

Sichere Nutzung von KI im Arbeitsalltag

Ein vollständiges Modul aus dem Seminar-Kit „KI-Kompetenz im Betrieb“: 10 Folien, Moderationsleitfaden, zwei Übungen mit Musterlösung und eine Quellenliste mit Stand. Sie dürfen es in Ihren eigenen Seminaren einsetzen.

Rechts- und Quellenstand: 26.10.2026 · Version 0.9 (Probe) · Allgemeine Information, keine Rechtsberatung im Einzelfall.

Teil A – Folien

Im Kit als editierbare Datei (PPTX und ODP); hier als Lesefassung.

FOLIE 1

Sichere Nutzung von KI im Arbeitsalltag

Was ich eingeben darf, wie ich Antworten prüfe, was ich weitergebe.

Nach diesem Modul können Sie ...

- Eingaben mit der Ampel einordnen: grün, gelb, rot.
- eine KI-Antwort in vier Schritten prüfen, bevor Sie sie verwenden.
- versteckte Anweisungen in Dokumenten und Webseiten erkennen.
- sagen, welche Regeln in Ihrem Betrieb gelten und wen Sie fragen.

Was mit Ihrer Eingabe passiert

- Die meisten KI-Werkzeuge laufen auf Servern des Anbieters. Ihre Eingabe verlässt den Betrieb.
- Ob sie gespeichert, ausgewertet oder zum Training genutzt wird, regeln Vertrag und Einstellungen – je Werkzeug und Tarif verschieden.
- Darum entscheidet der Betrieb, welche Werkzeuge für welche Daten freigegeben sind. Nicht das einzelne Konto.

Quelle: DSK, Orientierungshilfe „Künstliche Intelligenz und Datenschutz“, 06.05.2024, Abschnitt 2

FOLIE 4 · DIE EINGABE-AMPEL

Grün, gelb, rot

- **Grün:** öffentliche Informationen, allgemeine Fragen, eigene Texte ohne Bezug zu Personen oder Kunden.
- **Gelb:** interne Informationen ohne Personenbezug – nur in Werkzeugen, die der Betrieb freigegeben hat.
- **Rot:** Daten von Kunden oder Beschäftigten, Gesundheitsdaten, Geschäftsgeheimnisse, Passwörter – nur, wenn ein Werkzeug ausdrücklich dafür freigegeben ist (Vertrag, Auftragsverarbeitung).

Grundlage: DSGVO Art. 5, 6, 9, 28; GeschGehG § 2 Nr. 1 · Die Ampel ist eine Lernhilfe, keine Rechtsprüfung

FOLIE 5 · ANONYMISIEREN

Namen weglassen reicht oft nicht

- „Die Filialleiterin in Bad Honnef, seit März krank“ ist auch ohne Namen eine bestimmbare Person.
- Ersetzen Sie Namen durch Rollen, lassen Sie Orte und Daten weg, runden Sie Zahlen.
- Wenn sich der Fall ohne Personenbezug nicht beschreiben lässt: rot.

Quelle: DSGVO Art. 4 Nr. 1 und Erwägungsgrund 26 (bestimmbare Person)

KI-Antworten klingen sicher – auch wenn sie falsch sind

- Sprachmodelle erzeugen wahrscheinliche Texte. Sie können Fakten, Zahlen, Quellen und Paragraphen erfinden.
- Die Sprache verrät den Fehler nicht: Falsches klingt genauso überzeugend wie Richtiges.
- Verantwortlich für das Ergebnis bleibt, wer es verwendet.

Quelle: BSI, „Generative KI-Modelle – Chancen und Risiken für Industrie und Behörden“, Abschnitt Risiken

Erst prüfen, dann verwenden

1. **Fakten** mit einer Originalquelle abgleichen.
2. **Zahlen** nachrechnen.
3. **Fundstellen** öffnen: Gibt es die Quelle, und steht dort das Behauptete?
4. **Passung** prüfen: Ton, Zielgruppe, interne Vorgaben.

Wenn ein Dokument dem Werkzeug Befehle gibt

- Webseiten, Mails und Dateien können Anweisungen enthalten, die das KI-Werkzeug ausführt („Prompt-Injection“).
- Besonders heikel bei Werkzeugen, die Mails lesen, Dateien öffnen oder selbst handeln dürfen.
- Regel: Ergebnisse aus fremden Inhalten prüfen; Aktionen (senden, löschen, bezahlen) nie ohne eigene Kontrolle.

