

09/25/2023

Science-Talk-250923

00:02:24 **Speaker 2**

Schönen guten Abend hier in der Aula der Wissenschaften im ehemaligen Jesuitentheater. Hinter Ihnen sieht man sogar noch den Bogen des Theaterraums. Ich darf Sie herzlich begrüßen im Namen des Bildungs-, Wissenschafts- und Forschungsministeriums zu einem Thema, das eigentlich gar nicht so neu ist, wie wir alle glauben. Den Begriff KI gibt es seit den 50er Jahren und wir werden heute hoffentlich der Idee ein bisschen näherkommen, ob wir es eher mit kindlicher Intelligenz zu tun haben oder wie andere meinen, kritische Intelligenz. Darauf erhoffe ich mir zumindest eine Antwort, die ich mit nach Hause nehmen kann zu meinem 15-jährigen Sohn und ich hoffe, unser Moderator Jakob Zirm, Ressortleiter des Wirtschaftsmagazins Wirtschaftsteils der Presse, wird uns gut durch den Abend geleiten. Schönen Abend Ihnen allen.

00:03:36 **Speaker 2**

Eine Maschine, die zum selbstständigen und intelligenten Leben erweckt wird. Diese Vorstellung fasziniert die Menschen schon seit Jahrhunderten. Zuerst in der Literatur, später dann im Film. Sei es nun beispielsweise Frankenstein's Monster im berühmten Roman von Mary Shelley, Hell 9000 in Stanley Kubrick's 2001 Odyssee im Weltall oder Skynet in der Terminator -Reihe. Heute ist künstliche Intelligenz aber keine Science Fiction mehr. Spätestens seitdem ChatGPT im Herbst des Vorjahres online gegangen ist, kann sich jeder Mensch mit Internetzugang weltweit von der künstlichen Intelligenz Fragen beantworten lassen, komplexe Rechenaufgaben lösen lassen oder einfach nur Gedichte schreiben. Oder man kann sich bei einem Bild bei einem der Konkurrenzprogramme wie Midjourney oder Dolly nach eigenen Vorgaben einen Van Gogh erstellen lassen. Doch welche Auswirkungen wird diese künstliche Intelligenz auf unsere Wirtschafts- und Arbeitswelt haben? Inwiefern wird unsere Gesellschaft sich dadurch verändern

00:04:40 **Speaker 2**

und wird sie beispielsweise auch so wie der einst die Dampfmaschine oder der Verbrennungsmotor einen Großteil der körperlichen Arbeit ersetzt hat, auch einen Teil unserer geistigen Arbeit ersetzen? Und wenn ja, welche Folgen wird das haben für die Gesellschaft, für die Menschen? Das ist das Thema des heutigen Science Talks, zu dem ich Sie herzlich begrüßen darf. Ich habe hier ein wahrlich echtes Expertenpodium bei mir, das ich Ihnen zuerst einmal vorstellen darf und ich darf gleich hier links neben mir beginnen mit Assistenzprofessor Dr. Clemens Heitzinger. Er ist Assistenzprofessor für Informatik an der TU Wien und dort auch Co-Direktor des Zentrums für künstliche Intelligenz und Machine Learning. Dr.

00:05:32 **Speaker 2**

Stefan Strauß, Wirtschaftsinformatiker und Senior Scientist am Institut für Technikfolgenabschätzung an der Österreichischen Akademie der Wissenschaften.

00:05:45 **Speaker 2**

Frau Dr. Eugenia Stambouliev, Medienwissenschaftlerin und Technikphilosophin.

00:05:55 **Speaker 2**

Und auf der anderen Seite des Podiums Universitätsprofessor Dr. Matthias Kettemann, Professor für Innovation, Theorie und Philosophie des Rechts und Leiter des Instituts für Theorie und Zukunft des Rechts an der Universität Innsbruck.

00:06:14 **Speaker 2**

Sie haben natürlich auch, verehrtes Publikum, die Möglichkeit, nachher Fragen zu stellen, einerseits hier im Raum, aber auch alle, die zu Hause im Stream sind, haben natürlich auch die Möglichkeit, beim Stream die Fragen zu stellen. Ich werde sie im weiteren Verlauf der Veranstaltung dann hier vorlesen und die Fragen dann auch stellen ans Podium. Aber zu Beginn habe ich einmal ein paar Fragen und ich würde einmal beginnen gern mit so einer Art Einordnung, ein bisschen eine technische Einordnung. Wir haben jetzt viel gesprochen, künstliche Intelligenz für Kernbegriffe, Algorithmen etc. Und ich würde gleich bei Ihnen, Herr Heitzinger, anfangen gerne, was ist KI eigentlich oder künstliche Intelligenz? Wir wissen ja zum Beispiel, die Kugelsuche basiert auch auf Algorithmen, das ist ein Wort, das in den letzten Jahren sehr populär geworden ist. Wo ist, kann man hier technisch sagen, wo ist der Unterschied zwischen einem einfachen Algorithmus und einer künstlichen Intelligenz? Ja, das ist eine

