

K/A/I-UPDATE: Aktuelles zu generativer KI kompakt

Kai Spriestersbach



Wir befinden uns an einem Wendepunkt, an dem die Unterscheidung zwischen Mensch und Maschine sowie zwischen Wahrheit und Fiktion zunehmend schwieriger wird. Die generative KI steht im Zentrum dieser Entwicklung und hat das Potenzial, die Landschaft der Informationsverbreitung zu verändern. Eine kritische Frage ist, ob sie dabei eine hilfreiche Ressource oder ein Risiko für Fehlinformationen darstellt.

Ironischerweise könnte genau die Technologie, die ursprünglich zur Verbesserung unserer Informationsbeschaffung und -verarbeitung entwickelt wurde, nun eine Quelle für Desinformation sein. Ein aktuelles Beispiel sind Chatbot-Halluzinationen, die laut WIRED (www.wired.com/story/fast-forward-chatbot-hallucinations-are-poisoning-web-search/) das Internet beeinträchtigen. Darin wird berichtet, wie eine fiktive Forschungsarbeit von Claude Shannon, die Bing bei einem Nutzer herbeihalluziniert hat, bei einem anderen Nutzer von derselben Suchmaschine bereits als Fakt mit Quellenangabe präsentiert wurde, denn der erste Nutzer hatte für einen Blogartikel recherchiert und die Falschinformation in seinem Artikel veröffentlicht. Das Beispiel zeigt, wie schnell sich Fiktionen als vermeintliche „Fakten“ etablieren.

Dennoch ist es zu einfach, die Situation nur negativ zu sehen. Generative KI ist nicht per se schädlich für das Internet. Vielmehr hängt viel davon ab, wie wir diese Technologie nutzen und regulieren. Durch das Filtern von Spam aus den Trainingsdaten und das Hinzufügen mensch-

licher Überprüfung könnten wir die Qualität KI-erzeugter Inhalte wesentlich verbessern. Jüngste Studien deuten sogar darauf hin, dass durch wiederholte Verarbeitung derselben Trainingsdaten die Leistung einer KI deutlich gesteigert werden kann. Dies deutet darauf hin, dass wir uns in einer Phase des Lernens und der Anpassung befinden.

Die Herausforderung besteht darin, ein Gleichgewicht zu finden. Wir müssen die Risiken der generativen KI ernst nehmen und Maßnahmen ergreifen, um die Integrität der Informationen zu schützen. Gleichzeitig sollten wir das Potenzial dieser Technologie nicht überstürzt verwerfen. Generative KI könnte, richtig eingesetzt, ein wertvolles Werkzeug sein, um unser Wissen zu erweitern.

Ich lade Sie daher ein, mit mir gemeinsam tiefer in die Welt der generativen KI einzutauchen. Lassen Sie uns die damit verbundenen Geheimnisse und Herausforderungen erkunden und überlegen, wie wir diese Technologie nutzen können, ohne ihre Risiken zu ignorieren.

Foto: leonardo.ai

DER AUTOR



Kai Spriestersbach ist erfolgreicher Unternehmer, Tech-Journalist, SEO-Veteran und KI-Experte mit einem Masterabschluss in Webwissenschaften. Sein Sachbuch „Richtig Texten mit KI: ChatGPT, GPT-4, GPT-3 & Co.“ erschien im April 2023 im mvv Verlag.



Foto: leonardo.ai

Vergiftet generative KI das Netz?

Das Internet dient als zentrale Informationsquelle, doch die Glaubwürdigkeit dieser Informationen steht auf dem Spiel, insbesondere durch den Einsatz generativer KI. Ein Bericht von WIRED über „Chatbot-Halluzinationen, die die Websuche beeinträchtigen“ hebt dieses Problem hervor. So hat Microsofts Suchmaschine Bing falsche Informationen von Chatbots als Fakten dargestellt. Generative KI, die neue Inhalte erstellen kann, die nicht in ihren Trainingsdaten vorhanden sind, stellt somit eine Herausforderung für die Zuverlässigkeit der Websuche dar.

