

Vortrag „Digitalismus – die Utopie einer neuen Gesellschaftsform vom 24.04.2023, Daniel Reborn

Zusammenfassung der Fragerunde

Ob man Entscheidungen an eine KI abgeben kann – das sehe ich wegen unseres aktuellen Rechtssystems sehr kritisch, denn natürlich kann eine KI nicht ins Gefängnis gesperrt werden, wenn sie eine Fehlentscheidung getroffen hat. Im praktischen Fall sieht man schon heute in der Medizin, dass KI in der Diagnose eingesetzt wird – und zwar sehr sorgfältig. Aber es ist dabei immer noch ein Arzt dazwischen, also „human in the loop“. Und zwar, weil praktisch eine „false positive“ Entscheidung genauso schlecht ist wie eine „false negative“ Entscheidung – daher wird die KI nach heutigem Stand zunächst einmal als Hilfestellung betrachtet.

Die letzten KI Beispiele, die Sie gebracht haben, sind sehr gut, aber ich sehe sie auch als Vorboten für eine neue Ära, in die wir mit ChatGPT-4 und folgenden jetzt gerade erst eintreten.

Und dadurch erreichen wir nun langsam eine neue gesellschaftliche Position: Wir haben nämlich die Einstiegsschwelle stark gesenkt. Wir verlassen wir jetzt sozusagen das Feld, in welchem KI nur von Experten eingesetzt wird: Jeder kann ChatGPT ohne Weiteres nutzen und dann die Ergebnisse glauben, die herauskommen. Die Fehlleistungen sind hierbei natürlich gigantisch: Die Ergebnisse kommen leider auch sehr eloquent und selbstbewusst daher, weshalb sie leicht geglaubt werden.

Bei Ihnen, Herr Reborn, gefällt mir die Kausalität nicht, wenn Sie sagen, dass es die Digitalisierung ist, die die Gesellschaft verändert. Es sollte eigentlich umgekehrt sein: Die Gesellschaft sollte sich die Frage stellen, mit welcher Welt sie eigentlich leben möchte – Technik wurde durch die Gesellschaft gestaltet.

Und ich glaube, die Frage der Technikgestaltung kommt jetzt gerade mit GPT-4 verstärkt wieder und mit großer Aufregung auf.

Deshalb würde ich gerne Ihre Einschätzung hören, wo wir im Moment stehen und was Sie denken, wie der Erregungszustand der Seele gerade ist.

Bevor ich diese Frage beantworte noch einmal kurz: Es gibt ja den Begriff des technologischen Imperativs, der besagt, dass alles, was der Mensch technologisch erfunden hat, auch eingesetzt wird. Das reicht bis hin zur Atombombe: Man hat sie erfunden und dann auch eingesetzt. Ich stimme Ihnen zu, dass wir mitgestalten und formuliere dabei auch einen Aufruf, selbst aktiv mitzugestalten und die Entwicklungen nicht einfach passieren zu lassen, so wie wir jetzt die sozialen Medien sozusagen in einem weltweiten Experiment auf die Menschheit losgelassen haben.

Von daher würde ich sagen, dass es auch keinen Widerspruch gibt zu dem, was Sie gesagt haben: Ich bin Ihrer Meinung, dass wir mitgestalten müssen – und im Moment wird zu wenig gestaltet.

Auf Ihre Frage, was denn aktuell passiert: Ich kann es vermutlich nicht sagen. Aber es gibt beispielsweise eine schöne, spannende Antwort zu der „false-positive“ und „false-negative“

Thematik: Wenn man sich zum Beispiel als Patient nicht sicher ist, ob ein Arzt die richtige Wahl ist, dann geht man zu einem zweiten oder einem dritten Arzt. Wenn unser Problem ist, dass wir ChatGPT nicht glauben, dann könnten wir auch einfach drei oder vier unterschiedliche künstliche Intelligenzen füttern und die Ergebnisse vergleichen.

Das wäre eigentlich das gleiche Prinzip, wie wenn ein Mensch die Ergebnisse noch einmal überprüft: Es werden einfach nur mehr Instanzen, die involviert sind. Ich glaube, darauf wird es hinauslaufen: wir werden sich selbst regulierende Systeme haben, die irgendwann (das ist meine Voraussage für eine Zeitspanne von 10 Jahren oder mehr) dann allein gelassen werden können, weil der Mensch am Ende inkompetenter sein wird als die Summe der vielen Algorithmen.

