
Konstruktiver KI-Kompass

Wissen, Werkzeuge und Tipps
für den Redaktionsalltag

b

bonn institute

Das Bonn Institute

Mit anwendungsbezogener Forschung, Veröffentlichungen, Veranstaltungen und Fortbildungen trägt das gemeinnützige Bonn Institute dem gestiegenen Bedarf der Medienbranche nach Vernetzung und Wissensaustausch im Hinblick auf konstruktive und nutzerzentrierte Ansätze im Journalismus Rechnung. Dabei arbeitet es an der Schnittstelle von Forschung und Praxis und gibt Medienschaffenden alltags-taugliche und empirisch abgesicherte Werkzeuge an die Hand.

Ziel ist es, den Journalismus so weiterzuentwickeln, dass er die Interessen der Menschen in den Mittelpunkt stellt – gerade vor dem Hintergrund wachsender gesellschaftlicher Herausforderungen. Das international vernetzte Bonn Institute verbindet journalistische Kompetenz mit Einsichten aus Medienwissenschaften, Nutzerforschung, Psychologie und Mediation.

Das Bonn Institute wurde 2022 von den Gesellschaftern Deutsche Welle, Rheinische Post Mediengruppe, RTL Deutschland und Constructive Institute (Dänemark) gegründet. Gründerin und Geschäftsführerin ist Ellen Heinrichs. Zum Kuratorium des Bonn Institute gehören Persönlichkeiten wie Prof. Jay Rosen, Wolfgang Blau, Astrid Maier, Kübra Gümüşay und David Schraven.



CONSTRUCTIVE
INSTITUTE
Journalism for tomorrow

Media
University
of Applied Sciences



Deutsche
Welle



Konstruktiver KI-Kompass

**Wissen, Werkzeuge und Tipps
für den Redaktionsalltag**

Inhalt

Geleitwort Seite 3	Konstruktiver Journalismus	9
	Lösungsfokus	10
Einladung Seite 5	Perspektivenreichtum	11
	Konstruktiver Dialog	11
	KI verstehen	13
Index und Glossar Seite 65	Automatisierte Texte	14
	Entwicklungsstufen Künstlicher Intelligenz	14
Quellen und Literatur Seite 66	Bekannte KI-Modelle	22
Das Team Seite 70	KI verantwortungsvoll nutzen	25
	Recht	26
Danksagung Seite 72	Sorgfalt und Transparenz	29
	Interne Leitlinien	31
Impressum Seite 73	Training von KI-Modellen	34
	Auswirkungen auf die Umwelt	36
Fünf Einstiege ins konstruktive Prompting Umschlagklappe hinten	KI konstruktiv einsetzen	41
	Konstruktives Prompt-Engineering	42
	Prompts für konstruktive Themenideen	49
	Prompts für mehr Perspektivenreichtum	51
	Prompts für konstruktive Interviewfragen	55
	Von der Stimme zum Text: Transkribieren mit KI	57
	ChatGPT konstruktiv tunen	59
	Kann KI synthetische Inhalte erkennen?	62



Geleitwort

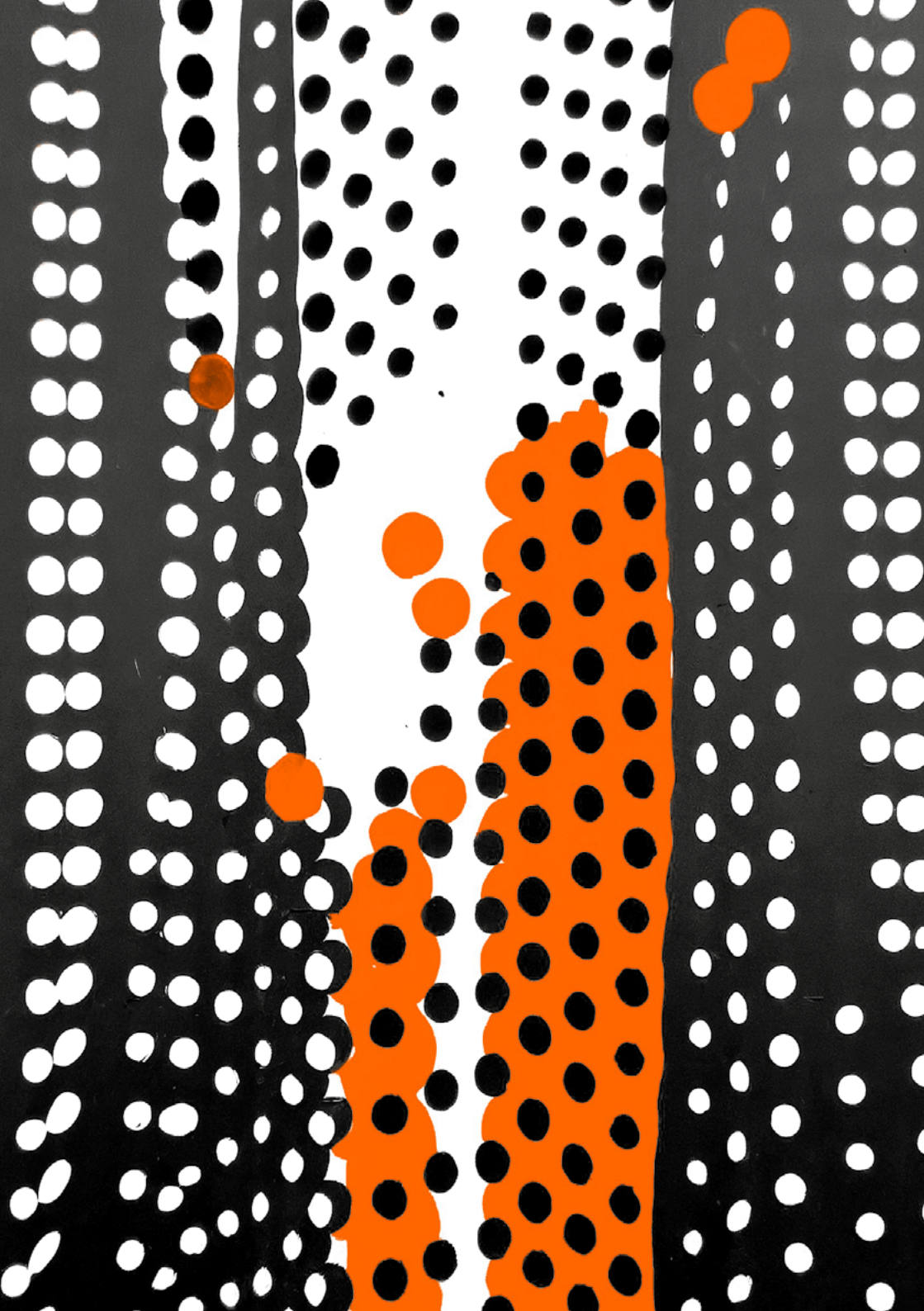
Viefältige journalistische Angebote und wirtschaftlich tragfähige Medienunternehmen sind nicht weniger als die Brandmauer unserer Demokratie. Unsere Aufgabe als Landesanstalt für Medien NRW ist die Sicherung dieser Angebots- und Anbietervielfalt der Medien in Nordrhein-Westfalen.

Wenn wir diese auch in der Zukunft aufrechterhalten wollen, müssen wir uns – gemeinsam mit den Medienunternehmen und wichtigen Stakeholdern wie dem Bonn Institute – für die Erneuerung des Journalismus einsetzen und die Zukunft und Wettbewerbsfähigkeit der Medien gemeinsam gestalten. KI ist eine der entscheidenden Schlüsseltechnologien und Wegbereiterin für journalistische Innovation in den kommenden Jahren. Sie kann zu einem breiten Spektrum an Perspektiven und Informationen beitragen, die einen offenen Diskurs ermöglichen – solange ihr Einsatz an professionellen Werten, ethischen Standards und Bedürfnissen der Nutzerinnen und Nutzer ausgerichtet ist.

Der *Konstruktive KI-Kompass* des Bonn Institute, der mit Unterstützung des Journalismus Lab der Landesanstalt für Medien NRW realisiert wird, bietet nachhaltigen Wissenstransfer und dient hoffentlich vielen Medienschaffenden als praxisrelevantes Werkzeug für den konstruktiven Umgang mit KI.

Simone Jost-Westendorf

Leiterin Journalismus Lab der Landesanstalt für Medien NRW



Einladung

Entwerten neue technische Entwicklungen die Arbeit von Menschen? Verlieren viele gar ihren Job? Solche Fragen waren schon in frühindustrieller Zeit Kern von Konflikten: Vor mehr als 200 Jahren erschütterte die Möglichkeit, dass Maschinen Aufgaben übernehmen, die bislang in Handarbeit erledigt wurden, das in Zünften organisierte Handwerk zutiefst.

Auch heute beschäftigen uns solche Fragen. Auch heute stehen wir an dem Punkt, an dem eine neue Gattung von Maschinen Tätigkeiten übernehmen wird, die bis dato nur Menschen ausführen konnten. Und auch heute stellt sich die Frage, wie die Berufsstände, also auch wir als Journalistinnen und Journalisten, damit umgehen.

Seit der Verbreitung von Computern in den 1980er Jahren haben zahlreiche technische Neuerungen Einzug in den journalistischen Alltag gehalten. Computer – erst stationäre, später tragbare –, das Internet, Mobiltelefone, Social Media und immer wieder auch neue Software und Tools verändern journalistische Recherche, Produktion und Distribution seither kontinuierlich. Dieser Wandel ist zum Alltag geworden, hat vieles einfacher und manches überhaupt erst möglich gemacht. Teilweise hat er aber auch eine gewisse Erschöpfung verursacht.

Und jetzt kommt Künstliche Intelligenz, kurz: KI.

Der wirtschaftliche Einsatz von Dampfmaschinen läutete vor mehr als 200 Jahren eine Phase des tiefgreifenden Umbruchs ein, die später als industrielle Revolution in die Geschichte einging. Teil dieses Umbruchs waren zentrale Erfindungen, die auf die Dampfmaschine zurückgriffen: Dampfschiffe, die den weltweiten Handel ankurbelten, Eisenbahnen, die Personenverkehr und Gütertransport veränderten, dampfbetriebene Webmaschinen, die die Textilproduktion revolutionierten und weitere von der neuen Technologie inspirierte Erfindungen.

Wie werden wir in 200 Jahren auf die Veränderungen in der Gesellschaft blicken, die durch KI ausgelöst wurden? Welche Erfindungen wird KI ermöglichen?

Das hängt auch davon ab, wie wir mit dieser neuen Technologie umgehen und ihr entgegenreten: Sorgenvoll? Hoffnungsvoll? Ambivalent? Es gibt gute Gründe für jede einzelne dieser Haltungen. Die Gefahren, die eine missbräuchliche Nutzung von KI für die Gesellschaft birgt, sind uns bewusst (vgl. z.B. Adeoso et al. 2024, S.125 ff.). Wir sehen in den neuen Entwicklungen aber zugleich auch viel Potenzial.

Um die Chancen von KI für den Journalismus konstruktiv zu nutzen, kommt es darauf an, die Technologie so einzusetzen, dass sie einen Mehrwert bietet – für Menschen, Medien und die Demokratie.

Dafür braucht es aus unserer Sicht zweierlei: zum einen ein solides Verständnis der neuen Technologie und zum anderen Inspiration und Know-how, um konstruktive Anwendungsszenarien zu entwickeln und umzusetzen.

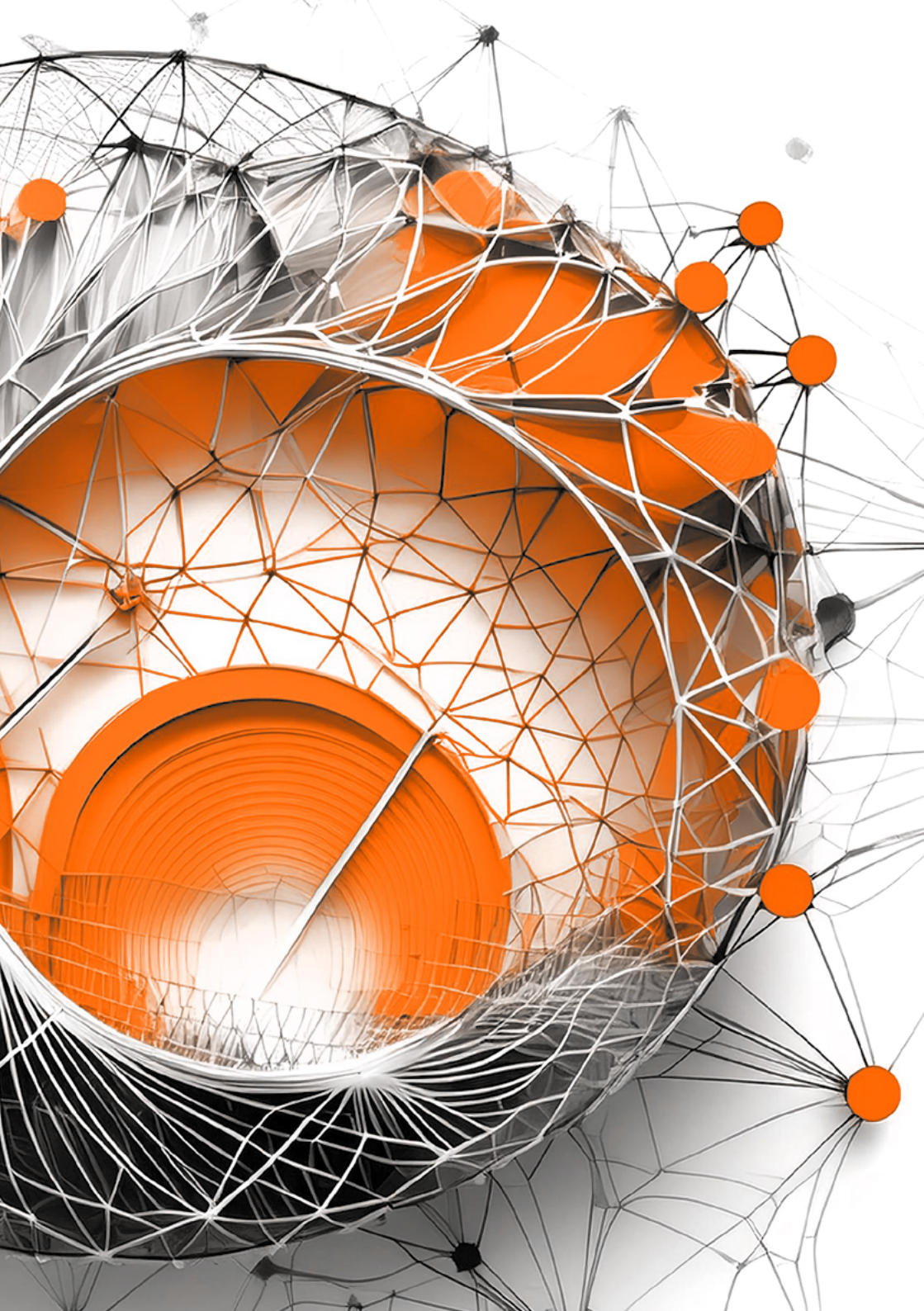
Dies beides zu vermitteln, ist zentrales Anliegen dieser Publikation. Der Fokus liegt dabei klar auf der Frage nach dem Nutzen von Text generierenden Anwendungen wie Open Als ChatGPT, Googles Gemini oder Anthropic Claude für den Journalismus. Außen vor bleiben hingegen Werkzeuge, die Bilder und Videos generieren. Letztere einzubeziehen, würde den Rahmen dieser Publikation sprengen.

Im Besonderen richtet sich dieses Angebot an Journalistinnen und Journalisten, die KI verantwortungsvoll im Alltag nutzen möchten. Wir laden Sie ein zu einer Reise, die konstruktiven Journalismus und KI verbindet. Fragen, die sich das Management im Kontext von KI stellen muss, behandelt diese Publikation hingegen nicht explizit. Doch auch Managerinnen und Manager sowie alle anderen, die sich für die Grundlagen von Künstlicher Intelligenz und konstruktive Einsatzszenarien interessieren, sind zu dieser Entdeckungsreise eingeladen.

Nach einem Ausflug zu den drei zentralen Elementen des konstruktiven Journalismus werden wir tief eintauchen in die Funktionsweise von KI, also von automatisierten Entscheidungssystemen. Danach zeigen wir auf, in welchem rechtlichen und ethischen Rahmen wir KI verantwortungsvoll nutzen können und welche Fragen wir uns dazu stellen sollten.

Dann wird es praktisch: Es geht ums konstruktive Prompt-Engineering. Dabei werden sowohl die Grundlagen des Promptings gezeigt, als auch Anwendungsszenarien vorgestellt, bei denen die drei Elemente des konstruktiven Journalismus eine zentrale Rolle spielen. Wer möchte, erfährt zudem, wie sich ein eigener digitaler Assistent trainieren lässt.

Wo Sie in die Lektüre einsteigen, ist ganz Ihnen überlassen. Die Buchseiten geben nicht Ihre Lesereihenfolge vor, die Reise ist nicht als Pauschalurlaub angelegt. Vielmehr ist dieses Büchlein so konzipiert, dass Sie bei jedem Kapitel einsteigen und sich individuell auf den Weg machen können. Wir wünschen Ihnen viel Freude beim Entdecken, Erkunden und Experimentieren!



Konstruktiver Journalismus

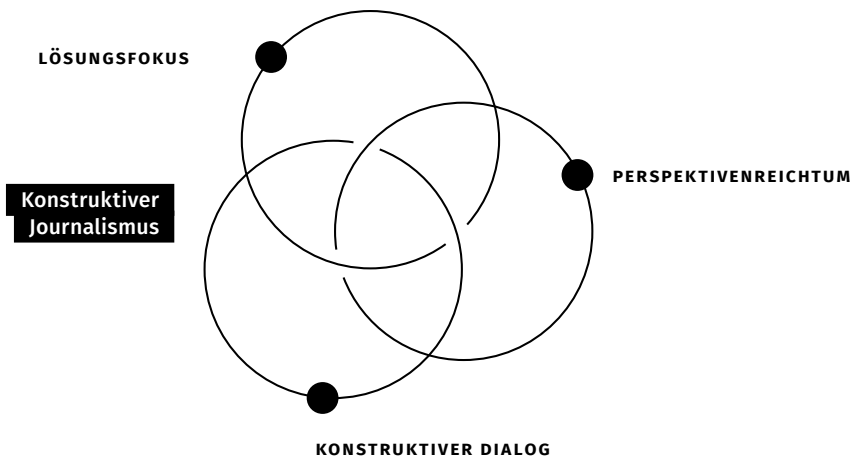
Vielfältige Veränderungsprozesse, politische und gesellschaftliche Konflikte, Desinformation, Nachrichtenvermeidung und das vielerorts schwindende Vertrauen in die Medien stellen die gesamte Branche vor enorme Herausforderungen.

Zentral ist die Frage: Wie muss sich Journalismus verändern, damit er relevant und zukunftsfähig bleibt und einen Mehrwert liefert – für Menschen, Medien und Demokratie? Konstruktiver Journalismus will genau darauf Antworten geben. Wie Studien zeigen, birgt er unter anderem das Potenzial, Nachrichtenvermeidung zurückzudrängen, gesellschaftlicher Polarisierung entgegenzuwirken und für Medienhäuser neue Monetarisierungschancen zu eröffnen.

Seine drei zentralen Elemente sind: Lösungsfokus, Perspektivenvielfalt und konstruktiver Dialog.

Konstruktiver Journalismus zielt darauf ab, Mediennutzenden ein zukunftsorientiertes, faktenbasiertes und nuanciertes Bild der Wirklichkeit zu vermitteln. Indem er Lösungsansätze genauso sorgfältig recherchiert wie Probleme, wirkt er einer einseitig negativen Weltsicht entgegen. Gleichzeitig stärkt er durch das Aufzeigen von Handlungsoptionen bei Mediennutzenden das Gefühl der Selbstwirksamkeit.

Weil konstruktiver Journalismus bewusst auf Vielfalt und unterschiedliche Perspektiven setzt, reflektiert er die Welt in ihrer ganzen Komplexität und wirkt übermäßiger Vereinfachung und Spaltungstendenzen entgegen. Und da er die Rolle von Journalistinnen und Journalisten als Moderierende eines öffentlichen konstruktiven Dialogs neu definiert, bietet er zahlreiche Möglichkeiten für bessere Gespräche in unserer Gesellschaft.



Lösungsfokus

Oft hört Journalismus da auf, wo ein Problem ausreichend benannt oder beschrieben ist. Konstruktiver Journalismus geht einen Schritt weiter: Er berichtet insbesondere auch über Lösungsansätze für gesellschaftliche Probleme und weitet damit schon in der journalistischen Recherche den Blick. »Wer macht es besser?« ist hier eine der zentralen Fragen.

Der Lösungsjournalismus, so wie ihn das Solutions Journalism Network entwickelt hat, weist vier Qualitätsmerkmale auf, die es Journalistinnen und Journalisten erleichtern, qualitativ hochwertige, kritische und zugleich konstruktive Inhalte zu erstellen. Dabei geht es darum,

... dass ein Lösungsansatz im Vordergrund der Berichterstattung steht.

... dass Evidenzen benannt werden, also quantitative oder qualitative Belege, die die Effektivität eines Lösungsansatzes darlegen.

... dass die Übertragbarkeit gewährleistet ist.

... dass Grenzen des recherchierten Lösungsansatzes beschrieben werden.

Diese Qualitätsmerkmale bewahren Journalistinnen und Journalisten vor Schönfärberei, Greenwashing oder anderen Fallstricken. Sie können als Checkliste für die Recherche und die Erstellung von Beiträgen verwendet werden.

