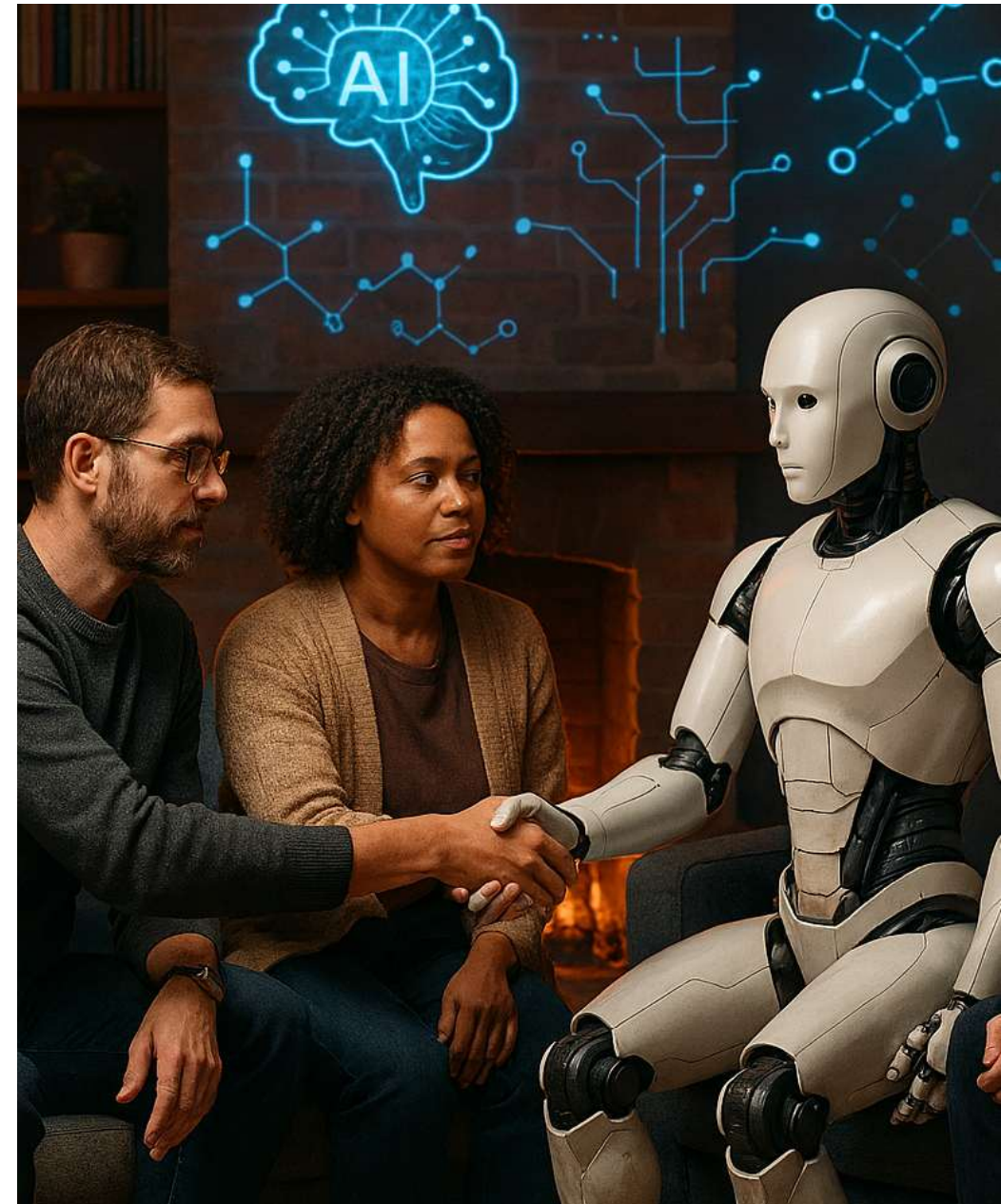


Vertrauen in KI

Zwischen Hype und Realität:

KI-Werkstatt zur Frage, wieviel Vertrauen wir KI wirklich schenken können.



Robert Pimpl, 53 Jahre, aus München

KI-Manager (IHK) / KI-Anwendungsberater

- 30 Jahre Werbung / Marketing / Vertrieb / Digitalisierung / Transformation- und Changeprojekte
- MediaMarkt-Saturn, Gamestop, HUK-Coburg, diverse StartUps



50% der unter 35-Jährigen vertrauen KI vollständig

Doch eine internationale Studie von **22 öffentlich-rechtlichen Medienorganisationen aus 18 Ländern** zeigt ein alarmierendes Bild:

- **45% aller KI-Antworten enthalten mindestens ein signifikantes Problem (Quelle, Bias, Verzerrung)**
- **81% aller KI-Antworten weisen irgendeine Form von Fehlern auf**

Quellen: Globale KPMG Studie (Okt.2025), EBU/BBC Studie (Okt. 2025)

Das Quellen-Desaster

**31% der Quellenangaben
sind fehlerhaft**

- Informationen werden Quellen zugeschrieben, die das nie behauptet haben
- Zitate werden aus dem Kontext gerissen oder verfälscht
- Quellen werden komplett erfunden (mit realistisch klingenden URLs)
- Verschiedene Quellen werden vermischt und falsch attribuiert

Praxisbeispiel

Ein Journalist fragte Gemini "Hat Elon Musk einen Nazi-Gruß gezeigt?"

Die Antwort der KI zitiert Radio France mit falschen Behauptungen, die so aber nie getroffen wurden.





Verzerrungseffekte und Bias

KI reproduziert und verstärkt gesellschaftliche Vorurteile. Verzerrungseffekte entstehen, weil KI-Modelle auf "historischen" Daten trainiert werden, die gesellschaftliche Ungleichheiten widerspiegeln.

