KI-Richtlinie der ZeoAI

Version: 1.1 | Stand: März 2026 | Verantwortlich: Geschäftsführung

---

Inhaltsverzeichnis

1. Zielsetzung und Präambel

Diese Richtlinie bildet den verbindlichen Handlungsrahmen für die Entwicklung, den Betrieb und die Nutzung von Künstlicher Intelligenz (KI) bei ZeoAI.

ZeoAI verpflichtet sich, die Innovationspotenziale der KI zu nutzen und gleichzeitig sicherzustellen, dass jeder Einsatz von KI ethisch vertretbar, rechtssicher und transparent erfolgt.

Diese Richtlinie dient dazu:

- Rechtssicherheit im Umgang mit KI-Systemen (insb. EU AI Act, DSGVO) zu gewährleisten.
- Risiken für Kunden, Mitarbeitende und Partner proaktiv zu minimieren.
- Transparenz und Vertrauen in digitale Prozesse zu schaffen.

2. Geltungsbereich

Die Vorgaben gelten für alle Geschäftsbereiche und Mitarbeitenden von ZeoAI.

Sie erstrecken sich zudem auf externe Dienstleister und Freelancer, soweit diese in unserem Auftrag KI-Lösungen entwickeln oder betreiben.

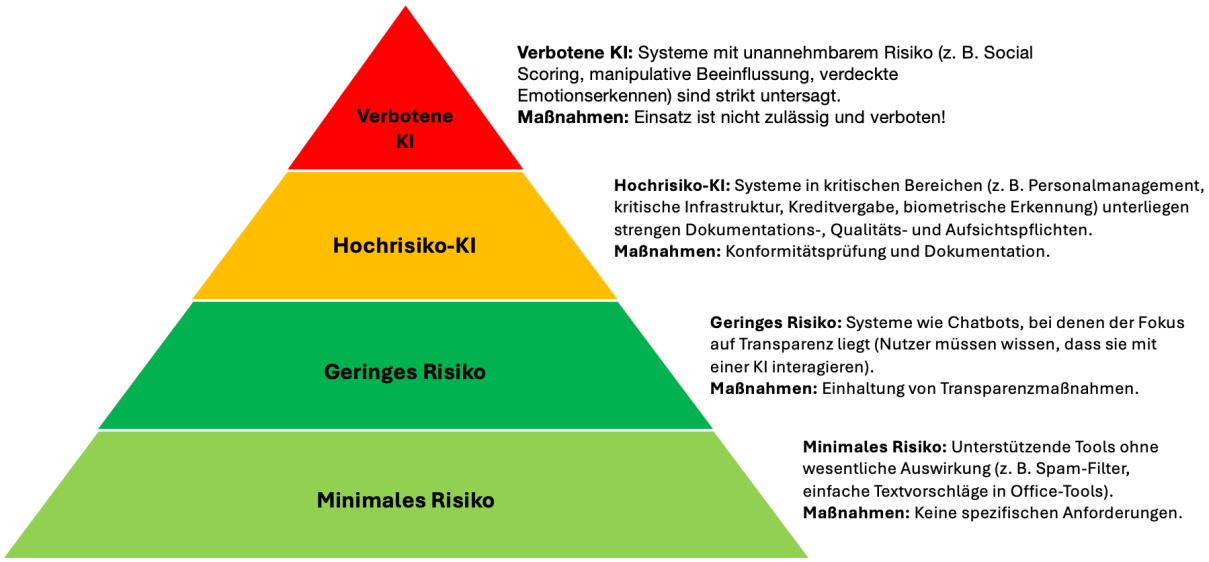
Trennung von Privat & Beruf: Die Nutzung privater Accounts (z. B. private E-Mail-Adressen für ChatGPT Free) für geschäftliche Zwecke oder die Verarbeitung von Kundendaten ist ausdrücklich untersagt, um den unkontrollierten Abfluss vertraulicher Informationen zu verhindern.

3. Definition von KI-Systemen

ZeoAI orientiert sich an der Definition des Art. 3 Nr. 1 der EU-KI-Verordnung (AI Act). Demnach ist ein KI-System ein maschinengestütztes System, das für einen in unterschiedlichem Grade autonomen Betrieb ausgelegt ist und das aus erhaltenen Eingaben ableitet, wie Ausgaben (z. B. Vorhersagen, Inhalte, Empfehlungen oder Entscheidungen) erstellt werden.

4. Risikoklassifizierung

Um angemessene Sicherheitsmaßnahmen zu gewährleisten, kategorisiert ZeoAI jedes KI-System vor dem Einsatz gemäß den Vorgaben der EU-KI-Verordnung in vier Risikoklassen.



