

TIMES MAGER

Rache



Von Judith von Sternburg

Der Nachbar ist ein freundlicher Mann, grüßt, winkt, stellt die Mülltonnen raus und so weiter. Macht was mit Werbung. Hat Familie. Außerdem hat er einen SUV, aus Diskretionsgründen wird hier die Farbe nicht genannt. Wäre es ein, sagen wir, sehr kleiner Kleinwagen, wäre die Farbwahl pflügend. Wie die Dinge liegen, fällt man vor Schreck hintenüber.

Sobald man sich wieder aufgegrappelt hat, stellt man fest, dass der SUV selbst für einen SUV sehr groß ist, auch ist er eckig, er ist ja gerechelt auf Zack. Da aus zweifellos einen Gesichtsausdruck haben, lässt sich durchaus sagen, dass der SUV des freundlichen Nachbarn die Passanten und Passantinnen grimmig fixiert. Man will ihn nicht nachts auf einer einsamen Straße treffen. Ederfalls nicht, wenn er angekippt ist oder wie auch immer dieses Teil in Gang gesetzt wird.

Natürlich sieht es im Dunkeln in Wahrheit still herum und frist keine Kinder. Dabel erinnert es, wie nachts alle Autos grau sind, an eine andere Art von Fahrzeug.

Vorurteile gegen Autos, sagt der Verwandte, der den freundlichen Nachbar ebenfalls mag, seien überhaupt nicht okay. Man sollte einen Menschen nicht nach seinem Auto beurteilen und die freie Wahl des Fahrzeugs respektieren.

Seltdem wurde nicht mehr darüber gesprochen.

Nun aber berichtet auch der englische „Guardian“ über jenes neue Auto der US-Marke Rezvani, die seit Monaten Ankündigungen macht für das neue Modell. Das Auto heißt „Vengeance“, Rache, Vergeltung. In der Werbung heißt es darum singend: „Dein ist die Rache“. Dabel soll sich aber nicht zum Beispiel Gott angesprochen fühlen, der keine schussicheren Autoscheiben braucht und eh keine Kinder zur Schule bringen muss, sondern lediglich der Käufer (ja, klar, oder die Käuferin), eine Person, die sauviel Geld haben kann. Und kommen wir jetzt zum Punkt: Das Auto traut sich nicht mehr und hat zweifelnd die Farbe „Gutwillig“. Könnte man sagen, es hat zweifelnd die Farbe und Anmutung eines Batmobils. Dann wäre es lustiger und verspielter. Die Spezialausstattung heißt jedoch nicht Batmobilpaket, sondern Militärpaket. Dann lassen sich die Türgriffe unter Strom setzen, stoßen die Rückspiegel Pifferspray aus und hinten können viel Bedarf intensichartig eine dunkle Wolke heraus. Bei Bedarf: Verfolgungsjagden oder so.

Is das doch ein Witz? Nein. Einschlägige Autoportale machen aber darauf aufmerksam, dass die Ausstattung mit „Militärpaket“ in Europa kaum eine Zulassung bekommen wird.

Der Vengeance, klar, ist ein streiner SUV. Das unvornehm und billige Gefühl, immer schon recht gehabt zu haben, es ist einfach nicht zu überreifen.

che, und der Mensch kann das nicht unterscheiden.

Is es nicht vielleicht irgendwann unerheblich, ob wir mit einer perfekten Simulation eines Bewusstseins interagieren oder mit einem anderen Bewusstsein?

Es wird rein faktisch nie dasselbe sein, weil maschinelle und menschliche kognitive Prozesse sich schon allein in ihrer Materialität unterscheiden. Was aber manche Wissenschaftler:innen und Philosoph:innen sagen, ist: Wenn wir die Möglichkeit nicht ganz ausschließen können, dass eine zukünftige Maschine etwas Bewusstseinsähnliches entwickeln kann, dann sollten wir vielleicht so handeln, als hätte sie ein Bewusstsein, rein aus einem Vorsichtsprinzip heraus.

Was meinen Sie dazu?

In einer fiktiven Zukunft ist prinzipiell vieles möglich. Der Mensch und sein Bewusstsein sind im Laufe der Evolution auf eine bestimmte Weise entstanden, und man kann niemals ausschließen, dass irgendetwas anderes irgendwann mal entsteht. Das öffnet Raum für Spekulationen, aber es hat nichts mit den Systemen zu tun, über die wir gerade reden. Ob sich aus Maschinen irgendwann etwas entwickeln kann, das mehr ist als eine Maschine – daran scheiden sich die Geister.