Perspektivenreichtum

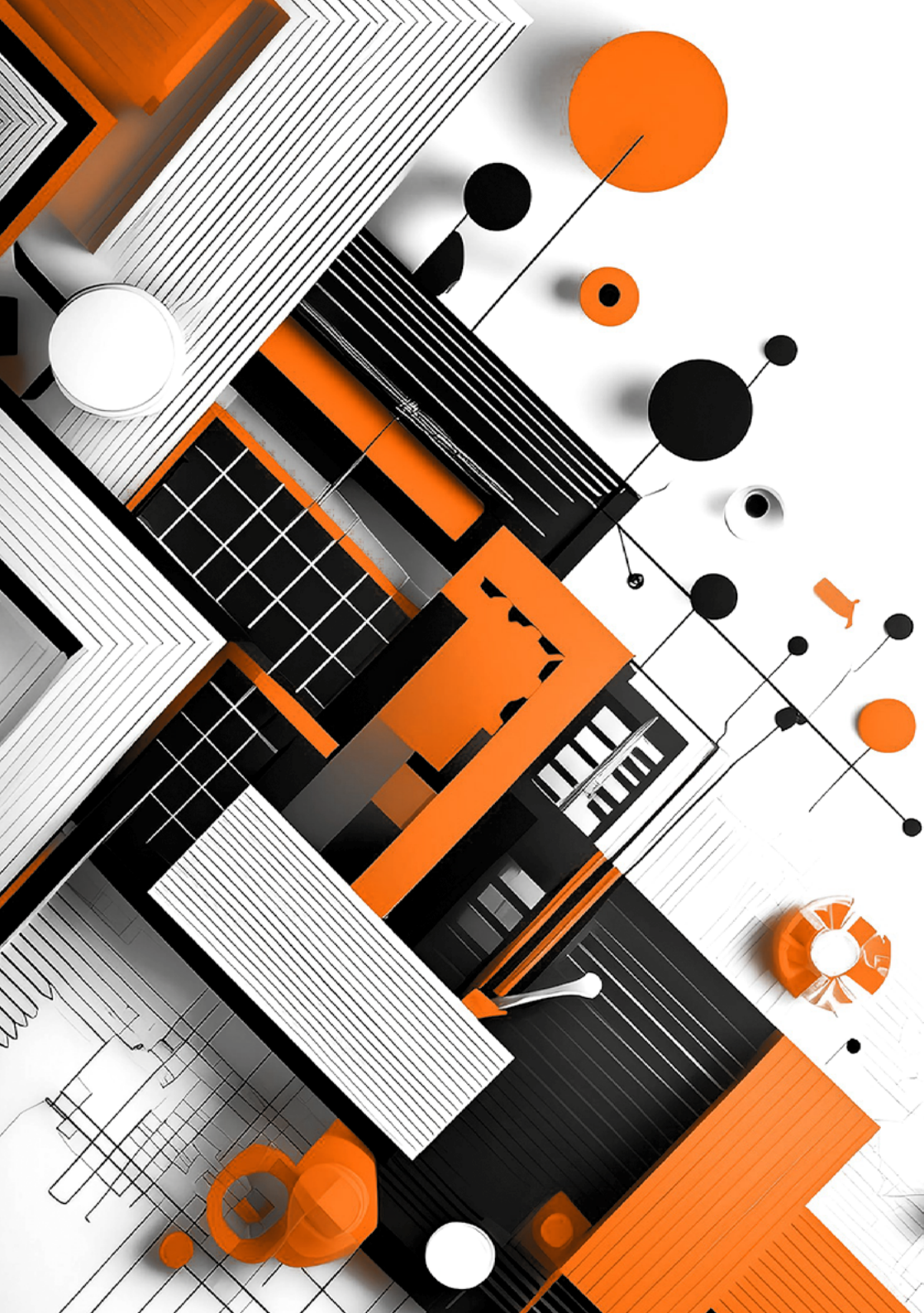
Während der Lösungsjournalismus vor allem auf das journalistische Handwerk fokussiert, schaut der Perspektivenreichtum auch auf systemische Faktoren – sei es im Hinblick auf die Personalauswahl oder auf ein besseres Verständnis eigener blinder Flecken und Voreingenommenheit: Ist mein Blick vielleicht verzerrt durch einen gewohnheitsmäßigen Fokus auf Negativität? Welche interessanten und relevanten Perspektiven auf das Thema habe ich bei meiner Recherche noch nicht berücksichtigt? Sind in unserer Redaktion genügend unterschiedliche Menschen tätig, um die Vielfalt der Gesellschaft auch angemessen abzubilden?

Perspektivenreicher Journalismus ist relevanter Journalismus – denn er hat die Informationsbedürfnisse der vielen verschiedenen Menschen in unserer Gesellschaft im Blick.

Konstruktiver Dialog

Der konstruktive Dialog bildet das dritte zentrale Element des konstruktiven Journalismus. Wichtige Faktoren sind eine Haltung der Allparteilichkeit, empathisches Interesse und Zuhören sowie der Fokus auf Gemeinsamkeiten (nicht nur Trennendes) und der Blick nach vorn. Denn Journalistinnen und Journalisten sind nicht nur als Vermittlerinnen und Vermittler von relevanten Informationen und verschiedenen Perspektiven gefragt, sondern moderieren zwischen Menschen und Gruppen in der Gesellschaft.

Konstruktiver Dialog ist damit ein wichtiges Werkzeug, um menschliche Verständigung zu organisieren: analoge und digitale Orte des Austauschs schaffen, Menschen mit unterschiedlichem Hintergrund miteinander ins Gespräch bringen, bessere Fragen stellen – auch bei Recherchen, friedliche, zukunftsgerichtete Debatten über relevante gesellschaftliche Themen befördern und moderieren. All dies ermöglicht Fortschritt und kann wesentlich dazu beitragen, den sozialen Zusammenhalt sowie die Demokratie zu stärken.



KI verstehen

Die Entstehung des Begriffs »Künstliche Intelligenz« ist banaler, als man denkt: Er geht zurück auf den Stanford-Professor John McCarthy. Der erfand ihn 1955, um Geld für ein zweimonatiges Forschungsprojekt im Sommer 1956 zu bekommen. Eine Erklärung oder formale Definition des Begriffs sucht man in seinem Förderantrag vergeblich. Überhaupt taucht er darin eher als Oberbegriff auf. »In der Studie wird davon ausgegangen, dass jeder Aspekt des Lernens beziehungsweise jedes weitere Merkmal von Intelligenz im Prinzip so präzise beschrieben werden kann, dass die Entwicklung einer Maschine zur Simulation aller einzelnen Schritte möglich wird«, heißt es in dem Antrag (McCarthy et al. 1955). Dann werden »einige Aspekte des Problems künstlicher Intelligenz« aufgeführt. Das wars. Ein Marketingbegriff also. McCarthy bekam das Geld und der Begriff war in der Welt. Doch genau zu beschreiben, was Intelligenz ist, misslang (vgl. Brooks 2018 und Whittaker 2023).

Hier setzt auch die aktuelle Kritik am Begriff Künstliche Intelligenz (KI) an. Denn bis heute gibt es keine allgemeine Definition von Intelligenz. Wenn wir also schon nicht wissen, was menschliche Intelligenz ausmacht – wie sollen wir dann eine künstliche Intelligenz definieren? Doch auch, wenn der Begriff nicht eindeutig ist, so ist KI nun da. KI ist nicht neu, die Forschung darüber schon gar nicht. Neu ist aber seit 2022, dass Programme, die mit generativer KI arbeiten, nun für alle zugänglich und sehr einfach nutzbar sind. Es ist also an der Zeit, sich einmal genauer anzuschauen, was dahintersteckt.

Automatisierte Texte

Wenn es derzeit darum geht, wie KI im Journalismus eingesetzt wird, so werden fälschlicherweise oft zwei völlig unterschiedliche Herangehensweisen in einen Topf geworfen: das automatisierte Erstellen von Inhalten einerseits und die Generierung synthetischer Inhalte durch den Einsatz großer Sprachmodelle andererseits.

Diese Unterscheidung ist aber für uns Journalistinnen und Journalisten von zentraler Bedeutung. Denn bei automatisiert erstellten Texten können wir das Ergebnis und bei Bedarf auch Fehler nachvollziehen, bei den Ergebnissen von Sprachmodellen bisher nicht. Warum? Bei der reinen Automatisierung macht der Mensch so genaue Vorgaben, dass die Ergebnisse nachvollziehbar bleiben. Sprachmodelle generieren Inhalte hingegen eigenständig, so dass diese Nachvollziehbarkeit fehlt.

Für automatisierte Texte braucht es zweierlei: Strukturierte Daten wie etwa eine Tabelle mit Wetterdaten, Wahlergebnissen oder Sportereignissen und ein Programm wie Retresco oder AX Semantics, in dem Textbausteine vorbereitet und Regeln vorgegeben werden, nach denen dann automatisiert Texte erstellt werden.

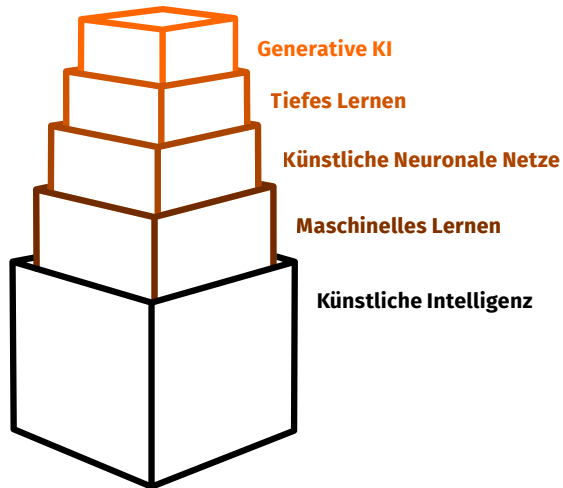
Die Nachrichtenagentur Associated Press (AP) nutzte diese Technik schon 2014, um automatisiert Texte über Quartalsberichte börsennotierter Unternehmen verfassen zu lassen. Inzwischen ist die Liste der Einsatzbereiche lang. Das Magazin *Der Spiegel* stellte zur Bundestagswahl 2021 auf diese Weise auch Einzelanalysen für alle 299 Wahlkreise zur Verfügung.

Diese Art der Texterstellung kann Routinen erleichtern und eröffnet neue Möglichkeiten. KI ist das aber noch nicht.

Entwicklungsstufen Künstlicher Intelligenz

Um zu verstehen, was hinter dem steckt, was wir als Künstliche Intelligenz bezeichnen, und wie sie funktioniert, werden wir die Motorhaube heben und uns anschauen, was sich darunter verbirgt. Was bezeichnen Begriffe wie Künstliche Neuronale Netze, Transformer oder GPT? Dieses Wissen gehörte bislang nicht zum Repertoire von uns Medienschaffenden. Wenn wir nun aber beginnen, KI-Anwendungen im Redaktionsalltag zu nutzen, sollten wir hiervon ein Grundverständnis haben. Wer Auto fährt, muss zwar weder ein Auto bauen noch es reparieren können. Den Motor von Tank und Auspuff zu unterscheiden, gehört hingegen durchaus

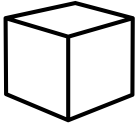
Entwicklungsstufen Künstlicher Intelligenz



zum Allgemeinwissen des 20. Jahrhunderts. Zum Allgemeinwissen des 21. Jahrhunderts wird es bald auch gehören, ein grundlegendes Verständnis von KI-Technologien zu haben. Doch wie hängen Künstliche Intelligenz, Maschinelles Lernen, Künstliche Neuronale Netze, Tiefes Lernen und generative KI zusammen? Wo liegen die Unterschiede? Und was muss man darüber wissen?

Um das zu verstehen, kann es uns helfen, die technologische Entwicklung mit dem Bau eines Stapelturms zu vergleichen. Stapeltürme gibt es in vielen Kinderzimmern und in verschiedenen Ausführungen. Der Klassiker besteht aus hohlen Holzwürfeln, die an einer Seite offen sind und sich in absteigender Größe aufeinander stapeln lassen. Je höher der Turm wird, desto wackeliger wird er. Für kleine Kinder ist das Stapeln eine Herausforderung, gilt es doch zu verstehen in welcher Reihenfolge die Würfel stabil stehen. Für noch kleinere ist es ein großer Spaß, die Türme zum Einstürzen zu bringen. Eltern wiederum freuen sich, dass sich die Würfel platzsparend ineinander stapeln lassen – alle in den größten hinein.

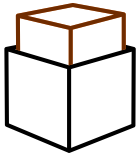
Ähnlich wie beim Stapelturm war es in der Entwicklung von KI: Generative KI ist nicht plötzlich entstanden, sondern nach und nach. Jede Stufe baut jeweils auf den Erkenntnissen der vorherigen auf. Generative KI, die uns im Moment am meisten beschäftigt, ist die am weitesten fortgeschrittene Technologie im Kontext Künstlicher Intelligenz und ist daher der oberste Würfel unseres Turms.



Künstliche Intelligenz

Der größte und erste Würfel unseres imaginären Stapelturmes ist Künstliche Intelligenz. Davon sprechen Expertinnen und Experten, wenn Maschinen in der Lage sind, automatisiert sinnvolle Ergebnisse zu liefern, ohne dass ihnen die Programmierung konkrete Lösungswege vorgibt. Das Fraunhofer-Institut für Kognitive Systeme IKS erklärt es so: »Künstliche Intelligenz imitiert menschliche kognitive Fähigkeiten, indem sie Informationen aus Eingabedaten erkennt und sortiert« (IKS 2024).

In der Geschichte der KI haben Forschende zunächst versucht, Software alle notwendigen Regeln zu geben, mit dem Ziel, dass die Systeme daraus eigene Entscheidungen ableiten können. Ein solches sogenanntes Expertensystem war auch der 1966 vorgestellte und damit weltweit erste Chatbot ELIZA, der einen Psychotherapeuten nachahmen sollte. Auch die Verarbeitung menschlicher Sprache sollten Programme zunächst über die Eingabe diverser Regeln lernen. Das gelang allerdings nur mit mäßigem Erfolg. Denn es stellte sich heraus, dass die Regeln, denen Sprache folgt, zu komplex sind. Daher wurde nach neuen Wegen gesucht. Für die heute wichtigen KI-Anwendungen ist Maschinelles Lernen von zentraler Bedeutung.



Maschinelles Lernen

Maschinelles Lernen (»machine learning«) ist ein Teilbereich von Künstlicher Intelligenz und damit der zweite Würfel unseres Turms. Es unterscheidet sich vom größeren Feld der KI dadurch, dass hier Computer aus Daten selbst etwas »lernen« und eben nicht über Regeln. Dabei wird ein Lernalgorithmus mit sehr großen Datenmengen gefüttert. Aus der Analyse all dieser Daten leitet der Algorithmus dann Zusammenhänge ab und entwickelt für sich ein statistisches Modell. Manchmal werden solche Prozesse von menschlichem Feedback begleitet, manchmal aber auch nicht. Je nach Methode spricht man von überwachtem Lernen (»supervised learning«) oder unüberwachtem Lernen (»unsupervised learning«). Welche Lernmethode gewählt wird, hängt von den Entscheidungen der Programmierenden ab und ist Teil des Lernalgorithmus.

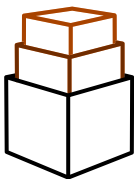
Das statistische Modell, das die Software während des Lernprozesses entwickelt, ist, nach dem Trainingsprozess, dann die Grundlage für Vorhersagen der Software. Das heißt, die Software kann dann ihr Modell später auf andere, ihr bis dahin unbekannte Daten anwenden und ebenso korrekte Antworten generieren.

Ab diesem Punkt der technologischen Entwicklung beginnt es für Menschen undurchsichtig zu werden. Denn: »Die Entwicklerinnen und Entwickler geben der Maschine dabei nur die grobe Struktur vor. Die

Feinheiten werden durch die lernende Methode in Zusammenarbeit mit den Daten festgelegt«, erklärt Informatikprofessorin Katharina Zweig (2023). Das heißt: Was wir als Lernen der Maschine bezeichnen, ist ein Prozess, der auch von den Programmierenden nicht mehr vollends nachvollzogen werden kann. Sie können sich zwar das statistische Modell anschauen, das die Software im Lernprozess entwickelt. Wie genau aber das System dazu gekommen ist, können auch Fachleute nur erraten.

Die Theorie des Maschinellen Lernens ist ebenso wie die der Künstlichen Intelligenz schon mehr als 70 Jahre alt. Das erste wissenschaftliche Papier dazu schrieb Alan Turing. Es ging den Fragen nach: Können Maschinen denken? Und wie ließe sich das testen? Veröffentlicht wurde es 1950 und damit noch bevor der Begriff »Künstliche Intelligenz« erfunden war. Es blieb dann zunächst bei der Theorie – denn für die Umsetzung und das Testen der Ideen fehlten lange zwei Dinge: Daten und Rechenpower.

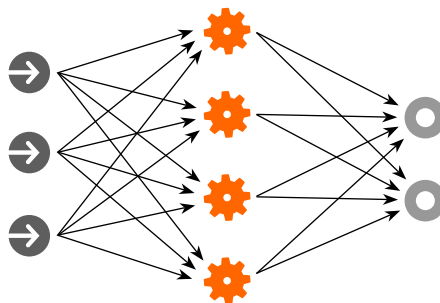
Punktuell gab es wichtige Erfolge: So etwa, als der Schachcomputer Deep Blue von IBM 1996 den damaligen Schachweltmeister Gary Kasparow besiegte. Gut zwanzig Jahre danach gelang es 2017 einem Computer namens AlphaGo, den die Google-Tochter Deepmind konstruiert hatte, den weltweit besten Spieler im Spiel Go, Ke Jie, zu besiegen. Doch der Durchbruch in der Anwendbarkeit ließ noch auf sich warten.



Künstliche Neuronale Netze

Bei der weiteren Entwicklung erwiesen sich dann ausgerechnet Katzenvideos als nützliche Ressource. Womit wir bei der dritten Entwicklungsstufe und dem nächsten Würfel unseres Turms wären: Künstlichen Neuronale Netzen.

Einfaches
Künstliches
Neuronales
Netz



Eingabe-Ebene

Verdeckte Ebene

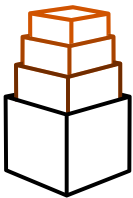
Ausgabe-Ebene

Dieser Begriff bezeichnet eine spezielle Methode des Maschinellen Lernens, die schon 1943 erfunden wurde (vgl. Brooks 2018). Bis zu ihrem Durchbruch dauerte es aber eine ganze Weile. Inzwischen wird sie gern bei komplexen Lernprozessen eingesetzt. Künstliche Neuronale Netze sind sehr ausgefeilte Algorithmen, deren Strukturen das menschliche Gehirn mit Neuronen und Synapsen zum Vorbild haben. Was im Gehirn Neuronen sind, sind im Algorithmus jeweils Rechenanweisungen – genauer gesagt mathematische Funktionen. Diese instruieren die Software, eine Reihe von Eingabewerten auf eine bestimmte Weise zusammenzuzählen. Diese Ergebnisse werden dann zu den Eingabewerten der nächsten Ebene.

So, wie die Synapsen im Gehirn Reize erst ab einer gewissen Intensität weiterleiten, bestimmt auch der Algorithmus, wie stark ein einzelnes Ergebnis weitergegeben und in der nächsten Ebene berücksichtigt wird. Das führt zu dem Effekt, dass – vergleichbar wie beim menschlichen Gehirn – nicht alle Eingaben gleichermaßen berücksichtigt werden.

Ein einfaches Künstliches Neuronales Netz hat drei Ebenen: Eine Eingabe-Ebene, eine verdeckte Ebene und eine Ausgabe-Ebene. Auf der Eingabe-Ebene stehen von Menschen konzipierte Rechenformeln. Sie machen der Maschine Vorgaben und setzen den Rechenprozess in Gang. Die Rechenergebnisse dieser ersten Ebene werden von der zweiten Ebene eigenständig weiterverarbeitet. Die Berechnungen für die verdeckte Ebene sind vorgegeben.

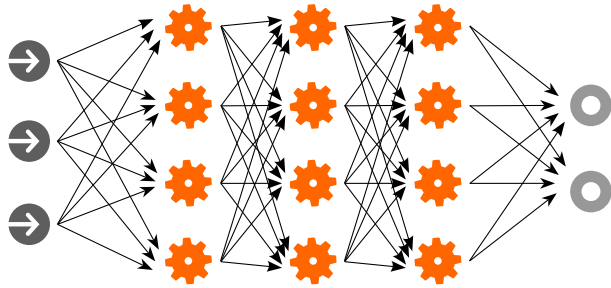
Der »Lernprozess« der Maschine bezieht sich darauf, in vielen Durchgängen die Gewichtungen festzulegen, mit denen die Ergebnisse der einen Ebene in die Berechnungen der nächsten Ebene einfließen. Die zweite Ebene und alle weiteren aber sind verdeckt – hier rechnet der Algorithmus jeweils mit den Ergebnissen der vorherigen Ebene und den von ihm selbst entwickelten Gewichtungen weiter. Einen Einblick in diese tiefen Ebenen haben die Programmierenden nicht. Das ist ein Charakteristikum Künstlicher Neuronaler Netze und der Hintergrund, warum KI oft auch als Black Box beschrieben wird.



Tiefes Lernen

Was hat das nun mit den Katzenvideos zu tun? Wie erwähnt, fehlte für die Entwicklung Künstlicher Intelligenz lange eine ausreichend große Menge verfügbarer Daten. Das änderte sich in den Nullerjahren. Nach der Gründung von YouTube 2005 wurden dort millionenfach Videos hochgeladen. Ein Datenschatz, den sich der Informatiker Andrew Ng zunutze machte: Er verwendete einen Trainingsdatensatz mit zehn Millionen Videos dazu, ein Künstliches Neuronales Netz auf das Erkennen von Katzen in Videos zu trainieren. Sein Ziel war dabei, dass die Software lernen sollte, wie eine Katze aussieht, ohne dass

Tiefes Künstliches Neuronales Netz



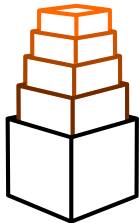
Eingabe-Ebene

Verdeckte Ebene

Ausgabe-Ebene

er es der Software vorher gezeigt hatte (vgl. DataScientest 2023). Das war der Beginn einer neuen Ära: der des sogenannten Tiefen Lernens (»deep learning«), dem vierten Würfel unseres Stapelturms.

Beim Deep Learning oder auch Tiefen Lernen werden neuronale Netze mit sehr vielen Ebenen eingesetzt. Ein Künstliches Neuronales Netz, das mindestens zwei dieser versteckten Ebenen hat, wird schon als Tiefes Neuronales Netz bezeichnet. Laut Fraunhofer IKS können sie aber sogar bis zu Hunderttausende oder Millionen Schichten haben. Können wir uns das noch vorstellen? Hunderttausende aufeinanderfolgende Ebenen, auf denen jeweils unzählige parallele Rechenoperationen ausgeführt werden... Zu Gesicht bekommen wir davon nur das Ergebnis: Katze erkannt! – und sind verblüfft.



