

Übersicht: Merkmale KI-generierter Inhalte

Bitte beachten Sie, keines der folgenden Merkmale ist für sich genommen ein verlässlicher Nachweis für KI-generierten Text. Sie können aber bei gehäuftem Auftreten einen Hinweis auf eine KI-gestützte Erstellung geben. Moderne Sprachmodelle unterscheiden sich zudem deutlich in ihrem Schreibstil von älteren Modellen, und durch menschliche Nachbearbeitung verschwinden viele dieser Merkmale.

1. Auffällige Gedankenstriche

- **Em-Dash:** —
- **En-Dash:** –
- Konstruktionen wie:
„Das Ergebnis ist bemerkenswert — allerdings nicht überraschend.“
- Besonders auffällig kann eine sehr häufige Verwendung des langen Gedankenstrichs sein, vor allem in deutschsprachigen Texten, in denen Autorinnen und Autoren normalerweise eher Kommas, Klammern oder den normalen Bindestrich verwenden würden.

Hinweiswert: mittel bis relativ hoch, wenn der individuelle Schreibstil der Person sonst keine Gedankenstriche dieser Art enthält.

2. Typografisch „perfekte“ Anführungszeichen

Häufig erscheinen automatisch korrekte typografische Zeichen:

- “...”
- ‘...’
- „...“
- ,...’

Auffällig kann insbesondere ein Wechsel zwischen englischen und deutschen typografischen Anführungszeichen sein. (Deutsch: „Text“; bei einem Zitat im Zitat: ,Text‘; Englisch: “Text” bei einem Zitat im Zitat: ‘Text’).

3. Überdurchschnittlich viele Doppelpunkte

KI-Texte verwenden häufig Konstruktionen wie:

- „Dabei sind drei Aspekte entscheidend:“
- „Das Ergebnis ist eindeutig:“
- „Besonders relevant ist folgender Punkt:“
- „Die zentrale Erkenntnis lautet:“

Oft folgt auf den Doppelpunkt eine Aufzählung.

4. Häufige Semikolons

;

Sprachmodelle verwenden teilweise Semikolons regelmäßiger als viele alltägliche menschliche Schreibende:

„Die Methode ist effizient; ihre klinische Relevanz bleibt jedoch unklar.“

Das ist insbesondere auffällig, wenn Semikolons im übrigen Schreibstil einer Person praktisch nicht vorkommen.

5. Sehr regelmäßige Klammerkonstruktionen

Beispiele:

- „... die Ergebnisse (insbesondere in der Subgruppe) zeigen ...“
- „... die Methode (siehe oben) ermöglicht ...“
- „... KI-Systeme (z. B. ChatGPT, Claude oder Gemini) ...“

KI tendiert dazu, Nebeninformationen sauber in Klammern einzubauen.

6. Typische Markdown-Zeichen

Ein besonders deutlicher Hinweis auf direkt aus einem Chatbot übernommenen Text können **stehen gebliebene Markdown-Zeichen** sein:

- ****Text**** – Fettdruck
- **Text** – Kursivschrift
- **#** Überschrift
- **##** Unterüberschrift
- **###** Unterüberschrift
- `Begriff`
- --- – Trennlinie

- > - Zitat
- [Text](URL) - Markdown-Link
- | Spalte 1 | Spalte 2 | - Markdown-Tabelle

Solche Artefakte sind wesentlich informativer als normale Satzzeichen.

7. Sehr systematische Bullet-Point-Strukturen

Typisch sind Aufzählungen wie:

- Erstens: ...
- Zweitens: ...
- Drittens: ...

oder

1. Ausgangslage
2. Problem
3. Lösung
4. Schlussfolgerung

Besonders auffällig ist die Kombination von **fett gesetztem Stichwort + Doppelpunkt + Erklärung**:

- **Relevanz:** ...
- **Methodik:** ...
- **Limitationen:** ...
- **Ausblick:** ...

