wenn die Basisanforderungen umgesetzt, dokumentiert und durch die zuständigen Stellen freigegeben wurden. Für die Verarbeitung von Sozialdaten, Personaldaten oder für IT-Services im KRITIS-Kontext gelten zusätzlich die jeweiligen erweiterten Anforderungen dieses Dokuments.

Die Sicherheitsanforderungen orientieren sich methodisch an einem risikoorientierten Vorgehen nach dem Stand der Technik. Sie berücksichtigen insbesondere Anforderungen an Systeme mit Netzwerkverbindungen zu internen Umgebungen, an extern betriebene oder extern untergebrachte Systeme, an Linux- und Kubernetes-Plattformen, an Detektion und Reaktion sowie an regulierte Daten und kritische Dienstleistungen.

1.1 Verankerung und Umsetzungslogik der Anforderungen

Die in dieser Vorgabe beschriebenen Sicherheitsanforderungen leiten sich aus regulatorischen, organisatorischen und technischen Anforderungen an einen sicheren Betrieb des AI Model Hub ab. Sie konkretisieren, welche Sicherheitsanforderungen bei Konzeption, Aufbau, Betrieb und Weiterentwicklung der Plattform zu berücksichtigen sind.

Die Umsetzung der Sicherheitsanforderungen erfolgt entlang eines strukturierten Vorgehens:

1. **Regulatorik und interne Vorgaben**
   Regulatorische Anforderungen, interne Sicherheitsvorgaben und relevante Standards, insbesondere § 392 SGB V, der B3S GKV/PV sowie Art. 32 DSGVO, werden im Informationssicherheitsmanagementsystem der SBK konsolidiert und in verbindliche Sicherheitsanforderungen für das AI Model Hub überführt.

2. **Sicherheitsanforderungen**
   Die Sicherheitsanforderungen in dieser Vorgabe beschreiben, welche Sicherheitsziele und Mindestanforderungen für das AI Model Hub zu konzipieren und umzusetzen sind.
3. **Konzepte**
   Die konkrete Umsetzung der Anforderungen MUSS in geeigneten Konzepten geplant und beschrieben werden. Hierzu gehören insbesondere das AI Model-Hub-Betriebskonzept und das AI Model-Hub-Sicherheitskonzept.
4. **Dokumentation**
   Die Umsetzung MUSS nachvollziehbar dokumentiert werden. Erforderliche Nachweise, Konfigurationen, Prüfprotokolle, Betriebsdokumentationen und Freigaben MÜSSEN so vorliegen, dass die Einhaltung der Anforderungen überprüfbar ist.
5. **Sicherheitsabnahme**
   Vor der produktiven Nutzung MUSS eine Sicherheitsabnahme erfolgen. Dabei wird geprüft, ob die erforderlichen Konzepte, Umsetzungsnachweise und offenen Risiken vorliegen, bewertet und freigegeben sind.

Die Sicherheitsanforderungen sind damit als verbindlicher Rahmen für Planung, Umsetzung, Dokumentation und Abnahme des AI Model Hub zu verstehen.

---

2. Begriffe

Begriff	Bedeutung
AI Model Hub	Der komplette technologische Stack, angefangen bei der Hardware, über Betriebssystem, Container hin zur Generative KI.
Besonders schützenswerte Daten	Sozial- und Gesundheitsdaten, Personaldaten.
Sozialdaten	Sozialdaten im Sinne der jeweils einschlägigen sozialrechtlichen Vorgaben.
Personaldaten	Personenbezogene Daten von Beschäftigten, Bewerbenden oder vergleichbaren Personengruppen.

Begriff	Bedeutung
KRITIS-relevanter IT-Service	IT-Service, der für die Erbringung einer kritischen Dienstleistung erforderlich ist oder dessen Störung die kritische Dienstleistung beeinträchtigen kann.

---

3. Geltungsbereich

Diese Vorgabe gilt für alle Komponenten, Prozesse und Schnittstellen des AI Model Hub. Dazu gehören insbesondere:

Bereich	Einordnung im AI Model-Hub
Daten	• Daten und Speicher • Datensicherung und Wiederherstellung • Protokollierung • Sichere Löschung • Notfallbetrieb und Wiederanlauf
Konfigurationen und Einstellungen	• Kubernetes-Konfigurationen • Netzwerkkonfigurationen • Sicherheitsrichtlinien • Deployment-Konfigurationen • Außerbetriebnahme und sichere Löschung
Identitäten und Benutzer	• Administrations- und Managementzugänge
Applications	• Virtualisierung • Kubernetes- und Container-Plattform • GenAI Modelle • Laufzeitumgebung
Netzwerk	• VPN-Anbindung • Netzwerk und Netzwerksegmentierung • Firewall
Betriebssystem	• Betriebssystem (z.B. Linux)
Physische Hosts	• GPU-Server Hardware

Physisches Netzwerk	• Switch Hardware • Router Hardware • Firewall Hardware • Internet Anbindung
Physisches Rechenzentrum	• Housing-Standort/Colocation Rack

Zum Zeitpunkt der Erstellung dieser Vorgabe ist ausschließlich die Bereitstellung von GPU-Rechenleistung vorgesehen. Innerhalb des AI Model Hub werden hierzu GPU-Rechencontainer bereitgestellt, die von einer außerhalb der Plattform betriebenen Middleware über definierte API-Schnittstellen angesprochen werden. Das AI Model Hub übernimmt in diesem Use Case ausschließlich die technische Ausführung der GPU-basierten Rechenoperationen und die Bereitstellung der hierfür erforderlichen Modellendpunkte. Eine direkte Nutzung der Plattform durch Benutzer, der Betrieb eigenständiger Fachanwendungen sowie eine direkte Anbindung von Endgeräten oder Fachverfahren sind nicht Bestandteil des betrachteten Use Cases.

Die Anforderungen sind auf diesen technischen Bezugsrahmen abgestimmt. Sie erheben keinen Anspruch darauf, abweichende oder zusätzliche Technologien, Betriebsmodelle oder Nutzungsszenarien vollständig abzudecken. Werden die Architektur, der Technologie-Stack, die Betriebsverantwortung oder die vorgesehenen Anwendungsfälle wesentlich geändert oder erweitert, MUSS vor der Umsetzung geprüft werden, ob zusätzliche oder angepasste Sicherheitsanforderungen erforderlich sind. Dies gilt insbesondere bei der Einführung zusätzlicher Plattformkomponenten, neuer Schnittstellen, weiterer Netzverbindungen, externer Dienste, neuer KI-Modelle, Trainings- oder Fine-Tuning-Verfahren sowie zusätzlicher Datenarten oder Anwendungsfälle.

Die Prüfung MUSS vor der produktiven Nutzung der geänderten oder erweiterten Komponenten erfolgen. Erforderliche Sicherheitsanforderungen MÜSSEN in diese Vorgabe oder in ergänzende Sicherheitsdokumente aufgenommen werden.

---

4. Anforderungsklassen

Die Anforderungen sind in drei Klassen gegliedert und kumulativ anzuwenden.

Klasse	Anwendung
Basisanforderungen	Gelten immer für den Aufbau und Betrieb des AI Model-Hubs.

Klasse	Anwendung
Anforderungen für Besonders schützenswerte Daten	Gelten zusätzlich, wenn Sozialdaten, Gesundheitsdaten, Beschäftigtendaten, Bewerberdaten oder sonstige besonders schützenswerte personenbezogene Daten (nach Art. 9 DSGVO) verarbeitet werden oder verarbeitet werden können.
Anforderungen für KRITIS	Gelten zusätzlich, wenn auf der Plattform IT-Services betrieben werden, die für die Erfüllung kritischer Dienstleistungen (siehe B3S GKV/PV) erforderlich sind oder diese unterstützen.

---

5. Basisanforderungen

Das AI Model Hub MUSS von Beginn an unter Berücksichtigung der Informationssicherheit konzipiert, aufgebaut und weiterentwickelt werden. Sicherheitsanforderungen MÜSSEN bereits bei der Architekturplanung, der Auswahl von Hardware und Software, der Netzsegmentierung, der Gestaltung von Schnittstellen, dem Berechtigungsmodell sowie den Betriebs- und Administrationsprozessen berücksichtigt werden.

Sicherheitsmaßnahmen dürfen nicht erst nachträglich ergänzt werden. Änderungen an Architektur, Komponenten, Diensten oder Betriebsmodellen MÜSSEN daher bereits in der Planungsphase auf ihre sicherheitsrelevanten Auswirkungen geprüft werden.