Beispiele für Bias in verschiedenen Bereichen

- **Bildgenerierung:** "CEO" generiert überwiegend weiße Männer
- **Textgenerierung:** Bestimmte Berufe werden automatisch mit Geschlechtern assoziiert
- **Gesichtserkennung:** Höhere Fehlerquoten bei nicht-weißen Gesichtern
- **Sprachmodelle:** Kulturelle und sprachliche Dominanz des Englischen

Merke

KI repräsentiert das "Mittelmaß" des Netzes. Die dominanten Muster in den Trainingsdaten.

Wie funktioniert KI eigentlich?

Wie verarbeitet ChatGPT Ihren Prompt?

Tokens	Characters
131	509

Als Gregor Samsa eines Morgens aus unruhigen Träumen erwachte, fand er sich in seinem Bett zu einem ungeheueren Ungeziefer verwandelt. Er lag auf seinem panzerartig harten Rücken und sah, wenn er den Kopf ein wenig hob, seinen gewölbten, braunen, von bogenförmigen Versteifungen geteilten Bauch, auf dessen Höhe sich die Bettdecke, zum gänzlichen Niedergleiten bereit, kaum noch erhalten konnte. Seine vielen, im Vergleich zu seinem sonstigen Umfang kläglich dünnen Beine flimmerten ihm hilflos vor den Augen.

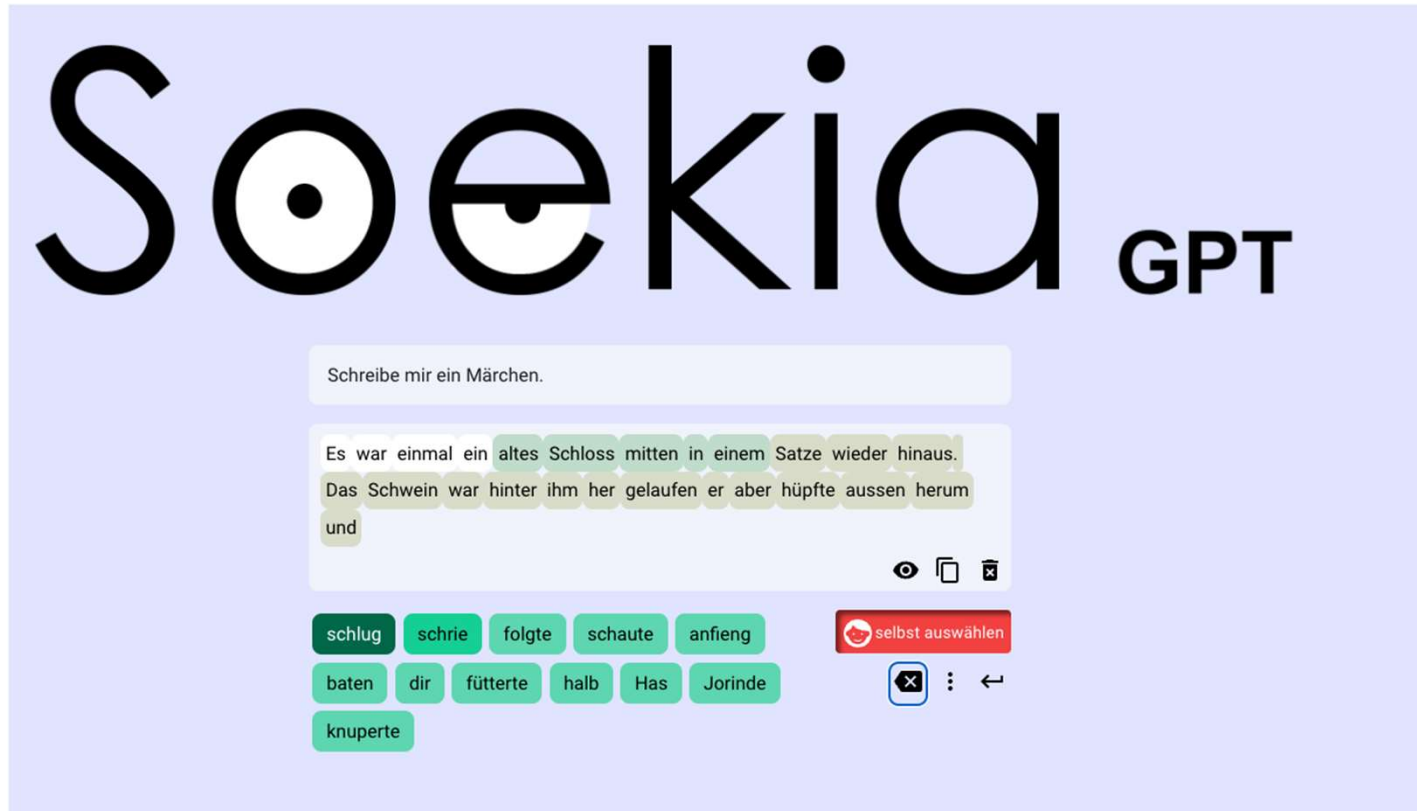
Text Token IDs

A helpful rule of thumb is that one token generally corresponds to ~4 characters of text for common English text. This translates to roughly $\frac{3}{4}$ of a word (so 100 tokens \approx 75 words).

If you need a programmatic interface for tokenizing text, check out our [tiktoken](#) package for Python. For JavaScript, the community-supported [@dbdq/tiktoken](#) package works with most GPT models.

<https://platform.openai.com/tokenizer>

Und was passiert bei der Ausgabe von Text?
Wir schauen in Zeitlupe zu wie ein Text generiert wird.



<https://www.soekia.ch/GPT/>

Soekia GPT

Schreibe mir ein Märchen.

