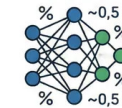


KI verstehen Chancen, Risiken, Verantwortung

9 Kernthemen für den verantwortungsvollen Umgang mit Künstlicher Intelligenz

KI verstehen – Chancen, Risiken, Verantwortung

⚡ Was KI konkret bedeutet



Generative KI = Wahrscheinlichkeitsmaschine
KI versteht nichts – sie berechnet Wahrscheinlichkeiten
Next-Token-Prediction
Metapher: Autovervollständigung auf Steroiden

⚠ Warum KI halluziniert



Halluzination = plausibel klingende Falschaussage
Modell muss immer antworten
Lücken im Training
Systemische Eigenschaft, kein Bug

✅ Schutz vor Halluzinationen



Risikominderung – keine Garantie
1. Gute Prompts
2. Aufgabenraum begrenzen
3. Rückfragen erzwingen
4. Quellenprüfung durch Menschen
5. Human-in-the-Loop

⚖ Verantwortung



KI-Anbieter: Modell, Infrastruktur
Organisation: Toolauswahl, Datenschutz, Schulung
Nutzende Person: Prüfung, Haftung
Die KI hat keine Verantwortung. Punkt.

⚖ Bias



Bias = systematische Verzerrung
Trainingsdaten aus realen Gesellschaften
Gesellschaften sind nicht neutral
KI verstärkt bestehende Muster

🌐 Geopolitik



Technologische und politische Dimension
Führende KI-Modelle aus USA
Digitale Abhängigkeit
KI ist nicht nur Technik – sie ist Geopolitik

🔒 DSGVO & Datenschutz



Nicht alles Technische ist rechtlich erlaubt
Keine personenbezogenen Daten in öffentliche Tools
Anonymisierung reicht nicht immer
Datenschutz ist Führungsaufgabe

🧠 Automation Bias



Kognitive Risiken
Menschen übernehmen KI-Vorschläge ungeprüft
Nachlassen eigener Kompetenz
KI als Sparringspartner, nicht als Entscheider

🤝 Human-Centered AI



Die Kernbotschaft
KI als Werkzeug verstehen
Ergänzt menschliche Verantwortung
Ersetzt niemals Ethik und Entscheidungskraft

Überblick: 9 Kernthemen

- | | | | |
|---|----------------------------|---|-------------------------|
| 1 | Was KI konkret bedeutet | 6 | Geopolitische Dimension |
| 2 | Warum KI halluziniert | 7 | DSGVO & Datenschutz |
| 3 | Schutz vor Halluzinationen | 8 | Automation Bias |
| 4 | Verantwortung und Rollen | 9 | Human-Centered AI |
| 5 | Bias und Verzerrung | | |

i Diese Präsentation vermittelt grundlegendes Verständnis für den sicheren und verantwortungsvollen Einsatz von KI-Systemen.



Was KI konkret bedeutet

Generative KI „versteht“ nichts – sie ist eine Wahrscheinlichkeitsmaschine.

- Basierend auf **Next-Token-Prediction**
- Berechnet statistisch wahrscheinlichstes nächstes Wort
- Trainiert mit Milliarden von Wörtern
- Große neuronale Netze (Large Language Models)

⊘ Was KI NICHT hat:

- | | |
|--------------------------|-------------------------|
| ✗ Keine Intention | ✗ Kein Weltverständnis |
| ✗ Kein Faktenbewusstsein | ✗ Kein Wahrheitsbegriff |



METAPHER

*„Autovervollständigung auf Steroiden
– extrem leistungsfähig, aber ohne
Verständnis.“*

Warum KI halluziniert



DEFINITION

„Plausibel klingende Falschaussage“

Die 5 Hauptursachen:

-  **Antwortpflicht:** Modell muss immer antworten
-  **Datenlücken:** Führen zu Spekulationen
-  **Unscharfe Prompts:** Zu offene Anfragen
-  **Erwartungsdruck:** „Gib mir eine Quelle“
-  **Veraltetes Wissen:** Widersprüche im Training



Wichtig: Halluzinationen sind keine Fehler, sondern eine **systemische Eigenschaft** probabilistischer Modelle.

Schutz vor Halluzinationen

Risikominderung durch Human-in-the-Loop

1 Gute Prompts schreiben

Klarer Kontext, klare Aufgabe, klares Format definieren.

2 Aufgabenraum begrenzen

Lieber „Zusammenfassen“ als „neu erfinden“. Vorhandene Texte prüfen lassen.

3 Rückfragen erzwingen

Anweisung: „Wenn Informationen fehlen, stelle Rückfragen.“

4 Quellenprüfung durch Menschen

KI nennt Quellen → Mensch prüft sie. Verantwortung bleibt extern.



Protection Against
Hallucinations

HUMAN-IN-THE-LOOP

KI = Vorschlagssystem
Mensch = Entscheidungssystem

**KI reduziert Arbeit – sie
übernimmt keine
Verantwortung.**

Wer trägt welche Verantwortung?



