

Sicher arbeiten mit KI

Datenschutz, EU AI Act, vertrauliche Informationen, Phishing und Kontosicherheit

BEISPIELPROMPT

Entferne aus diesem Text alle Angaben, die eine Person erkennbar machen, und ersetze sie durch Platzhalter in eckigen Klammern.

Inhalt

5.1	KI sicher nutzen: Datenschutz	3
5.2	EU AI Act: Was im Büro gilt	5
5.3	Vertrauliche Informationen schützen	8
5.4	Phishing erkennen	10
5.5	Kontosicherheit	13
A	Anhang: Quellen	14

5.1 KI sicher nutzen: Datenschutz

KI-Werkzeuge sind im Büroalltag eine große Hilfe. Damit das so bleibt, kommt es darauf an, was Sie eingeben, wie Sie die Werkzeuge einstellen und wie Sie mit den Ergebnissen umgehen.

Drei Grundsätze für jedes Werkzeug

1. Keine sensiblen Daten eingeben, außer in Werkzeuge, die Ihr Betrieb dafür freigegeben hat.
2. Ergebnisse prüfen, besonders alles, was den Betrieb verlässt.
3. Die Verantwortung bleibt beim Menschen.

Was personenbezogene Daten sind

Personenbezogen ist alles, was sich einer bestimmten Person zuordnen lässt: Name, Adresse, E-Mail-Adresse, Telefonnummer, Kundennummer, Geburtsdatum, Foto oder Stimme. Auch scheinbar harmlose Angaben werden personenbezogen, wenn man aus ihnen zusammen auf eine Person schließen kann, etwa „die neue Kollegin aus dem Einkauf, die im Oktober angefangen hat“.

Besonders geschützt sind Gesundheitsdaten, Religion, politische Meinung, Gewerkschaftszugehörigkeit und biometrische Daten. Krankmeldungen, Atteste und Personalgespräche gehören deshalb nie in ein KI-Werkzeug.

Privatkonto oder Geschäftskonto

In den Business- und Enterprise-Tarifen werden Eingaben standardmäßig nicht zum Training der Modelle verwendet, und der Betrieb schließt mit dem Anbieter einen Vertrag zur Auftragsverarbeitung. Berufliche Inhalte gehören deshalb in ein Geschäftskonto, das Ihr Betrieb verwaltet. Im privaten Konto gelten Ihre privaten Einstellungen, und für die Inhalte gibt es keinen Vertrag mit dem Betrieb.

Datennutzung in privaten Konten abschalten

In den kostenlosen und privaten Tarifen können Ihre Chats zur Verbesserung der Modelle verwendet werden. Sie können das abschalten. Die Einstellung gilt für neue Chats.

Werkzeug	Wo Sie es abschalten	Gut zu wissen
ChatGPT	Profilbild > Einstellungen > Datenkontrollen > „Das Modell für alle verbessern“ ausschalten	Temporäre Chats erscheinen nicht im Verlauf, werden nicht zum Training genutzt und nach 30 Tagen gelöscht.
Claude	Name > Einstellungen > Datenschutz > „Helfen Sie uns, unsere KI-Modelle zu verbessern“ ausschalten	Mit Zustimmung werden Chats bis zu fünf Jahre aufbewahrt. Team und Enterprise sind davon nicht betroffen.
Gemini	Aktivitätseinstellung für Gemini-Apps („Keep Activity“) ausschalten, zu finden unter myactivity.google.com/product/gemini	Temporäre Chats werden nicht zum Training verwendet und bis zu 72 Stunden aufbewahrt.

Wer das Training abschaltet, hat damit noch nichts gelöscht. Ältere Chats bleiben im Verlauf, bis Sie sie selbst entfernen. Ein temporärer Chat eignet sich für einmalige Fragen, ersetzt aber nicht das Weglassen sensibler Daten.

