

Vortrag „KI fürchten oder beherrschen?“

Vom 23.09.24, Thomas Matzner

Zusammenfassung der Fragerunde

Frage: Ich würde gerne etwas zur Intuition sagen, da bin ich hängengeblieben: Ich finde, das ist ein Aspekt, den man gründlich beleuchten müssten. Denn erstens ist die Begrifflichkeit recht unscharf, obwohl die meisten sicherlich auf den ersten Blick zu wissen glauben, wovon sie reden. Intuition ist ja im Menschen verwurzelt – wir haben über Generationen ein Gefühl dafür erworben, wie wir etwas einschätzen müssen; und Intuition ist hochkomplex. Ich glaube nicht, dass man sie einfach mit irgendwelchen Regelwerken wegwischen kann (ich habe allerdings Ihre Literatur nicht gelesen) – aber das ist, meine ich, auf keinen Fall die Himmelsrichtung, in die wir Menschen gehen sollten – unter anderem, weil ChatGPT ja nicht semantisch verstehen, sondern nur Wahrscheinlichkeiten auswerfen kann.

Matzner: Was ChatGPT angeht, habe ich jetzt gerade als aktuellen Fall ein Zitat auf LinkedIn gesehen. ChatGPT wurde gefragt: „Was hättest du 2020 getan, wenn du die Weltherrschaft gehabt hättest, um die Corona Krise zu beherrschen?“. ChatGPT hat eine ganze Reihe an vernünftigen Antworten darauf gegeben – und ich habe nur angemerkt, dass man daraus bitte nicht schließen soll, ChatGPT sei vernünftiger als wir, weil ChatGPT ja auf der Basis des heutigen Wissens diese Antwort gegeben hat – und heute wissen wir natürlich mehr als alle entscheidenden Personen Anfang 2020.

Eigentlich spielt Intuition auch bei einem Chatbot eine Rolle, denn Sie können natürlich ChatGPT danach fragen, welche Schlüsse die SPD wohl aus den Wahlergebnissen in Brandenburg ziehen wird. Das kann man intuitiv entscheiden.

Ich nehme meine Literatur im Übrigen sehr ernst; und zwar deshalb, weil es immer wieder EthikerInnen gibt, die ihr Wissensgebiet überschreiten und zu PsychologInnen werden. Und zwar nicht mehr nur die Fragen „Was soll ich tun? Wie sollen sich Menschen verhalten, was sind gute Verhaltensweisen?“ beantworten, sondern auch festsetzen, *wie* sich Menschen verhalten. Da bin ich sehr vorsichtig – ich halte mich raus und lese entsprechende Literatur. Zwei meiner Vorbilder sind hier Daniel Kahnemann und Norbert Bischof.

Diese Ergebnisse, die Kahnemann im entsprechenden Kapitel referiert, sind schon sehr zwingend – schauen Sie sich die Ergebnisse gerne direkt an, wenn Sie jetzt noch zweifeln.

Sie haben auch gesagt, dass es unklar ist, was Intuition überhaupt ist, und auch das hat Kahnemann behandelt. „Intuition“ nach der Definition der Psychologie ist das, was ich „Mustererkennung“ genannt habe. Dabei lassen wir eine aktuelle Situation gegen unsere

Erfahrungen und Erinnerungen laufen, und die Situation, die ihr am ähnlichsten ist, setzen wir als Handlungsmuster für die aktuelle Situation ein. Das war jetzt sehr laienhaft ausgedrückt, genauer sollte man das bei Kahnemann nachlesen.

Frage: Mich hat die Formulierung „...gegen die Wirklichkeit testen“ getriggert; ich sehe da eine Verbindung zu den generativen Möglichkeiten von KI (die heute sicherlich wegen ChatGPT im Hintergrund stehen). Ich selbst komme aus dem Gebiet Visual Computing/Grafik und es ist ja beeindruckend, was heute möglich ist. Im Prinzip kann ich jede Art von Wirklichkeit generieren, wenn ich die Wirklichkeit durch visuelle oder akustische Sensorik definiere.

Meine Frage ist letztendlich: Was ist es denn für eine Wirklichkeit, gegen die wir testen können?

Matzner: Ich verwende den Begriff „Wirklichkeit“ und übrigens bewusst nicht den Begriff „Realität“ – denn „Wirklichkeit“ ist etwas weniger anspruchsvoll als „Realität“. „Realität“ ist eine wissenschaftlich-empirisch als wahr erklärte Wirklichkeit. Die Wirklichkeit kann z.B. die subjektive Wahrnehmung von Temperatur – warm, kalt – sein, während die Realität die gemessene Temperatur in Grad beschreibt. Die Wirklichkeit sagt etwas darüber aus, wie die Realität auf ein Individuum wirkt.

