

KI-gestützte Finanzanalyse

Auf diesem Blog wird oft die These vertreten, dass KI die Anleger bei der Vermögensentwicklung unterstützen kann. Allerdings ist es wichtig, hier stärker zu differenzieren – was wir mit diesem Beitrag erreichen wollen.

Die aktuelle Debatte über künstliche Intelligenz im Finanzbereich¹ wird häufig zugespitzt geführt. Entweder gilt KI als Effizienzrevolution – oder als gefährliche Fehlquelle. Eine Studie der HHL Leipzig, der University of Applied Sciences St. Pölten (USTP) und Nexxar bringt nun empirische Struktur in die Diskussion. Sie zeigt: Die Qualität von KI-Antworten zu Unternehmensdaten hängt stark vom Datenformat und der maschinellen Lesbarkeit der Quellen ab.

Struktur entscheidet über Sichtbarkeit

Die Untersuchung „AI Optimized Reporting“ analysiert, welche Quellen ein Large Language Model wie ChatGPT heranzieht und wie sich HTML- gegenüber PDF-Berichten auswirken. Das Ergebnis ist eindeutig: HTML-Berichte werden häufiger korrekt erkannt und verarbeitet als reine PDF-Dokumente (USTP 2025).²

Damit verschiebt sich eine klassische kapitalmarkttheoretische Frage in die digitale Sphäre. Informationsasymmetrie entsteht nicht nur durch fehlenden Zugang zu Daten, sondern durch deren technische Verwertbarkeit.

Wer seine Daten maschinenlesbar strukturiert, beeinflusst die digitale Wahrnehmung. Wer sie im „Papierformat auf dem Screen“ belässt, verliert Sichtbarkeit.

Doch hier endet die eigentliche Analyse nicht – hier beginnt sie.

Das Kernproblem: Sprachliche Kohärenz ist keine Bewertung

Large Language Models sind probabilistische Sprachmodelle. Sie rekonstruieren Muster in Texten. Sie verifizieren keine Daten. Sie bewerten keine Unsicherheit.

Das Phänomen der sogenannten „Halluzination“ – das plausible Erfinden von Fakten – ist in der wissenschaftlichen Literatur gut dokumentiert (Ji et al. 2023). Auch Tests unabhängiger Institutionen zeigen relevante Fehlerquoten bei komplexen Sachverhalten (EBU 2025).

¹ Wir beziehen uns hier auf den FAZ-Artikel vom 11.02.26: <https://www.faz.net/aktuell/finanzen/chatgpt-als-quelle-wenn-die-ki-anleger-in-die-irre-fuehrt-accg-110834870.html>

² **USTP (2025):** University of Applied Sciences St. Pölten, *AI Optimized Reporting*. Forschungsprojekt zur Untersuchung der Nutzung und Quellenbasis von Large Language Models (LLMs) bei der Verarbeitung von Unternehmensinformationen. Die öffentlich zugängliche Projektseite beschreibt Forschungsdesign, Fragestellungen und erste Ergebnisse zum Einfluss von HTML- gegenüber PDF-Berichten auf die Sichtbarkeit und Antwortqualität von KI-Systemen. Ein vollständiger, peer-reviewter Abschlussbericht ist derzeit noch nicht veröffentlicht. Verfügbar unter: <https://www.ustp.at/de/studium/digital-business-innovation/digital-business-communications/projekte/studie-ai-optimized-reporting>

Im Finanzkontext ist das besonders heikel. Denn Unternehmensbewertung ist kein Extraktionsproblem, sondern ein Unsicherheitsproblem.

- Ein EBIT ist keine Wahrheit, sondern eine Kennzahl unter Annahmen.
- Ein Cashflow ist kein Urteil, sondern ein Indikator im Kontext.
- Eine Strategie ist kein Fakt, sondern eine Erwartung.

KI kann Kennzahlen strukturieren. Aber sie gewichtet keine Szenarien.

