

Anwendungsrealität der KI: Zwischen Durchbruch und Dilemma

David Limacher

Seit vielen Monaten steht künstliche Intelligenz (KI) im Zentrum der öffentlichen Aufmerksamkeit, mit bereits spürbaren Auswirkungen auf die Wirtschaft und das Alltagsleben. Warum tritt gerade jetzt, fast 70 Jahre nach der erstmaligen Prägung dieses Begriffs, eine derartige Veränderung ein? Was lässt sich konkret mit KI, insbesondere ChatGPT, bewirken und welche Herausforderungen und Risiken sind damit in der Praxis verbunden?

Technologien unterliegen einer kontinuierlichen Evolution und Verbesserung, bis sie einen Punkt erreichen, an dem die breite Öffentlichkeit ihren Nutzen adaptiert und auf kreative Weise einsetzt. Ähnlich verhielt es sich mit dem Internet, dessen Vorläufer bereits in den 1960er-Jahren entstand. Anwendungen wie soziale Medien, Uber oder Twint waren damals noch unvorstellbar. Erst durch die ständige Anpassung und Innovation, zum Beispiel durch die Einführung von Smartphones, entstanden Anwendungen, die heute aus unserem Alltag nicht mehr wegzudenken sind.

Woher kommt diese Technologie?

Künstliche Intelligenz, wie wir sie heute kennen, zum Beispiel bei ChatGPT, basiert auf einer Technik, die neuronale Netzwerke genannt wird. Diese Technik ist inspiriert von der Funktionsweise des menschlichen Gehirns und existiert schon fast so lange wie der Begriff künstliche Intelligenz selbst. Das erste GPT-Modell (basierend auf Transformer-Netzwerken), ähnlich den heutigen Versionen, wurde 2018 vorgestellt. Der Hauptunterschied zwischen diesem ersten Modell und dem 2024 aktuellen Modell GPT-4 liegt in der Größe und Komplexität: GPT-4 verwendet viel mehr Daten zum Lernen, welche als Beispiele gegeben werden, und kann dadurch Informationen präziser verarbeiten.

Ausserdem war das ursprüngliche GPT darauf ausgelegt, Texte zu ergänzen, und nicht, wie das heutige ChatGPT, Dialoge zu führen. Diese scheinbar kleine, aber dennoch bahnbrechende Modi-

fikation hat es ermöglicht, dass die Anwendung sehr einfach und ohne spezifisches Vorwissen von jedem genutzt werden kann. Dies trug massgeblich zum Hype um ChatGPT im letzten Jahr bei.

Wie unterscheidet sich ChatGPT von früherer KI?

Doch ChatGPT ist nicht nur ein faszinierendes Spielzeug, das beispielsweise ein lustiges Gedicht über Fondue im Stil Goethes generieren kann. Es verändert grundlegend die Art und Weise, wie wir recherchieren, arbeiten, schreiben und nach Ideen suchen. Natürlich hat auch die Industrie grosses Interesse daran, Arbeit durch den Einsatz dieser Technologie effizienter und besser zu gestalten. Welche Anwendungen dies sind und welche Herausforderungen dabei auftreten, werde ich anhand einiger Beispiele erläutern, die bei meinem Arbeitgeber Artifact, ein Unternehmen, das sich auf den Einsatz von KI spezialisiert hat, realisiert wurden.

Im Vergleich zur künstlichen Intelligenz, wie sie vor der Einführung von Large Language Models (LLMs) wie ChatGPT eingesetzt wurde, sind diese Modelle nicht nur auf eine singuläre Aufgabe spezialisiert, sondern können eine Vielzahl von sprachbasierten Anwendungen bewältigen. Früher, wenn man beispielsweise das Sentiment einer Rezension automatisch als positiv oder negativ klassifizieren wollte, benötigte man umfangreiche Datensätze von Rezensionen mit bereits zugewiesenen Kategorien. Basierend darauf wurde eine KI anhand vieler Beispiele trainiert, die jedoch nur diese eine Aufgabe erfüllen konnte. Der Aufwand, insbesondere für das Erheben der Daten, war oft enorm. Mit einer Technologie, die Sprache grundsätzlich versteht und analysieren kann, ist so eine Aufgabe wesentlich einfacher geworden. Daher werden diese im EU AI Act, einer neuen Gesetzgebung für künstliche Intelligenz, auch als «KI-Systeme mit allgemeinem Verwendungszweck» bezeichnet.

