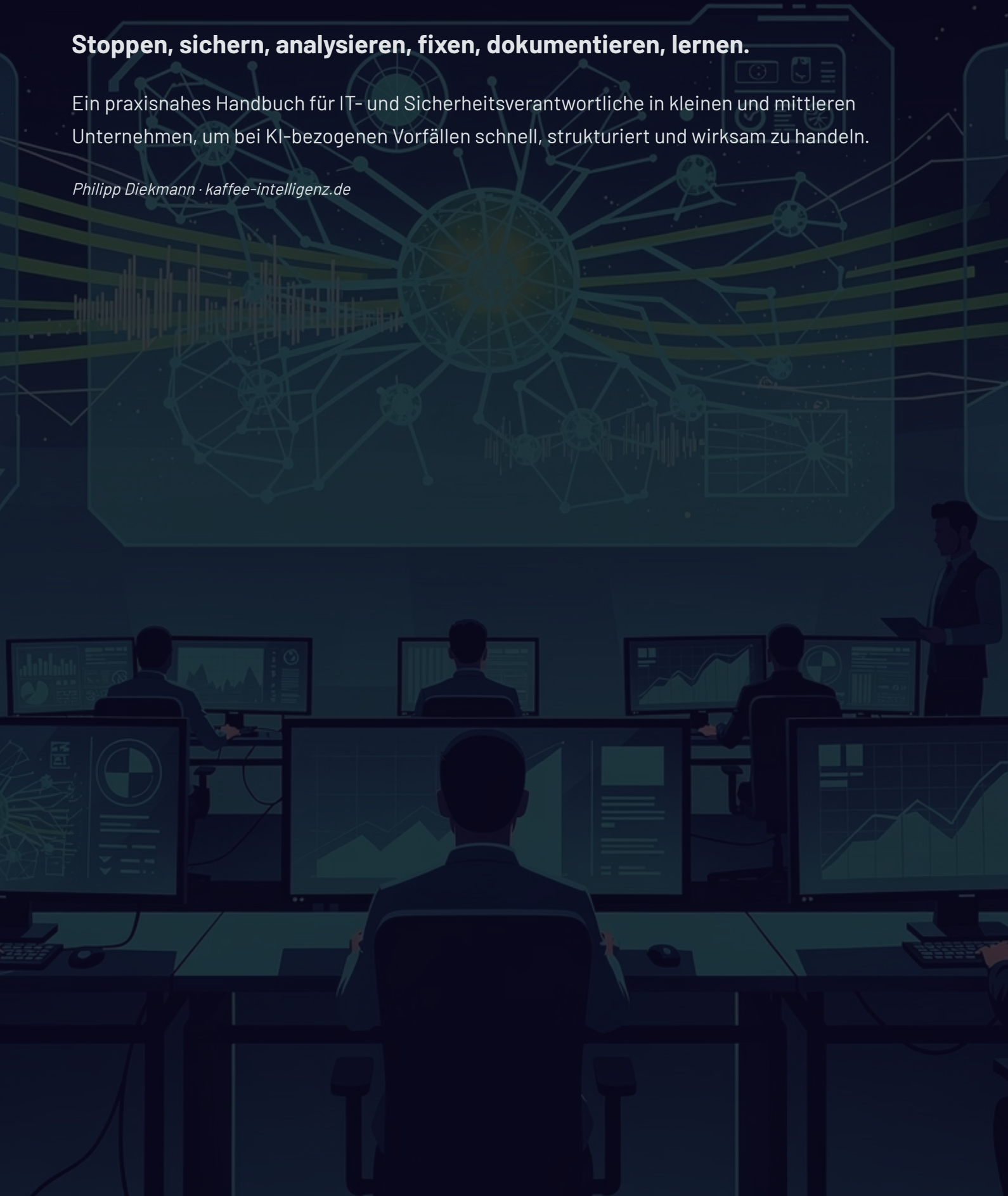


Runbook – KI Incident Response (KMU-Edition)

Stoppen, sichern, analysieren, fixen, dokumentieren, lernen.

Ein praxisnahes Handbuch für IT- und Sicherheitsverantwortliche in kleinen und mittleren Unternehmen, um bei KI-bezogenen Vorfällen schnell, strukturiert und wirksam zu handeln.