Das muss man beim Einsatzzweck und der moralischen Bewertung überlegen. Für diejenigen, die ethisch bewandert sind: Ich gehe sehr weitgehend konsequentialistisch vor – das heißt, ich stelle zuerst Handlungsfolgen zusammen und versuche dann, diese zu bewerten; die von einer Person wahrgenommene Handlungsfolge ist ein Einfluss auf deren Wirklichkeit.

Beantwortet das Ihre Frage?

Nachfrage: Zum Teil. Wie sehen Sie denn in diesem Kontext der Bewertung die generativen Ansätze?

Matzner: Wenn jemand ein Video erschafft, kommt es wiederum auf den Einsatzzweck an; meine Vortragsfolien illustriere ich auch mit KI-generierten Bildern. Das halte ich für relativ unkritisch, weil es ein Bild ist, das ich exakt haben möchte, um eine Idee zu verwirklichen. Ich sage damit aber nicht, dass das Bild einen realen Sachverhalt wiedergibt – wenn ich ein Bild von einer Schildkröte verwende, behaupte ich damit ja nicht, dass es exakt diese Schildkröte tatsächlich gibt.

Es kommt also sehr auf den Einsatzzweck an, und wir werden uns überlegen müssen, welchen Medien, AutorInnen und BetreiberInnen wir auch glauben, dass ihre Bilder nicht alle künstlich erzeugt wurden. Wenn ich heutzutage ein Bild sehe, muss ich damit rechnen, dass es künstlich erzeugt wurde. Es muss also bei der Betrachtung immer nach Einsatzzweck und Absicht des Erzeugenden gefragt werden – besonders danach, ob die Absicht einer wirklichkeitsgemäßen Information gegeben ist.

Frage: Ich habe bei diesen Themen immer ein kleines Problem, herauszufinden, was eigentlich genau Künstliche Intelligenz ist.

Ich habe beruflich seit 40 Jahren Software Engineering gemacht – und zwar ziemlich ernsthaftes Software Engineering im Bereich Bankensoftware, Weltraumfahrt und den dazugehörigen Algorithmen; und ich kann eigentlich zu allem, was Sie eben gesagt haben, bestätigen, dass wir das in den 80er und 90er Jahren auch so gemacht haben.

Deshalb meine Frage: Was ist bei Ihnen spezifisch die Künstliche Intelligenz?

Matzner: Sie haben – ohne, dass ich es explizit so formuliert habe, eine meiner Kernideen direkt erfasst. Mein Fach heißt ja Informatikethik; als ich mein Buch geschrieben habe und auch während meiner Lehrveranstaltung hatte ich ursprünglich gar keinen großen Schwerpunkt auf KI; den habe ich nun gelegt, weil KI die Öffentlichkeit aktuell sehr bewegt und ich Antworten auf die entsprechenden Fragen habe. Ich habe neulich einen recht gut geschriebenen Einleitungsartikel gelesen, in dem ein Abschnitt „Risiken von KI“ enthalten war. Wenn ich dort überall das Wort „artificial Intelligence“ durch „automation“ ersetze, macht der Text immer noch Sinn.

Wenn ich überall das Wort „automation“ durch „human or automated solution or mixture of both“ ersetze, macht er auch immer noch Sinn.

Mein nächster Vortrag – das habe ich Herrn Mörike schon gesagt – wird heißen: „Warum wir den Begriff der KI am besten abschaffen sollten“. KI ist ein historischer Terminus; es sind die Techniken, die vor über 70 Jahren, also Anfang der 1950er Jahre, bereits erdacht wurden. Also zu einem Zeitpunkt, als die damaligen Maschinen von ihrer damaligen Rechenkraft und Speicherkapazität diese Aufgaben noch lange nicht ausführen konnten – zu diesem Zeitpunkt hat man aber schon begonnen, über diese zukünftigen Möglichkeiten nachzudenken. Bei KI handelt es sich einfach um eine Reihe von Programmier-Techniken, die nicht einmal alle etwas miteinander zu tun haben.

Alles, was man in den 1950er und 1960er Jahren für hypothetisch und Science Fiction hielt, weil man es noch nicht umsetzen konnte, hat man damals KI genannt. Wenn man aber heute von KI spricht, würde man direkt nachfragen, welches Verfahren damit konkret gemeint ist. Es kommt auch noch dazu, dass das Wort „Intelligenz“ in Künstlicher Intelligenz dazu beiträgt, dass Maschinen vermenschlicht werden.

Natürlich springe ich gewissermaßen auf den medialen KI-Zug auf, der KI häufig verurteilt; aber ich möchte anmerken, dass häufig, wenn der KI-Einsatz medial kritisiert wird, bei genauerem Durchlesen eines entsprechenden Artikels bei mir der Gedanke aufkommt, dass es sich vielleicht um gar keine KI-Software handelt, sondern um eine einfache Regelsoftware.

