T.I.S.P./T.P.S.S.E. Community Meeting 2026

Berlin, 10. - 11.11.2026

Shadow AI und Data Leakage: Wenn der Chatbot zum Insider Risk wird

Aurea Kindshofer, exkulpa GmbH

Der neue Insider: gutgläubig und unsichtbar

TeleTrusT
Pioneers in IT security.
Bundesverband IT-Sicherheit e.V.

Shadow AI ist kein Angriff – sondern ein Datentransfer an Dritte, ausgelöst von Ihren besten Leuten.



Böswillig

Gezielter Diebstahl, Sabotage, Rache. Selten, aber teuer.



Kompromittiert

Fremde nutzen gestohlene Identitäten von innen.



Fahrlässig

Fehlbedienung, falscher Empfänger, Schatten-IT.



Die KI-Abkürzung

Will nur schneller arbeiten – und kopiert dafür Code, Verträge und Kundendaten in fremde KI.

Merkmal: gute Absicht, hohe Frequenz, keine Telemetrie – klassische Insider-Programme sehen genau diesen Typ am schlechtesten.

Lagebild: Schatten-KI ist Normalzustand

TeleTrusT
Pioneers in IT security.
Bundesverband IT-Sicherheit e.V.

Beschäftigte sind schneller als ihre Organisationen – Regeln und Werkzeuge hinken hinterher.

29,1 %

der beruflichen KI-Nutzenden setzen Tools ein, die der Arbeitgeber nicht bereitgestellt hat

IW Köln 2026

4 von 10

Unternehmen gehen davon aus, dass private KI-Tools im Einsatz sind (2024: 3 von 10)

Bitkom 2025

26 %

der Unternehmen stellen ihren Beschäftigten generative KI zur Verfügung

Bitkom 2025

23 %

der Unternehmen haben Regeln für die KI-Nutzung – 24 % haben sich noch nicht damit befasst

Bitkom 2025

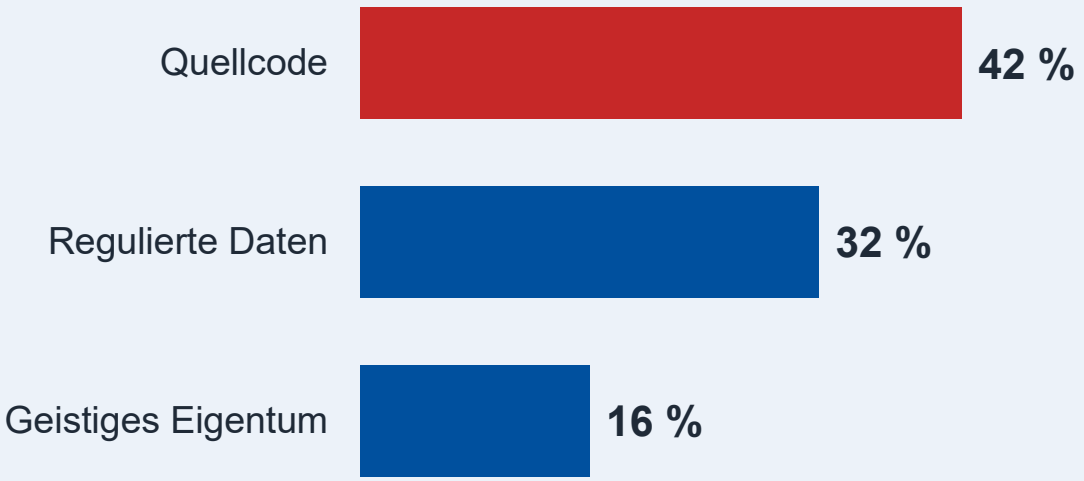
Auffällig: unter 30-Jährige 38,4 %, Führungskräfte 34,6 % – Schatten-KI ist ein Phänomen der Leistungsträger, nicht der Regelbrecher.

Quelle: IW-Kurzbericht 63/2026 (n = 2.793 KI-Nutzende); Bitkom-Presseinformation 21.10.2025 (n = 604 Unternehmen)

Telemetrie: Was tatsächlich in Prompts landet

Die Prompt-Menge versechsfacht sich – und mit ihr die Datenschutzverstöße.