Und genau hier liegt die Gefahr der Scheingenauigkeit: Sprachlich überzeugende Antworten erzeugen Vertrauen – auch dann, wenn keine Bewertungslogik offengelegt wird.

Architektur schlägt Euphorie

Die richtige Frage lautet daher nicht: „Ist KI zuverlässig?“

Sondern: „In welcher Architektur wird sie eingesetzt?“

Generative Modelle ohne Datenanbindung operieren mit Trainingswissen. Dieses ist zeitlich begrenzt („knowledge cutoff“) und nicht quellengebunden.

Retrieval-Augmented Generation (RAG)³ verändert diese Logik grundlegend. Vor der Textgenerierung werden definierte, qualitativ hochwertige Quellen abgefragt – etwa strukturierte XBRL-Daten⁴, regulatorische Veröffentlichungen oder geprüfte Geschäftsberichte.

Erst auf dieser Grundlage wird eine Antwort formuliert.

Damit entsteht:

- Quellenklarheit
- Dokumentierbarkeit

³ **RAG (Retrieval-Augmented Generation)** bezeichnet eine Architektur für KI-Systeme, bei der ein Sprachmodell nicht ausschließlich auf sein internes Trainingswissen zurückgreift, sondern vor der Antwortgenerierung gezielt externe, definierte Datenquellen abrufen. In einem ersten Schritt („Retrieval“) werden relevante Dokumente oder Daten aus einer angebundenen Datenbank gesucht und bereitgestellt. Erst danach erzeugt das Modell („Generation“) seine Antwort auf Basis dieser konkreten Quellen. RAG erhöht damit Transparenz, Aktualität und Nachvollziehbarkeit von KI-Antworten und reduziert das Risiko sogenannter „Halluzinationen“ (also plausibel klingender, aber faktisch falscher Aussagen). In professionellen Analysekontexten ist RAG eine zentrale Voraussetzung für belastbare KI-gestützte Auswertungen.

⁴ **XBRL (eXtensible Business Reporting Language)** ist ein internationaler, XML-basierter Standard zur strukturierten elektronischen Übermittlung von Finanz- und Unternehmensdaten. Er ermöglicht es, einzelne Bilanz- und GuV-Positionen, Cashflow-Angaben oder Anhanginformationen mit eindeutigen semantischen Tags zu versehen. Dadurch können Computerprogramme diese Daten automatisch erkennen, vergleichen und weiterverarbeiten. XBRL wird weltweit von Aufsichtsbehörden und Börsen genutzt, etwa im Rahmen regulatorischer Berichtspflichten (z. B. IFRS-Reporting oder SEC-Filings in den USA). Der Vorteil liegt in der maschinenlesbaren, standardisierten Struktur, die eine präzise Analyse ohne manuelle Texterfassung erlaubt.

- Nachvollziehbarkeit

Nicht das Modell entscheidet über Qualität, sondern die Datenarchitektur.

Quantifizierung statt Sprachsuggestion

Der entscheidende methodische Unterschied liegt jedoch noch tiefer. Viele KI-Nutzer erhalten Texte. Professionelle Analyse benötigt Modelle.

Ein Sprachmodell erzeugt kohärente Sätze.

Ein Bewertungsmodell definiert:

- Variablen
- Annahmen
- Gewichtungen
- Szenarien
- Sensitivitäten

Es macht explizit, was implizit sonst verborgen bleibt.

Erst wenn KI in eine solche quantitative Logik eingebettet wird – etwa durch strukturierte KPI-Extraktion, dokumentierte Scoring-Systeme oder Szenariosimulationen – entsteht ein belastbarer Analyseprozess.

Hier entscheidet sich, ob KI:

Informationsverdichtung betreibt

oder

Entscheidungsunterstützung leistet.

Das ist kein semantischer Unterschied. Das ist ein methodischer.