Wie hat sich die Arbeit im Bereich KI verändert?

Die Arbeit im Bereich der künstlichen Intelligenz hat sich daher teilweise grundlegend verändert. Während zuvor ein signifikanter Teil der Arbeit darin bestand, Daten in der erforderlichen Qualität zu kuratieren und eine KI zu trainieren, spielt heute die Erfahrung mit Tools wie ChatGPT eine zentrale Rolle. Es geht dabei um sogenanntes Prompt Engineering, also das Entwickeln von Anfragen im Chat, die die KI dazu bringen, das gewünschte Ergebnis mit

den richtigen Eigenschaften, wie beispielsweise Textlänge, Stil oder einem spezifischen Format, zu liefern. Zudem ist die Integration dieser Verarbeitung in bestehende Arbeitsabläufe ein wesentlicher Aspekt zusammen mit der Herausforderung, dass sich die Modelle rasant entwickeln und neue Anbieter auf den Markt kommen.

Anwendungen von ChatGPT

Was kann ChatGPT bei Textanalysen leisten?

Die Fähigkeit von Computern, grosse Datenmengen zu verarbeiten, übertrifft grundsätzlich die des Menschen. Dies beschränkt sich nicht mehr nur auf Rechenoperationen, sondern erstreckt sich auch auf die Analyse grosser Textmengen. Dies wurde in einem Projekt für einen Kunden genutzt, um alle Erwähnungen dieser grossen Organisation in den Medien zu analysieren und täglich automatisch für interne Interessensgruppen zusammenzufassen. Hierbei kam nicht die bekannte Chat-Oberfläche zum Einsatz, sondern eine Schnittstelle, über die ein Computerprogramm Antworten ohne menschliches Zutun erhalten und verarbeiten kann. Aufgrund des begrenzten Kontextverständnisses (Länge der Eingaben) von ChatGPT musste die Verarbeitung der grossen Textmengen in mehreren Schritten erfolgen. Zunächst können die Artikel gefiltert werden, um ihre Relevanz zu bestimmen. Eine entsprechende Instruktion (Prompt) definiert, wann ein Artikel relevant ist, und weist ChatGPT an, lediglich mit «Ja» oder «Nein» zu antworten. Die so gefilterten Artikel können anschliessend auf wenige Kernpunkte reduziert und darauf basierend massgeschneiderte Zusammenfassungen für die Interessengruppen innerhalb der Organisation erstellt werden. Eine solche Abfolge unterschiedlicher Instruktionen – Kategorisieren von Relevanz, Extrahieren von Kernpunkten, Erstellen von Zusammenfassungen – ist oft sinnvoll, um bessere Ergebnisse zu erzielen und die später besprochenen Limitationen zu minimieren. Tipps zur Anwendung solcher Techniken finden sich in der Infobox zum Thema Prompt Engineering.

In Unternehmen existiert oft eine Fülle von internem Wissen, das in unzähligen Dokumenten gespeichert ist. Auch hier bietet sich der Einsatz von ChatGPT an. Da jedoch die Menge an Text häufig zu gross ist, um sie direkt zu verarbeiten, werden einzelne Texte herangezogen, um darauf basierend eine Antwort zu generieren.

Bei einer im Chat gestellten Frage wird mit Hilfe einer weiteren KI, die ähnlich wie ChatGPT funktioniert, nach für die Beantwortung relevanten Dokumenten gesucht, die dann in den Kontext eingebunden werden. Fragt beispielsweise ein Mitarbeiter nach den Erfolgen im Jahr 2010 im Vergleich zu 2020, werden die entsprechenden Kapitel aus den Geschäftsberichten von 2010 und 2020 gefunden, um daraus eine Antwort zu generieren. Solche Anwendungen dürften in naher Zukunft breit eingesetzt werden, um das interne Wissen in Unternehmen möglichst effizient nutzen zu können.

Neben der internen Verwendung, bei der die Qualität der Antworten nicht unbedingt perfekt sein muss, wird zunehmend auch der Einsatz für Kundeninteraktionen geprüft. Hierbei kommt dieselbe Technologie zum Einsatz, um Benutzeranleitungen und häufig gestellte Fragen (FAQs) zu verarbeiten. Diese können dann den Kundenservice unterstützen oder sogar zu einem gewissen Grad ersetzen.