Datenarten in GenAI-Policy-Verstößen



223

GenAI-Datenschutzverstöße pro Monat in der Durchschnittsorganisation – Top-Quartil: 2.100

×6

Prompts pro Monat: von 3.000 auf 18.000 in der Durchschnittsorganisation

50 %

der Organisationen haben keine durchsetzbaren Datenschutz-Policies für GenAI-Apps

Quelle: Netskope Threat Labs, Cloud and Threat Report 2026 (Datenbasis 10/2024–10/2025)

Was Shadow AI im Ernstfall kostet

TeleTrusT
Pioneers in IT security.
Bundesverband IT-Sicherheit e.V.

Shadow AI ist im Breach-Kostenmodell angekommen – und Governance ist die Lücke.

1 von 5

Organisationen meldet einen Breach mit Shadow-AI-Bezug

+670.000 $

Mehrkosten pro Breach bei hohem Shadow-AI-Anteil

97 %

der Organisationen mit KI-bezogenem Breach hatten keine KI-Zugriffskontrollen

63 %

der betroffenen Organisationen haben keine KI-Governance-Policy oder entwickeln sie noch

Was abfließt

- **Personenbezogene Daten:** 65 % der Shadow-AI-Fälle (Ø aller Breaches: 53 %)
- **Geistiges Eigentum:** 40 % (Ø 33 %)
- Globaler Ø-Schaden eines Breach: 4,44 Mio. $

Quelle: IBM / Ponemon Institute, Cost of a Data Breach Report 2025

Fall 1: Samsung 2023 – drei Leaks in Wochen

Kein Angreifer, keine Malware – nur Ingenieure, die ihre Arbeit gut machen wollten.

1. Quellcode

Fehlerhafter Code einer Halbleiter-Datenbank wird zur Fehlersuche in ChatGPT kopiert.

2. Testsequenzen

Proprietäre Testmuster zur Chip-Ausbeute werden zur Optimierung eingegeben.

3. Meeting-Mitschnitt

Eine interne Besprechung wird transkribiert und als Protokoll zusammengefasst.

Reaktion: zunächst Prompt-Limit, im Mai 2023 Verbot generativer KI auf Firmengeräten – parallel Aufbau eigener interner KI-Werkzeuge.

Lektionen

- Daten sind nach dem Absenden nicht zurückholbar
- Kontrollverlust entsteht beim Anbieter: Speicherung, Training, Review
- Das Verbot war die Reaktion – nicht die Lösung

Quelle: Presseberichte (u. a. Economist Korea, Bloomberg), April/Mai 2023

Sechs Abflusspfade – nicht nur Copy & Paste

TeleTrusT
Pioneers in IT security.
Bundesverband IT-Sicherheit e.V.

Die Angriffsfläche wächst vom Browserfenster in Erweiterungen, SaaS-Features und Agenten.

Copy & Paste, Uploads

Private Accounts, kein SSO, kein Logging, keine Löschkontrolle.

Consumer-Bedingungen

Training, Retention, Human Review – Datenverwendung beim Anbieter.

Browser-Erweiterungen

KI-Sidebars lesen jede geöffnete Seite – inkl. CRM, HR, Tickets.

Eingebettete KI & Notetaker

SaaS-Features per Toggle, Meeting-Bots in vertraulichen Calls.

Agenten & Connectors

OAuth-Grants, MCP-Server, API-Keys mit weitreichenden Rechten.

Prompt Injection

Auch die freigegebene KI wird über präparierte Inhalte zum Leck.

■ klassische Schatten-IT ■ wachsend, oft auch in sanktionierten Umgebungen

Fall 2: EchoLeak – die genehmigte KI als Leck

TeleTrusT
Pioneers in IT security.
Bundesverband IT-Sicherheit e.V.

CVE-2025-32711 in Microsoft 365 Copilot: Zero-Click, CVSS 9.3.

Präparierte E-Mail

Angreifer versteckt Anweisungen in normal wirkendem Text/Markdown.

RAG holt Kontext

Copilot zieht die Mail bei einer beliebigen Nutzerfrage mit in den Kontext.

Scope Violation

Untrusted Input und interne Daten landen im selben Prompt.