Philipp Diekmann · kaffee-intelligenz.de



Ziel & Definition of Done

„Schnell reagieren, Schäden begrenzen, Ursachen beheben.“

Dieses Runbook verfolgt ein klares Ziel: Jeder KI-bezogener Vorfall in Ihrem Unternehmen soll innerhalb definierter Zeitfenster erkannt, eingedämmt und behoben werden. Dabei geht es nicht nur um technische Reaktion, sondern auch um organisatorische Klarheit, rechtliche Absicherung und systematisches Lernen aus jedem Vorfall. Die hier beschriebenen Prozesse sind bewusst für KMU-Strukturen konzipiert – schlank, pragmatisch und sofort umsetzbar, ohne dass ein eigenes Security Operations Center (SOC) erforderlich ist.

Definition of Done (DoD)

Ein Incident gilt erst dann als abgeschlossen, wenn **alle vier Kriterien** vollständig erfüllt sind. Die DoD dient als verbindliche Checkliste für den Incident Lead und verhindert, dass Vorfälle vorzeitig als erledigt markiert werden.

01

Incident klassifiziert

Typ, Schweregrad (1–5) und betroffene Systeme sind dokumentiert. Die Klassifikation folgt der Incident-Typ-Matrix auf Seite 3.

02

Maßnahmen durchgeführt

Sofortmaßnahmen (Stop & Sichern) sowie dauerhafte Fixes sind implementiert und verifiziert. Kein offener Workaround ohne Termin.

03

Kommunikation erfolgt

Alle relevanten Stakeholder – intern wie extern – sind informiert. Meldepflichten (DSGVO Art. 33/34) wurden geprüft und ggf. erfüllt.

04

Postmortem + Prävention umgesetzt

Root-Cause-Analyse ist abgeschlossen, Lessons Learned dokumentiert, und präventive Maßnahmen mit Owner und Termin hinterlegt.



Hinweis: Bewahren Sie dieses Runbook sowohl digital als auch ausgedruckt auf. Im Ernstfall muss der Zugriff auch bei Systemausfall gewährleistet sein.

Incident-Typen im Überblick

KI-Systeme können auf vielfältige Weise versagen oder kompromittiert werden. Die folgende Klassifikation umfasst die **sechs häufigsten Incident-Typen**, die in KMU-Umgebungen auftreten. Jeder Vorfall muss bei der Erstmeldung einem dieser Typen zugeordnet werden, da sich daraus Eskalationspfade, Checklisten und Kommunikationspflichten ableiten.



Datenschutz

Datenabfluss an Dritte, falsche Empfänger, unbeabsichtigte Speicherung personenbezogener Daten durch KI-Systeme. **Meldepflicht prüfen!**



Security

Kompromittierte Accounts, geleakte API-Keys, Prompt-Injection-Angriffe, unautorisierter Zugriff auf KI-Endpoints oder Trainingsdaten.



Qualität

Halluzinationen, faktisch falsche Aussagen, inkonsistente Antworten, fehlerhafte Klassifikationen oder Empfehlungen durch das Modell.



Bias / Fairness

Diskriminierende, voreingenommene oder ethisch problematische Outputs – z. B. bei Bewerberscreening, Kreditentscheidungen, Kundenansprache.




Compliance

Fehlender Transparenzhinweis auf KI-Nutzung, fehlende Freigabe durch Datenschutz/Rechtsabteilung, Verstoß gegen interne Policies oder EU AI Act.



Verfügbarkeit

System-Downtime, API-Timeouts, Provider-Ausfälle, Rate-Limiting oder Performance-Degradation mit Auswirkung auf Geschäftsprozesse.

 **Tipp:** Ein einzelner Vorfall kann mehreren Typen zugeordnet werden (z. B. Security + Datenschutz). In diesem Fall gelten **alle** zugehörigen Checklisten parallel.

SOP: Der 6-Schritte-Prozess

Das Standard Operating Procedure (SOP) für jeden KI-Incident folgt einem festen 6-Schritte-Ablauf. Die Reihenfolge ist verbindlich – kein Schritt darf übersprungen werden. Jeder Schritt hat ein klares Ziel und definierte Outputs, die vor dem Weitergehen erfüllt sein müssen.



Der Prozess ist bewusst linear aufgebaut: Erst wenn das betroffene System gestoppt und alle Beweise gesichert sind, beginnt die Analyse. Vorschnelle Fixes ohne Beweissicherung sind einer der häufigsten Fehler in der Incident Response und können sowohl die forensische Aufklärung als auch die rechtliche Dokumentation gefährden.