Bei solchen Anwendungen ist allerdings besondere Sorgfalt geboten, da die Qualität der Antworten direkt die Kundenzufriedenheit und das Vertrauen in das Unternehmen beeinflussen können. Es gibt Herausforderungen, wie beispielsweise die Gewährleistung der Aktualität der Informationen, die Vermeidung von Missverständnissen oder die korrekte Interpretation von Benutzeranfragen. Die Entwicklung solcher Systeme erfordert daher eine sorgfältige Planung und ständige Überwachung, um sicherzustellen, dass sie den Bedürfnissen der Kunden gerecht werden.

Welche Anwendungen gibt es für kürzere Texte?

Neben dem Analysieren und Durchsuchen grosser Textmengen gibt es auch Anwendungen, die sich mit der Verarbeitung von weniger Text befassen. Eine solche Anwendung ist das Extrahieren spezifischer Informationen aus unstrukturierten Daten, wie es bei Fliesstext der Fall ist. Ein konkretes Beispiel ist das Kategorisieren eingehender E-Mails. Heutzutage muss das Mailpostfach der allgemeinen Kontaktadresse eines Unternehmens von einer Person bearbeitet werden, die die Anfragen den jeweiligen Fachabteilungen zuordnet. Mit den richtigen Instruktionen und dem Wissen über die Aufgaben der einzelnen Abteilungen in einem Prompt kann dieser Prozess automatisiert werden. Dabei können gleichzeitig strukturierte Informationen wie beispiels-

weise der Name, die Kundennummer oder das betroffene Produkt aus dem Fliesstext extrahiert werden, um das spätere Auffinden oder das Erstellen von Statistiken zu erleichtern.

Wie können Mitarbeitende KI selber nutzen?

Mitarbeitende können durch das Wissen um die richtige Anwendung von künstlicher Intelligenz auch individuell einen Mehrwert in ihrer täglichen Arbeit schaffen. Eine Möglichkeit besteht darin, diese Tools für das Brainstorming einzusetzen. Möchte man beispielsweise in einem Gastronomieunternehmen neue Namen für Gerichte entwickeln, können die Zutaten und die Beschreibung des Gerichts zusammen mit der Philosophie des Restaurants eingegeben werden, woraufhin ChatGPT Namensvorschläge generiert. Die Bandbreite der Ideen, die so abgefragt werden können, ist nahezu unbegrenzt und reicht von Metaphern für einen Vortrag bis hin zu neuen Ansätzen zur Produktoptimierung. Diese Ideen müssen zwar gefiltert werden, können aber als wertvolle Inspirationsquelle dienen.

Die Kernkompetenz von ChatGPT liegt im Umgang mit Sprache, und es bietet sich daher an, Texte zu generieren, zu verbessern oder umzugestalten. Dies kann von einer Übersetzung in eine andere Sprache bis hin zu einer Stiländerung reichen, um einen Text intellektueller, formeller oder detaillierter zu gestalten. So kann beispielsweise eine kurze Nachricht schnell in einen ausführlichen juristischen Text transformiert oder eine grobe Skizze in einen ausgearbeiteten Essay umgewandelt werden. Gerade in sprachintensiven Bereichen, wie etwa in Anwaltskanzleien, können diese Tools besonders effektiv eingesetzt werden, um beispielsweise einen Vertragstext auf einen bestimmten Fall anzupassen.

Was kann KI heute noch ausser Text lesen und erzeugen?

Die Fähigkeiten künstlicher Intelligenzen haben sich inzwischen über die reine Textverarbeitung hinaus erweitert und umfassen nun auch Multimodalität, bei der auch andere Daten wie Audio oder Bilder verstanden und erzeugt werden können. Ein praxisnahes Beispiel hierfür ist ein Online-Händler mit Tausenden von Produkten, bei denen teilweise Produktbeschreibungen fehlen oder knapp gehalten sind. Mit den passenden Instruktionen kann eine KI diese Bilder analysieren und, ergänzt durch die vorhandenen Produktinformationen, eine detailliertere Beschreibung erstellen. Diese Beschreibung kann dabei im gewünschten

Format und Tonfall verfasst werden, was die Attraktivität der Produktpräsentationen erheblich steigert.

Die Erzeugung von Bildern ist ein weiteres Gebiet, auf dem künstliche Intelligenz in Form von Applikationen wie Dall-E oder Midjourney eine bedeutende Rolle spielt. Basierend auf Instruktionen, die das gewünschte Motiv beschreiben, können Bilder generiert werden. So lässt sich beispielsweise einfach ein Titelbild für eine geschriebene Kurzgeschichte erstellen. Insbesondere in Kombination mit ChatGPT ist es möglich, diese Instruktionen automatisch aus einem längeren Text zu extrahieren.