Das AI Model Hub MUSS standardmäßig mit sicheren Voreinstellungen betrieben werden. Nicht benötigte Funktionen, Dienste, Schnittstellen, Benutzerkonten, Berechtigungen und Kommunikationsverbindungen MÜSSEN deaktiviert oder gesperrt sein.

Zugriffe und Funktionen dürfen nur im erforderlichen Umfang freigeschaltet werden. Unsichere Standardkonfigurationen MÜSSEN vor der Inbetriebnahme angepasst werden. Neue Komponenten und Dienste dürfen erst produktiv genutzt werden, wenn ihre sicherheitsrelevanten Einstellungen geprüft und freigegeben wurden.

5.1 Governance, Freigabe und Verantwortlichkeiten

5.1.1 Sicherheitsfreigabe vor Inbetriebnahme

Das AI Model Hub DARF erst produktiv genutzt oder an interne Systeme angebunden werden, wenn eine dokumentierte Sicherheitsfreigabe durch Informationssicherheit, IT-Betrieb, Datenschutz bei Bedarf, Netzwerkverantwortliche und Serviceverantwortliche vorliegt.

Nachweis:

5.1.2 Dokumentierte System- und Serviceverantwortung

Für das AI Model Hub MUSS eine eindeutige fachliche, technische und sicherheitsbezogene Verantwortung festgelegt werden. Dazu gehören Systemverantwortung, Plattformverantwortung, Applikationsverantwortung, Informationssicherheitskontakt, Datenschutzkontakt bei Bedarf und Betriebsverantwortung.

Nachweis:

5.1.3 Verbindliche Betriebsvereinbarungen

Für Betrieb, Entstörung, Wartung, Fernzugriff, Sicherheitsvorfälle, Änderungen, Notfallbetrieb und Kommunikation MÜSSEN verbindliche Betriebsvereinbarungen und sicherheitsrelevante Vorgaben bestehen. Diese MÜSSEN mindestens Betriebszeiten, Servicezeiten, SLAs, RTO, RPO, Mitwirkungspflichten, Eskalationswege und Kommunikationskanäle enthalten.

Nachweis:

5.1.4 Dokumentierter Schutzbedarf

Für die Plattform, die angebundenen Use Cases, Datenarten, Schnittstellen, Modellartefakte, Logs und Backups MUSS der Schutzbedarf dokumentiert werden. Die Sicherheitsarchitektur MUSS am höchsten relevanten Schutzbedarf ausgerichtet sein.

Nachweis:

5.1.5 Architektur- und Sicherheitskonzept

Vor Aufbau der Plattform MUSS ein Sicherheitskonzept erstellt werden. Es MUSS mindestens (Sicherheits)-Architektur, Netzsegmente, Datenflüsse, Vertrauensgrenzen, Administrationswege, Schnittstellen, Kubernetes-Design, Logging, Detektion, Backup, Notfallbetrieb und Betriebsprozesse beschreiben.

Nachweis: AI Model-Hub-Sicherheitskonzept

5.2 Housing und physische Sicherheit

5.2.1 Geeigneter Rechenzentrumsstandort

Die GPU-Server MÜSSEN in einer geeigneten Rechenzentrumsumgebung mit geregelter physischer Sicherheit, Zutrittskontrolle, Stromversorgung, Klimatisierung, Brandfrüherkennung, Brandschutz und Betriebsüberwachung betrieben werden.

Nachweis: C5 Testat, ISO 27001 des Housing Dienstleisters

5.2.2 Physischer Zugriffsschutz auf Rack und Server

Der physische Zugriff auf Rack, Server, Netzwerkports, Managementports und Datenträger MUSS auf autorisierte Personen beschränkt sein. Zutritte MÜSSEN nachvollziehbar protokolliert werden.

Nachweis: C5 Testat, ISO 27001 des Housing Dienstleisters, Zutrittskonzept, Zutrittsprotokolle, Berechtigungsliste, Rack-Dokumentation.

5.2.3 Begleiteter Zugriff durch Housing-Anbieter

Mitarbeitende des Housing-Anbieters DÜRFEN nur im vertraglich geregelten Umfang Zugriff auf Rack, Server oder Verkabelung erhalten. Wartung und Eingriffe MÜSSEN dokumentiert, genehmigt und nachvollziehbar sein.

Nachweis: AI Model-Hub-Betriebskonzept

5.2.4 Umgang mit Datenträgern und Hardwaretausch

Datenträger, GPU-Server und defekte Komponenten DÜRFEN den kontrollierten Bereich nur nach dokumentierter Freigabe verlassen. Datenträger mit schützenswerten Daten MÜSSEN vor Rückgabe, Entsorgung oder Austausch sicher gelöscht oder physisch vernichtet werden.

Nachweis: AI Model-Hub-Betriebskonzept

5.3 Netzwerk, VPN und Zonen

5.3.1 Netzsegmentierung

Das AI Model Hub MUSS durch geeignete Netzsegmentierung in voneinander getrennte Sicherheitszonen unterteilt werden. Mindestens voneinander zu trennen sind:

- Management
  - Hardware
  - Server- und Plattform
- Produktivumgebung
- Testumgebung
- Entwicklungsumgebung

Nachweis: AI Model-Hub-Betriebskonzept

5.3.2 Default-Deny-Kommunikation

Alle Netzwerkverbindungen MÜSSEN nach dem Default-Deny-Prinzip geregelt werden. Es DÜRFEN ausschließlich technisch und fachlich erforderliche Kommunikationsbeziehungen freigegeben werden. Berechtigungen und Netzwerkfreigaben MÜSSEN nach dem Least-Privilege- und Need-to-Communicate-Prinzip vergeben werden.

Nachweis: AI Model-Hub-Sicherheitskonzept

5.3.3 Anbindung an das BITMARCK-Rechenzentrum

Die Anbindung an das Bitmarck Rechenzentrum MUSS über eine, gegen unbefugte Nutzung, Manipulation und Mitlesen geschützte Verbindung erfolgen. Die konkrete technische Umsetzung MUSS beschrieben und begründet werden.

Nachweis: AI Model-Hub-Sicherheitskonzept, AI Model-Hub-Betriebskonzept

5.3.3.1 Site-to-Site-VPN

Kommt für die Anbindung ein Site-to-Site-VPN zum Einsatz, MÜSSEN Authentisierung, Verschlüsselung, Schlüsselaustausch und Zertifikatsnutzung dem Stand der Technik entsprechen.

Für die Umsetzung zu beachten:

- IPsec: BSI TR-02102-3 "Kryptographische Verfahren: Verwendung von Internet Protocol Security (IPsec) und Internet Key Exchange (IKEv2)" Version: 2026-01
- TLS: BSI TR-03116-4 Kryptographische Vorgaben für Projekte der Bundesregierung Teil 4 Stand 2025

Nachweis: AI Model-Hub-Sicherheitskonzept, AI Model-Hub-Betriebskonzept

5.3.3.2 MPLS

Kommt für die Anbindung eine MPLS-Verbindung zum Einsatz, MUSS diese als geschützte, providerseitig getrennte Kommunikationsverbindung konzipiert und dokumentiert werden. Die Netztrennung, Übergabepunkte, Routing-Beziehungen, Sicherheitszonen und Verantwortlichkeiten zwischen Betreiber, BITMARCK und beteiligten Dienstleistern MÜSSEN nachvollziehbar beschrieben sein.

Da MPLS keine Ende-zu-Ende-Verschlüsselung ersetzt, MÜSSEN zusätzlichen Schutzmaßnahmen gegen Mitlesen und Manipulation ergriffen werden. Dafür MUSS eine durch den Leitungsanbieter bereitgestellte MPLS-Verbindung mit integrierter Verschlüsselung (z.B. durch Overlay-Verschlüsselung (IPsec)) beauftragt werden.

Eine unverschlüsselte Übertragung über MPLS ist unzulässig!

Für die Umsetzung zu beachten:

- IPsec: BSI TR-02102-3 "Kryptographische Verfahren: Verwendung von Internet Protocol Security (IPsec) und Internet Key Exchange (IKEv2)" Version: 2026-01
- TLS: BSI TR-03116-4 Kryptographische Vorgaben für Projekte der Bundesregierung Teil 4 Stand 2025

5.3.4 Internetzugriff

Die gesamte AI Model Hub DARF keinen unreglementierten direkten Internetzugriff besitzen. Externe Kommunikationsbeziehungen MÜSSEN über definierte Übergänge (z.B. Proxie der Bitmarck) geführt sowie nach dem Least-Privilege- und Need-to-Communicate-Prinzip freigegeben werden.