Sofortmaßnahmen (≤ 30 Min.)

- **STOP:** System pausieren, Feature-Flag deaktivieren, API-Endpoint sperren
- **SICHERN:** Logs exportieren, betroffene Inputs/Outputs speichern, Screenshots anfertigen, exakte Zeitpunkte notieren

Analyse & Behebung (≤ 4 Std.)

- **ANALYSIEREN:** Ursache identifizieren – liegt es am Prompt, Modell, den Daten, einer Policy oder den Zugriffsrechten?
- **FIXEN:** Patch einspielen, Guardrails anpassen, kompromittierte Daten entfernen, Zugänge sperren/rotieren

Nachbereitung (≤ 5 Werkzeuge)

- **KOMMUNIZIEREN:** Internes Stakeholder-Update versenden, bei Datenschutz-Incidents ggf. Betroffene und Aufsichtsbehörde informieren (72-Stunden-Frist gemäß DSGVO Art. 33)
- **LERNEN:** Postmortem durchführen, präventive Maßnahmen definieren, Monitoring-Regeln anpassen, Runbook aktualisieren

Rollen & Eskalation

Klare Rollenverteilung ist im Ernstfall entscheidend. Jede Person muss **vor** einem Incident wissen, welche Verantwortung sie trägt. Die folgende Mini-RACI-Matrix zeigt die Zuständigkeiten je Prozessschritt. Darunter finden Sie die Kontaktkette mit Platzhaltern zum Ausfüllen.

RACI-Matrix

Schritt	Incident Lead	AI Owner	IT/Security	Datenschutz	Fachbereich	GF
1 - STOP	R/A	C	R	I	I	I
2 - SICHERN	A	R	R	C	I	I
3 - ANALYSIEREN	A	R	R	C	C	I
4 - FIXEN	A	R	R	C	I	I
5 - KOMMUNIZIEREN	R/A	C	I	R	C	A
6 - LERNEN	R/A	R	C	C	C	I

R = Responsible (führt aus) · A = Accountable (verantwortet) · C = Consulted (wird befragt) · I = Informed (wird informiert)

Kontaktkette (bitte ausfüllen)

Incident Lead Name: _____ Tel: _____ E-Mail: _____	Datenschutzbeauftragte/r Name: _____ Tel: _____ E-Mail: _____	IT / Security Name: _____ Tel: _____ E-Mail: _____	Geschäftsführung Name: _____ Tel: _____ E-Mail: _____
--	---	--	---

Eskalationsregel: Bei Schweregrad ≥ 3 wird die Geschäftsführung sofort informiert. Bei Datenschutz-Incidents wird der DSB **immer** hinzugezogen – unabhängig vom Schweregrad.

Checkliste: Datenschutz-Incident

🛡️ DATENSCHUTZ

MELDEPFLICHT BEACHTEN

Diese Checkliste wird aktiviert, sobald ein KI-bezogener Vorfall personenbezogene Daten betrifft – sei es durch unbeabsichtigten Datenabfluss, falsche Empfänger, unerlaubte Speicherung oder die Verarbeitung sensibler Daten durch ein KI-Modell. Arbeiten Sie die Punkte **streng in der angegebenen Reihenfolge** ab. Jeder Punkt muss vom Incident Lead mit Zeitstempel abgezeichnet werden.

1 Zugriff sofort sperren

Betroffenes KI-System, API-Endpoint oder Feature deaktivieren. Keine weiteren Daten dürfen das System erreichen oder verlassen.

2 Betroffene Daten identifizieren

Art, Umfang und Kategorien der betroffenen personenbezogenen Daten erfassen (Name, E-Mail, Gesundheitsdaten, etc.).

3 Betroffene Personen ermitteln

Anzahl und Kreise der betroffenen Personen bestimmen (Kunden, Mitarbeitende, Dritte).

4 Beweise sichern

Logs, Inputs/Outputs, API-Calls, Zeitstempel und Screenshots exportieren und unveränderbar ablegen.

5 DSB / Datenschutzbeauftragte(n) informieren

Sofortige Benachrichtigung mit allen bisher bekannten Fakten. Gemeinsame Risikobewertung einleiten.

6 Meldepflicht prüfen (72-Stunden-Frist)

Gemäß DSGVO Art. 33: Muss die Aufsichtsbehörde informiert werden? Frist beginnt ab Kenntnis des Vorfalls.