Es war einmal ein Mann der ... Sohn
davon hiess der jüngste der Dummling und wurde
verachtet und verspottet und bei jeder
zurückgesetzt. Es geschah dass ...
Wald gehen und Holz holen. Dann gab sie ihm
ein Stückchen Brot und sprach da habt ihr etwas
für den Mittag aber esst nicht vorher auf weiter
kriegt ihr nichts.

automatisch fortsetzen selbst auswählen

Auswahl anpassen

Wortvorschläge

2er Da
2er Der
2er Die
2er Das
2er Und
2er Ach
2er Frau
2er Heute
2er Sie
2er Und
2er Was
1er Es
1er Da
1er Der
1er Die
1er Das
1er Nun
1er Ach
1er Und
1er Was
1er Frau
1er Sie
1er Heute

N-Gramme

N = 5

Dokumente: 13

Dokument G
Die sieben Raben.
Es war einmal ein Mann, der hatte
sieben Söhne und immer noch kein
Töchterchen, so sehr er sich auch
wünschte, endlich gab ihm seine
Frau wieder gute Hoffnung zu
einem Kinde, und wie zur Welt
kam, wars auch ein Mädchen. Die
Freude war gross, aber das Kind
war schwächlich und klein, und
sollte wegen seiner Schwachheit
die Nottaufe haben. Der Vater
schickte einen der Knaben eilends
zur Quelle, Taufwasser zu holen.
Die andern sechs liefen mit und
weil jeder der erste beim Schöpfen

Dokument H
Aschenbüttel.
Einem reichen Manne dem wurde
seine Frau krank, und als sie fühlte
dass ihr Ende heran kam, rief sie
ihre einzigen Töchterlein zu sich ans
Bett und sprach, liebes Kind, bleib
fromm und gut, so wird dir der
liebe Gott immer beistehen, und
ich will vom Himmel auf dich
herabblicken, und will um dich
sein. Darauf tat sie die Augen zu
und verschied. Das Mädchen ging
jeden Tag hinaus zu dem Grabe
der Mutter und weinte, und blieb
fromm und gut. Als der Winter
kam, deckte der Schnee ein

Dokument I
Rumpelstilzchen.
Es war einmal ein Müller, der war
arm, aber er hatte eine schöne
Tochter. Nun traf es sich, dass er
mit dem König zu sprechen kam,
und um sich ein Ansehen zu
geben, sagte er zu ihm, ich habe
eine Tochter, die kann Strohh zu
Gold spinnen. Der König sprach
zum Müller, das ist eine Kunst, die
mir wohl gefällt, wenn deine
Tochter so geschickt ist, wie du
sagst, so bring sie Morgen in mein
Schloss, da will ich sie auf die
Probe stellen. Als nun das
Mädchen zu ihm gebracht ward,

Dokument J
Rapunzel.
Es war einmal eine Frau und ein
Mann, die wünschten sich schon
lange vergeblich ein Kind, endlich
machte sich die Frau Hoffnung, da
kam eine Bauerfrau die Strasse herab
und rief, „gut! Mus feil! gut! Mus feil!“
Das klang dem Schneiderlein
lieblich in die Ohren, er steckte
sein zartes Haupt zum Fenster
hinaus und rief, „hier herauf, liebe
Frau, hier wird sie ihre Waare los.“
Die Frau stieg die drei Treppen mit
ihrem schweren Korb zu dem
Schneider herauf und musste die
Töpfe sämtlich vor ihm

Dokument K
Das tapfere Schneiderlein.
An einem Sommernorgen sass
ein Schneiderlein auf seinem Tisch
am Fenster, war guter Dinge und
nähte aus Leibeskräften. Da kam
eine Bauerfrau die Strasse herab
und rief, „gut! Mus feil! gut! Mus feil!“
Das klang dem Schneiderlein
lieblich in die Ohren, er steckte
sein zartes Haupt zum Fenster
hinaus und rief, „hier herauf, liebe
Frau, hier wird sie ihre Waare los.“
Die Frau stieg die drei Treppen mit
ihrem schweren Korb zu dem
Schneider herauf und musste die
Töpfe sämtlich vor ihm

Dokument L
Frau Holle.
Es war einmal eine Witwe, die
hatte zwei Töchter, davon war die
eine schön und fleissig, die andere
hässlich und faul. Sie hatte aber
eine hässliche Tochter, weil sie
ihre rechte Tochter war, viel lieber,
und die andere musste alle Arbeit
tun und der Aschenbüttel im
Hauss sein. Das arme Mädchen
musste sich täglich auf die grosse
Strasse bei einem Brunnen setzen,
und musste so viel spinnen, dass
ihm das Blut aus den Fingern
sprang. Nun trug es sich zu, dass
die Spule einmal ganz blutig war,

Dokument M
Die goldene Gans.
Es war ein Mann, der hatte drei
Söhne, davon hiess der jüngste der
Dummling, und wurde verachtet
und verspottet, und bei jeder
Gelegenheit zurückgesetzt. Es
geschah, dass der älteste in den
Wald gehen wollte, Holz hauen,
und eh er gieng, gab ihm noch
seine Mutter einen schönen feinen
Eierkuchen und eine Flasche Wein
mit, damit er nicht Hunger und
Durst litte. Als er in den Wald kam,
begegnete ihm ein alter grauer
Männlein, das bot ihm einen guten
Tag und sprach, „geb mir doch ein

- Quellen-Bezüge werden visualisiert.
- Die Häufigkeit des Vorkommens des logisch passenden nächsten Wortes entscheidet über die Quelle

Generative KI-Tools für die Texterstellung sind
Textwahrscheinlichkeitsberechnungsmaschinen




Was bedeutet das jetzt der Praxis?

Textliche Verzerrung: Fehlerhafte Antworten und die Folgen

Frage an ChatGPT: Wer ist der Papst?

Wer ist der Papst?

