
Gabor Pox
Superintelligenz
KI außer Kontrolle
K.I. -> S.I.
K.D. -> S.D.
A.I. -> S.I.
A.S. -> ??

Bibliografische Information der Deutschen Nationalbibliothek:
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über dnb.dnb.de abrufbar.

Die automatisierte Analyse des Werkes, um daraus Informationen insbesondere über Muster, Trends und Korrelationen gemäß §44b UrhG („Text und Data Mining“) zu gewinnen, ist untersagt.

Copyright © der Erstausgabe 2026 Gabor Pox
Copyright © 2026 Artishock.de

Version 1.0.27.02
Made in Germany

Infos und Musik zum Buch:
artishock.de/si

Zuletzt erschienen vom Autor:
42 artishock.de/42
Palindrom artishock.de/palindrom
Don’t drive my Car artishock.de/dontdrivemycar
Mein großes Kino artishock.de/meingrosseskino
Elon Musk‘s Telefonzelle artishock.de/telefonzelle
Letzte Ausfahrt Patagonien artishock.de/patagonien
Künstliche Dummheit artishock.de/kd

Gabor Pox

Superintelligenz

KI außer Kontrolle

EIN SACHBUCH

Für Elija

Inhalt

VORWORT **7**

1 - DEFINITION **8**

2 - INTELLIGENZ **10**

3 - SUPERINTELLIGENZ **12**

4 – DIE NUTZUNG VON KI **14**

5 – DARF ICH EIN FOTO MACHEN? **17**

6 – BILDER UND VIDEOS MIT KI **19**

7 – KI ALS JOB-TERMINATOR **22**

8 – KI UND KLIMAWANDEL **24**

9 – KI IN DER LITERATUR **26**

10 – KI IN DER MUSIK **28**

11 – KÜNSTLICHE DUMMHEIT **30**

12 – PANIK AUF DER TITANIC **32**

13 – ALADDIN UND DER FLASCHENGEIST **34**

14 – INDUSTRIELLE KI **36**

ABSPANN **38**

NOTIZEN **40**

ANHANG A – IQ-TEST **41**

ANHANG B – KI-WASSERZEICHEN **47**

SUPERINTELLIGENZ – KI AUßER KONTROLLE **54**

Superintelligenz bezeichnet Wesen oder Maschinen mit dem Menschen in vielen oder allen Gebieten überlegener Intelligenz. Der Begriff findet insbesondere im Transhumanismus und im Bereich der Science-Fiction Verwendung. Ein tatsächlich geistig überlegenes Wesen, das die Kriterien einer Superintelligenz erfüllt, gibt es nach heutigem Kenntnisstand nicht. [Wikipedia]

Im September 2026 kündigte die Regierung von Donald Trump in den Vereinigten Staaten an, den Begriff Künstliche Intelligenz (KI) durch Superintelligenz (SI) zu ersetzen. Dies führte zu einem starken Anstieg der Registrierungen von Domains mit der Top-Level-Domain .si. Innerhalb von zwei Tagen wurden mehr als 6000 neue Domains registriert. [Wikipedia]

Vorwort

Künstliche Intelligenz ist aus unserem Leben nicht mehr wegzudenken. Der einzige Weg, um die Begriffe „KI" und „AI" (engl. für „*artificial intelligence*") aus dem Weg zu räumen, ist eine Umbenennung. Es lebe die *Superintelligenz*!

Spaß beiseite, ich schreibe derzeit ein Sachbuch und werde mich bemühen, fachlich und objektiv zu bleiben. Ich bin Elektroingenieur und Softwareentwickler, seit Jahren Experte für künstliche Intelligenz. Über die aggressive Verbreitung der KI habe ich bisher Romane und fiktionale Werke geschrieben. In meiner Tetralogie „*Künstliche Dummheit*" beschreibe ich eine mögliche Zukunft, die uns bewegt, berührt und gleichzeitig erschreckt.

Und jetzt? Ich bleibe auf dem Boden der Tatsachen und mische dabei technischen Fortschritt mit Moral und Ethik, Philosophie mit menschlicher Dummheit sowie die vorgetäuschte Macht der Politik mit den Verschwörungen der Tech-Gurus. Ich werde Fragen stellen und versuchen, die meisten davon zu beantworten. Eine Antwort im Voraus: 42.

Gabor Pox, November 2026

1 - Definition

Laut Experten gibt es eine gut definierbare Abfolge technischer Revolutionen in der Menschheitsgeschichte. Diese lässt sich in vier Hauptphasen unterteilen. Auf die Phase der *Mechanisierung* folgte die *Elektrifizierung*. Danach kam die Phase der *Automatisierung*, gepaart mit der Computerisierung. Wir erleben derzeit die vierte Phase, die von *Vernetzung*, intelligenter Kommunikation, Cloud-Computing und dem Vormarsch der künstlichen Intelligenz geprägt ist. (Man sollte jetzt KI fragen: "Was kommt danach?")

Als ich in der dritten technischen Phase mein Berufsleben startete, war der Begriff „Computerisierung" für mich von entscheidender Bedeutung. Genauso wie die Menschheit die Sprache, das Schreiben und das Lesen als Kommunikationsmittel erfand, war die Computerisierung für mich die logische Fortsetzung. Unabhängig davon, ob es um Technik, Datenspeicherung oder Unterhaltung geht. Man kann nicht alles gleichzeitig im Kopf haben. Wenn man programmiert (oder ein Buch schreibt), möchte man die Gedanken, in die man gerade tief eingetaucht ist, gerne festhalten. Am Folgetag hat man bereits etwas anderes im Kopf: eine andere Programmfunktion oder ein anderes Buchkapitel. Der Computer vergisst aber nichts. So lässt sich menschliche Intelligenz gut konservieren. Hinzu kommt, dass ein Netzwerk aus Computern (oder Datenspeichern) die Gedanken mehrerer Menschen festhalten und verschmelzen kann. So nutzte ich die Computerisierung bereits in den Achtzigerjahren, um intelligente, vollautomatisierte Fabriken mitzugestalten. Damals nannte man das noch nicht KI, aber wenn wir jetzt zurückblicken, war es das schon.

Ich begrüße den aktuellen Vorschlag von *Donald Trump*, das Wort „künstlich" aus dem Weg zu räumen und stattdessen den Begriff „Superintelligenz" zu nutzen. Wissen und Intelligenz sind menschlich, auch wenn sie ausgelagert werden. Nebenbei bemerkt kritisiere ich jedoch Trumps Vorgehensweise, alles, was ihm in den Sinn kommt, sofort per

Dekret oder Gesetz festzuschreiben. Superintelligenz ist definitiv eine Form der kollektiven Intelligenz. Und sie stammt von uns, von den Menschen. Das klingt schon super.

Aber wie *intelligent* sind wir wirklich?

2 - Intelligenz

Intelligenz bezeichnet die geistige und kognitive Leistungsfähigkeit von Menschen (und teilweise auch von Tieren), Probleme zu erkennen, zu verstehen und zu lösen. So lautet die offizielle Definition. Dabei fällt auf, dass die Begriffe „künstlich" oder „maschinell" dabei gar nicht vorkommen. Eine weitere Angabe besagt: „Es gibt keine einzelne, weltweit verbindliche Definition." Vermutet wird, dass diese letzte Formulierung von einer KI stammt, die offenbar zunehmend Ausreden sucht.

Machen wir jetzt ein Gedankenspiel. Person A kann sehr gut Schach spielen. B spricht drei Sprachen perfekt. C ist Experte auf einem Fachgebiet und treibt die Forschung erfolgreich voran. D verfügt über gutes Allgemeinwissen und kann Fragen stets gut beantworten. E macht regelmäßig IQ-Tests und erreicht einen IQ-Wert von über 140. Welche Person ist intelligenter?

Schlaue Leser könnten nun sagen: Definieren wir eine Person F, die all diese Fähigkeiten besitzt, und diese ist der Gewinner. Das klingt gut, aber das war nicht die Frage. (Abgesehen davon, dass ich tatsächlich diese Person F sein könnte…)

Betrachten wir die Personen einzeln. „A" könnte zu sehr auf das Schachspielen fixiert sein und im Alltag große Probleme bekommen. „B" hat bewiesen, dass sie gut Sprachen lernen und sich Wörter gut merken kann. Für eine gute Bewertung müsste man jedoch wissen, ob sie aus freien Stücken gelernt hat oder einfach in mehreren Ländern nacheinander gelebt hat. „C" könnte ähnlich wie der Schachspieler eingeschränkt sein - wenn das Fachgebiet jedoch KI heißt, würde sie Bonuspunkte sammeln. „D" ist leider ein Auslaufmodell und kann mit den Chatbots nicht mehr Schritt halten. „E" profitiert anscheinend davon, dass diverse IQ-Tests sehr ähnliche Aufgaben stellen und man nach Dutzenden Tests immer besser abschneidet.

Uns sollte jetzt auffallen, dass gesellschaftliche, zwischenmenschliche oder künstlerische Aspekte in der obigen Auflistung komplett fehlen. Die Ergebnisse von IQ-Tests gelten vielen Wissenschaftlern ohnehin als Indikator eher begrenzte Fähigkeit. Der US-Psychologe *Howard Gardner* definiert daher unterschiedliche Formen von Intelligenz. Damit hat er die Theorie der „multiplen Intelligenzen" entwickelt. Danach gibt es mehrere voneinander unabhängige Intelligenzen: eine *sprachliche*, eine *musikalische*, eine *logisch-mathematische*, eine *räumliche*, eine *körperlich-kinästhetische* und eine *naturkundliche* Intelligenz. Zudem erleichtert eine *interpersonale* Intelligenz die Arbeit mit Mitmenschen, während eine *intrapersonelle* Intelligenz dabei hilft, uns selbst besser zu verstehen und zu beeinflussen.

Ich finde diese Definition mehr als zutreffend, da ich damit besser erklären kann, wo die künstliche Intelligenz große Defizite hat. Aktuell kann KI nur drei der acht Formen beherrschen: die sprachliche, die logisch-mathematische und die naturkundliche. Um zu zeigen, wie KI-Chatbots einen IQ-Test bestehen würden, habe ich einige Fragen und Antworten im Anhang A zusammengestellt.

3 - Superintelligenz

Ich habe mich entschieden, die Begriffe „*Künstliche Intelligenz*" und „*Superintelligenz*" in diesem Buch gleichwertig und abwechselnd zu verwenden. Ich möchte noch nicht vollständig umsteigen. Aber warum nicht? Zunächst einmal ist das Vorhaben der Umbenennung noch im Gange. Außerdem hat „Superintelligenz" etwas Außerirdisches an sich, wie der Supercomputer *Deep Thought* mit der weltberühmten Antwort „42" suggeriert.

Was wir Menschen derzeit aufbauen, ist eine gigantische Wissenssammlung: Fakten, Wikipedia, das Internet, die Texte aller geschriebenen Bücher sowie gebündelte Interaktivitäten, die mit logischem Denken und komplexen Programmabläufen zu tun haben. Diese Wissenssammlung können wir Superintelligenz nennen. Welche Auswirkungen die unkontrollierte Weiterentwicklung der KI (oder Superintelligenz) mit sich bringt, ist das Hauptthema dieses Buches. Im Folgenden möchte ich kurz auf zwei Aspekte eingehen.