Es gibt die Idee des Mind-Uploadings – ein beliebtes Motiv im Science-Fiction-Genre –, mit der auch Fantasien von der eigenen Unsterblichkeit ehergehen. Halten Sie es für möglich, dass sich ein Mensch in der KI irgendwann perfekt simulieren lässt?

Ganz nüchtern betrachtet geht es hier erst mal um rein statistische Maschinen, die aus alten Datenmaterialien neuen Output produzieren. Natürlich könnte man eine Simulation von mir erstellen, wenn man alle Daten von mir hätte, wie ich kommuniziere, wie ich mich verhalte und so weiter. Aber natürlich ist das nur eine Simulation meines Verhaltens – ich wäre dann trotzdem *ich*. Diese Debatten erscheinen mir oft als eine Obsession vornehmlich reicher Männer, die sich mit ihrer Endlichkeit nicht abfinden können und sich selbst amidst ihren Privilegien für alle Ewigkeit erhalten wollen. Ich halte das nicht nur für wenig wissenschaftler, sondern auch für eine irreführende Debatte.

Was sind die wichtigeren Fragen? KI wird bereits heute in vielfältigen Kontexten eingesetzt, die uns alle – wenn auch auf sehr unterschiedliche Weise – betreffen, von sozialen Medien über medizinische Diagnostik bis hin zur Kreditvergabe. Hier ergeben sich zahlreiche Herausforderungen etwa in Bezug auf Fragen der Autonomie, der Gerechtigkeit und der Privatsphäre. Ich halte es für erheblich dringender geboten, sich diesen Fragen zu widmen. Und – um noch einmal zu betonen – diese Fragen zu beantworten ist ein Menschensache. In Anbetracht des immensen Stromverbrauchs von Sprachmodellen sollte man sich immer auch über die diese Ressourcen sinnvoll eingesetzt werden.

INTERVIEW: LISA BERINS

„Alle Systeme, die basierend auf alten Daten neue Ausgaben produzieren, spiegeln die Gegenwart der Welt, so wie sie in den verwendeten Daten repräsentiert wird, wider und schreiben sie damit in die Zukunft fort“

ChatGPT wurde von OpenAI entwickelt, die mit Microsoft zusammenarbeiten. Googles Sprachsystem – heißt LaMDA. Chatbots sollen bald in Suchmaschinen integriert werden. Könnte man sich vorstellen, dass es in naher Zukunft eine Konkurrenz von KI-gesteuerten Wissenschaftssystemen gibt, die vielleicht dann jeweils eine andere Ausrichtung haben?

Durch die Auswahl der Daten, mit denen ein System lernt, und verschiedene methodische Entscheidungen kann man natürlich auch den späteren Output beeinflussen. Dadurch ist es durchaus möglich, zum Beispiel konservativere oder liberalere, rechte oder linke Tendenzen in diese Sprachmodelle zu importieren. Natürlich können unterschiedliche Firmen unterschiedliche Sprachmodelle entwickeln, aber am Ende spielt es natürlich auch einfach eine Rolle, welches Modell leistungsstärker ist. Ich denke, es werden sich Sprachmodelle durchsetzen, die am passendsten und überzeugendsten auf die unterschiedlichsten Anfragen der Nutzer:innen antworten.

Die Lernfähigkeit dieser Systeme scheint enorm, sie können Schlussfolgerungen ziehen, auf ethische Fragestellungen antworten, Smalltalk halten. Da durch selber denken. Ist es komplett ausgeschlossen, dass irgendwann in diesem System ein Bewusstsein existiert?

Dass ein System wie ChatGPT oder LaMDA ein Bewusstsein hätte, ist Unstinn. Diese irreführende Diskussion geht zurück auf den Turing-Test. Dort wird insinuiert, dass eine Maschine dann denken könnte, wenn ein Mensch nicht mehr zwischen dem Output einer Maschine und dem eines Menschen unterscheiden könnte. Diese Ununterscheidbarkeit scheint ja bei den von ChatGPT produzierten Texten durchaus gegeben. Aber natürlich sagt die Tatsache, dass ein Mensch nicht unterscheiden kann, ob ein Text von einem Menschen oder einer Maschine produziert wurde, rein gar nichts über die (Denk-)Kompetenz der Maschine aus, sondern nur über die Kompetenz des Menschen, zwischen beidem zu unterscheiden. ChatGPT denkt nicht, es versteht nicht, es simuliert nur Spra-

ZUR PERSON

Judith Simon ist Professorin für Ethik in der Informationstechnologie an der Universität Hamburg. Sie beschäftigt sich mit der Verschränkung ethischer, erkenntnistheoretischer und politischer Fragen im Kontext von Big Data, Künstlicher Intelligenz und Digitalität.