Frage: Ich denke auch, dass der Begriff „KI“ ein schwieriger ist, da er so viele Dinge verinnerlicht, die meiner Meinung nach getrennt sein sollten, wie es mit dem Begriff „Intuition“ und der Entscheidungsfindung angesprochen wurde. Ich denke, insgesamt müsste man, aus der Arbeitswelt kommend, vor allem ein Assistenz- oder

Delegationssystem betrachten, wenn man von KI spricht. Dabei sind dann natürlich, wenn es um Entscheidungskompetenzen geht, ethische Werte sehr wichtig.

Sie sagten eingangs schon, dass es wichtig ist, zwischen Fake und Realität zu unterscheiden – und dass das für so manche User eine Herausforderung ist.

Matzner: Wir werden uns an neue Medien gewöhnen müssen, so wie wir uns an die jetzt bestehenden und früher neuen ja auch gewöhnen mussten. Ich habe neulich in einer Diskussion einen interessanten Gedanken gehört: Selbst die europaweite Hexenverfolgung hätte ohne die Verbreitung von hexenfeindlicher Literatur – ermöglicht durch die Druckerpresse – nicht so koordiniert erfolgen können.

Mein Eindruck ist: Wir werden bei neuen Technologien immer erst eine Schockphase haben, bei der man merkt, dass es auch negative Konsequenzen geben könnte – und dann werden wir uns Wege ausdenken müssen, um eine solche neue Quelle zu testen und herauszufinden, wie verlässlich sie ist.

Frage: Zu generativen Bildern: Man kann nicht jedes Bild oder Video erstellen, das man sich wünscht. Trainingsbilder geben ja viele Eckdaten vor. Ich glaube, hier steckt auch eine gewisse Gefahr.

Es stellt sich in diesem Zusammenhang unter anderem auch die Frage, wer die zugrunde liegenden Bilder gelabelt hat – es gibt also schon sehr viele Prädispositionen bei den generativen Bildern, wegen derer man trotz Prompt und Negativprompt trotzdem nicht genau das kreieren kann, was man sich wünscht, weil die Trainingsdaten zu großen Einfluss nehmen.

Matzner: Das führt uns wieder zurück zur Frage nach dem Einsatzzweck. Ich beispielsweise verwende Midjourney zum Illustrieren meiner Vortragsfolien mit dekorativen Bildern, die ich so selber nicht malen könnte; ich kann mir dabei zwar vorstellen, was ich abgebildet haben möchte, aber ich kann es nicht selber malen. Dafür ist mir das System – so, wie es sich verhält – auf jeden Fall gut genug. Und wie gesagt: Am Ende bin ich selbst der Richter darüber, ob ich mit diesem Bild an die Öffentlichkeit gehen möchte oder nicht.

Bei der Bilderkennung ist das natürlich etwas anderes, wenn beispielsweise das System nur mit reichen, schönen, hellhäutigen Menschen trainiert wurde – da gab es ja einmal den Fall, bei welchem dunkelhäutige Menschen aufgrund dieser Trainingsdaten gar nicht erst als Mensch klassifiziert wurden.

Man muss sich bei diesen generativen Systemen klar machen, dass sie keine Universalgenies sind.

Bemerkung: Bei durch KI generierten Bildern lassen sich oftmals eine große Menge an künstlerischen, historischen und ähnlichen Referenzen erkennen, auf die natürlich nicht explizit hingewiesen wird. Es ist dann an der nutzenden Person, sich darüber Gedanken zu machen, ob man das entsprechende Bild mit diesen Referenzen verwenden möchte oder nicht. Hier spielen natürlich auch kommerzielle Aspekte eine Rolle – es gibt einfach sehr

viele Verflechtungen und verschiedene Bias-Verzerrungen bei einer unreflektierten Rezeption.

Matzner: Jetzt muss ich Ihnen gestehen, dass dieser ganze Teil – nicht nur generative KI, sondern der ganze Teil, der sich mit neuen Veröffentlichungen befasst – an der Schnittstelle zur Medienethik sitzt. Vieles von dem, was Sie angesprochen haben, muss sicherlich erforscht werden – ich stehe dafür aber nicht in der ersten Reihe.

Frage: Mich hätte noch interessiert, was bei der von Ihnen genannten Literatur von Herrn Kahnemann als Kernargument bezüglich annähernden Regeln versus Intuition steht.

Matzner: Das überzeugende Hauptargument ist dabei, dass schon Mitte des letzten Jahrhunderts eine ganze Reihe an Tests gemacht wurden, die ergeben haben, dass selbst einfache Regelwerke (beispielsweise zur Prognose von Straftätern oder Vergabe von Krediten) verlässlichere Ergebnisse hervorbringen als menschliche Intuition; und das sind Ergebnisse aus dem letzten Jahrhundert.