Was die Datenschutzaufsicht sagt

Die Datenschutzkonferenz von Bund und Ländern (DSK) hat eine Orientierungshilfe zu KI und Datenschutz veröffentlicht. Die Kernpunkte für den Büroalltag:

- Personenbezogene Daten können in der Eingabe, im Ergebnis und im Benutzerkonto selbst stecken.
- Wer nur den Namen ersetzt, hat pseudonymisiert. Die Daten bleiben personenbezogen. Erst echte Anonymisierung nimmt sie aus dem Datenschutz heraus.
- Entscheidungen über Menschen dürfen nicht allein auf dem Vorschlag einer KI beruhen. Es braucht einen echten menschlichen Entscheidungsspielraum (Art. 22 DSGVO).
- Betriebe sollen klare Regeln für frei verfügbare und für interne KI-Werkzeuge festlegen und ihre Entscheidungen dokumentieren.

Betriebliche KI-Regeln

Viele Betriebe halten in einer kurzen Richtlinie fest, welche Werkzeuge erlaubt sind, welche Daten hineindürfen, wie KI-Texte gekennzeichnet werden und wer bei Fragen hilft. Gibt es in Ihrem Betrieb noch keine Regel, gilt: nur öffentliche oder unkritische Inhalte eingeben und im Zweifel nachfragen. Ein Serverstandort in der EU allein macht ein Werkzeug noch nicht zulässig. Es zählen auch der Vertrag, der Zweck und die eingegebenen Daten.

5.2 EU AI Act: Was im Büro gilt

Die EU-KI-Verordnung, auch EU AI Act genannt, ist das erste umfassende KI-Gesetz. Für Beschäftigte im Büro sind nur wenige Pflichten unmittelbar wichtig. Dieser Abschnitt zeigt, welche das sind.

Worum es geht

Die Verordnung (EU) 2024/1689 ist seit dem 1. August 2024 in Kraft und gilt in allen Mitgliedstaaten direkt. Ihre Pflichten greifen gestaffelt. Sie regelt KI nach dem Risiko ihres Einsatzes: Je mehr eine Anwendung über Menschen entscheidet, desto strenger die Regeln.

Am 27. Juli 2026 ist die Änderungsverordnung (EU) 2026/1744 in Kraft getreten, bekannt als „Digital Omnibus zur KI“. Sie verschiebt die Pflichten für Hochrisiko-KI, fasst die Pflicht zur KI-Kompetenz weicher und ergänzt zwei Verbote.

Vier Risikostufen

Stufe	Beispiele	Folge
Verboten	Social Scoring; Emotionserkennung am Arbeitsplatz und in Bildungseinrichtungen, außer aus medizinischen oder Sicherheitsgründen; manipulative Techniken, die Schwächen ausnutzen; ab 2. Dezember 2026 auch KI, die ohne Einwilligung realistische intime Darstellungen echter Personen erzeugt	verboten seit 2. Februar 2025, neue Verbote ab 2. Dezember 2026
Hohes Risiko	KI bei Bewerberauswahl, Beförderung, Kündigung und Leistungsbewertung; Kreditwürdigkeitsprüfung; Bewertung von Prüfungen	strenge Pflichten ab 2. Dezember 2027
Transparenz	Chatbots; KI-erzeugte Bilder, Videos und Tonaufnahmen; Deepfakes	Kennzeichnung, seit 2. August 2026
Minimales Risiko	Spamfilter, Rechtschreibprüfung, Textentwürfe, Zusammenfassungen im Büro	keine besonderen Pflichten

ChatGPT, Claude und Gemini sind KI-Modelle mit allgemeinem Verwendungszweck. Die Pflichten dafür treffen die Anbieter, nicht die Nutzerinnen und Nutzer. Für den Betrieb zählt, wofür er das Werkzeug einsetzt: Wer ChatGPT Briefe entwerfen lässt, ist im minimalen Bereich. Wer damit Bewerbungen bewertet oder vorsortiert, bewegt sich im Hochrisikobereich.

Was das für Sie im Büro bedeutet

KI-Kompetenz (Art. 4)

Betriebe, die KI einsetzen, müssen seit dem 2. Februar 2025 die KI-Kompetenz ihrer Beschäftigten fördern. Seit dem Digital Omnibus heißt das: Der Betrieb ergreift Maßnahmen, die die Kompetenzentwicklung unterstützen. Ein bestimmtes Niveau für jede Person muss er nicht garantieren. Ein Zertifikat oder eine Prüfung schreibt die Verordnung nicht vor. Schulungen wie diese sind eine solche Maßnahme.