Normalerweise liefert eine Websuche eine Liste relevanter Ergebnisse basierend auf komplexen Technologien im Hintergrund. Doch generative KI könnte diese Prozesse stören, indem sie Algorithmen irreführt, die ursprünglich für ein von Menschen dominiertes Web konzipiert waren.

Ein Beispiel hierfür ist die fälschliche Information durch Bing über eine nicht existente Forschungsarbeit von Claude Shannon. Die Suchmaschine behauptete irrtümlich, Shannon hätte

eine Arbeit mit dem Titel „A Short History of Searching“ verfasst. Tatsächlich wurden diese falschen Informationen von zwei Chatbots generiert und von einem Nutzer online gestellt, woraufhin Bing diese als Tatsachen präsentierte.

Ein Nutzerexperiment offenbarte die Unwahrheiten: Bei einer erneuten Anfrage an Bing wurden die ursprünglichen falschen Antworten der Chatbots als gesicherte Fakten dargestellt. Dies zeigt, wie anfällig Suchmaschinen für Täuschungen durch generative KI sein können.

Experten weisen darauf hin, dass es – nicht nur für Suchmaschinen – schwierig bis unmöglich sein kann, KI-generierte Texte automatisch zu identifizieren. In der letzten Ausgabe #83 habe ich hierzu bereits einen ausführlichen Artikel geschrieben.

Mit der zunehmenden Erstellung von SEO-Seiten, Social-Media-Beiträgen und Blogposts mittels KI könnte sich dieses Problem verschärfen. Die aktuelle Entwicklung erinnert an einen algorithmischen Uroboros, bei dem sich die KI selbst auffrisst. Dies unterstreicht, die Bedeutung, die Risiken und die Auswirkungen von generativer KI auf die Websuche zu verstehen und

Schritte zu unternehmen, um die Integrität unserer täglich genutzten Informationen zu schützen.

Qualität als entscheidender Faktor in Sachen Trainingsdaten

In der Bewertung der KI-Entwicklung sollten wir uns darauf konzentrieren, dass Qualität wichtiger ist als Quantität. Die Ära der Large Language Models (LLMs) zeichnet sich nicht nur durch ihre Größe, sondern auch durch die Qualität der Trainingsdaten aus.

Die Vorbereitung großer Sprachmodelle, wie durch Microsofts Phi-1 und das dazugehörige Paper „Textbooks Are All You Need“ (arxiv.org/pdf/2306.11644.pdf) illustriert, ist ein entscheidender Prozess. Neue Basismodelle werden normalerweise auf umfangreichen Textdaten wie Artikeln, Büchern und Websites trainiert, um ein grundlegendes Verständnis von Grammatik, Syntax und Semantik zu entwickeln.

„Doch der wahre Durchbruch liegt in der Qualität der Daten: Sind diese nicht nützlich für einen menschlichen Lernenden, sind sie es auch nicht für ein LLM.“

Dieser Ansatz führte zu beeindruckenden Ergebnissen. Phi-1, das auf „Lehrbuchqualität“-Daten trainiert wurde, zeigte erhebliche Verbesserungen in der Python-Code-Erzeugung. Dies beweist, dass qualitativ hochwertige Trainingsdaten die Leistung von LLMs steigern können, selbst bei reduzierter Datengröße und Rechenaufwand.

Forschungen zeigen zudem, dass Modelle, die mit ausschließlich durch LLM generierten Daten trainiert wurden, besser abschneiden als solche, die auf herkömmlichen Daten basieren. Dies verdeutlicht, dass nicht die Menge, sondern die Qualität der Daten



Foto: leonardo.ai

entscheidend ist und dass der Output von Sprachmodellen qualitativ hochwertig sein kann.

Das Feintuning mit speziell erstellten, künstlichen Trainingsdaten ermöglicht den Modellen, komplexe algorithmische Aufgaben zu bewältigen und externe Bibliotheken zu nutzen. Dies deutet darauf hin, dass der Verfeinerungsprozess dem Modell hilft, das im Pre-Training erworbene Wissen zu organisieren und zu festigen.