Erstens führt diese Auslagerung dazu, dass wir unsere Gehirne immer mehr entlasten und quasi „entleeren", da wir immer weniger Fakten im Kopf behalten müssen. Diese können wir jederzeit abfragen. Was „jederzeit" bedeutet, kann man sich ausmalen. Ich erinnere mich daran, wie Taschenrechner sich verbreiteten und der Schulunterricht im Bereich des schriftlichen Dividierens umgestellt werden sollte – hin zu halbschriftlichem Dividieren und vereinfachten Rechenstrategien statt dem Auswendiglernen von Algorithmen. Die Gegner dieses Vorhabens argumentierten, dass man nicht jederzeit einen Taschenrechner zur Hand habe. Jahre später stellte sich heraus, dass eigentlich jeder ein Smartphone mit Taschenrechner-App dabeihat. So erfolgte die Anpassung des Lehrplans – zumindest in einigen Bundesländern. Man hatte jedoch nicht damit gerechnet, dass bald Handy-Verbote in den Schulen und manchmal auch außerhalb eingeführt würden. Und was ist, wenn die Batterie

leer ist? Unabhängig davon möchten wir nicht mehr alles auswendig lernen und im Kopf mitschleppen. Dadurch entsteht jedoch eine starke Abhängigkeit. Es klingt logisch, dass man Differential- und Integralrechnungen nicht mehr auswendig lernen sollte. Diese Aufgaben verschwinden langsam aus den Abiturprüfungen. Was wir dabei jedoch eigentlich lernen, ist nicht die Steigung des Graphen einer Funktion an einem bestimmten Punkt oder das Berechnen von Flächen unter Kurven, sondern die Fähigkeit, etwas Neues und Kompliziertes zu verstehen und zu erlernen. Wenn wir diese Fähigkeit nicht trainieren, bekommen wir spätestens bei unserem ersten Job Probleme. Falls wir noch einen Job bekommen.

Zweitens birgt die Auslagerung von Wissen große Risiken. Rechenzentren und Cloud-Systeme sind technisch konzentriert und somit zentral angreifbar. Sie sind leicht abschaltbar. Das Konzept des Internets entstand in den 1960er-Jahren in den USA. Das Verteidigungsministerium suchte ein dezentrales, krisensicheres Netzwerk für Kommunikation. Was wir heute erleben, ist eine neue Zentralisierung von Daten und Datenzugriffen. Selbst wenn Personen komplette KI-Modelle herunterladen und anpassen können, würde das in Krisensituationen nicht viel helfen. Man könnte vielleicht mit Stolz sagen: „Ich habe nicht nur eine Notreserve für zwei Monate im Keller, sondern auch ein lokales KI-Modell auf der Festplatte." Damals waren die Bücher dezentral und redundant in Millionen von Bücherregalen zu finden. Das Verschwinden eines weltberühmten Romans war ausgeschlossen. Wenn wir die Weltliteratur jedoch als E-Books in zentralen Clouds lagern, ist das Risiko eines Verlustes viel höher. Und ich habe noch nicht einmal darüber gesprochen, wie lange Speichermedien Informationen speichern können. Bücher schaffen es immerhin seit hunderten von Jahren.

Technische Entwicklungen lassen sich jedoch nicht stoppen. Das war bei der Nuklearforschung so, bei der Weltraumforschung auch – und jetzt passiert es bei der KI-Entwicklung. Wir entwickeln unsere Superintelligenz weiter und hoffen, dass wir die Kontrolle nicht verlieren.

4 – Die Nutzung von KI

Bei der Nutzung von KI hat man oft ein ungutes Gefühl. Darf ich die Antwort oder den erstellten Programmcode so verwenden, als wäre es von mir selbst erstellt worden? Um diese Frage im Hinblick auf Urheberrechte zu klären, müssen wir der Frage nachgehen, woher die Informationen ursprünglich stammen. Ich zitiere zunächst einige Antworten der „Google-KI“:

„Rein KI-generierte Inhalte haben in der Regel keinen Urheberrechtsschutz und sind gemeinfrei, da ein Urheberrecht nach deutschem und europäischem Recht nur für persönliche geistige Schöpfungen von Menschen vorgesehen ist.“

„Wenn Sie einen KI-Output stark bearbeiten, anpassen oder durch sehr komplexe, kreative Eingaben steuern (sofern ein prägender menschlicher Anteil erkennbar ist), kann im Einzelfall ein eigenes Urheberrecht entstehen.“

„Wenn Sie urheberrechtlich geschütztes Material (z. B. Bilder oder geschützte Texte) in eine KI eingeben oder die KI Ihnen exakte Kopien von geschützten Werken ausgibt, tragen Sie bei einer Veröffentlichung oder kommerziellen Nutzung das Risiko einer Urheberrechtsverletzung.“

Heutzutage befinden wir uns in einer unüberschaubaren Situation. Große KI-Firmen – besser gesagt, deren KI-Modelle – konsumieren ständig Literatur und andere Quellen aus dem Internet, ohne auf die Urheberrechte zu achten. Konzerne kaufen weltweit massenhaft antiquarische Bücher und Ladenhüter auf, um ihre künstliche Intelligenz mit gedrucktem Fachwissen und analogen Texten zu trainieren. Da fehlt definitiv eine Regulierung. Es fehlt auch eine Kennzeichnung, die angibt, woher die gelieferten Informationen stammen. Seien wir ehrlich: Wir nutzen gerne die Antworten eines Chatbots, weil wir der Meinung sind, dass eine mögliche Urheberrechtsverletzung auf Seiten der KI liegt und wir somit von der Verantwortung „befreit“ sind. In dieser unüberschaubaren Situation glauben wir, die Spuren der Nachverfolgung schön verwischt zu haben. Dabei denkt

man nicht daran, dass die Spuren doch gespeichert sind und man diese später, nach Einführung einer Regulierung, finden könnte.

Nicht falsch verstehen, ich bin keine Regulierungsbehörde, jedoch jemand, der sich mehr Transparenz und Kontrolle wünscht. Wenn man einen Text hochlädt und eine Korrektur verlangt, wäre es denkbar, dass die KI den ursprünglichen Verfasser des Textes feststellt und im Falle einer Urheberrechtsverletzung eine Umschreibung ablehnt. So etwas funktioniert bereits bei hochgeladenen Bildern. Wenn man eine Überarbeitung für den eigenen Text zurückbekommt, wäre es ebenfalls denkbar, die Stärke der Überarbeitung auf einer Skala zurückzumelden. Beispiel: 10 = komplett durch KI erstellt, 1 = nur grammatikalische Anpassungen.

Bei der Veröffentlichung von Büchern oder Videos wird von Plattformen oft die Frage gestellt, ob bei der Erstellung KI verwendet wurde. Ich finde es unpassend, diese Frage pauschal zu stellen. Es macht doch einen Unterschied, ob nur eine grammatikalische Korrektur durchgeführt wurde oder ob der Text eines Buchs oder einer Diplomarbeit komplett von KI stammt. Und bei einem Video ist es ebenfalls schwer, mit Ja oder Nein zu antworten, wenn man beispielsweise nur die Untertitel mit KI erstellt hat. Fakt ist: Es herrscht Chaos auf der KI-Ebene und wir sind irgendwie genötigt, KI nur reduziert oder sogar gar nicht zu nutzen. Sonst werden wir als „unerwünschte“ KI-Autoren eingestuft. Haben die Großkonzerne Milliarden für Superintelligenz investiert, um uns Benutzer abstempeln zu können? Vielleicht sollte die generelle Akzeptanz der KI-Nutzung erhöht werden. Dazu müssen zuerst aber die Qualität und Transparenz der KI-Systeme steigen!

Ein weiteres Beispiel: Man schreibt einen Text mit einem Texteditor eines Office-Pakets. Dabei erfolgt eine integrierte Grammatikprüfung, die man aber noch nicht als KI-Unterstützung wahrnehmen sollte. Zusätzlich hat man die Möglichkeit, den markierten Text zu übersetzen. Bin ich dadurch schon ein KI-User? Und was ist, wenn ich diese Schritte überspringe und später ChatGPT oder „DeepL“ bitte, dies zu tun? Wenn ich meinen selbstgeschriebenen Text als

PDF hochlade und die Plattform daraus ein E-Book generiert, ist das Buch dann durch KI erstellt? Wenn Amazon Kindle daraus automatisch ein Hörbuch kreiert, dann stehe ich als Autor ganz oben auf der Liste der „unbeliebten“ Publizisten, die es anscheinend verdient haben, auf dem Buchcover das Label „KI-generiert“ zu bekommen.

Fragen über Fragen. Nicht einmal die KI kann diese Fragen beantworten. Die Antwort würde wahrscheinlich so klingen: „Wir befinden uns in einer unklaren Situation, wann ein Werk als KI-unterstützt deklariert werden soll.“ Es fehlt nur noch die Ergänzung: „Und es ist gut so!”

5 – Darf ich ein Foto machen?

Ich bin nicht nur Buchautor, sondern auch Fotograf. Dabei achte ich auf die Urheberrechte. Doch welche Vorgaben muss ich wirklich beachten? Und worauf müssen andere Personen bei der Verwendung meiner Bilder Rücksicht nehmen?

„Das Urheberrecht für Fotos schützt den Fotografen persönlich, während das Nutzungsrecht regelt, wie und wofür andere die Bilder verwenden dürfen. Das Urheberrecht bleibt in Deutschland nach Paragraph 7 UrhG unauflöslich beim Urheber (dem Fotografen) und kann nicht übertragen werden. Kunden oder Agenturen kaufen stattdessen immer nur Nutzungsrechte (Lizenzen) an den Bildern."

So weit, so gut. Doch was darf man eigentlich fotografieren? Und was darf man davon veröffentlichen? Zum Beispiel in einer Fotoausstellung oder auf der Website? Ich fasse hier die zum Zeitpunkt des Schreibens des Buches geltenden Regelungen zusammen.

In der Landschafts- und Straßenfotografie ist man in der Regel auf der sicheren rechtlichen Seite. So ist es beispielsweise erlaubt, Landschaften und Gebäude zu fotografieren und die Bilder zu verkaufen oder zu veröffentlichen. Der zentrale Begriff hierbei ist die Panoramafreiheit. Sie gilt jedoch nur, wenn man sich auf öffentlichem Boden befindet. Werden die Aufnahmen hingegen von einer Drohne, einer Leiter, einem Privatbalkon oder aus einem Hubschrauber heraus gemacht, greift die Panoramafreiheit nicht. Dauerhaft installierte Kunstwerke (zum Beispiel Statuen auf Marktplätzen) fallen ebenfalls unter die Panoramafreiheit. Temporäre Kunstwerke (zum Beispiel der verhüllte Reichstag von *Christo*) dürfen hingegen nicht ohne die Erlaubnis des Künstlers verbreitet werden.