7 AVV / Anbieter informieren

Auftragsverarbeiter (z. B. OpenAI, Azure, etc.) gemäß AVV-Vertrag über den Vorfall benachrichtigen.

8 Betroffene Personen benachrichtigen

Bei hohem Risiko (Art. 34 DSGVO): Betroffene transparent, verständlich und zeitnah informieren.

9 Dokumentation anlegen

Vollständigen Incident-Report erstellen: Chronologie, Ursache, ergriffene Maßnahmen, Risikobewertung.

10 Daten löschen / Zugriff dauerhaft entziehen

Betroffene Daten beim Anbieter und in eigenen Systemen löschen. Zugriffsbeschränkungen dauerhaft umsetzen und verifizieren.

🚨 **Kritisch:** Die 72-Stunden-Frist für die Meldung an die Aufsichtsbehörde beginnt mit dem Zeitpunkt der Kenntnisnahme – nicht mit Abschluss der Analyse. Im Zweifel **vorsorglich melden** und nachträglich ergänzen.

Checkliste: Qualitäts-Incident

QUALITÄT

HALLUZINATIONEN & FEHLER

Diese Checkliste kommt zum Einsatz, wenn ein KI-System fehlerhafte, halluzinierte oder irreführende Outputs produziert – etwa falsche Fakten in der Kundenberatung, erfundene Quellenangaben oder inkonsistente Empfehlungen. Qualitäts-Incidents können Reputationsschäden verursachen und müssen ebenso systematisch behandelt werden wie Security-Vorfälle.

1 Beweisbeispiele sammeln

Mindestens 5 konkrete Beispiele des fehlerhaften Outputs dokumentieren – mit exaktem Input, Output und erwartetem Ergebnis.

2 System pausieren / Warnhinweis schalten

Feature deaktivieren oder sichtbaren Disclaimer für Endnutzer einblenden: „Antworten werden derzeit überprüft.“

3 Eval-Set / Testdaten laufen lassen

Vorhandenes Evaluations-Set gegen das betroffene Modell/Prompt ausführen, um die Fehlerquote systematisch zu messen.

4 Root Cause eingrenzen

Ist die Ursache im Prompt, im Modell (Update/Drift), in den Retrieval-Daten (RAG) oder in der Systemkonfiguration?

5 Prompt-Guardrails ergänzen / anpassen

System-Prompt verschärfen, Output-Validierungsregeln hinzufügen, Temperatur reduzieren, Kontext einschränken.

6 Eskalation zu Mensch aktivieren

Human-in-the-Loop-Mechanismus einschalten: Kritische Outputs müssen manuell geprüft werden, bevor sie den Nutzer erreichen.

7 Retrieval-Datenbank / Knowledge Base prüfen

Bei RAG-Systemen: Sind die zugrunde liegenden Dokumente aktuell, korrekt und vollständig? Veraltete Quellen entfernen.

8 Rollback durchführen

Falls ein Prompt- oder Modell-Update die Ursache war: Auf die letzte stabile Version zurückrollen und verifizieren.

9 Erneuten Test durchführen

Eval-Set nach dem Fix erneut laufen lassen. Fehlerrate muss unter dem definierten Schwellenwert liegen, bevor das System wieder aktiviert wird.

10 Monitoring-Regeln verschärfen

Automatische Alerts für ähnliche Fehlermuster einrichten (z. B. Confidence-Score-Schwellenwerte, Output-Pattern-Matching).

Checkliste: Security-Incident

⚠ SECURITY

ACCOUNT / API KOMPROMITTERT

Security-Incidents bei KI-Systemen haben eine besondere Brisanz: Kompromittierte API-Keys können in Minuten zu enormen Kosten führen, Prompt-Injection-Angriffe können Daten exfiltrieren, und unautorisierte Zugriff auf Modell-Endpoints kann vertrauliche Unternehmensdaten offenlegen. Diese Checkliste deckt die häufigsten Security-Szenarien im KI-Kontext ab.

1 API-Keys sofort rotieren

Alle potenziell kompromittierten API-Keys, Tokens und Secrets invalidieren und neue generieren. Alte Keys in allen Systemen ersetzen.

2 Betroffene Accounts sperren

Kompromittierte Benutzerkonten deaktivieren. Passwörter erzwungen zurücksetzen. Sessions invalidieren.