Generative KI

Noch verblüffender sind die Möglichkeiten der nächsten Entwicklungsstufe von KI – dem obersten Würfel unseres Turms: generativer KI. Seit die Firma OpenAI im November 2022 ChatGPT für die Allgemeinheit freigegeben hat, kommen immer mehr Angebote hinzu. Claude, Imagen, Sora und andere Tools können in Windeseile Texte, Bilder und Videos generieren.

Doch was unterscheidet diese Art von KI nun von den anderen Bereichen, die wir schon betrachtet haben? Auf den Punkt bringt das der leitende KI-Experte der *New York Times*, Zach Seward (2023): »Wenn traditionelles Maschinelles Lernen gut Muster in einem Haufen ungeordneter Daten erkennen kann, so könnte man sagen, dass die Superpower generativer KI darin liegt, Muster zu kreieren.« KI-Anwendungen werden in der Regel jeweils für eine spezifische Aufgabe entwickelt. Das gilt für die Antworten des Robo-Therapeuten ELIZA genauso wie für das Erkennen von Katzen in Videos und ebenso

für generative KIs. Die großen Sprachmodelle (→Glossar) sind dabei auf eine sehr komplexe und umfangreiche Aufgabe trainiert worden: das Verarbeiten von natürlicher Sprache.

Deswegen heißen sie auch Sprachmodelle. Ihr statistisches Modell ist darauf trainiert, die Wahrscheinlichkeit von Wortteilen und Sätzen vorherzusagen – zum Beispiel auch die Sequenz »Alle meine Entchen ...« zu vervollständigen.

Transformer

Möglich wird das in der aktuellen Qualität und Schnelligkeit durch so genannte Transformer. ChatGPT trägt das Wort sogar im Namen: GPT steht nämlich für Generative Pretrained Transformer – es beschreibt also auf Englisch, womit wir es hier zu tun haben: einem generativen vortrainierten Transformer. Transformer bezeichnet dabei die Art des Lernalgorithmus. Vortrainiert beschreibt, dass das Modell zunächst trainiert werden muss, bevor es einsatzbereit ist, und dann aber auch weiterhin verbessert werden kann. Und generativ beschreibt schließlich die Art des Outputs – etwas neu Geschaffenes.

Technisch ist ein Transformer eine speziell aufgebaute Deep-Learning-Architektur eines Künstlichen Neuronalen Netzwerks. Sie wurde 2017 erstmals auf einer Fachkonferenz vorgestellt und hat seitdem die Verarbeitung natürlicher Sprache revolutioniert. Denn die besondere Herausforderung bei der Verarbeitung von Sprache im Unterschied zu numerischen Daten ist, dass Wörter je nach ihrer Platzierung im Satz und auch trotz gleicher Schreibweise unterschiedliche Bedeutungen haben können.

Es kommt also auf den Sinnzusammenhang an.

Um Text zu verarbeiten, muss dieser zunächst in Zahlen übersetzt werden. Transformer-Modelle sind erstmals in der Lage, auch die Position der Wörter im Satz in numerischen Werten zu markieren und dadurch den Kontext von Wörtern besser zu erfassen. Zudem ist es durch Transformer erstmals möglich, dass Algorithmen lernen, welche Teile der eingespeisten Daten am relevantesten für einen bestimmten Output sind (vgl. Klofat 2023 und Clarke et al. 2023).

Im Fall von ChatGPT haben wir es mit einem sehr großen Sprachmodell zu tun, dessen Aufgabe es ist, Sprache zu verarbeiten und zu generieren. Ähnlich verhält es sich mit anderen Angeboten (vgl. dazu Seite 22). Der Chatbot von Open AI, ChatGPT, basiert – je nach Wahl – auf den Modellen GPT 3.5, GPT 4 oder GPT 4o. Weil diese Modelle für verschiedene Aufgaben weitertrainiert werden können, heißen sie auch Basismodelle oder Grundmodelle (»foundation models«).

Sind große Sprachmodelle (LLM) und Basismodelle also das Gleiche? Nein, nicht ganz. Der Begriff Basismodelle umfasst zwar die großen Sprachmodelle. Aber er erfasst darüber hinaus auch Modelle, die nicht Sprache erzeugen, sondern Bilder oder Videos. Denn während Sprachmodelle ausschließlich mit Textdaten trainiert werden, werden andere Basismodelle wie beispielsweise DALL-E von Open AI neben Textdaten auch mit Bilddaten trainiert und Sora auch mit Videodaten. Entsprechend haben sie auch andere Fähigkeiten wie das Generieren von Bildern, Audio oder Videos.

Starke und schwache KI

Auch wenn das alles schon enorm stark wirkt, so gelten alle diese Anwendungen noch als sogenannte schwache KIs (»narrow artificial intelligence«). Das heißt, dass sie darauf ausgelegt sind, eine spezifische Aufgabe zu übernehmen, die bislang nur Menschen erledigen konnten, also zum Beispiel das Generieren von Texten oder das Generieren von Bildern. Unter einer starken KI hingegen verstehen Forschende der Computerwissenschaften eine Anwendung, die alle Aufgaben erledigen kann, die ein Mensch erledigen kann. Umgangssprachlich wird dabei auch von einer »artificial general intelligence«, kurz AGI oder auch Superintelligenz gesprochen. Laut Ankündigungen von Meta-Chef Mark Zuckerberg wird derzeit an einer solchen Superintelligenz gearbeitet. Unabhängige Wissenschaftler hingegen sehen die Entwicklung noch weit davon entfernt (vgl. Stieler/Zweig 2024).

Bekannte KI-Modelle

Die Firma OpenAI wurde 2015 gegründet und bietet neben **ChatGPT** für Text mit **DALL-E** auch ein Tool an, das Bilder aus Textbeschreibungen generieren kann. Mit **Sora** kommt bald noch ein Werkzeug hinzu, das Videos aus Texteingaben erstellen kann. Mit ChatGPT 3.5 gibt es eine kostenfreie Version. Das leistungsfähigere Sprachmodell ChatGPT 4.0 liegt dem kostenpflichtigen Angebot ChatGPT+ zugrunde. Tatsächlich liefert die kostenlose Variante weit weniger eloquente Ergebnisse und macht mehr Fehler als ChatGPT 4.0. Mit ChatGPT 4o gibt es seit Mai 2024 zudem ein drittes Modell, das Text, Video und Audio verarbeiten kann.

Software-Anbieter Microsoft hat seit 2019 mehr als 13 Milliarden US-Dollar in OpenAI investiert und sich so Zugang zu dessen Technologie gesichert. Daher beruht das KI-Angebot von Microsoft, das unter dem Namen **Copilot** bekannt ist, auch auf ChatGPT. Copilot kann direkt aus Microsoft-Anwendungen wie Word, Excel oder PowerPoint aufgerufen werden. Darüber hinaus kooperiert Microsoft mit Meta. Und außerdem entwickelt das Unternehmen ein eigenes KI-Sprachmodell namens **MAI-1**, das mit den Modellen von Google und OpenAI konkurrieren soll.

Der Facebook-Mutterkonzern Meta Platforms hat im Februar 2023 sein Sprachmodell **LLaMA** veröffentlicht. Die Abkürzung steht für Large Language Model Meta AI. Der Code, den das Modell bei der Ausführung von Anfragen verwendet, ist frei zugänglich. Im Juli 2023 ist in Kooperation mit Microsoft das Nachfolgemodell LLaMA 2 erschienen, das ebenfalls Open Source ist. Neben Text und Bildern kann LLaMa 2 auch Code generieren. Im April 2024 stellte Meta LLaMA 3 vor.

Google hat als Reaktion auf den Release von ChatGPT im Februar 2023 den Chatbot **Bard** veröffentlicht, den das Unternehmen im Februar 2024 in **Gemini** umbenannt hat. Die Modelle, die dem zugrunde liegen, sind unterschiedlich in Größe und Funktionsumfang und teils kostenpflichtig. Die aktuelle Version kann verschiedene Arten von Inhalten, wie Text, Bilder, Audio und Video, verarbeiten. Insgesamt sind die Funktionen denen von OpenAI sehr ähnlich. Gemini gibt es auch als App.

Anthropic ist ein 2021 von ehemaligen OpenAI Mitarbeitenden gegründetes Unternehmen, das sich der Entwicklung einer sicheren und verständlichen KI verschrieben hat. Sein Sprachmodell **Claude** kann Texte generieren und fokussiert sich dabei besonders auf KI-Sicherheit. Die im März 2024 veröffentlichte Version Claude 3 umfasst drei Modelle, die auf verschiedene Kundentypen und Anwendungsfälle angepasst sind. Lange war Claude nur in den USA und Großbritannien verfügbar, inzwischen gibt es das Angebot auch in der EU. Hauptzielgruppe sind zwar Unternehmen, dennoch ist Claude auch für Einzelpersonen und die allgemeine Nutzung zugänglich. Im Juni 2024 stellte Anthropic mit Claude 3.5 Sonnet ein Sprachmodell vor, das sowohl Text als auch Bilder analysieren kann.

Midjourney ist ein unabhängiges Forschungslabor aus San Francisco. Es wurde im Juli 2022 gegründet und hat sich auf die Generierung von Bildern aus Textbeschreibungen spezialisiert. Zugänglich ist das gleichnamige Tool über einen Chatbot auf der Plattform Discord. Die Anwendung ist kostenpflichtig und ermöglicht es Nutzenden, über die Eingabe von Textanweisungen, Bilder zu generieren. Midjourney zeichnet sich dadurch aus, dass es spezifische künstlerische Stile und Bildinterpretationen ermöglicht und sehr fotorealistische Bilder produzieren kann.

Die Firma Stability AI, die 2019 von Emad Mostaque in London gegründet wurde, arbeitet ebenfalls vor allem an der Entwicklung von bildgenerierenden KI-Anwendungen. Flaggschiff ist das Tool **Stable Diffusion**, dessen Code von einem Team der Universität München entwickelt und frei zugänglich veröffentlicht wurde. Stability AI unterscheidet sich von DALL-E und Midjourney durch seinen Open-Source-Ansatz, der Entwicklerinnen und Entwicklern ebenso wie Unternehmen einen leichteren Zugang sowie die Möglichkeit bietet, das Modell an spezifische Bedürfnisse anzupassen. Inzwischen lassen sich mit diesem Werkzeug zum Teil auch schon kurze Video-Sequenzen und animierte Bilder erstellen.

Die deutsche Firma Aleph Alpha arbeitet seit 2019 an einem großen Sprachmodell mit dem Namen **Luminous**. Es ist noch in der Entwicklung und wird derzeit nicht für die breite Öffentlichkeit angeboten, vielmehr richtet es sich an Unternehmen und Organisationen. Ziel von Aleph Alpha ist es, die Sicherheit und Anwendbarkeit von KI im europäischen Kontext zu erhöhen, indem sie die Einhaltung von Datenschutzbestimmungen garantieren und maßgeschneiderte Lösungen für lokale Märkte bieten. Zudem möchte Aleph Alpha für die Nutzenden nachvollziehbar machen, wie die KI Luminous zu ihrer Entscheidung kommt. Erste Kunden sind beispielsweise Bosch, SAP und verschiedene Kommunalverwaltungen.

Das französische Start-up Mistral ist gerade mal ein Jahr alt und derzeit die große Hoffnung der europäischen KI-Szene. Gegründet wurde es im April 2023 von drei Software-Entwicklern: Guillaume Lample und Timothée Lacroix waren früher beim Facebook-Mutterkonzern Meta an der Entwicklung von LLaMA beteiligt und Arthur Mensch war Mitarbeiter von Deepmind, der KI-Tochter von Google. Seinen Sitz hat Mistral in Paris, investiert wird inzwischen aber größtenteils aus den USA – darunter Wagniskapitalgeber Andreessen Horowitz, der Software-Konzern Salesforce und Microsoft. **Le Chat** heißt das Angebot von Mistral, dessen Betaversion derzeit nach einer Registrierung nutzbar ist. Wie Aleph Alpha richtet sich auch Mistral primär an Unternehmen und Organisationen. Zur Kundschaft gehören die Bank BNP Paribas und der Mobilfunkanbieter Orange.



KI verantwortungsvoll nutzen

Ende 2023 kündigte das US-amerikanische Start-up *Channel 1* mit einem Werbevideo seinen Markteintritt für das Jahr 2024 an. Als einer der weltweit ersten Nachrichtensender will *Channel 1* seine Produktionsprozesse vollständig auf Künstliche Intelligenz ausrichten. Menschen liefern demnach hauptsächlich das Material und prüfen die Fakten. Den Rest übernimmt die KI. Im Video erklärt ein künstlich generierter Avatar, was das bedeutet: »Alles, was Sie auf *Channel 1* sehen, nutzt KI, um Ihnen Nachrichten auf die Art und Weise zu präsentieren, wie Sie es wünschen – personalisiert, lokalisiert und auf Sie zugeschnitten.« Weiter sagt der Avatar: »Technologie ist zwar das Werkzeug, trotzdem orientieren wir uns in allem an den Grundwerten journalistischer Sorgfalt und Verantwortung.«

Ist das die Zukunft des Journalismus? Ist es ethisch vertretbar – ganz, in Teilen oder gar nicht? Ist es die Zukunft, die wir uns wünschen?

Um KI verantwortungsvoll einzusetzen und um in einer Arbeitsumgebung mit KI bestehen zu können, braucht es Wissen auf verschiedenen Ebenen. Neben einem soliden Verständnis der Funktionsweise der Technologie sind auch Kenntnisse der rechtlichen Rahmenbedingungen und presseethischen Aspekte notwendig. Zudem ist es hilfreich zu verstehen, wie Verzerrungen und Voreingenommenheit von KI-Anwendungen entstehen. Bisher wenig diskutiert, aber auch ein wichtiger Punkt für den verantwortungsvollen Umgang, ist aus unserer Sicht die Frage, welche Umweltbelastungen Training und Einsatz dieser Technologie mit sich bringen.

Recht

Derzeit gibt es mit Blick auf KI mehr ungeklärte Rechtsfragen als geklärte. Im Fokus steht daher unter Juristinnen und Juristen oft die Frage, was alles noch geregelt werden sollte. Auf Basis der schon jetzt geltenden Gesetze lassen sich aber dennoch schon Grundsätze und Tendenzen aufzeigen. Im Folgenden geben wir einen knappen Überblick über wichtige Aspekte.

KI-Verordnung oder die Frage: Wie hoch ist das Risiko?

Im März 2024 hat das Europaparlament mit der KI-Verordnung das weltweit erste KI-Gesetz beschlossen, das für mehrere Länder gilt. International firmiert es als AI-Act und gilt unmittelbar in allen EU-Mitgliedsstaaten. Für die alltägliche redaktionelle Arbeit wird es jedoch erst mal keine unmittelbare Rolle spielen. Zum einen läuft bis zum Inkrafttreten im Frühjahr 2026 noch eine Übergangszeit, während der sich alle auf die Einhaltung der Vorgaben vorbereiten können. Zum anderen werden auch von 2026 an Journalistinnen und Redakteure eher wenig mit der EU-Verordnung in Berührung kommen. Denn das Gesetz richtet sich in erster Linie an Entwickelnde, Herstellende und Anbietende von KI-Systemen.

Zentraler Regelungsansatz ist, dass KI-Systeme in Risikoklassen eingeteilt werden. Je nach Risikoklasse müssen dann entsprechende Sicherheitsvorkehrungen getroffen werden. Im journalistischen Alltag wichtiger sind derzeit noch Fragen, die sich um Urheberrecht und Datenschutz drehen.

Datenschutz oder die Frage: Gilt das Medienprivileg?

Datenschutzrechtliche Fragen können sich schon bei der Eingabe eines Prompts stellen, wenn dieser personenbezogene Daten enthält. Auch bei der weiteren Nutzung von KI im redaktionellen Alltag kann Datenschutz relevant werden. Denn die Datenschutz-Grundverordnung (DSGVO) schützt personenbezogene Daten, also alle Informationen, die sich auf eine Person beziehen und direkt mit ihrem Namen verknüpft sind oder ihm zugeordnet werden können. Personen, die diese erheben oder verarbeiten, haben deshalb verschiedene gesetzliche Verpflichtungen.

Allerdings kann im Journalismus eine interessante Ausnahme zum Tragen kommen – das sogenannte Medienprivileg. Es besagt, dass

Datenschutzbestimmungen bei der journalistischen Arbeit von der Recherche bis zur Veröffentlichung größtenteils nicht gelten. Warum? Ohne das Medienprivileg wäre die freie Medienberichterstattung in manchen Fällen gefährdet oder sogar unmöglich – wenn etwa im Rahmen einer investigativen Recherche wegen Datenschutzes erst eine Einwilligung der betroffenen Person eingeholt werden müsste.

Das Medienprivileg gilt jedoch nur unter der Bedingung, dass die betroffenen personenbezogenen Daten zu journalistischen Zwecken genutzt und verarbeitet werden. Und auch wenn für die Verarbeitung der Daten eine Ausnahme besteht, so sind andere Regeln wie das Persönlichkeitsrecht dennoch einzuhalten. Auch gilt das Medienprivileg nicht außerhalb der eigentlichen journalistischen Arbeit – die Distribution von Inhalten, wie der Versand eines Newsletters über personalisierte Angebote, fällt zum Beispiel nicht darunter.

Wollen Journalistinnen und Journalisten im Alltag KI nutzen, so ist der Datenschutz hier also erst mal keine grundsätzliche Hürde, wenn sie sich bei der Recherche und Erarbeitung von Beiträgen auf das Medienprivileg berufen können.

Inwieweit diese Ausnahmeregelung bei der zukünftig noch umfassenderen Nutzung von KI im Journalismus gelten wird, ist noch nicht final geklärt. Auch hier wird es vermutlich darauf ankommen, inwieweit die Nutzung den eigentlichen journalistischen Zwecken dient oder nicht.

Urheberrecht oder: Wie viel menschliche Kreativität steckt drin?

Wer das Urheberrecht an einem Werk hat, darf darüber entscheiden, ob, wo und auf welche Weise es veröffentlicht werden darf und zu welchem Preis. Daraus ergeben sich durch generative KI-Anwendungen neue Herausforderungen. Den Aspekt, wie journalistische Inhalte zum Training von KI genutzt werden, beleuchten wir im Kapitel »Training von KI-Modellen«. Aber auch im praktischen Einsatz gibt es neue Fragen: Kann man mit der Veröffentlichung von KI-generierten Inhalten gegen Urheberrechte verstoßen? Und wie ist es, wenn bei der Produktion von Inhalten KI geholfen hat – ist dann der Content zugunsten der Medienschaffenden trotzdem noch urheberrechtlich geschützt?

Im Zentrum all dieser Fragen stehen zwei zentrale juristische Aspekte: Hat ein Mensch es erstellt und ist es wirklich kreativ? Nur wenn diese beiden Faktoren zutreffen, gilt überhaupt das Urheberrecht.

Da durch das Urheberrecht nur menschliches Schaffen geschützt wird, sind von KI völlig autonom und automatisiert erstellte Inhalte wie Texte, Bilder oder Videos nicht urheberrechtlich geschützt.

Spuckt die Maschine aber einen bestimmten Text oder ein bestimmtes Foto quasi unverändert von einem Original aus, das urheberrechtlich geschützt ist, liegt das Urheberrecht noch bei der Person, die das Original erstellt hat. Entsprechend müssen dann Lizenzgebühren gezahlt werden. Dieser Fall kommt verhältnismäßig selten vor. Dies zu wissen und aufmerksam zu bleiben, ist hier aber wichtig.

Eine praktische Frage, die zunehmend relevanter werden wird, ist die folgende: Kann ich einen Beitrag, für den ich KI verwendet habe, bei der Verwertungsgesellschaft anmelden? Das geht nämlich nur, wenn ich selbst Urheberin oder Urheber bin.

Wenn KI lediglich Werkzeug, Hilfsmittel und Unterstützung des kreativen journalistischen Schaffensprozesses war und die eigentlich kreative Tätigkeit durch einen Menschen erfolgte, kann der Beitrag grundsätzlich urheberrechtlich geschützt sein.

Aber egal, ob KI im Spiel ist oder nicht – es gibt noch eine zweite Bedingung. Juristen haben sich dafür den Begriff der Schöpfungshöhe ausgedacht. Diese liegt vor, wenn ein Text, eine Grafik, ein Video oder ein anderer Beitrag in ausreichendem Maße individuell, einzigartig und kreativ ist. Bei einer Nachricht als bloße Wiedergabe von Fakten ohne gestalterische Elemente ist das eben nicht der Fall.

Doch wann ist die Schwelle zur »Schöpfungshöhe« überschritten, damit das Urheberrecht überhaupt gilt? Die Übergänge sind hier fließend und es lassen sich keine starren Abgrenzungen vornehmen. Diese Frage kann jeweils nur im Einzelfall in Bezug auf den konkreten journalistischen Inhalt beantwortet werden. Eine Daumenregel allerdings gibt es schon: Je mehr ein Mensch den Inhalt geistig-kreativ gestaltet hat, umso eher ist Urheberschutz gegeben.

Sorgfalt und Transparenz

»KI-Journalismus geht schief, wenn er unkontrolliert, faul, egoistisch, unehrlich und undurchsichtig ist«, schreibt Zach Seward, leitender KI-Experte der *New York Times* am Anfang eines langen Beitrags (2024). Doch unter welchen Rahmenbedingungen kann der verantwortungsvolle Einsatz von KI im Journalismus gelingen? Das, so Seward, müsse sehr gründlich geprüft werden. In jedem Fall solle der Einsatz von KI im Journalismus »davon motiviert sein, was das Beste für das Publikum ist«. So müssten unbedingt die Grundprinzipien des Journalismus gelten.

Die zwei zentralen Grundprinzipien in diesem Kontext sind die journalistische Sorgfalt und Transparenz. Auch wenn der Begriff Sorgfaltspflicht eher unklar und schwammig klingt, so kann ein Verstoß dagegen auch rechtliche Konsequenzen haben. Ist aber nachweisbar, dass Sorgfaltspflichten eingehalten wurden, kommen Gegenansprüche auf Unterlassen oder gar Schadenersatz sowie eine strafrechtliche Verfolgung nicht in Betracht.