Nachweis: AI Model-Hub-Sicherheitskonzept

5.3.5 Administrationszugriff

Administration MUSS ausschließlich über definierte Administrationswege erfolgen. Direkte Administration aus normalen Client-Netzen oder aus dem Internet DARF NICHT erlaubt werden.

Nachweis: AI Model-Hub-Sicherheitskonzept

5.3.6 Management-Interfaces

Hardware-Management Interfaces (z.B. iDRAC, iLO, IPMI usw) und vergleichbare Administrationsschnittstellen MÜSSEN gemäß Anforderung 5.3.1 Netzsegmentierung in einem separaten Managementnetz betrieben werden. Zugriffe MÜSSEN mittels starker Authentisierung, vorzugsweise unter Verwendung von Multi-Faktor-Authentisierung, abgesichert werden. Sämtliche Administrationszugriffe MÜSSEN protokolliert werden.

Nachweis: AI Model-Hub-Sicherheitskonzept

5.4 Betriebssystem, Serverhärtung

5.4.1 Betriebssysteme

Für den Betrieb der AI Model Hub MÜSSEN Betriebssysteme mit Long-Term-Support (LTS/LTSR) eingesetzt werden. Für alle eingesetzten Betriebssysteme MUSS ein aktiver Hersteller-, Distributor- oder Community-Security-Support bestehen. End-of-Life-Betriebssysteme DÜRFEN NICHT betrieben werden. Die Auswahl des Betriebssystems MUSS anhand dokumentierter Kriterien begründet werden.

Nachweis: AI Model-Hub-Betriebskonzept

5.4.2 Härtung der Systeme

Für alle Komponenten der AI Model Hub MÜSSEN dokumentierte Härtungsstandards (z.B. CIS) definiert, umgesetzt und regelmäßig überprüft werden. Dies gilt insbesondere für Betriebssysteme, Virtualisierungs- und Containerplattformen, Managementschnittstellen sowie Plattform- und Infrastrukturdienste. Nicht benötigte Funktionen, Dienste, Benutzerkonten, Softwarekomponenten, Netzwerkports und Protokolle MÜSSEN deaktiviert oder entfernt werden.

Für die Umsetzung zu beachten:

- BSI TR-02102-1 "Kryptographische Verfahren: Empfehlungen und Schlüssellängen" Version: 2026-01

Nachweis: AI Model-Hub-Sicherheitskonzept, AI Model-Hub-Betriebskonzept

5.4.3 Sichere Basiskonfiguration der GPU-Server

GPU-Treiber, CUDA-Bibliotheken, Firmware, BIOS, BMC und sonstige hardwarebezogene Komponenten MÜSSEN in das Patch- und Schwachstellenmanagement einbezogen werden.

Nachweis: AI Model-Hub-Betriebskonzept

5.4.4 Host-Firewall

Auf allen Servern MUSS eine lokale Host-Firewall oder eine vergleichbare technische Zugriffskontrolle aktiviert sein. Die Konfiguration MUSS nach dem Least-Privilege- und Need-to-Communicate-Prinzip erfolgen. Es DÜRFEN ausschließlich die für den Betrieb erforderlichen Netzwerkverbindungen, Ports und Protokolle freigegeben werden.

Nachweis: AI Model-Hub-Sicherheitskonzept, AI Model-Hub-Betriebskonzept

5.4.5 Zeitsynchronisation

Alle Systeme MÜSSEN eine vertrauenswürdige und einheitliche Zeitsynchronisation nutzen. Der Status der Zeitsynchronisation MUSS überwacht werden.

Nachweis:

5.5 Virtualisierungs-, Container- und Orchestrierungsplattformen

5.5.1 Virtualisierungs-, Container- und Orchestrierungsplattformen

Vor Inbetriebnahme MUSS festgelegt werden, wie Anwendungen, Umgebungen, Namespaces, Cluster, Netze, CPU, GPU, Speicher und Datenbestände voneinander getrennt werden. Anwendungen mit unterschiedlichem Schutzbedarf oder deutlich unterschiedlichen Angriffsflächen SOLLTEN nicht denselben Cluster teilen.

Nachweis: AI Model-Hub-Sicherheitskonzept, AI Model-Hub-Betriebskonzept

5.5.2 Kubernetes-Identitäten und Rechte

Kubernetes MUSS jede administrative und automatisierte Aktion authentisieren und autorisieren. RBAC MUSS restriktiv nach Least Privilege umgesetzt werden. Cluster-Admin-Rechte DÜRFEN nur einem sehr kleinen, namentlich bekannten Personenkreis zugeordnet werden.

Nachweis:

5.5.3 Service Accounts

Pods DÜRFEN NICHT den Default-Service-Account nutzen. Service Accounts MÜSSEN anwendungsbezogen vergeben und auf die minimal erforderlichen Rechte beschränkt werden. Tokens DÜRFEN nur eingebunden werden, wenn sie tatsächlich benötigt werden.

Nachweis:

5.5.4 Network Policies

Zwischen Namespaces, Pods und Services MUSS eine restriktive Netzwerkkommunikation umgesetzt werden. Standardmäßig MUSS Kommunikation zwischen Namespaces untersagt sein. Erforderliche Verbindungen MÜSSEN explizit freigegeben werden.

Nachweis:

5.5.5 Admission Control

Der Cluster MUSS technische Richtlinien durchsetzen, die unsichere Pod-Konfigurationen verhindern. Dazu gehören mindestens Verbot privilegierter Container, Einschränkung von HostPath, HostNetwork, HostPID, HostIPC, Root-Containern, unnötigen Linux Capabilities und unsicheren Image-Quellen.

Nachweis:

5.5.6 Container Images

Container Images DÜRFEN nur aus freigegebenen Registries bezogen werden. Images MÜSSEN versioniert, nachvollziehbar gebaut, auf Schwachstellen geprüft und vor produktivem Einsatz freigegeben werden. Aus betrieblicher Sicht MUSS einen dokumentierter Prozess zur

Bewertung und Freigabe von Container Images für den produktiven Einsatz definieren und umsetzen sein.

Nachweis:

5.5.6.1 Image-Quellen

Aus betrieblicher Sicht MUSS festgelegt werden, aus welchen Quellen Container Images für das AI Model Hub bezogen werden dürfen. Zulässige Image-Quellen MÜSSEN nachvollziehbar bewertet und freigegeben sein.

Eine Image-Quelle darf nur als zulässig eingestuft werden, wenn Herkunft, Betreiber, Bezugsweg, Aktualität, Integrität und Supportfähigkeit der bereitgestellten Images nachvollziehbar sind. Der Bezug von Images aus öffentlichen, privaten oder herstellerspezifischen Registries MUSS kontrolliert erfolgen.

Container Images aus unbekannten, nicht freigegebenen oder nicht nachvollziehbaren Quellen DÜRFEN NICHT für den produktiven Einsatz verwendet werden.

Nachweis:

5.5.7 Keine unkontrollierten Modell- oder Paketdownloads

Modelle, Container, Python-Pakete, Helm Charts, Operators und sonstige Artefakte DÜRFEN NICHT unkontrolliert aus dem Internet nachgeladen werden. Externe Artefakte MÜSSEN vor Einsatz geprüft, versioniert und intern bereitgestellt werden.

Nachweis:

5.5.8 Secrets Management

Secrets, Tokens, API-Keys, Zertifikate und Passwörter DÜRFEN NICHT in Images, Git-Repositories, Kubernetes-Manifests oder Logs abgelegt werden. Sie MÜSSEN über ein geeignetes Secrets-Management bereitgestellt, rotiert und protokolliert werden.

Nachweis:

5.5.9 Ressourcenbegrenzung

Für Pods MÜSSEN CPU-, RAM-, Storage- und GPU-Ressourcen begrenzt werden. Kubernetes-Quotas und Limits MÜSSEN verhindern, dass einzelne Workloads die Plattform destabilisieren.

Nachweis:

5.5.10 Trennung von Entwicklungs-, Test- und Produktionsumgebungen

Entwicklung, Test, Abnahme und Produktion MÜSSEN logisch und technisch getrennt werden. Produktivdaten DÜRFEN nicht in Entwicklungs- oder Testumgebungen verwendet werden.

Nachweis:

5.6 API, Middleware und LLM-Zugriff

5.6.1 API-Absicherung

Der Zugriff der Middleware auf die LLM-Dienste MUSS über definierte, authentisierte und autorisierte API-Schnittstellen erfolgen. Jede API MUSS eine dokumentierte Schnittstellenbeschreibung, Authentisierung, Autorisierung, Transportverschlüsselung, Rate Limits und Logging besitzen.