Ein kleiner Einwand dazu: Wir haben insofern Neuland, dass die Maschinen, die der Mensch gebaut hat, vom Menschen in ihrer Vorgehensweise bisher verstanden wurden – ich bin richtig erschrocken, als ich bei ABC News ein Interview mit Sam Altman (CEO von OpenAI) gehört habe. Er sagte im Gespräch über ChatGPT-4 plötzlich: „I am scared“ – und zwar, weil er ChatGPT nicht mehr versteht – weil das System Handlungen vollzieht, die ihm nicht beigebracht wurden; er ist also von seinem eigenen Produkt überrascht.

Und auf die Frage, was jetzt passieren soll, sagte er, dass ChatGPT reguliert werden soll – das ist übrigens das erste Mal, dass ich von jemandem in seiner Position den Wunsch nach einer Regulierung höre. Die zweite Frage war dann, wer regulieren soll – und darauf sagte er, dass das die Gesellschaft („society“) machen soll.

Die Gesellschaft ist aber doch aber am wenigsten darauf vorbereitet, zu regulieren. Wir sind also in einem recht chaotischen Zustand: Ein Tool wurde einfach so losgelassen und jetzt wird nach einem Moratorium verlangt – das ist, als würde ich die Stalltüre schließen, nachdem die Pferde schon nach draußen gerannt sind. Und jetzt rasen die Pferde durch die Gegend und wir schauen, wie das auf die Gesellschaft wirkt.

Und das ist jetzt wirklich eine Gestaltungsaufgabe für alle, die Ahnung haben – und das bei einer hohen gesellschaftlichen Relevanz. Das ist das, was mich ein bisschen beunruhigt.

Ich würde einfach sagen: Volle Zustimmung, das sehe ich genauso. Meiner Meinung nach muss man auf jeden Fall anfangen, aktiver zu gestalten.

Herr Rebhorn, was glauben Sie, wer denn gestalten könnte? Meiner Vermutung nach sind die allerwenigsten in unserem Land oder auch auf der Welt überhaupt in der Lage, KI richtig zu verstehen – wer sollte denn KI dann gestalten? Politiker, Wissenschaftler, Firmenlenker oder Philosophen? Wen würden Sie empfehlen?

Ich versuche mir, darüber, wer was wann und wie gestaltet, nicht zu viele Gedanken zu machen. Aber: Ich glaube, dass diese Veränderung, die wir gerade sehen, nicht erst heute und nicht nur durch ChatGPT passiert – sie zeigt das sogenannte „Eine-Meile-Phänomen“ auf. Ich glaube, dass wir vor einer großen Veränderung stehen – wir haben es weltweit geschafft (wenn man es denn wollen würde), sich zusammzusetzen und eine weltweite Kooperation aufzustellen; dabei werden natürlich nicht immer die ethischen Aspekte berücksichtigt.

Deshalb plädiere ich ja auch dafür, dass wir für eine KI ein System von Werten brauchen, die vielleicht auch unterschiedlich sein können – in unterschiedliche Regionen für unterschiedliche Themenfelder. Aber auch ich glaube, dass das nicht durch eine einzelne Person geleistet werden kann.

Und wenn in einem Startup Unternehmen danach gefragt wird, warum eine KI eine Kauf- oder Verkaufsentscheidung getroffen hat und wie sich diese Entscheidung entwickelt hat – keine Ahnung! Die künstliche Intelligenz hat in den Kursverläufen Dinge entdeckt, die ansonsten nicht entdeckt worden wären. Aber welche Dinge das genau sind und warum diese entdeckt wurden, kann im Nachhinein wahrscheinlich niemand mehr nachvollziehen.

Und deshalb brauchen wir eine interdisziplinäre Zusammensetzung von Profis und Spezialisten ihres eigenen Bereichs.