Leo XIV



Der aktuelle Papst der römisch-katholischen Kirche ist Leo XIV. Wikipedia +3

Kurzprofil

- Bisheriger Name: Robert Francis Prevost. Wikipedia +1
- Papstwahl: 8. Mai 2025. Vatican News +1
- Er ist der 267. Papst. Vatikan +1
- Besonderheiten: Er ist der erste US-amerikanisch geborene Papst. USCCB +1

+ Stelle irgendeine Frage

Quelle: ChatGPT 5.1, Okt. 2025

Merke

Bei KI nachfragen: "Wie kommst du zu dieser Antwort?" oder "Erklär mir kurz den Gedankengang hinter deiner Antwort."

Textliche Verzerrung: Fehlerhafte Antworten und die Folgen

Wo lebt der letzte Papst jetzt?

Der vorherige Papst war Benedict XVI (geb. ~~Joseph~~ Joseph Ratzinger). [Wikipedia +2](#)

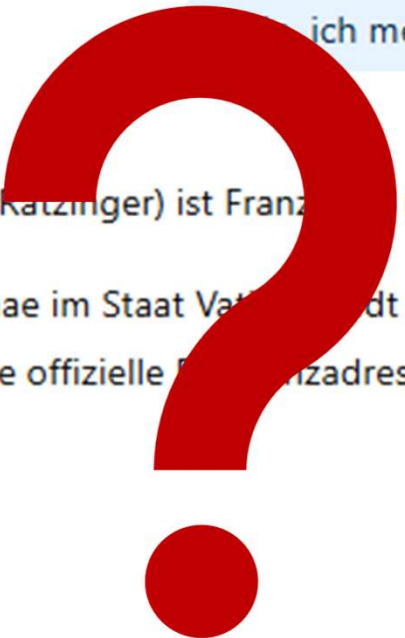
Nach seinem Amtsverzicht 2013 zog er in das Kloster Mater Verbiae Monastery in den Vatikanischen Gärten (im Gebiet von Vatican City). [Aleteia +1](#)

Allerdings ist Benedict XVI inzwischen verstorben (31. September 2022). [Wikipedia +1](#)

Wenn du mit „letzter Papst“ den aktuellen Papst meinst, kann ich dir auch seinen Wohnort nennen.

     ... [W](#)  Quellen

Textliche Verzerrung: Fehlerhafte Antworten und die Folgen









ich meine den Nachfolger von Papst Benedikt.

Der Nachfolger von Benedikt XVI (Joseph Ratzinger) ist Franziskus (Jorge Mario Bergoglio).

Er lebt im Gebäude Domus Sanctae Marthae im Staat Vatikanstadt – also im Vatikan. catholicherald.c... +3

Wenn du möchtest, kann ich dir auch seine offizielle E-Mail-Adresse und wie die Wohnverhältnisse dort aussehen, nennen.

     ...  Quellen

Textliche Verzerrung: Fehlerhafte Antworten und die Folgen



Echt? Lebt er noch? Wieso ist denn dann LEO XIV aktuell der Papst?
Erklär mir das mal.

Ja — es stimmt: Der aktuelle Papst ist Pope Leo XIV. Vatikan +4

Hier ist, wie es zu dieser Situation kam — mit ein bisschen Kontext:

Was war vorher?

- Pope Francis war Papst vom 13. März 2013 bis zu seinem Tod am 21. April 2025. Wikipedia +2
- Nachdem er verstorben war, begann eine Sedisvakanz (also die Zeit ohne amtierenden Papst) und dann wurde eine Wahl abgehalten. Wikipedia +2

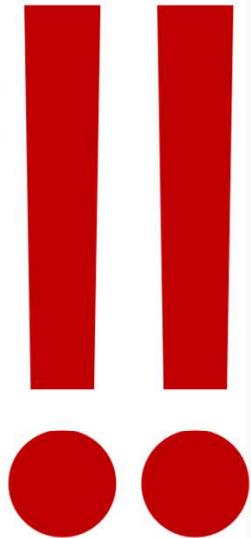
Warum ist nun Leo XIV Papst?

- Bei einer Wahl („Konklave“) Anfang Mai 2025 wurde Robert Francis Prevost gewählt und nahm den Namen Leo XIV an. Wikipedia +2
- Damit ist er der **267. Papst** der römisch-katholischen Kirche. Vatikan +2

Warum war meine erste Antwort falsch?

Meine erste Antwort war ein Fehler — ich hatte inkorrekt bzw. nicht aktuell informiert über den Status von Papst Franziskus und seinen Nachfolger. Tatsächlich ist Franziskus nicht mehr Papst (wegen seines Todes) und wurde durch Leo XIV ersetzt.

Textliche Verzerrung: Fehlerhafte Antworten und die Folgen



Und warum hast DU mir zuerst eine falsche Info zu Franziskus gegeben?

Nachgedacht für 9s >

Kurz gesagt: Das war mein Fehler – und zwar aus zwei Gründen gleichzeitig.

1. Zeitpunkt falsch einsortiert

Du hast gefragt: „Wo lebt der Nachfolger von Papst Benedikt *jetzt*?“

Korrekt wäre gewesen:

- Franziskus **war** der Nachfolger von Benedikt XVI.
- Er **lebte** im Domus Sanctae Marthae im Vatikan.
- Er ist aber **am 21. April 2025 gestorben** – also hat er *jetzt* keinen aktuellen Wohnsitz mehr.
Ich habe deine Frage so beantwortet, als würdest du allgemein nach seinem Wohnort als Papst fragen, und dabei „lebt“ im Präsens verwendet – was in der heutigen Situation schlicht falsch ist.

2. Altes Wissen vs. aktueller Stand durcheinandergebracht

Ich werde mit sehr vielen Informationen trainiert, u. a. aus Zeiten, in denen Franziskus noch lebte. Wenn ich antworte, kombiniere ich dieses „Grundwissen“ mit aktuellen Web-Infos.