Sie ist Mitglied des Deutschen Ethikrates, sowie verschiedener anderer Gremien für wissenschaftliche Politikberatung. Sie war Mitglied der Datenethikkommission der Bundesregierung (2018-2019). fr

Ja, wenn man so hochperformante Systeme wie ChatGPT lösst, ist man natürlich mit verantwortlich für all die positiven wie auch negativen Dinge, die damit geschehen können. Dadurch sind alle an der Entwicklung beteiligten Personen natürlich verantwortlich für ihr Tun – aber eben nicht nur sie, sondern auch die Manager, die entscheiden, ein solches System zu entwickeln – und es dann frei zugänglich zu machen. Spannend fand ich, dass Google nach dem Release von ChatGPT direkt meldete, dass das Unternehmen auch ein solch leistungsstarkes System besäße, es aber aus ethischen Gründen nicht öffentlich gemacht habe. Jetzt ist das Programm der Firma OpenAI auf der Welt, und es wird vieles umwälzen.

Mit besser werdender Technik sourcen wir also unsere eigene Bildung weiter aus – und machen uns damit von der künstlichen Intelligenz abhängig. Ist das nicht alarmierend?

Zumindest wird das in dem Moment ein Problem, wo wir Kompetenzen verkümmern lassen, die wir für unabkömmlich erachten, oder wenn wir uns so stark auf digitale Technologien verlassen, dass bei einem Ausfall nichts mehr geht. Was passiert zum Beispiel, wenn in einer immer stärker roboterunterstützten Chirurgie die Fertigkeiten des Operierens nicht mehr genügend eingeübt werden und dann die Maschinen – kurzfristig oder dauerhaft – ausfallen? Auf der anderen Seite kann man auch sagen: Heute weiß kaum einer mehr, wie man Feuer ohne Feuerzeug oder Streichholz macht. Auch in Zukunft werden also bestimmte menschliche Fertigkeiten verloren gehen – und das wird teilweise auch verschmerzbar sein, wenn sie einfach nicht mehr so relevant sind.

Was wäre fatal, wenn es verloren ginge?

Eine schwierige Frage. Ich unterrichte Ethik und Philosophie in der Informatik, und mein Ziel ist es, meine Studierenden zum Nachdenken zu bringen über all das, was wir oft als gegeben annehmen. Eine meiner – zugegeben etwas vagen – Sorgen ist, dass vielleicht diese Fähigkeit des grundsätzlichen Hinterfragens verloren gehen könnte, wenn wir zunehmend von Systemen umgeben sind, die die Zukunft aus der Vergangenheit freischieben und dies noch dazu verschleiern. Ähnliche Fragen stellen sich vielleicht auch in der Kunst. Es ist unmöglich verkraftbar, wenn ein Chatbot Krimis schreibt. So gern ich Krimis lese – der Plot ahnt sich ja doch oft. In anderen Bereichen der Kultur hingegen stellt sich schon die Frage, ob die Fähigkeit, einen Unterschied zu machen, verloren geht in einer zunehmend statistisch vorgeprägten Welt. Und diese grundsätzliche Frage, einen Unterschied machen zu können, besteht natürlich nicht nur in Kunst und Kultur, sondern auch in jedem anderen Bereich unseres Lebens.

Besonders an den Universitäten wird ChatGPT einiges umwerfen, vor allen Dingen in Bezug auf Prüfungen und Hausarbeiten ... Die Universitäten mussten sehr schnell auf ChatGPT reagieren, es wurden unmittelbar Diskussionsforen zu Prüfungsfragen losgetrieben. Vordringend geht es um die Frage, wie wir Prüfungen fair gestalten können, einen Unterschied machen zu können, besteht natürlich nicht nur in Kunst und Kultur, sondern auch in jedem anderen Bereich unseres Lebens.

Das ist ein Problem, das ich sehr gerne mit den Studierenden diskutiere. Es geht um die Frage, wie wir Prüfungen fair gestalten können, einen Unterschied machen zu können, besteht natürlich nicht nur in Kunst und Kultur, sondern auch in jedem anderen Bereich unseres Lebens.

Anstatt wissenschaftlich gesicherter Fakten verbreiten solche Systeme dann Algorithmen-Unfug. Sie generieren sozusagen alternative Fakten – und das klingt schon sehr gefährlich, oder?

Das ist ein Problem, das ich sehr gerne mit den Studierenden diskutiere. Es geht um die Frage, wie wir Prüfungen fair gestalten können, einen Unterschied machen zu können, besteht natürlich nicht nur in Kunst und Kultur, sondern auch in jedem anderen Bereich unseres Lebens.