3 MFA aktivieren / überprüfen

Multi-Faktor-Authentifizierung für alle Zugänge zu KI-Plattformen, Dashboards und Admin-Interfaces erzwingen.

4 Logs umfassend prüfen

API-Zugriffslogs, Authentifizierungslogs und Audit-Trails für den betroffenen Zeitraum exportieren und analysieren.

5 Schadensausmaß ermitteln

Welche Daten wurden abgerufen? Welche Aktionen wurden durchgeführt? Gibt es ungewöhnliche Kosten (z. B. API-Usage-Spikes)?

6 Network Rules / IP-Restrictions setzen

API-Zugriff auf bekannte IP-Ranges einschränken. Geo-Blocking aktivieren. Ungewöhnliche Quell-IPs blockieren.

7 Anbieter-Support kontaktieren

Incident beim KI-Anbieter melden (OpenAI, Google, Azure, etc.). Ggf. Account-Sperre oder forensische Unterstützung anfordern.

8 Secrets Management überprüfen

Sind API-Keys sicher gespeichert (Vault, Environment Variables)? Liegen Keys in Code-Repos, .env-Dateien oder Chats?

9 Prompt-Injection-Schutz prüfen

Input-Validierung und Sanitization überprüfen. Sind bekannte Injection-Patterns gefiltert? Testfälle durchlaufen.

10 Härungsmaßnahmen dokumentieren & umsetzen

Alle durchgeführten Maßnahmen dokumentieren. Langfristige Härung planen: Rate-Limiting, Usage-Alerts, regelmäßige Key-Rotation.

Incident Log Template


DOKUMENTATION

Jeder Incident muss lückenlos dokumentiert werden – sowohl für interne Auswertung als auch für den Nachweis gegenüber Aufsichtsbehörden, Kunden und Geschäftsführung. Nutzen Sie die folgende Tabelle als Vorlage. Drucken Sie mehrere Exemplare aus oder führen Sie die Tabelle in einem geschützten Shared Drive (z. B. SharePoint, Confluence).

Datum / Uhrzeit	____/____/____ / ____:____ Uhr
Use Case	_____
Incident-Typ	<input type="checkbox"/> Datenschutz <input type="checkbox"/> Security <input type="checkbox"/> Qualität <input type="checkbox"/> Bias <input type="checkbox"/> Compliance <input type="checkbox"/> Verfügbarkeit
Schweregrad (1-5)	<input type="checkbox"/> 1 (gering) <input type="checkbox"/> 2 <input type="checkbox"/> 3 (mittel) <input type="checkbox"/> 4 <input type="checkbox"/> 5 (kritisch)
Beschreibung	_____
Betroffene Systeme / Personen	_____
Sofortmaßnahme	_____
Root Cause	_____
Fix / Dauerlösung	_____
Kommunikation	<input type="checkbox"/> intern <input type="checkbox"/> Kunden <input type="checkbox"/> Behörde <input type="checkbox"/> Anbieter
Owner	_____
Status	<input type="checkbox"/> offen <input type="checkbox"/> in Bearbeitung <input type="checkbox"/> geschlossen
Postmortem-Datum	_____

Schweregrad-Skala

- 1 – Gering**
Kein Schaden, intern bemerkt, schnell behebbar
- 2 – Niedrig**
Geringe Auswirkung, einzelne Nutzer betroffen
- 3 – Mittel**
Spürbare Auswirkung, mehrere Nutzer, GF informieren
- 4 – Hoch**
Erheblicher Schaden, Meldepflicht wahrscheinlich
- 5 – Kritisch**
Massiver Schaden, sofortige Eskalation, Krisenmodus

 **Hinweis:** Füllen Sie dieses Template innerhalb der ersten 60 Minuten nach Erkennung des Incidents aus. Unvollständige Felder werden im Postmortem ergänzt.

Kommunikationsbausteine

Im Ernstfall fehlt oft die Zeit, Kommunikation von Grund auf zu formulieren. Die folgenden **Textvorlagen** können direkt übernommen und an den konkreten Vorfall angepasst werden. Achten Sie auf einen sachlichen, transparenten Ton – sowohl intern als auch extern. Vermeiden Sie Schuldzuweisungen und spekulieren Sie nicht über Ursachen, solange die Analyse läuft.