Die Qualität der Trainingsdaten macht also den Unterschied zwischen einem guten und einem hervorragenden Sprachmodell aus.

„Es ist die Qualität der Daten, die das Potenzial und die Fähigkeiten des Modells bestimmt.“

Ein weiteres Paper (arxiv.org/pdf/2305.16635.pdf) zeigt Möglichkeiten auf, den ökologischen Fußabdruck von LLMs zu verringern: Kleinere Modelle, die weniger Training benötigen, können einen erheblichen Unterschied machen.

Dies eröffnet die Chance, gleichzeitig die Leistung zu verbessern und die Umweltbelastung zu reduzieren.

„Wir stehen also vor einer doppelten Gewinnchance: Verbesserung der Leistung bei gleichzeitiger Reduzierung der Umweltbelastung.“

Die Chancen für kleinere Akteure in der KI-Welt

Im Bereich der KI dominieren oft Großunternehmen mit umfangreichen finanziellen Mitteln die Szene. OpenAI und Anthropic beispielsweise, unterstützt durch Milliarden an Venturecapital, beherrschen regelmäßig die Nachrichten. Doch die zunehmende Bedeutung von Qualität über Quantität eröffnet jetzt auch kleineren Unternehmen und Forschungsgruppen neue Möglichkeiten, auf dem globalen Markt konkurrenzfähig zu sein.

In Ländern wie Deutschland, wo Ressourcen möglicherweise nicht so reichlich wie im Silicon Valley verfügbar sind, besteht nun eine realistischere Chance, in der Entwicklung fortschrittlicher KI-Modelle mitzuhalten. Entscheidend ist nicht mehr allein die Menge an Kapital, sondern der intelligente Einsatz von Ressourcen.

Der Schlüssel liegt darin, nicht die größte Menge an Daten zu sammeln, sondern die richtigen, qualitativ hochwertigen Daten. Kleine, fokussierte Datensätze können genauso effektiv sein wie umfangreiche, aber weniger spezialisierte Datenkollektionen. Dies ermöglicht es auch Akteuren mit begrenzten Mitteln, effiziente und leis-

tungsstarke Modelle zu entwickeln.

In Deutschland, wo Präzision und Qualität fest verwurzelt sind, gibt es bereits zahlreiche innovative Unternehmen und Forschungseinrichtungen, die diese Gelegenheit nutzen. Sie fokussieren sich auf die Entwicklung von Nischenmodellen für spezifische Problemstellungen oder auf maßgeschneiderte Modelle für bestimmte Industrien oder Märkte.

Diese Entwicklung deutet auf eine Demokratisierung der KI-Technologie hin und zeigt, dass wir in einer Ära sind, in der Diversität und Spezialisierung ebenso wichtig sind wie Größe und universelle Anwendbarkeit. Die neuen Vorbilder in der KI-Ära könnten aus kleineren Laboren und Start-ups stammen, die mit Leidenschaft, Fachwissen und einem Fokus auf Qualität arbeiten.

Die Rolle von Open-Source-Modellen in der KI-Entwicklung

Ein weiterer wichtiger Aspekt in der Diskussion über KI sind Open-Source-Basismodelle. In den letzten Jahren wurden zahlreiche Large Language Models als Open Source veröffentlicht. Einige dieser Modelle sind auch für die kommerzielle Nutzung lizenziert (wie Apache 2.0, MIT, OpenRAIL-M) und stehen der Allgemeinheit zur Verfügung. Beispiele hierfür sind T5, UL2, Cerebras-GPT, Open Assistant (Pythia-Familie), GPT-J, GPT-NeoX, Bloom, StableLM-Alpha, Falcon, MPT, LLaMA 2, OpenLM, Mistral 7B, Mixtral 8x7B und OpenHermes. Diese Modelle bieten für Unternehmen und Forschungsgruppen, die nicht über große Budgets verfügen, vollkommen neue Möglichkeiten. So reichen sie von Millionen bis zu Milliarden Parametern und decken damit eine

Vielzahl von Anwendungen ab.

Eine Übersicht findet sich auf github.com/eugeneyan/open-llms.