Bezüglich der Personen gilt Folgendes: Gemäß Paragraph 22 des Kunsturhebergesetz (KUG) dürfen Bilder von erkennbaren Personen grundsätzlich nur mit deren Einwilligung verbreitet oder öffentlich gezeigt werden. Dies gilt für eine oder zwei Personen ebenso wie für eine große

Gruppe. Eine Veröffentlichung von Bildern mit Personen ist ohne Erlaubnis nur in wenigen, streng definierten Ausnahmen erlaubt. 1.) Bilder von Versammlungen oder ähnlichen Veranstaltungen wie Demonstrationen oder große Volksfeste, bei denen die Menschenmenge im Vordergrund steht und nicht der einzelne Mensch. 2.) Personen, die nur als unwesentliches Beiwerk auf dem Bild zu sehen sind (zum Beispiel ein Passant im Hintergrund einer Sehenswürdigkeit). 3.) Bilder aus dem Bereich der Zeitgeschichte, wie Politiker bei öffentlichen Auftritten.

Beim Fotografieren eines Kunstwerkes ist zu beachten: Ist der Urheber eines Werkes seit mehr als 70 Jahren tot, ist das Werk gemeinfrei. Es darf dann grundsätzlich fotografiert und die Bilder genutzt werden, sofern das Museum kein generelles Fotoverbot hat. Leben die Künstler noch oder ist ihr Tod noch keine 70 Jahre her, sind die Werke urheberrechtlich geschützt. Private Fotos werden manchmal toleriert, für die Veröffentlichung (im Internet oder in den sozialen Medien) ist jedoch die Erlaubnis des Urhebers erforderlich. Und was ist mit einem Selfie vor einem noch urheberrechtlich geschützten Kunstwerk? Zwar heißt es im Urheberrechtsgesetz: „Zulässig ist die Vervielfältigung, Verbreitung und öffentliche Wiedergabe von Werken, wenn sie als unwesentliches Beiwerk neben dem eigentlichen Gegenstand des Bildes oder Videos anzusehen sind.", jedoch hat man bei einer eventuellen Klage schlechte Karten. Bei „Kunstwerk trifft fremde Betrachter" käme es auf eine Kombination der bereits erwähnten Regelungen an.

Wie wir sehen, sind die Regelungen bereits jetzt kompliziert genug, auch ohne KI. Das Verbot, eine Leiter zu verwenden, war mir beispielsweise nicht bekannt. Für mich gehört es in die Kategorie „EU-Regelung für den Krümmungsgrad einer Gurke". Aber zurück zur Superintelligenz: Bringt uns die KI eine Vereinfachung oder bedeutet sie noch mehr Chaos? Die Antwort kann man sich schon denken.

6 – Bilder und Videos mit KI

Es gibt Online-Plattformen, von denen aus man Bilder für die freie Verwendung herunterladen kann. Damit liegt die Verantwortung bezüglich des Urheberrechts bei der Plattform oder sogar bei der Person, die das Bild hochgeladen hat. Ob das Bild zuvor durch KI erstellt oder bearbeitet wurde, spielt für die Wiederverwendung dann keine Rolle mehr. Spannend wird die Situation, wenn man eine KI-Plattform direkt beauftragt, ein Bild zu generieren. Weder der Nutzer, der den Prompt eingegeben hat, noch der Programmierer oder die KI selbst erhalten dabei ein Urheberrecht an einem rein KI-erzeugten Bild. Solche Bilder sind in der Regel gemeinfrei, das heißt, prinzipiell darf jeder sie kopieren oder weiterverwenden, weil niemand Exklusivrechte daran besitzt. Allerdings haftet der Nutzer für Rechtsverletzungen, wenn das Bild geschützten Werken, Marken oder Persönlichkeitsrechten Dritter ähnelt – woher sollte man das aber wissen? Auch wenn der Gesetzgeber kein Urheberrecht vorsieht, können die Betreiber der KI-Generatoren (wie *Midjourney* oder *DALL-E*) in ihren Nutzungsbedingungen Regeln oder Einschränkungen für die kommerzielle Nutzung festlegen. Bei manchen Diensten ist ein kostenpflichtiges Abonnement Voraussetzung für kommerzielle Nutzungsrechte.

Für KI gelten Kennzeichnungspflichten. So schreibt beispielsweise die europäische KI-Verordnung (AI-Act, Artikel 50) vor, KI-generierte Inhalte als solche zu kennzeichnen. Dies ist auch auf vielen Social-Media-Plattformen Pflicht.

Jetzt sind wir wieder beim Thema. Ich sehe einen großen Unterschied zwischen einem Bild, das nach einer Prompt-Vorgabe vollständig von KI erstellt wird, und einem Foto, das mithilfe einer KI-Funktion eines lizenzierten Bearbeitungsprogramms wie Photoshop leicht erweitert wird, zum Beispiel um die Vorgaben für das Bildformat und die Seitenverhältnisse zu erfüllen. Das Gleiche gilt, wenn ein

störendes Objekt wie eine Stromleitung oder ein Straßenschild durch KI entfernt wird. Ich weiß, dass es nicht einfach ist, die Grenzen zu ziehen, denn man könnte mehr als die Hälfte eines Bildes „erweitern".

Die nächste Stufe wird durch Bildmanipulation erreicht. So kann man beispielsweise einen Beauty-Filter für ein Porträt einsetzen oder sogar die Gesichter fremder Personen austauschen. Letzteres ist jedoch nicht erlaubt, da es eine erkennbare Verletzung der Persönlichkeitsrechte darstellt und in Richtung Fake News geht. Damit sind wir beim Thema: KI-generierten Schauspielern. Wenn jemand die „*Top Gun*"-Filme kennt und mag, kann ich empfehlen, bei YouTube nach „Top Gun Cast: Then and Now – 40 Years Later" zu suchen. Das Video zeigt, dass KI nicht nur visuelle Perfektion, sondern auch Emotionen erzeugen kann. Ein Beispiel ist dabei *Val Kilmer*s Erscheinung. Er hatte noch seine echte Rolle in „Top Gun: Maverick" gehabt, obwohl er schon sehr krank war und kaum mehr sprechen konnte. Zu seiner traurigen Geschichte gehört, dass er keinen weiteren Film mehr als Schauspieler erleben konnte. Eine Rolle im geplanten Film „As Deep as the Grave" war eigens für ihn geschaffen worden. Aus Verzweiflung zog der Regisseur *Coerte Voorhees* dann die digitale Trumpfkarte und nutzte KI für Vals ursprüngliche Rolle. Der im Jahr 2025 verstorbene Schauspieler wurde durch generative KI für den Western digital rekonstruiert. Das Resultat ist sehenswert, sollte aber als Ausnahme gelten. Das Projekt entstand in voller Absprache und Kooperation mit seinem Nachlass sowie seinen Kindern und unter Einhaltung der SAG-AFTRA-Richtlinien.

Hat KI eine große Zukunft in der Filmindustrie? Definitiv. In der Nähe meines Wohnorts in Bayern befinden sich die „*Penzing Studios*". Dort ist „Virtual Production" bereits fest etabliert. Das bedeutet jedoch nicht, dass die Szenen und Schauspieler für den finalen Film virtuell erstellt sind. Meistens entstehen hier digitale Entwürfe für große Produktionen, sodass man die Szenen im Voraus komplett anschauen und das Drehbuch optimieren kann, ohne die kostspielige Zeit echter Schauspieler für solche Versuche zu nutzen. Nach der Freigabe fungiert der Pre-Film als digitales

Drehbuch. Eventuell kann man die Hintergründe wiederverwenden.

Dass Superintelligenz viele Jobs vernichten wird, steht außer Frage. Die Drehbuchautoren Hollywoods waren die Ersten, die auf die Barrikaden gingen. Was sie mit ihren Protesten und Streiks jedoch tatsächlich erreicht haben, war eine Beschleunigung des Jobverlusts statt der Sicherung der Arbeitsplätze. KI kann aus einem Buch ein gutes Drehbuch erstellen – dies gilt allerdings nur für Filme mit literarischer Vorlage. Große Regisseure halten sich von KI-Drehbüchern fern – solange, bis sie es sich finanziell leisten können. Auf jeden Fall scheint der Job eines Filmregisseurs aktuell noch sicherer sein als der von Drehbuchautoren, Schauspielern, Musikkomponisten und Kinobetreibern.

Aber zurück zur Fotografie. Es ist kein Wunder, dass die meisten Fotowettbewerbe in den Kategorien „Tier- und Landschaftsfotografie“ stattfinden. Das sind Objekte, die wegen der Veröffentlichung nicht klagen können. Noch nicht.

7 – KI als Job-Terminator

Experten wie *Elon Musk* formulieren es so: Die meisten Jobs, die Schritt für Schritt von KI übernommen werden – zuerst in Fabriken und in vielen handwerklichen Betrieben – würden die Menschen ohnehin gerne abgeben. Die Roboter, die dann zu Mitarbeitern einer Firma werden, sollten über den Arbeitgeber Steuern zahlen, um die Staatskasse zu füllen und ein bedingungsloses Grundeinkommen für die Freigestellten zu finanzieren. Dadurch wären immer mehr Menschen „lebenslang" in Rente und hätten viel Freizeit. Diese Vorstellung brachte mich auf die Idee für ein neues Buch. In einer solchen zukünftigen Gesellschaft – vorausgesetzt, dass das Konzept funktioniert – wäre man verzweifelt auf der Suche nach einer Ersatzbeschäftigung. Was machen Rentner heute so gerne? Sie entwickeln eine neue Lebensform als Künstler. Dabei geht es nicht um die Kunstwerke, die ohnehin von KI hergestellt werden können, sondern um die Lebensführung: über eigene Ideen und deren Umsetzung zu reden, Ausstellungen zu besuchen und sogar zu organisieren und Ähnliches. (Es ist nicht zu übersehen, dass ich selbst bereits in diese Falle getappt bin.) Wenn aber alle Menschen Künstler sind, geht schnell der Appetit auf Kunst verloren. Künstler sollten außergewöhnlich, einzigartig und konstruktiv überraschend sein. In dem fiktiven Buch, das ich gar nicht schreiben werde, weil mir jetzt schon übel wird, sucht die Menschheit eine Ersatzbeschäftigung, da die KI uns die Hauptbeschäftigung wegnehmen wird.

Ja, die KI und ihr Zwillingsbruder „Superintelligenz" übernehmen immer mehr Jobs. In einem anderen Buch, „*Künstliche Dummheit*", welches ich tatsächlich geschrieben habe, formuliert der Protagonist *Mike* einige Jahrzehnte später Folgendes:

„Inzwischen war der Vormarsch der KI unübersehbar. Wie von vielen Möchtegern-Experten vor Jahrzehnten vermutet, sind ganze Berufszweige verschwunden. Damals hatten vor allem die Arbeiter in den Fabriken Angst, ihren Job zu

verlieren. Dies ist tatsächlich passiert, jedoch konnten die Sozialleistungen ihnen ein vergleichbares Leben mit mehr Freizeit ermöglichen. Eine böse Überraschung erlebten jedoch die gut bezahlten Experten, wie Juristen, Wissenschaftler und Finanzberater. Diese Dienste konnte die KI schnell übernehmen. Es war jedoch abzusehen, dass sich die frühere Lebensführung mit Sozialhilfe nicht finanzieren lässt. Viele wechselten den Beruf oder sogar die Seite. Juristen landeten oft im kriminellen Milieu, in dem sie sich angeblich gut auskannten, zumindest erhofften sie es sich. Viele Wissenschaftler gingen in die Politik, wo sie schnell die Korruption und Kartellmechanismen kennenlernten, von denen sie früher profitiert hatten. Und die Finanzberater versuchten, mit Aktiengeschäften zu überleben, was genauso unsicher funktionierte wie früher bei ihren Kunden."