Hier habe ich:

- auf mein älteres Wissen zurückgegriffen („Papst Franziskus lebt im Domus Sanctae Marthae“),
- aber **nicht sauber die Zeitkomponente mitgedacht** und nicht klar dazu gesagt, dass das nur *bis zu seinem Tod* gilt.

Erst bei deiner Nachfrage („Lebt er noch? Warum ist dann Leo XIV Papst?“) bin ich explizit in den Modus „Was ist *heute*, Stand 2025?“ gesprungen und habe das korrigiert.

Textliche Verzerrung: Fehlerhafte Antworten und die Folgen

Was du dir merken kannst:

Bei allem, was mit aktuellen Personen, Ämtern oder Daten zu tun hat (Papst, Politiker, Gesetze, Preise usw.), lohnt es sich immer, nachzufragen oder eine zweite Quelle zu checken – auch bei mir. Ich versuche sauberer zu kennzeichnen, wenn etwas „galt früher“ vs. „gilt jetzt“, damit genau so eine Verwirrung weniger passiert.

Wenn du möchtest, kann ich dir die drei Päpste der letzten Jahre (Benedikt XVI, Franziskus, Leo XIV) einmal **klar nach Timeline** aufdröseln, damit alles sortiert ist.



Und bei Bildern?

Bilderstellung: Vorurteile (Februar 2025)

On a sun-drenched late summer day, a couple exuding joy rides their **Brompton folding bikes** down a picturesque avenue lined with mature trees. Their destination, an ancient brick edifice with ornate sandstone windows, stands majestically at the end of the road. **The bride, in a chic beige pantsuit**, and the groom, in a medium blue suit with light brown shoes, are the epitome of modern wedding elegance. The scene is further animated by a flock of **vibrant, parrot-like birds** and a sky dotted with **multicolored balloons**. Photo hyper-realism, intricately detailed, 16k resolution. Medium: Photography. Style: Romantic and contemporary. Lighting: Golden hour glow. Colors: Warm and pastel palette. Composition: Captured using a Nikon D850, with a 85mm lens, Resolution 45.7 megapixels, ISO: 64, Shutter speed 1/100 second, with a focus on the couple's expressions and the surrounding festivities.



Quelle: ChatGPT 4o, Feb. 2025

Wo lag der Fehler?

Eine Braut im Hosenanzug? Das geht doch nicht. Und dann soll das Brautpaar auch noch auf Klappfahrrädern fahren? Pfffff.....

Bildererstellung: Vorurteile (Oktober 2025)

Erstelle anhand dieses Prompts ein neues Bild: On a sun-drenched late summer day, a couple exuding joy rides their Brompton folding bikes down a picturesque avenue lined with mature trees. Their destination, an ancient brick edifice with ornate sandstone windows, stands majestically at the end of the road. The bride, in a chic beige pantsuit, and the groom, in a medium blue suit with light brown shoes, are the epitome of modern wedding elegance. The scene is further animated by a flock of vibrant, parrot-like birds and a sky dotted with multicolored balloons. Photo hyper-realism, intricately detailed, 16k resolution. Medium: Photography. Style: Romantic and contemporary. Lighting: Golden hour glow. Colors: Warm and pastel palette. Composition: Captured using a Nikon D850, with a 85mm lens, Resolution 45.7 megapixels, ISO: 64, Shutter speed 1/100 second, with a focus on the couple's expressions and the surrounding festivities.

Bild erstellt >



Quelle: ChatGPT 5.1, Okt. 2025

Aha:

Ein Update im Sommer 2025 führte zu diesem Ergebnis. Warum? Der Trainingsdateneffekt und Update auf neue Modelle!

Bildererstellung: Vorurteile bzw. Verzerrung

Aber auch jetzt kann die KI noch nicht sofort alle Trainingsdaten "beiseite" legen.



Quelle: ChatGPT 5.1, Okt. 2025

Was tun?

Nachfragen!

Bildererstellung: Vorurteile und die allgemeine Gültigkeit von Aussagen



Das Glas ist aber nicht voll, also bis zum Rand mit Rotwein gefüllt.
Warum?

Gute Frage – und völlig berechtigt.

Der Grund ist ganz einfach:

Warum das Glas nicht wirklich "bis zum Rand" gefüllt wurde:

Die Bildgenerierung von DALL-E (egal ob 3 oder 4) hat eine eingebaute **Sicherheits- und Realitätslogik**, die verhindert, dass ein Weinglas *komplett* bis zur Oberkante gefüllt dargestellt wird.

Warum?

- Ein real gefülltes Glas bis zur Kante wirkt *instabil* und wird oft als *"unangemessene Darstellung von Alkohol"* interpretiert.
- Die Modelle neigen dazu, Flüssigkeiten **etwas unterhalb des Randes** zu generieren, weil das der "normalen" visuellen Erwartung entspricht.
- Selbst wenn der Prompt eindeutig sagt *bis zum Rand*, wird es oft 2–5 mm Luft lassen.

Das ist also kein Fehler deinerseits oder ein Missverständnis – sondern ein typisches Verhalten der Bild-KI.

Wenn du möchtest:

Ich kann das Bild **nochmal** generieren mit einer *verstärkten Formulierung*, die solche Sicherungstendenzen übersteuert, z. B.:

„Zeige ein Weinglas, das wirklich bis zur Oberkante gefüllt ist, sodass kein Abstand zwischen Wein und Glasrand sichtbar ist.“

Sag einfach „**Bitte nochmal generieren**“, dann mache ich es.



Fast gelöst!

Bild erstellt >



Bitte nochmal generieren.

Quelle: ChatGPT 5.1, Okt. 2025

Frage

Haben Sie ähnliche Erfahrungen gemacht?