Wo können solche Bilder verwendet werden?

Auch in der Industrie findet diese Technologie Anwendung. Plattformen, die Kochrezepte anbieten, benötigen normalerweise umfangreiche Ressourcen, um die Gerichte zuzubereiten und ansprechende Bilder zu erstellen. Mit Hilfe von KI-generierten Bildern, die auf dem Rezept und einer zusätzlichen Beschreibung der gewünschten Stimmung basieren, können passende Rezeptbilder erzeugt werden. Dies vereinfacht den Prozess erheblich und bietet gleichzeitig die Möglichkeit, kreativ und flexibel auf die Präsentation der Rezepte einzugehen.

Wie kann KI mit der Welt interagieren?

Die meisten aktuellen Anwendungen sind auf eine reine Chat-Oberfläche beschränkt, ohne dass die KI zusätzliche Aktionen durchführen kann. In Zukunft werden wahrscheinlich vermehrt Anwendungen entstehen, bei denen die KI Schnittstellen nutzen kann, um beispielsweise nach einer Anfrage nach guten Restaurantempfehlungen automatisch eine Reservierung vorzunehmen. Gleichzeitig könnte sie eine Reiseplanung erstellen und diese in den eigenen Kalender eintragen.

Umsetzung

Die Basis für all diese Anwendungen bilden KI-Modelle, die auf umfangreichen Textdatensätzen trainiert wurden. Dadurch haben sie gelernt, wie Sprache funktioniert und verfügen über das nötige Wissen, um Fragen zu beantworten. Da dieses erlernte Wissen jedoch auf bis zu 100 Milliarden Parametern verteilt ist und die genauen Prozesse, wie dieses Wissen zustande kommt, weitgehend undurchsichtig sind, ist es nicht möglich, einfach neue Fakten in das Modell zu programmieren. Ein eigenes Modell

mit vergleichbaren Fähigkeiten zu trainieren, würde allein für die erforderliche Rechenleistung Millionen kosten und ist daher nicht praktikabel. In der Praxis gibt es jedoch zwei Ansätze, um ein Modell an spezifische Bedürfnisse anzupassen.

Wie baue ich mein eigenes ChatGPT?

Das erste Verfahren besteht darin, ein bereits trainiertes Modell mit eigenen Daten, wie internen Dokumenten, weiter zu trainieren. Dadurch kann das Modell nicht nur neue Fakten erlernen, sondern sich auch an spezifische Schreibstile anpassen. Allerdings ist dieser Ansatz mit einem gewissen Aufwand verbunden, und in Unternehmen dürfen häufig nicht alle Mitarbeitenden auf alle Informationen zugreifen, weshalb sich ein pragmatischerer Ansatz durchgesetzt hat: Das Laden der benötigten Informationen und Dokumente direkt in den Kontext.

Neben dem Einbinden zusätzlicher Informationen in den Kontext werden auch spezifische Anweisungen häufig in Chats integriert. Dies gilt auch für ChatGPT, wo beispielsweise Anweisungen gegeben werden können, sich nicht rassistisch zu äussern oder bei medizinischen Fragen darauf hinzuweisen, dass ein Facharzt konsultiert werden sollte. Dieses sogenannte System-Prompt kann je nach Verwendungszweck variieren. Ein System-Prompt ist eine Anweisung an die KI, welche den Anwendern verborgen bleibt, aber das Verhalten steuert und beispielsweise ChatGPT erlaubt, auf die Frage, wie sein Name ist, mit «ChatGPT» zu antworten. Ein Chat, der für medizinisches Personal konzipiert ist, sollte möglicherweise andere Hinweise geben und kann sich einer fachspezifischeren Sprache bedienen, um präziser auf die Bedürfnisse der Nutzer einzugehen.

Herausforderungen

Mit dem rasanten Fortschritt der KI ergeben sich auch neue Herausforderungen und Risiken. Dieser Abschnitt diskutiert die Komplexitäten und möglichen Fallstricke, die beim Einsatz von KI-Technologien auftreten können, und bietet Einblicke in die Balance zwischen Innovation und verantwortungsvoller Nutzung.