Für die Umsetzung zu beachten:

- BSI TR-02102-3 "Kryptographische Verfahren: Verwendung von Internet Protocol Security (IPsec) und Internet Key Exchange (IKEv2)" Version: 2026-01

Nachweis:

5.6.2 API Direktzugriff

Middleware, Fachanwendungen und Benutzer DÜRFEN nicht direkt auf interne Model-Serving-Pods zugreifen. Der Zugriff MUSS über freigegebene API-Endpunkte, Gateways oder Service-Schichten geführt werden.

Nachweis:

5.6.3 Prompt- und Output-Schutz

Für LLM-Nutzung MÜSSEN technische und organisatorische Schutzmaßnahmen gegen Missbrauch, Prompt Injection, unbeabsichtigte Offenlegung, unsichere Ausgaben und unzulässige Datenverarbeitung umgesetzt werden.

Nachweis:

5.6.4 Use Cases

Vor produktiver Nutzung MÜSSEN die zulässigen Use Cases, Datenklassen, Nutzergruppen und technischen Schnittstellen dokumentiert und freigegeben werden. Nicht freigegebene Nutzungen sind UNZULÄSSIG.

Nachweis:

5.7 Identitäts- und Berechtigungsmanagement

5.7.1 Vergabeprozess und Least Privilege

Alle Benutzerkonten, Rollen und Berechtigungen MÜSSEN nach einem standardisierten Prozess für die Provisionierung, Pausierung und Entzug gemanaged und freigegeben werden. Alle Berechtigungen auf Linux, Kubernetes, Registry, Monitoring, Backup, Modell-Repositories, API-Gateways, Middleware und Anwendungen MÜSSEN nach Least Privilege und Need-to-Know vergeben werden.

Nachweis:

5.7.2 Personalisierte Administration

Administrative Tätigkeiten MÜSSEN über personalisierte administrative Konten erfolgen. Gemeinsame Administrationskonten DÜRFEN NICHT verwendet werden.

Nachweis:

5.7.3 MFA für privilegierte Zugriffe

Privilegierte Zugriffe MÜSSEN mit Multi-Faktor-Authentisierung abgesichert werden. Dies gilt insbesondere für die Betriebssystem- und Kubernetes-Administration.

Nachweis:

5.7.4 Regelmäßiger Berechtigungsreview

Alle Benutzerkonten, Rollen und Berechtigungen MÜSSEN aktuell und gemäß der Funktion / den Geschäftstätigkeiten der Ausführenden zugeteilt sein. Benutzerkonten und Berechtigungen MÜSSEN mindestens jährlich im Rahmen des IKS und zusätzlich anlassbezogen überprüft werden. Privilegierte Benutzerkonten (administrative Konten) und deren Berechtigungen MÜSSEN halbjährlich im Rahmen des IKS überprüft werden.

Nachweis:

5.7.5 Technische Konten (Service Accounts)

Technische Konten wie z.B. Service Accounts MÜSSEN eindeutig einem Zweck, einer Anwendung und einer verantwortlichen Person zugeordnet sein. Interaktive Anmeldung SOLLTE deaktiviert werden, sofern technisch möglich.

Nachweis:

5.8 Protokollierung, Detektion und Sicherheitsüberwachung

5.8.1 EDR/XDR auf Servern

Das AI Model Hub MUSS an ein EDR- oder XDR-System angebunden werden.

Nachweis:

5.8.2 SIEM-Anbindung

Sicherheitsrelevante Logs der GPU-Server, Linux-Systeme, Kubernetes-Komponenten, API-Gateways, Registry, VPN, Firewalls, PAM, Secrets-Management und Middleware-Schnittstellen MÜSSEN an ein SIEM oder eine vergleichbare zentrale Sicherheitsüberwachung angebunden werden.

Nachweis:

5.8.3 Manipulationsschutz für Protokolle

Sicherheitsrelevante Protokolle MÜSSEN gegen Manipulation, unbefugten Zugriff und unbefugtes Löschen geschützt werden. Die Aufbewahrung MUSS dem Schutzbedarf und den regulatorischen Anforderungen entsprechen.

Protokolldaten gelten als sicherheitsrelevant, wenn sie für die Nachvollziehbarkeit von Zugriffen, administrativen Tätigkeiten, sicherheitsrelevanten Ereignissen, Störungen, Änderungen, Fehleranalysen, Sicherheitsvorfällen oder regulatorischen Nachweisen erforderlich sind.

Nachweis:

5.8.4 Detektionsregeln für AI Model-Hub

Für die Plattform MÜSSEN spezifische Detektionsregeln definiert werden. Dazu gehören mindestens auffällige administrative Aktionen, ungewöhnliche API-Nutzung, anomale Modellzugriffe, unerwartete Egress-Verbindungen, Änderungen an Kubernetes-RBAC, Secrets-Zugriffe, Image-Änderungen und verdächtige Pod-Konfigurationen.

Nachweis:

5.8.5 Alarmierungs- und Reaktionsprozess

Für Sicherheitsalarme MUSS ein dokumentierter Prozess existieren. Zuständigkeiten, Rufbereitschaften, Eskalationen, Reaktionszeiten und technische Sofortmaßnahmen MÜSSEN geregelt sein.

Nachweis:

5.8.6 Security Operations Center (SOC) / Sicherheitsüberwachung

Ein Security Operations Center (SOC) bzw. eine gleichwertige Funktion zur kontinuierlichen Sicherheitsüberwachung MUSS etabliert sein, um sicherheitsrelevante Ereignisse, Anomalien und Bedrohungen frühzeitig zu erkennen, zu bewerten und zu behandeln. Die SOC Prozesse können auf die Dienstleister ausgelagert werden. In diesem Fall müssen die SOC Schnitstellen technisch, prozessual und organisatorisch aus der SBK Perspektive harmonisiert werden. Die relevanten Protokoll- und Ereignisdaten (u. a. aus Systemen, Netzwerken, Anwendungen, IAM, Sicherheitskomponenten und kritischen Diensten) MÜSSEN zentral erfasst, analysiert und hinsichtlich ihrer Kritikalität bewertet werden. Es MÜSSEN definierte Prozesse für Erkennung, Bewertung, Eskalation, Reaktion und Nachverfolgung von Sicherheitsereignissen vorhanden sein. Die Leistungsfähigkeit und Wirksamkeit der SOC-Prozesse MUSS regelmäßig überprüft und im Rahmen von Tests, Übungen und kontinuierlichen Verbesserungsmaßnahmen weiterentwickelt werden.

Nachweis:

5.9 Schwachstellen-, Patch- und Konfigurationsmanagement

5.9.1 Schwachstellenscans

Die Plattform MUSS regelmäßig auf Schwachstellen geprüft werden. Mindestens einzubeziehen sind externe und interne Angriffsflächen, Linux-Hosts, Kubernetes-Komponenten, Container Images, API-Endpunkte, Registries und relevante Middleware-Komponenten.

Nachweis:

5.9.2 Bewertung und Behebung von Schwachstellen

Schwachstellen MÜSSEN zeitnah bewertet, priorisiert und entsprechend ihrer Kritikalität behoben oder wirksam kompensiert werden. Bei kritischen Schwachstellen MÜSSEN

Sofortmaßnahmen wie Isolation, Deaktivierung betroffener Funktionen oder zusätzliche Filterung geprüft werden.

Nachweis:

5.9.3 Patch-Management

Für Betriebssystem, Kernel, Firmware, GPU-Treiber, Kubernetes, Container Runtime, Operators, Images, Libraries, API-Komponenten und Middleware MUSS ein dokumentierter Patch-Prozess bestehen.

Nachweis:

5.9.4 Umgang mit End-of-Life

Komponenten ohne Security-Support DÜRFEN NICHT produktiv betrieben werden. Abweichungen MÜSSEN risikobewertet, zeitlich befristet, kompensiert und durch Informationssicherheit freigegeben werden.

Nachweis:

5.9.5 Konfigurationsmanagement

Alle wesentlichen Plattformkomponenten MÜSSEN in einem Konfigurationsmanagement geführt werden. Dazu gehören Versionen, Patchstände, Verantwortliche, Standort, Schutzbedarf, Abhängigkeiten, RTO, RPO und Supportstatus.