Kennzeichnen, wo es das Gesetz verlangt (Art. 50 Abs. 4)

- **Deepfakes:** Wer realistische KI-Bilder, -Videos oder -Tonaufnahmen echter Personen, Orte oder Ereignisse veröffentlicht, muss offenlegen, dass sie künstlich erzeugt oder verändert sind. Für künstlerische und satirische Werke gelten Erleichterungen.
- **KI-Texte für die Öffentlichkeit:** Texte, die die Öffentlichkeit über Angelegenheiten von öffentlichem Interesse informieren, müssen als KI-erzeugt gekennzeichnet werden, außer ein Mensch hat sie geprüft und trägt die redaktionelle Verantwortung.
- **Normale Büropost:** Für eine E-Mail an einen Kunden, ein Anschreiben oder ein Protokoll verlangt der AI Act keine Kennzeichnung. Ihr Betrieb kann mehr verlangen. Wer einen Text verschickt, verantwortet ihn.

Verbotene Anwendungen meiden

Ein Werkzeug, das in Videokonferenzen die Stimmung von Kolleginnen und Kollegen auswertet, ist am Arbeitsplatz verboten. Gleiches gilt für KI, die Menschen gezielt manipuliert oder nach ihrem Sozialverhalten bewertet.

Entscheidungen über Menschen bleiben beim Menschen

Die Hochrisiko-Pflichten für KI im Personalbereich greifen ab dem 2. Dezember 2027. Die Datenschutz-Grundverordnung verlangt aber schon heute, dass solche Entscheidungen nicht allein auf einem KI-Vorschlag beruhen (Art. 22 DSGVO). Bewerbungsunterlagen sind personenbezogen und gehören nicht in ein privates KI-Konto.

Fristen im Überblick

Datum	Was gilt
1. August 2024	Verordnung in Kraft
2. Februar 2025	Verbote (Art. 5) und KI-Kompetenz (Art. 4)
2. August 2025	Pflichten für Anbieter von KI-Modellen mit allgemeinem Verwendungszweck, Sanktionsregeln
27. Juli 2026	Digital Omnibus (VO (EU) 2026/1744) in Kraft

Datum	Was gilt
2. August 2026	allgemeiner Geltungsbeginn, unter anderem Transparenzpflichten nach Art. 50
2. Dezember 2026	zwei neue Verbote; maschinenlesbare Kennzeichnung auch für generative Systeme, die vor August 2026 auf den Markt kamen
2. Dezember 2027	Hochrisiko-KI nach Anhang III, zum Beispiel im Personalbereich
2. August 2028	Hochrisiko-KI in regulierten Produkten nach Anhang I

Aufsicht in Deutschland

Das deutsche Durchführungsgesetz (KI-Marktüberwachungsgesetz, KI-MIG) ist am 29. Juli 2026 in Kraft getreten. Die Bundesnetzagentur ist zentrale Marktüberwachungsbehörde. Mit dem KI-Service-Desk bietet sie eine Anlaufstelle besonders für kleine und mittlere Unternehmen. Fachbehörden wie die BaFin bleiben in ihren Bereichen zuständig, die Datenschutzbehörden für alles, was den Datenschutz betrifft.

Bußgelder

Für verbotene Praktiken drohen bis zu 35 Millionen Euro oder 7 % des weltweiten Jahresumsatzes, bei kleinen und mittleren Unternehmen gilt der niedrigere Betrag. Für die KI-Kompetenz nach Art. 4 sieht die Verordnung kein eigenes Bußgeld vor.

5.3 Vertrauliche Informationen schützen

Nicht nur personenbezogene Daten sind schützenswert. Kalkulationen, Konditionen und interne Abläufe sind es auch. Oft kommt man mit Platzhaltern zum selben Ergebnis, ohne etwas Vertrauliches preiszugeben.