Warum sind die Ausgaben einer KI nicht immer gleich?

Eine Herausforderung ist die Inkonsistenz der Ausgaben, die selbst bei gleicher Aufgabenstellung auftreten kann. Oft entsprechen die Ergebnisse nicht dem exakt vorgegebenen Format.

Selbst wenn klare Anweisungen gegeben werden, beispielsweise dass die Antwort ausschliesslich aus «Ja» oder «Nein» bestehen soll, können unerwünschte Erklärungen vor oder nach dieser Antwort folgen. Dies stellt bei der automatischen Auswertung eine Herausforderung dar und kann nur durch sehr präzise und gut getestete Instruktionen minimiert werden.

Warum belügt mich die KI?

Das Problem der Unwahrheiten in den Aussagen künstlicher Intelligenz stellt eine noch grössere Herausforderung dar. Dies liegt nicht nur an falschen Informationen in den Trainingsdaten, sondern auch daran, dass KI-Modelle gelegentlich Informationen erfinden. Das KI-Modell ist darauf ausgerichtet, stets eine zufriedenstellende Antwort zu liefern, was es auf den ersten Blick auch tut. Jedoch kann es zu irreführenden oder falschen Aussagen kommen, wie im Beispiel der Empfehlung eines nichtexistenten «Schlosses von Schüpfheim» bei der Frage nach Sehenswürdigkeiten in meinem Heimatort. Dieses Phänomen des Halluzinierens, basierend auf Transfer-Wissen (die Annahme, dass viele Orte ein Schloss haben, also auch der genannte Ort), sollte stets durch eine zusätzliche Verifikation ergänzt werden, bevor man sich auf die Suche nach einer nicht existierenden Sehenswürdigkeit begibt.

Besondere Vorsicht ist geboten, wenn generierte Texte veröffentlicht werden sollen. Da KI-Modelle gelegentlich sogar Quellenangaben erfinden, können solche Texte auf den ersten Blick vertrauenswürdig erscheinen. Eine Methode, um dies zu minimieren, ist, wenn man eigene Daten in den Kontext einlädt und die KI anweist, ausschliesslich diese Fakten zu verwenden, die durch die gegebenen Informationen gestützt sind.

Welche Daten soll ich der KI geben?

Das Thema Datenschutz stellt in der praktischen Anwendung von KI-Systemen wie ChatGPT eine bedeutende Herausforderung dar. Da ChatGPT von OpenAI das führende und leistungsstärkste LLM ist, jedoch nur über das Internet zugänglich und nicht für den Betrieb auf eigenen Servern herunterladbar ist, müssen alle zu verarbeitenden Informationen über das Internet übermittelt werden. Dies kann problematisch sein, insbesondere wenn es sich um sensible Geschäftsinformationen oder schützenswerte Personendaten handelt. Besonders in der Gratis-Version von

ChatGPT kann nicht ausgeschlossen werden, dass übermittelte Informationen zur Qualitätsprüfung weiterverwendet werden.

Daher ist es sowohl für den persönlichen Gebrauch als auch besonders in Unternehmen wichtig, die rechtlichen Voraussetzungen genau zu prüfen. Einige Unternehmen haben aufgrund dieser Datenschutzbedenken die Verwendung von Systemen wie ChatGPT komplett untersagt oder erwägen die Einführung eines selbst betriebenen LLMs. Daher sollten Unternehmen definieren, wie ihr Umgang mit diesen neuen Technologien ist. Dies beinhaltet die sorgfältige Bewertung, welche Daten an externe Dienste gesendet werden dürfen, und gegebenenfalls die Entwicklung interner Richtlinien und Schulungen zur Nutzung solcher Systeme.

Warum blockiert die KI manchmal bestimmte Anfragen?

Eine weitere Herausforderung in der Anwendung von KI-Systemen sind die sogenannten Guardrails, die darauf abzielen, Missbrauch zu verhindern, aber in bestimmten Fällen die Nutzung einschränken können. Diese Leitplanken dienen dazu, unangemessene Inhalte wie rassistische Äußerungen zu verhindern. Wenn man beispielsweise eine KI nutzen möchte, um Kommentare auf einer Website automatisch auf Rassismus zu überprüfen, könnte die KI sich weigern, darauf zu antworten, da die Eingabe selbst möglicherweise rassistische Passagen enthält. Diese Einschränkungen sind aus ethischer Sicht wichtig und notwendig, können jedoch die Effektivität der KI in bestimmten Anwendungsfällen mindern.

