

Alpha-Plus oder Epsilon-Minus

Was zählt Urteilskraft in der schönen neuen KI-Welt? Und besitzt das von einer App komponierte beste Osterhasenlied aller Zeiten eine Aura? Philosophische Befragungen von Marie-Luise Frick

Unsere Lebenswelt wird sich in allen Bereichen verändern, raunte es, als Ende 2022 erstmals ein computerlinguistisches Wahrscheinlichkeitsmodell zur freien Nutzung zur Verfügung gestellt wurde. Was durch Open AI mit ChatGPT eingeläutet wurde, deuteten Technik-Pessimisten rasch als ein Zeitalter des Überflüssigwerdens – überflüssige Arbeitskraft, überflüssige Denkleistung, überflüssige Kreativitätsanstrengungen – und des Kontrollverlustes. Die Euphoriker hingegen sprachen von Prozessoptimierung, Produktivitätssteigerung und dem Abbau von Ungleichheiten zwischen Könnern und Amateuren, zwischen Begabten und denen, die sich abmühen in der Texterstellung, Illustration oder Komposition. Beide Lager waren sich, zusammen mit den Unschlüssigen dazwischen, jedoch einig, dass das Rad der technologischen Entwicklung sich nicht zurückdrehen lassen wird – zu viel Kapital, zu viel brain power, zu viel Ehrgeiz treiben dieses Rad schließlich an. Ein Hype ging um und hyperbolisch waren folglich auch beide – die Hoffnungen und Befürchtungen.

Inzwischen ist die Zahl der generativen KI-Modelle weiter angewachsen. Mit ihrer Hilfe lässt sich von Text und Code über Bilder bis zu Songs alles generieren und die frühen Versionen von generativer KI, die mit ihren Wissenslücken und Halluzinationen noch für Lacher gesorgt haben, haben trainiert. Oder anders gesagt: Sie haben so lange so viele Daten gefressen, bis sie ihre Wahrscheinlichkeitsrechnungen richtig hinbekommen. Und weil bei solcher Fülle schnell einmal etwas im Bauch durcheinanderkommen kann, haben die besseren KI-Modelle noch control-layers draufbekommen und gelernt: Zuerst denken, dann sprechen. Wenn üblicherweise auf das Wort x das Wort y folgt, wenn meist dies und jenes zusammenhängt, wenn bisher immer solches galt, dann muss es wohl so stimmen. Gerechnet, geraten oder gewusst? Macht das noch einen Unterschied für Hunderte Millionen Nutzerinnen

und Nutzer von US-amerikanischen, europäischen und inzwischen auch chinesischen KI-Orakeln?

Viele von ihnen mögen mittlerweile zwar grob verstanden haben, wie große Sprachmodelle algorithmisch ticken (genau wissen es oftmals nicht einmal die Entwickler), also wie der Auffindungszusammenhang der ausgespuckten Ergebnisse beschaffen ist. Ob sie sich aber darüber Gedanken machen, inwiefern der Rechtfertigungszusammenhang der KI-Ergebnisse außerhalb dieser Modelle angesiedelt ist, darf als fraglich gelten. Tiefe Wahrheitsfragen und ihre Verstrickungen lauern hier: Ist wahr dasjenige, welches übereinstimmt mit dem, was Menschen üblicherweise oder bisher so meinen, sagen, behaupten? Korrespondenztheorie, anyone? Die Welt der KI-Modelle ist Menschenwelt, eine Dauerschleife von Beschreibungen, fensterlos. Sie ist nicht die Wirklichkeit jenseits unseres Wähnens. Can't think outside a black box. Und diese Menschenwelt, wie ugly sie doch manchmal ist! Voller bias, politisch inkorrekt, Abweichungen von der guten Erziehung. Da müssen dann Filter her. Moral priming von KI-Modellen ist jedenfalls ein neues Berufsfeld, das der Angst, KI würde uns alle arbeitslos machen, entgegenhält. Vielleicht wird es so wichtig werden wie der aufstrebende Tätigkeitsbereich der Prompting-Coaches: Personen, die ungeübten Nutzerinnen und Nutzern dabei helfen, mit den richtigen Eingabeaufforderungen die besten Antworten aus der KI herauszukitzeln. Wer willst du sein in der schönen, neuen KI-Welt? Alpha-Plus oder Epsilon-Minus?