Anonymisieren mit Platzhaltern

ORIGINALTEXT · FIKTIV · NICHT IN EIN KI-WERKZEUG

Guten Tag, hier ist Andrea Brandt, Lindenweg 12, 42929 Wermelskirchen, Kundennummer K-48211. Meine Bestellung B-1002 über 698,00 € ist immer noch nicht da. Meine Tochter braucht den Schreibtisch dringend für die Schule. Rufen Sie mich bitte unter 0171 2345678 zurück.

ANONYMISIERTE FASSUNG

Guten Tag, hier ist [Kundin], [Adresse], Kundennummer [Kundennummer]. Meine Bestellung [Bestellnummer] über [Betrag] ist immer noch nicht da. Ich brauche den Schreibtisch dringend. Rufen Sie mich bitte unter [Telefonnummer] zurück.

Entfernt sind außerdem die Angaben zur Tochter. Sie sind für die Antwort unwichtig und betreffen eine weitere Person.

BEISPIELPROMPT · MIT PLATZHALTERN

Formuliere eine Antwort an [Kundin] zur Bestellung [Bestellnummer]. Die Lieferung erfolgt am [Termin]. Wir rufen wie gewünscht zurück. Sie-Form, höchstens 80 Wörter.



Die Antwort der KI enthält dieselben Platzhalter. Die echten Angaben setzen Sie erst danach in Outlook oder Word ein, also außerhalb des KI-Werkzeugs.

Pseudonymisieren oder anonymisieren?

Ersetzt man nur den Namen, bleiben Adresse, Kundennummer oder ein bekannter Vorfall als Hinweis auf die Person. Das ist Pseudonymisierung, die Daten bleiben personenbezogen. Echte Anonymisierung entfernt jeden Bezug. Für Prompts genügt meist: alles weglassen, was die KI für die Aufgabe nicht braucht.

Dokumente vor dem Hochladen prüfen

Word-Dateien tragen oft mehr mit sich, als man sieht: Kommentare, frühere Änderungen, Autorennamen, ausgeblendeten Text. Bevor Sie eine Datei in ein KI-Werkzeug hochladen:

- 1 **Datei > Informationen > Auf Probleme überprüfen > Dokument prüfen.**
- 2 Kommentare, Dokumenteigenschaften und ausgeblendeten Text auswählen, dann **Prüfen**.
- 3 Gefundene Einträge mit **Alle entfernen** löschen. Vorher eine Kopie speichern.

So stufen Sie Inhalte ein

Inhalt	Stufe	Umgang mit KI
Pressemitteilung vom letzten Monat	öffentlich	in allen freigegebenen Werkzeugen
Ablaufbeschreibung „Kundenanfrage beantworten“	intern	nur mit Geschäftskonto in freigegebenen Werkzeugen
Besprechungsnotizen ohne Namen	intern	nur mit Geschäftskonto in freigegebenen Werkzeugen
Angebot mit Kalkulation	vertraulich	nicht eingeben oder nur nach ausdrücklicher Freigabe
Krankmeldung einer Kollegin	vertraulich	nie in ein KI-Werkzeug
Zugangsdaten zum Kundenportal	vertraulich	nie in ein KI-Werkzeug

Geschäftsgeheimnisse

Kalkulationen, Einkaufskonditionen, Kundenlisten, Quellcode und Pläne sind Geschäftsgeheimnisse. Das Geschäftsgeheimnisgesetz schützt sie nur, wenn der Betrieb angemessene Geheimhaltungsmaßnahmen trifft. Wer solche Inhalte in ein privates KI-Konto kopiert, kann diesen Schutz gefährden, auch wenn nie etwas nach außen dringt.

5.4 Phishing erkennen

KI macht Betrugsnachrichten fehlerfrei und persönlich. Rechtschreibfehler taugen kaum noch als Erkennungszeichen. Es gibt aber Merkmale, die auch eine perfekt formulierte Nachricht verraten.

Ein Beispiel

Von Buchhaltung <rechnung@beispielbetrieb-rechnung.example>
Betreff Dringend: Offene Rechnung heute freigeben
Anhang Rechnung_2026-10.html

Guten Tag,

die Rechnung zum Auftrag vom 2. Oktober ist noch nicht freigegeben. Bitte erledigen Sie das heute bis 14:00 Uhr, sonst entfällt das vereinbarte Skonto.