Nachweis:

5.10 Verschlüsselung und Schlüsselmanagement

5.10.1 Verschlüsselung bei Übertragung

Die Kommunikation zwischen Middleware, API-Gateway, Kubernetes-Services, Administrationskomponenten, Monitoring, Registry, Backup und internen Systemen MUSS verschlüsselt werden. Die kryptografischen Verfahren MÜSSEN dem Stand der Technik entsprechen.

Für die Umsetzung zu beachten:

- BSI TR-03116-4 Kryptographische Vorgaben für Projekte der Bundesregierung Teil 4 Stand 2025

Nachweis: TLS-Konzept, Zertifikatsinventar, technische Prüfung.

5.10.2 Verschlüsselung gespeicherter Daten

Schützenswerte Daten, Modellartefakte, Konfigurationen, Secrets, Backups, Logs und Datenträger MÜSSEN entsprechend Schutzbedarf verschlüsselt werden. Schlüssel DÜRFEN nicht gemeinsam mit den verschlüsselten Daten abgelegt werden.

Für die Umsetzung zu beachten:

- BSI TR-02102-1 "Kryptographische Verfahren: Empfehlungen und Schlüssellängen" Version: 2026-01

Nachweis:

5.10.3 Zertifikats- und Schlüsselrotation

Zertifikate, API-Schlüssel, Tokens und kryptografische Schlüssel MÜSSEN inventarisiert, geschützt, regelmäßig erneuert und bei Verdacht auf Kompromittierung unverzüglich rotiert werden.

Für die Umsetzung zu beachten:

- BSI TR-02102-1 "Kryptographische Verfahren: Empfehlungen und Schlüssellängen" Version: 2026-01

Nachweis:

5.11 Backup, Restore und Wiederherstellung

5.11.1 Datensicherung

Für kritische Plattformdaten MUSS eine Datensicherung eingerichtet werden. Dazu gehören Kubernetes-Konfigurationen, etcd oder vergleichbare Steuerdaten, Deployment-Manifeste, Registry-Daten, Modellartefakte, Konfigurationen, relevante Logs und Betriebsdokumentation.

Nachweis: AI Model-Hub-Betriebskonzept

5.11.2 Wiederherstellungstests

Restore-Verfahren MÜSSEN regelmäßig getestet werden. Der Test MUSS nachweisen, dass die Plattform oder definierte Kernkomponenten innerhalb der vorgesehenen Wiederherstellungsziele wiederhergestellt werden können.

Nachweis: AI Model-Hub-Betriebskonzept

5.11.3 Schutz der Backups

Backups MÜSSEN gegen unbefugten Zugriff, Manipulation, Verschlüsselung durch Ransomware und unbeabsichtigtes Löschen geschützt werden. Für kritische Daten SOLLTEN unveränderbare oder besonders geschützte Backup-Kopien vorgesehen werden.

Nachweis: AI Model-Hub-Betriebskonzept

5.12 Change-, Release- und Betriebsprozesse

5.12.1 Change-Management

Änderungen an Netz, VPN, Firewall, Linux, Kubernetes, Images, Modellartefakten, API, Middleware-Anbindung, IAM, Logging und Security Policies MÜSSEN über einen dokumentierten Change-Prozess erfolgen.

Nachweis:

5.12.2 Notfalländerungen

Für dringende Sicherheitsänderungen MUSS ein beschleunigter Prozess existieren. Dieser MUSS Nachdokumentation, nachträgliche Prüfung und Risikobewertung enthalten.

Nachweis:

5.12.3 Betriebsdokumentation

Die Plattform MUSS so dokumentiert sein, dass Betrieb, Entstörung, Wiederanlauf und Sicherheitsreaktion auch bei Ausfall einzelner Wissensträger möglich sind.

Nachweis: AI Model-Hub-Betriebsdokumentation

5.12.4 Schulung des Betriebspersonals

Administrierende und Betriebsverantwortliche MÜSSEN für Linux, Kubernetes, GPU-Plattform, KI-spezifische Risiken, Logging, Incident Response und Datenschutzanforderungen geschult sein.

Nachweis:

5.13 Dienstleister, Lieferkette und Housing-Anbieter

5.13.1 Lieferanten- und Dienstleistersteuerung

Leistungen, Verantwortlichkeiten, Schnittstellen und Sicherheitsanforderungen von Housing-Anbietern, Hardwarelieferanten, Wartungsdienstleistern, Softwarelieferanten, Managed Service Providern (MSP) sowie sonstigen Dritten MÜSSEN vertraglich geregelt werden. Anforderungen der Informationssicherheit MÜSSEN verbindlich vereinbart, nachweisbar umgesetzt und bei Einsatz von Unterauftragnehmern entsprechend weitergegeben werden.

Nachweis:

5.13.2 Meldepflichten und Sicherheitsvorfälle bei Dienstleistern

Dienstleister MÜSSEN verpflichtet werden, sicherheitsrelevante Ereignisse, Schwachstellen, Störungen und Änderungen mit möglicher Auswirkung auf die Plattform oder kritische Dienstleistungen unverzüglich zu melden.

Nachweis:

5.13.3 Exit- und Rückbauplanung

Für das AI Model Hub MUSS eine dokumentierte Exit-Strategie vorliegen. Diese MUSS den kontrollierten Rückbau, die Migration oder den Weiterbetrieb von Plattform, Daten, LLM Modelle und Betriebsprozessen bei Anbieterwechsel, Standortwechsel, Vertragsende oder schwerwiegenden Sicherheitsereignissen sicherstellen. Die Exit-Strategie MUSS regelmäßig überprüft und bei wesentlichen Änderungen aktualisiert werden (min. einmal im Jahr).

Nachweis: AI Model Hub Exit-Strategie Konzept.

6. Anforderungen für Besonders schützenswerte Daten

Die folgenden Anforderungen gelten zusätzlich zu den Basisanforderungen, sobald Besonders schützenswerte Daten verarbeitet werden oder realistischerweise in Prompts, Dateien, Logs, RAG-Datenbestände, Embeddings, Modellkontexte oder Ausgaben gelangen können.

6.1 Zulässigkeit, Zweckbindung und Use-Case-Steuerung

6.1.1 Datenschutz- und Rechtsfreigabe vor Verarbeitung

Vor Verarbeitung von Sozialdaten oder Personaldaten MUSS eine Freigabe durch Datenschutz, Recht bei Bedarf, Informationssicherheit und Fachbereich erfolgen. Die Verarbeitung DARF erst erfolgen, wenn Zweck, Rechtsgrundlage, Datenarten, Rollen, Speicherorte, Löschfristen und technische Schutzmaßnahmen dokumentiert sind.

Nachweis:

6.1.2 Keine offene Datenverarbeitung ohne Datenklassenfreigabe

Die Plattform DARF NICHT für beliebige Freitexte, Dokumente oder Datenklassen geöffnet werden. Zulässige Datenklassen sowie die Verarbeitungsregeln MÜSSEN je Use Case explizit festgelegt werden.

Nachweis:

6.1.3 Sozialdatenverarbeitung im Auftrag

Wenn durch den Housing-Anbieter, Dienstleister, Plattformbetreiber oder Unterauftragnehmer eine Verarbeitung von Sozialdaten im Auftrag vorliegt oder nicht ausgeschlossen werden kann, MUSS dies vorab rechtlich und datenschutzrechtlich bewertet werden. Erforderliche Vereinbarungen, Anzeigen und Genehmigungen MÜSSEN vor Beginn der Verarbeitung abgeschlossen sein.

Nachweis:

6.1.4 Personaldaten und Beschäftigungskontext

Use Cases mit Personaldaten MÜSSEN gesondert freigegeben werden. KI-Ausgaben DÜRFEN ohne gesonderte Freigabe nicht für arbeitsrechtlich relevante Entscheidungen, Leistungsbewertungen, Auswahlentscheidungen oder vergleichbare Personalmaßnahmen genutzt werden.

Nachweis:

6.2 Datenminimierung, Protokollierung und Speicherung

6.2.1 Datenminimierung

Es DÜRFEN nur die Daten eingebracht werden, die für den freigegebenen Zweck erforderlich sind. Pseudonymisierung, Maskierung oder Aggregation MUSS geprüft und vorrangig umgesetzt werden.

Nachweis:

6.2.2 Kein Training oder Fine-Tuning mit Echtdaten

Besonders schützenswerte Daten DÜRFEN NICHT für Training, Fine-Tuning, Embedding-Aufbau, Evaluation oder Benchmarking genutzt werden, sofern hierfür keine explizite Datenschutz-, Rechts- und Sicherheitsfreigabe vorliegt.