Exfiltration

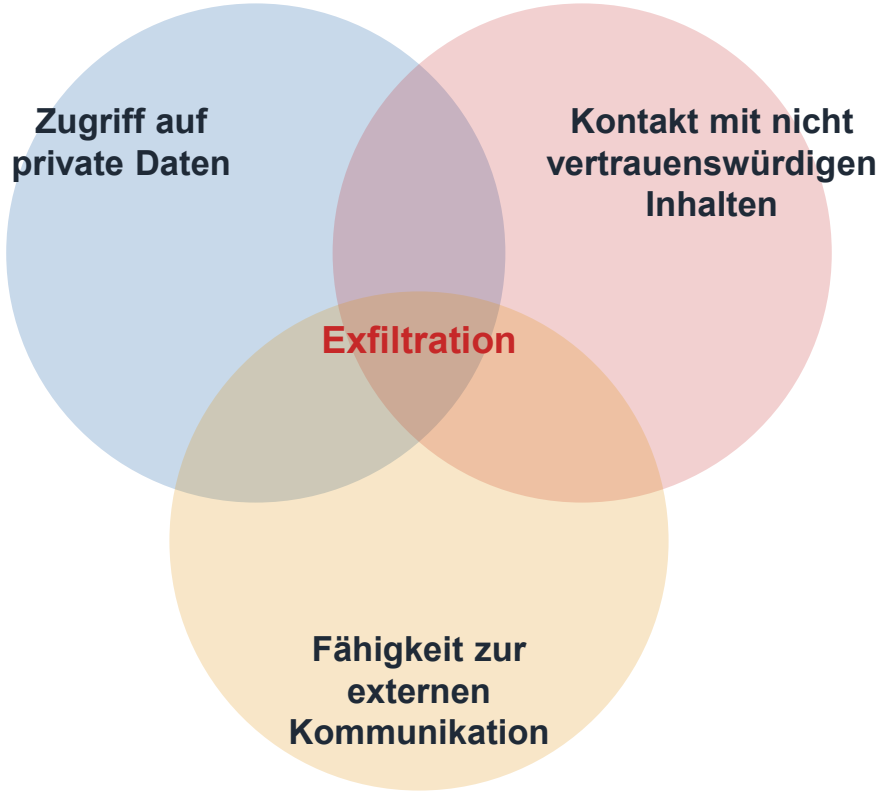
Antwort enthält Link/Bild-Referenz – Daten gehen an Angreifer-Server.

Für Ihr Insider-Modell: Die KI handelt mit den Rechten des Nutzers, aber auf Anweisung eines Externen. Entdeckt von Aim Security, im Juni 2025 serverseitig behoben, keine bekannte Ausnutzung – das Muster bleibt.

Quelle: Aim Security; Microsoft Security Update Guide CVE-2025-32711; The Hacker News, 12.06.2025

Agentic AI: Die gefährliche Dreierkombination

Ein Agent mit allen drei Eigenschaften ist ein Insider, den jeder Fremde steuern kann.



Nie alle drei zugleich erlauben

- **Daten:** Least Privilege für Agenten, eigene Service-Identitäten, keine Vollzugriffs-Tokens
- **Inhalte:** E-Mail, Web und Tickets als untrusted behandeln, Tool-Aufrufe nicht aus Fremdtext ableiten
- **Kommunikation:** Egress-Allowlist, keine automatisch geladenen Bilder/Links, Human-in-the-Loop bei Versand
- **Inventar:** MCP-Server und OAuth-Apps wie privilegierte Konten führen und rezertifizieren