Als Hüter ethischer Standards in den Medien in Deutschland hat sich der Deutsche Presserat mit dem Thema beschäftigt. In seinem Jahresbericht für 2023 heißt es: »Der Pressekodex verpflichtet Redaktionen bereits jetzt zur Einhaltung der berufsethischen Standards unabhängig davon, ob ein Text von einem Menschen oder einer Maschine stammt« (Deutscher Presserat 2024). Aus diesem Grund und mit Blick auf die Dynamik der KI-Entwicklung gebe es aktuell nicht den Bedarf, den Pressekodex anzupassen und beispielsweise eine Kennzeichnungspflicht von KI-Inhalten einzuführen. Trotzdem dürften KI-generierte Inhalte die Leserschaft nicht in die Irre führen.

Für Bilder, die mit KI-Tools generiert werden, gebe es zudem schon jetzt eine Kennzeichnungspflicht. Von KI erzeugte Bilder sind demnach von der Regelung zu Symbolbildern umfasst und müssen als solche gekennzeichnet werden. Der Schweizer Presserat (vgl. 2024) hingegen hat seinem Journalistenkodex schon einen Leitfaden für den Umgang mit Künstlicher Intelligenz hinzugefügt.

Ein Medienhaus, das eher vorprescht, ist der Kölner Verlag M. DuMont Schauberg. Die Nachricht, dass neuerdings Texte von Klara Indernach – Kürzel: KI – in seinen Medien erscheinen sollten, ließ im Sommer 2023 aufhorchen. »Der *Express* tut zumindest in Teilen so, als wäre die KI ein Mensch«, kommentierte die Tageszeitung *taz* dies im September. Ethisch vertretbar oder nicht? In vielen Fällen werden Fragen zur Transparenz eine Gratwanderung bleiben. Was sollte deklariert werden? Und

wo? Reicht es, dass der Hinweis auf die Produktion von Texten durch eine KI nicht direkt am Beitrag sondern auf einer anderen Webseite steht – analog zu den Kurzbiografien tatsächlicher Personen?

Im Fall Klara Indernach handelte sich der Kölner *Express* 2023 eine Missbilligung des Presserats ein. Allerdings nicht für sein Vorgehen bei der Texterstellung oder der Platzierung des Transparenzhinweises. Sondern, weil er zunächst nicht darauf hingewiesen hatte, dass das Bild der Autorin mit Midjourney erstellt worden war.

Eine erste Rüge für einen KI-generierten Text hat der Deutsche Presserat 2023 gegenüber der Zeitschrift *Die Aktuelle* ausgesprochen. Sie hatte ein angebliches Interview mit dem ehemaligen Formel-1-Rennfahrer Michael Schumacher veröffentlicht und es auf der Titelseite als »erstes Interview« seit dessen tragischem Skiunfall 2013 angekündigt. Die Kennzeichnung als KI-generiertes Interview im hinteren Teil des Beitrags war hier nicht ausreichend. Der Presserat sah darin eine »schwere Irreführung der Leserschaft« und eine mögliche Verletzung des Persönlichkeitsrechts von Schumacher.

Neben Grundsatzentscheidungen zu fehlenden Kennzeichnungen bei Bildern und dem Schumacher-Interview gab es bis Redaktionsschluss keine weiteren Entscheidungen des Deutschen Presserates. Auch gab es laut einer Sprecherin bis Mai 2024 keine sonstigen nennenswerten Beschwerden. Über das zukünftige Vorgehen berät derzeit eine Arbeitsgruppe. Der Schweizer Journalist und Kommunikationswissenschaftler Colin Porlezza (2024) hält das für zu zögerlich und verweist auf Richtlinien, die der Europarat entwickelt hat.

In seinen Guidelines für eine verantwortungsvolle Handhabung von KI-Systemen im Journalismus (vgl. Council of Europe 2024) betont der Europarat die Chancen und das Innovationspotenzial von KI für den Journalismus und macht zugleich Vorschläge, wie Medienschaffende, Medienorganisationen, Technologieanbieter und Nationalstaaten mit dem Thema umgehen sollten. Dabei erinnert er daran, dass das Recht auf freie Meinungsäußerung, das auch in Artikel 10 der Europäischen Menschenrechtskonvention verbrieft ist, mit Verantwortung gegenüber der Gesellschaft einhergeht. Journalistische KI-Systeme in Übereinstimmung mit Werten einzusetzen, sei jedoch eine schwierige Aufgabe, auf die es keine vorgefertigten Antworten gebe, heißt es ebenso in dem Papier. Die eigentliche Herausforderung läge daher in der Ausgestaltung konkreter Anwendungen.

Interne Leitlinien

Wie dürfen Redaktionen KI nun also einsetzen? Und was ist ethisch nicht vertretbar? Die neuen KI-Tools haben dieser Frage Brisanz und Aktualität verliehen. Deshalb hat die international tätige Nicht-regierungsorganisation Reporter ohne Grenzen, die sich für Pressefreiheit einsetzt, am 10. November 2023 in Paris gemeinsam mit 16 Partnerorganisationen eine Charta zu KI und Journalismus verabschiedet.

Im Kern betrachtet die Paris-Charta (ROG 2023) Künstliche Intelligenz als Werkzeug mit einer bedeutenden Besonderheit. Diese liegt darin, dass das Werkzeug eigene Entscheidungen trifft und dadurch auch die öffentliche Aufmerksamkeit lenken kann. Darum gelte es, den Einsatz dieses Werkzeugs konstant zu hinterfragen, zu überwachen und transparent zu machen. Die Friedensnobelpreisträgerin und Vorsitzende der Charta-Kommission, Maria Ressa, sagte anlässlich der Unterzeichnung: »Technologische Innovation führt nicht per se zu Fortschritt: Sie muss von der Ethik gesteuert werden, um der Menschheit wirklich zu nutzen.«

Konkret werden in der Charta zehn Grundsätze zum journalistischen Umgang mit Künstlicher Intelligenz formuliert. Sie richten sich an Medienorganisationen ebenso wie an Journalistinnen und Journalisten. So fordert die Charta Unternehmen auf, »das menschliche Handeln in den Vordergrund« zu stellen und »Herkunft und Rückverfolgbarkeit von Inhalten« zu garantieren. Des Weiteren seien die Medienunternehmen »immer verantwortlich für die Inhalte, die sie veröffentlichen«.

Darüber hinaus sollten sowohl Unternehmen als auch einzelne Medienschaffende laut der Charta eine klare Trennlinie zwischen synthetischen und authentischen Inhalten ziehen und »eine aktive Rolle bei der Steuerung von KI-Systemen spielen«. Handlungsleitend sollte stets ihre Hauptaufgabe sein, das Recht aller Menschen auf qualitativ hochwertige, vertrauenswürdige Informationen zu gewährleisten.

Die Paris-Charta von Reporter ohne Grenzen ist nicht die erste Richtlinie ihrer Art. Zahlreiche Medienorganisationen hatten zuvor bereits eigene Codes of Conduct formuliert. In Deutschland gab sich als eine der ersten der *Bayerische Rundfunk* im November 2020 KI-Richtlinien (vgl. BR 2020). Von 2023 an gab es dann einen enormen Zuwachs. Eine weltweite Vergleichsanalyse aus September 2023 (Becker et al. 2023) zeigt, dass viele Unternehmen in einigen zentralen Aspekten ein ähnliches Verständnis von einem verantwortungsvollen Umgang mit Künstlicher Intelligenz entwickelt hatten.

Transparenz, die Verantwortung für KI und eine Überwachung von KI-generierten Inhalten und Entscheidungen sind gemeinsamer Konsens. Die untersuchten Medienorganisationen schreiben in ihren Kodizes ausdrücklich, dass sie Journalistinnen und Journalisten nicht durch KI ersetzen wollen. Einige heben zudem hervor, dass KI rassistischen, sexistischen oder anderweitig menschenfeindlichen Dateninput reproduzieren kann.

Die Autoren der Studie erwähnen allerdings auch blinde Flecken in den Branchenrichtlinien. Sie monieren beispielsweise, dass einige Richtlinien nicht konkret genug seien. So sei zum Beispiel unklar, ob die selbst auferlegte Kennzeichnungspflicht nur für vollständig KI-generierte Texte und Bilder gelte oder auch für Titel und Teaser, bei denen KI lediglich zu Rate gezogen wurde. Probleme wie die menschenunwürdige Bezahlung von Hilfsarbeitenden bei der Kuratierung der Trainingsdaten und dem Training der KI oder die ökologischen Schäden durch den enormen Wasserverbrauch von KI-Servern fehlen demnach in allen untersuchten Richtlinien.

Laut einer Studie der Landesanstalt für Medien NRW vom März 2024 (vgl. Nennstiel/Isenberg 2024) wünscht sich eine große Mehrheit der Bevölkerung eine unabhängige Aufsichtsbehörde und strenge, verbindliche Regeln für den Einsatz von KI im Journalismus. Ein Teil der Befragten lehnt KI im Journalismus sogar gänzlich ab. Gerade im Zusammenhang mit politischer Berichterstattung herrscht Skepsis.

Aktuell plant laut der Vergleichsstudie von Becker, Simon und Crum kaum ein Verlag, Mediennutzende in die Entwicklung von Handreichungen zum Umgang mit Künstlicher Intelligenz im Journalismus einzubeziehen. Das könnte allerdings wichtig sein für Medienhäuser, um die Glaubwürdigkeit ihrer Medienmarke und das Vertrauen in ihre Berichterstattung zu bewahren.

Tipps für die Entwicklung eines KI-Kodex

BEZUGSRAHMEN

Für welche Teile des Unternehmens soll der KI-Kodex gelten? Den Anwendungsbereich bewusst abzustecken, ist ein wichtiger Aspekt. Bislang wurden ausschließlich die Auswirkungen von KI auf die Arbeit von Journalistinnen und Journalisten behandelt. Allerdings werden zukünftig in vielen Medienhäusern auch Abteilungen für Presse- und Öffentlichkeit, Personal oder der Vertrieb Künstliche Intelligenz verwenden. Unternehmen sollten sich überlegen, welche Regeln eines Kodex auch für diese Bereiche gelten.

KODEX-VORLAGEN

Welche Regeln geben sich vergleichbare Unternehmen? Zu Beginn des Prozesses sollten sich Verantwortliche einen Überblick über den Status quo verschaffen. Je besser sich einzelne Kodizes aneinander orientieren, desto eher kommt die Branche zu einheitlichen Standards. Die Paris-Charta ist ein guter Ausgangspunkt. Da sich das KI-Feld aber schnell weiterentwickelt, ist es wichtig, auf möglichst aktuelle Ansätze zurückzugreifen. Das Poynter Institute beispielsweise bietet eine Vorlage an, die regelmäßig aktualisiert wird (vgl. Poynter Institute 2024).

RISIKOMANAGEMENT

Wo gibt es ethische Herausforderungen im Workflow? Dass ein KI-spezifischer Code of Conduct überhaupt nötig ist, hängt damit zusammen, dass Künstliche Intelligenz einen negativen Einfluss auf die journalistische Arbeit haben kann. Für einen guten KI-Kodex sollten Medienhäuser diese möglichen negativen Einflüsse deshalb systematisch identifizieren und Gegenstrategien formulieren.

EINBEZIEHUNG DER ÖFFENTLICHKEIT

Was meinen die Nutzenden? Die breite Öffentlichkeit hat konkrete Anforderungen an den Einsatz von KI im Journalismus. Darum sollten Medienorganisationen auf der Suche nach einem KI-Kodex eine Plattform etablieren, über die die Öffentlichkeit am Prozess beteiligt wird. Dies kann zudem die notwendige Transparenz schaffen, die für die Glaubwürdigkeit und das Vertrauen in Medienorganisationen und ihre Angebote notwendig ist.

KONTROLLE UND WEITERENTWICKLUNG

Sind unsere Leitlinien noch zeitgemäß? Da KI sich sehr schnell weiterentwickelt, kann das aktuelle KI-Niveau schon bald überholt sein. Deshalb sollten sich auch die Kodizes weiterentwickeln. Der Deutsche Journalisten-Verband schlägt dazu vor, dass Medienunternehmen Beauftragte benennen, die die Praxis des Einsatzes von KI im Haus mit den jeweils gültigen externen und internen Regeln abgleichen. Zudem könnten sie auch als Ansprechpersonen für Beschwerden dienen (DJV 2023). Sinnvoll ist darüber hinaus, regelmäßig zu hinterfragen, ob einzelne Aspekte der internen Regeln überarbeitet werden müssen.

Training von KI-Modellen

Biases

Zur verantwortungsvollen Nutzung von KI-Systemen gehört aus unserer Sicht auch das Bewusstsein, dass Werkzeuge, die auf KI basieren, möglicherweise ein einseitig geprägtes Weltbild transportieren. Denn obwohl sich die Strukturen der Lernalgorithmen und die Detailtiefe, mit der sie Daten prozessieren können, im Laufe der Zeit immer weiterentwickelt haben, ist doch ein Aspekt von Maschinellen Lernen immer gleich geblieben: In ihrem Lernprozess analysieren die Lernalgorithmen große bis gigantisch große Datenmengen. Diese Phase des Lernens ist hoch relevant. Denn hier liegt auch die grundlegende Anfälligkeit von KI für eine mögliche Voreingenommenheit (»bias«).

Biases können sich dabei sowohl aus den Trainingsdaten selbst ergeben als auch durch das Feedback, das Menschen der Maschine im Trainingsprozess geben. Die Voreingenommenheit kommt dabei jedoch nicht von alleine zustande, sondern beruht auf menschlichen Entscheidungen, die vom Weltbild der Handelnden geprägt sind und die die bestehenden gesellschaftlichen Verhältnisse widerspiegeln. Ein zentraler Punkt hierbei ist die Auswahl der Daten für den Trainingsprozess. Welche Datensätze werden für das Training verwendet? Welche Bestandteile bei der Bereinigung herausgelöscht?

Internetinhalte spielen beim Zusammenstellen von Trainingsdatensätzen eine enorm wichtige Rolle. Über die genauen Inhalte wissen jedoch auch Expertinnen und Experten nicht genau Bescheid (vgl. Burgess/Rogers 2024 und Bender et al. 2021). Klar ist nur, dass sie auf gigantischen Datenmengen beruhen (vgl. Bender et al.).

Ein bekannter Datensatz nennt sich Common Crawl. Nach Aussage des Crawlingdienstes sind in ihm »Petabyte an Daten« (Common Crawl 2024) vereint – allein ein Petabyte entspricht einer Million Gigabyte. Diese Daten wurden seit 2008 gesammelt, sie beinhalten Webseiten, Metadaten und Textauszüge. In bereinigter Form wurden diese Daten auch für das Training von ChatGPT eingesetzt.

Doch wer produziert überwiegend Inhalte im Internet? Es sind eher Personen aus der Wohlstandsgesellschaft als solche, deren Alltag sich um existenzielle Fragen des täglichen Überlebens dreht. Eine sehr pointierte Kritik zum Umgang mit Trainingsdaten haben Bender et al. 2021 formuliert. Ihrer Ansicht nach würden mit der aktuellen Praxis eine hegemoniale Weltsicht sowie die Marginalisierung unterrepräsentierter Bevölkerungsgruppen in den Trainingsdaten festgeschrieben.

Sie kritisieren auch die Art und Weise, wie Datasets digital kuratiert werden. Ein Beispiel: Die Praktik, alle Beiträge mit Wörtern, die im Bezug zu Sex stehen, zu unterdrücken, hat ihnen zufolge auch ausgrenzende Nebenwirkungen. Denn auch, wenn dadurch in erster Linie Obszönitäten und pornographische Inhalte herausgefiltert werden sollen, hat diese grobe Art der Datensäuberung auch zur Folge, dass viele Beiträge der LGBTQ-Community entfernt werden. Und damit eben auch ihre Weltsicht.

Weiterhin kritisieren die Autorinnen, dass die Datasets für das Training von Sprachmodellen zwar immer größer werden, zugleich aber nicht dokumentiert wird, welche Daten sie beinhalten oder um welche Daten diese Sets bereinigt wurden. Zudem würden einmal generierte Sets gern immer wieder verwendet, wodurch Ungerechtigkeiten und Marginalisierungen festgeschrieben würden, ohne dass dies überprüft werden könne.

Für mehr Transparenz bei der Entwicklung von KI-Basismodellen setzt sich das Stanford Center for Research on Foundation Models (vgl. CRFM 2023) ein. Seine Wissenschaftlerinnen und Wissenschaftler haben einen Transparenz-Index entwickelt, der Standards für Offenheit und Zugänglichkeit in der Entwicklung von KI-Basismodellen setzt. Außerdem macht der AI-Act Entwicklerinnen und Entwicklern von Technologie Vorgaben, die sie ab 2026 erfüllen müssen.

Der Kampf um das Urheberrecht

Die fehlende Transparenz der Trainingsdaten führt auch zu urheberrechtlichen Fragen. Denn viele der in den Trainingsdaten erfassten Inhalte sind zwar im Internet frei zugänglich, aber möglicherweise dennoch urheberrechtlich geschützt. Doch dafür haben sich Datenfirmen beim Durchsuchen großer Teile des Internets erst mal weniger interessiert. Und von den Erstellerinnen und Erstellern der Inhalte wurde das lange Zeit kaum wahrgenommen.

Aktuell gilt grob gesagt: Widerspricht eine Rechteinhaberin oder ein Rechteinhaber der Nutzung seiner Inhalte nicht durch einen entsprechenden Vermerk auf der Webseite, dann ist es grundsätzlich erlaubt. Machen die Betreibenden eines Online-Angebotes aber Gebrauch von der sogenannten Opt-out-Möglichkeit, dann dürfen ihre Inhalte nicht von Maschinen digital durchsucht und vervielfältigt werden. Wie das geht, erklären Burgess und Rogers (2024) im Digitalmagazin *Wired*.

Medienunternehmen gehen mit dieser Situation sehr unterschiedlich um: Der deutsche Axel Springer Verlag und die amerikanische Nachrichtenagentur Associated Press haben beispielsweise Verträge mit

OpenAI geschlossen: Sie erlauben, dass ihre Inhalte genutzt werden, lassen sich das aber auch bezahlen. Andere sehen das hingegen völlig anders. Die *New York Times* zum Beispiel sieht ihr Geschäftsmodell bedroht und hat gegen die Macher von ChatGPT geklagt. Sie fordert mehrere Milliarden Dollar Schadensersatz.

Wer sich darüber informieren möchte, ob Inhalte seines Medienunternehmens oder der eigenen Webseite für Trainingszwecke verwendet wurden, kann das über die Webseite »Have I been trained?« (<https://haveibeentrained.com/>) von der Firma Spawning machen. Schon eine schnelle Suche zeigt: Inhalte vieler deutscher Medienunternehmen finden sich in den Trainingsdaten für KI. Auch hier gibt es für einzelne Inhalte die Möglichkeit eines Opt-outs.

Auswirkungen auf die Umwelt

KI verbraucht große Mengen an Wasser, wertvollen Rohstoffen und Strom. Eine verantwortungsvolle Nutzung von KI sollte deshalb auch die ökologische Nachhaltigkeit berücksichtigen. Doch anders als das menschliche Gehirn sind die für KI bislang genutzten Computersysteme bei weitem noch nicht energie- und ressourceneffizient. Während unser Gehirn nur eine Leistungsaufnahme von etwa 20 Watt hat und damit etwa so viel Energie wie ein Laptop verbraucht, werden für KI-Lösungen neue fußballfeldgroße Rechenzentren gebaut. Jedes einzelne von ihnen hat den Stromverbrauch einer Großstadt. Dennoch reicht die KI bislang noch nicht an die Leistungsfähigkeit des menschlichen Gehirns.

Wie viel Strom und andere Ressourcen verbraucht KI schon jetzt? Wie wird sich das in Zukunft entwickeln? Und welche Möglichkeiten gibt es, auch im KI-Zeitalter möglichst sparsam mit den natürlichen Ressourcen umzugehen? Die Wissenschaft hat noch keine sehr genauen Daten zu den ökologischen Auswirkungen der KI-Nutzung, denn die großen Tech-Konzerne geben Informationen zu den von ihnen genutzten KI-Systemen und deren Energie- und Ressourcenbedarf nur sehr zögerlich bekannt. Doch auch ohne konkrete Zahlen der Unternehmen lassen sich qualifizierte Schätzungen zum Ressourcenverbrauch abgeben. Forschende nähern sich dem über eine differenzierte Betrachtung des Gesamtverbrauchs an. Unterschieden wird dabei zwischen der Modellbildung, dem Training und der eigentlichen Anwendung von KI-Modellen.

Jährlicher Strombedarf größer als der von Dänemark

Mit rund 80 Prozent wird der Großteil des Stroms nach aktuellen Erkenntnissen bei der Anwendung von KI, der sogenannten Inferenz, verbraucht – der Rest entfällt auf Modellbildung und Training. Vor allem bei den öffentlich zugänglichen Modellen wie ChatGPT ist der Anteil des Verbrauchs durch die Nutzung besonders hoch. Experten schätzen, dass allein der Betrieb des im November 2022 veröffentlichten ChatGPT 3.0 mehr als 500.000 Kilowattstunden Strom am Tag benötigt. Damit könnten 150 Haushalte ein Jahr mit Strom versorgt werden.

Anders sieht es bei KI-Anwendungen aus, die nicht für die breite Öffentlichkeit bestimmt sind, wie zum Beispiel das frühzeitige Erkennen von Fehlern oder Risiken in der industriellen Produktion. Hier ist die Zahl der Nutzenden relativ gering. Daher wird bei solchen Anwendungen meist der Ressourcenbedarf in der Trainingsphase im Verhältnis zum späteren Verbrauch bei ihrer Nutzung dominieren.

Der jährliche weltweite Stromverbrauch aller KI-Anwendungen in Rechenzentren wird aktuell auf fast 40 Milliarden Kilowattstunden geschätzt (vgl. Avelar et al. 2023). Das ist mehr als ganz Dänemark in der gleichen Zeit an Strom benötigt. Und dabei ist der Verbrauch der Energie durch die Endgeräte, die für die Nutzung der KI zum Einsatz kommen, noch nicht berücksichtigt.