Nachweis:

6.2.3 Prompt-, Chat- und API-Logging

Die Protokollierung innerhalb der AI Model Hub MUSS auf das erforderliche Maß beschränkt werden.

Vollständige Prompts, Antworten, Dateien, Chatverläufe oder API-Payloads DÜRFEN nur protokolliert werden, wenn dies fachlich, betrieblich, sicherheitstechnisch oder regulatorisch erforderlich ist.

Für die Plattform MÜSSEN geeignete Loglevel definiert werden. Die Schutzmaßnahmen MÜSSEN sich höchst möglichen Detailgrad orientieren. Werden sensitive Inhalte protokolliert, MÜSSEN Zugriff, Zweck, Aufbewahrung, Löschung und Schutzmaßnahmen geregelt sein.

Nachweis:

6.2.4 Schutz von Embeddings und Vektordatenbanken

Embeddings und Vektordatenbanken MÜSSEN wie die zugrunde liegenden Ausgangsdaten geschützt werden. Eine Einstufung als anonym DARF nur nach belastbarer Prüfung erfolgen.

Nachweis:

6.2.5 Löschung und Aufbewahrung

Für Prompts, Antworten, hochgeladene Dateien, temporäre Dateien, Caches, Embeddings, Logs, Modellartefakte und Backups MÜSSEN Lösch- und Aufbewahrungsfristen definiert und technisch umgesetzt werden.

Nachweis:

6.3 Zugriffsschutz und Mandantentrennung

6.3.1 Application- und Use-Case-Trennung

Use Cases mit Besonders schützenswerte Daten MÜSSEN von anderen Workloads getrennt werden. Die Trennung MUSS mindestens auf Namespace-, Berechtigungs-, Secret- und Logging-Ebene erfolgen. Bei hohem Schutzbedarf SOLLTE ein eigener Cluster geprüft werden.

Nachweis:

6.3.2 Zugriff nur für berechtigte Rollen

Zugriff auf Sozialdaten, Personaldaten, Prompts, Outputs, Logs, Backups, Embeddings und Modellkontexte DARF nur berechtigten Rollen gewährt werden. Administrative Zugriffe MÜSSEN protokolliert und regelmäßig überprüft werden.

Nachweis:

6.3.3 Schutz vor Datenabfluss

Für Use Cases mit Sozialdaten oder Personaldaten MÜSSEN Maßnahmen gegen Datenabfluss umgesetzt werden. Dazu gehören kontrollierter Egress, API-Rate-Limits, DLP-Prüfungen soweit technisch möglich, Output-Kontrollen, Download-Beschränkungen und Monitoring auffälliger Abfragen.

Nachweis:

6.4 KI-spezifische Schutzmaßnahmen

6.4.1 Schutz vor Prompt Injection

Wenn LLMs auf interne Dokumente, RAG-Datenbestände, Tools oder Schnittstellen zugreifen, MÜSSEN Schutzmaßnahmen gegen Prompt Injection und indirekte Prompt Injection umgesetzt werden. Externe oder nutzergelieferte Inhalte DÜRFEN keine Systemanweisungen oder Berechtigungen überschreiben.

Nachweis:

6.4.2 Tool- und API-Aufrufe durch LLMs

LLM-gestützte Tool-Aufrufe DÜRFEN nur für freigegebene Funktionen möglich sein. Jeder Tool-Aufruf MUSS autorisiert, protokolliert und auf den fachlichen Zweck begrenzt werden. Schreibende oder auslösende Aktionen MÜSSEN gesondert freigegeben werden.

Nachweis:

6.4.3 Human Review bei sensiblen Ergebnissen

Bei Verarbeitung von Sozialdaten oder Personaldaten MÜSSEN KI-Ergebnisse, die für fachliche Entscheidungen, Kommunikation mit Betroffenen oder interne Bewertungen genutzt werden, durch fachlich zuständige Personen geprüft werden.

Nachweis:

---

7. Anforderungen für KRITIS

Die folgenden Anforderungen gelten zusätzlich zu den Basisanforderungen, wenn das AI Model Hub für IT-Services genutzt wird, die für die Erbringung kritischer Dienstleistungen erforderlich sind oder diese unterstützen. Dies umfasst auch vorgelagerte, nachgelagerte und unterstützende Systeme, wenn eine Störung die kritische Dienstleistung beeinträchtigen kann.

7.1 KRITIS-Geltungsbereich und Risikomanagement

7.1.1 KRITIS-Scope-Bewertung

Vor Nutzung für KRITIS-nahe oder KRITIS-relevante IT-Services MUSS bewertet und dokumentiert werden, ob die Plattform, einzelne Workloads, Schnittstellen, Datenbestände, Abhängigkeiten oder Betriebsprozesse Bestandteil des KRITIS-Anwendungsbereichs und somit der kritischen Dienstleistungen sind.

Nachweis: Eingliederung in den SBK KRITIS Scope

7.1.2 Risikobewertung für kritische Dienstleistungen

Für KRITIS-relevante Nutzung MUSS eine Risikobewertung durchgeführt werden, die Auswirkungen auf Verfügbarkeit, Integrität, Authentizität und Vertraulichkeit der kritischen Dienstleistung betrachtet. Es sind reale Risiken zu identifizieren und Risikominimierungsmaßnahmen abzuleiten. Die Maßnahmen zur Risikomitigierung sind nach einem definierten Umsetzungsplan einzuführen.

Nachweis: Risikomatrix, Behandlungsplan

7.1.3 Abhängigkeiten und Single Points of Failure

Abhängigkeiten (u.a. von Housing-Anbieter, Netzwerk, VPN, Strom, Kühlung, GPU-Hardware, Kubernetes-Control-Plane, Registry, DNS, IAM, SIEM, Backup, Middleware und Fachverfahren) MÜSSEN in Bezug auf deren Kritikalität zum Leistungsgegenstand analysiert werden. Single Points of Failure MÜSSEN vermieden oder risikobasiert behandelt werden. Dies gilt insbesondere im BCM und IT-SCM Kontext.

Nachweis: BIA

7.1.4 Asset Management

Die im KIH eingesetzten Assets (u. a. Hardware, Software, Anwendungen, Daten) MÜSSEN in Bezug auf deren Kritikalität für die kritischen Dienstleistungen identifiziert, inventarisiert, klassifiziert und dokumentiert werden. Dabei müssen die Abhängigkeiten / Beziehungen zwischen den kDL-Prozessen bis hin zu den unterstützenden Systemen und weiteren IT-Schichten transparent darstellt werden.

Nachweis: Asset Verzeichnisse (z.B. CMDB, UAMNavi)

7.2 Verfügbarkeit, Resilienz und Notfallbetrieb

7.2.1 BIA und Wiederherstellungsziele

Für KRITIS-relevante Workloads MUSS eine Business-Impact-Analyse oder gleichwertige Bewertung in der Zusammenarbeit mit BCM und IT-SCM vorliegen. RTO, RPO, MTPD und Prioritäten für Wiederanlauf MÜSSEN definiert sein. Die Wiederanlaufplanung des KIH ist in den Gesamtwiederanlaufplan der SBK zu integrieren.

Nachweis: BIA, Wiederherstellungspläne, Wiederanlaufpläne, Notbetrieb, Notfallbeseitigungsplan

7.2.2 Hochverfügbarkeit und Redundanz

Die Plattformarchitektur MUSS den Verfügbarkeitsanforderungen der kritischen Dienstleistung entsprechen. Redundanz von Netzwerk, VPN, Firewalls, Stromversorgung, Storage, Control Plane und kritischen Services MUSS geprüft und risikobasiert umgesetzt werden.

Dabei MUSS auch bewertet werden, ob die Verfügbarkeits-, Wiederanlauf- und Wiederherstellungsanforderungen mit einem einzelnen Housing-Standort erfüllt werden können.

Sofern dies nicht angemessen möglich ist, MUSS eine geeignete Standortredundanz, Georedundanz oder eine gleichwertige Ersatzlösung vorgesehen werden.

Die Entscheidung für oder gegen Standortredundanz beziehungsweise Georedundanz MUSS dokumentiert, risikobewertet und mit den Vorgaben aus BIA, RTO, RPO, MTPD und Disaster-Recovery-Plan abgestimmt werden.