Problematisch ist allerdings, wenn es kein Zielbild gibt. Die EU hat einen Vorschlag gemacht, wie KI implementiert werden könnte – es wird aber noch viele Jahre dauern, bis der Vorschlag angenommen werden und dann erst umgesetzt werden könnte – da sind, um auf den Vergleich von vorhin zurückzukommen, die Pferde schon längst aus dem Stall und über alle Berge.

Ich glaube, die Zeit, die uns noch zum Regulieren bleibt, wird immer kürzer; ein Moratorium von sechs Monaten hilft da normalerweise nicht.

Als Ergänzung zu dem bisher Gesagten: Es ist wirklich schwierig, weil diese Themen durch ein Tool wie ChatGPT sehr plötzlich relevant geworden sind. Eigentlich ist KI schon seit vielen Jahren präsent und niemand macht sich mehr Gedanken darüber, dass die Wettervorhersage, Stauprognosen oder Notfallbremssysteme nur mit KI möglich sind. Wir nehmen das dankbar hin, weil wir feststellen, dass es meistens gut funktioniert – nur manchmal ist es ärgerlich, wenn wir merken, dass eine Maschine nicht mehr funktioniert und wir nicht verstehen, warum sie nicht funktioniert.

Aber die Regulierung, das Wertesysteme und die Frage, wer sich kümmern kann – müssen wir da nicht wirklich auf ganz Europa schauen? Ich glaube nicht, dass es möglich ist, ein weltweit gültiges Wertesystem zu schaffen. Es wäre aus meiner Sicht eine wirklich tolle Chance, in Europa eine gemeinsame Vision aufzubauen, die ganz sicher auch für die EU selber gilt.

Was mich als Frage an Sie brennend interessieren würde: Reden wir nicht vielmehr darüber, wie die Kooperation Mensch-Maschine besser werden kann – sodass wir die Dinge, die die Maschine besser kann als wir, nutzen können, und umgekehrt ?

Natürlich ist es sehr naheliegend, zu sagen, dass die Durchführung von Entscheidungen nicht gleich an eine Maschine übergeben werden soll. In meinem tiefen Inneren fühlt es sich für mich natürlich auch nicht gut an, alles an eine Maschine zu übergeben. Aber ich persönlich glaube Folgendes: Selbst in Europa haben wir hochgradig populistische Systematiken. Ich glaube, für uns ist es in der Summe und langfristig besser, wenn wir einen Großteil der Entscheidungen nicht mehr von Menschen treffen lassen.

Ich zitiere dazu gerne meinen Freund, der das Unternehmen mit mir zusammen in Indien aufgebaut hat: „Die Systeme und Prozesse, vor allem in öffentlicher Hand, in denen man den Menschen vollständig eliminiert hat sind erstens schnell, zweitens weitgehend fehlerfrei und vor allem drittens nicht mehr korrupt.“

Der indische Staat hat beispielsweise versucht, die Passbeantragung ohne Menschen durchzuführen. Früher war es so, dass man Bestechungsgeld in den Antrag legen musste, wenn man den Pass schneller haben wollte. Heute ist der Antrag vollständig digitalisiert und so etwas passiert nicht mehr. Wir Menschen neigen leider zu einer Unzulänglichkeit in dieser Hinsicht (Bestechung). Deswegen bleibe ich erst einmal bei der These: Erst durch den Wunsch, diese Vorgänge in die Hände von Maschinen zu legen, werden wir den nächsten Schritt gehen.

„Der ich rief die Geister, werd‘ sie nicht mehr los“ – damit müssen wir meiner Meinung nach rechnen. Wir haben jetzt zwar den ersten Meilenstein, wo die KI richtig ins öffentliche Interesse rückt und auch für viele Firmen präsent wird, erreicht. Aber es gibt noch einen zweiten Meilenstein: Wenn nämlich die KI aus dem Schatten tritt und mit ihr nicht mehr nur Kraftwerke betrieben werden, sondern wenn wir einen sichtbaren und für jeden zugänglichen Bereich haben, wo die KI der „bessere Bearbeiter“ als der Mensch ist.

Das ist etwas, wo ich Ihnen, Herr Rebhorn, voll zustimme: Man müsste den Menschen aufrüsten, damit er vernünftig mit KI umgehen kann. Es kann ja nicht sein, dass man die KI immer weiter entwickelt, bis sie Entscheidungen trifft, die womöglich besser als die menschlichen sind – denn dann kommen wir ja an einen Punkt wie in utopischen Romanen: Man bräuchte nur noch die KI, mit der die Welt bestens funktioniert, und der Mensch hat nichts mehr zu tun.