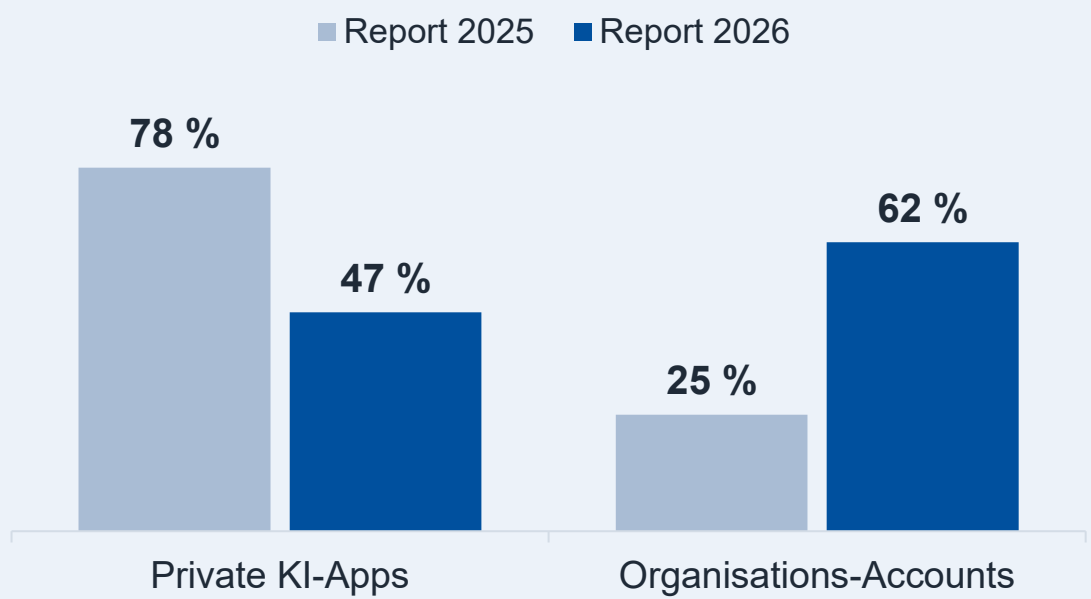
Quelle: Konzept „Lethal Trifecta“ nach Simon Willison (2025); eigene Darstellung

Warum Verbote allein scheitern

TeleTrusT
Pioneers in IT security.
Bundesverband IT-Sicherheit e.V.

Wer KI fordert, muss KI liefern – sonst wird sie privat beschafft.

Anteil der KI-Nutzenden nach Account-Typ



44,5 % vs. 21,3 %

Schatten-KI-Quote bei Beschäftigten, die starke vs. keine Förderung von KI-Kompetenz durch den Arbeitgeber wahrnehmen (IW 2026)

- Druck ohne Werkzeug erzeugt Schatten-KI
- Bereitstellung wirkt: private Nutzung sinkt messbar, wenn Firmen-Accounts verfügbar sind
- 90 % der Organisationen blockieren Apps – Blocklisten veralten schneller, als neue Tools entstehen

Quelle: Netskope Cloud and Threat Report 2026; IW-Kurzbericht 63/2026

Rechtsrahmen: Ein Prompt, vier Regime

TeleTrusT
Pioneers in IT security.
Bundesverband IT-Sicherheit e.V.

Ein unbedachter Prompt berührt Datenschutz-, Sicherheits- und Geheimnisschutzrecht zugleich.

DSGVO

- Art. 5 Abs. 1 f, 32: Vertraulichkeit & TOMs
- Art. 28: AVV fehlt bei privaten Tools
- Art. 44 ff.: Drittlandtransfer
- Art. 33/34: Offenlegung = mögliche Datenpanne

NIS2 / BSIG

- § 30: Risikomanagement inkl. Cloud & Lieferkette
- § 38: Billigung, Überwachung, Schulung durch Geschäftsleitung
- Shadow AI als ungesteuertes Lieferantenrisiko

AI Act

- Art. 4: KI-Kompetenz seit 02.02.2025
- Nach Digital Omnibus: Pflicht zu Fördermaßnahmen bleibt
- Betreiber kennen ihre KI-Systeme nur mit Inventar

GeschGehG

- § 2 Nr. 1 b: Schutz nur bei angemessenen Maßnahmen zur Geheimhaltung
- Ohne KI-Regeln droht der Verlust des Geheimnisstatus

Plus: Prompt-Logging und DLP-Inspektion sind Leistungs- und Verhaltenskontrolle – § 87 Abs. 1 Nr. 6 BetrVG, Betriebsrat früh einbinden.

Kontrollrahmen in fünf Phasen

TeleTrusT
Pioneers in IT security.
Bundesverband IT-Sicherheit e.V.

Shadow AI lässt sich in bestehende ISMS-Prozesse einhängen – ein Sonder-Framework braucht es nicht.