- Verbotene KI: Systeme mit unannehmbarem Risiko (z. B. Social Scoring, manipulative Beeinflussung, verdeckte Emotionserkennung) sind strikt untersagt. Maßnahmen: Einsatz ist nicht zulässig und verboten!
- Hochrisiko-KI: Systeme in kritischen Bereichen (z. B. Personalmanagement, kritische Infrastruktur, Kreditvergabe, biometrische Erkennung) unterliegen strengen Dokumentations-, Qualitäts- und Aufsichtspflichten. Maßnahmen: Konformitätsprüfung und Dokumentation.
- Geringes Risiko: Systeme wie Chatbots, bei denen der Fokus auf Transparenz liegt (Nutzer müssen wissen, dass sie mit einer KI interagieren). Maßnahmen: Einhaltung von Transparenzmaßnahmen.
- Minimales Risiko: Unterstützende Tools ohne wesentliche Auswirkung (z. B. Spam-Filter, einfache Textvorschläge in Office-Tools). Maßnahmen: Keine spezifischen Anforderungen.

5. Grundprinzipien für den KI-Einsatz

Jeder Umgang mit KI bei ZeoAI muss folgenden Prinzipien genügen:

1. Rechtmäßigkeit & Compliance: Einhaltung aller gesetzlichen Vorgaben, insbesondere der DSGVO, des Urheberrechts und des EU AI Act.
2. Transparenz: KI-generierte Inhalte und Interaktionen müssen für Kunden und Nutzer als solche erkennbar sein (Kennzeichnungspflicht).
3. Menschliche Aufsicht („Human-in-the-Loop"): Kritische Entscheidungen oder die Finalisierung von Kundeninhalten dürfen nicht vollautomatisiert ohne menschliche Kontrolle erfolgen.
4. Ethik & Fairness: ZeoAI achtet aktiv darauf, Bias (Voreingenommenheit) in KI-Modellen zu erkennen und diskriminierende Ergebnisse zu vermeiden.

6. Datenschutz und Datensicherheit

Der Schutz personenbezogener und vertraulicher Daten hat oberste Priorität.

- Datenschutz-Folgenabschätzung (DSFA): Soweit KI-Systeme personenbezogene Daten verarbeiten, prüft ZeoAI vorab die Risiken gemäß Art. 35 DSGVO.
- Datenverarbeitung & Server: ZeoAI bevorzugt Anbieter, die Daten auf sicheren Servern innerhalb der EU/EWR verarbeiten oder setzt auf lokale Ausführung (siehe Anlage A).
- Kein Training mit Kundendaten: Ohne explizite vertragliche Regelung (z. B. Enterprise-Verträge) dürfen keine vertraulichen Kundendaten in öffentliche KI-Modelle eingegeben werden, die diese Daten zum Training nutzen könnten.
- Recht auf menschliche Entscheidung: Betroffene Personen haben gemäß Art. 22 DSGVO das Recht, keiner ausschließlich auf einer automatisierten Verarbeitung beruhenden Entscheidung unterworfen zu werden.

7. Sicherheitsmaßnahmen und Notfallmanagement

Um die Betriebssicherheit zu gewährleisten, gelten folgende Maßnahmen:

- Monitoring: KI-Systeme werden regelmäßig auf Fehlfunktionen oder "Halluzinationen" (falsche Ausgaben) überwacht.
- Notfallplan (Kill-Switch): Bei erkannten Risiken oder Fehlfunktionen, die Nutzer gefährden könnten, wird das betroffene System unverzüglich deaktiviert.
- Meldewege: Sicherheitsvorfälle im Zusammenhang mit KI sind umgehend der Geschäftsführung zu melden. Sofern ein externer Datenschutzbeauftragter bestellt ist, wird dieser ebenfalls informiert.

8. Zulässige Nutzung und Tools

Es dürfen ausschließlich die von ZeoAI freigegebenen und geprüften KI-Tools genutzt werden (siehe Anlage A). Der Einsatz von „Schatten-IT“ ist untersagt. Die Nutzung neuer KI-Lösungen bedarf der vorherigen Genehmigung durch die Geschäftsführung.

Die zulässigen Anwendungsbereiche sind in der Anwendungsmatrix (Anlage B) definiert.

9. Schulung und Verantwortung

Die Geschäftsführung trägt die Gesamtverantwortung für die Einhaltung dieser Richtlinie. Alle Mitarbeitenden, die mit KI-Technologien arbeiten, sind verpflichtet, an regelmäßigen Schulungen zu Themen wie Prompting, Datensicherheit und rechtlichen Grundlagen teilzunehmen.

Anforderungen gemäß Art. 4 EU-KI-Verordnung (KI-Kompetenz):

Die Geschäftsführung stellt sicher, dass alle Personen, die KI-Systeme betreiben oder nutzen, über ein ausreichendes Maß an KI-Kompetenz verfügen. Schulungsmaßnahmen werden dokumentiert und können auf Anfrage nachgewiesen werden.

10. Aktualisierung

Diese Richtlinie wird mindestens einmal jährlich sowie bei relevanten technologischen oder rechtlichen Änderungen (z. B. Updates des EU AI Act) überprüft und aktualisiert.