Für die Umsetzung zu beachten:

- BSI Kriterien für die Standortwahl von Rechenzentren

Nachweis:

7.2.3 Melde- und Eskalationswege

Eskalations- und Meldewege für sicherheitsrelevante Ereignisse, Störungen und Notfälle MÜSSEN definiert, dokumentiert und etabliert sein. Die Alarmierungsprozesse MÜSSEN klare Zuständigkeiten, Kommunikationswege, Eskalationsstufen sowie zeitliche Anforderungen für die Bearbeitung und Weiterleitung von Ereignissen enthalten. Relevante interne und externe Beteiligte (u. a. Betrieb, IT, Informationssicherheit, BCM, Dienstleister und zuständige Meldestellen) MÜSSEN in Abhängigkeit von Kritikalität und Auswirkung eines Ereignisses berücksichtigt werden.

Nachweis: Alarmierungs- und Kommunikationsplan

7.2.4 Notfallhandbuch und Disaster-Recovery-Plan

Für KRITIS-relevante Nutzung MUSS ein Notfallhandbuch und ein Disaster-Recovery-Plan vorhanden sein. Diese MÜSSEN konkrete Handlungsanweisungen für Ausfall, Kompromittierung, Datenverlust, Ransomware, VPN-Ausfall, Cluster-Ausfall und Housing-Störung enthalten. Im Notfallhandbuch sind für alle relevanten Komponenten die Wiederanlauf- sowie die Wiederherstellungspläne zu erstellen. Die relevanten Schnittstellen, Verantwortlichen und Dienstleister MÜSSEN in die Notfallbehandlung eingebunden werden.

Nachweis: BIA, Wiederherstellungspläne, Wiederanlaufpläne, Notbetrieb, Notfallbeseitigungsplan

7.2.5 Regelmäßige Notfallübungen

Wiederanlauf, Restore, Isolation kompromittierter Systeme und Kommunikation (Wirksamkeit der Eskalations- und Meldewege) im Sicherheitsvorfall MÜSSEN regelmäßig geübt werden. Erkenntnisse MÜSSEN dokumentiert und in Verbesserungsmaßnahmen überführt werden.

Nachweis: Übungsplan im Rahmen des BCM / IT-SCM

7.2.6 Kapazitäts- und Lastmanagement

Für KRITIS-relevante Workloads MUSS sichergestellt sein, dass GPU-, CPU-, RAM-, Storage-, Netzwerk- und API-Kapazitäten überwacht und ausreichend dimensioniert sind. Überlastsituationen DÜRFEN nicht unkontrolliert auf kritische Dienstleistungen durchschlagen.

Nachweis: Betriebsdokumentation, Kapazitätsplanung und -überwachung

---

7.3 Systeme zur Angriffserkennung

7.3.1 Vollständige SzA-Abdeckung

Für KRITIS-relevante Nutzung MÜSSEN Systeme zur Angriffserkennung die maßgeblichen informationstechnischen Systeme, Komponenten und Prozesse abdecken. Dazu gehören Housing-Anbindung, VPN, Firewalls, Linux-Hosts, Kubernetes, API-Gateways, Registry, IAM, PAM, Middleware-Schnittstellen und relevante Workloads.

Technische und organisatorische Maßnahmen zum Schutz vor Schadsoftware MÜSSEN dabei auch etabliert, betrieben und regelmäßig überprüft werden.

Nachweis: Sicherheitssystem-Architektur, SzA Konzeption, SzA Systemeinstellungen

7.3.2 Protokollierung, Detektion und Reaktion

Die Angriffserkennung MUSS Protokollierung, Detektion und Reaktion umfassen. Es ist festzulegen und zu dokumentieren, welche geeigneten Parameter und Merkmale aus dem laufenden Betrieb kontinuierlich und automatisiert erfasst und ausgewertet werden.

Nachweis: SzA Konzeption, SzA Systemeinstellungen

7.3.3 24/7-Reaktionsfähigkeit

Für KRITIS-relevante Sicherheitsereignisse MUSS eine angemessene Reaktionsfähigkeit sichergestellt werden. Rufbereitschaften, Eskalationswege und Entscheidungsbefugnisse MÜSSEN dokumentiert und getestet sein.

Nachweis: Abgestufter Alarmierungs- und Kommunikationsplan

7.3.4 Bedrohungsinformationen

Informationen zu aktuellen Angriffsmustern, Schwachstellen, Exploits und Herstellerwarnungen MÜSSEN fortlaufend eingeholt, bewertet und in Schutzmaßnahmen überführt werden.

Nachweis: z.B. über die Verwendung der täglichen BSI Schwachstellenreports

7.4 Erhöhte Prüf- und Nachweisanforderungen

7.4.1 Regelmäßige Sicherheitsprüfungen

KRITIS-relevante Plattformbestandteile MÜSSEN regelmäßig durch technische Sicherheitsprüfungen überprüft werden. Dazu gehören Schwachstellenscans, Konfigurationsprüfungen, Kubernetes-Sicherheitsprüfungen, API-Sicherheitstests und bei erhöhtem Risiko Penetrationstests.

Es müssen für den laufenden Betrieb Kennzahlen zur Informationssicherheit definiert und in regelmäßigen Abständen überwacht werden. Die Überprüfung kann in den SBK IKS Prozess integriert werden.

Nachweis: Planung, Durchführung und Dokumentation der Sicherheitstests

7.4.2 Angriffssimulation für exponierte oder kritische Services

Für besonders kritische oder exponierte KI-Services SOLLTE eine Angriffssimulation oder ein Red-Team-Test durchgeführt werden. Der Umfang MUSS risikobasiert festgelegt werden.

Nachweis: Planung, Durchführung und Dokumentation der Sicherheitstests

7.4.3 Auditierbarkeit

Die Umsetzung der KRITIS-Anforderungen MUSS prüfbar sein. Relevante Konzepte, Nachweise, Protokolle, Risikoentscheidungen, Ausnahmen und Maßnahmenstände MÜSSEN versioniert und geordnet bereitgestellt werden können.

Nachweis: Arbeitsanweissungen / Vorgaben, operative Betriebsdokumentation, Konfigurationseinstellungen, relevante LogFiles etc.

7.4.4 Ausnahmen nur mit Risikogenehmigung

Abweichungen von diesen Anforderungen DÜRFEN nur dokumentiert, risikobewertet, kompensiert, befristet und durch eine autorisierte Stelle genehmigt werden.

Nachweis: dokumentierte Bewertungen und Freigaben

8. Mindestnachweise für die Sicherheitsabnahme

Für die produktive Freigabe der AI Model Hub MÜSSEN die für den jeweiligen Anwendungsfall relevanten Anforderungen konzeptionell beschrieben, nachvollziehbar umgesetzt und dokumentiert sein.

Die Mindestnachweise dienen dazu, die Umsetzung der Anforderungen aus dieser Vorgabe nachvollziehbar zu machen. Sie MÜSSEN zeigen, wie die Anforderungen in der Zielarchitektur, im Betrieb, in den Sicherheitsmaßnahmen, in den Dienstleisterbeziehungen und in der technischen Umsetzung berücksichtigt wurden.

8.1 AI Model-Hub-Betriebskonzept

Das AI Model-Hub-Betriebskonzept MUSS beschreiben, wie das AI Model Hub betrieben werden soll. Es MUSS insbesondere darstellen, wie die betrieblichen Anforderungen aus dieser Vorgabe umgesetzt werden. Dazu gehören insbesondere Betriebsmodell, Zuständigkeiten, Betriebsprozesse, Administration, Wartung, Monitoring, Patchmanagement, Datensicherung, Wiederherstellung, Störungsbearbeitung, Dienstleistereinbindung und Notfallverfahren.

Das Betriebskonzept beschreibt damit die geplante betriebliche Umsetzung der Anforderungen. Detailregelungen dürfen in mitgeltenden Dokumenten beschrieben werden, sofern diese eindeutig referenziert werden.

8.2 AI Model-Hub-Sicherheitskonzept

Das AI Model-Hub-Sicherheitskonzept MUSS beschreiben, wie die Sicherheitsanforderungen an das AI Model Hub angemessen und nach dem Stand der Technik umgesetzt werden sollen.

Es MUSS insbesondere darstellen, wie die Anforderungen aus dieser Vorgabe in der Architektur, im Schutzbedarf, im Risikomanagement, im Identitäts- und Berechtigungsmanagement, in der Netzwerk- und Kommunikationssicherheit, in der Plattformabsicherung, in der Protokollierung, in der Angriffserkennung, im Incident Management, im Datenschutz, in der Lieferkette sowie bei der sicheren Außerbetriebnahme berücksichtigt werden.