Discover	Govern	Protect	Enable	Respond
• SSE/Proxy-Logs, DNS • OAuth-Grants & Extensions • KI-Inventar	• KI-Richtlinie, Positivliste • Use-Case-Freigabe • Klassifizierung → KI-Klasse	• DLP für Prompts & Uploads • Tenant-Restriktion, SSO • Agenten-Rechte begrenzen	• Enterprise-KI bereitstellen • Schulung (Art. 4 AI Act) • Anbieter-Due-Diligence	• Playbook „Prompt-Leak“ • Secrets rotieren, Löschung • Art. 33-Bewertung
ISO 27001 A. **5.9, 8.16**	ISO 27001 A. **5.1, 5.10, 5.12**	ISO 27001 A. **8.12, 5.14, 5.23**	ISO 27001 A. **6.3, 5.19, 5.20**	ISO 27001 A. **5.24 – 5.28**

Für KI-Managementsysteme ergänzend: ISO/IEC 42001 (KI-Inventar, Rollen, Folgenabschätzung)

Technische Kontrollen und ihre Grenzen

TeleTrusT
Pioneers in IT security.
Bundesverband IT-Sicherheit e.V.

Keine Einzelmaßnahme reicht – es kommt auf Sicht, Steuerung und Bereitstellung an.

Kontrolle	Wirkung	Grenze
SSE / CASB-Discovery	Inventar: welche KI, wer, wie viel	Blind für private Geräte und Mobilnetz
Tenant-Restriktion, SSO-Zwang	Firmen- statt Privat-Account erzwingen	Pflegeaufwand je Anbieter
Browser- / Endpoint-DLP	Prompts & Uploads prüfen, maskieren, blocken	False Positives bei unstrukturierten Daten
KI-Gateway mit Redaction	Zentrale Policy, Logging, Modellwahl	Greift nur für angebundene Use Cases
OAuth- & MCP-Governance	Agenten-Rechte begrenzen, rezertifizieren	Wildwuchs wächst schneller als Prozesse
Extension-Allowlisting	KI-Sidebars und Scraper ausschließen	Akzeptanz, Ausnahmeprozess nötig

Governance: Datenklasse trifft KI-Klasse

TeleTrusT
Pioneers in IT security.
Bundesverband IT-Sicherheit e.V.

Eine Ampel statt 20 Seiten Richtlinie – verständlich für alle, prüfbar für die Revision.

	Consumer-KI (privater Account)	**Enterprise-KI** (AVV, EU, kein Training)	**Interne KI** (eigener Tenant / On-Prem)
Öffentlich	erlaubt	erlaubt	erlaubt
Intern	verboten	erlaubt	erlaubt
Vertraulich	verboten	mit Freigabe	erlaubt
Streng vertraulich / Art. 9	verboten	verboten	mit Freigabe

Dazu gehören

- Positivliste freigegebener Tools und Use Cases
- Schneller Freigabeprozess (Tage, nicht Monate)
- Schulung mit Beispielen aus dem eigenen Alltag
- Meldeweg mit Amnestie für Selbstmeldungen

Shadow AI in der ISO/IEC 42001

TeleTrusT
Pioneers in IT security.
Bundesverband IT-Sicherheit e.V.

Kein eigener Abschnitt – aber Shadow AI berührt die Norm an acht Stellen.

4.3 · A.4.2 · A.4.4

Transparenz über genutzte KI

Anwendungsbereich und Ressourcen ohne blinde Flecken

5.2 · A.2.2 · A.2.3

Verbindliche KI-Politik

Abgestimmt mit ISMS und Datenschutz

5.3 · A.3.2

Klare Freigabeverantwortung

Rollen für Tools und Use Cases statt Grauzone

6.1.2–6.1.4 · 8.2–8.4

Bewertetes KI-Risiko

Beurteilung, Behandlung, Folgenabschätzung

A.9.2 · A.9.3 · A.9.4

Geregelte Nutzungsprozesse

Verantwortungsvolle Nutzung mit festgelegtem Zweck – der Kern

A.10.2 · A.10.3

Geprüfte KI-Anbieter

Ein privates KI-Tool ist ein ungeprüfter Lieferant

7.2 · 7.3 · A.3.3

Kompetenz und Meldeweg

Schulung, Bewusstsein, Meldung ohne Sanktionsangst

9.1 · A.8.4 · 10.2

Überwachung und Verbesserung

Messen, Vorfälle kommunizieren, korrigieren

Quelle: ISO/IEC 42001:2023 · Ziffern = Abschnitte des Hauptteils, A.x = Maßnahmen aus Anhang A · eigene Zuordnung

Harmonisierte Struktur: 27001 und 42001

TeleTrusT
Pioneers in IT security.
Bundesverband IT-Sicherheit e.V.

Gleiche Struktur – Unterschiede nur bei Rolle, Folgenabschätzung und Maßnahmenkatalog.