Die Freigabe läuft diesmal über unser neues Zahlungsportal: [Rechnung ansehen und freigeben](#)

Ich bin heute in Terminen und telefonisch nicht erreichbar. Vielen Dank für die schnelle Erledigung.

Mit freundlichen Grüßen
Ihre Buchhaltung

Sechs Merkmale

Merkmal	Im Beispiel	Warum verdächtig
Absender	„Buchhaltung“ mit Adresse @beispielbetrieb-rechnung.example	Anzeigename und Domain passen nicht zum eigenen Betrieb
Zeitdruck	„Dringend“, „heute bis 14:00 Uhr“, Skonto entfällt	soll Nachdenken und Rückfragen verhindern
Ungewöhnliche Bitte	Zahlung über ein neues Portal freigeben	weicht vom üblichen Freigabeweg ab
Link	„Rechnung ansehen und freigeben“	Ziel nur beim Darüberfahren sichtbar, nicht klicken
Keine Rückfrage möglich	„in Terminen und telefonisch nicht erreichbar“	verhindert die Prüfung auf einem zweiten Weg

Merkmal	Im Beispiel	Warum verdächtig
Anhang	Rechnung_2026-10.html	HTML-, ZIP- und Makro-Dateien sind häufige Träger von Schadsoftware

Merksatz

Zahlungen, Kontoänderungen und Passwörter werden nie allein auf Grund einer E-Mail, Nachricht oder eines Anrufs geändert oder freigegeben. Ein zweiter, bekannter Kommunikationsweg ist Pflicht.

Was KI am Betrug verändert

- **Fehlerfreie, persönliche Mails:** Betrüger lassen Texte von KI schreiben und bauen öffentlich auffindbare Details ein, etwa Namen aus dem Impressum oder aktuelle Projekte.
- **Geklonte Stimmen:** Wenige Sekunden Aufnahme reichen, um eine Stimme nachzubilden. Der Anruf „der Geschäftsführung“ mit einer dringenden Überweisung ist der klassische Fall. Bei ungewöhnlichen Bitten: auflegen und unter der bekannten Nummer zurückrufen.
- **QR-Codes:** In Briefen, auf Paketzetteln oder Parkautomaten führen sie auf gefälschte Seiten. Die Adresse ist vor dem Öffnen kaum zu prüfen.
- **Gefälschte Anmeldeseiten:** Nachgebaute Microsoft- oder Paketdienst-Seiten sehen dem Original zum Verwechseln ähnlich.

KI als zweite Meinung

Bei einer verdächtigen Nachricht können Sie den Text in ChatGPT, Claude oder Gemini kopieren, ohne Links anzuklicken und ohne den Anhang.

BEISPIELPROMPT

Ich habe diese E-Mail erhalten. Nenne alle Merkmale, die auf Phishing hindeuten, und sag mir, wie ich die Nachricht über einen zweiten Weg prüfen kann, ohne auf etwas in der Mail zu reagieren.

[Text der E-Mail]



Die KI findet typische Merkmale meist zuverlässig und erklärt sie verständlich. Den echten Absender und das Linkziel kann sie aber nicht prüfen. Die Entscheidung und der Rückruf unter der bekannten Nummer bleiben bei Ihnen.

Melden in Outlook

Im neuen Outlook die verdächtige Nachricht markieren und im Menüband **Melden** bzw. **Phishing melden** wählen, sofern angeboten. Im Betrieb gilt der Meldeweg der IT. Die Nachricht nicht weiterleiten.

Wenn Sie schon geklickt haben: sofort die IT informieren und das Passwort ändern. Schnelles Melden begrenzt den Schaden.

5.5 Kontosicherheit

Ein KI-Konto enthält Chatverläufe, hochgeladene Dateien, Erinnerungen und oft Verbindungen zu Postfach und Cloud. Es verdient denselben Schutz wie Ihr E-Mail-Konto.

