

Manual für die inhaltliche Kategorisierung der Artikel zum Thema KI

Aufgabe

Wir möchten ein Korpus erstellen, das Texte beinhaltet, die sich mit Künstlicher Intelligenz auseinandersetzen.

Ordne die Texte (txt-Dateien) in der Datei "corpus_full" in einen der 4 Ordner ein, die den unten stehenden Kategorien entsprechen:

- a. 1_hauptthema
- b. 2_nebenthema
- c. 3_kein_thema
- d. 4_unklar

Die Texte müssen für die Kategorisierung nicht komplett gelesen werden. Es reicht aus, wenn ihr die Texte selektiv überfliegt und die Themen der einzelnen Absätze begreift. Die Texte wurden mit Suchbegriffen zum Thema KI gefunden. Daher ist es sinnvoll, nach diesen Ausdrücken auch im Text zu suchen (Shortcut: **Strg + F**), um die besonders relevanten Sätze und Absätze zu finden. Die Suchausdrücke, mit denen ein Text gefunden wurde, sind oben in der Datei vermerkt.

Die Ausdrücke in der Datei sind (oft) Lemmata, also die Grundform des Wortes. Um die richtige Stelle zu finden, muss man evtl. verschiedene Wortformen ausprobieren (z.B. *künstliche Intelligenz*, *künstlichen Intelligenz*, ...). Siehe Screenshot.

Es kann sein, dass es noch andere KI-Begriffe im Text gibt, die nicht in den Search Terms stehen. Daher lohnt es sich oft, den Text genauer durchzulesen – vor allem, um zwischen Hauptthema und Nebenthema zu unterscheiden.

AUßERDEM kann es bei den Suchausdrücken zu falsch-positiven Ergebnissen kommen. Denn nicht alle Suchausdrücke sind *nur* KI-bezogen; manche sind **ambig** (wie auch bei Moralwörtern 😊). Zum Beispiel ist *Claude* eine KI, aber auch ein französischer Name.

Ist ein Text nur ambig, ist es natürlich 3_kein_thema. Aber stellt bitte sicher, dass es wirklich nicht irgendwo doch um KI geht!

```
URL: https://www.informatik-aktuell.de/management-und-recht/digitalisierung/10-fallen-bei-der-unternehmenstransformation.html
Title: 10 Fallen bei der Unternehmenstransformation
Description: Wollen Unternehmen trotz Volatilität und Komplexität wettbewerbsfähig bleiben, müssen sie einen Wandel durchführen. Auf diese Stolperfallen sollten Sie auf dem steinigen Weg der Transformation achten.
Search terms: künstlich Intelligenz
```



```
+++++
```

10 Fallen bei der Unternehmenstransformation

Das Gelingen der digitalen Unternehmenstransformation ist für Unternehmen zum Überlebensfaktor geworden. Dank der Digitalisierung

Was gehört alles zu KI? (auch äquivalente englische Begriffe; z. B. AI, Robotics usw.)

KI definieren wir nach der Definition des EU-Parlaments:

Künstliche Intelligenz ist die Fähigkeit einer Maschine, menschliche Fähigkeiten wie logisches Denken, Lernen, Planen und Kreativität zu imitieren. KI ermöglicht es technischen Systemen, ihre Umwelt wahrzunehmen, mit dem Wahrgenommenen umzugehen und Probleme zu lösen, um ein bestimmtes Ziel zu erreichen. Der Computer empfängt Daten (die bereits über eigene Sensoren, zum Beispiel eine Kamera, vorbereitet oder gesammelt wurden), verarbeitet sie und reagiert.

KI-Systeme sind in der Lage, ihr Handeln anzupassen, indem sie die Folgen früherer Aktionen analysieren und autonom arbeiten.

In eigenen Worten: KI imitiert menschliche Intelligenz. Die Probleme, die mit KI gelöst werden, sind oft vielschichtiger und es gibt meist keine 'perfekte' Lösung (weil es keine Regel gibt, nach der man etwas lösen kann). Stattdessen 'lernt' eine KI, etwas iterativ immer besser zu machen.

Zum Beispiel: Sprachverarbeitung, Bilder generieren und erkennen, Schach spielen, Wetter vorhersagen.

Beispiele für KI:

Konzepte: Chatbots, (Large) Language Models, maschinelles Lernen, Deep Learning, Neuronales Netz, Bildgenerierung, Recommender-Algorithmen (z.B. von TikTok, Spotify), Computergegner in Videospielen, Schachbots, Retrieval Augmented Generation (RAG), Transformer, autonome Waffensysteme, fiktionale KIs (R2D2), intelligente Systeme, Gesichtserkennung, Autonome Roboter

Konkrete Namen: OpenAI, ChatGPT, Copilot, Claude, Bard, BERT, Gemini, Deepseek, Dall-E, Midjourney, Stockfish, Deep Blue, Boston Dynamics

Das keine / nicht automatisch KI:

Algorithmus: Die meisten Algorithmen sind 'dumm' und führen einfache Regeln aus. Sie 'lernen' nicht; passen ihr Handeln nicht an.

Cybersicherheit: Eines von vielen informatischen Themen, die zwar mit KI zu tun haben *können*, aber für sich genommen ist Cybersicherheit eine ganze Menge von Dingen, von denen die meisten nicht 'intelligent' sind, zum Beispiel Firewalls oder Passwörter.