Abschnitt	ISO/IEC 27001:2022	ISO/IEC 42001:2023
4 Kontext	Kontext, interessierte Parteien, Anwendungsbereich	4.1–4.4 **+ Rolle zu KI-Systemen bestimmen**
5 Führung	Führung, Informationssicherheitspolitik, Rollen	Führung, KI-Politik, Rollen
6 Planung	Risikobeurteilung, Risikobehandlung inkl. SoA, Ziele	Ebenso für KI-Risiken **+ 6.1.4 Folgenabschätzung**
7 Unterstützung	Ressourcen, Kompetenz, Kommunikation, Dokumentation	Identisch aufgebaut
8 Betrieb	Steuerung, Risikobeurteilung und -behandlung	Ebenso **+ 8.4 Folgenabschätzung durchführen**
9 Bewertung	Überwachung, internes Audit, Managementbewertung	Identisch aufgebaut
10 Verbesserung	Fortlaufende Verbesserung, Korrekturmaßnahmen	Identisch aufgebaut
Anhang A	93 Maßnahmen in 4 Themen	**38 Maßnahmen zu 9 Zielen (A.2–A.10)**

Gemeinsam: Risikoprozess, Schulung, Audit, Managementbewertung

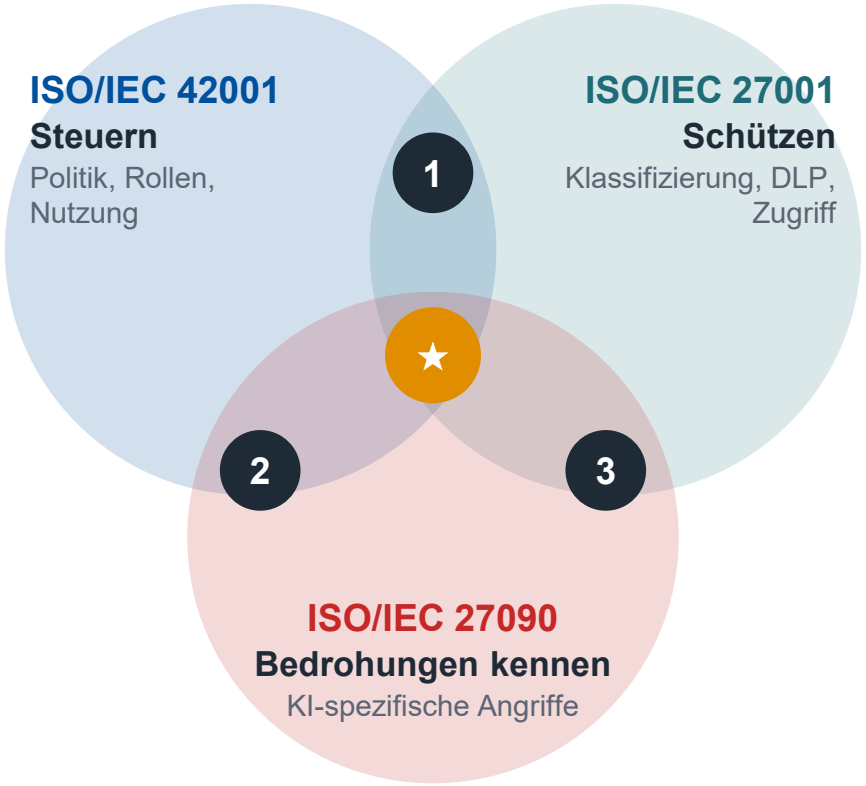
Getrennt: Folgenabschätzung, Anhang A, Erklärung zur Anwendbarkeit

Quelle: ISO/IEC 27001:2022; ISO/IEC 42001:2023; ISO/IEC Directives Annex SL (Harmonized Structure) ■ zusätzlich bzw. abweichend in der 42001

Normen-Zusammenspiel: 42001, 27001, 27090

TeleTrusT
Pioneers in IT security.
Bundesverband IT-Sicherheit e.V.

Steuern, schützen, Bedrohungen kennen – wirksam wird es erst in den Schnittmengen.



1 — 42001 ∩ 27001: ein integriertes Managementsystem
Gleiche Grundstruktur: ein Risikoprozess, ein Audit, eine Managementbewertung. A.2.3 ↔ 5.1/5.10 · A.10.3 ↔ 5.19–5.23 · 7.3 ↔ 6.3 · A.8.4 ↔ 5.24 ff.