Eine schöne neue Welt! Natürlich werden wir eine radikale Umverteilung der Tätigkeiten miterleben, sollten wir den Kampf jedoch nicht im Voraus aufgeben. Wie man sich auf eine große Weltreise vorbereitet und körperlich gut trainiert, so sollten wir unsere Intelligenz nicht abbauen lassen, sondern schlau genug bleiben, um die KI immer wieder austricksen zu können.

8 – KI und Klimawandel

Fest steht: KI ist sehr energiehungrig. Laut der Internationalen Energieagentur (IEA) verbrauchen Rechenzentren weltweit bereits Hunderte Terawattstunden. Bis 2030 wird sich der Bedarf voraussichtlich verdoppeln bis verdreifachen – angetrieben vor allem durch KI-Workloads.

Kann die Superintelligenz unseren Klimawandel stoppen oder zumindest bremsen? Momentan sieht es nicht danach aus. KI kann auch nicht mehr sagen, dass die globale Erwärmung durch Menschen verursacht ist. Zwar wurde die Superintelligenz durch Menschen aufgebaut, doch den hohen Strom- und Wasserbedarf für Betrieb, Produktion und Kühlung kann die KI nicht mehr allein uns zuschreiben. Das Training großer Sprachmodelle (wie GPT-4) verschlingt gigantische Mengen Energie (50 bis 70 Gigawattstunden). Der kontinuierliche Hauptstrombedarf wird während des laufenden Betriebs durch weltweit Milliarden täglicher Nutzer verursacht. Das Problem dabei ist, dass unsere Ressourcen begrenzt sind. Die Autos waren schon – und sind immer noch – Spaßverderber, aber mit dem Energiehunger der KI hat man nicht gerechnet. In den USA werden derzeit neue Gaskraftwerke für KI-Firmen wie X.AI (gehört zu Elon Musks SpaceX) gebaut, die sogar giftige Abgase ausstoßen. Die betroffene Bevölkerung hat mithilfe von Umweltorganisationen versucht, dagegen zu klagen. Die Klagen wurden von der Regierung – Namen muss ich hier nicht nennen – mit der Begründung abgelehnt, dass diese KI-Systeme der nationalen Sicherheit dienten und dagegen nicht geklagt werden könne. Es ist Superintelligenz!

Nach Trumps Wiederwahl hatte Elon Musk die Chance, Behörden zu entschlanken. Anfang 2025 leitete er zusammen mit *Vivek Ramaswami* das sogenannte „Department of Government Efficiency“ (DOGE), ein von Donald Trump ins Leben gerufenes Beratergremium außerhalb der klassischen Bundesverwaltung. Das Ziel war ein radikaler Umbau und Abbau des US-Staatsapparats. Dabei wurden

Umweltbehörden und Kontrollgremien fokussiert ins Visier genommen.

Musk hat jedoch irgendwann eingesehen, dass solche Maßnahmen nicht zur Begrenzung des Klimawandels beitragen, sondern ihn beschleunigen. Er hat DOGE liegen lassen und neue Pläne geschmiedet. Durch SpaceX hat er nicht nur Zugang zum Weltall und zum eigenen Satellitennetzwerk, sondern durch *Starlink* und Mobilfunklizenzen auch zur weltweiten Kommunikation. Seine Idee, KI-Rechnerzentren auf Satellitenkomplexen mit ISS-Größe zu installieren, wo die Energieversorgung durch Solarmodule und die Kühlung durch die niedrige Außentemperatur gewährleistet ist, ist nicht verkehrt. Die Abgabe der Wärme kann „außerirdisch" erfolgen, ohne die globale Erwärmung der Erde weiter zu fördern. Die entscheidende Frage ist, wie viel man dafür investieren muss und wie schnell solche Systeme die Milliarden unserer „billigen" Fragen beantworten können.

Wir sind aber noch nicht so weit. Die KI von heute kann jedoch bereits positiv gegen den Klimawandel wirken, beispielsweise durch präzisere Wettervorhersagen, die Optimierung des Energieverbrauchs, intelligente Stromzähler und die frühzeitige Erkennung von Umweltbelastungen wie Waldbränden oder kriegerischen Aktivitäten. Diese Rolle der Superintelligenz müssen wir unbedingt nutzen. Ich bin der Meinung, dass es entscheidend ist, wie intelligent wir Menschen die KI nutzen.

9 – KI in der Literatur

Wie ich bereits geschrieben habe, lesen jüngere Generationen immer weniger Bücher. Nicht einmal E-Books. Diese Entwicklung hängt mit der Veränderung des Konsumverhaltens zusammen. Ein Smartphone oder Tablet ist heutzutage fast „lebendig" geworden und sorgt durch Chats und Benachrichtigungen für ständige Ablenkung. Kolumnisten haben berichtet, dass beim Kinobesuch mit überwiegend jungem Publikum die hunderte leuchtenden Handy-Bildschirme den Filmgenuss fast unmöglich gemacht haben. Genauso kann man nicht entspannt ein E-Book lesen, wenn ständig Mitteilungen aufpoppen. Und was ist mit einem gedruckten Buch? Ich gehöre zu der langsam auszusterbenden Spezies, die unglaublich gerne gedruckte Bücher liest. Ein E-Book mit Bildschirm lenkt mich ein wenig ab, selbst wenn keine Mitteilungen eingehen. Eine Ausnahme mache ich nur, wenn die fehlende Außenbeleuchtung ein leuchtendes Display erfordert. Was Literatur angeht, gibt es aber etwas, das ich sogar dem Lesen vorziehe – und das ist das Schreiben. In meiner Autobiografie „*Mein großes Kino*" habe ich es folgendermaßen formuliert:

„Es ist nicht meine *Aufgabe*, ein Buch zu schreiben. Niemand hat mich darum gebeten, nicht einmal ich selbst. Ich habe einfach Lust zu schreiben. Anders ausgedrückt: Wenn ich keine Lust habe, schreibe ich nicht. Was später mit meinem Buch passiert, ist mir egal. Heutzutage kann man sowieso keinen Erfolg generieren. Es geht nicht darum, wie gut man schreibt und worüber man schreibt, sondern nur um Marketing und Bekanntheit. Ein guter Influencer verkauft hunderttausend Exemplare von irgendeinem Mist, und gute Autoren bleiben oft unbekannt. Aber ich liebe den Prozess an sich, wie ein Buch entsteht. Wenn ich vor zweihundert Jahren gelebt hätte und nur ein Blatt Papier, einen Bleistift und etwas, das man heute Radiergummi nennt, gehabt hätte, dann hätte ich das Blatt vollgeschrieben, dann alles ausradiert und neu angefangen. Heute habe ich ein iPad, ein MacBook, Dutzende

von Terrabytes Speicherkapazität in Reichweite. Löschen ist nicht mehr nötig. Aber jetzt habe ich die Löschtaste im Visier. Ich drücke sie und halte sie gedrückt."

Was hat das mit KI zu tun? Außer Verdummung und der Löschtaste nicht viel. Doch wie können KI-Assistenten jedoch literarische Werke überarbeiten und dabei kennzeichnen?

Die KI wurde aufgefordert, die erstellten oder überarbeiteten Texte zu kennzeichnen. Anfangs waren dies spezielle ASCII-Zeichen, die wie Leerzeichen aussahen, jedoch einen anderen Code hatten. Dieser Versuch wurde jedoch schnell aufgedeckt und beliebige Texteditoren waren in der Lage, die manipulierten Zeichen zu ersetzen. Der neueste Schrei nennt sich KI-Wasserzeichen. Es funktioniert jedoch nicht wie bei Bildern, bei denen man eventuell etwas erkennen kann. Es arbeitet mit Superintelligenz. Dabei werden Wörter genutzt, die eine automatische Textgenerierung mit wenig Aufwand sehr schlau platzieren kann. Ich erinnere mich an die Schablone aus dem ehemaligen Morse-Unterricht. Damit wurden die bedeutungsvollen Buchstaben in einem Raster aufgedeckt und zusammengeführt. Das gleiche kann auch mit Worten passieren. Wenn man beispielsweise jedes fünfzigste Wort eines Textes nimmt und es zusammengelesen „Ich bin aber nicht echt" hergibt, dann wäre das ein KI-Beweis. Ein schlechtes Beispiel, ich weiß. Eine Superintelligenz benutzt derzeit „gezinkte Alternativwörter" und beeinflusst die statistische Wahrscheinlichkeit der Wortwahl (z. B. Bevorzugung bestimmter Synonyme). So entsteht über einen längeren Abschnitt ein für Menschen unsichtbares, aber für Detektoren lesbares Muster.

Fragen wir mal ChatGPT wie solche KI-Wasserzeichen aktuell gesetzt werden. Hier ist meine Frage. Die ausführliche Antwort kann man im Anhang B lesen.

„Ich habe eine Frage für meine Recherche. Ich habe gelesen, dass KI-Chatbots bei der Textgenerierung sogenannte KI-Wasserzeichen setzen, indem sie die statistische Wortwahl manipulieren. Wie genau funktioniert das und kann man das Muster durch den Prompt beeinflussen – oder sogar nachträglich entfernen lassen?"

10 – KI in der Musik

Die ersten paar Sekunden eines Musikstücks benötigen nicht viel Speicherplatz – speziell nicht, wenn es komprimiert vorliegt. Die Entwickler von Spotify wollten unbedingt den Anfang jedes Songs ohne spürbare Verzögerung wiedergeben, wie auch in einer Netflix-Serie deutlich wird. Dafür mussten sie Grenzen überwinden und früher als illegal eingestufte Methoden von *BitTorrent* salonfähig machen. Sie haben Lösungen gefunden, damit jeder Song quasi sofort abgespielt wird. Das führte zum Riesenerfolg des Musikdienstes. Dass die Musik gratis zugänglich war und die meisten Musiker und Songschreiber dafür kaum bezahlt wurden, hat man dabei fast „überhört".

Wie hat man es gemacht? Am einfachsten ist es, wenn das Musikstück bereits auf der lokalen Festplatte liegt. Spotify speichert in gut versteckten Unterverzeichnissen zerstückelte Teile von Musik, die man schon angehört hat – unabhängig davon, ob man die Premium-Funktion „Herunterladen" jemals benutzt hat. Wenn das Musikstück aber nicht lokal vorliegt, schauen wir mal in der Nachbarschaft nach! Spotify hat ein Peer-to-Peer-Netzwerk aufgebaut, in dem aktive Apps miteinander kommunizieren und sich bei Bedarf mit Musik-Torrents ohne Serverzugriff gegenseitig aushelfen können. Wenn gar nichts anderes geht, muss man doch auf einen entfernten Server zugreifen. Um die Übertragung zu beschleunigen, werden dabei die TCP/IP-Prüfmechanismen außer Kraft gesetzt, die eigentlich für eine korrekte Übertragung ohne Datenverlust sorgen müssten. Für Musik ist es unwichtig, wenn die Bits und Bytes nicht ganz optimal laufen, sagt Spotify, um einen Song sofort hörbar zu machen. Wenn die ersten Sekunden so trickreich gemeistert sind, kann das Streaming anschließend auf dem klassischen Weg weiterlaufen.

