# Playbook: Schatten-KI & Datenschutz

# 10 konkrete Schritte, die ich als Entscheider jetzt gehen würde

> Schatten-KI entsteht nicht durch schlechte Mitarbeitende – sondern durch fehlende Leitplanken.

Dieses Playbook zeigt Ihnen pragmatisch und umsetzbar, wie Sie KI-Nutzung sichtbar machen, DSGVO-konform steuern und Enablement statt Verbote umsetzen. Keine Panikmache, sondern klare Handlungsschritte für die Praxis.

# Warum das Thema jetzt auf Ihren Tisch gehört

Die unkontrollierte KI-Nutzung in Unternehmen ist kein Randthema mehr. Sie berührt gleichzeitig Datenschutz, Reputation und Effizienz – und der EU AI Act verschärft die Anforderungen an dokumentierte Prozesse und klare Verantwortlichkeiten.

### Risiko

Datenschutzverletzungen, vertrauliche Informationen in fremden Systemen, fehlende Nachweise für Aufsichtsbehörden

### Vertrauen

Reputationsschäden bei Kunden und Partnern, wenn sensible Daten unkontrolliert verarbeitet werden

### Geschwindigkeit

Effizienzgewinne bleiben ungenutzt, wenn Teams aus Unsicherheit auf KI verzichten oder heimlich arbeiten

**EU AI Act:** Neue Anforderungen an Rollen, Prozesse und Dokumentation machen strukturiertes KI-Management zur Pflicht.

# Schatten-KI: Typische Situationen aus dem Alltag

Schatten-KI entsteht dort, wo Mitarbeitende praktische Lösungen suchen – ohne böse Absicht, aber mit echten Risiken. Diese Szenarien kennen Sie wahrscheinlich aus Ihrem Unternehmen:

### E-Mail-Zusammenfassungen

Lange Kundenkonversationen werden in ChatGPT eingefügt, um schnell eine Zusammenfassung zu erhalten

### Ticket-Analyse

Support-Mitarbeitende lassen KI-Tools Kundenbeschwerden kategorisieren – inklusive persönlicher Daten

### Angebotsentwürfe

Vertrieb nutzt KI für Angebotstexte und fügt versehentlich Preisinformationen oder Kundennamen ein

### Meeting-Notizen

Protokolle strategischer Meetings werden durch KI optimiert – mit vertraulichen Projektinhalten

### HR-Texte

Stellenbeschreibungen oder interne Kommunikation enthält unbeabsichtigt sensible Organisationsdetails

Was häufig schiefgeht: Copy-Paste sensibler Inhalte, unklare Tool-Policies, fehlende Schulung zu sicheren Prompts.

# Schnelltest: Wie hoch ist Ihr Schatten-KI-Risiko?

Beantworten Sie diese 10 Fragen ehrlich. Jedes „Nein" erhöht Ihr Risiko. Am Ende sehen Sie Ihre Risikozone und wissen, wo Sie ansetzen müssen.

1. **Freigegebene Tools?**
   Gibt es eine Liste erlaubter KI-Tools mit klaren Nutzungsregeln?
2. **Logging aktiv?**
   Können Sie nachvollziehen, welche Tools in welchen Bereichen genutzt werden?
3. **Schulung durchgeführt?**
   Haben Mitarbeitende ein Training zu sicherer KI-Nutzung erhalten?
4. **Datenklassifizierung?**
   Wissen Teams, welche Daten in KI-Tools dürfen und welche nicht?
5. **DPA/AVV vorhanden?**
   Sind Auftragsverarbeitungsverträge mit Tool-Anbietern abgeschlossen?
6. **Prompt-Guidelines?**
   Existieren Regeln, wie Mitarbeitende Prompts sicher formulieren?
7. **Verantwortlichkeiten geklärt?**
   Ist definiert, wer für KI-Governance zuständig ist?
8. **Incident-Prozess?**
   Gibt es einen Plan für den Umgang mit KI-bedingten Vorfällen?
9. **Betriebsrat eingebunden?**
   Wurden Mitbestimmungsrechte bei KI-Einführung berücksichtigt?
10. **Review-Zyklus?**
    Werden KI-Richtlinien regelmäßig überprüft und aktualisiert?