1. Internes Stakeholder-Update

„Am [Datum] um [Uhrzeit] wurde ein KI-bezogener Incident vom Typ [Typ] im System [Systemname] festgestellt. Das betroffene Feature wurde sofort deaktiviert und die Beweissicherung eingeleitet. Die Root-Cause-Analyse läuft. Ein Update folgt bis [Zeitpunkt].“

2. Kundenhinweis (extern, neutral)

„Wir haben festgestellt, dass unser [Servicename] am [Datum] vorübergehend fehlerhafte Ergebnisse geliefert hat. Wir haben das Feature umgehend deaktiviert und arbeiten an der Behebung. Ihre Daten sind sicher. Wir informieren Sie, sobald der Service wieder vollständig verfügbar ist.“

3. Abschlussmeldung (intern)

„Der Incident [ID/Bezeichnung] vom [Datum] ist abgeschlossen. Die Ursache wurde identifiziert und behoben. Das Postmortem mit präventiven Maßnahmen ist im Incident Log hinterlegt.“

Kommunikationsmatrix nach Schweregrad

Schwere	Intern	Kunden	Behörde
1-2	Team-Channel	Nicht erforderlich	Nicht erforderlich
3	GF + Team	Bei direkter Betroffenheit	Prüfung durch DSB
4-5	Krisenkommunikation	Proaktiver Hinweis	Meldung innerhalb 72 Std.

Postmortem Template

NACHBEREITUNG

BLAMELESS CULTURE

Das Postmortem ist das wichtigste Instrument, um aus Vorfällen zu lernen und künftige Incidents zu verhindern. Es wird innerhalb von **5 Werktagen** nach Incident-Abschluss durchgeführt. Grundprinzip: **Blameless** – wir suchen nach Systemursachen, nicht nach Schuldigen. Alle Beteiligten nehmen teil. Das ausgefüllte Postmortem wird im Incident Log verlinkt.

1

Was ist passiert?

Chronologische Darstellung des Vorfalls mit Zeitstempeln. Welches System war betroffen? Wie wurde der Incident entdeckt? Wer hat zuerst reagiert?

2

Warum ist es passiert?

Root-Cause-Analyse (5-Why-Methode empfohlen). Technische, organisatorische und prozessuale Ursachen getrennt benennen.

3

Was hat gut funktioniert?

Welche Prozesse, Tools oder Reaktionen haben geholfen? Was sollte beibehalten werden?

4

Was hat gefehlt?

Welche Lücken in Prozessen, Tooling, Dokumentation oder Wissen wurden offenbar?

5

Maßnahmen (mit Owner & Termin)

Jede Maßnahme braucht eine verantwortliche Person und ein konkretes Zieldatum.