Human-in-the-Loop ist kein Rückschritt

Auch mit RAG-Architektur und quantitativer Logik bleibt die menschliche Urteilskraft eine zentrale Konstante. Studien zur Leistungsfähigkeit von LLMs in der Finanzanalyse zeigen, dass sie standardisierte Aufgaben durchaus bewältigen können, bei komplexer Kontextinterpretation jedoch an Grenzen stoßen (Zhao et al. 2024).

Kapitalmärkte sind keine deterministischen Systeme. Sie sind Erwartungsräume unter Unsicherheit.

Bewertung bedeutet:

- Wahrscheinlichkeiten einschätzen
- Extremrisiken berücksichtigen
- Narrativ und Zahl trennen
- Strukturelle Brüche erkennen

Das sind normative Entscheidungen.

Differenzierung statt Alarmismus

Die aktuelle Diskussion sollte daher weder euphorisch noch alarmistisch geführt werden.

Richtig ist:

- Unstrukturierte Daten führen zu Fehlern.
- Generative Modelle ohne Quellenbindung sind anfällig.
- Sprachliche Plausibilität kann analytische Schwächen verdecken.

Ebenso richtig ist: Mit

- strukturierter Datenanbindung (RAG),
- geprüften Datenquellen,
- klar definierter Bewertungslogik,
- expliziter menschlicher Kontrolle

wird KI die Qualität und Geschwindigkeit fundierter Analyse erheblich erhöhen können.

Fazit

KI ersetzt keine Urteilstkraft. Aber sie kann helfen, Urteile systematischer, transparenter und reproduzierbarer zu machen. Wer KI nur fragt, erhält Texte. Wer sie methodisch einbettet, erhält strukturierte Entscheidungsgrundlagen.

Im Kapitalmarkt entscheidet nicht die Sprachfähigkeit eines Modells – sondern die Architektur, in die es eingebunden ist.

Quellen

1. USTP – „AI Optimized Reporting“

Projektseite der *University of Applied Sciences St. Pölten* zur Studie (Studienbeschreibung):

 <https://www.ustp.at/de/studium/digital-business-innovation/digital-business-communications/projekte/studie-ai-optimized-reporting>

(Stand: Feb. 2026 – Projektseite mit Beschreibung, Zielsetzung und ersten Ergebnissen)

2. Ji et al. (2023) – Hallucination in Natural Language Generation

Übersichtsarbeit zu Halluzinationen in KI-Systemen – peer-reviewed, umfassend, zitierfähig:

 **Survey of Hallucination in Natural Language Generation (arXiv Preprint)**
<https://arxiv.org/abs/2202.03629>

Z. Ji, J. Lee, Z. Frieske et al., “Survey of Hallucination in Natural Language Generation,” *ACM Computing Surveys*, 2023.

3. EBU (European Broadcasting Union) – KI-Fehleranfälligkeit

Bericht über Fehlerquoten bei KI-Antworten zu Nachrichteninhalten (journalistische Auswertung, nicht peer-reviewed):

 <https://orf.at/stories/3409217/>

„Europäische Rundfunkunion: KI-Systeme geben Nachrichteninhalte oft falsch wieder“, *ORF.at*

Und ergänzend:

 <https://www.heise.de/news/Europaeische-Rundfunkunion-KI-Systeme-geben-Nachrichteninhalte-oft-falsch-wider-10796779.html>

Heise Online-Bericht zur Studie der EBU

4. Zhao et al. (2024) – KI in der Finanzanalyse (Preprint)

Diskussionspapier zur Leistungsfähigkeit von KI für finanzielle Auswertung (Preprint, nicht peer-reviewed):

 https://www.researchgate.net/publication/380862587_AI-Driven_Financial_Analysis_Exploring_ChatGPT%27s_Capabilities_and_Challenges

„AI-Driven Financial Analysis: Exploring ChatGPT’s Capabilities and Challenges“