Vor kurzem war ich auf einem Workshop zu Künstlicher Intelligenz und Schulbildung an der Uni Graz. Ein Lehrer präsentierte eine Umfrage unter 150 Schülerinnen und Schülern der Oberstufe, der zufolge 40 Prozent mehrmals pro Woche KI-Tools nützen, 10 Prozent sogar täglich. Am meisten genutzt wird dabei ChatGPT (95 Prozent), dessen generierte Informationen für 38

Prozent der Befragten vertrauenswürdig sind, für 8 Prozent sehr vertrauenswürdig. Okay, denke ich mir, so ist das nun einmal heute. Immerhin haben diese jungen Leute den entscheidenden Teil ihrer Schulzeit ohne KI erlebt, sie können dann vielleicht kritischer beurteilen, was an künstlicher Intelligenz intelligent ist und was nicht. So, wie Menschen meiner Generation anders mit Smartphones umgehen, weil sie ein Aufwachsen ohne Handy noch erinnern können, als jene, die von Anfang an hineingeworfen werden in digitale Weltensimulationen. Dann denke ich an all die Maturantinnen und Maturanten, die uns die Schulen mit glänzenden weißen Flecken auf ihren Allgemeinwissenslandkarten an die Hochschulen schicken, und durchschaue mein resignatives Wunschdenken als das, was es ist. Nein, denke ich dann: Ich glaube nicht ehrlich daran, dass jene 46 Prozent, die KI-Ergebnisse für vertrauenswürdig halten, beurteilen können, wann sie welchen Ergebnissen vertrauen dürfen und wann nicht. Macht mich das zu einer Kulturpessimistin?

Es sei, schrieb Descartes 1644 in seinen *Principia Philosophiae*, „zweifellos vorzuziehen [...], daß ein Mensch sich seiner eigenen Augen bedient, um seine Schritte zu lenken, und sich mit Hilfe derselben an den Schönheiten der Farben und des Lichts erfreut, als daß er blindlings der Führung eines anderen folgt.“ Wie schön, wie schön unpraktisch. Ist es nicht weit zu viel verlangt, ständig alles kritisch zu prüfen und zu hinterfragen? Kann man so überhaupt leben? Nein, müsste auch in größter Verbundenheit mit den Idealen der Aufklärung allen Imperativen des Selbstdenkens und Selbst-in-Erfahrung-Bringens entgegengerufen werden. Das geht sich in einer globalen Wissensgesellschaft hinten und vorne nicht mehr aus! Der Aufruf „Habe Mut, dich deines eigenen Verstandes ohne Leitung eines anderen zu bedienen“ ist noch nicht affiziert von prometheischer Scham des antiquierten Menschen. Er ist heute gleichbedeutend mit „Habe den absurden Eigenwillen, dich deines Verstandes ohne die nützliche und wohltätige Hilfe von Milliarden anderer, nicht zwangsläufig dümmere Menschen, lebenden und toten, zu bedienen“. Wer diesen tollen Mut nicht aufbringen kann, wird nicht umhinkommen, einen Weg

zu furchen zwischen Vertrauensseligkeit und Argwohn, zwischen Energieeinsatz in innovative Produktion und Energiesparen durch indolente Reproduktion. Menschen sind schließlich keine Maschinen. Noch. Sie müssen das anwenden, was Maschinen, ohne Bewusstsein, nicht entwickeln können: Urteilkraft. Eine ganz besondere Kraft, da sie selbst Kraft kostet, bevor sie Kraft in Macht umwandelt: im Drehmoment des Erwägens, im Umdrehen und Wieder-neu-Erwägen, im Abwägen, im Urteilsfällen schließlich. Dort, wo dies zur Last wird, versprechen KI-Algorithmen heute schon existenzielle Entlastung. Je nach eingespeisten Regeln prognostizieren sie in einigen Ländern bereits mehr oder weniger treffsicher Straffälligkeitsrisiken oder schlagen Angriffsziele vor, wie aktuell im Gaza-Krieg. Human on the loop. Human out of the loop. Human off the hook? Ist das der Endpunkt der neuen, schönen KI-Welt, in der wir alle so unendlich frei sein werden, weil Verantwortlichkeit zur Unkenntlichkeit zerstäubt?

Wie viele Geistes-Arbeiterinnen und -Arbeiter stelle auch ich mir die Frage, wie viel von meinen kognitiv-kreativen Prozessen – so undurchschaubar in vielem – bereits heute ein KI-Modell übernehmen könnte? Ohne dass es auffällt. Wem auffällt? Mir oder anderen? Und die Zukunft erst! Würde man ein KI-Modell nur mit allen je von mir publizierten Texten (und von mir aus auch allen je vergebenen Likes und Smileys mit samt ihren Kontexten) füttern, würde es dann sprechen wie ich, denselben Ton treffen, den gleichen Stil haben? Zu neuen Problemen und Fragen sich verhalten so in etwa wie ich, selbst wenn ich schon hunderte Jahre tot bin? Wäre dieses „in etwa“ gut genug? Bin ich Gläubiger eines abgehalfterten Genie-Begriffs, wenn ich daran zweifle? Bin ich narzisstisch gar? Jetzt wird es ernst.