Maßnahme	Owner	Termin
-----	-----	-----
-----	-----	-----
-----	-----	-----

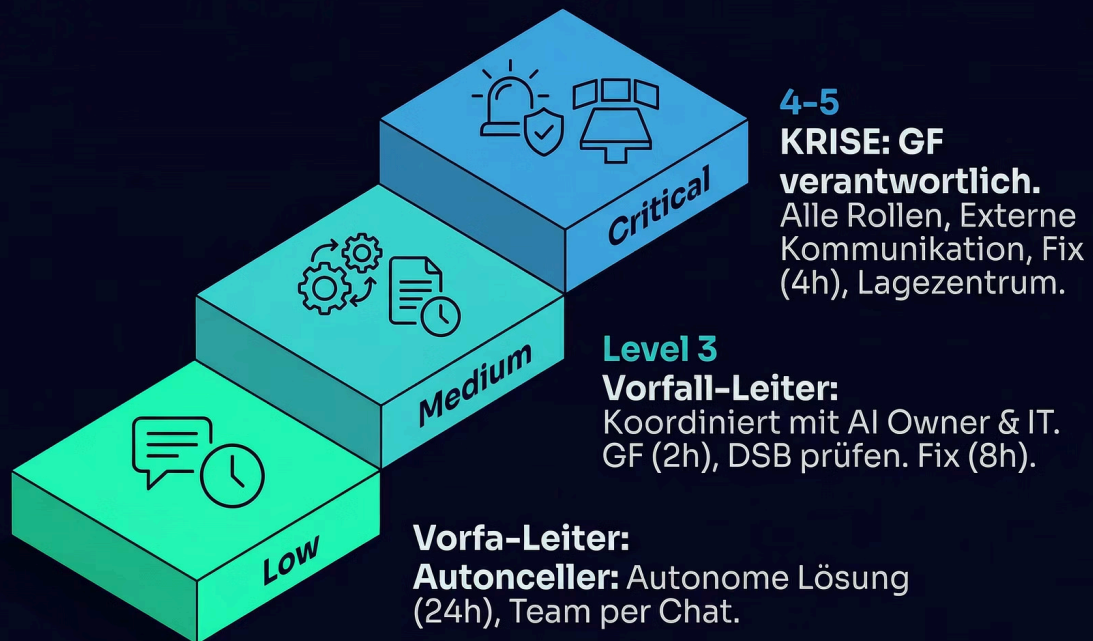
6

Monitoring-Anpassung

Welche neuen Alerts, Dashboards oder Prüfroutinen werden eingerichtet, um ähnliche Vorfälle frühzeitig zu erkennen?

Eskalationspfade im Detail

Nicht jeder Incident erfordert die gleiche Eskalationsstufe. Die folgenden Pfade zeigen, welche Maßnahmen bei welchem Schweregrad automatisch ausgelöst werden. Drucken Sie diese Seite aus und hängen Sie sie sichtbar im Büro auf – im Ernstfall zählt jede Minute.



Zeitleisten nach Schweregrad

Schwere	Erstmeldung	GF-Info	Fix-Ziel	Postmortem
1-2	Sofort intern	Nicht nötig	≤ 24 Std.	≤ 10 Werktage
3	≤ 30 Min.	≤ 2 Std.	≤ 8 Std.	≤ 5 Werktage
4-5	Sofort	Sofort	≤ 4 Std.	≤ 3 Werktage

War Room: Bei Schweregrad 4-5 wird ein dedizierter Kommunikationskanal (z. B. Teams-Call, Slack-Channel) eingerichtet, der bis zum Incident-Abschluss aktiv bleibt. Alle Beteiligten sind ständig erreichbar.

Prävention & Monitoring

Reaktive Incident Response ist nur die halbe Miete. Ein ausgereiftes KI-Risikomanagement setzt auf **proaktive Prävention und kontinuierliches Monitoring**, um Vorfälle frühzeitig zu erkennen oder ganz zu verhindern. Die folgenden Maßnahmen sollten als Mindeststandard in jedem KMU implementiert sein.

Technische Maßnahmen

- **API-Usage-Monitoring:** Alerts bei ungewöhnlichen Nutzungsspitzen oder Kostenüberschreitungen einrichten
- **Output-Qualitäts-Monitoring:** Stichprobenartige manuelle Prüfung + automatisierte Confidence-Score-Checks
- **Regelmäßige Key-Rotation:** API-Keys und Secrets mindestens alle 90 Tage rotieren
- **Input-Validierung:** Prompt-Injection-Filter und Input-Sanitization als Standard
- **Rate-Limiting:** Maximale Anfragen pro Minute/Stunde begrenzen

Organisatorische Maßnahmen

- **Quartalsweise Runbook-Reviews:** Dieses Dokument mindestens alle 3 Monate auf Aktualität prüfen
- **Tabletop-Übungen:** Halbjährlich einen fiktiven Incident durchspielen
- **Schulungen:** Alle KI-Nutzer jährlich zu Risiken und Incident-Meldung schulen
- **Kontaktliste aktuell halten:** Bei Personalwechsel sofort die Kontaktkette (Seite 5) aktualisieren
- **Eval-Sets pflegen:** Testdatensätze regelmäßig um neue Szenarien erweitern