SEBASTIAN ENGELS

nen, menschlich zu sein, ist es also umso entscheidender, dass man diese Gefahren vermeidet.

Wie könnte das geschehen? Könnte man diesen Programmen beibringen, um etwa moralische oder politische Grenzen liegen?

Es gibt Methoden, mittels derer Entwickler:innen ihre Systeme in Bezug auf Biases und Verzerrungen zunächst testen und dann diesen entgegenwirken können. Das ist aber alles andere als trivial und Gegenstand intensiver Forschung. Softwaressysteme sind durchaus in der Lage, wenn auch nicht fehlerfrei, politische Haltung in Texten zu erkennen – auch das ist ja letzten Endes eine Form der Mustererkennung. Man kann dann technisch versuchen, die Wahrscheinlichkeit der Produktion solch problematischer Inhalte zu verringern. Sindre die Veränderungen positiver Natur – und falls ja für wen?

„Für ChatGPT ist ein wissenschaftlicher Artikel nur ein Textgenre unter vielen. Der Artikel und die Quellen können frei erfunden sein“

Die Entwicklerinnen und Entwickler halten eine große Verantwortung in ihren Händen.

„ChatGPT versteht nicht, es simuliert nur Sprache“

Die Professorin für Ethik in der Informationstechnologie Judith Simon über Funktionsweisen von Sprachsystemen, potenzielle Gefahren für die Gesellschaft und die Frage nach einem künstlichen Bewusstsein

Gibt man dem System dann einen Prompt – das heißt, eine Frage beziehungsweise Anweisung, welche Art von Text zu welchem Thema produziert werden soll –, werden neue Texte basierend auf diesen zuvor erlernten Mustern produziert. Das funktioniert mittlerweile so gut, dass man oft nicht mehr unterscheiden kann, ob der Text von diesen Sprachsystemen oder von Menschen produziert wurde.

Ein klassisches Beispiel war hier der Chatbot Tay von Microsoft, der innerhalb von Stunden aufgrund rassistischer und sexistischer Kommentare wieder vom Netz genommen wurde. Hier lernte der Chatbot dieses Verhalten von Twitter-Nutzern, die teilweise wohl auch bewusst die Schwachstellen von Tay ausnutzten. Auch Vorgänger von ChatGPT hatten nachweislich dieses Problem, dennoch hat OpenAI sich entschieden, das Tool frei verfügbar zu machen.

Bedeutet das, dass diese Sprachmodelle in der Gesellschaft vorhandenen Ungleichheiten und Ungerechtigkeiten verfestigen?

Alle Systeme, die basierend auf alten Daten trainiert werden, produzieren, spiegeln die Gegenwart der Welt, so wie sie in den verwendeten Daten repräsentiert wird, wider und schreiben sie damit in die Zukunft fort. Das gilt für ChatGPT genauso wie für



„Hintergründig fordert uns ChatGPT auch heraus, neu darüber nachzudenken, was Bildung ist und sein soll“, sagt Judith Simon.

Frau Simon, ChatGPT wird seit Wochen gehypt, und gleichzeitig wird scharf davor gewarnt. Können Sie diese gespaltene Reaktion nachvollziehen?

Die Reaktionen auf ChatGPT sind sicher deswegen so gespalten, weil das Tool ein so großes Potenzial hat. Entsprechend ergeben sich große Möglichkeiten, aber gleichzeitig auch sehr viele Herausforderungen. Es ist natürlich so, dass Medien, neue Technologien – von der Schrift, über den Buchdruck zum Computer –, oft große Sorgen ausgelöst haben, dass alles den Bach runtergeht. Letztlich haben Menschen und Technologien sich dann aber gemeinsam weiterentwickelt. Allerdings finden diese Veränderungen heute erheblich rasanter statt.

Chatbots wie ChatGPT greifen auf alle möglichen Daten aus dem Netz zu, auf Texte aus Onlinenetzen und sozialen Netzwerken. Wie verarbeiten die Chatprogramme diese Daten?

Diese großen Sprachmodelle erkennen im Grunde genommen Muster in all diesen Daten, mit denen sie gefüttert wurden. Dabei geht es zum einen um Wahrscheinlichkeiten von Wortkombinationen, also welches Wort folgt auf welches mit welcher Wahrscheinlichkeit. Aber es geht auch darum zu erkennen, durch welche Muster sich bestimmte Textgenres, etwa Kriminalromane, Essays, Zeitungsartikel oder wissenschaftliche Artikel, auszeichnen.