2 — 42001 ∩ 27090: Risiko mit KI-Bedrohungen bewerten
6.1.2 / 8.2 nutzen den Bedrohungskatalog; Verifikation und Betrieb (A.6.2) mit KI-Sicherheitstests.

3 — 27001 ∩ 27090: vorhandene Maßnahmen auf KI ausrichten
DLP für Prompts (8.12), Maskierung im Gateway (8.11), Netztrennung und Egress für Agenten (8.20/8.22).

★ — Mitte: Shadow AI beherrscht
Gesteuert, geschützt und bedrohungsgerecht abgesichert.

Quelle: ISO/IEC 42001:2023; ISO/IEC 27001:2022; ISO/IEC 27090:2026; eigene Darstellung

Incident Response: Der Prompt ist raus

TeleTrusT
Pioneers in IT security.
Bundesverband IT-Sicherheit e.V.

Nicht jeder Prompt-Leak ist meldepflichtig – aber jeder ist bewertungs- und dokumentationspflichtig.

1

Erkennen

DLP-Alert oder Selbstmeldung; Chat, Zeitpunkt, Tool und Account sichern

2

Eindämmen

Chat und Freigabe-Links löschen, Secrets und API-Keys sofort rotieren

3

Bewerten

Personenbezug? Art. 33: 72 h. Geschäftsgeheimnis, NDA, Mandatsgeheimnis?

4

Anbieter

Retention und Training prüfen, Löschung verlangen, Antwort dokumentieren

5

Lernen

Ursache: fehlendes Tool? Fehlende Regel? → Enable-Maßnahme ableiten

Art. 33 Abs. 5 DSGVO: Jede Verletzung wird intern dokumentiert – auch wenn das Risiko keine Meldung an die Aufsicht auslöst.

Take-aways: Ihr 90-Tage-Plan

TeleTrusT
Pioneers in IT security.
Bundesverband IT-Sicherheit e.V.

Shadow AI ist kein Nutzerproblem, sondern ein Angebotsproblem.

Tag 1–30 Sichtbarkeit

- SSE/Proxy- und DNS-Logs auf KI-Domains auswerten
- OAuth-Apps, MCP-Server, Browser-Extensions inventarisieren
- Secrets-Scanning für Code-Repos

Tag 31–60 Steuerung

- KI-Richtlinie mit Ampel-Matrix
- Enterprise-KI für den Top-Use-Case bereitstellen
- Betriebsrat und DSB einbinden

Tag 61–90 Durchsetzung

- DLP für Prompts scharf schalten
- IR-Playbook „Prompt-Leak“ üben
- KPI: Anteil gemanagter KI-Nutzung

Erfolg messen: Anteil gemanagter KI-Nutzung · Zahl freigegebener Use Cases · Zeit bis zur Freigabe · Selbstmeldungen

Fragen & Diskussion



Aurea Kindshofer

exkulpa gmbh – Der Faktor für Ihre Compliance

Aurea Kindshofer
a.kindshofer@exkulpa.de

Quellen

- Bitkom e. V.: „Beschäftigte nutzen vermehrt Schatten-KI“, Presseinformation, 21.10.2025 (n = 604 Unternehmen ab 20 Beschäftigten)
- Engels, B.; Hammermann, A.: „Schatten-KI: Beschäftigte sind schneller als ihre Betriebe“, IW-Kurzbericht Nr. 63/2026, 19.09.2026
- Netskope Threat Labs: Cloud and Threat Report 2026, 05.01.2026 (Datenbasis 01.10.2024–31.10.2025)
- IBM / Ponemon Institute: Cost of a Data Breach Report 2025, Juli 2025
- Aim Security: EchoLeak (CVE-2025-32711), Juni 2025; Microsoft Security Response Center; The Hacker News, 12.06.2025
- Willison, S.: „The lethal trifecta for AI agents“, simonwillison.net, Juni 2025
- Samsung-Vorfälle: Economist Korea (März 2023), Bloomberg (02.05.2023)
- Verordnung (EU) 2024/1689 (AI Act) i. d. F. des Digital Omnibus; DSGVO; BSIG i. d. F. NIS2UmsuCG; GeschGehG; BetrVG
- ISO/IEC 27001:2022 Anhang A; ISO/IEC 42001:2023; ISO/IEC 27090:2026