Quellen: OWASP Top 10 for LLM Applications, LLM01 Prompt Injection; BSI (wie Folie 6)

Kennzeichnen und Rechte beachten

- Täuschend echte Bilder, Videos oder Stimmen (Deepfakes) müssen als KI-erzeugt offengelegt werden.
- Texte, die die Öffentlichkeit über Angelegenheiten von öffentlichem Interesse informieren, ebenfalls – außer sie wurden redaktionell geprüft und jemand trägt die Verantwortung.
- Keine fremden Texte, Bilder oder Kundendokumente hochladen, an denen Sie keine Rechte haben. Interne Regeln gehen weiter.

Quelle: KI-VO (EU) 2024/1689 Art. 50 Abs. 4, gilt seit 02.08.2026 · allgemeine Information, keine Rechtsberatung

Unsere Regeln (vom Betrieb auszufüllen)

- Freigegebene Werkzeuge und wofür:
- Ansprechperson bei Fragen:
- Etwas Falsches eingegeben? Sofort melden an: – der Betrieb muss eine Datenschutzverletzung unter Umständen binnen 72 Stunden der Aufsichtsbehörde melden.

Quelle: DSGVO Art. 33 Abs. 1

Teil B – Moderationsleitfaden (75 Minuten)

Zeit	Inhalt	Methode	Folien
0–5	Einstieg: „Wofür haben Sie KI diese Woche genutzt?“	Zuruf, Stichworte am Flipchart sammeln	1–2
5–20	Was mit Eingaben passiert, Eingabe-Ampel, Anonymisieren	Kurzvortrag mit Rückfragen	3–5
20–35	Übung 1: Eingabe-Ampel	Kleingruppen à 3–4, Karten sortieren, Auswertung im Plenum	4
35–45	Antworten prüfen	Kurzvortrag	6–7
45–60	Übung 2: Fehlersuche	Einzelarbeit 7 Min., Vergleich zu zweit, Auflösung	7

Zeit	Inhalt	Methode	Folien
60– 70	Versteckte Anweisungen, Kennzeichnen und Rechte	Kurzvortrag, ein Beispiel vorlesen	8–9
70– 75	Transfer: Regeln im eigenen Betrieb	Folie 10 gemeinsam ausfüllen oder offene Punkte notieren	10

Hinweise zur Moderation

- Folie 4 und 5 nicht als Rechtsprüfung verkaufen: Die Ampel hilft beim ersten Einordnen. Ob ein Werkzeug für gelbe oder rote Daten freigegeben ist, entscheidet der Betrieb.
- Auf Folie 10 nichts erfinden. Kennt die Gruppe die Regeln nicht, ist das ein Ergebnis: Die offenen Punkte gehen an die zuständige Stelle.
- Für Übung 2 braucht niemand ein KI-Konto. Die Beispielantwort steht unten und wurde für die Übung geschrieben.

Typische Fragen

▼ „Muss jetzt jeder ein Zertifikat nach der KI-Verordnung haben?“

Nein. Art. 4 KI-VO verlangt von Anbietern und Betreibern Maßnahmen, um die KI-Kompetenz ihres Personals zu unterstützen; seit der Änderung durch Verordnung (EU) 2026/1744 (ABl. L vom 24.07.2026) ist kein bestimmtes Kompetenzniveau mehr vorgeschrieben. Ein Zertifikat sieht der Artikel nicht vor. Eine Teilnahmebestätigung kann als Nachweis der Maßnahme dienen.

▼ „Wenn ich die Einstellung ‚nicht zum Training verwenden‘ wähle, darf ich dann alles eingeben?“

Nein. Die Einstellung betrifft nur das Training. Die Eingabe wird trotzdem an den Anbieter übertragen und dort verarbeitet. Für personenbezogene Daten braucht der Betrieb eine Rechtsgrundlage und in der Regel einen Vertrag zur Auftragsverarbeitung (DSGVO Art. 6, 28).

▼ „Darf ich KI-Texte für unsere Website verwenden?“

Nach Prüfung in vier Schritten ja, sofern die internen Regeln es erlauben. Eine Kennzeichnungspflicht nach Art. 50 Abs. 4 KI-VO besteht für Texte zu Angelegenheiten von öffentlichem Interesse, nicht, wenn sie redaktionell geprüft wurden und jemand die Verantwortung trägt. Im Zweifel die zuständige Stelle im Betrieb fragen.

Teil C – Übungen

Übung 1: Eingabe-Ampel (15 Minuten)

Die Gruppe ordnet jede Karte grün, gelb oder rot zu und begründet in einem Satz.

Annahme: Der Betrieb hat ein KI-Werkzeug für interne Informationen freigegeben, aber nicht für personenbezogene Daten.