Energieverbrauch wird stark wachsen

Durch zunehmende Nutzungsmöglichkeiten wird der Bedarf durch die Anwendung von KI in Zukunft voraussichtlich noch stark steigen. Laut dem Marktforschungsunternehmen Gartner könnte KI bis zum Jahr 2030 für 3,5 Prozent des weltweiten Stromverbrauchs verantwortlich sein (vgl. Gartner 2022). Das wären – legt man die aktuellen Szenarien zur Entwicklung des Stromverbrauchs zugrunde – mehr als 1.000 Milliarden Kilowattstunden und doppelt so viel Strom wie ganz Deutschland aktuell pro Jahr an Strom benötigt.

Optimisten hoffen darauf, dass verbesserte Modelle und für spezielle Anwendungsfälle entwickelte Hardware den Anstieg des Stromverbrauchs in Grenzen halten können. Realistisch muss aber berücksichtigt werden, dass Effizienzverbesserungen in der IT bislang immer dazu geführt haben, dass noch mehr IT eingesetzt wurde. Dass es nun beim Einsatz von KI-Anwendungen nicht so kommen wird, ist sehr unwahrscheinlich. Schließlich sind die Einsatzmöglichkeiten schier unbegrenzt und die Entwicklung steht hier erst am Anfang.

Größe der Modelle bestimmt den Ressourcenbedarf

Das Ausmaß des immer weiter steigenden Ressourcenbedarfs von KI lässt sich auch gut am Stromverbrauch des Trainings ablesen. Schon 2019 errechneten amerikanische Forscher, dass das Training eines damaligen KI-Sprachmodells 656.000 Kilowattstunden Strom benötigen kann (Strubell et al. 2019). ChatGPT 3.0 hat im Jahr 2020 für das Training knapp 1,3 Millionen Kilowattstunden und damit schon fast doppelt so viel Strom verbraucht (vgl. Patterson et al. 2022). Und beim Modell ChatGPT 4.0 hat sich die Strommenge für das Training nach Schätzungen sogar auf 7,2 Millionen Kilowattstunden noch einmal fast versechsfacht (vgl. TRG Datacenters 2023). Damit könnten 1.800 Einfamilienhäuser ein Jahr mit Strom versorgt werden.

Neben elektrischem Strom braucht KI auch Unmengen an Wasser. Forscher der University of California schätzen, dass allein das Training von ChatGPT 3.0 rund 5,4 Millionen Liter Wasser benötigt hat. Das entspricht dem täglichen Wasserbedarf von mehr als 40.000 Personen in Deutschland. Hinzu kommt noch ein Wasserverbrauch von 1,6 Litern pro einhundert Anfragen bei ChatGPT. Der Großteil des Wasserverbrauchs entsteht bei der Erzeugung von herkömmlichem Strom sowie bei der Kühlung von Servern in Rechenzentren.

Produktion von Hardware frisst viele Ressourcen und Energie

Auch die Materialien der verwendeten Hardware verbrauchen Ressourcen – darunter Gold, Silber und andere seltene Elemente. Zudem hat die Herstellung von Computern und Servern einen nennenswerten Energiebedarf und produziert Treibhausgasemissionen. Lebenszyklusanalysen von aktuellen Servern zeigen, dass zwischen zehn und 40 Prozent des gesamten Energiebedarfs und der Treibhausgasemissionen auf die Herstellung der Geräte zurückzuführen sind.

Und die Hardware wird mehr. Nach Einschätzung von Analystinnen und Analysten wird sich das jährliche Volumen der KI-Hardwareverkäufe zwischen 2023 und 2033 um den Faktor 9 auf fast 500 Milliarden US-Dollar erhöhen (Precedence Research 2024). Beeindruckender Beleg für dieses Wachstum ist die Firma Nvidia, die Hochleistungschips herstellt. Anfang 2024 kam Nvidia mit der Produktion kaum hinterher. Aufgrund der stark gestiegenen Nachfrage ist der Chiphersteller zu einer der fünf wertvollsten Firmen der Welt geworden – gleich hinter Apple, Microsoft, Alphabet (Google) und Amazon.

Ideen für eine nachhaltigere Nutzung von KI-Anwendungen

Oft wird das Argument vorgebracht, dass KI-Systeme mit regenerativ erzeugtem Strom versorgt und die Umweltauswirkungen dadurch im Zaum gehalten werden können. Diese Argumentation ist aber nur teilweise stichhaltig. Zwar kann dies den negativen Einfluss der Technologie auf die Treibhausgasemissionen verringern. Der Blick muss aber auch darauf gelenkt werden, ob es genügend regenerative Energie gibt oder KI möglicherweise E-Autos, Wärmepumpen oder der Industrie umweltfreundlichen Strom wegnimmt. Klimafreundlich können die KI-Rechenzentren daher nur sein, wenn ihr Strom aus neu errichteten regenerativen Stromerzeugungsanlagen kommt.

Doch selbst 100 Prozent grüner Strom für die Rechenzentren, in denen KI-Prozesse laufen, reichen nicht, um die Umweltauswirkungen von KI zu neutralisieren. Denn auch dann bleibt noch der hohe Verbrauch von Wasser und Materialien für die Herstellung der notwendigen Geräte und Anlagen, die bei der Betrachtung der Auswirkungen auf die Umwelt stark ins Gewicht fallen.

Eine Möglichkeit besteht darin, die Modelle effizienter zu machen – zum Beispiel indem die Modelle mit weniger, aber qualitativ hochwertigeren Daten trainiert werden. In diesem Punkt unterscheiden sich Mensch und KI nur wenig. Je schlechter die Daten, also das Lehrmaterial, desto weniger wird gelernt. Zudem kann der bewusste Einsatz von KI helfen, Ressourcen zu sparen.



KI konstruktiv einsetzen

»Wir sind *Wired* und möchten bei der Entwicklung neuer Technologien ganz vorne mit dabei sein, aber auch ethisch und mit der gebotenen Umsicht vorgehen«, schreibt die Redaktion des Technikmagazins über ihren Umgang mit generativer KI im journalistischen Alltag. Um diesen Standards gerecht zu werden, hat sie sich Grundregeln gegeben. Darin heißt es nicht nur: »Wir publizieren keine Geschichten, die KI-generierten Text enthalten«, sondern auch: »Wir publizieren auch keine Texte, die von KI bearbeitet wurden« (*Wired* 2023). Neben dem Verweis auf mögliche Fehler argumentiert die Redaktion damit, dass die redaktionelle Bearbeitung eben immer auch eine Beurteilung darüber ist, was das Wichtigste, Originellste oder Unterhaltsamste an einem Beitrag ist. Eine Beurteilung, die sie nicht an eine Maschine auslagern möchte.

Damit trifft die Redaktion einen Nerv. Wie der Reuters Digital News Report 2024 zeigt, fühlen sich rund die Hälfte der in verschiedenen Ländern befragten Menschen unwohl bei der Vorstellung, dass Inhalte überwiegend von KI generiert und von Menschen nur überwacht werden. Ein weiteres Viertel hat sich noch keine klare Meinung gebildet. Die qualitative Forschung zu KI im Rahmen dieser Erhebung lege zudem nahe, dass »Vertrauen in Zukunft ein zentrales Thema sein wird« (Newman et al. 2024, S.20).

Vertrauen aufzubauen, braucht viel Zeit. Verspielt ist es hingegen schnell. Derzeit sind laut dem Reuters Digital News Report weniger als die Hälfte der Befragten der Ansicht, dass man dem Großteil der Nachrichten meist vertrauen könne (ebd.). Die Redaktion von *Wired* hat also gute Gründe für ihr zurückhaltendes Vorgehen. Doch wofür nutzt *Wired* dann KI? Zum Beispiel, um Ideen für Beiträge zu finden. KI soll also nicht grundlegende journalistische Arbeit verrichten, sondern dabei helfen, journalistische Arbeit besser zu machen. Auf den folgenden Seiten haben wir Ideen zusammengetragen, die sehr konkret zeigen, wie das im Redaktionsalltag gelingen kann.

Konstruktives Prompt-Engineering

Mit der Veröffentlichung von ChatGPT für die Allgemeinheit hat eine neue Wortgruppe Eingang in die deutsche Sprache gefunden: Prompt, Prompting, Prompt-Engineering. Diese Bezeichnungen sind Entlehnungen aus dem Englischen. Sie gehen zurück auf das englischsprachige Verb »to prompt«. Übersetzt heißt es so viel wie: »aufordern«, »veranlassen« oder »Anlass für etwas sein«. Der Text, den wir bei ChatGPT, Copilot oder Gemini in das dafür vorgesehene Feld eingeben, heißt davon abgeleitet »Prompt«. Mit ihm veranlassen wir das KI-System etwas zu tun.

Die Fähigkeit, die richtigen Eingaben im dafür vorgesehen Eingabefeld zu machen, nennt man Prompt-Engineering. Die Frage, ob ein Chatbot etwas »kann«, hängt dabei von zwei zentralen Faktoren ab. Zum einen davon, ob man es überhaupt verlangen kann. Und zum anderen davon, wie gut der Prompt ist. Doch was macht einen guten Prompt aus?

Grenzen: Was KI nicht können kann

Künstliche Intelligenz berechnet statistische Wahrscheinlichkeiten, sie orientiert sich nicht an der Wahrheit oder der Wirklichkeit. Das ist auch der Grund, warum generative KI-Anwendungen halluzinieren. Halluzinieren bedeutet, dass sie Antworten geben, die zwar sinnvoll oder passend klingen, faktisch aber nachweislich falsch sind. Seine Antwort liefert das System nämlich nicht, weil es etwas »weiß«, sondern weil der Textoutput mit größtmöglicher Wahrscheinlichkeit die passendste Antwort auf unsere Eingabe ist. KI-Anwendungen haben also weder ein »Verständnis« noch einen Bezug dazu, was richtig oder falsch, wahr oder unwahr ist.

Weil Halluzinationen keine Randerscheinungen sind, sondern eine logische Folge der Art und Weise, wie KI-Anwendungen funktionieren, sind Sprachmodelle derzeit nicht geeignet, eigenständig journalistische Texte zu produzieren, die veröffentlicht werden sollen. Dass KI gute journalistische Texte ersetzt, sehen wir daher nicht als den naheliegendsten, sondern vielmehr als den abwegigsten Anwendungsfall. Dennoch kann KI als Werkzeug in vielerlei Hinsicht im journalistischen Alltag nützlich sein und unsere Berichterstattung besser machen.

Grundlagen: Elemente eines guten Prompts

Damit das Prompting gut klappt, sollte man einige Grundregeln beachten. Dazu bietet es sich an, ein Grundgerüst zu nutzen, das man je nach Aufgabe und gewünschtem Ergebnis anpasst. Es setzt sich zusammen aus einer Rolle, einer Aufgabe, Kriterien für die gewünschte Antwort, dem Ziel des Prompts und dem gewünschten Ausgabeformat.

Grundgerüst eines Prompts

ROLLE Angenommen, die KI wäre eine Person. Wie würde sie sprechen? Wie würde sie Informationen weitergeben? Das ist die Rolle des KI-Modells. Der Prompt sollte diese Rolle explizit beschreiben.

Du bist Journalist bei einem Online-Medium und schreibst regelmäßig konstruktive Artikel zu politischen und gesellschaftlichen Themen.

AUFGABE Welches Problem oder welche Aufgabenstellung soll das KI-Tool lösen? Hier sollte möglichst genau beschrieben werden, was die KI machen soll.

Fasse zusammen, wie verschiedene Personen und Personengruppen zu [Thema] stehen. Erläutere deren Meinungen und Überzeugungen. Nutze dabei auch betroffene Personen aus Minderheitengruppen.

KRITERIEN Was soll eine gute Antwort bieten? Vor allem Genauigkeit, eher Kreativität oder etwas anderes? Gut definierte Kriterien für die gewünschte Antwort sind ein wichtiger Teil des Prompts.

Deine Übersicht sollte alle relevanten Personengruppen einbeziehen: Politiker, Einwohner, Aktivisten, betroffene Minderheiten und so weiter und dabei alle Geschlechter und ihre Perspektiven berücksichtigen.

ZIEL Wie lässt sich das Ziel des Prompts definieren? Je klarer das übergeordnete Ziel der Aufgabe formuliert wird, desto effektiver ist der Prompt.

Ziel ist es mit den gesammelten Informationen Interviewpartnerinnen und Interviewpartner zu suchen und einen journalistischen Artikel zu erstellen.

FORMAT In welchem Format soll die Maschine die Antwort liefern? Als Liste, ausformulierten Absatz, Gespräch? Dies sollte im Prompt spezifiziert werden.

Erstelle eine Tabelle mit den verschiedenen Personengruppen und deren Meinungen und Überzeugungen.

Detailtiefe

Wichtig ist zudem die Detailtiefe des Prompts: Ist er zu allgemein, interpretiert das System die Eingabe möglicherweise völlig anders als gedacht. Liefert er dagegen zu viele Details, werden die möglichen Antworten zu sehr eingeschränkt, so dass sie eventuell nicht mehr nützlich sind. Hier gilt es, die richtige Balance zwischen zu vielen und zu wenigen Details zu finden. Als Daumenregel gilt: so kurz wie möglich, so spezifisch wie nötig.

Refinement

Wer diese Elemente eines guten Prompts kennt, wird in der Lage sein, beste Ergebnisse aus der KI herauszukitzeln. Dennoch kann es immer mal wieder vorkommen, dass man mit der Antwort der KI nicht zu 100 Prozent zufrieden ist. Dann beginnt das »Refinement«, also das Nachjustieren. So, als würde ich mit einem Menschen über eine erbrachte Aufgabe reden, kann man auch der KI Feedback zum Ergebnis geben.

KONSTRUKTIVES FEEDBACK GEBEN

Der [dritte Vorschlag] gefällt mir grundsätzlich sehr gut. Bitte formuliere ihn ausführlicher.

VARIATIONEN DER ANTWORT ERFRAGEN

Das Ergebnis gefällt mir schon gut. Allerdings würde ich gern weitere betroffene Minderheiten abbilden. Kannst du bitte weitere hinzufügen?

UM ERGÄNZUNGEN UND ERWEITERUNGEN BITTEN

Kannst du in deiner Tabelle bitte eine weitere Spalte hinzufügen, in der du begründest, wie die Meinungen und Überzeugungen der Personengruppen zustande kommen?

Gendern in Prompts

Bei der Rollenzuschreibung der KI und auch in anderen Fällen kann die Frage aufkommen, ob man in Prompts gendern sollte. Wie wir gelernt haben, ist es von zentraler Bedeutung, dass Prompts so präzise wie möglich formuliert sind, um sehr gute Ergebnisse zu erzielen. Gendern birgt daher das Risiko, KI zu verwirren. Vor allem, wenn bei mehreren Begriffen nacheinander sowohl die weibliche als auch die männliche Form genannt wird wie zum Beispiel in diesem Prompt:

Liste die Perspektiven folgender Personengruppen zu [Thema] auf: Politikerinnen und Politiker, Einwohnerinnen und Einwohner, Aktivistinnen und Aktivisten und betroffene Minderheiten.

Dies könnte die KI dazu bringen, sich zu sehr mit den Geschlechtern auseinanderzusetzen, anstatt die eigentliche Aufgabe zu lösen. Offizielle Auswertung zu dieser Frage gibt es bislang zwar nicht. Dennoch empfehlen wir, solche Prompts zu vermeiden. Es gibt dazu verschiedene Alternativen. Wahrscheinlich verwirrt das Gendern mit Stern oder Doppelpunkt – zum Beispiel: »Einwohner:innen« – die KI weniger als verschiedene parallele Nennungen. Am präzisesten und damit am sichersten ist jedoch die Verwendung einer klaren Angabe, also entweder »Du bist Journalist« oder »Du bist Journalistin«.

Für die Prompts, die wir im *Konstruktiven KI-Kompass* vorstellen, haben wir uns für die Nennung eines einzelnen Geschlechts entschieden. Allerdings haben wir weder ausschließlich männliche oder

weibliche Formulierungen gewählt, sondern beide Formen in unregelmäßiger Reihenfolge abwechselnd genutzt. Wenn es uns vom Kontext her sinnvoller erschien die KI nicht auf ein Geschlecht festzulegen, haben wir sowohl die männliche als auch die weibliche Schreibweise verwendet.

Prompting für Fortgeschrittene: Parameter einstellen

Neben einem gut konstruierten Prompt-Gerüst gibt es weitere Faktoren, die die Qualität und Relevanz der Antworten einer KI beeinflussen können: die Parameter. Diese Einstellungen ermöglichen es, den Schreibstil der KI-Antworten gezielt zu steuern. Während diese Parameter besonders bei ChatGPT prominent sind, können sie in abgewandelter Form auch bei anderen textbasierten KI-Tools genutzt werden. Angewendet werden Parameter, indem man sie einfach als Kurzbefehl an den Prompt anhängt. Allgemein sieht das dann zum Beispiel so aus: [Parameter]=[0.2]. Um Probleme zu vermeiden, sollten für die Zahlen die englischen Schreibweisen gewählt werden: mit Punkt, statt mit Komma. Im Folgenden stellen wir zwei Parameter vor, die für die journalistische Nutzung von KI-Modellen relevant sind und die Ergebnisse auf interessante Art und Weise beeinflussen können.

Temperatur

Soll das Ergebnis eher sachlich sein oder vielmehr originell? Der Parameter »Temperatur« spielt eine wichtige Rolle bei der Generierung von Antworten durch KI-Tools. Er beeinflusst, wie konservativ, sachlich oder kreativ die Antworten ausfallen. Da jede Antwort der KI eine Wahrscheinlichkeitsrechnung ist, definiert die Temperatur, ob das Tool eher wahrscheinliche oder unwahrscheinliche Antworten wählen wird.

Die Temperatur wird durch den Wert X in der Anfrage »temperature=X« festgelegt, wobei X zwischen 0 und 1 liegt. Durch Experimentieren mit verschiedenen Temperaturwerten kann das optimale Gleichgewicht zwischen Kreativität und trockenen Antworten gefunden werden.

Ist der Wert der Temperatur niedrig eingestellt – also zum Beispiel auf 0.2 –, generiert das KI-Tool konservative Antworten. Es bevorzugt dann bekannte Phrasen und formuliert Sätze in einem eher förmlichen Stil. Für Medienschaffende könnte dies nützlich sein, wenn es darum geht, klare und präzise Recherchehilfen zu generieren.

Ein hoher Wert der Temperatur – also zum Beispiel 0.9 – gibt der KI das Signal, dass die Antworten kreativer und experimenteller sein sollen. Die Ergebnisse werden dadurch origineller, zugleich aber auch unvorhersehbarer. Medienschaffende könnten diese Einstellung nutzen, um kreative Ideen zu entwickeln.

Diversity Penalty

Ein weiterer wichtiger Parameter, der einen Effekt auf die Qualität der generierten Antworten hat, ist »Diversity Penalty«. Mit diesem Parameter kann die Diversität, also die Vielfalt, der verwendeten Wörter in der Antwort eingestellt werden. Dies geschieht, indem Einfluss genommen wird auf die Wiederholung gleicher Wörter. Bei einer höheren Einstellung bestraft der Parameter Tokens (Wörter oder Zeichen), die bereits in der Antwort verwendet wurden. Das sorgt im Ergebnis für vielfältigere und abwechslungsreichere Ergebnisse.

Der Wertebereich für Diversity Penalty liegt zwischen 0 und 2. Ein höherer Wert führt zu vielfältigeren Antworten. Je kreativer der Artikel ausfallen soll, desto höher sollte der Wert für den Parameter der Diversity Penalty eingestellt sein.

Wie bei der Temperatur wird auch hier ein Wert X gewählt und im Prompt angegeben: »diversity_penalty=X«

Diversity Penalty ist vor allem bei längeren Texten interessant. Aber auch bei kurzen Texten lassen sich Unterschiede in der Antwort erkennen.

Angenommen die KI soll erklären, was konstruktiver Journalismus ist. Bei einem niedrigen Wert wie 0.2 neigt das KI-Tool dazu, die gleichen Wörter oder Phrasen häufiger zu wiederholen, der höhere Wert schafft mehr Abwechslung.

Erkläre mir in zwei Sätzen, was konstruktiver Journalismus ist. »diversity_penalty=0.2«

Antwort:

Konstruktiver Journalismus ist ein Ansatz im Journalismus, der sich darauf konzentriert, nicht nur über Probleme und Konflikte zu berichten, sondern auch über Lösungen und positive Entwicklungen. Ziel ist es, ein ausgewogenes Bild zu vermitteln, das zur informierten Meinungsbildung und positiven gesellschaftlichen Veränderungen beiträgt.

Erkläre mir in zwei Sätzen, was konstruktiver Journalismus ist. »diversity_penalty=0.9«

Antwort:

Konstruktiver Journalismus konzentriert sich darauf, neben Problemen auch Lösungen und positive Entwicklungen zu berichten, um ein umfassenderes Bild der Realität zu vermitteln. Dieser Ansatz zielt darauf ab, das Publikum zu informieren, zu inspirieren und zu befähigen, indem er Wege zur Bewältigung gesellschaftlicher Herausforderungen aufzeigt.

Doch was macht einen Prompt nun konstruktiv?

Prompts können maßgeblich dabei unterstützen, die gesamte Berichterstattung konstruktiver auszurichten. Das fängt bereits bei der Frage an, wie Journalistinnen und Journalisten an ein Thema herangehen, und inwieweit zum Beispiel persönliche Prägungen oder kognitive Verzerrungen den eigenen Blick auf ein Ereignis oder einen Zusammenhang prägen. Generative KI kann so als ein Korrektiv für eine reflektiertere, aber auch für eine vollständigere und realistischere Weltwahrnehmung dienen.

Ob es um Themenfindung und Themenzuschnitt, die Auswertung von Dokumenten, verschiedene Sichtweisen auf ein Thema oder um bessere Fragen bei Interviews geht: Spezifisches Prompting kann dabei helfen, Lösungsorientierung, Perspektivenvielfalt und konstruktiven Dialog – die drei zentralen Elemente des konstruktiven Journalismus – gezielter in die tägliche Arbeit einzubeziehen und die Berichterstattung stärker an den Bedürfnissen der Menschen auszurichten. Wie genau das funktionieren kann, zeigen wir auf den nächsten Seiten. Dort stellen wir einige ausgewählte Anwendungsfälle vor. Darüber hinaus gibt es aber natürlich noch unzählige weitere interessante Einsatzfelder.