8. Dreierstrukturen

LLMs strukturieren Argumente auffallend häufig in **Dreiergruppen**:

„Die Methode ist schnell, zuverlässig und kostengünstig.“

oder:

„Drei Aspekte sind entscheidend: Transparenz, Nachvollziehbarkeit und Verantwortung.“

Einzelne Dreierstrukturen sind völlig normal; ihre permanente Wiederholung kann jedoch charakteristisch wirken.

9. Sehr regelmäßige Zwischenüberschriften

Beispielsweise:

Hintergrund
Zentrale Herausforderungen
Mögliche Lösungsansätze
Praktische Implikationen
Fazit

KI produziert häufig eine beinahe „lehrbuchartige“ Gliederung, selbst wenn der Text eigentlich keine solche Struktur benötigt.

10. Pfeile und andere Visualisierungssymbole

Gelegentlich werden Zusammenhänge mit Zeichen dargestellt wie:

- →
- ?
- ?
- ✓
- ?
- ?
- ??
- ?

Beispielsweise:

Datenerhebung → Analyse → Interpretation → Publikation

Besonders Emojis wie ?, ?? oder ? sind für den typischen Chatbot-Stil charakteristisch, sofern sie nicht bewusst zum Kommunikationsformat gehören.

11. Ungewöhnliche Unicode-Zeichen

Beim direkten Kopieren aus KI-Systemen können Zeichen auftreten, die auf der Tastatur normalerweise nicht bewusst eingegeben werden:

- geschütztes Leerzeichen
- schmales geschütztes Leerzeichen
- typografischer Apostroph ’
- geschützter Bindestrich -
- En-Dash –

- Em-Dash —
- Ellipse ... statt drei Punkten ...

Solche Zeichen können bei einer **technischen Textanalyse** interessant sein. Sie entstehen allerdings auch durch Word, Webseiten und automatische Textformatierung.

12. Sehr konsequente Parallelkonstruktionen

Beispielsweise:

„Die erste Studie untersucht ... Die zweite Studie analysiert ... Die dritte Studie bewertet ...“

oder:

„Auf individueller Ebene ... Auf institutioneller Ebene ... Auf gesellschaftlicher Ebene ...“

Sprachmodelle produzieren besonders gerne sprachlich symmetrische Konstruktionen.

13. Häufige Kontrastkonstruktionen

Typische Formulierungen sind:

- „nicht nur ..., sondern auch ...“
- „einerseits ..., andererseits ...“
- „während ..., zeigt sich zugleich ...“
- „Obwohl ..., ist dennoch ...“
- „Dabei geht es nicht darum, ..., sondern vielmehr darum, ...“

Eine hohe Dichte solcher rhetorisch ausgewogenen Konstruktionen kann KI-typisch wirken.

14. Charakteristische Einleitungssphrasen

Häufig finden sich Formulierungen wie:

- „Es ist wichtig zu beachten, dass ...“
- „Dabei ist zu berücksichtigen, dass ...“
- „In diesem Zusammenhang ist hervorzuheben, dass ...“
- „Ein zentraler Aspekt ist ...“
- „Von besonderer Bedeutung ist ...“
- „Vor diesem Hintergrund ...“
- „Im Folgenden werden ...“

- „Grundsätzlich lässt sich festhalten ...“

Sie sind einzeln völlig normal. KI-Texte enthalten sie jedoch teilweise in ungewöhnlich hoher Frequenz.

15. Charakteristische Schlussformulierungen

Besonders häufig:

- „Zusammenfassend lässt sich sagen, dass ...“
- „Insgesamt zeigt sich, dass ...“
- „Abschließend ist festzuhalten, dass ...“
- „Letztlich wird deutlich, dass ...“
- „Dies unterstreicht die Bedeutung von ...“
- „Damit wird deutlich, dass ...“

Typisch ist auch ein letzter Satz, der die vorherigen Aussagen noch einmal **allgemein und positiv zusammenfasst**, ohne neue Information zu enthalten.