Das Sicherheitskonzept beschreibt damit die geplante sicherheitstechnische Umsetzung der Anforderungen. Es MUSS vor der produktiven Inbetriebnahme durch Informationssicherheit beziehungsweise IT-Sicherheit bewertet werden. Bei der Verarbeitung von Sozialdaten oder Personaldaten MUSS zusätzlich der Datenschutz beteiligt werden. Bei einer Nutzung für kritische Dienstleistungen MÜSSEN die KRITIS-relevanten Anforderungen und Risiken ausdrücklich berücksichtigt werden.

8.3 Ergänzende Architekturdokumentation

Sicherheitsrelevante Architekturentscheidungen MÜSSEN nachvollziehbar dokumentiert werden. Dies kann insbesondere in Form von Architectural Decision Records oder einer vergleichbaren Entscheidungsdokumentation erfolgen.

Die Architekturdokumentation MUSS deutlich machen, welche Entscheidung getroffen wurde, welche Anforderung oder welches Risiko damit adressiert wird, welche Alternativen betrachtet wurden und warum die gewählte Lösung ausgewählt wurde.

Dies betrifft insbesondere Entscheidungen, durch die Anforderungen dieser Vorgabe konkret erfüllt werden, zum Beispiel zur Netzwerkarchitektur, VPN-Anbindung, Proxy-Nutzung, Kubernetes-Architektur, Mandantentrennung, Administrationszugängen, Protokollierung, Schlüssel- und Secret-Management, Modellbereitstellung sowie zu Schnittstellen zur Middleware.

8.4 Betriebsdokumentation

Die Betriebsdokumentation MUSS die tatsächliche Umsetzung der konzipierten Lösungen nachvollziehbar abbilden.

Sie MUSS den produktiven Aufbau, die produktiven Konfigurationen, die relevanten Schnittstellen, die Betriebsverfahren und die Zuständigkeiten so dokumentieren, dass Betrieb, Administration, Fehleranalyse, Sicherheitsprüfung, Wiederherstellung und Notfallbearbeitung nachvollzogen werden können.

Soweit Informationen in anderen Systemen oder Dokumenten gepflegt werden, zum Beispiel CMDB, Netzplan, Firewall-Dokumentation, Kubernetes-Dokumentation, Monitoring-System, Secret-Management, Ticketsystem oder Vertragsunterlagen, MUSS die Betriebsdokumentation eindeutig darauf verweisen.

8.5 Vertrags- und Dienstleisterdokumentation

Soweit externe Dienstleister in Aufbau, Betrieb, Wartung, Housing, Helping Hands, Fernwartung oder Support eingebunden sind, MÜSSEN die relevanten Verantwortlichkeiten, Schnittstellen, Sicherheitsanforderungen, Leistungsgrenzen, Meldewege und Kontrollmöglichkeiten vertraglich oder in geeigneten Dienstleisterunterlagen geregelt sein.

Die Dokumentation MUSS nachvollziehbar machen, welche Leistungen durch interne Stellen erbracht werden, welche Leistungen durch externe Dienstleister erbracht werden und wie die Einhaltung der Sicherheitsanforderungen überprüft werden kann.

9. Ausnahmen

Grundsätzlich sind Abweichungen von den AI Model-Hub Sicherheitsanforderungen nur in begründeten Ausnahmefällen zulässig.

- Für jede Ausnahme MUSS dokumentiert werden, welche Anforderung betroffen ist und aus welchen fachlichen, technischen, betrieblichen oder wirtschaftlichen Gründen die Anforderung nicht oder nicht vollständig umgesetzt werden kann.
- Für jede Ausnahme MÜSSEN geeignete Kompensations- oder Mitigationsmaßnahmen definiert und umgesetzt werden, um die verbleibenden Risiken angemessen zu reduzieren.
- Alle Ausnahmen MÜSSEN zentral dokumentiert und nachvollziehbar nachverfolgt werden. Hierzu MUSS ein zentrales Ausnahmeregister geführt werden.
- Für jede Ausnahme MUSS eine Bewertung der daraus resultierenden Risiken erfolgen. Die Bewertung ist durch den Informationssicherheitsbeauftragten (ISB) und IT-Sicherheit durchzuführen. Dabei MUSS dokumentiert werden, ob die Ausnahme aus Sicht der Informationssicherheit empfohlen werden kann oder nicht.

Das Ausnahmeregister MUSS mindestens folgende Informationen enthalten:

- betroffene Anforderung
- Beschreibung und Begründung der Ausnahme
- definierte Kompensations- bzw. Mitigationsmaßnahmen
- verantwortliche Stelle
- Datum der Erfassung und letzte Überprüfung
- Bewertung der Risiken durch,
    - Informationssicherheitsbeauftragten (ISB)
    - IT-Sicherheit
    - bei Bedarf Datenschutzbeauftragter (DSB)

Ausnahmen MÜSSEN regelmäßig überprüft werden (min. einmal im Jahr), um festzustellen, ob die Gründe für die Abweichung weiterhin bestehen oder die Anforderung zwischenzeitlich umgesetzt werden kann.

---

10. Abnahme

Das AI Model Hub gilt aus Sicht der Informationssicherheit/IT-Sicherheit nur dann als freigabefähig, wenn die für den jeweiligen Anwendungsfall geltenden Anforderungen nachvollziehbar berücksichtigt, geplant, umgesetzt oder wirksam kompensiert und dokumentiert wurden.

Die Sicherheitsabnahme erfolgt auf Grundlage der jeweils anwendbaren Anforderungsklasse:

- **Basisanforderungen** gelten für jede Nutzung der AI Model Hub.
- **Anforderungen für Besonders schützenswerte Daten** gelten zusätzlich, sobald entsprechende Daten verarbeitet werden.
- **Anforderungen für KRITIS** gelten zusätzlich, sobald die Plattform für IT-Services eingesetzt wird, die zur Erbringung einer kritischen Dienstleistung erforderlich sind.

Für die Freigabe MÜSSEN die jeweils geltenden Anforderungen in den relevanten Konzepten, Verträgen und Betriebsdokumentationen nachweisbar erfüllt sein.

Eine Anforderung gilt als erfüllt, wenn

- die geplante Umsetzung im jeweiligen Konzept beschrieben ist,
- sicherheitsrelevante Architekturentscheidungen nachvollziehbar dokumentiert sind,
- die technische oder organisatorische Umsetzung nachvollziehbar dokumentiert ist,
- die erforderlichen Nachweise vorliegen,
- die Zuständigkeiten und Betriebsprozesse geregelt sind,
- keine offenen Punkte bestehen, die einer sicheren produktiven Nutzung entgegenstehen,
- und Informationssicherheitsbeauftragten (ISB) und IT-Sicherheit bestätigt haben, dass die konzipierten Lösungen
    - die jeweiligen Anforderungen tatsächlich erfüllen und
    - dem Stand der Technik entsprechen.

Kann eine Anforderung zum Zeitpunkt der Abnahme nicht vollständig erfüllt werden, MUSS dies dokumentiert, begründet und bewertet werden (siehe 9. Ausnahmen).

Die Sicherheitsabnahme ersetzt nicht die fachliche, technische, datenschutzrechtliche oder betriebliche Freigabe. Sie bestätigt ausschließlich, dass die aus Sicht der Informationssicherheit beziehungsweise IT-Sicherheit relevanten Anforderungen für den vorgesehenen Use Case angemessen berücksichtigt und nachweisbar behandelt wurden.

---

11. Quellenbasis

Diese Vorgabe wurde inhaltlich an folgenden Dokumenten und Themenfeldern ausgerichtet:

- Sicherheitsstandards für Systeme im BITMARCK-Verbund
- Mindeststandards des BSI nach § 44 Abs. 1 Satz 1 BSIG
- BSI IT-Grundschutz-Kompendium 2023 / BSI Grundschutz++ (GS++)
- BSI-Standard 200-1 Managementsysteme für Informationssicherheit (ISMS)
- BSI-Standard 200-2 IT-Grundschutz-Methodik

- BSI-Standard 200-3 Risikomanagement
- BSI-Standard 200-4 Business Continuity Management
- B3S-GKV/PV
    - Orientierungshilfe zum Einsatz von Systemen zur Angriffserkennung
    - Orientierungshilfe zu Inhalten und Anforderungen an branchenspezifische Sicherheitsstandards (B3S)
    - Konkretisierung der Anforderungen an die gemäß § 8a Absatz 1 und Absatz 1a BSIG
    - Anforderungen nach § 8a Absatz 5 BSIG
- BSI C5:2020 / C5:2026
- OWASP ASVS/LLM TOP 10
- EU KI-Verordnung
- ISO 27001
- ISO 42001