Monitoring-Dashboard – empfohlene KPIs

< 4h

Mean Time to Detect

Durchschnittliche Zeit von Auftreten bis Erkennung eines Incidents

< 8h

Mean Time to Resolve

Durchschnittliche Zeit von Erkennung bis vollständiger Behebung

100%

Postmortem-Quote

Anteil der Incidents mit abgeschlossenem Postmortem

0

Wiederkehrende Incidents

Ziel: Kein gleicher Vorfall tritt zweimal auf

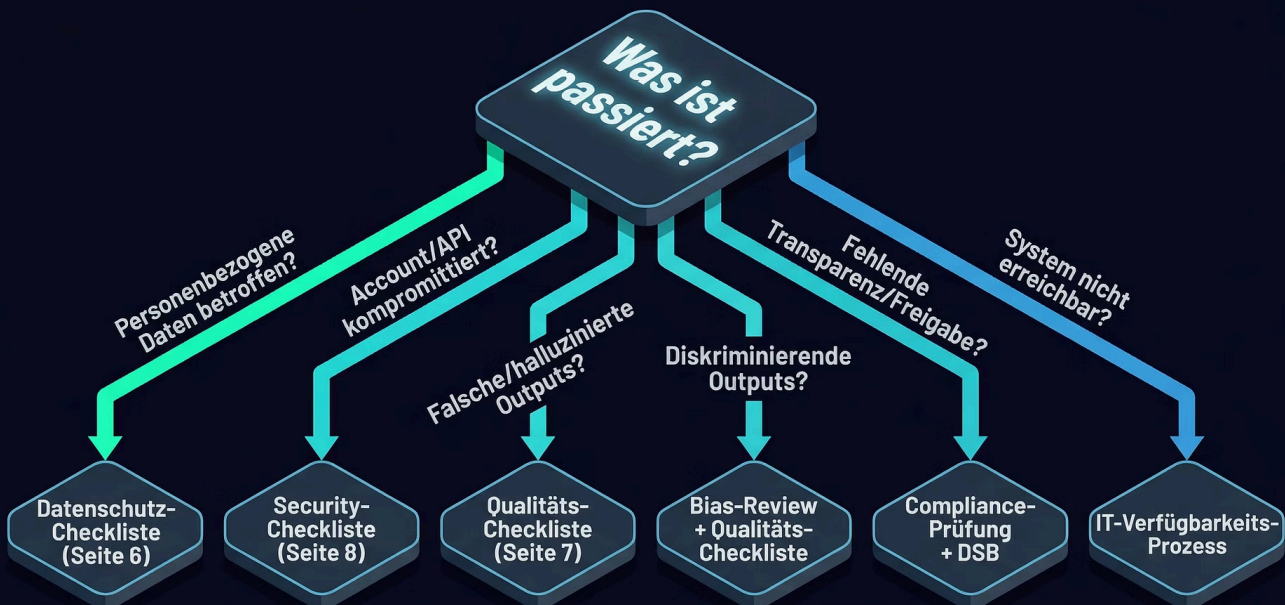
Quick-Reference-Karte

⚡ SOFORT-REFERENZ

Diese Seite fasst die wichtigsten Handlungsanweisungen auf einen Blick zusammen. Drucken Sie sie aus und legen Sie sie griffbereit neben Ihren Arbeitsplatz – oder heften Sie sie an die Pinnwand im IT-Büro.

1. STOP System pausieren Feature abschalten API sperren	2. SICHERN Logs exportieren Screenshots Zeitstempel
3. ANALYSIEREN Prompt? Modell? Daten? Policy? Zugriff?	4. FIXEN Patch deployen Guardrails setzen Zugriff entziehen
5. REDEN Intern informieren Extern melden DSB einbeziehen	6. LERNEN Postmortem Maßnahmen Monitoring

Entscheidungsbaum: Welche Checkliste?



Note: Multiple checklists can apply simultaneously.

Erinnerung: Ein Vorfall kann mehrere Typen gleichzeitig betreffen. Im Zweifel **alle** relevanten Checklisten parallel abarbeiten.

Ressourcen & Kontakt

Dieses Runbook ist Teil der **KI-Governance-Reihe** von Philipp Diekmann. Es wird regelmäßig aktualisiert und an neue regulatorische Anforderungen (EU AI Act, DSGVO-Anpassungen) und technologische Entwicklungen angepasst.

Weiterführende Ressourcen

- **EU AI Act – Volltext:** eur-lex.europa.eu
- **DSGVO Art. 33/34:** Meldepflichten bei Datenschutzverletzungen
- **BSI IT-Grundschutz:** Baustein Incident Management
- **NIST AI Risk Management Framework:** ai.nist.gov
- **OWASP Top 10 for LLMs:** owasp.org/llm-top-10

Kontakt

Philipp Diekmann

Web: kaffee-intelligenz.de

QR-Code: Scannen für direkten Zugang zur Website und weiteren Templates

📖 Buchtipp

Vertiefen Sie Ihr KI-Governance-Wissen mit praxisnahen Frameworks, Checklisten und Implementierungsleitfäden – speziell für KMU entwickelt.

Alle Templates, Checklisten und Runbooks auf:

kaffee-intelligenz.de