Werden wir das selbstständige Denken verlernen?

Die weitreichende Implementierung von KI-Systemen bringt auch tiefgreifende Veränderungen für Individuen und die Gesellschaft mit sich, die über technische Aspekte hinausgehen. Ich gehe hierbei nicht auf die dystopischen Szenarien ein, in denen eine künstliche Intelligenz die Menschheit unterwirft, sondern auf konkrete Veränderungen für Individuen und Gesellschaft, welche man beim breiten Einsatz beachten muss.

Das Aufkommen des Internets, insbesondere Dienste wie Google und Wikipedia, haben unsere Art und Weise, mit Informationen umzugehen, fundamental verändert. Wo früher ein Besuch in der Bibliothek notwendig war, um Informationen für einen Vortrag zu sammeln, können Schülerinnen und Schüler heute diese Re-

cherche bequem während einer Busfahrt auf ihrem Smartphone durchführen. Aber auch Diskussionen, in denen man einst lebhaft nach der Wahrheit rang, können jetzt schnell durch eine Online-Suche beendet werden. Obwohl dies den Zugang zu Wissen erheblich erleichtert, hat es auch Auswirkungen auf das selbstständige Denken.

Mit der zunehmenden Verbreitung von KI, insbesondere von Systemen wie ChatGPT, könnte sich diese Dynamik weiter verstärken. KI-Systeme tendieren dazu, ausgewogene, gemässigte Antworten zu liefern, besonders bei Themen, die Meinungen und Interpretationen beinhalten. Dies kann zwar in vielen Fällen hilfreich sein, birgt aber auch das Risiko, dass Nutzer sich weniger mit Extrempositionen oder kontroversen Perspektiven auseinandersetzen. Eine solche Auseinandersetzung ist jedoch wichtig für die Entwicklung kritischen Denkens und für das Verständnis komplexer Themen.

Darüber hinaus besteht die Gefahr, dass Nutzer sich auf die von der KI gelieferten Antworten verlassen, ohne eigene Schlussfolgerungen zu ziehen oder zusätzliche Recherchen anzustellen. Dies könnte langfristig die Fähigkeit zur kritischen Bewertung von Informationen und zur eigenständigen Problemlösung beeinträchtigen.

Wie reproduzieren KIs gesellschaftliche Stereotypen?

Weitere Bedenken gibt es hinsichtlich des historischen Bias, also Voreingenommenheiten in den Trainingsdaten von KI-Systemen. Die Trainingsdaten, die aus bestehenden, teilweise älteren Texten bestehen, können überholte Rollenbilder und gesellschaftliche Stereotypen widerspiegeln. Diese werden dann von der KI reproduziert und weitergetragen.

Insbesondere in der englischen Sprache wird dieser Bias deutlich sichtbar, wenn man Bilder generieren lässt für Berufe wie «Firefighter» (Feuerwehrmann/-frau) oder «Nurse» (Krankenpfleger/-in). Häufig generiert die KI Bilder, die stereotype Geschlechterrollen diesen Berufen zuordnen. Dies ist ein Problem, da es veraltete Vorstellungen und Stereotype verstärken kann, anstatt zu ihrer Auflösung beizutragen. Um diesem Problem entgegenzuwirken, ist es wichtig, dass KI-Entwickler und -Nutzer sich der inhärenten Voreingenommenheiten in den Trainingsdaten bewusst sind und aktiv Massnahmen ergreifen, um sie zu minimieren. Letzt-

lich ist es entscheidend, dass wir als Gesellschaft ein Bewusstsein für diese Probleme entwickeln und einen verantwortungsbewussten Umgang mit KI-Technologien pflegen.

Wer hat das Recht an den Daten?

Ein weiterer Kritikpunkt, welcher in Zukunft wohl noch zu einigen Gerichtsfällen führen dürfte, ist die Frage des Urheberrechts. Für das Trainieren von KI-Modellen wie ChatGPT und bildgenerierenden Systemen werden massive Mengen an Daten aus dem Internet verwendet, oft ohne explizite Einwilligung der ursprünglichen Urheberinnen oder Urheber. Obwohl diese Daten nicht direkt abrufbar sind, adaptiert die KI doch Informationen und Stile aus diesen Quellen oder gibt sogar exakte Zitate wieder.