Die Verfügbarkeit dieser Open-Source-Modelle stellt eine signifikante Veränderung dar. Früher war das Pre-Training von LLMs aufgrund der hohen Kosten und des Aufwands nur großen Akteuren möglich. Heute können Unternehmen und Forschungsgruppen diese Modelle als Ausgangspunkt nutzen und für ihre speziellen Bedürfnisse anpassen und verfeinern. Dies spart Kosten und Zeit und ermöglicht eine größere Diversität und Spezialisierung in der KI-Entwicklung.

Für deutsche Unternehmen und Forschungseinrichtungen eröffnet sich dadurch eine große Chance. Sie können auf diesen Basismodellen aufbauen und ihre Ressourcen für die Entwicklung maßgeschneiderter Lösungen einsetzen, die speziell auf ihre Anforderungen zugeschnitten sind. Dies kann besonders in Bereichen mit spezifischem Fachwissen und Spezialkenntnissen einen Wettbewerbsvorteil darstellen.

Es ist eine spannende Zeit für die KI-Entwicklung, nicht nur in Deutschland, sondern weltweit. Es findet eine Verschiebung statt, weg von großen Konzernen hin zu einem breiteren Ökosystem, in dem Qualität, Innovation und Präzision im Vordergrund stehen. Diese Entwicklung verspricht technologische Fortschritte und eine inklusivere, vielfältigere KI-Zukunft.

OpenAI eröffnet GPT Store und ChatGPT Team Plan

OpenAI hat kürzlich den offiziellen Start des GPT Stores bekannt gegeben, der es Nutzern ermöglicht, eine Vielzahl von GPTs zu entdecken, und Entwicklern erlaubt, mit ihren erstellten



Foto: leonardo.ai

KI-Assistenten Gewinne zu erzielen. Der GPT Store bietet bereits zum Start über drei Millionen von der Community erstellte GPTs in verschiedenen Kategorien sowie eine Trend-Bestenliste und handverlesene Empfehlungen. Ein Umsatzbeteiligungsprogramm für US-Entwickler soll im ersten Quartal starten und die Creator basierend auf der Nutzerinteraktion mit ihren GPTs belohnen.

Zusätzlich kündigte OpenAI einen neuen ChatGPT Team Plan an. Dieser kostet 25 US-Dollar pro Monat und Nutzer (bei jährlicher Zahlung) oder 30 US-Dollar pro Monat (bei monatlicher Zahlung) und bietet einen gemeinsamen Arbeitsbereich sowie Werkzeuge zur Nutzerverwaltung. Weiterhin wurden weitere Personalisierungsoptionen und ein Langzeitgedächtnis für ChatGPT eingeführt, wodurch GPT zwischen Chats Erinnerungen behalten und sich im Laufe der Zeit verbessern kann.

Rabbit präsentiert R1: das AI-betriebene Mobilgerät ohne Apps

Die Zukunft der mobilen Technologie scheint sich von Apps zu entfernen und sich der KI-gesteuerten Hardware zuzuwenden. Nach dem ersten „AI Pin“ genannten Gerät zweier ehemaliger Apple-Mitarbeiter wurde mit dem R1-Gerät von Rabbit ein weiteres Produkt dieser vollkommen neuen Geräteklasse vorgestellt. Ähnlich zum AI Pin von hu.ma.ne wird das R1 als eigenständiges Mobilgerät von einem „Large Action Model“ angetrieben. Das Gerät verfügt über einen Touchscreen, eine Sprechaste, ein analoges Scrollrad, Mikrofone, Lautsprecher und eine drehbare Kamera. Es ist außerdem mit Bluetooth, WLAN und einem SIM-Karten-Slot ausgestattet.

Durch Drücken der einzigen Taste auf dem Gerät kann man nach Themen wie Philosophie, Aktienkursen und Filmen fragen, worauf es mit Sprache, Text und Bildern auf dem Bildschirm

antwortet. Das R1 integriert sich auch in Apps wie Spotify und Uber, sodass Benutzer beispielsweise ein Lied abspielen oder ein Taxi bestellen können. Außerdem kann es Aufgaben erledigen, für die normalerweise mehrere Apps erforderlich wären, wie das Planen und Buchen einer Reise.