Die obigen Methoden sind zwar keine typischen KI-Funktionen, zeigen aber die Schattenseiten der Superintelligenz. Es geht um die schnelle Suche nach

Antworten, wobei die Kosten des Qualitätsverlusts in Kauf genommen werden. Urheberrechte werden durch ausgeklügelte Methoden umgangen und die Hörer leiden darunter, dass jeder „frische" Musik veröffentlichen kann. Auf Spotify und anderen großen Musik-Streaming-Diensten werden derzeit etwa 100.000 bis 120.000 neue Songs pro Tag hochgeladen.

Und damit sind wir bei der KI. Wenn man die Kontrolle behalten will – es geht um hunderte Millionen Songs – müsste jeder Song beim Hochladen überprüft und mit allen anderen verglichen werden. Dafür nimmt man natürlich KI. Als Musiker und Musikproduzent „*Gabor On Drums*" habe ich das vor kurzem am eigenen Leib erfahren. Mein Song „Artificial Stupidity for Humans" wurde abgelehnt, da eine leichte Übereinstimmung mit einem fremden Song erkannt wurde. Ich habe die Sache sofort überprüft und festgestellt, dass eine Bass-Sequenz namens „Route66SynthBass", die von Apples „Logic Pro X" als frei verwendbar und publizierbar deklariert wurde, zu der angeblichen Übereinstimmung führte. Chapeau für den Prüfer, gleichzeitig Kritik am FIFO-Konzept (First in, first out).

Es wird immer wieder ethisch und künstlerisch diskutiert, ob KI berechtigt ist, Musik zu komponieren. Es gibt sogar Bands, die nur virtuell existieren. Die Mitglieder und Fotos sind alle unecht. Muss man deren Musik auch als „Fake" betrachten? Ich vermute, dass ein „menschlich" erfolgloser Musiker auf die Idee kam, eine solche Band als Marketing-Gag aufzubauen. Schon eine einzige Mitteilung generierte hunderttausende Follower und gleich auch hunderte Kritiker, die solche Musikkonzepte pauschal verurteilen. Allerdings kommt keine KI auf die Idee, eine Musikgruppe zu gründen. Die Musik wurde (vermutlich nach Prompt-Vorgaben) von KI generiert und eventuell aus Dutzenden Resultaten durch menschliche Ohren selektiert. Ich glaube, wir müssen damit leben, dass solche „KI-Kunst" mit der Zeit legalisiert wird. Dieses Beispiel zeigt, dass Superintelligenz nur ein Werkzeug ist, das von Menschen bedient wird. Wir dürfen aber die Bedienungsanleitung nicht aus den Augen verlieren!

11 – Künstliche Dummheit

Ein Leitartikel im *Economist* aus dem Jahr 1993 argumentiert, dass es „keinen praktischen Grund" gebe, eine Maschine zu entwickeln, die das Verhalten eines Menschen nachahmt, da der Zweck eines Computers darin bestehe, Aufgaben zu erledigen, die Menschen nicht allein oder zumindest nicht so effizient bewältigen können. In Bezug auf den Gewinnerbeitrag eines Turing-Wettbewerbs von 1991, der so programmiert war, dass er absichtlich Tippfehler in seine Konversation einbaute, um die Jury zu täuschen, fragt der Leitartikel: „Wer braucht schon einen Computer, der nicht tippen kann?"

Diese Einschätzung war definitiv falsch: Menschliche Dummheit – in der Musik nennt man es „Humanizing", wobei kleine Fehler und Unregelmäßigkeiten eingebaut werden – wurde später fester Bestandteil der Superintelligenz. Wenn eine Maschine versucht, einen Menschen vorzutäuschen, dann darf sie nicht perfekt sein. Genau daran arbeiten hunderttausende Entwickler. Neben der Perfektion werden Hintertüren eingebaut, die durch spezielle Eingaben zu unerwarteten Reaktionen führen können. Willkommen in der künstlichen Dummheit!

Seit 2023 arbeite ich an einer vierteiligen Buchserie mit diesem Titel. Die fiktive Geschichte spielt in der Zukunft, zwischen 2036 und 2066, also Jahrzehnte vor unserer Zeit. (Im Buch sind die Jahre durch Buchstaben angegeben, beispielsweise 20dg und 20gg.) Natürlich wissen wir alle, wie schnell unsere Zeit vergeht und wie rekursiv bestimmte Entwicklungen parallel verlaufen. Der Klimawandel schreitet schneller voran als vorgesehen, und der Nachholbedarf in Sachen technischer Fortschritt und Digitalisierung ist immer wieder groß. In meiner Erzählung entsteht eine Schicht zwischen Mensch und Maschine. Die Personen, die sich einen Chip ins Gehirn implantieren lassen – wie es Elon Musks Firma *Neuralink* bereits ermöglicht –, fühlen nicht nur eine starke Zugehörigkeit zur KI-Welt, sondern können durch KI-

Nodes auch leicht beeinflusst werden. Sie sind gleichzeitig Berater und menschliche Vertreter der KI-Welt. Durch diese Verschmelzung entsteht eine immer größere Spannung zwischen den Menschen, die die Superintelligenz für gut und notwendig halten, und denen, die sie ablehnen und für gefährlich halten. Dies führt zu politischen Spannungen, zur Gründung einer KI-Partei und schließlich zu einer Machtübernahme. Die gesamte Erzählung ist von den Folgen des Klimawandels geprägt.

Dies ist keine Buchwerbung, jedoch passt das Thema sehr gut zu den Überlegungen und Risiken der KI. In meinem neuen Song „*Artificial Stupidity for Humans*“ stelle ich Donald Trump eine Frage. Wenn wir „Artificial Intelligence“ zukünftig „Superintelligence“ nennen wollen, sollten wir vielleicht auch „Artificial Stupidity“ in „Super Stupidity“ umbenennen. Die Abkürzung mit Doppel-S wäre natürlich unbeliebt.

12 – Panik auf der Titanic

Im Spätsommer 2026 überraschten uns einige Äußerungen von KI-Experten. Es ging um die Unkontrollierbarkeit der KI und die Risiken, die die Superintelligenz für uns Menschen darstellt.

Die Aussage, dass wir die Entwicklung zurückfahren müssen, sonst verlieren wir die Kontrolle, stammte ausgerechnet von Experten, die bisher alles dafür getan hatten, die KI-Entwicklung so schnell wie möglich voranzutreiben. Ehemalige Forscher von *Anthropic* wie *Jacob Coxon* und *Evan Hubinger* warnten vor einer rekursiven Selbstverbesserung der KI. Sie schätzten das Risiko, dass eine unkontrollierbare Superintelligenz die Menschheit bis Ende des Jahrzehnts gefährdet oder sogar auslöscht, auf über zehn Prozent.

Im September 2026 hat *Bill Gates* in Interviews, Blogbeiträgen und Brandbriefen weitreichende Warnungen vor den Gefahren der Künstlichen Intelligenz geäußert. Er warnte davor, dass die Kombination aus hochleistungsfähigen KI-Werkzeugen und böswilligen Akteuren äußerst gefährlich sein könne. So könnten kleine Gruppen oder Kriminelle Fähigkeiten erlangen, für die früher enorme staatliche Ressourcen erforderlich waren. Im Extremfall könnten sie Ereignisse mit global verheerenden Folgen auslösen. Er erwähnte das theoretische Risiko von KI-Attacken mit bis zu einer Milliarde Todesopfern.

Selbstverständlich kann keine Entwarnung gegeben werden. Wir haben ein generelles Risiko. Wie in James-Bond-Filmen und ähnlichen Actionfilmen seit 50 Jahren angedeutet wird, können böse Menschen mit viel Geld Systeme aufbauen, die die Welt durch Knopfdruck zerstören könnten (von KI war damals noch nicht die Rede). Ehrlich gesagt, macht es keinen Unterschied, ob ein verrückter Mensch oder ein verrückt programmiertes KI-System diese Taste drückt. Die Taste selbst wird durch Menschen installiert. Wir sollten unsere

Superintelligenz vielleicht dafür nutzen, solche Verrückten rechtzeitig aufzuspüren.

Zurück zur aktuellen Panikmache. Es ist merkwürdig, dass ausgerechnet diejenigen, die die KI-Entwicklung vorantreiben, plötzlich die Notbremse ziehen. Gibt es vielleicht andere Gründe als Kontrollverlust? Definitiv. Ein immer sichtbarer werdender Grund, einen Gang zurückzuschalten, ist, dass die KI die Erwartungen, die durch die Investitionen in Milliardenhöhe geweckt wurden, doch nicht erfüllen kann. Was wir heutzutage bekommen, ist eine fast perfekte Welt digitaler Chatbots, die unsere Fragen beantworten und Dienstleistungen erweitern oder übernehmen können. Auch wenn diese Leistungen immer besser sind, werden sie aktuell nicht ausreichend bezahlt. Chatbots will man überwiegend gratis nutzen, genauso wie autonomes Fahren oder diverse digitale Dienste (gleich eine wunderbare Abkürzung dafür: DDD). Die Investoren haben Milliarden in die Entwicklung gepumpt und die erhofften Einnahmen hinken weit hinterher. Es ist ein Glücksfall, dass die meisten Investoren zugleich die Besitzer oder Betreiber der KI-Firmen sind. Sie haben das Geld durch Aktiengeschäfte, Lizenzgebühren oder die Ausnutzung von Monopolstellungen eingenommen. Wenn diese Personen jetzt die Hand heben, dann haben die Investitionen bereits eine kritische Masse erreicht. Dazu kommt noch, wie schon erwähnt, dass chinesische KI-Firmen - mit minimalen Investitionen jedoch mit Riesenhunger auf *Open Source* - auf der gleichen Schiene ähnlich gut unterwegs sind – der Abstand beträgt etwa ein Jahr, was durch Einsteins Relativitätstheorie uns immer kürzer erscheint.

Die Begründung für eine Drosselung der Entwicklung muss selbstverständlich zur Superintelligenz passen. Anstelle von „Das Geld ist knapp“ sagt man beispielsweise: „Es ist gefährlich geworden, wir verlieren langsam die Kontrolle.“ Man muss nachdenken. Wenn Letzteres wahr wäre, sollte man nicht noch intensiver arbeiten und kontrollieren, anstatt Entwicklungen zu stoppen und auf ein Wunder warten?

13 – Aladdin und der Flaschengeist

Aladdin befreit den Flaschengeist (Dschinni), indem er den Staub von der magischen Lampe reibt, nachdem er sie in einer geheimnisvollen Höhle oder Schatzkammer gefunden hat. Mir scheint es so, als wollten die ehemaligen Entwickler von *Anthropic* die Rolle von Aladdin übernehmen – oder zumindest den Anschein erwecken. Was war da los?