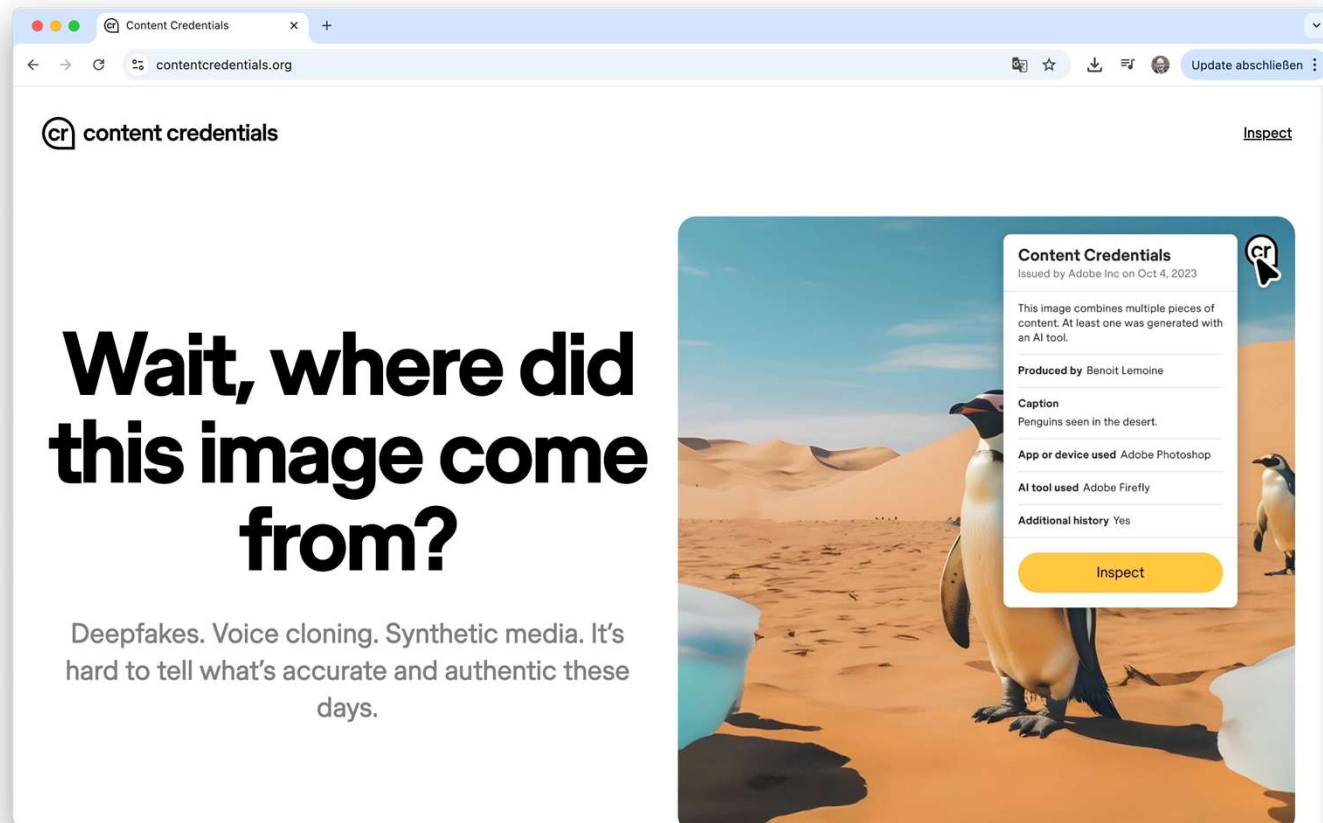
Und jetzt noch bei Videos?