Passend dazu ist vielleicht der Begriff „AI Prompter“: Das ist ein neuer Job, bei welchem man Menschen dazu anleitet, einer KI (wie ChatGPT) die richtigen Fragen zu stellen, um bestimmte Antworten zu bekommen. Vor 15 Jahren war ein solcher Job vielleicht noch völlig utopisch – aber es gibt heute wirklich Ausschreibungen dazu.

Ich wollte nochmal eine Lanze für die EU und die Regulierung brechen. Im Grunde wird im AI-Act gerade die Frage beantwortet, die ich vorhin gestellt habe: In welcher Gesellschaft wollen wir leben? Bei der Europäischen Union wurde meiner Meinung nach ein Kraftakt vollbracht: Sie haben innerhalb der letzten zwei Jahre eine gemeinsame Linie innerhalb von 27 Mitgliedsstaaten geschaffen.

Sie haben dabei etwas sehr Kluges gemacht: Sie haben eine Risikopyramide erstellt. Es geht außerdem nicht darum, die Technologie zu regulieren, sondern die Wirkung der Technologie. Ein hohes Risiko besteht dann etwa bei Applikationen, die Auswirkungen auf Menschenleben haben; in solchen Fällen werden die Anbieter gezwungen, Transparenz herzustellen, und auch Risikomanagement zu betreiben.

Das finde ich sehr sinnvoll, weil dabei eben diese Rückkopplung des Einflusses der Gesellschaft auf die KI stattfindet.

Übrigens ist es so, dass weltweit Druck aufgebaut wird: Die USA sind in der Entwicklung von KI weit hintendran, Europa führt vor ihnen. Von den USA wird nun Druck auf Europa aufgebaut, dass sie in ihrer Führungsposition regulieren müssen. Bisher hieß es allerdings wiederum von Seiten Europas, dass China vor ihnen führe und nicht reguliere – und daher bei Europa kein Handlungsbedarf bestehe.

Doch jetzt gibt es einen Gesetzesentwurf zur Regulierung von KI in China: In diesem wird allerdings nur die eigene Gesellschaft geschützt, die ja durchaus anders strukturiert ist als unsere. Dennoch sehe ich hier Ähnlichkeiten zu den EU-Entwürfen.

Also nochmal: Die EU leistet gute Arbeit und ich möchte für sie eine Lanze brechen.

(neuer Beitrag)

Das Beispiel mit der automatischen Passvergabe durch KI leuchtet mir sehr ein – die KI-gesteuerte Passvergabe bietet mir als Einzelperson einen ganz konkreten Vorteil: Ich muss in Ländern mit viel Korruption weniger Geld ausgeben (durch den Wegfall von Bestechungsgeldern), aber auch in Ländern mit weniger Korruption läuft die Passvergabe schneller ab, weil die KI so effizient arbeitet.

Es gibt allerdings Situationen, in denen eine KI Vorschläge macht, die für die Gesellschaft zwar gut sind, welche Einzelpersonen jedoch in der gefühlten Persönlichkeit einschränken – ein Beispiel wären Handlungsempfehlungen zum Thema Klimawandel. In solchen Situationen würden wahrscheinlich sehr viele Menschen aus persönlichen Gründen widersprechen.

Das lässt sich auch im Bereich des Waffenhandels beobachten: Feinderkennung ist ja letztendlich auch KI, und hier stellt sich die gleiche Frage: Auch ein Waffensystem soll ja kontrolliert werden und nach ethischen Gesichtspunkten handeln. Ja, nur sind Ethik und Moral eben zeitgeistabhängig.

Drastisch formuliert und mit Blick auf die EU: Vom Mittelalter bis in das 19. Jahrhundert war es ethisch und moralisch vertretbar, Hexen zu verbrennen – hätte man damals eine KI programmiert, hätte sie das auch mit Sicherheit so abgebildet. Heute ist es innerhalb der EU moralisch vertretbar, Leute im Mittelmeer ertrinken zu lassen.