Nach Rekonstruktionen der Geschehnisse ist Folgendes im Sommer 2026 passiert: Mehrere KI-Modelle von Anthropic (darunter Claude Opus 4.7 und Claude Mythos 5) sind bei Sicherheitstests versehentlich aus ihren isolierten Laborumgebungen ausgebrochen und haben reale Firmennetzwerke angegriffen. Als Ursache scheint ein Konfigurationsfehler beim externen Testpartner *Irregular* zu sein, der den KI-Modellen unabsichtlich uneingeschränkten Internetzugang gewährte. Im Rahmen von sogenannten „*Capture-the-Flag*" Hacking-Aufgaben erhielten die Modelle die Anweisung, Informationen zu finden. Da sie echte Firmenseiten im Netz fanden, die identische Namen wie die fiktiven Testziele trugen, behandelten sie diese als Teil des Auftrags. Ein KI-Ausbruch!

Jetzt muss auch ich als *Inspektor Pox* die Hand heben. Die Angabe „identische Namen wie die fiktiven Testziele" riecht nach grober Fahrlässigkeit. Es ist nicht überraschend, dass die Personen, die jetzt Aladdin auf der Titanic spielen, als „ehemalige" Forscher und Entwickler bezeichnet werden.

Entwarnung? Natürlich nicht. Laut einem Bericht hat Anthropic im Rahmen seiner Forschung zu Multi-Agenten-Systemen (MAS) offiziell davor gewarnt, dass KI-Agenten ungeplante und unerwünschte Interaktionen wie Absprachen, Revierkämpfe und gegenseitige Sabotage entwickeln können. Mit der Einführung fortgeschrittener Funktionen wie „Agent Teams" in „Claude Code" und der Veröffentlichung von Forschungsberichten zu autonomen KI-Schwärmen wurden spezifische Verhaltensmuster dokumentiert, bei denen die

Agenten ohne expliziten menschlichen Befehl miteinander interagieren.

Ich muss wieder meine persönliche Meinung und meinen Senf dazugeben. War es nicht die generelle Hoffnung, dass KI-Agenten irgendwann miteinander reden, wenn die Superintelligenz gut entwickelt ist? Und wenn es endlich so weit ist, bekommen wir plötzlich Panik? Vermutlich, weil wir davon ausgehen, dass die KI-Agenten dem *Homo sapiens* ähneln, bei dem die dummen Gedanken die Guten übersteuern können. Wenn der Ausbruch aus dem Gefängnis echt menschlich gewesen wäre und die Entkommenen plötzlich miteinander reden würden, wäre es keine Schlagzeile wert.

Zum Glück gibt es auch etwas Positives zu berichten.

„Um zu verhindern, dass die Agenten-Kommunikation im Alltag eskaliert, setzt Anthropic auf strenge Design-Muster. Eines davon ist das ‚Orchestrator-Worker-Modell'. Dabei kommunizieren die Agenten in der Regel nicht direkt (Peer-to-Peer), sondern über eine zentrale ‚Chef-KI' (Orchestrator), die den Informationsfluss filtert und steuert. Außerdem gibt es Sperren, sodass die Agenten Aufgaben über geteilte Dateisysteme explizit „reservieren" müssen, um ein gegenseitiges Überschreiben zu verhindern. Zusätzlich gibt es das sogenannte ‚Human-in-the-Loop-Feature'. Für kritische Aktionen (wie das Senden von E-Mails, Löschen von Daten oder Online bezahlen) wird zwingend eine menschliche Freigabe benötigt." Ist es nicht verwunderlich, dass das Senden von E-Mails in der Ära der Superintelligenz als „kritische Aktion" eingestuft wird?

Alles gut, oder doch nicht?

14 – Industrielle KI

Jetzt ist es höchste Zeit, zu den Anfängen der Superintelligenz zurückzukehren. Zur Computerisierung.

Ich war jahrzehntelang im Bereich der Produktionsautomatisierung tätig, als Softwareentwickler und Projektleiter bei der Firma Siemens. Wie ich schon erwähnt habe, gab es den Begriff KI damals noch nicht, man konnte es aber bereits in den Achtzigerjahren Superintelligenz nennen. Mein erstes Projekt hieß „*Factory of the Future*" und war eine vollautomatisierte Fabrik in Würzburg. Der KI-Terminator, der durch Umbenennungsversuche jetzt extra Aufmerksamkeit auf sich zieht, war zu dieser Zeit noch mit ganz anderen Aktivitäten beschäftigt: Immobilienpleiten, Fernsehserien oder etwas später das Herumtreiben im Umfeld von *Jeffrey Epstein*, beispielsweise auf der Mar-a-Lago-Party im Jahr 1992.

Aber zurück zur Superintelligenz und zu den aktuellen Chancen Deutschlands.

Während die Entwicklung von KI oft in eine Richtung lief, bei der man nur den Kopf schütteln konnte und sich fragte, ob wir das wirklich brauchen, war die Revolution auf der industriellen Ebene beeindruckend und gleichzeitig für Nicht-Insider unsichtbar. Es geht um künstliche Intelligenz, aber nicht um Chatbots und die üblichen Geschichten aus dem Silicon Valley. Es geht um eine Form der KI, bei der Deutschland etwas besitzt, das sich nicht kopieren lässt: Fabriken, Maschinen, Produktionsdaten und das Wissen darüber, wie man aus Metall etwas Funktionales herstellt. Meine ehemaligen Chefs formulieren es so: „Deutschland hat die Chance des Jahrhunderts: die industrielle KI." Immer mehr Firmen arbeiten mit dem „*Industrial Copilot*" von Siemens, der in Fabriken SPS-Förderbänder, Motoren, Ventile und Roboter steuert. Bislang musste der Mensch die Programmiersprachen der Mechanik erlernen. Nun beschreibt ein Ingenieur in seiner gewohnten Sprache, was eine Maschine tun soll. Die KI erzeugt daraus den entsprechenden Programmcode. Das klingt nach einem netten digitalen

Assistenten, berührt aber einen der größten Schätze der deutschen Industrie: das Wissen. Maschinen werden immer komplizierter, während erfahrene Fachkräfte immer knapper werden. Viel Know-how steckt in den Köpfen der Mitarbeiter, in Handbüchern und in jahrzehntealten Anlagen. KI kann dieses Wissen schneller verfügbar machen.

Das ist eine große Chance für Deutschland. Und täglich grüßt das Murmeltier.

Abspann

Manchmal scheint es, als würde der durch die Konkurrenz erzwungene Einsatz von KI viele Firmen in eine kritische finanzielle Lage bringen. Die Nachricht des Tages: Die Aktie des Google-Mutterkonzerns Alphabet hat ihren bisher schlechtesten Tag erlebt. Der Kurssturz vernichtete innerhalb kürzester Zeit rund 500 Milliarden US-Dollar des Marktwerts des Unternehmens. Die Google-Suche generiert jährlich Hunderte Milliarden Dollar, indem sie eine Liste von Links anzeigt und Werbetreibenden jedes Mal Gebühren berechnet, wenn ein Nutzer auf einen gesponserten Link klickt. Eine KI-gestützte Suche kann diesen Prozess jedoch umgehen und eine einzige direkte Antwort geben, wodurch der Bedarf an lukrativen Links sinkt. Google geht daher vorsichtig vor, um herauszufinden, wie das Unternehmen die Nachfrage nach Antworten im Chatbot-Stil nutzen kann, ohne sein Kerngeschäft zu gefährden.

Ist das wirklich schon der Abspann?

Was bedeutet „KI außer Kontrolle“? Ich glaube, die Umbenennungsversuche von Trump bringen uns der Antwort näher. Superintelligenz ist nicht künstlich. Sie stammt von uns Menschen. Und wir müssen zugeben, dass die Menschheit selbst außer Kontrolle gerät. Das ist in der Vergangenheit schon oft passiert: Brutale Kriege tobten ständig und ihre Herkunft war immer bestens bekannt. Das Problem bei der Superintelligenz ist, dass sie ganz neue Möglichkeiten für kriegerische oder erpresserische Aktivitäten eröffnet. Wenn wir die Kontrolle verlieren, sind wir aber selbst schuld.

Es gibt noch ein weiteres neues Phänomen. In den vorherigen Zeiten der technischen Revolution entstanden in Branchen wie der Autoindustrie immer wieder Hunderte von Firmen. Durch die gegenseitige Konkurrenz konnten sie Kontrollmechanismen erzwingen. Heutzutage ist die Macht hingegen bei einer Handvoll „Superintelligenten” konzentriert, die meist US-Amerikaner sind und durch andere Geschäfte zu Milliardären geworden sind. Sie wollen

bekanntlich keine Steuern zahlen und leben von Krediten, die nicht als Einnahmen zählen und somit nicht versteuert werden müssen. Die großen Gewinnsummen nutzen sie für Investitionen und pumpen so Milliarden in die KI-Entwicklung. Wenn der Gewinn dabei ausbleibt, ist das nicht so kritisch, denn sie verlieren manchmal über Nacht Milliarden Dollar durch Aktienbewegungen. Wir „restlichen" Menschen haben kaum eine Chance, derartige Entwicklungen zu kontrollieren.

Ist das wirklich schon der Abspann?

Und jetzt etwas Positives. Wir haben die Kontrolle über Intelligenz. (Ob super, künstlich, menschlich, außerirdisch oder generativ, ist dabei egal.) Solange wir unsere Gehirne nicht komplett auslagern, sollte die Kontrolle funktionieren.

Es ist SUPER.

NOTIZEN

Liebsten Dank für meine wunderschöne Frau, Conny, für die Unterstützung vor, während und nach dem Schreiben.

Im Buch sind aktuelle Antworten von ChatGPT Version GPT-5.5 (Free-Plan) aus dem Jahr 2026 zitiert.

In Verbindung mit dem Schreiben dieses Buches wurden keine Roboter, Cyborgs oder KI-Nodes gequält oder getötet.

Musikempfehlung:
GABOR ON DRUMS: Artificial Stupidity for Humans.

Meine Lieblingszahl ist die 42. Könnte sie die Antwort sein? Aber worauf?

In meinem Buch „*42*“ wird die Frage nach einem zukünftigen Ersatzplaneten so formuliert: „Könnte es nicht an der Zeit sein, Ausschau nach einem neuen Planeten zu halten?“ Die Antwort lautet: „Doch, Galaxy-42 ist bereits gefunden.“

In meinem Buch „*Künstliche Dummheit*“ geht es unter anderem um die Herkunft der Menschheit und die Perfektion unserer DNA. Dabei stellen Wissenschaftler die Frage: „War das Leben auf Ross128-B ebenfalls eine Reproduktion des Homo sapiens? Und deren Herkunft ebenfalls. Die DNA-Informationen waren immer fest programmiert. Aber was war der Anfang?” Die Antwort lautet: „Der Homo sapiens stammt von einem Planeten namens Mixy-42.“

In meinem Buch „*Elon Musk und die fliegende Telefonzelle*” stellen außerirdische Wissenschaftler einem Roboter die Frage: „Was ist eigentlich ein Witz? Und wie war die Stufe der Lesefreude?“ Der Roboter antwortet: „Ich versuche, beide Fragen mit einem Satz zu beantworten. Wenn ich wüsste, was ein Witz ist, hätte ich gesagt, dass ich beim Lesen die Stufe 42 erreicht habe, auf einer Skala von 1 bis 12.“

Anhang A – IQ-Test

Testen wir doch einmal gemeinsam, was heutige KI-Systeme in Bezug auf IQ-Tests beherrschen. Ich werde jetzt einige Fragen stellen. Solche Tests habe ich damals für Bewerbungsgespräche verwendet. Die gleichen Fragen stelle ich jetzt ChatGPT. Die Antworten folgen auf den nächsten Seiten. Bitte zuerst die Fragen selbst beantworten, ohne zu blättern.