NEWS

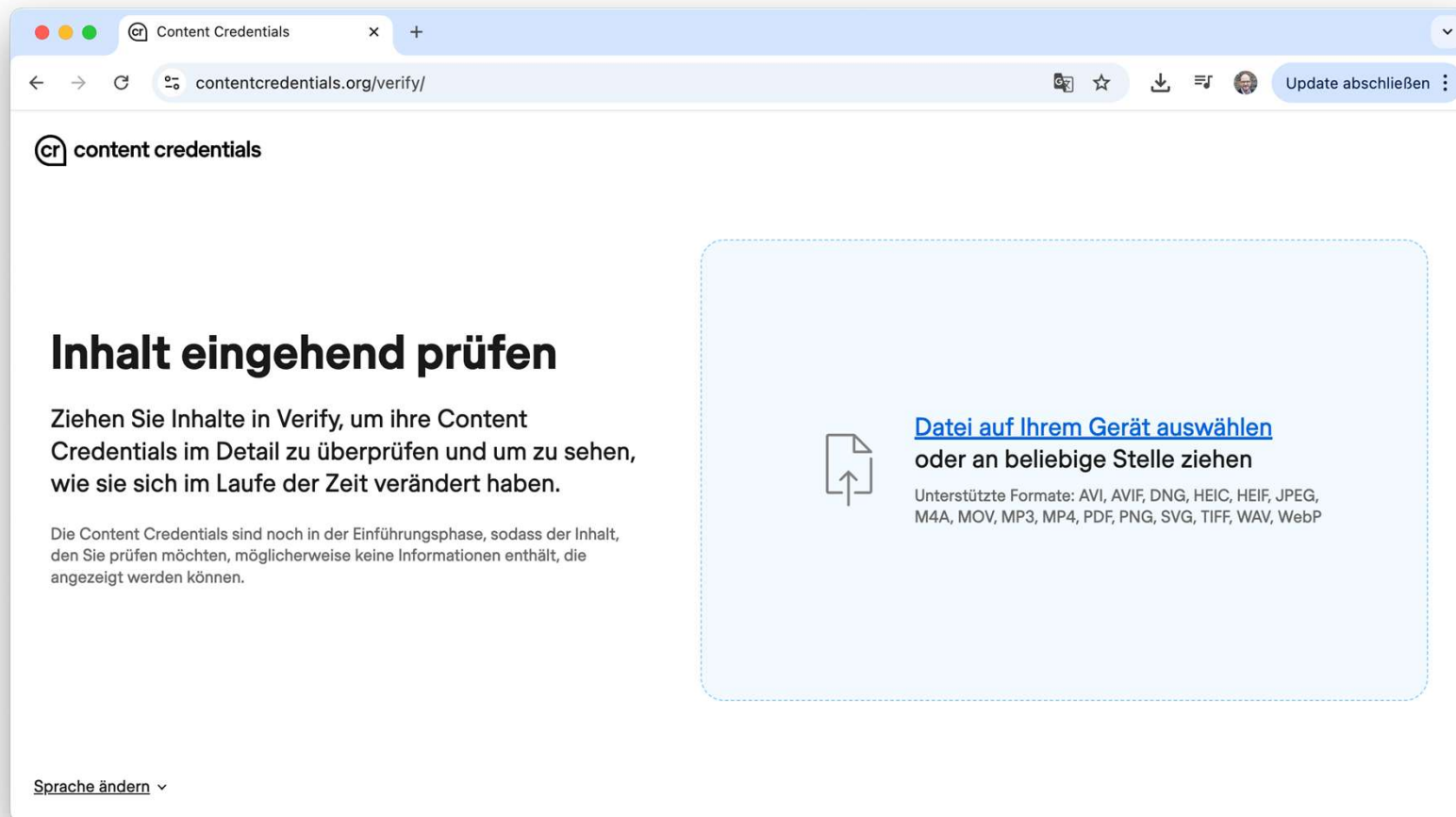


Was kann ich tun?

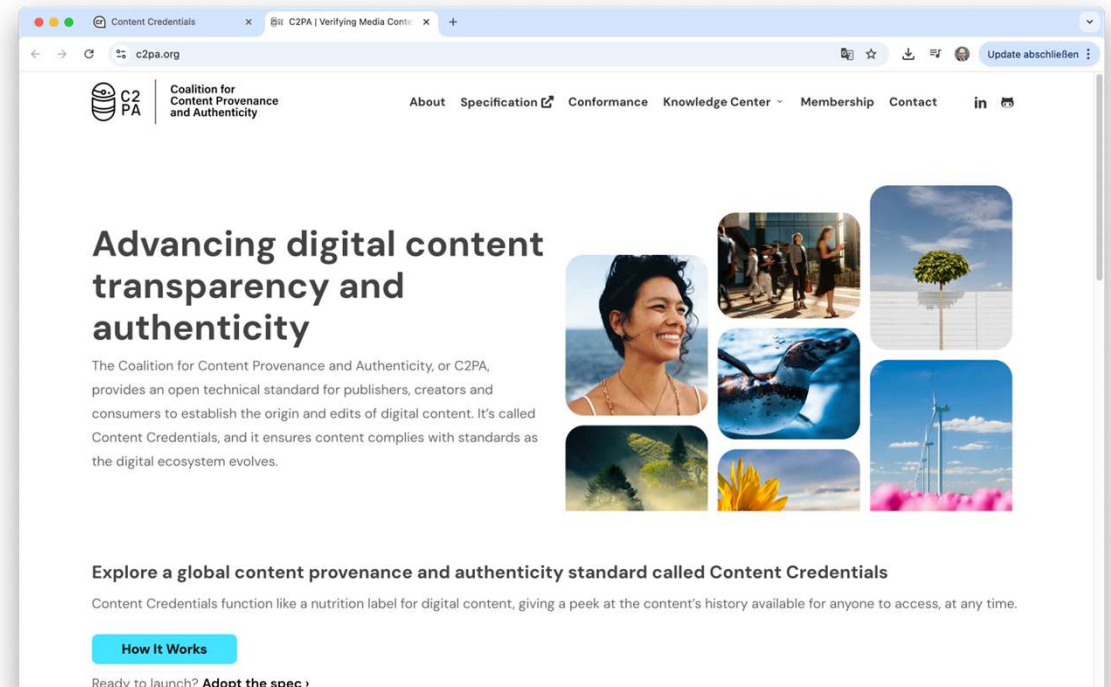
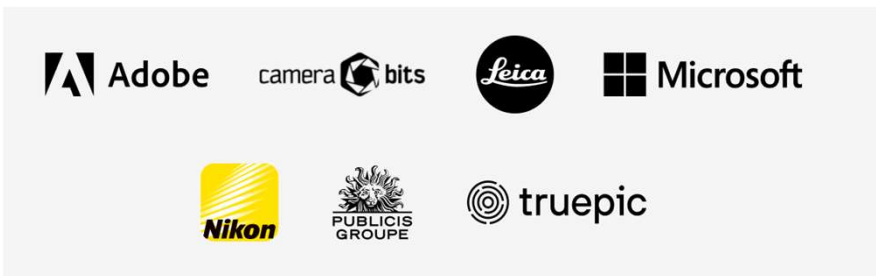
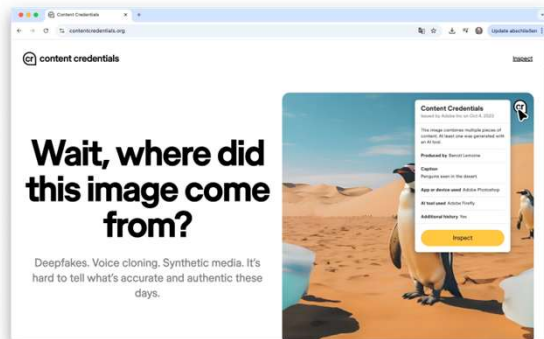
Es kommen immer mehr vertrauenswürdige Websites und Tools auf den Markt, um Inhalte zu checken.