16. Inhaltlich wenig informative Metasätze

Beispiele:

- „Dieses Thema ist von großer Bedeutung.“
- „Diese Frage ist von besonderer Relevanz.“
- „Die Ergebnisse liefern wichtige Erkenntnisse.“
- „Dies eröffnet neue Perspektiven.“
- „Weitere Forschung ist erforderlich.“
- „Die Thematik ist komplex und vielschichtig.“

Solche Sätze sind nicht falsch, tragen aber häufig wenig konkrete Information. **Eine hohe Dichte solcher „glatten“, aber informationsarmen Aussagen ist wesentlich aussagekräftiger als einzelne Satzzeichen.**

17. Übermäßig ausgewogene Argumentation

KI formuliert häufig nach dem Schema:

Vorteil → Einschränkung → ausgewogene Schlussfolgerung

Beispiel:

„KI bietet erhebliche Potenziale zur Effizienzsteigerung. Gleichzeitig bestehen Herausforderungen hinsichtlich Datenschutz und Transparenz. Entscheidend ist daher ein verantwortungsvoller und reflektierter Einsatz.“

Diese fast perfekte argumentative Balance ist ein typisches LLM-Muster.

18. Ungewöhnlich gleichmäßige Satz- und Absatzstruktur

Mögliche Auffälligkeiten sind:

- fast alle Absätze ähnlich lang;
- jeder Absatz beginnt mit einer klaren Kernaussage;
- danach folgen zwei bis drei Erläuterungssätze;
- der letzte Satz fasst den Absatz zusammen oder bildet eine Überleitung;
- kaum Satzfragmente oder stilistische Brüche;
- wenige wirklich ungewöhnliche Formulierungen.

Der Text wirkt dadurch häufig **ungewöhnlich sauber, kontrolliert und gleichförmig**.

19. „Pseudo-Präzision“

Sprachmodelle produzieren gerne Formulierungen wie:

- „insbesondere“
- „maßgeblich“
- „wesentlich“
- „zentral“
- „relevant“
- „differenziert“
- „systematisch“
- „nachvollziehbar“
- „evidenzbasiert“
- „ganzheitlich“
- „robust“

Problematisch wird dies, wenn diese Begriffe häufig auftreten, ohne dass konkretisiert wird, **was** beispielsweise „robust“ oder „systematisch“ bedeutet.

20. Verdächtige Literatur- und Zitationsmuster

Bei wissenschaftlichen Texten sollte besonders auf Folgendes geachtet werden:

- auffällig gleichmäßig über Absätze verteilte Zitate
- zahlreiche Quellen hinter einer einzelnen allgemeinen Aussage; sogenannte **Reference Clusters**, z. B. [12–19] ohne das einzelne Referenzen erklärt werden
- Quellen, die die konkrete Aussage nur teilweise unterstützen
- formal plausible, aber nicht existente Literatur
- falsche DOI
- falsche Seitenzahlen
- richtige Publikation, aber falscher Inhalt
- sehr aktuelle oder sehr passende Literatur ohne erkennbare Suchstrategie.

Dies ist bei wissenschaftlichen Arbeiten häufig **aussagekräftiger als die sprachliche Oberfläche**.

Praktische Einordnung

Für die Beurteilung eines Textes würde ich die Merkmale ungefähr so gewichten:

Schwache Hinweise:

Semikolon, Doppelpunkt, typografische Anführungszeichen, Gedankenstrich, Ellipse.

Mittlere Hinweise:

ungewöhnlich viele Em-Dashes, Dreierstrukturen, stereotype Übergangsformulierungen, extrem gleichmäßige Absätze, übermäßige Zwischenüberschriften, informationsarme Metasätze.

Stärkere Hinweise:

stehen gebliebenes Markdown, typische Chatbot-Formatierungen, plötzlich gegenüber früheren Arbeiten stark veränderter Schreibstil, ungewöhnliche Unicode-Artefakte in Kombination mit weiteren Merkmalen, erfundene oder nicht passende Referenzen sowie eine Kombination vieler der oben genannten sprachlichen Muster.

Entscheidend ist daher nicht ein einzelnes „KI-Satzzeichen“, sondern das Muster. Eine seriöse Beurteilung sollte möglichst den aktuellen Text mit früheren authentischen Textproben derselben Person vergleichen und zusätzlich Inhalt, Quellen, Argumentationsführung und Entstehungsprozess berücksichtigen.