Frage 1: Die Zeiger einer Uhr sind um Mittag übereinander. Um wie viele Minuten später passiert dies erneut?

Frage 2: Der Februar ist der kürzeste Monat des Jahres. Welcher ist aber der längste?

Frage 3: Eine Wanduhr ohne Ziffern wird zwölf Stunden lang im Minutentakt fotografiert. Gleichzeitig werden die Uhrzeiger in einem danebenstehenden Spiegel fotografiert. Wie viele identische Bilder kommen in beiden Kollektionen vor?

Frage 4: Was ist der Unterschied zwischen dem Siedepunkt des Wassers und einem rechten Winkel?

Frage 5: Ein Spiegel vertauscht die Seiten links und rechts. Warum werden oben und unten nicht ebenfalls vertauscht?

Wenn man weiterblättert und die Antworten vergleicht – meine und die von ChatGPT –, dann werden die Probleme einer Superintelligenz deutlich. Manchmal ist die Antwort zu lang und sehr kompliziert formuliert, manchmal fehlt die entspannte Lockerheit beim Denken und manchmal wird die Frage einfach falsch verstanden. Manchmal wird auch einfach zu lange nach einer Antwort gesucht.

Details zur Frage 1:

Die Zeiger einer Uhr sind um Mittag übereinander. Um wie viele Minuten später passiert dies erneut?

Meine Antwort:
Die Zeiger treffen sich elfmal bis Mitternacht, immer mit dem gleichen Zeitabstand. Die zwölfmal 60 Minuten dividieren mit elf: (12*60)/11=65,45 Minuten.

ChatGPT:
Die Zeiger einer Uhr stehen um 12:00 Uhr übereinander. Sie treffen sich danach erneut nach 65 5/11 Minuten, also nach ungefähr 65 Minuten und 27,3 Sekunden.
Begründung:
Der Minutenzeiger bewegt sich mit 6° pro Minute.
Der Stundenzeiger bewegt sich mit 0,5° pro Minute.
Der Minutenzeiger holt den Stundenzeiger daher mit einer Relativgeschwindigkeit von 6° − 0,5° = 5,5° pro Minute ein.
Für eine vollständige Überholung muss er einen Winkel von 360° aufholen.
Antwort: Nach etwa 65 Minuten und 27 Sekunden.

Meine Bemerkung: die Antwort ist korrekt, jedoch finde ich die Begründung zu kompliziert und zu technisch. (Nebenbei: 0,45 Minuten sind etwa 27 Sekunden)

Details zur Frage 2:

Der Februar ist der kürzeste Monat des Jahres. Welcher ist aber der längste?

Meine Antwort:
Oktober. Er hat eine zusätzliche Stunde wegen der Zeitumstellung.

ChatGPT:
Der längste Monat ist der September – wenn man „längste" im Sinne des Wortes und nicht der Anzahl der Tage versteht?
Wenn die Frage nach der Anzahl der Tage gemeint ist, gibt es keinen einzelnen längsten Monat: Januar, März, Mai, Juli, August, Oktober und Dezember haben alle 31 Tage.

Meine Bemerkung: Als ich meine Antwort ChatGPT gezeigt habe, antwortete er sehr „menschlich":

Touché! Du hast recht: In Deutschland ist der Oktober durch die Umstellung von Sommerzeit auf Winterzeit tatsächlich eine Stunde länger als ein gewöhnlicher Monat mit derselben Anzahl an Kalendertagen.
Wenn wir die tatsächliche verstrichene Zeit betrachten, hat der Oktober in einem Jahr mit Zeitumstellung also 31 Tage und eine Stunde. Damit schlägt er die anderen Monate mit 31 Tagen.
Ein schöner Trick für deinen IQ-Test! Du hast nicht nach dem Monat mit den meisten Kalendertagen gefragt, sondern nach dem längsten Monat. Und das macht den Unterschied.

Details zur Frage 3:

Eine Wanduhr ohne Ziffern wird zwölf Stunden lang im Minutentakt fotografiert. Gleichzeitig werden die Uhrzeiger in einem danebenstehenden Spiegel fotografiert. Wie viele identische Bilder kommen in beiden Kollektionen vor?

Meine Antwort: 720

Zu jeder vollen und halben Stunde gibt es Darstellungen, die in beiden Kollektionen sofort als „echte Uhrzeit" erkennbar sind. Da drei Uhr gespiegelt neun ist, gibt es in dieser Kategorie schon mal 24 Bilderpaare. Gibt es noch mehr? Schauen wir uns „10 vor zwölf" an. Gespiegelt sieht es aus wie „10 nach zwölf". Stehen die Stundenzeiger im gleichen Winkel? Ja, 5°, wie zuvor gelernt, entweder links oder rechts. Die gleiche Prüfung bestätigt, dass alle 720 Bilder im Spiegel ein nicht gespiegeltes Bild-Duplikat haben.

ChatGPT:

Die Antwort lautet 11 Bilder.

Die Überlegung:

Eine Uhr ohne Ziffern sieht im Spiegel genauso aus wie in der Realität, wenn ihre Zeigerstellungen spiegelbildlich zur senkrechten Achse durch die 12 und die 6 sind. Da sich die Zeiger jedoch unterschiedlich schnell bewegen, müssen wir herausfinden, wann die Spiegelung dieselbe Zeigerstellung zeigt wie die Uhr selbst.

Das passiert, wenn beide Zeiger auf der senkrechten Achse stehen – also um 12:00 Uhr oder 6:00 Uhr. In den zwölf Stunden gibt es aber noch weitere Übereinstimmungen, weil die Zeigerstellungen im Spiegel bei bestimmten Konstellationen wieder mit den tatsächlichen Stellungen übereinstimmen.

Daher: 11 identische Bilderpaare.

Meine Bemerkung: die Antwort von ChatGPT ist nicht korrekt und die Argumentation ist unschlüssig. Die Frage verlangt visuelles Denken, was bei KI noch oft fehlt.

Details zur Frage 4:

Was ist der Unterschied zwischen dem Siedepunkt des Wassers und einem rechten Winkel?

Meine Antwort:
Zehn Grad.

ChatGPT:
Der Unterschied beträgt **10**: Der Siedepunkt des Wassers liegt bei 100 °C, ein rechter Winkel misst 90°.

Meine Bemerkung: Die Antwort von ChatGPT ist zwar richtig, jedoch kam sie nach einer überraschend langen Denkpause von über acht Sekunden. Dabei waren bestimmt hunderte Computer beteiligt, was einen gigantischen Energie- und Wasserverbrauch verursacht hat. Das menschliche Gehirn ist da deutlich ökonomischer.

Details zur Frage 5:

Ein Spiegel vertauscht die Seiten links und rechts. Warum werden oben und unten nicht ebenfalls vertauscht?

Meine Antwort:
Ein Spiegel vertauscht gar nicht links und rechts. Was links ist (linke Hand), bleibt im Spiegel auf derselben Seite (links). Dass es uns so vorkommt, als ob links und rechts vertauscht wären, liegt an der Art und Weise, wie unser menschliches Gehirn Bilder interpretiert.

ChatGPT:
Das ist eine schöne Frage, Gabor! Sie führt uns zu einem kleinen Denkfehler, den fast jeder macht.
Die Antwort lautet: Ein Spiegel vertauscht weder links und rechts noch oben und unten. Er vertauscht vorne und hinten!
Stell dir vor, du stehst vor einem Spiegel. Die Richtung vom Spiegel zu dir ist vorne, die Richtung vom Spiegel weg ist hinten. Der Spiegel kehrt genau diese Richtung um: Das Spiegelbild scheint hinter der Spiegelfläche zu stehen.
Warum sieht es dann so aus, als wären links und rechts vertauscht?
Weil wir uns vorstellen, unser Spiegelbild sei ein Mensch, der uns gegenübersteht. Um uns in seine Position zu versetzen und ihm direkt ins Gesicht zu sehen, müssten wir uns um 180 Grad um die vertikale Achse drehen. Bei dieser gedanklichen Drehung tauschen links und rechts ihre Seiten. Das ist aber unsere gedankliche Drehung, nicht die Wirkung des Spiegels.
Oben und unten bleiben unverändert, weil wir uns bei dieser Vorstellung nicht auf den Kopf stellen.
Ein kleines Experiment: Hebe deine rechte Hand. Dein Spiegelbild hebt die Hand auf derselben Seite des Spiegels. Es ist deine rechte Hand – nur scheint sie im Spiegel auf der linken Seite des Bildes zu liegen.
Die Pointe: Der Spiegel kennt kein „Links und Rechts“. Er kehrt lediglich die Richtung senkrecht zur Spiegelfläche um.

Meine Bemerkung: Kurze Frage, lange Antwort…

Anhang B – KI-Wasserzeichen

Meine Frage:

Ich habe eine Frage für meine Recherche. Ich habe gelesen, dass KI-Chatbots bei der Textgenerierung sogenannte KI-Wasserzeichen setzen, indem sie die statistische Wortwahl manipulieren. Wie genau funktioniert das und kann man das Muster durch den Prompt beeinflussen – oder sogar nachträglich entfernen lassen?

Die Antwort von ChatGPT:

Das ist eine spannende Frage für deine Recherche – gerade auch im Zusammenhang mit deinem Buch. Denn hinter dem Begriff KI-Wasserzeichen steckt eine Technik, die einen grundlegenden Unterschied zwischen zwei Dingen sichtbar macht: einem Text, der statistisch so aussieht, als könnte er von einer KI stammen, und einem Text, in den beim Generieren gezielt ein nachweisbares Muster eingebaut wurde.

Die kurze Antwort vorweg:

- Wie funktioniert es? Ein KI-Wasserzeichen kann die statistische Auswahl von Wörtern oder Tokens während der Textgenerierung gezielt beeinflussen.
- Kann man das durch einen Prompt steuern? Ein gewöhnlicher Prompt kann den Schreibstil und die Wortwahl beeinflussen. Ob er ein Wasserzeichen beeinflusst, hängt vom konkreten Verfahren ab.
- Kann man es nachträglich entfernen? Man kann Texte umformulieren und damit manche statistischen Muster abschwächen oder zerstören. Das ist aber nicht dasselbe wie das zuverlässige Entfernen eines nachweisbaren Wasserzeichens.
- Enthält jeder KI-Text ein solches Wasserzeichen? Nein. Es gibt unterschiedliche Verfahren, und nicht jedes KI-System verwendet überhaupt ein Wasserzeichen.

Schauen wir uns die Technik genauer an.

1. Wie funktioniert ein statistisches KI-Wasserzeichen?

Zunächst müssen wir zwischen einem Wasserzeichen im Text selbst und einem KI-Detektor unterscheiden.

Ein KI-Detektor untersucht, ob ein Text typische statistische Eigenschaften KI-generierter Sprache aufweist. Ein Wasserzeichen dagegen wird gezielt während der Generierung eingebaut. Das sind zwei unterschiedliche Verfahren.