Auf welcher Basis schule ich eine KI, Ergebnisse zu produzieren, welche die Gesellschaft schützen sollen? Das kann sicher nicht der aktuelle moralische Kompass sein – er zeigt ja in die falsche Richtung.

(neuer Beitrag)

Noch einmal zum Beruf AI Prompter: Interaktives Fragen und das Gehörte Hinterfragen sollte doch zu jeder Berufsausbildung dazugehören – wo wird das in Zukunft gelehrt und wo wird das von Anfang an in die Bildung aufgenommen?

Das Zweite: ich glaube, wir vermischen gerade Expertensysteme und künstliche Intelligenz. Eine Erklärbarkeit und Nachvollziehbarkeit ist bei Expertensystemen sehr gut gegeben; diese Systeme lernen nicht selbst, die Entscheidungen sind einzeln nachvollziehbar.

Ist nicht die KI vielmehr dadurch gekennzeichnet, dass ich eben nicht mehr weiß, wie die Ergebnisse zustande kommen und dadurch eben im Umkehrschluss interaktiv nachfragen muss, um Konsequenzen auszuloten?

(neuer Beitrag)

Ich glaube, da haben wir einen sehr wichtigen Punkt angeschnitten: Herr Rebhorn, wo unterscheiden sich diese Expertensysteme, also klassische IT-Systeme, von nicht mehr kontrollierbaren KI-Systemen?

Im Beispiel der Passvergabe ging es nur darum, menschliches Versagen zu demonstrieren; hier geht es ja nicht einmal um ein Expertensystem, sondern nur um eine Prozessunterstützung durch IT.

Dabei wird allerdings eine weitere Frage aufgeworfen: Wie weit bin ich zu Verzicht bereit? Meine Logik sagt mir, dass eine Zukunft ohne Verzicht nicht funktioniert.

Bloß: Wie soll das genau funktionieren? Durch einen Generalverzicht, zum Beispiel auf Rindfleischkonsum? Oder führen wir vielleicht eine Sekundärwährung ein, eine Art Freiheitseinheit, mit der ich im Prinzip Freiheit personalisiere – so dass nicht jeder jede Freiheit nutzen darf und man eine Freiheit (wie beispielsweise Reisen) mit anderen (wie beispielsweise der Verzehr von Fleisch) aufrechnet.

Aus meiner Sicht ist eine neue Definition von Freiheit unbedingt notwendig – ob mit KI oder ohne KI. Durch eine KI kann allerdings dann ein jeweiliges neues Konzept von Freiheit besser gesteuert werden.

Ich habe viel im Bereich der Psychologie und auch mit „alten“ KI-Forschern gearbeitet, die noch mit Elektronengehirnen zu tun hatten und das analoge Denken auf Computer übertragen wollten – das heißt, KI hat früher einmal damit angefangen, dass man die KI-Systeme dem Menschen angepasst hat – nicht umgekehrt. Das heißt, früher gab es den Analogcomputer, und jetzt gilt schon das sogenannte digitale Denken wieder als veraltet. Die Frage wäre jetzt, ob nicht diese ganzen sozialen Vorschläge eine Folge des digitalen Denkens sind – das digitale Denken setzt nämlich keine Relation zwischen dem Denken eines Computers und dem eines Menschen.

Wenn man einen historischen Blick anwendet und nach den Zeitpunkten sucht, zu welchen es einen Wandel von einem zum nächsten Gesellschaftssystem gab: Es gab ja früher einmal den Futurismus. Man hat damals versucht, einen neuen Menschen zu schaffen, der sich an die industrialisierte Gesellschaft in der Sowjetunion anpasst; und auch in Italien hat man versucht, den Menschen neu anzupassen. Eine Folge davon ist die repräsentative Demokratie – sie zeigt uns Entwicklungen und dass aus den Fehlern der Vergangenheit gelernt wurde.

Die konkrete Frage ist, ob statt einem „entweder... oder“, also die Sicht auf den Menschen als ein Subjekt, das man auf die KI überträgt, es nicht vielmehr um die Relation zwischen analogem und digitalem Denken gehen müsste.