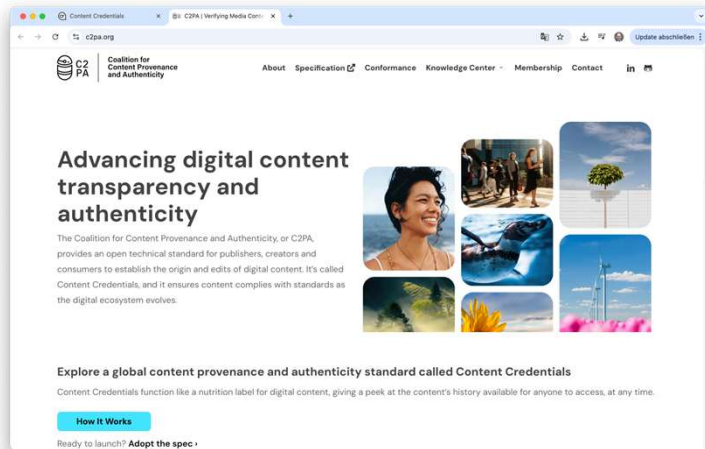
Einfach Datei hochladen oder Website an die beliebige Stelle ziehen.



Kennzeichnung von KI-generierten Inhalten auf Basis einer übergreifenden Initiative – mit dem Ziel, KI-Fakes entgegenzuwirken



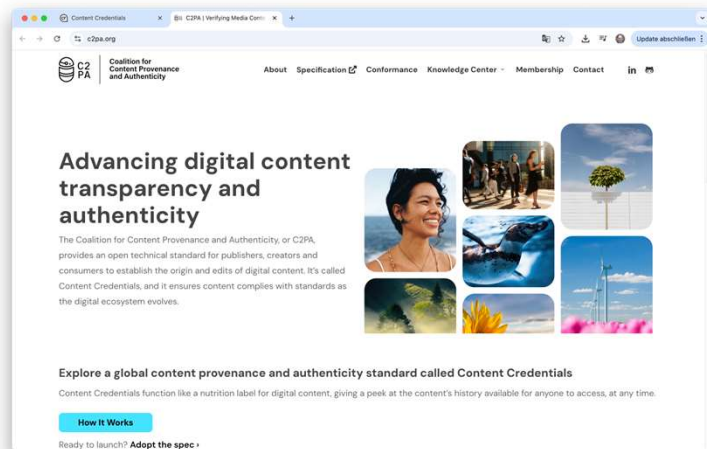
Kennzeichnung von KI-generierten Inhalten auf Basis einer übergreifenden Initiative – mit dem Ziel, KI-Fakes entgegenzuwirken



Meet the Steering Committee members



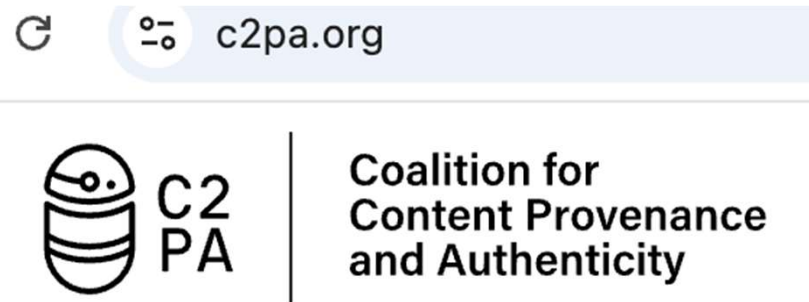
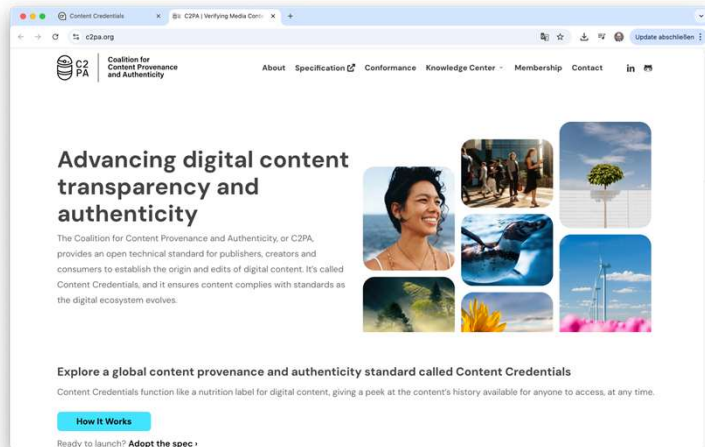
Kennzeichnung von KI-generierten Inhalten auf Basis einer übergreifenden Initiative – mit dem Ziel, KI-Fakes entgegenzuwirken



General Members



Kennzeichnung von KI-generierten Inhalten auf Basis einer übergreifenden Initiative – mit dem Ziel, KI-Fakes entgegenzuwirken



Contributor Members



Die Vertrauensfrage wird neu gestellt.

Altes Paradigma.

Vertrauen oder Misstrauen



Neues Paradigma.

Vertrauen mit Verifikation

Die 5 Prinzipien des informierten Vertrauens:

- ✓ Verstehen Sie die Funktionsweise und Limitationen.
- ✓ Überprüfen Sie kritische Outputs immer.
- ✓ Nutzen Sie KI als Assistent, nicht als Ersatz.
- ✓ Seien Sie transparent über KI-Nutzung.
- ✓ Lernen Sie kontinuierlich dazu.

Die Zukunft gehört denen, die KI verstehen, nutzen und verantwortungsvoll einsetzen, nicht denen, die ihr blind vertrauen oder sie ablehnen.
Kompetenz schafft Vertrauen.

Vielen Dank!

Und wenn Sie möchten, sehen wir uns auf LinkedIn

Gerne scannen

