

KI als Black Box: Was das für High-Stakes-Einsätze bedeutet

MANAGEMENT SUMMARY & KERNAUSSAGE

KI-Modelle sind Black Boxes, weil niemand nachvollziehen kann, wie ein Sprachmodell sein komprimiertes Wissen in einer konkreten Antwort abrufen. Bei High-Stakes-Einsätzen wie Diagnosen oder Kreditentscheiden ist das ein reales Haftungs- und Governance-Risiko. Der EU AI Act (Art. 13) verlangt seit 2. August 2026 ausreichende Transparenz für Hochrisiko-Systeme – die Schweiz geht mit dem Ratifizierungsprozess zur Europarat-KI-Konvention einen sektorspezifischen Weg statt EU-Blanko-Übernahme.

KONTEXT & RELEVANZ FÜR ENTSCHEIDER

KI-Modelle sind Black Boxes, weil niemand – auch die Entwickler nicht – heute lückenlos nachvollziehen kann, wie ein Sprachmodell sein komprimiertes Wissen in einer konkreten Ausgabe abrufen. Für dein Unternehmen heisst das konkret: Bei High-Stakes-Einsätzen wie medizinischen Diagnosen, Kreditentscheiden oder juristischer Risikoeinschätzung setzt du ein System ein, das mehr Daten gesehen hat als jeder Mensch je verarbeiten könnte – aber seine Arbeit nicht zeigen kann.

Geoffrey Hinton, einer der Begründer des Deep Learning, hat den fundamentalen Unterschied zwischen menschlichem und künstlichem Lernen so präzise wie kaum jemand sonst beschrieben. Das menschliche Gehirn lernt über rund zwei Milliarden Lebenssekunden mit einer relativ kleinen Datenmenge. Ein LLM lernt in wenigen Trainingsläufen mit einem Vielfachen dieser Datenmenge – bei deutlich weniger Verbindungen als das Gehirn besitzt.

KERN-HANDLUNGSEMPFEHLUNGEN

- 01 Medizinische Diagnose- und Triage-Systeme
- 02 Gerichtliche Risikoprognosen und Rückfallbewertungen
- 03 Kreditvergabe, Scoring und Versicherungsentscheide
- 04 Autonome Steuerung kritischer Infrastruktur

SCHWEIZER UMSETZUNGSKRITERIEN (REVD SG & IT-ARCHITEKTUR)

- ☑ Vollständige Konformität mit dem neuen Schweizer Datenschutzgesetz (revDSG).
- ☑ Keine Abhängigkeiten: Zero Lock-In auf proprietäre geschlossene Modelle.
- ☑ Schnittstellen zu Schweizer ERPs (Abacus, Bexio, M365) im operativen Betrieb.

Unverbindliche Prozess-Analyse

Lass uns in 20 Minuten prüfen, wo KI-Agenten in deinem Unternehmen sofort messbare Entlastung schaffen.

ki-outsourcing.ch