**8-10 Ja: Grün**
Solide Basis, Feintuning empfohlen

**4-7 Ja: Gelb**
Handlungsbedarf, schnell Lücken schließen

**0-3 Ja: Rot**
Hohes Risiko, sofort starten

# Leitplanken statt Verbote: Was darf in KI-Tools?

Die Datenklassifizierung ist Ihr wichtigstes Werkzeug. Sie gibt Mitarbeitenden klare Orientierung: Was ist unbedenklich? Was braucht besondere Vorsicht? Was ist tabu?

### Grün: Öffentlich

Allgemein zugängliche Informationen, Marketingtexte, öffentliche Produktbeschreibungen

**Beispiel:** Blogbeitrag optimieren, FAQ-Entwurf erstellen

### Gelb: Intern

Interne, nicht-sensible Inhalte wie allgemeine Prozessbeschreibungen oder Teamnotizen

**Beispiel:** Meeting-Agenda strukturieren, Checkliste erstellen

### Orange: Vertraulich

Vertrauliche Inhalte nur in freigegebenen, vertraglich geregelten Umgebungen

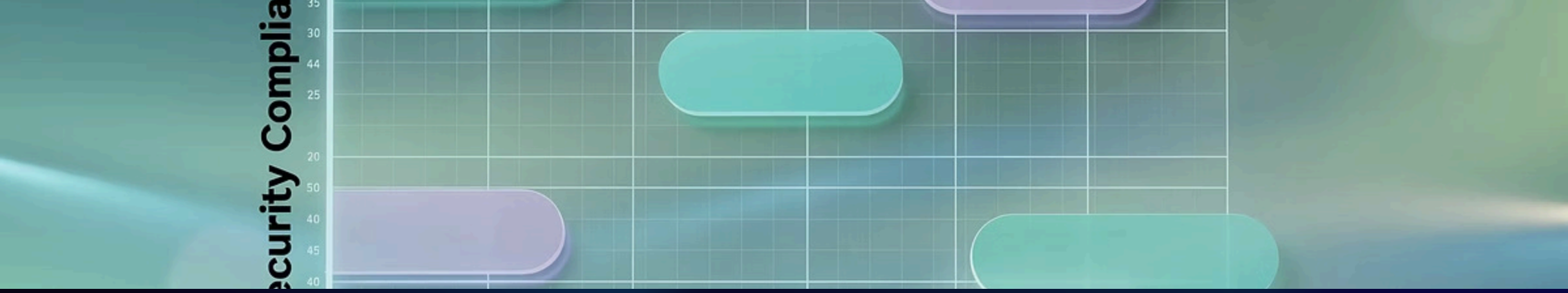
**Beispiel:** Projektdetails in Enterprise-Tool mit AVV

### Rot: Geschützt

Personenbezogene Daten, Kundendaten, Geschäftsgeheimnisse nur mit klaren Prozessen oder gar nicht

**Beispiel:** Gehaltsdaten, Vertragsdetails, Patientenakten

Praxis-Tipp: Drucken Sie diese Klassifizierung als Poster und hängen Sie sie in Arbeitsbereichen auf.



# Tool-Strategie: Offiziell erlauben statt heimlich verbieten

Verbote treiben Schatten-KI erst richtig voran. Besser: Definieren Sie freigegebene Tools mit klaren Rahmenbedingungen und bieten Sie sichere Alternativen.

### Freigegebene Tools

- Liste erlaubter KI-Services (z.B. ChatGPT Enterprise, Microsoft Copilot, interne Lösungen)
- Klare Nutzungsbedingungen je Datenklasse
- Regelmäßige Überprüfung und Erweiterung

### Vertrags-Checkpunkte

- Auftragsverarbeitungsvertrag (AVV/DPA) vorhanden?
- Datenverarbeitung außerhalb EU geregelt?
- Speicherorte und Löschfristen dokumentiert?
- Modelltraining mit Unternehmensdaten ausgeschlossen?

### Vor- und Nachteile

**Vorteile:** Rechtssicherheit, klare Orientierung, weniger Schatten-IT, bessere Kontrolle

**Nachteile:** Initialer Verhandlungsaufwand, laufende Verwaltung, möglicherweise höhere Kosten

**Alternative:** Sichere interne KI-Umgebung (z.B. Azure OpenAI) mit voller Datenkontrolle – höhere Investition, maximale Sicherheit

# Governance Light: Wer macht was?

KI-Governance muss nicht schwerfällig sein. Diese schlanke RACI-Struktur zeigt, wer entscheidet, wer berät, wer umsetzt – ohne Bürokratie-Monster.