(Statistische, theoretische, ...) Modelle: Manche, aber nicht alle, Modelle sind KI. Ein *Sprachmodell* ist KI, weil es die menschliche Fähigkeit, Sprache zu verstehen und zu produzieren, imitiert. Ein Statistikmodell ist grenzwertig, aber in der Regel ein simpler Algorithmus. Ein neues Handy-Modell oder ein Unterrichtsmodell ist natürlich keine KI. Hier müsst ihr von Fall zu Fall unterscheiden und am besten nachschlagen oder uns fragen!

Usw...

Wenn Unsicherheit besteht, ob ein Konzept zu KI gehört, dann googlen. Falls immer noch unklar Text in Kat 4_unsicher schieben und Beispiel mit uns besprechen.

Kategorien

1. KI ist Hauptthema des Artikels (1_hauptthema):

a. **ALTERNATIV**: Der Fokus des Artikels liegt auf KI. Das erkennt man an verschiedenen Kriterien, nämlich:

- i. KI wird in der Überschrift und/oder in der Description thematisiert
- ii. KI wird in der Mehrzahl der Absätze thematisiert
- iii. KI wird an besonders wichtigen Stellen, z.B. im Fazit oder in der Einleitung, thematisiert.

ACHTUNG: Gerade, wenn es nur im Teaser thematisiert wird, kann es immer noch nur Nebenthema sein. Es müssen mehrere Kriterien gegeben sein.

Entscheidungshilfe: Wenn man die Phrasen über KI aus dem Text nimmt, ergibt der Text keinen Sinn mehr.

Beispiel:

Titel: Voice-User-Interface-Design – wie ein Chatbot zu sprechen lernt

Description: Wie kann es VUI-Designern gelingen, ein Spracherlebnis zu gestalten, das mehr bietet als die monotone Computerstimme einer Telefonhotline. Die Erwartungen der Nutzenden sind hoch!

Link: <https://www.informatik-aktuell.de/betrieb/kuenstliche-intelligenz/voice-user-interface-design-ein-chatbot-lernt-sprechen.html#top>

2. KI wird erwähnt (2_nebenthema):

- a. KI wird in einem oder mehreren Absätzen thematisiert, Das Hauptthema des Artikels ist allerdings nicht KI.
- b. Es wird eine inhaltliche Aussage über KI getätigt. Eine Nennung von KI reicht nicht aus (vgl. Kat. 3_kein_thema)

Entscheidungshilfe: Wenn man die Phrasen über KI aus dem Text nimmt, ergibt der Text noch Sinn, aber es fehlt ein interessanter Aspekt des Hauptthemas.

Beispiel:

Abschnitt über KI:

Artikelthema/Überschrift: *Der Schlüssel zu einer guten Entwicklererfahrung*

[...] 75 Prozent der Unternehmen stellen Schulungen und Ressourcen für KI bereit, erklärten die für die Studie befragten Entwickler. Ein ähnlich großer Anteil gab indes an, dass sie selbst nach Ressourcen suchen müssten. Das deutet darauf hin, dass die vom Unternehmen bereitgestellten Ressourcen möglicherweise unzureichend sind.

Unternehmen sollten ihren Entwicklungsteams ständig aktualisierte Ressourcen zur Verfügung stellen, da sich KI schnell weiterentwickelt und wächst.” [...]

Link: <https://www.informatik-aktuell.de/entwicklung/methoden/der-schluessel-zu-einer-guten-entwicklererfahrung-keep-it-simple.html>

Satz über KI:

Artikelthema/Überschrift: *Vesuvius Challenge belohnt Entschlüsselung antiker Papyrusrolle*

[...] Allerdings wurden die Herculaneum-Texte mit kohlenstoffbasierter Tinte verfasst, die

unter den mit Teilchenbeschleunigern aufgenommenen CT-Scans kaum fluoresziert. Dennoch ging Seales davon aus, dass sich die Schrift durch Unterschiede in der Textur sichtbar machen lassen sollte – mit Hilfe von KI-Modellen, die darauf trainiert wurden.
[...]

3. KI wird nicht thematisiert (3_kein_thema):

- a. Lediglich kurze Nennung von KI: Es wird keine inhaltliche Aussage über KI getätigt. KI wird stattdessen in einer Aufzählung erwähnt, als Beispiel aufgeführt etc.
- b. KI kommt überhaupt nicht vor (weder Thematisierung noch Nennung)

Entscheidungshilfe: Wenn man die Phrasen über KI aus dem Text nimmt, ergibt der Text noch Sinn, und es ändert sich inhaltlich nichts.

Beispiel “Nennung” - KI wird als Beispiel aufgeführt:

“Eine Halbierung der CO2-Emissionen innerhalb weniger Monate. Von solchen Effizienzgewinnen träumen andere Industrien. In den kommenden Jahren und Jahrzehnten könnten sich die Effizienzgewinne verlangsamen. Rebound-Effekte sowie neue Anforderungen durch Technologien wie KI und Blockchain lassen Bedenken aufkommen, dass der Sektor sich automatisch auf einer klimaneutralen oder "grünen" Reise befindet.”

Link: <https://www.informatik-aktuell.de/management-und-recht/digitalisierung/nachhaltigkeit-im-it-sektor-eine-strategische-perspektive.html>

Beispiel “Keine inhaltliche Aussage”:

Auch Kanada verurteilte die angekündigten US-Zölle. Premierminister Justin Trudeau sagte am Rande eines Treffens zu künstlicher Intelligenz in Paris, seine Regierung werde kanadische Arbeiter schützen.

Link: <https://www.zeit.de/politik/ausland/2025-02/europaeische-union-donald-trump-zoelle-stahl-aluminium-gegenmassnahmen>

4. Unsicher (4_unklar):

- a. Das Dokument kann nicht zweifelsfrei einer der drei anderen Kategorien zugewiesen werden.