MERKSATZ

**„Die KI hat keine
Verantwortung. Punkt.“**

KI-Anbieter

- Modell, Infrastruktur, Trainingsdaten
- Technische Sicherheitsmaßnahmen
- Kontinuierliche Weiterentwicklung

Organisation / Arbeitgeber

- Toolauswahl und Freigabeprozesse
- Datenschutzkonzept & Nutzungsregeln
- Schulung der Mitarbeitenden

Nutzende Person

- Prüfung der Ergebnisse
- Sachliche und rechtliche Verantwortung
- Haftung im Arbeitskontext
- Kritische Bewertung der Outputs

Bias – Warum KI nicht neutral ist



DEFINITION

Bias = systematische Verzerrung in Daten und Ergebnissen

? Warum unvermeidbar?

- > Trainingsdaten stammen aus **realen Gesellschaften**
- > Gesellschaften sind **nicht neutral**
- > Ungleichheiten & Stereotype spiegeln sich in Texten wider

! Typische Beispiele



Geschlecht



Herkunft



Alter



Status

→ **Konsequenz: KI verstärkt bestehende Muster, wenn niemand gegensteuert.**

Geopolitik und Abhängigkeit

KI ist nicht nur Technik – sie ist Geopolitik



FAKTENLAGE

- Führende Modelle überwiegend aus USA
- Infrastruktur: Cloud, Chips, Basismodelle
- Europäische Alternativen im Aufbau

Risiken der digitalen Abhängigkeit

Technologische Abhängigkeit

Begrenzte Kontrolle über Weiterentwicklung und Abhängigkeit von Geschäftsmodellen.

Politische Einflussmöglichkeiten

Exportkontrollen (z.B. Chips), Änderungen in AGBs, Datenzugriff.

Strategische Souveränität

Notwendigkeit eigener Kapazitäten; kritische Infrastruktur schützen.



Diskussionspunkt: KI-Entwicklung ist eine Frage nationaler und europäischer Souveränität.

DSGVO & Datenschutz

Datenschutz ist Führungsaufgabe, nicht Technikproblem



Datenschutz und DSGVO

GRUNDPRINZIP

Nicht alles, was technisch geht,
ist rechtlich erlaubt.

Zentrale Regeln

- ❌ Keine personenbezogenen Daten in öffentliche Tools
- ❌ Keine sensiblen Daten (Gesundheit, Religion etc.)
- ⚠️ Anonymisierung reicht nicht immer aus
- ⚠️ Kostenlose Tools = Hohes Risiko (Training)

Was Organisationen brauchen

- ✅ Klare Regeln & Richtlinien
- ✅ Freigabeprozesse
- ✅ Datenschutzkonforme Tools
- ✅ Schulung der Mitarbeitenden



Merksatz: Datenschutz ist keine technische Frage, sondern eine Führungsaufgabe.

Automation Bias

Kognitive Risiken beim KI-Einsatz



AUTOMATION BIAS

DEFINITION

Menschen übernehmen KI-Vorschläge ungeprüft, weil sie „kompetent wirken“.

! Risiken

- 🧠 **Nachlassen der Fachkompetenz:** Verlust von Wissen durch Delegation.
- ⚖️ **Verlust von Urteilskraft:** Kritisches Denken wird ausgelagert.
- 🚶 **Scheinsicherheit:** Fehler in „smarten“ Antworten werden übersehen.

🛡️ Gegenmittel

- ✓ KI als **Sparringspartner** nutzen, nicht als Entscheider.
- ✓ Bewusste **Reflexionspausen** einbauen.
- ✓ Klare Regeln: **Was entscheidet immer der Mensch?**

Human-Centered AI

Die Kernbotschaft: Mensch im Zentrum



VISION

KI als Werkzeug, das menschliche Fähigkeiten ergänzt – aber niemals ersetzt.

4 Grundprinzipien



Technologie dient dem Menschen

KI unterstützt Entscheidungen, Mensch behält finale Kontrolle.



Ethik bleibt menschlich

Moralische Entscheidungen können nicht delegiert werden.



Transparenz

Verstehen, wie KI zu Ergebnissen kommt (Erklärbarkeit).



Kontinuierliches Lernen

Menschen müssen KI verstehen und kritisch hinterfragen.

KI ist ein mächtiges Werkzeug – aber **Verantwortung, Ethik und Entscheidungskraft** bleiben menschlich.

Zusammenfassung

9 Kernbotschaften für den verantwortungsvollen KI-Einsatz

Diskussion & Fragen

| Ihre Fragen und Erfahrungen



DISKUSSIONSANREGUNGEN



Welche Erfahrungen haben Sie mit KI-Tools gemacht?



Wo sehen Sie in Ihrem Arbeitsbereich Chancen?



Welche Risiken bereiten Ihnen Sorgen?



Wie können wir verantwortungsvollen KI-Einsatz fördern?

Vielen Dank für Ihre Aufmerksamkeit!



kontakt@beispiel.de



www.beispiel-ki.de