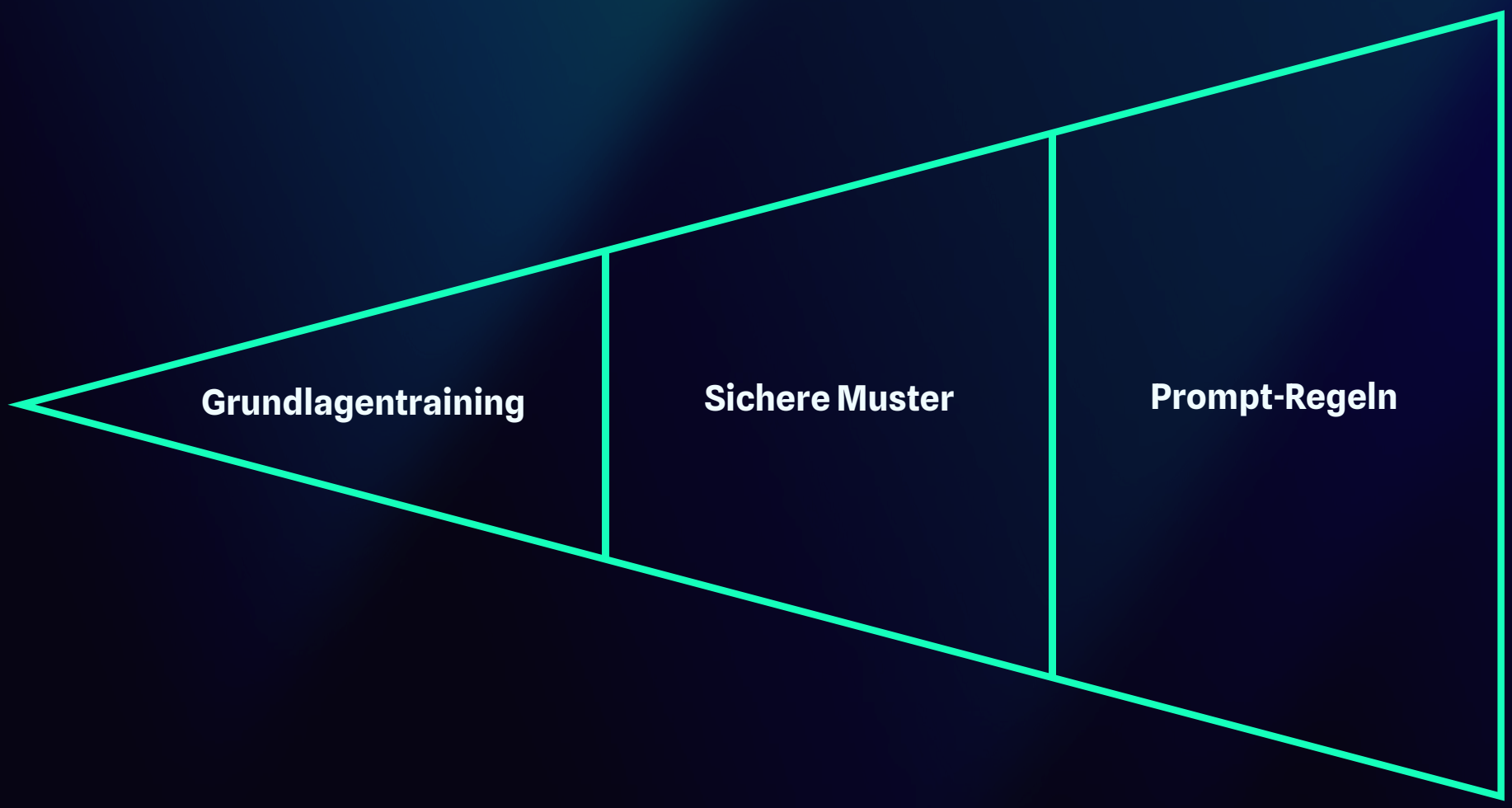
| Aufgabe | Management | IT/Security | Datenschutz | Fachbereich | HR/Training |
|---|---|---|---|---|---|
| Tool-Freigabe entscheiden | A | R | C | C | I |
| Datenschutz-Prüfung | I | C | R/A | C | I |
| Schulung durchführen | A | C | C | C | R |
| Richtlinien erstellen | A | R | C | C | I |
| Monitoring umsetzen | A | R | C | I | I |
| Incident-Management | I | R | C | C | I |

**Legende:** R = Responsible (führt aus) • A = Accountable (entscheidet) • C = Consulted (wird einbezogen) • I = Informed (wird informiert)

**Betriebsrat:** Bei KI-Einführung frühzeitig einbinden – Mitbestimmung bei Leistungs- und Verhaltenskontrolle beachten.

# Enablement: Schulung und sichere Prompts

Mitarbeitende wollen KI nutzen – geben Sie ihnen das Wissen dazu. Ein 60-Minuten-Training und klare Prompt-Regeln reduzieren Risiken drastisch.