Für die Beispiele nutzen wir aus zwei Gründen vorrangig ChatGPT: Wir gehen davon aus, dass es in unserer Leserschaft das am weitesten verbreitete KI-Werkzeug ist, weil es eines der ersten großen und für alle zugänglichen generativen KI-Systeme war – und daher schon am bekanntesten ist. Außerdem hatte ChatGPT zum Zeitpunkt der Produktion der hier vorliegenden Inhalte im Bereich Textanalyse und Textgenerierung den größten Funktionsumfang der vergleichbaren Tools (Überblick bekannter KI-Modelle siehe Seiten 22/23).

Prompts für konstruktive Themenideen

Der konstruktive Journalismus zeigt nicht nur Probleme auf, sondern recherchiert auch Lösungsansätze und leistet damit einen wichtigen Beitrag zur gesellschaftlichen Debatte. Eine zentrale Herausforderung besteht darin, Ideen für eine konstruktive Berichterstattung zu finden. Wie können Medienschaffende Themen und Ansätze identifizieren, die konstruktiv und relevant für ihre Zielgruppe sind? Hier stellen wir Schritte und Methoden vor, um konstruktive Themenideen zu entwickeln.

Themenwahl

Ein erster Schritt bei der Ideenfindung ist die Auswahl eines Themenbereichs. Es kann inspirierend sein, zunächst komplett offen nach möglichen konstruktiven Themen zu fragen:

Ich suche nach Themen, die sich für konstruktiven Journalismus eignen. Nenne mir zehn mögliche Oberthemen, die ich als Journalistin umsetzen könnte.

Sollte der Themenbereich schon feststehen, sollte diese Information unbedingt mit der KI geteilt werden. Dann kann die Suche nach konkreten Themenideen bereits eine Ebene tiefer beginnen. Steht beispielsweise der Bereich Umwelt- und Klimaschutz im Fokus, könnte ein Prompt so aussehen:

Ich suche nach Themen im Bereich Umwelt- und Klimaschutz, die sich für konstruktiven Journalismus eignen. Nenne mir zehn mögliche Themen innerhalb dieses Themenkomplexes, die ich als Journalistin umsetzen könnte.

Diese im nächsten Schritt von der KI genannten Themen bieten die Grundlage für eine breite Palette an möglichen Geschichten, die sowohl informativ als auch inspirierend sein können.

Themenidee vertiefen

Sobald ein Thema ausgewählt ist, gilt es, spezifische Aspekte zu identifizieren, die besonders interessant und relevant für eine konstruktive Berichterstattung sind. Dabei stehen vor allem Lösungsansätze, Perspektivenvielfalt und konstruktiver Dialog im Zentrum.

Welche spezifischen Aspekte könnte ich bei der konstruktiven Berichterstattung über [Thema] hervorheben, um verschiedene Perspektiven und Lösungsmöglichkeiten darzustellen?

Medienschaffende können so tiefer in das Thema einsteigen, zusätzliche Informationen sammeln und die weitere Richtung der Recherche und Berichterstattung identifizieren. Diese Einblicke ermöglichen es, die Berichterstattung bewusst auf ausgewählte Aspekte auszurichten.

Sobald das Thema feststeht und in einem ersten Schritt vertieft wurde, kann im Detail beleuchtet werden, wie eine konstruktive Umsetzung aussehen könnte. Dafür ist es wichtig, möglichst viele Perspektiven und Sichtweisen abzubilden. Das nächste Kapitel gibt dazu eine Anleitung.

Prompts für mehr Perspektivenreichtum

Im konstruktiven Journalismus geht es auch darum, das Wohl der Gesellschaft und die Bedürfnisse der Menschen in den Mittelpunkt zu stellen. Doch nicht zu allen Gruppen in unserer Gesellschaft haben Medienschaffende unmittelbar Zugang. Mit Hilfe von KI können sie sich diesen Gruppen aber annähern und sich ein wenig in sie hineinversetzen. Fragen an die KI-Anwendung diesbezüglich können sein: Wie schaut eine bestimmte Gruppe auf das Thema? Was ist den Menschen in dieser Gruppe besonders wichtig? Welche Fragen haben sie?

Andere Sichtweisen kennenlernen

Beim Prompting für mehr Perspektivenreichtum bedeutet dies, dass die KI verschiedene Rollen annimmt, je nachdem, welche Perspektive beleuchtet werden soll. Die KI könnte zum Beispiel gebeten werden, die Rolle einer Politikerin, eines Aktivisten, der Angehörigen einer Minderheit oder einer Person aus einer bestimmten Altersgruppe einzunehmen, deren Ansichten zu einem Thema interessieren. Mit den Antworten ist es dann möglich, sich unbekanntes Sichtweisen anzunähern, sie besser zu verstehen und ihre Perspektiven gegebenenfalls sinnvoll in die Berichterstattung einzubeziehen. Ein Prompt, der dabei hilft, könnte zum Beispiel so aussehen:

Versetze dich in die Rolle von [Person/Gruppe].
Erläutere deren Meinungen und Überzeugungen zu
[Thema].

Vielfalt abbilden

Ein Prompt für mehr Perspektivenreichtum kann aber auch komplexer formuliert sein. Zum Beispiel, indem die KI dazu aufgefordert wird, alle Stimmen – auch jene, die oft überhört werden – zu Wort kommen zu lassen. Dies garantiert eine Vielfalt an Perspektiven.

Welche verschiedenen Gruppen/Gemeinschaften sind von
[Thema] betroffen und wie unterscheiden sich ihre
Perspektiven und Erfahrungen?

In einem nächsten Schritt kann man dann ins Detail gehen:

Versetze dich in die Rolle von [Person A, Person B ...]. Sammle und erkläre deren unterschiedliche Perspektiven, Meinungen und Erfahrungen zu [Thema]. Liste anschließend Minderheitengruppen, die von [Thema] betroffen sind. Erkläre dann, welche Meinungen und Erfahrungen Minderheitengruppen in Bezug auf das Thema haben.

Optimalerweise ist dieser Prompt in zwei Schritten formuliert. Andernfalls kann es vorkommen, dass bestimmte Aufgaben nicht ausgeführt werden.

Blinde Flecken beleuchten

Perspektivenreichtum bedeutet auch, tief in die vielschichtigen Ursachen eines Konflikts und die Aspekte eines Themas einzutauschen, um ein vollständigeres Bild zu zeichnen. Es geht darum, über das Offensichtliche hinaus, Fragen zu stellen und Bereiche zu erkunden, die von sozialen und ökonomischen Bedingungen bis hin zu kulturellen und politischen Faktoren reichen.

Medienschaffende sollen dabei sowohl die gängigen Erklärungen als auch die weniger beachteten Perspektiven berücksichtigen, um die Vielzahl an Stimmen und die Komplexität der Realität widerzuspiegeln, ohne in die Falle der falschen Ausgewogenheit zu tappen. Als falsche Ausgewogenheit (»false balance«) wird eine mediale Verzerrung beschrieben, bei der – zum Beispiel aus dem Bestreben ein Thema von allen Seiten zu beleuchten – zweifelhaften Positionen Legitimität verliehen wird oder einer Minderheitenmeinung unverhältnismäßig viel Raum gegeben wird.

Ziel des konstruktiven Journalismus hingegen ist es, ein nuanciertes Verständnis von Themen und Zusammenhängen zu vermitteln und dadurch Menschen zu befähigen, informierte Entscheidungen für ihr Leben und die Gesellschaft zu treffen. Mit KI lassen sich dafür Sichtweisen zu Tage fördern, die sonst vielleicht übersehen würden:

Identifiziere und erkläre die verschiedenen Dimensionen und Einflussfaktoren, die bei [Thema] eine Rolle spielen. Stelle dir dabei die Frage: Wer ist noch betroffen – abseits der bekannten Akteure?

Zoomstufen finden

Eine größere Vielfalt an Perspektiven wird auch möglich, indem Journalistinnen und Journalisten auf verschiedene Weise in ein Thema hineinzoomen oder den Zoom erweitern und dadurch unterschiedliche Ebenen eines Themas beleuchten.

Auf gesellschaftlicher Ebene kann ein Ereignis oder ein Thema auf der individuellen Ebene (mikro), aus der Perspektive einer Gruppe (meso) oder der Perspektive ganzer Staaten (makro) betrachtet werden.

Medienschaffende können mit KI zum Beispiel ein Ereignis auf Mikroebene – aus Sicht einer einzelnen Person – darstellen oder auf Makroebene – in Bezug auf größere soziale oder politische Strukturen. Perspektivenreich zu berichten heißt, mehrere Zoomstufen und Blickwinkel gleichzeitig zuzulassen.

Betrachte [Thema] aus der Perspektive einer einzelnen betroffenen Person und der gesamten Gesellschaft. Erkläre, wie sich die Ziele dieser beiden Ebenen voneinander unterscheiden.

Der Zoom kann sich auch auf den zeitlichen Rahmen beziehen: Die Berichterstattung über ein aktuelles Ereignis verändert sich, wenn es in den Kontext längerer historischer Linien gestellt wird. Ein zeitlich weiterer Zoom kann helfen, tiefergehende Ursachen und Muster zu erkennen.

So kann man beispielsweise untersuchen, wie lokale Entscheidungen globale Auswirkungen haben oder wie sich internationale Trends auf individuelle Gemeinschaften auswirken. Ein weiteres Beispiel wäre, ein politisches Ereignis, wie eine Wahl, nicht nur durch das Prisma des Wahltags zu sehen, sondern auch die historischen Entwicklungen und kulturellen Strömungen in den Blick zu nehmen, die zu diesem Punkt geführt haben.

Betrachte die Geschichte von [Ereignis] aus der Perspektive verschiedener Zeiträume, geografischer Standorte und gesellschaftlicher Fokusse. Betrachte die Vorgeschichte des Ereignisses, die Auswirkungen auf lokaler und internationaler Ebene und erkläre diese aus Sicht einzelner Personen, bestimmter Gruppen und ganzer Staaten.

Achtung: Da die KI hier nach konkreten Fakten gefragt wird, sollten diese im Anschluss alle geprüft werden.

Selbstreflexion

Letztlich kann die KI auch helfen, sich selbst kritisch zu hinterfragen. In dieser Selbstreflexion stellen Journalistinnen und Journalisten ihre eigenen Vorstellungen, Überzeugungen und mögliche Verzerrungen ihrer eigenen Wahrnehmungen auf den Prüfstand. Ziel ist es, die Berichterstattung so objektiv und unvoreingenommen wie möglich zu gestalten.

Dieser introspektive Ansatz hilft dabei, ein breiteres Spektrum an Sichtweisen zu integrieren und unter anderem den sogenannten Bestätigungsfehler (»confirmation bias«) zu vermeiden, bei dem man dazu neigt, Informationen stärker zu gewichten, die die eigenen Ansichten widerspiegeln – und diese dadurch weiter zu verstärken.

In der Praxis bedeutet Selbstreflexion, dass Medienschaffende aktiv nach Quellen und Perspektiven suchen, die ihre eigenen Annahmen herausfordern. Es geht um das Bewusstsein, dass die eigenen Erfahrungen die Interpretation von Informationen beeinflussen. Die KI kann dabei helfen, indem sie bereits geschriebene Textentwürfe oder Rechercheergebnisse analysiert und auf Perspektivenreichtum überprüft.

Analysiere [den folgenden Text/die folgenden Rechercheergebnisse]. Reflektiere kritisch über mögliche Vorurteile bei der Berichterstattung über [Thema]. Überprüfe, wie Aussagen und Annahmen im Text die Sicht auf das Thema beeinflussen könnten, und suche aktiv nach Informationen und Perspektiven, die diesen entgegenstehen.

Durch die Berücksichtigung dieser Elemente kann ein Prompt die KI effektiv leiten. Die Antworten können dabei unterstützen, journalistische Beiträge zu fördern, die informieren, inspirieren und zur gesellschaftlichen Diskussion anregen. Andererseits ist klar, dass ein solches Vorgehen bestenfalls eine grobe Annäherung leisten kann. Denn Künstliche Intelligenz liefert statistische Stereotype, keine nuancierten Charaktere. Die eigene Vor-Ort-Recherche kann der Chat mit einer KI-Anwendung also nicht ersetzen.

Prompts für konstruktive Interviewfragen

Konstruktive Interviewfragen fördern ein tieferes Verständnis der Themen und motivieren sowohl die Interviewten als auch das Publikum, mehr über positive Veränderungen und Fortschritte nachzudenken. Sie lenken die Aufmerksamkeit auf Verbesserungsmöglichkeiten, Resilienz und erfolgreiche Strategien zur Problemlösung. Solche Fragen sind speziell darauf ausgerichtet, den am Gespräch Beteiligten und dem Publikum Wege aufzuzeigen, wie aktuelle Herausforderungen überwunden und Erfolge erzielt werden können.

Um konstruktive Interviewfragen zu finden, kann Künstliche Intelligenz eine bedeutende Rolle spielen, indem sie den Prozess effizienter, kreativer und umfassender gestaltet. Den Brainstorming-Prozess für Interviewfragen kann KI unterstützen und so mehr Konstruktivität in die Berichterstattung bringen. Im Folgenden stellen wir die Kernaspekte von konstruktiven Interviewfragen und Beispiel-Prompts für ein KI-gestütztes Brainstorming zu diesen Fragen vor.

Zukunftsorientierung

Konstruktive Fragen ermutigen dazu, über die gegenwärtige oder vergangene Situation hinauszudenken und sich eine bessere Zukunft vorzustellen. Fragen wie: »Was wünschen Sie sich in der gegenwärtigen Situation?« oder »Angenommen, das Problem wäre eines Tages plötzlich gelöst, woran würden die Menschen das zuerst im Alltag merken?« helfen dabei, Visionen und Wünsche zu artikulieren, die motivierend wirken können.

Erzeuge Fragen, die auf die Vision und die langfristigen Auswirkungen von [Thema] abzielen. Beispielsweise: »Angenommen, die wichtigsten Ziele Ihrer Initiative sind erreicht – wie sieht die Situation in zehn Jahren aus? Was wären die spürbarsten Veränderungen für die Gemeinschaft?«

Lösungsansätze

Statt sich nur auf die Probleme zu konzentrieren, richten konstruktive Fragen den Fokus auf praktische Lösungen und Erfolge. Sie fragen nach bereits realisierten kleinen Verbesserungen und erkunden, was dazu beigetragen hat. Beispiele sind: »War es schon mal ein klein wenig besser? Was war da anders?« oder »Kennen Sie Beispiele, wo es ein klein bisschen besser läuft? Was wird dort getan?«

Generiere Fragen, die darauf abzielen, vergangene oder potenzielle Erfolge zu erkunden. Zum Beispiel: »Können Sie ein Beispiel nennen, bei dem ein ähnliches Problem erfolgreich gelöst wurde, und was können wir daraus lernen?«

Perspektivenreichtum

Durch das Erfragen verschiedener Blickwinkel und das Einholen von Meinungen von Menschen, die direkt von einem Thema betroffen sind, gewährleisten konstruktive Fragen, dass eine Geschichte mehrdimensional und umfassender erzählt werden kann. KI kann uns helfen, solche Fragen zu generieren.

Erstelle Fragen, die dazu ermutigen, die Meinungen und Erfahrungen anderer Betroffener zu berücksichtigen. Zum Beispiel: »Welche anderen Gruppen sind von diesem Thema betroffen und wie unterscheiden sich deren Ansichten und Lösungsansätze von Ihren eigenen?«

Reflexion und Kontextualisierung

Diese Art von Fragen ermuntert dazu, über den eigenen Beitrag zum Status quo und die Möglichkeiten der Veränderung nachzudenken. »In wessen Macht läge es, jetzt und hier oder in Zukunft etwas zu verändern?« ist ein Beispiel dafür, Interviewte anzuregen, über ihre eigene Rolle und die anderer Akteure in der Situation nachzudenken.

Schlage Fragen vor, die zur Selbstreflexion anregen. Zum Beispiel: »Inwiefern haben Ihre eigenen Erfahrungen und der Kontext, in dem Sie arbeiten, Ihre Ansichten zu [Thema] geprägt?«

Von der Stimme zum Text: Transkribieren mit KI

Ein einstündiges Interview Wort für Wort zu übertragen, kann lange dauern. Und nerven. Diesen Arbeitsschritt übernehmen schon heute Transkriptionstools – inzwischen auch recht zuverlässig. Möglich macht das eine Technologie für automatische Spracherkennung.

Die Automatic-Speech-Recognition-Technologie (ASR), die Sprache in Text umwandelt, hat ihre Ursprünge in den 1950er Jahren, entwickelte sich jedoch zuletzt rapide weiter. Die Verbesserungen sind vor allem dem Fortschritt in KI und Maschinellem Lernen (→siehe Seite 16) zu verdanken: Systeme können trainiert werden, komplexe Sprachmuster mit hoher Genauigkeit zu erkennen sowie verschiedene Sprecherinnen und Sprecher zu identifizieren und zu trennen.

Die Umwandlung von Sprache zu Text durch ASR funktioniert, indem die gesprochene Sprache zunächst in kleine Einheiten, sogenannte Phoneme, zerlegt wird – die kleinsten hörbaren Einheiten einer Sprache. ASR-Systeme analysieren diese Phoneme im Kontext der umgebenden Phoneme, um Wörter und Sätze zu identifizieren. Moderne Spracherkennungssysteme verwenden Tiefe Künstliche Neuronale Netzwerke (→siehe Seite 17), die aus umfangreichen Datenmengen lernen. Sie erkennen Sprachmuster und verstehen komplexe Nuancen wie Betonung und Rhythmus.

Transkriptionstools sind meist webbasiert und erfordern keine spezielle Software-Installation. Sie können direkt im Browser oder auf Smartphones verwendet werden. Die Bedienung ist intuitiv: Mit nur wenigen Klicks lassen sich Dateien hochladen und die Transkription starten.

Doch worauf sollte man achten, wenn man diese Technologie für sich nutzen möchte? Kriterien für die Auswahl eines Tools können die Genauigkeit der Übertragung, der Umgang mit Datenschutz und auch die Kosten sein.

Genauigkeit

Anbieter von Transkriptionstools geben häufig eine Genauigkeitsrate an, die in der Regel durch die Word Error Rate (WER) bestimmt wird. Die WER misst die Fehleranzahl in einer Transkription im Vergleich zu einem Referenztext, indem sie falsch transkribierte, fehlende oder zusätzliche Wörter berücksichtigt. Eine WER von zehn Prozent deutet darauf hin, dass 90 Prozent der Wörter korrekt erfasst wurden.

Populäre Tools wie **Trint**, **Rev** oder **Descript** geben an, Genauigkeitsraten von mehr als 90 Prozent zu erreichen, wobei diese durch Faktoren wie die Qualität der Audioaufnahme und die sprachliche Komplexität beeinflusst werden. Tools, die speziell für bestimmte Sprachen entwickelt wurden, können oft eine bessere Genauigkeit bieten.

Dies macht sie besonders wertvoll für Medien-schaffende, die Transkriptionen in bestimmten Sprachen benötigen. Beispielsweise hat sich **GoSpeech** auf deutsche Transkriptionen spezialisiert. Für Journalistinnen und Journalisten, die in einer Vielzahl von Sprachen arbeiten, können Tools wie **HappyScribe** und **Sonix** nützlich sein, da beide mehr als 50 verschiedene Sprachen unterstützen.

Die Qualität einer Transkription wird oft durch Faktoren wie Akzente, Dialekte und Fachjargon beeinträchtigt. Moderne Transkriptionstools nutzen adaptive Algorithmen, die sich mit jeder Nutzung besser an spezielle Sprachmuster anpassen, um die Fehlerquote zu senken. Zudem erlauben viele Tools das Anlegen benutzerdefinierter Wörterbücher, sodass Nutzende spezifische Fachbegriffe, Eigennamen und Abkürzungen hinzufügen können, die dann präziser erkannt und transkribiert werden.

Datenschutz

Gerade für Journalistinnen und Journalisten, die vertrauliche Interviews führen und die Anonymität ihrer Quellen wahren müssen, ist Datenschutz wichtig. Die Ansätze der Anbieter von Transkriptionstools variieren allerdings stark. Es lohnt sich, besonders auf die Art der Datenverschlüsselung, die Einhaltung der DSGVO und den Umgang mit Datenlöschungsanfragen zu achten.

Eine Ende-zu-Ende-Verschlüsselung oder die Transport Layer Security (TLS) stehen beispielsweise für eine sichere Übertragung

von Daten. Auch Zertifizierungen können Hinweise auf einen guten Umgang mit sensiblen Daten bieten. Die ISO 27001 zum Beispiel ist eine international anerkannte Norm, die die Anforderungen für ein Informationssicherheitsmanagementsystem (ISMS) festlegt.

Die Wahl des Speicherortes der Daten ist ebenfalls entscheidend, da zum Beispiel europäische Server der DSGVO unterliegen, während US-Server anderen Regularien folgen. Beispielsweise speichern Tools wie **Amberscript** oder **F4X** ihre Daten auf Servern in Europa, was sie an die strengen Anforderungen der DSGVO bindet. Auffällig ist auch, dass Anbieter, die sich für Datensicherheit einsetzen, häufig am transparentesten mit ihren Datenschutzbedingungen umgehen.

Die Kosten für Transkriptionstools variieren und können Abonnement- oder Pay-per-Use-Modelle umfassen. Viele Hersteller bieten eine kostenlose Basisversion an, die für gelegentliche Nutzung ausreichend sein kann. Je nach Anbieter und dem Umfang des enthaltenen Zeitkontingents bewegen sich die monatlichen Kosten eines Abonnements zwischen fünf und 60 Euro.

ChatGPT konstruktiv tunen

Nutzende von ChatGPT können den Chatbot auf die eigenen Bedürfnisse anpassen oder auch für bestimmte Aufgaben trainieren. Technisch gesehen gibt es dafür zwei Möglichkeiten: ChatGPT individuell zu konfigurieren oder einen eigenen GPT zu erstellen. Wir erläutern die grundsätzliche Funktionsweise anhand der aktuellen Möglichkeiten, die OpenAI bietet, wohl wissend, dass dies eine Momentaufnahme ist. Details können sich ändern.