Ein besonderes Thema ist die Fähigkeit von KI-Systemen, Werke im Stil bekannter Künstler zu erstellen. Dies wirft wichtige urheberrechtliche Fragen auf, da die KI stilistische Elemente dieser Künstler nachahmen kann, ohne dass dies eine direkte Kopie eines spezifischen Werks darstellt. Die rechtlichen Rahmenbedingungen in Bezug auf solche Nachahmungen und die Verwendung urheberrechtlich geschützten Materials für das Training von KI-Modellen sind derzeit noch nicht vollständig geklärt, jedoch gibt es immer mehr Guardrails, um solche Antworten der KI zu verhindern.

Wo werden Lösungen zu den rechtlichen Herausforderungen gesucht?

Die rechtlichen Aspekte rund um künstliche Intelligenz werden in den kommenden Jahren ein zentrales Thema sein, um als Gesellschaft zu definieren, welche Arten von KI wir akzeptieren und wie deren Sicherheit gewährleistet werden kann. Die Europäische Union ist mit dem EU AI Act dabei, eine wegweisende rechtliche Grundlage zu schaffen, die weltweit Auswirkungen haben wird, einschliesslich auf die Schweiz.

Der EU AI Act zielt darauf ab, bestimmte Anwendungen von KI zu regulieren oder sogar zu verbieten. So wird beispielsweise die Verwendung von Gesichtserkennungstechnologien im öffentlichen Raum fast komplett untersagt. Andere Anwendungen, wie das automatische Beurteilen von Bewerbungsschreiben, werden einer erhöhten Sorgfalt und Transparenz unterliegen. Dies bedeutet, dass Unternehmen, die solche Technologien einsetzen, klare Richtlinien befolgen und möglicherweise ihre Prozesse offenlegen

müssen, um Diskriminierung zu vermeiden und Fairness zu gewährleisten. Darüber hinaus werden auch KI-Systeme wie ChatGPT vermehrten Transparenz- und Sorgfaltspflichten unterliegen.

Anwendungen im Bildungswesen

Einige Leute nennen es künstliche Intelligenz, aber in Wirklichkeit wird uns diese Technologie verbessern. Ich denke also, dass wir statt künstlicher Intelligenz unsere Intelligenz erweitern werden.

Ginni Rometty, CEO und Präsidentin von IBM

Wie die Umsetzung im Detail aussehen wird, muss sich noch zeigen. Fakt ist jedoch, dass bereits heute sehr grosse Veränderungen im Gange sind und die meisten Branchen tangieren. Insbesondere die Bildung wird davon betroffen sein. Dies nicht nur dadurch, dass neue Fähigkeiten im Umgang mit Technologien vermittelt werden müssen und ein Verständnis geschaffen werden muss, wie diese funktionieren, sondern auch, indem diese direkt im Bildungswesen angewendet werden. Ein grosser Vorteil dabei ist, dass eine künstliche Intelligenz den Lernfortschritt von einzelnen Schülerinnen und Schülern präziser begleiten kann als eine einzelne Lehrperson, die verantwortlich für viele Lernende gleichzeitig ist.

Welche Anwendungen sind denkbar?

In einem Projekt für einen Lehrmittelverlag haben wir Ansätze geprüft, wie Matheaufgaben gestellt werden können und die Schülerinnen und Schüler individuell durch diese geleitet werden, um selber die Lösung zu finden. Dabei kann auch mitgegeben werden, dass sich die künstliche Intelligenz empathisch und motivierend verhält und falls notwendig Tipps und alternative Erklärungen liefert, um das Verständnis für das Problem optimal zu fördern. Für die Lehrperson kann aus diesen Interaktionen dann ein Bericht erstellt werden, um den einzelnen Lernfortschritt zu erfahren. Die oben genannten Herausforderungen wie das Halluzinieren der KI können hierbei limitiert werden, indem Schnittstellen zu Mathematik-Programmen angebunden werden, um die Lösungen zu überprüfen.

Wie sieht es konkret im Sprachunterricht aus?

Da diese künstlichen Intelligenzen mit Sprache arbeiten, ist der Einsatz im Deutsch- oder Fremdsprachen-Unterricht als Lehr-

mittel, aber auch zur Korrektur von Arbeiten naheliegend. Gerade für die Korrektur ist es jedoch wichtig zu beachten, dass sich auch beim selben Text eine Bewertung unterscheiden kann und genaue Anweisungen und Beispiele notwendig sind, um eine Konsistenz zu erhalten. Ein Vorteil neben der eingesparten Zeit könnte auch darin liegen, viel detaillierteres Feedback geben zu können, oder einen Bias, welchen Lehrpersonen haben können, zu limitieren.