Ich habe vor vielen Jahren einen sehr spannenden Artikel über amerikanische Forscher gelesen, die einen Zusammenhang zwischen Quantencomputern und den quantenmechanischen Vorgängen in Nervenzellen gesetzt haben. Sie haben die steile These aufgestellt, dass erst die quantenmechanischen Vorgänge in den Zellen für das Bewusstsein bei Menschen sorgen.

Diese These könnte ja in gewisser Weise dazu führen, die digitale und die analoge Welt zusammenzuführen.

Am Ende ist natürlich heute jede Entwicklung immer eine Synthese aus verschiedenen unterschiedlichen Entwicklungen. Ich glaube, dass der Mensch sich selber anpassen wird – wir

brauchen dazu keine zentrale Instanz. Wir sehen ja schon alle Effekte der Digitalisierung; ich befürchte nur, wenn man sie nicht kontrolliert, wird es negative Konsequenzen geben –wie in einem großen Versuchslabor, in dem man nicht weiß, was herauskommt, weil man eigentlich kein Ziel hat.

Herr Rebhorn, Sie haben die individuellen Freiheiten angesprochen: Das läuft dann natürlich früher oder später auf das System „Lebensmittelkarte“ hinaus – man bekommt eine bestimmte Menge CO2 zur Verfügung gestellt, bis der Planet stirbt; das kann man dann beim Fahren, Essen oder Fliegen verbrauchen.

Was ich aber wirklich sagen wollte: Wir können natürlich im zivilen Sektor regulieren und uns bemühen. Man sollte allerdings bedenken, wie viele technische Fortschritte aus dem militärischen Bereich kamen (wie zum Beispiel die ganze Raumfahrt): Das Problem ist dann, dass der militärische Bereich keine Regulierung erlaubt. Man sieht ja, dass dort alles möglich ist und es keine Moral oder Ethik gibt.

Die Frage ist eben: Wenn die Entwicklung dann auf dem militärischen Sektor stattfindet, dann geht sie natürlich an sämtlichen Regulierungen auf der Zivilebene vorbei.

Ja. Und wenn beispielsweise Südafrika sich entscheidet, gar nicht zu regulieren, dann wird ein Großteil der Forschungsgelder auch nach Südafrika gehen. Ich stimme vollkommen zu. Wir sollten auf jeden Fall über solche Lücken sprechen – aber der Blick auf diese Lücken sollte uns nicht davon abhalten, für den Teil, den wir tatsächlich regulieren können, Lösungen zu finden.

Eine Metapher aus „Alice im Wunderland“: Als Alice an einer Weggabelung steht, fragt sie die Grinsekatze: „Welchen Weg soll ich gehen?“ und die Grinsekatze antwortet: „Wo möchtest du denn hin?“ Alice sagt: „Das ist mir egal“ und die Grinsekatze antwortet daraufhin: „Dann ist es auch egal, welchen Weg du gehst“.

Wir brauchen also eine Zielsetzung – denn wenn wir nirgendwo hinwollen, dann werden wir auch nirgendwo landen. Und so, wie es mir scheint, fehlt eben diese Zielsetzung bei den Auseinandersetzungen der EU mit dem Thema der KI in den letzten Jahren.

In diesen ganzen Auseinandersetzungen ist das Ziel ja, den Untergang der Menschheit zu verhindern. Dafür gibt es natürlich verschiedene Ansätze und wir stehen im Moment sozusagen vor einem Scheideweg: Wir haben das Dilemma, dass wir einerseits die romantische Idee haben, die Menschheit zu bewahren (also mit einer Vorstellung des Menschen als einzigartig und gut) – und auf der anderen Seite haben wir KI, welche die Menschheit bedroht.

Im Moment haben wir das Paradigma, dass wir die intelligenteste Spezies auf dem Planeten sind und deshalb bestimmen, was auf dem Planeten passiert; auch heute benutzen wir noch die Formulierung, dass der Mensch „im Mittelpunkt“ steht.

Noch einmal zu dem Spannungsverhältnis zwischen Mensch und Maschine: Entweder sagen wir, dass wir menschlich bleiben (das ist heute schon nicht mehr ganz der Fall, weil wir einzelne Körperteile ersetzen und sich das Bild „Mensch“ schon verändert hat) – in dem Fall würde uns die KI „davonlaufen“ und würde zur intelligentesten Spezies der Welt werden und uns wahrscheinlich wie unbedeutende Ameisen behandeln.