Die Kamera des R1 ist selbstverständlich mit Computer-Vision ausgestattet, was bedeutet, dass sie zum Beispiel ein Rezept basierend auf dem, was sie im Kühlschrank sieht, generieren kann. Sie kann auch Dokumente bearbeiten, wenn man die Kamera des Geräts auf einen Laptop richtet. Besonders interessant ist die Entwicklung der Funktion „Teach Rabbit“, die es Benutzern ermöglicht, dem Gerät beizubringen, bestimmte Aufgaben in ihrem Namen zu erledigen.

Innovative AI-Entwicklung auf Apple-Hardware

Apple ML Research hat kürzlich MLX vorgestellt, ein Open-Source-Framework, das die Entwicklung und Implementierung von AI-Modellen auf Apple-Geräten vereinfacht und optimiert. MLX ist speziell auf die neueste Apple-Silizium-Technologie zugeschnitten und verbessert die AI-Performance über das gesamte Apple-Produktspektrum. Es verwendet vertraute Programmierwerkzeuge wie Python, was die Modellentwicklung erleichtert und gleichzeitig eine effizientere und schnellere Modellschulung ermöglicht.

Eine Besonderheit von MLX ist die Möglichkeit, auf Apples M2- und M3-Produkten ein „Mini“-Datacenter für Large Language Models (LLMs) zu erstellen. Dies ermöglicht es, fortschrittliche und rechenintensive Modelle direkt auf lokalen Geräten auszuführen. Es gibt außerdem Gerüchte, dass Apple an einer neuen Siri-Version mit erweiterten AI-Funktionen arbei-



Foto: leonardo.ai

tet, die Informationen über Gespräche geräteübergreifend speichern kann.

MLX könnte für Apple einen bedeutenden Schritt in Richtung AI auf Geräten bedeuten und das Unternehmen im AI-Wettlauf wieder ins Rennen bringen. Die benutzerfreundliche und für Apple-Silizium optimierte Gestaltung von MLX unterstreicht Apples Ambitionen im AI-Segment für 2024.

Perplexity: eine neue Ära der Online-Suche?

In einer spannenden Entwicklung im Bereich der künstlichen Intelligenz hat Perplexity AI, eine auf AI basierende Suchmaschine, beeindruckende 73,6 Millionen US-Dollar in einer Finanzierungsrunde gesammelt, was das Unternehmen auf eine Bewertung von 520 Millionen US-Dollar bringt. Die Investorenliste liest sich wie das Who's who der Tech-Industrie: Jeff Bezos, Nvidia, Shopify-CEO Tobi Lutke, der ehemalige GitHub-CEO Nat Friedman sowie weitere Investoren und Risikokapitalfonds.

Perplexity AI setzt sich zum Ziel, eine ernst zu nehmende Herausforderung für Google im Markt der Online-

Suchmaschinen zu sein. Aber was macht Perplexity so besonders?

Hier sind einige Schlüsselmerkmale:

1. Direkte Suchantworten: Anstelle einer Liste von Links bietet Perplexity direkte Antworten auf Suchanfragen.
2. Konversationelle Suche: Ein gesprächsbasierendes Suchmodell ermöglicht eine natürlichere und interaktivere Sucherfahrung.
3. Werbefrei: In einer Welt, in der die Anzahl der Werbeanzeigen in Suchmaschinen zunimmt, strebt Perplexity ein besseres, werbefreies Nutzererlebnis an.

Unsere eigenen Tests mit Perplexity haben beeindruckende Ergebnisse geliefert. Die Suchmaschine bietet präzise und einfache Antworten auf die meisten Fragen. Diese neue Finanzierungsrunde positioniert Perplexity nun als ernsthaften Konkurrenten in einem Markt, der lange von Google dominiert wurde. Das macht diese Entwicklung zu einer faszinierend zu beobachtenden – nicht nur für Technikenthusiasten, sondern für jeden, der im digitalen Raum aktiv ist.