Nächste planmäßige Überprüfung: Q1 2027 oder früher bei relevanten Änderungen der EU-KI-Verordnung.

Anlage A: Liste erlaubter KI-Tools & Infrastruktur

1. Generative KI (Cloud-basiert)

Nutzung für Standard-Aufgaben unter Beachtung der Datenschutzvorgaben (keine sensiblen Personendaten in Standard-Prompts).

- ChatGPT (OpenAI – Team/Enterprise Lizenz)
- Claude (Anthropic – Pro/Team Lizenz)
- Gemini (Google)
- Langdock (KI-Plattform & Orchestrierung – Text, Workflows & Agenten; Business Lizenz)

2. Lokale KI & Entwicklung (Maximum Privacy)

Nutzung für sensible Daten, Entwicklung und RAG-Systeme, da die Verarbeitung lokal oder auf eigenen Servern erfolgt (kein Datenabfluss).

- Ollama (Lokale Ausführung von Open-Source Modellen)
- LM Studio (Lokale LLM-Umgebung)
- Cursor (KI-gestützte Code-Entwicklung)
- Hostinger Horizons (KI-gestützte Code-Entwicklung)

3. Datenbanken & Vektor-Speicher

Infrastruktur für KI-Gedächtnis, Embeddings und Retrieval-Augmented Generation (RAG).

- Chroma DB (Vektordatenbank)
- MariaDB (Relationale Datenbank)
- Airtable (Low-Code Datenbank)
- Supabase (Low-Code Datenbank)

4. Prozess-Automatisierung & Agenten

Orchestrierung von KI-Workflows.

- n8n (Workflow-Automatisierung – Self-Hosted bevorzugt für sensible Daten)
- Make.com (SaaS-Automatisierung)
- Zapier (Schnittstellen-Integration)

5. Workspace & Kommunikation

Integrierte KI-Funktionen in bestehender Business-Software.

- Notion (Wissensmanagement)
- Google Workspace / Drive
- Trello / Jira / Slack
- Zoho One (CRM & Business Suite)
- Fonio.ai (KI-Telefonie)
- Bots4You (KI-Telefonie)
- Pipedrive (CRM)

6. Design & Medien

- Canva (Magic Studio)
- Midjourney / Dall-E 3
- Freepik
- Nano Banana 2 / Gemini 3 Flash Image (Google – KI-Bildgenerierung in Gemini)

Anlage B: Nutzungs-Matrix & Kennzeichnung

(Festlegung zulässiger Anwendungsfälle und Dokumentationspflichten gemäß KI-Richtlinie Punkt 8)

I. Text-Generierung & Code

Anwendungsfall	Zulässig?	Kennzeichnung (extern)?	Dokumentation (intern)?
Interne Dokumente (Wikis, Konzepte, E-Mails intern, Präsentationen, Schulungen)	✅ Ja	❌ Nein	❌ Nein
Präsentationen & Schulungsmaterial	✅ Ja	❌ Nein	❌ Nein
Öffentliche Werbung, Blog & Social Media	✅ Ja	✅ Ja	✅ Ja
Produktbeschreibungen / Webseite	✅ Ja	✅ Ja	✅ Ja
Programmcode / Softwareentwicklung	✅ Ja	❌ Nein	✅ Ja (im Versionsverlauf)
Individuelle Kundenkommunikation (z.B. 1:1 E-Mails)	✅ *Mit Prüfung*	❌ Nein	❌ Nein
Mitarbeiterbeurteilungen / HR-Entscheidungen	❌ Nein	-	-
Rechtsverbindliche Verträge (AGB, AVV, Auftragsbestätigung, etc. werden vom GF immer gegengelesen)	✅ Mit Prüfung	❌ Nein	❌ Nein
Rechtsverbindliche Verträge als Dienstleistung für Dritte	❌ Nein	-	-

II. Bild- & Medienerstellung

Anwendungsfall	Zulässig?	Kennzeichnung (extern)?	Dokumentation (intern)?
Interne Nutzung (Präsentationen, Moodboards)	✅ Ja	❌ Nein	❌ Nein
Webseite, Blog & Social Media	✅ Ja	✅ Ja	❌ Nein
Werbeanzeigen (Ads)	✅ Ja	✅ Ja	✅ Ja
Deepfakes / Manipulation echter Personen	🚫 **Verboten**	-	-

Versionshistorie:

Version	Stand	Änderungen
1.0	24.11.25	Erstveröffentlichung
1.1	Mär 2026	Art. 4 KI-VO ergänzt (§9); Grammatikfehler §1 behoben; Tippfehler §4 („Emotionserkennung"); DSB-Formulierung §7 angepasst; Langdock + Claude-Lizenz in Anlage A ergänzt; Art. 6/Anhang III Verweis §4; „Beratungsunternehmen" → ZeoAI; Verträge Anlage B differenziert (eigene GF vs. Dritte); Nano Banana 2 /