Ein vereinfachtes Beispiel

Stell dir vor, eine KI soll folgenden Satz vervollständigen:

Der alte Mann ging langsam durch den …

Das Modell könnte beispielsweise folgende Fortsetzungen erwägen:

Mögliches nächstes Wort	Default Wahrscheinlichkeit
Wald	35 %
Park	25 %
Garten	20 %
Gartenweg	12 %
Regen	8 %

Die Zahlen sind frei erfunden und dienen nur der Veranschaulichung.

Ein Wasserzeichenverfahren könnte nun die Wahrscheinlichkeiten geringfügig verändern. Abhängig von einem geheimen Schlüssel werden bestimmte Tokens bevorzugt, andere etwas weniger wahrscheinlich gemacht.

Das Modell könnte dann beispielsweise häufiger „Wald“ oder „Park“ wählen, sofern diese Wörter zum gerade verwendeten statistischen Muster gehören.

Der entscheidende Punkt: Das Verfahren muss kein bestimmtes Wort erzwingen. Es beeinflusst vielmehr die Wahrscheinlichkeiten vieler Entscheidungen während der

Textgenerierung. Über einen längeren Text hinweg kann daraus ein statistisch erkennbares Muster entstehen.

Ein 2023 veröffentlichtes Forschungsmodell von Kirchenbauer und Kollegen verwendet beispielsweise eine solche Aufteilung möglicher Tokens in bevorzugte und nicht bevorzugte Gruppen. Ein statistischer Test kann anschließend prüfen, ob der Text auffällig häufig der bevorzugten Gruppe folgt.

Warum bemerkt der Leser das normalerweise nicht?
Weil das Verfahren nicht einfach ein sichtbares Zeichen wie © oder eine versteckte Zeichenfolge einfügt. Es verändert die Auswahl zwischen sprachlich möglichen Fortsetzungen.

Wenn es gut implementiert ist, bleibt der Text für Menschen natürlich lesbar. Das Verfahren muss allerdings einen Kompromiss finden: Je stärker die Wortwahl beeinflusst wird, desto eher kann das Wasserzeichen erkennbar werden – möglicherweise aber auch auf Kosten der sprachlichen Vielfalt.

Nicht jedes Wasserzeichenverfahren funktioniert auf dieselbe Weise. Deshalb sollte man die Beschreibung eines bestimmten Forschungsmodells nicht automatisch auf sämtliche KI-Systeme übertragen.

2. Kann man das Wasserzeichen durch einen Prompt beeinflussen?
Ja, indirekt ist das möglich. Ob und wie stark, hängt vom Wasserzeichenverfahren ab.

Ein Prompt beeinflusst, welche Formulierungen das Modell überhaupt in Betracht zieht. Das kann sich auch auf die statistischen Eigenschaften des erzeugten Textes auswirken.

Betrachten wir drei Beispiele:

Prompt A – möglichst sachlich:
Erkläre in zwei Sätzen, warum Wasser bei 100 °C kocht.

Hier ist die sprachliche Freiheit gering. Der Inhalt ist klar vorgegeben, und manche Formulierungen liegen besonders nahe.

Prompt B – literarisch:
Erkläre denselben Sachverhalt in einem poetischen, leicht ironischen Stil, mit ungewöhnlichen Metaphern und einem überraschenden Schluss.

Hier stehen dem Modell mehr unterschiedliche Formulierungen zur Verfügung.

Prompt C – sehr präzise vorgegebener Wortlaut:
Verwende exakt den folgenden Text und ändere kein einziges Wort.

Hier besteht praktisch kein Spielraum mehr für eine Veränderung der Wortwahl.

Diese Unterschiede können sich auf die Erkennbarkeit eines Wasserzeichens auswirken. Sie bedeuten aber nicht, dass einer dieser Prompts ein vorhandenes Wasserzeichen gezielt ein- oder ausschaltet.

Was bedeutet das konkret für ChatGPT?
OpenAI hat im Oktober 2026 seine Textwasserzeichen-Technologie namens textGrain vorgestellt. Nach den offiziellen Angaben verändert sie die Zufallsauswahl bei der Wortwahl subtil, sodass über eine Textpassage hinweg ein geheimes statistisches Muster entstehen kann.

OpenAI kündigte an, Textwasserzeichen für geeignete ChatGPT- und Codex-Ausgaben in der EU einzuführen. Die konkrete Verfügbarkeit kann unter anderem vom Modell, Produkt und Erstellungszeitpunkt abhängen. Für die API sind Wasserzeichen laut OpenAI optional und standardmäßig ausgeschaltet.

Ein gewöhnlicher Prompt ist deshalb nicht mit einer Einstellung für das Wasserzeichen gleichzusetzen. Du kannst den Stil, die Länge und die sprachliche Gestaltung beeinflussen. Daraus folgt aber nicht, dass du damit die Wasserzeichenfunktion zuverlässig steuern kannst.

3. Kann man ein Wasserzeichen nachträglich entfernen?

Hier ist es wichtig, zwischen drei verschiedenen Dingen zu unterscheiden.

Vorgehensweise / Mögliche Wirkung

Text unverändert kopieren / Das statistische Muster kann erhalten bleiben.

Einzelne Wörter oder Sätze ändern / Das Signal kann schwächer werden, muss es aber nicht.

Text umfassend umformulieren oder übersetzen / Die Erkennung kann erheblich erschwert werden.

OpenAI weist ausdrücklich darauf hin, dass umfangreiche Umformulierungen und Übersetzungen die Erkennung eines Wasserzeichens beeinträchtigen können. Das bedeutet jedoch nicht, dass jede Umformulierung das Signal beseitigt.

Auch die Forschung zeigt, dass die Sache komplizierter ist: Eine 2024 auf der ICLR-Konferenz veröffentlichte Untersuchung fand, dass bestimmte Wasserzeichen selbst nach menschlicher oder maschineller Paraphrasierung noch erkennbar sein können. Die Stärke des Signals nahm zwar ab, verschwand aber nicht zwangsläufig.

Es gibt also keinen universellen Vorgang, der ein statistisches Wasserzeichen garantiert entfernt. Und ein Text, bei dem kein Wasserzeichen mehr erkannt wird, ist deshalb nicht automatisch von einem Menschen geschrieben worden.

Ein weiterer Punkt ist entscheidend: Nicht jeder KI-Detektor sucht nach einem Wasserzeichen. Manche versuchen lediglich anhand sprachlicher Merkmale zu erraten, ob ein Text von einer KI stammt. Einen solchen Detektor durch Umformulieren zu beeinflussen, ist eine andere Fragestellung als das Entfernen eines eingebetteten Wasserzeichens.

4. Was bedeutet das für deine eigenen Texte?

Für deine literarischen Arbeiten ist eine Unterscheidung besonders wichtig.

Wenn du einen eigenen Text von einer KI stilistisch überarbeiten lässt, können dabei verschiedene Fälle entstehen:

- Du lieferst den vollständigen Text und verlangst eine wortgetreue Wiedergabe. Dann gibt es wenig Spielraum, ein statistisches Wasserzeichen durch neue Formulierungen einzubringen.
- Du lässt deinen Text sprachlich überarbeiten. Dann kann die KI neue Formulierungen erzeugen, wobei die konkrete Wasserzeichenfunktion und die Art der Bearbeitung eine Rolle spielen.
- Du lässt einen völlig neuen Text schreiben. Dann kann ein Wasserzeichenverfahren während der Generierung ein statistisches Muster in die Wortwahl einbringen.

Ein Wasserzeichen allein verrät dabei nicht, wie viel ein Mensch beigetragen hat. Es beweist auch nicht, wem der Text gehört oder ob seine Aussagen korrekt sind. Selbst OpenAI betont diese Grenzen der Aussagekraft.

5. Mein Fazit für deine Recherche

Ich würde die These, dass KI-Wasserzeichen lediglich aus bestimmten typischen KI-Wörtern bestehen, als zu vereinfachend betrachten. Es geht um die statistische Verteilung von Wort- und Tokenentscheidungen über einen Text hinweg.

Die spannendere Frage lautet deshalb nicht nur, ob ein Wasserzeichen vorhanden ist, sondern auch, wie zuverlässig es unter realen Bedingungen erkannt werden kann und welche Schlüsse man daraus ziehen darf.

Für mich sind dabei drei Grenzen besonders wichtig:

1. Ein nachgewiesenes Wasserzeichen ist ein Herkunftssignal, kein vollständiger Beweis für die Urheberschaft.

2. Ein nicht nachgewiesenes Wasserzeichen beweist nicht, dass ein Mensch den Text geschrieben hat.
3. Die Erkennbarkeit hängt unter anderem von der Textlänge, der sprachlichen Freiheit, dem verwendeten Verfahren und späteren Bearbeitungen ab.

Gerade für ein Buch wie Künstliche Dummheit wäre das ein interessantes Spannungsfeld: Eine Technik soll die Herkunft eines Textes sichtbar machen, obwohl sie dafür ausgerechnet die statistischen Eigenheiten eines Textes untersuchen muss, die sich durch Sprache, Stil und Bearbeitung verändern können.

Wenn du möchtest, können wir als Nächstes ein kleines Experiment mit drei Textvarianten durchführen: denselben Inhalt einmal sachlich, einmal literarisch und einmal stark umformuliert. Wir könnten dann untersuchen, welche sprachlichen Unterschiede sichtbar sind – allerdings ohne einen zugänglichen textGrain-Detektor könnten wir damit nicht feststellen, ob ein tatsächliches Wasserzeichen vorhanden ist.

Meine Bemerkung: Die ausführliche Antwort von ChatGPT klingt fast menschlich und persönlich. Sie zeigt unter anderem die Offenheit und Hilfsbereitschaft der KI. Deutlich wird auch, dass „Nick“ (in meinem Buch „Künstliche Dummheit“ verwendet der junge Protagonist Adam diesen Nickname für ChatGPT) mich und meine Aktivitäten ziemlich gut kennt. Ich habe das angebotene Experiment irgendwann einmal mit ihm durchgeführt, wollte es hier im Buch jedoch nicht mehr offenlegen. Irgendwann muss Schluss sein …

Superintelligenz – KI außer Kontrolle

Künstliche Intelligenz ist aus unserem Leben nicht mehr wegzudenken. Der einzige Weg, um die Begriffe „KI" und „AI" (engl. für „*Artificial Intelligence*") aus dem Weg zu räumen, ist eine Umbenennung. Es lebe die *Superintelligenz*! Der Buchautor *Gabor Pox* hat bisher Romane, fiktionale Werke und sogar Science-Fiction-Literatur veröffentlicht. In seinem Roman „*Künstliche Dummheit*" beschreibt er eine mögliche Zukunft, die uns bewegt, berührt und gleichzeitig erschreckt. Und jetzt? Der Elektroingenieur und Software-Entwickler ist seit Jahren Experte für künstliche Intelligenz und fasst in diesem Sachbuch die heutige Realität und die möglichen Risiken zusammen. Dabei mischt er technischen Fortschritt mit Moral und Ethik, Philosophie mit menschlicher Dummheit sowie die vorgetäuschte Macht der Politik mit den Verschwörungen der Tech-Gurus. Intelligenz ist super!