Dieser Enablement-Ansatz verwandelt unsichere Nutzer in kompetente Anwender, die Risiken selbst erkennen und vermeiden.

### Basistraining (60 Min)

- Datenklassifizierung verstehen
- Freigegebene Tools kennen
- Typische Fehler vermeiden
- Sichere Prompt-Patterns anwenden

### Safe Prompt Patterns

- Anonymisieren: Namen durch Platzhalter ersetzen
- Abstrahieren: Konkrete Fälle verallgemeinern
- Synthetische Beispiele: Eigene fiktive Daten erstellen

### Prompt-Guidelines (Do/Don't)

✓ **Do:** „Erstelle eine Stellenbeschreibung für eine Projektleitung im Tech-Bereich"

**Don't:** „Erstelle eine Stellenbeschreibung für Sarah Müller als Nachfolgerin von Tom Schmidt in unserem KI-Projekt X"

### Champions-Netzwerk

KI-Champions in Fachabteilungen bieten Office Hours, sammeln Fragen und geben Best Practices weiter

Vor-/Nachteile: Initialer Schulungsaufwand vs. massive Reduktion von Fehlverhalten und Vorfällen.

# Betrieb: Monitoring, Feedback, Lernen

Governance lebt vom kontinuierlichen Lernen. Ein klarer Incident-Flow sorgt dafür, dass Vorfälle konstruktiv behandelt werden – ohne Schuldzuweisungen, mit echten Verbesserungen.

**Verdacht melden** | **Klassifizierung** | **Eindämmung** | **Kommunikation** | **Lessons**

Dieser Prozess macht aus jedem Vorfall eine Chance zur Verbesserung. Niedrigschwellige Meldemöglichkeiten sind dabei entscheidend.

### Was loggen?

- Tool-Zugriffe (aggregiert, nicht personenbezogen)
- Datenklassen-Verstöße
- Schulungsteilnahme

### Wie melden?

- Niedrigschwellige Kanäle (E-Mail, Formular, Ansprechperson)
- Keine Bestrafungskultur
- Anonyme Meldung möglich

### Was ist ein Incident?

- Sensible Daten in nicht-freigegebenen Tools
- Verstoß gegen Datenklassifizierung
- Unkontrollierte Tool-Nutzung

### Lessons Learned

- Pflichtschritt nach jedem Incident
- Was können wir verbessern?
- Schulungen anpassen, Guidelines erweitern

# 30-Tage-Plan: So starten Sie sofort

Sie brauchen keine monatelange Strategiephase. Dieser 4-Wochen-Plan bringt Sie von Null auf ein funktionierendes KI-Operating-Model – pragmatisch und sofort umsetzbar.

### 1 Woche 1: Bestandsaufnahme

- Shadow-KI-Umfrage in allen Bereichen durchführen
- Tool-Inventory erstellen (welche Tools werden genutzt?)
- Quick-Risikoanalyse mit Schnelltest (siehe Seite 4)

### 2 Woche 2: Grundlagen schaffen

- Datenklassifizierung definieren und kommunizieren
- Quick-Policy erstellen (1 Seite, klar verständlich)
- Erste Tool-Freigaben mit IT/Datenschutz abstimmen

### 3 Woche 3: Enablement starten

- 60-Min-Basisschulung für Pilotgruppen durchführen
- Prompt-Guidelines veröffentlichen und verteilen
- Champions-Netzwerk aufbauen

### 4 Woche 4: Operating Model aktivieren

- Monitoring-Tools einrichten (Logging, Dashboards)
- Incident-Prozess kommunizieren und testen
- Erste Review-Runde: Was läuft? Was muss angepasst werden?

**Nächster Schritt:** Kommentieren Sie „PLAYBOOK" – dann teile ich eine 1-Seiten-Policy-Vorlage als Text, die Sie direkt anpassen können.

---

*Philipp Diekmann | kaffee-intelligenz.de | linkedin.com/in/philippdiekmann*