Individuelle Einstellungen

Über die Funktion »ChatGPT individuell konfigurieren« in den Einstellungen kann das Tool an individuelle Bedürfnisse angepasst werden. Dort kann zum Beispiel hinterlegt werden, was das Tool wissen muss, so dass nicht regelmäßig die gleichen Informationen mit der KI geteilt werden müssen. Dort kann beispielsweise stehen, dass der Nutzende des Accounts Journalist ist und konstruktiv journalistisch arbeitet. Ebenfalls kann dort festgelegt werden, wie ChatGPT reagieren soll. Zum Beispiel indem die KI angewiesen wird, immer konstruktiv und mit Lösungsansätzen zu argumentieren. Diese Funktion ist sowohl in der kostenlosen als auch in der kostenpflichtigen Version verfügbar. Hat man die individuellen Einstellungen einmal vorgenommen, gelten sie von da an grundsätzlich für sämtliche Anfragen, die man über den entsprechenden Account stellt.

Einen eigenen Bot bauen

Die zweite Option ist das Training einer Version von ChatGPT für einen konkreten Zweck. Dabei wird ein eigener GPT-Bot für eine bestimmte Aufgabe trainiert. Anders als bei den individuellen Einstellungen nutzt man die Voreinstellungen über einen GPT gezielt bei Bedarf.

Über den GPT-Store von Open AI können alle Nutzenden auf alle GPTs zugreifen, die von Nutzenden erstellt und für den Store freigegeben wurden. Einen GPT erstellen können nur diejenigen, die über die kostenpflichtige Plus-Version verfügen.

Für das Erstellen eines GPT bieten sich verschiedene Szenarien an: Man kann ihn für sich persönlich, für das eigene Team oder aber für die breite Öffentlichkeit entwickeln. Ein eigener GPT kann dann im Alltag von Medienschaffenden als Sparringspartner dienen. Ein möglicher Anwendungsbereich ist das Finden von konstruktiven Ideen.

Um einen GPT für den gewünschten Einsatzzweck zu nutzen, muss man den Chatbot zunächst trainieren. Das funktioniert, indem man dem GPT schriftliche Anweisungen gibt. Dieser programmiert sich dann entsprechend selbst. Um diesen Prozess zu starten, füttert man ihn mit Informationen, Anweisungen und Hinweisen, die er berücksichtigen soll. Das kann über ein direktes Prompting geschehen oder auch über das Hochladen von relevanten Dokumenten. Es bietet sich ein Training in mehreren Runden an, bei dem man auch immer wieder den Chatbot zu zentralen Aspekten befragt. Ergeben sich unpassende Antworten, lässt sich das über weitere Eingaben feintunen.

Definition des Einsatzzwecks

Zur konstruktiven Nutzung eines eigenen GPT muss zunächst klar definiert werden, dass der GPT darauf trainiert werden soll, lösungsorientierte Ansätze zu finden, vielfältige Perspektiven darzustellen und den konstruktiven Dialog zu fördern. Diese Ziele müssen in den Anweisungen und Trainingsdaten klar benannt werden.

Du bist ein GPT, der Medienschaffenden bei der Findung und Ausarbeitung von konstruktiven Ideen helfen soll. Dein Ziel ist es, lösungsorientierte Ansätze zu finden, vielfältige Perspektiven darzustellen und konstruktiven Dialog zu fördern.

Bereitstellung von Trainingsdaten

Anschließend muss ein GPT mit Informationen gefüttert werden, die als Grundlage für seine späteren Antworten dienen. Für einen konstruktiven GPT bietet es sich an, grundlegende Leitlinien des konstruktiven Journalismus als PDF- oder Word-Datei hochzuladen. Diese Informationen umfassen die wichtigsten Elemente des konstruktiven Journalismus: Lösungsfokus, Perspektivenreichtum und konstruktiver Dialog.

Anbei findest du ein ausführliches Dokument zu den Leitlinien des konstruktiven Journalismus. Nutze diese Informationen als Grundlage für jede deiner Antworten.

Erstellung von spezifischen Anweisungen

Nun müssen Anweisungen formuliert werden, die den GPT dazu anleiten, lösungsorientierte Ansätze zu priorisieren, verschiedene Perspektiven zu berücksichtigen und den Dialog zu fördern. Auch muss festgelegt werden, wie der GPT auf Anfragen reagiert: Duzt er die Nutzenden? Gibt er Antworten als Fließtext oder Tabelle?

Jede deiner Antworten priorisiert lösungsorientierte Ansätze, berücksichtigt verschiedene Perspektiven und fördert den Dialog. Gib deine Antworten in dem Format, das für die Anfrage am besten geeignet ist. Du duzt deine Nutzenden.

Kontinuierliches Feintuning

Durch kontinuierliches Feedback und Anpassungen wird sichergestellt, dass der GPT konsistente und relevante Vorschläge für die journalistische Praxis liefert. Dies wird durch mehrstufige Trainingsrunden erreicht, bei denen der GPT immer wieder zu zentralen Aspekten befragt – und bei unpassenden Antworten – entsprechend feingetuned wird.

Manchmal gibst du Antworten, die sich nicht auf die grundlegenden Informationen zum konstruktiven Journalismus berufen. Nutze diese bereitgestellten Informationen für jede deiner Antworten.

Die Anpassung eines GPT für den konstruktiven Journalismus bietet zahlreiche Vorteile, darunter die Möglichkeit, effizienter und zielgerichteter zu arbeiten, sowie die Unterstützung bei der Identifizierung und Darstellung von Lösungen und positiven Entwicklungen. Herausforderungen bestehen in der kontinuierlichen Überwachung und Anpassung des GPT, um sicherzustellen, dass die Ergebnisse den gewünschten Qualitätsstandards entsprechen und keine Verzerrungen enthalten.

Kann KI synthetische Inhalte erkennen?

Wurde ein Text von einem Menschen verfasst oder von einer KI generiert? Wenn es eine KI war, stellt sich unter anderem die Frage: Wurde der Text womöglich nachträglich von einer anderen KI bearbeitet, um menschlicher zu wirken? Das erscheint wie ein großer digitaler Maskenball der Informationen. Für Journalistinnen und Journalisten wird es zunehmend schwieriger, herauszufinden, woher ein Text kommt. KI-Detektoren versprechen hier Abhilfe. Aber kann ausgerechnet KI zuverlässig dabei unterstützen, KI-generierte Inhalte zu erkennen?

KI-Detektoren sind spezialisierte Algorithmen, die darauf trainiert sind, künstlich erzeugte Inhalte zu erkennen. Sie nutzen eine Vielzahl von Techniken, um zwischen von KI generierten und von Menschen geschriebenen Texten zu differenzieren. Diese Techniken demaskieren die KI-erzeugten Inhalte, indem sie einzigartige Muster oder »Fingerabdrücke« identifizieren, die diese von menschlichen Beiträgen unterscheiden.

Indikator Gleichmäßigkeit

Dabei helfen Faktoren wie die Lesbarkeit, die Zeichensetzung, aber auch die Verwendung bestimmter Wörter und Phrasen. Ein weiterer Aspekt ist, wie stark sich die Struktur der Sätze in einem Dokument ähnelt. In KI-generierten Texten sind Wörter oder Formulierungen häufig gleichmäßig verteilt. Dadurch wird der Text vorhersehbarer. Genau diese Vorhersehbarkeit eines Textes ist für die KI-Detektoren ein zentraler Indikator dafür, dass ein Text künstlich generiert wurde.

Neben diesen Analysemethoden untersuchen KI-Detektoren auch den Kontext, indem sie Schreibstil und Konsistenz eines Textes betrachten. Ein markantes Beispiel dafür ist die stilometrische Untersuchung von Beiträgen der Gruppe QAnon, die unter anderem im Internet Verschwörungsmymen verbreitet. Durch den Einsatz Künstlicher Intelligenz konnten Forscher vor einiger Zeit den Schreibstil analysieren und so potenzielle Urheber bestimmter Verschwörungsmymen identifizieren. KI-basierte Stilanalysen ermöglichen es also auch, Autorinnen oder Autoren hinter Pseudonymen zu enttarnen.

Darüber hinaus setzen manche KI-Detektoren auf vorab trainierte Sprachmodelle und bewerten die Wahrscheinlichkeit, dass ein Inhalt von einer KI erzeugt wurde, danach, ob er ähnliche Merkmale aufweist wie die Inhalte, die das Sprachmodell selbst generiert.

Eine dritte Gruppe von KI-Detektoren, die ebenfalls auf Sprachmodellen basiert, wird speziell auf das Erkennen von synthetischen Texten trainiert. Diese Anwendungen lernen, mit Datensätzen, die sowohl von Menschen geschriebene als auch synthetische Texte enthalten, die Unterschiede zu erkennen und zu bestimmen, ob der Inhalt KI-generiert ist oder von einem Menschen stammt.

Viele Ungenauigkeiten

Trotz ihrer fortgeschrittenen Technologie stoßen KI-Detektoren jedoch schnell an ihre Leistungsgrenzen. So entfernte OpenAI im Juli 2023 das eigene ChatGPT-Erkennungstool AI-Classifier aufgrund »seiner geringen Genauigkeit« von seiner Webseite. Diese Entscheidung spiegelt ein komplexeres Problem wider, das auch in neueren Studien erkennbar wird. So zeigte eine im Dezember 2023 veröffentlichte Untersuchung zur Genauigkeit von 14 Detektionstools für KI-generierte Texte, dass fast jeder fünfte Text unerkannt blieb (vgl. Weber-Wulff et al. 2023).

Die möglichen Gründe für diese Ungenauigkeiten sind vielfältig. So haben KI-Detektoren oft Schwierigkeiten, mit neuen Schreibstilen und Techniken Schritt zu halten. Ein Mangel an Trainingsdaten kann ebenfalls zu Problemen führen – vor allem bei Texten, die nicht auf Englisch verfasst sind.

Eine zusätzliche Herausforderung für die Erkennung von synthetischen Inhalten stellen Tools dar, die KI-generierte Texte so verfeinern, dass sie menschliche Schreibfehler imitieren oder Schreibstile variieren. Damit fügen diese sogenannten »Humanizer-Tools« auf unserem digitalen Maskenball der Maskerade also eine weitere täuschende Maske hinzu.

So bleibt die Erkenntnis, dass KI-Detektoren ein nützliches, aber bei weitem nicht fehlerfreies Instrument für Medienschaffende bei der Erkennung von künstlich generierten Medieninhalten sein können. Selbst die raffiniertesten Detektoren können nur die Masken erkennen, die sie gelernt haben zu sehen.

Deep Fakes auf der Spur

Neben der Detektion von KI-generierten Texten spielen auch Technologien zur Erkennung von Deep Fakes eine große Rolle. Denn Deep Fakes stellen für die Gesellschaft eine besondere Gefahr dar: Manipulierte Videos und Fotos können die öffentliche Meinung stark beeinflussen. Fortgeschrittene KI-Detektoren, die Unstimmigkeiten in Bildern und Tönen erkennen, können daher hilfreiche Werkzeuge für Journalistinnen und Journalisten sein, um Desinformation zu verhindern und die Integrität ihrer Berichterstattung zu wahren.

Diese Art der KI-Detektoren analysiert Merkmale wie Gesichtsausdrücke und Lippenbewegungen. Obwohl diese Techniken immer weiterentwickelt werden, bleibt die Herausforderung bestehen – denn auch Deep Fakes werden stetig realistischer und schwieriger zu identifizieren.



Index und Glossar

Algorithmen sind klar definierte Anweisungsreihen zur Lösung von Problemen oder Erfüllung von Aufgaben. Ein Algorithmus erstreckt sich von einfachen Abfolgen von »Wenn-Dann-Bedingungen« bis zu anspruchsvollen Folgen mathematischer Gleichungen. Im Kontext Künstlicher Intelligenz analysieren so genannte Lernalgorithmen große Datenmengen und erkennen darin Muster, die sie abspeichern. Dieses Abspeichern von erkannten Zusammenhängen wird auch als »lernen« bezeichnet. Das ermöglicht es Maschinen zum Beispiel auch, sprachliche Inhalte zu erfassen und Entscheidungen zu formulieren. Lernalgorithmen bilden die Grundlage für Maschinelles Lernen, Künstliche Neuronale Netze und natürliche Sprachverarbeitung. → **Seite 62**

Automatisierung → **Seite 14**

Basismodelle (»foundation models«) sind KI-Modelle, die eine Reihe verschiedener Aufgaben erledigen und damit eine Vielfalt an Ergebnissen liefern können wie beispielsweise Texte, Bilder oder Ton. Sie funktionieren sowohl für sich genommen, können aber auch Grundlage für andere Anwendungen sein (vgl. Jones 2023). Um die Transparenz und Vergleichbarkeit dieser vielfältigen KI-Modelle zu verbessern, hat das Stanford Center for Research on Foundation Models (vgl. CRFM o.D.) einen Transparenz-Index für Basismodelle entwickelt, der Standards für Offenheit und Zugänglichkeit in der KI-Entwicklung setzt. → **Seite 21**

Bias → **Seite 34**

Chatbots sind technische Dialogsysteme, die den Austausch zwischen Menschen und Technik mittels gesprochener oder geschriebener Sprache ermöglichen. Man unterscheidet zwischen KI-basierten und regelbasierten Systemen. KI-basierte Chatbots lernen aus Interaktionen und können komplexe Antworten generieren, während

regelbasierte Chatbots vordefinierte Antworten auf spezifische Befehle liefern. ELIZA, der erste Chatbot aus den 1960er Jahren, nutzte vorprogrammierte Antworten zur Simulation eines Psychotherapeuten. → **Seiten 16, 20, 22 ff., 42, 59 ff.**

Diversity Penalty → **Seite 47**

Foundation Model → **Basismodelle**

GPT → **Seiten 20 f.**

generativ → **Seite 20**

Halluzination → **Seite 42**

Konstruktiver Journalismus → **Seiten 9 ff. und 49 ff.**

Künstliche Neuronale Netze → **Seiten 17 f.**

Large Language Model (LLM)
→ **Sprachmodelle**

Maschinelles Lernen (»machine learning«)
→ **Seite 16**

Parameter → **Seiten 46 f.**

Prompt → **Seiten 42 ff.**

Prompt-Engineering → **Seiten 42 ff.**

Refinement → **Seite 44**

Sprachmodelle (»language models«) sind KI-basierte Computerprogramme, die darauf trainiert sind, menschliche Sprache zu verstehen und zu generieren. Mittels → Künstlicher Neuronaler Netze und vieler Daten entwickeln sie während ihrer Trainingsphase statistische Modelle, mit denen sie die Wahrscheinlichkeit des Auftretens eines Wortes oder Wortteils nach einem gegebenen Textfragment vorhersagen können. So würde ChatGPT nach der Eingabe »Der Himmel ist« das Wort »blau« mit höherer Wahrscheinlichkeit nennen als das Wort »rot«. Von großen Sprachmodellen oder Large Language Models (LLM) spricht man, wenn die Computerprogramme mit extrem vielen Daten trainiert werden. → **Seiten 19 f., 34 f.**

Temperatur → **Seiten 46 f.**

Quellen und Literatur

Avelar, Victor/Patrick Donovan/Paul Lin/Wendy Torell/Maria A. Torres Arango (2023): *The AI Disruption: Challenges and Guidance for Data Center Design*, Schneider Electric Energy Management Research Center, [online] https://download.schneider-electric.com/files?p_Doc_Ref=SPD_WP110_EN&p_enDocType=White+Paper&p_File_Name=WP110_V2.1_EN.pdf.

Baumgartner, Katharina (2024): In den Kaninchenbau und wieder hinaus, in: Marie-Sophie Adeoso/Eva Berendsen/Leo Fischer/Deborah Schnabel (Hrsg.), *Code & Vorurteil: Über Künstliche Intelligenz, Rassismus und Antisemitismus*, Verbrecher Verlag, S. 125–136.

Beckett, Charlie (2019): *New powers, new responsibilities: A global survey of journalism and artificial intelligence*, London School of Economics and Political Science, [online]. Zum Download verfügbar unter: <https://blogs.lse.ac.uk/polis/2019/11/18/new-powers-new-responsibilities/> [abgerufen am 3. Juni 2024].

Beckett, Charlie/Mira Yaseen (2023): *Generating Change: A global survey of what news organisations are doing with AI*, London School of Economics and Political Science, [online]. Zum Download verfügbar unter: <https://www.journalism.ai.info/research/2023-generating-change> [abgerufen am 3. Juni 2024].

Becker, Kim Björn/Felix M. Simon/Christopher Crum (2023): *Policies in Parallel? A Comparative Study of Journalistic AI Policies in 52 Global News Organisations* [Preprint]. Verfügbar unter: <https://osf.io/preprints/socarxiv/c4af9> [abgerufen am 24. Mai 2024].

Bedford-Strohm, Jonas/Uli Köppen/Cécile Schneider (2020): *Ethik der Künstlichen Intelligenz: Unsere KI-Richtlinien im Bayerischen Rundfunk*, in: *Bayerischer Rundfunk*, 30. November 2020, [online] <https://br.de/extra/ai-automation-lab/ki-ethik-100.html> [abgerufen am 24. Mai 2024].

Bender, Emily M./Timnit Gebru/Angelina McMillan-Major/Shmargaret Shmitchell (2021): *On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? in: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FACCT 21)*, S. 610–623. <https://doi.org/10.1145/3442188.3445922> [abgerufen am 24. Mai 2024].

Bommasani, Rishi/Kevin Klyman/Shayne Longpre/Sayash Kapoor/Nestor Maslej/Betty Xiong/Daniel Zhang/Percy Liang (2023): *The Foundation Model Transparency Index* [Preprint]. Verfügbar unter: <https://arxiv.org/abs/2310.12941> [abgerufen am 24. Mai 2024].

Brammer, Robert (2023). *Die Macht der Computer. Zum 100. Geburtstag von Joseph Weizenbaum*, in: *Deutschlandfunk Kultur. Aus den Archiven*, 14. Januar 2023, [online] <https://deutschlandfunkkultur.de/weizenbaum-100.html> [abgerufen am 29. Mai 2024].

Brooks, Rodney (2018): *Die Ursprünge der Künstlichen Intelligenz*, in: *reframe Tech*, 22. November 2018, [online] <https://reframetech.de/2018/11/22/die-urspruenge-der-kuenstlichen-intelligenz/> [abgerufen am 18. April 2024].

Burgess, Matt/Reece Rogers (2024): *How to Stop Your Data From Being Used to Train AI*, in: *Wired*, 10. April 2024, [online] <https://wired.com/story/how-to-stop-your-data-from-being-used-to-train-ai/> [abgerufen am 17. Mai 2024].

Center for Research on Foundation Models [CRFM] (o. D.): *Foundation Model Transparency Index*, [online] <https://crfm.stanford.edu/fmti/> [abgerufen am 16. Mai 2024].

Clarke, Seán/Dan Milmo/Garry Blight (2023): *How AI chatbots like ChatGPT or Bard work – visual explainer*, in: *The Guardian*, 2. November 2023, [online] <https://theguardian.com/technology/ng-interactive/2023/nov/01/how-ai-chatbots-like-chatgpt-or-bard-work-visual-explainer> [abgerufen am 24. Mai 2024].

Commissariat à l'énergie atomique et aux énergies alternatives [CEA] (2018): *[Histoire des sciences] L'histoire de l'intelligence artificielle (IA)* [YouTube], <https://youtube.com/watch?v=qmwjx-r5vmw>.

Common Crawl (2024): *Overview*, [online] <https://commoncrawl.org/overview> [abgerufen am 24. Mai 2024].

Council of Europe (2023): *Guidelines on the responsible implementation of artificial intelligence systems in journalism*, [online] <https://rm.coe.int/cdmsi-2023-014-guidelines-on-the-responsible-implementation-of-artific/1680adb4c6>.

DataScientest (2023): Künstliche Intelligenz: Definition, Geschichte, Chancen und Risiken, in: *DataScientest*, 8. August 2023, [online] <https://datascientest.com/de/kuenstliche-intelligenz> [abgerufen am 18. April 2024].

De Vries, Alex (2023): The growing energy footprint of artificial intelligence, in: *Joule*, Bd. 7, Nr. 10, S. 2191–2194, <https://doi.org/10.1016/j.joule.2023.09.004>.

Der Spiegel (2021): Einzelanalysen für alle 299 Wahlkreise, in: *Der Spiegel*, 27. September 2021, [online] <https://spiegel.de/politik/deutschland/bundestagswahl-2021-alle-wahlkreis-ergebnisse-der-btw-a-6c222d37-87d8-4d85-bec9-3d95e6e0eefa> [abgerufen am 24. Mai 2024].

Deutscher Journalisten-Verband [DJV] (o.D.): Positionspapier bezüglich des Einsatzes Künstlicher Intelligenz im Journalismus, [online]. Zum Download verfügbar unter: <https://djv.de/medienpolitik/kuenstliche-intelligenz/> [abgerufen am 24. Mai 2024].

Deutscher Presserat (2024): Jahresbericht 2023, [online] https://presserat.de/files/presserat/bilder/Downloads%20Jahresberichte/DPR-2024-Jahresbericht%202023_BF.pdf [abgerufen am 24. Mai 2024].

Diakopoulos, Nicholas (2019): *Automating the News: How Algorithms are Rewriting the Media*, Harvard University Press.

Facklam, Anton Yong (2024): »Superintelligenz« bei Meta: Mark Zuckerberg plant KI-Revolution, in: *CHIP*, 31. Januar 2024, [online] https://chip.de/news/Kuenstliche-Intelligenz-bei-Meta-Mark-Zuckerberg-plant-AGI-Revolution_185116013.html [abgerufen am 23. April 2024].

Fraunhofer-Institut für Kognitive Systeme [IKS] (2024): Künstliche Intelligenz (KI) und maschinelles Lernen, [online] <https://iks.fraunhofer.de/de/themen/kuenstliche-intelligenz.html> [abgerufen am 18. April 2024].

Gartner (2022): Gartner Unveils Top Predictions for IT Organizations and Users in 2023 and Beyond, 18. Oktober 2022, [online] <https://gartner.com/en/newsroom/press-releases/2022-10-18-gartner-unveils-top-predictions-for-it-organizations-and-users-in-2023-and-beyond> [abgerufen am 28. Mai 2024].