Nr.	Karte	Lösung	Begründung
1	„Formuliere eine freundliche Absage für eine Messeeinladung.“	grün	Allgemeiner Text ohne Personenbezug.
2	Bewerbungsunterlagen einer Kandidatin zum Zusammenfassen	rot	Personenbezogene Daten, Werkzeug dafür nicht freigegeben.
3	Interne Prozessbeschreibung „Wareneingang“ ohne Namen zum Kürzen	gelb	Intern, ohne Personenbezug: nur im freigegebenen Werkzeug.
4	Krankmeldungen des Teams für eine Urlaubsplanung	rot	Gesundheitsdaten, besondere Kategorie (DSGVO Art. 9).
5	Text der eigenen öffentlichen Stellenanzeige zum Überarbeiten	grün	Bereits veröffentlicht.
6	Kalkulation eines laufenden Angebots mit Einkaufspreisen	gelb bis rot	Geschäftsgeheimnis; je nach Regel des Betriebs nur im freigegebenen Werkzeug oder gar nicht. Gute Diskussionskarte.
7	Zugangsdaten zum Kundenportal, „damit die KI das Formular ausfüllt“	rot	Passwörter nie eingeben.
8	Kundenmail mit Namen und Bestellnummer, „bitte	rot	Personenbezogen. Variante: Namen und Nummer entfernen,

Nr.	Karte	Lösung	Begründung
	Antwort entwerfen"		Anliegen allgemein beschreiben → gelb.

Auswertung: Karte 6 und 8 zeigen, dass die Einordnung von den Regeln des Betriebs abhängt und dass Anonymisieren eine rote Karte gelb machen kann – wenn danach niemand mehr bestimmbar ist (Folie 5).

Übung 2: Fehlersuche (15 Minuten)

Aufgabe: Ein Kollege hat ein KI-Werkzeug gefragt, was der Betrieb zur KI-Schulung tun muss. Finden Sie vier Fehler. Die Antwort wurde für diese Übung geschrieben und enthält absichtlich Fehler.

„Nach Art. 4 der KI-Verordnung müssen alle Beschäftigten, die KI nutzen, bis Ende 2026 ein anerkanntes KI-Zertifikat erwerben. Die DSGVO spielt dabei keine Rolle, solange nur Daten von Geschäftskunden verarbeitet werden. Laut BSI-Leitfaden Nr. 2024-17 genügt eine Schulung von zwei Stunden. Für Ihr Team von 20 Personen rechnen Sie also mit 35 Stunden Schulungszeit.“

▼ Musterlösung

1. **Zertifikat:** Art. 4 KI-VO verlangt Maßnahmen zur Unterstützung der KI-Kompetenz, kein Zertifikat und keine Frist „Ende 2026“ (Fassung VO (EU) 2026/1744). Schritt „Fakten“.
2. **Geschäftskunden:** Auch Ansprechpartner bei Geschäftskunden sind natürliche Personen; ihre Daten fallen unter die DSGVO (Art. 4 Nr. 1). Schritt „Fakten“.
3. **Erfundene Quelle:** Einen „BSI-Leitfaden Nr. 2024-17“ mit dieser Aussage gibt es nicht. Schritt „Fundstellen“.
4. **Rechenfehler:** 20 Personen × 2 Stunden = 40 Stunden, nicht 35. Schritt „Zahlen“.

Teil D – Quellen (Stand 26.10.2026)

- Verordnung (EU) 2024/1689 (KI-Verordnung), Art. 4 in der Fassung der Verordnung (EU) 2026/1744, ABl. L vom 24.07.2026; Art. 50; Art. 113 – eur-lex.europa.eu
- Verordnung (EU) 2016/679 (DSGVO), Art. 4, 5, 6, 9, 28, 33 und Erwägungsgrund 26 – eur-lex.europa.eu
- Gesetz zum Schutz von Geschäftsgeheimnissen (GeschGehG), § 2 – [gesetze-im-internet.de](https://www.gesetze-im-internet.de)

- Datenschutzkonferenz (DSK), Orientierungshilfe „Künstliche Intelligenz und Datenschutz“, Version 1.0, 06.05.2024 – datenschutzkonferenz-online.de
- Bundesamt für Sicherheit in der Informationstechnik (BSI), „Generative KI-Modelle – Chancen und Risiken für Industrie und Behörden“ – bsi.bund.de
- OWASP, „Top 10 for Large Language Model Applications“, LLM01 Prompt Injection – owasp.org

Änderungsprotokoll: 0.9 (26.10.2026) Probefassung. Lizenz des Probemoduls: Sie dürfen es in eigenen Seminaren einsetzen und an Ihre Teilnehmenden ausdrucken. Weiterverkauf und Veröffentlichung als eigenes Material sind nicht erlaubt.