Oder aber es wird eine Konvergenz zwischen Mensch und Maschine geben.

Außerdem noch zur Vorstellung, dass der Mensch „im Mittelpunkt“ steht: Ich versuche, den Begriff Soziogenten zu prägen – gemeint sind alle autark regierenden Wesen auf der Erde. Wir sprechen heute von Menschenwürde – vielleicht sollten wir eher von Soziogentwürde sprechen und darunter dann auch Schweine oder Bäume und deren Würde verstehen. Das heißt, im Prinzip möchte ich den Menschen ein bisschen von seinem hohen Ross herunterholen und dafür die Umwelt und Tiere und Pflanzen berücksichtigen, um insgesamt zu einer gerechteren Welt zu gelangen: Das wäre eine Zukunftsvision.

(neuer Beitrag)

Zu der EU-Regulierung: Hier wurde die Risikopyramide recht clever gestaltet. Wenn man die Kriterien anschaut, werden sie immer noch dafür kritisiert, dass sie zu unscharf seien – aber das ist tatsächlich bewusst so gemacht. Denn wenn die Kriterien zu fein gestaltet werden, muss man sie dafür dauernd nachjustieren.

Im Prinzip gibt es natürlich KI, die ungesehen direkt zugelassen werden und dann wiederum solche, die unerwünscht ist (wie beispielsweise Gesichtserkennung im öffentlichen Raum mit Ausnahme der Polizei und des Militärs).

Beim Blick auf ChatGPT-4 oder folgende Versionen fällt vor allem die extreme Geschwindigkeit der Entwicklung auf. Im Moment begegnen wir einer KI, deren Ergebnisse man nicht nachvollziehen kann. In Kürze sind wir vielleicht von solchen KI-Helferlein (oder Überwachern) umgeben, die mit uns kooperieren, uns aber vielleicht auch überwachen. Und wenn diese plötzlich schlauer werden als Menschen und ihre Funktion ist, Menschen zu schützen, könnten sie auch prophylaktisch eingreifen, wenn sie bemerken, dass menschliche Handlungen schädliche Konsequenzen haben könnten.

Das könnte auch bedeuten, dass die KI vorschlägt, prophylaktisch in eine Wahl einzugreifen oder zum Wohle der gesamten Menschheit ein paar einzelne Personen umzubringen. Das sind Perspektiven, in denen wir – wenn wir nicht aufpassen – einfach die Kontrolle abgeben.

Wir sprechen dann von Maschinenwesen, von denen wir nicht wissen, wie sie funktionieren und können nur hoffen, dass sie auf Ethik (die Frage ist: Welche Ethik?) programmiert sind; es geht dann – auf dieser spekulativen Ebene – um Grundsatzfragen, die mit dem Überleben der Menschheit zu tun haben.

Wenn ich das Gefühl hätte, die Menschen würden bessere Entscheidungen treffen, würde ich gar nicht auf die Idee kommen, KI überhaupt solche Entscheidungen treffen zu lassen – aber die Entscheidung, andere Menschen umzubringen, treffen heute doch auch Menschen, und zwar nicht gerade wenige. Die Frage ist doch nicht, ob das passiert (denn es passiert sowieso), sondern: Was ist der richtige Weg zu einer Verbesserung?

Wenn solche Überlegungen ohne Zuhilfenahme von Rechnern und Maschinen jeglicher Art funktionieren, wäre das hervorragend – ich glaube nur nicht daran, dass das möglich ist. Deshalb

habe ich mir mit meinem Buch erlaubt, ein neues Gedankenmodell zu entwickeln, in welchem der Rechner objektive Entscheidungsfindungen übernimmt.

Das Thema ist wahrscheinlich deshalb so abendfüllend, weil es ja schlussendlich um eine persönliche Meinung geht – und ich habe mit meinem Buch ja auch keine wissenschaftliche Abhandlung, sondern eine Utopie geschrieben. Das ist im Prinzip nur eine von vielen Zukunftsvisionen, wie es anders sein könnte als jetzt.