Zur Osterzeit letzten Jahres saß ich mit einem befreundeten Musiker und KI-Experten zusammen und wie öfters nach langen Arbeitstagen ging es redlich unernst zu. Bis wir uns schließlich mit einer generativen Musik-KI unterhielten. Unser Prompt sah in etwa so aus: Der Freund wollte unbedingt Lyrics, in denen ein Easter Bunny vorkommt, „a satanic one“. Für mich war dann Dubstep/Reggae die passende Genrewahl. Nach we-

nigen Sekunden erklang es: das beste Oster-Lied ever! Wir hören es heute noch ab und an, um laut zu lachen und auch weil es, das muss gesagt sein, ein echt grooviger Song ist. Ein ungekannter Hit. Niemand würde nur vom Zuhören erraten, dass nie ein Mensch daran komponiert hat, nie ein Mensch dafür ein Instrument in die Hand genommen oder ein Studio betreten hat. Wenn bei Liedern einzig zählt, was sie als Ergebnis – als Konsumgut – darstellen, was bedeutet das für die Musikkunst?

Die Maschine kreiert, der Mensch performt höchstens, wenn überhaupt? Wer braucht noch Live-Musik, wer braucht überhaupt noch Musikerinnen und Musiker oder Bands, wenn sich jede und jeder mit dem richtigen prompting – oder dem richtigen Coach dafür – per Knopfdruck total private Lieblingslieder erzeugen und vorspielen lassen kann? Braucht es eine „KI-Steuer“, wie manche Kulturschaffende fordern – mit welchen künstlerischen Erzeugnissen werden generative KI-Modelle denn gefüttert, wenn nicht mit denen von Menschen, die dafür nicht bloß gepromptet, sondern kreiert haben? Wird es bald Charts für die besten „self-made“ KI-Hits geben? So wie Ausstellungen von KI-generierten Bildern? Wahrscheinlich gibt es das längst (ich tippe mal schnell auf Südkorea) und dann denke ich an die vielen Bands, die ich je live gesehen habe. An die Faszination für die Biografien von Lieblingsmusikern, für die Inspirationsgenese ihrer Songs, für das Ereignishafte einer zündenden Idee, einer genialen Begegnung. Wie jene von Jim Morrison und Ray Manzarek im Sommer 1965 in Venice Beach, wie man in dessen Memoiren so schön nachlesen kann. Moonlight. Gänsehaut. Drive. Kein KI-generierter Hit wird darum je wissen, egal, wie gut er imitiert und micro-sampling betreibt.

Und würde ich in ein Museum gehen, in dem ausschließlich künstlich generierte Werke gezeigt würden? Würde ich anders durch so ein Museum gehen als letztens durch das Rijksmuseum, wo die Gemälde der Alten immer auch über eine Zeit sprechen, eine Gesellschaft, die einmal wirklich war und die uns bis heute prägt? Würde ich anders durch so ein Museum gehen als durch ein Museum für zeitgenössische

Kunst, das alles zeigt, viel Interessantes, selbst Schönes, nur seltenst das Erhabene? Was also ist Kunst? Ich bin nicht sicher, ob man diese Frage im Zeitalter der generativen KI wieder aufgreifen wird (müssen) oder ob sie sich nun endgültig erledigt hat. Vor 90 Jahren hat Walter Benjamin seinen viel zitierten, seltener gelesenen Aufsatz „Das Kunstwerk im Zeitalter seiner technischen Reproduzierbarkeit“ im Pariser Exil verfasst. Das Kunstwerk hat es also schon lange nicht mehr leicht. Drucke, Massenproduktion von Gebilden, Filme. Für Benjamin war mit den neuen technologischen Möglichkeiten etwas in Gefahr geraten, das er als den „empfindlichsten Kern“ eines Kunstwerkes ansah: seine Echtheit. Ein echtes Kunstwerk habe historische Zeugenschaft, habe deshalb Autorität. „Was im Zeitalter der technischen Reproduzierbarkeit des Kunstwerkes verkümmert, ist seine Aura“, schrieb er. Ich stutze. Hat ein Kunstwerk eine Aura? Erhält es eine solche erst, wenn ich weiß, dass es ein genuines Kunstwerk ist (was einen Ausweg aus Benjamins Problemstellung versperrt), oder erkenne ich echte Kunst an ihrer Aura? Wer coacht mich dann bitte im Aura-Aufspüren? Darf es auch eine KI sein? Und was beißt sich hier gerade in den Schwanz?

Bevor meine Gedanken sich im Überschlag zerschlagen, fallen mir zum Glück die Worte des besagten Freundes ein, der bei einem unserer jüngsten Gespräche Benjamins Thesen mit der Weisheit der Nachgeborenen und mit der Gelassenheit der digital natives betrachtet. „Moderne Technologien lösen alte Techniken immer nur teilweise ab“, meint er. Klar verändere generative KI den Kunstprozess, aber letztlich werden sich Mensch und KI ergänzen und sogar neue Nischen würden sich bilden für altes Hand-Werk. Keine Panik, also. Ich wünsche mir, dass er recht hat, denke ich. Und fast wäre mir ein mäßig witziges Zitatspiel rausgerutscht. Come on AI, light my fire!

[mit Dank an R.L.]