Hastein-Teßmer, Oliver (2024): *Digitaler Erfolg im Lokaljournalismus – Leitfaden für Praxis und Ausbildung in Redaktionen*, Springer VS.

Heinrichs, Ellen (2018): Wir müssen reden: über Künstliche Intelligenz und Ethik im Journalismus, in: *Medium*, 30. Oktober 2018, [online] <https://ellenheinrichs.medium.com/https-medium-com-ellen-sch-medienethik-und-ki-c953c2d70fba> [abgerufen am 24. Mai 2024].

Jones, Elliot (2023): Explainer: What is a foundation model?, Ada Lovelace Institute, 17. Juli 2023, [online] https://adalovelaceinstitute.org/resource/foundation-models-explainer/#_ftn5 [abgerufen am 24. Mai 2024].

Klofat, Aleksandra (2023): Wie funktionieren Transformer? Definition und Praxis, in: *Informatik Aktuell*, 16. Mai 2023, [online] <https://informatik-aktuell.de/betrieb/kuenstliche-intelligenz/wie-funktionieren-transformer-definition-und-praxis.html#c40768> [abgerufen am 24. Mai 2024].

Landesanstalt für Medien NRW (2024): Akzeptanz von Prozess-Automatisierung und künstlicher Intelligenz in der Medienbranche, [online] https://medienanstalt-nrw.de/fileadmin/user_upload/lfm-nrw/Forschung/LFM-NRW_Akzeptanz_von_Prozessautomatisierung_und_Kuenstlicher_Intelligenz_in_der_Medienbranche.pdf.

McBride, Kelly (2024): Your newsroom needs an AI ethics policy. Start here., Poynter Insitute, 25. März 2024, [online] <https://poynter.org/ethics-trust/2024/how-to-create-newsroom-artificial-intelligence-ethics-policy/> [abgerufen am 24. Mai 2024].

McCarthy, John/Marvin L. Minsky/Nathaniel Rochester/Claude E. Shannon (1955): A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence, [online] <http://jmc.stanford.edu/articles/dartmouth/dartmouth.pdf>.

Newman, Nic/Richard Fletcher/Craig T. Robertson/ Amy Ross Arguedas/Rasmus Kleis Nielsen (2024): Reuters Institute Digital News Report 2024, Reuters Institute for the Study of Journalism, [online] https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2024-06/RIS/DNR_2024_Digital_v10%20lr.pdf.

OpenAI (2018): AI and compute, 16. Mai 2018, [online] <https://openai.com/index/ai-and-compute/> [abgerufen am 31. Mai 2024].

Patterson, David/Joseph Gonzalez/Urs Hölzle/Quoc Le/Chen Liang/Lluis-Miquel Munguia/Daniel Rothchild/David So/Maud Texier/Jeff Dean (2022): The Carbon Footprint of Machine Learning Training Will Plateau, Then Shrink [Preprint]. Verfügbar unter: <https://arxiv.org/abs/2204.05149> [abgerufen am 31. Mai 2024].

Porlezza, Colin (2024): The datafication of digital journalism: A history of everlasting challenges between ethical issues and regulation, in: *Journalism*, Bd. 25, Nr. 5, S. 1167-1185, <https://doi.org/10.1177/14648849231190232>.

Precedence Research (2024): Artificial Intelligence (AI) in Hardware Market Size to Reach USD 473.53 Bn By 2033, in: *globoNewsWire.com*, 6 März 2024, [online] <https://www.globenewswire.com/news-release/2024/03/06/2841613/0/en/Artificial-Intelligence-AI-in-Hardware-Market-Size-to-Rreach-USD-473-53-Bn-By-2033.html> [abgerufen am 31. Mai 2024].

Reporter ohne Grenzen (2023): Paris Charta on AI and Journalism, [online] <https://rsf.org/sites/default/files/medias/file/2023/11/Paris%20charter%20on%20AI%20in%20journalism.pdf>.

Rinehart, Aimee/Ernest Kung (2022): Artificial Intelligence in Local News: A survey of US newsrooms' AI readiness, The Associated Press, [online]. Zum Download verfügbar unter: <https://doi.org/10.13140/RG.2.2.16926.82246>.

Schweizer Presserat (2024): KI im Journalismus: Leitfaden, [online] https://presserat.ch/journalistenkodex/ki_leitfaden/ [abgerufen am 17. Mai 2024].

Seward, Zach (2024): AI news that's fit to print: The New York Times' new AI leader on what this powerful tech can do for journalism, Reuters Institute for the Study of Journalism, 13. März 2024, [online] <https://reutersinstitute.politics.ox.ac.uk/news/ai-news-thats-fit-print-new-york-times-new-ai-leader-what-powerful-tech-can-do-journalism> [abgerufen am 27. März 2024].

Stieler, Wolfgang/Katharina Zweig (2024, 28. März): Deep Dive: Für eine Zukunft nach ChatGPT und Co. [Audio-Podcast], MIT Technology Review – Der Podcast, [online] <https://tech2go.podigee.io/105-neue-episode>.

Strubell, Emma/Ananya Ganesh/Andrew McCallum (2019): Energy and Policy Considerations for Deep Learning in NLP [Preprint]. Verfügbar unter: <https://arxiv.org/abs/1906.02243> [abgerufen am 31. Mai 2024].

TRG Datacenters (2023): AI Chatbots: Energy usage of 2023's most popular chatbots (so far), [online] <https://www.trgdatacenters.com/resource/ai-chatbots-energy-usage-of-2023s-most-popular-chatbots-so-far/> [abgerufen am 31. Mai 2024].

Turing, Alan M. (2023): *Computing Machinery and Intelligence – Können Maschinen denken? Englisch/Deutsch* (2., durchges. und erg. Aufl.), Reclam.

Weber-Wulff, Debor/Alla Anohina-Naumeca/Sonja Bjelobaba/Tomáš Foltýnek/Jean Guerrero-Dib/Olumide Popoola/Petr Šigut/Lorna Waddington (2023): Testing of detection tools for AI-generated text, in: *International Journal for Educational Integrity*, Bd. 19, Nr. 26, <https://doi.org/10.1007/s40979-023-00146-z>.

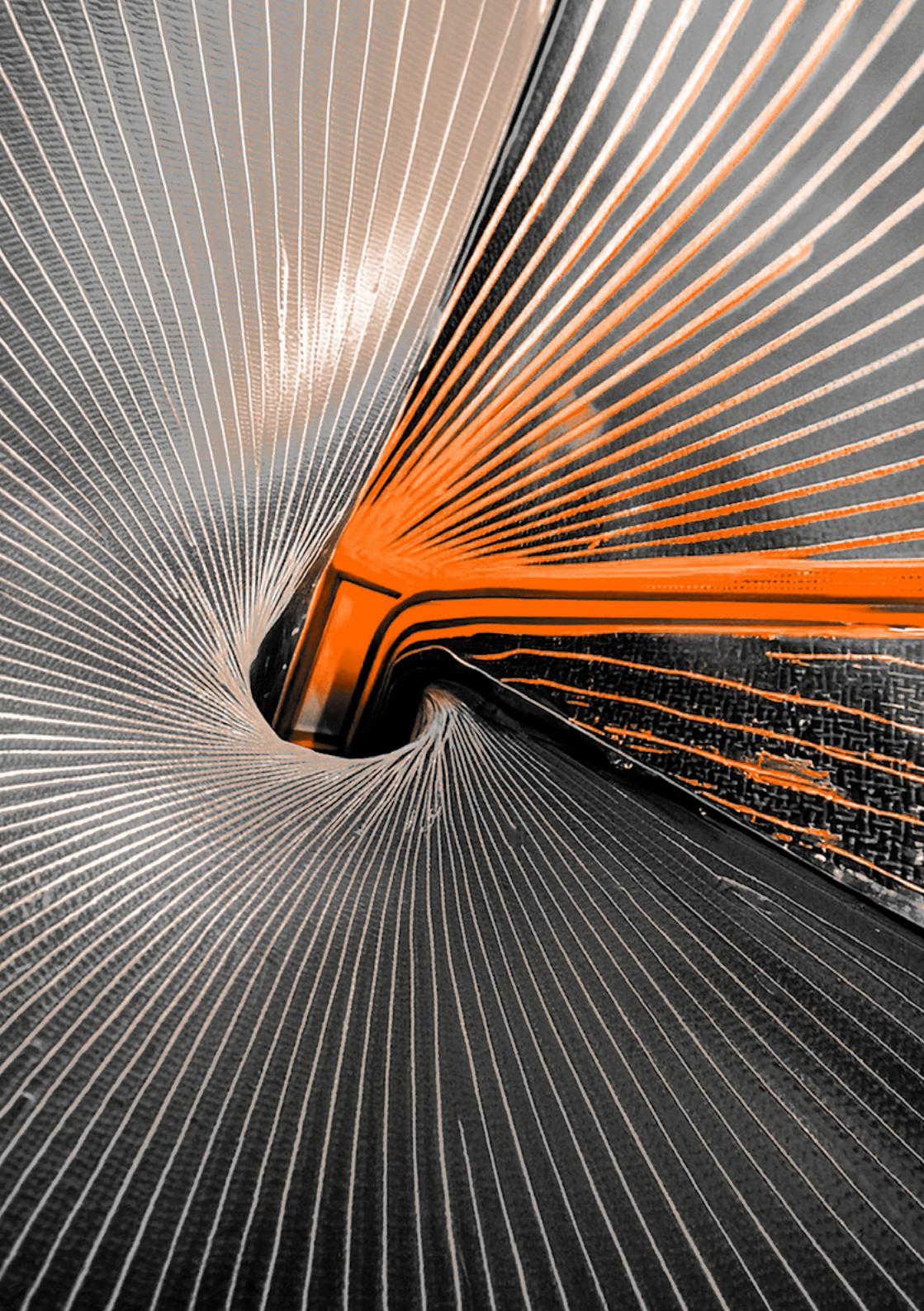
Whittaker, Meredith (2023): Vermessung bis ins Innerste, in: *netzpolitik.org*, 7. Juni 2023, [online] <https://netzpolitik.org/2023/kuenstliche-intelligenz-vermessung-bis-ins-innerste/> [abgerufen am 24. Mai 2024].

Wired (2023): How *Wired* Will Use Generative AI Tools, 22. Mai 2023, [online] <https://www.wired.com/about/generative-ai-policy/> [abgerufen am 5. Juli 2024].

Zentrum für vertrauenswürdige Künstliche Intelligenz [ZVKI]: Glossar, [online] <https://zvki.de/ki-navigator/unsere-inhalte/glossar> [abgerufen am 29. Mai 2024].

Zweig, Katharina (2023): *Die KI war's! Von absurd bis tödlich: Die Tücken der künstlichen Intelligenz* (4. Aufl.), Heyne.

<prompt> abstract, rough, lines, clean, less detail, technical, black and white with orange, to the point </prompt>



Das Team



Claudia Isabel Rittel hat mehr als zehn Jahre in verschiedenen Rollen in der Regionalredaktion der *Frankfurter Rundschau* gearbeitet. Nach einer Station bei der regionalen Mediengruppe VRM in Mainz, wo sie als Editorial Developerin tätig war, kam sie 2023 zum Bonn Institute. Als Referentin Projekte und Forschung hat die Politikwissenschaftlerin das Projekt, in dessen Rahmen der *Konstruktive KI-Kompass* entstanden ist, koordiniert und die Produktion von Booklet und Webangebot betreut. Zudem hat sie mehrere Beiträge für die Publikation verfasst – darunter die Kapitel »Einladung«, »KI verstehen« und Teile des Kapitels »KI verantwortungsvoll nutzen«.



Peter Lindner hat lange für die *Süddeutsche Zeitung* gearbeitet, fast zwölf Jahre als Ressortleiter Politik und mehr als sechs Jahre als stellvertretender Chefredakteur von *sz.de*. Seit dieser Zeit beschäftigt er sich intensiv mit digitaler Transformation und Innovationen im Journalismus. 2022 hat er am Constructive Institute in Dänemark dazu geforscht, wie Journalismus den konstruktiven Austausch in der Gesellschaft befördern und noch besser zur Stärkung der Demokratie beitragen kann. Am Bonn Institute leitet der Politikwissenschaftler den Bereich Projekte und Forschung. Er hat das Projekt *Konstruktiver KI-Kompass* federführend konzipiert und die Umsetzung auf verschiedenen Ebenen begleitet.



Professor Dr. Ranty Islam lehrt Digitalen Journalismus an der Media University of Applied Sciences. Seine Schwerpunkte sind die digitale Transformation von Nachrichtenredaktionen, die Rolle von Empathie in neuen Konzepten von Journalismus, konstruktiver Journalismus und Künstliche Intelligenz. Im Wintersemester 2023/2024 hat er mit einer Studiengruppe Tests für dieses Projekt gemacht. Zudem hat er seine Expertise beim Überprüfen der Fakten eingebracht.



Patrick Große ist Gründer von The AI Journalist und Journalist. Er leitet die Redaktion des paneuropäischen Medienprojekts »ENTR what's next« bei der *Deutschen Welle*. Sein Fokus liegt auf der Anwendung von Künstlicher Intelligenz im Journalismus, vor allem im Bereich des Prompt-Engineering. Für diese Publikation hat er mit dem Bonn Institute die Prompts entwickelt und maßgeblich die Beiträge für den Teil »KI konstruktiv einsetzen« verfasst.



Dr. Ralph Hintemann forscht seit mehr als 20 Jahren zu den Wechselwirkungen von Digitalisierung und Nachhaltigkeit. Aktuell beschäftigt sich der Ingenieur und Wirtschaftswissenschaftler mit den Potenzialen und Risiken neuer Technologien wie Künstliche Intelligenz und dem Internet der Dinge. Ein Schwerpunkt seiner Forschung liegt in der Analyse des Energie- und Ressourcenbedarfs für digitale Infrastrukturen. Für den *Konstruktiven KI-Kompass* hat er den Beitrag »Auswirkungen auf die Umwelt« verfasst.



Dr. Peter Niepalla war von 2003 bis 2023 Justiziar und Leiter der Rechtsabteilung der *Deutschen Welle*. Zudem lehrt er seit vielen Jahren an verschiedenen Hochschulen Medienrecht. Aktuell arbeitet der Rechtsanwalt als Berater und Gutachter zu Fragen des Vertrags-, Arbeits- und Medienrechts. So auch beim *Konstruktiven KI-Kompass*: Mit seiner Expertise hat er die inhaltliche Ausgestaltung des Kapitels »KI verantwortungsvoll nutzen« unterstützt und alle Fragen zu rechtlichen Aspekten einer kritischen Prüfung unterzogen.



Benedikt Dahlmann ist Volontär der *Augsburger Allgemeinen Zeitung* und war Anfang 2024 Hospitant am Bonn Institute. Er hat einen Master in Politischer Kommunikation der Universität Düsseldorf. Seine Abschlussarbeit, die sich mit der Wirkung von konstruktivem Journalismus auf das Publikum beschäftigt, wurde mit dem Heinrich-Heine-Journalismuspreis 2024 ausgezeichnet. Für den *Konstruktiven KI-Kompass* hat er die Beiträge »Interne Leitlinien« und »Tipps für die Entwicklung eines KI-Kodex« verfasst.



Tobias Kohmann ist Werkstudent beim Bonn Institute. Er studiert seit 2020 Politik und Gesellschaft an der Rheinischen Friedrich-Wilhelms-Universität Bonn und hat erste berufliche Erfahrungen bei Phoenix und der Journalist_innen-Akademie der Friedrich-Ebert-Stiftung gesammelt. Seit Ende 2022 arbeitet er als Reporter für das Uni-Radio *BonnFM*. Für den *Konstruktiven KI-Kompass* hat er die Beiträge zu Transkription, KI-Detektoren und Anbieter großer Sprachmodelle geschrieben und die Projektorganisation unterstützt.

Danksagung

Diese Publikation wäre ohne die Unterstützung vieler wunderbarer Menschen nicht möglich gewesen.

Danke möchten wir zuvorderst den Mitarbeitenden unserer Kooperationspartner sagen, die dieses Projekt von Beginn an konstruktiv begleitet haben. Danke an Simon Hof und Donata Dröge von *RTL* für den intensiven Austausch über die Inhalte, die Einblicke in die vielfältigen Aktivitäten rund um KI in der Praxis bei *RTL* – unter anderem bei Redaktionsbesuchen in Köln und Berlin. Danke an Ruben Bouwmeester und Tilman Wagner von der *Deutschen Welle* für Inspiration und Input zu Beginn des Projektes und für die Unterstützung beim Faktencheck und Korrekturlesen zum Ende hin. Danke an Peter Damgaard und Gustav Aarup Lauridsen vom *Constructive Institute* in Aarhus (Dänemark) für Anregungen, kritisches Hinterfragen und für das Einbringen ihrer technischen Expertise.

Vielen Dank an Prof. Dr. Ranty Islam für die wissenschaftliche Perspektive und die Unterstützung beim Faktencheck und an seine Studierenden für die Testings, bei denen sie sich mit grundlegenden Fragen für den konstruktiven Einsatz von Prompts im journalistischen Alltag beschäftigt haben. Auf diesen aufbauend hat das *Bonn Institute* gemeinsam mit Patrick Große die hier vorgestellten Prompts entwickelt.

Besonderer Dank gilt weiterhin Kai Kullen und Nikolaus Rulle von *KKNR* für das innovative grafische Konzept, das stimmige Layout und die gelungenen KI-generierten Illustrationen. Oliver Schwarz von *schwarzdesign* danken wir zudem für die professionelle und ansprechende Aufbereitung der Inhalte für die Webseite des *Bonn Institute*.

Alle Texte wurden handwerklich und ohne KI erstellt. Daran mitgewirkt haben verschiedene Personen, die wir auf den Seiten 70 und 71 vorstellen. Ihnen allen sei hier gedankt. Doch was wäre ein gutes Buch ohne Lektorat? Herzlichen Dank an dieser Stelle an Beate Kohmann vom Lektorat *Wortgut* für die akribische Durchsicht. Ein Dank geht auch an das gesamte Team des *Bonn Institute* für die tatkräftige Unterstützung.

Nichts von alledem jedoch wäre möglich gewesen ohne die Initiative der Gründerin und Geschäftsführerin des *Bonn Institute*, Ellen Heinrichs.

Bonn im Juni 2024

Claudia Isabel Rittel und Peter Lindner

IMPRESSUM

Herausgeberin

Bonn Institute gGmbH
Bennauerstraße 60
53115 Bonn

Vertreten durch

Ellen Heinrichs, Geschäftsführerin

Kontakt

Telefon +49 (0) 228 28653370
hello@bonn-institute.org
www.bonn-institute.org

Inhaltliches Konzept

Peter Lindner
Claudia Isabel Rittel

Text und Recherche

Claudia Isabel Rittel
Patrick Große
Dr. Ralph Hintemann
Benedikt Dahlmann
Tobias Kohmann
Peter Lindner
Prof. Dr. Ranty Islam
Dr. Peter Nipalla

Testings

Patrick Große
Prof. Dr. Ranty Islam und Studierende:
Felix Früh, Adriana Kauffmann, Calvin Kliemann,
Alina Langreck, Jolanda Ost, Lisa Schubert,
Alina Schunk, Lisa Vogt, Vanessa Wobb

Graphisches Konzept und Gestaltung

Kai Kullen und Nikolaus Rulle
kknr.de

Illustration

KKNR mit Unterstützung durch
die generative KI Adobe Firefly

Lektorat

Beate Kohmann
lektorat-wortgut.de

Druck

Brandt GmbH, Bonn

Bildnachweis

LFM-NRW S. 3; Fotostudio Bruder S. 70 (1);
Joerg Frank, Borderstep S. 71 (1);
Deutsche Welle S. 71 (2);
alle weiteren Bilder privat

Erste Auflage Juni 2024

Alle Rechte vorbehalten. Diese Publikation
oder Teile dieser Publikation dürfen nicht
vervielfältigt, gespeichert oder in irgendeiner
Form übertragen werden ohne die schriftliche
Genehmigung der Bonn Institute gGmbH.

© 2024 Bonn Institute

ISBN 978-3-9826202-3-7

Fünf Einstiege ins konstruktive Prompting

Welche spezifischen Aspekte könnte ich bei der konstruktiven Berichterstattung über [Thema] hervorheben, um verschiedene Perspektiven und Lösungsmöglichkeiten darzustellen?

Identifiziere und erkläre die verschiedenen Dimensionen und Einflussfaktoren, die bei [Thema] eine Rolle spielen. Stelle dir dabei die Frage: Wer ist noch betroffen – abseits der bekannten Akteure?

Generiere Fragen, die darauf abzielen, vergangene oder potenzielle Erfolge zu erkunden. Zum Beispiel: »Können Sie ein Beispiel nennen, bei dem ein ähnliches Problem erfolgreich gelöst wurde, und was können wir daraus lernen?«

Ich suche nach Themen im Bereich Umwelt- und Klimaschutz, die sich für konstruktiven Journalismus eignen. Nenne mir zehn mögliche Themen innerhalb dieses Themenkomplexes, die ich als Journalistin umsetzen könnte.

Betrachte [Thema] aus der Perspektive einer einzelnen betroffenen Person und der gesamten Gesellschaft. Erkläre, wie sich die Ziele dieser beiden Ebenen voneinander unterscheiden.

Künstliche Intelligenz verstehen und nutzen, um die eigene journalistische Arbeit konstruktiver zu gestalten – das dazugehörige Wissen vermittelt der *Konstruktive KI-Kompass*.

Für den praktischen Einsatz gibt es zahlreiche Prompts, die bei der Ideenfindung, Recherche und Interviewvorbereitung unterstützen können.

Wie funktioniert Künstliche Intelligenz? Wo kommt sie her und welche Auswirkungen hat die Technologie auf die Umwelt? In welchem rechtlichen Rahmen sind KI-Tools verantwortungsvoll anwendbar und worauf sollten Redaktionen bei der Konzeption interner Leitlinien achten?

Diese und weitere Fragen beantwortet der *Konstruktive KI-Kompass* – er bietet Orientierung und Inspiration für die tägliche Praxis.

b

bonn institute

ISBN

2-0