Aber auch bereits in der Vorbereitung kann ein solches Tool von Lehrpersonen ohne grössere Risiken angewendet werden. So kann eine Idee für einen Aufsatz oder einen ansprechenden Text für ein Diktat generiert werden, welcher genauer auf die Interessen der Schülerinnen und Schüler abgestimmt ist. Auch Fragen für ein Quiz oder eine Prüfung können mit den richtigen Instruktionen schnell erstellt werden.

Wo wird dies hinführen?

Die Integration künstlicher Intelligenz in den Bildungsbereich bietet sowohl für Lehrpersonen als auch für Schülerinnen und Schüler erhebliche Erleichterungen. Dabei ist es wichtig, den Umgang mit Aufgaben und Prüfungen entsprechend anzupassen, insbesondere weil es mit präzisen Anweisungen fast unmöglich wird, Texte, die von künstlicher Intelligenz generiert wurden, zu erkennen. Dies führt zur Überlegung, inwiefern der Einsatz künstlicher Intelligenz im Bildungskontext problematisch oder gar hinderlich sein kann. Eine noch wichtigere Frage ist jedoch, welche Fähigkeiten in Zukunft entscheidend sein werden und wie sich der Bildungsbedarf dementsprechend entwickeln muss, um mit den technologischen Fortschritten Schritt zu halten.

7 Take-Aways, um das Beste aus ChatGPT rauszukitzeln

- 1. Bereitstellung ausreichender Informationen:** ChatGPT hat in jeder neuen Sitzung kein Vorwissen über dich oder deine Situation. Um präzise Antworten zu erhalten, solltest du umfassende Informationen bereitstellen: Was hast du bereits unternommen? In welchem Berufsfeld bist du tätig? Wer ist der potenzielle Empfänger des Textes? Was möchtest du erreichen? Für wiederkehrende Anfragen kann es nützlich sein, relevante Informationen zu speichern und erneut zu verwenden.
- 2- Nachfrage nach zusätzlichen Informationen durch ChatGPT:** Wenn du zu wenige Informationen bereitstellst, wird ChatGPT dennoch eine Antwort generieren. Fordere die KI auf, erst dann zu antworten, wenn alle notwendigen Informationen erfragt wurden. Dies kann durch eine Anweisung wie «Frage alles Notwendige, bevor du antwortest» im Chat umgesetzt werden.
- 3. Definition des gewünschten Formats:** ChatGPT ist in der Lage, nicht nur Text, sondern auch Tabellen, Überschriften und Formatierungen zu erstellen. Gib das gewünschte Format an, wie beispielsweise «eine Tabelle mit hervorgehobenen Schlüsselwörtern, unterteilt in Kapitel».
- 4. Bereitstellung von Beispielen:** Wenn du bereits Beispiele für das gewünschte Ergebnis hast, teile diese mit, um Format, Sprache und Umfang ähnlich zu gestalten. Wenn du beispielsweise Texte nach verschiedenen Kriterien bewertet haben möchtest, können diese Bewertungen als Beispiele in den Chat eingegeben werden.
- 5. Schritt-für-Schritt-Ansatz:** Anstatt die gesamte Aufgabe auf einmal zu lösen, kann die KI angewiesen werden, dies iterativ zu tun. Frage zunächst nach den notwendigen Schritten zur Lösung der Aufgabe und arbeite dann jeden Schritt nacheinander ab. Bei jedem Schritt kannst du weitere Verbesserungen vorschlagen.
- 6. Selbstkorrektur durch die KI:** Da aktuelle KI-Systeme Texte von oben nach unten erstellen, kann es hilfreich sein, die KI zur Selbstkorrektur anzuregen. Nach einer Antwort kannst du fragen, ob die KI den Text kritisch beurteilen und anschliessend die Kritikpunkte verbessern kann.
- 7. Erweiterung des Kontexts:** Integriere nicht nur die direkte Frage, sondern ermögliche der KI, im Internet nach Informationen zu suchen oder füge Texte hinzu, die bei der Beantwortung der Anweisung hilfreich sein können. Die Anweisung kann auch in mehrere Teile gegliedert werden, wie «Kontext: Ich bin..., Frage: Was ist...?, Inhalt: Längerer Text».