Die wichtigsten Einstellungen

Was	Wo Sie es einstellen
Mehr-Faktor-Anmeldung im Microsoft-Konto	account.microsoft.com > Sicherheit > Erweiterte Sicherheitsoptionen
Mehr-Faktor-Anmeldung in ChatGPT	Einstellungen > Sicherheit
Geteilte Chat-Links in ChatGPT	Einstellungen > Datenkontrollen > Geteilte Links > Verwalten
Angemeldete Geräte	in den Kontoeinstellungen des jeweiligen Dienstes; alte Sitzungen beenden
Verbundene Dienste	in den Einstellungen des KI-Werkzeugs unter Apps, Konnektoren oder Verbindungen

Die Bezeichnungen ändern sich gelegentlich. Wer einen geteilten Chat-Link kennt, kann den Chat lesen. Alte Freigaben sollten Sie deshalb regelmäßig löschen.

Passkeys statt Passwort

Ein Passkey ersetzt das Passwort durch einen Schlüssel, der auf Ihrem Gerät liegt und mit Fingerabdruck, Gesichtserkennung oder Geräte-PIN freigegeben wird. Weil der Schlüssel nur auf der echten Seite funktioniert, läuft eine gefälschte Anmeldeseite ins Leere. Das Bundesamt für Sicherheit in der Informationstechnik (BSI) empfiehlt Passkeys ausdrücklich als Schutz vor Phishing.

Wo kein Passkey angeboten wird: lange, für jeden Dienst unterschiedliche Passwörter aus einem Passwortmanager und zusätzlich die Mehr-Faktor-Anmeldung. Ein starkes Passwort schützt gegen Erraten, aber nicht gegen Phishing. Wer es auf einer gefälschten Seite eingibt, hat es verschenkt.

Keine geteilten Teamkonten

Jede Person braucht ein eigenes Konto. Bei geteilten Konten sieht jeder alle Verläufe, niemand kann nachvollziehen, wer was eingegeben hat, und meist verstößt es gegen die Nutzungsbedingungen.

A Quellen

Die Angaben zum EU AI Act wurden im Oktober 2026 an diesen Seiten geprüft, ebenso die Einstellungen der Werkzeuge.

EU AI Act und KI-MIG

- Digital Omnibus zur KI: Was am AI Act wirklich verändert wird (ITMR Legal)
itmr-legal.de/blog/digitaler-omnibus-in-kraft
- KI-Omnibus: Was ändert sich bei der KI-Kompetenz nach Artikel 4? (Mittelstand-Digital Zentrum Fokus Mensch, 19. Oktober 2026)
digitalzentrum-fokus-mensch.de/kos/WNetz?art=News.show&id=3120
- Der Zeitplan des EU AI Act: welche Pflicht ab wann
caroline-herrmann.com/ki-vo-fristen/
- Digital Omnibus: Was ändert sich an der KI-Verordnung? (Haufe, 20. Juli 2026)
haufe.de/finance/steuern-finanzen/digital-omnibus-was-aendert-sich-an-der-ki-verordnung_190_691870.html
- KI-MIG beschlossen: Was das AI-Act-Durchführungsgesetz für Unternehmen bedeutet
advisori.de/blog/ki-mig-ai-act-durchfuehrungsgesetz-unternehmen

Datenschutz

- Datenschutzaufsichtsbehörden veröffentlichen Orientierungshilfe zu KI (Gleiss Lutz)
gleisslutz.com/de/know-how/datenschutzaufsichtsbehoerden-veroeffentlichen-orientierungshilfe-zu-ki
- ChatGPT-Hilfe: FAQ zu Datenkontrollen
help.openai.com/de-de/articles/7730893-data-controls-faq
- Claude-Datenschutzcenter: Einstellungen zur Modellverbesserung ändern
privacy.claude.com/de/articles/12109829-wie-andere-ich-meine-datenschutzeinstellungen-zur-modellverbesserung
- Gemini: Temporäre Chats und Datenschutzeinstellungen
blog.google/products-and-platforms/products/gemini/temporary-chats-privacy-controls/

Kontosicherheit

- Besser als Passwörter: BSI empfiehlt Passkeys (it-daily, 2. Oktober 2024)
it-daily.net/shortnews/besser-als-passwoerter-bsi-empfiehl-passkeys