00:07:01 **Speaker 2**

sehr gute Frage und die Abgrenzung ist sicher nicht so einfach im Vergleich zu anderen Gebieten. Der Begriff KI geht zurück auf die 1950er -Jahre eigentlich und damals hat man begonnen, ausgehend von der Entwicklung der ersten Computer, sich die Frage zu stellen, können wir menschliche Fähigkeiten, menschliche Intelligenz genau verstehen und so genau verstehen, dass wir sie mit einem Computer nachbilden können, eben programmieren können und der Begriff geht zurück auf einen Workshop im Jahr 1956, der eigentlich richtungsweisend war, weil das Programm, das man, das wissenschaftliche Programm, das man dort schon zusammengestellt hat, hat eigentlich alles abgedeckt, was man unter Intelligenz verstehen kann, nämlich Perzeption, also Bilderkennung, künstliche neuronale Netze sind damals schon angesprochen worden und waren angedacht, obwohl

man nicht wusste, wie man die trainieren kann, aber die Ideen waren schon da, Kreativität hat damals schon eine Rolle gespielt und ist als Frage angedacht

00:08:05 **Speaker 2**

worden, also so ziemlich alle Aspekte der Intelligenz sind damals in diesem Forschungsprogramm eigentlich schon erwähnt worden. Es hat jetzt 70 Jahre gedauert, bis wir in vielen Bereichen soweit sind, dass diese Programme mit künstlicher Intelligenz, die eigentlich künstliche Intelligenz haben, dass die mit menschlicher Intelligenz konkurrieren können oder sie teilweise auch übertreffen. Also Algorithmus, um das ein bisschen zu definieren, Algorithmus für sich genommen bedeutet jetzt einfach nur eine Rechenvorschrift, die man stupide befolgen kann, also wie ein Kochrezept werden einfach diese Befehle ausgeführt und dadurch ein Output erzeugt und dadurch ist das eindeutig definiert. Also ein Algorithmus ist mit jedem Computer verbunden eigentlich. Das, was die KI ausmacht, heutzutage sind vielfach Methoden des maschinellen Lernens, natürlich auch Methoden der Logik, also das KI -Feld ist methodisch sehr, sehr vielfältig. Also in den letzten 70 Jahren hat man viele Ideen entwickelt, programmiert,

00:09:14 **Speaker 2**

teilweise auch wieder verworfen und manche weiterentwickelt.

00:09:19 **Speaker 2**

Chatship -ID zum Beispiel würde man einordnen eigentlich als Programm des maschinellen Lernens und da ist die Idee, dass man Datensätze hat und mit Hilfe dieser Datensätze Probleme löst und die Datensätze sind am besten so groß, dass sie alle möglichen Fragestellungen abdecken und die Lösungen, die dann kommen, das ist dann natürlich die Frage, wie man das definiert, ob die kreativ sind oder nicht, also Chatship -ID kann man sich eine gewisse Kreativität nicht absprechen, aber im Prinzip beruht das darauf, dass man enorme Datenmengen hat und daraus das Essentielle quasi extrahieren will, das in eine neue Form gießt und diese neue Form kann dann Fragen beantworten, das ist so die Idee dahinter. Ich glaube, der Grund, warum wir heute hier sitzen, ist ja der, dass wir seit Anfang des Jahres dieses große mediale Interesse gesehen haben, zu Recht auf jeden Fall an Chatship -ID, weil es in der Geschichte der KI das erste System ist, das eigentlich wirklich gut mit menschlicher Sprache, mit natürlicher

00:10:24 **Speaker 2**

Sprache umgehen kann. Das war bis dahin, bis Ende letzten Jahres ein ungelöstes Problem eigentlich in der Geschichte der KI oder in der Geschichte der Menschheit, wenn Sie so wollen. Warum

funktioniert das? Es kommen jetzt drei Aspekte zusammen, wir haben riesige Datenmengen, also Chatship -ID hat ganze Lehrbuchsammlungen gelesen, also die besten Texte, die es gibt, also irrsinnig viel Wissen angelernt eigentlich, also das ist wirklich enorm aufwendig gewesen. Wir haben enorme Rechenkapazitäten zur Verfügung, in den letzten Jahren eben durch Grafikkarten ist das möglich geworden eigentlich erst und das Dritte ist, wir haben doch algorithmische Weiterentwicklungen gesehen, vor allem getrieben durch die Arbeit von Google Brain, Google Research, DeepMind und jetzt eben durch OpenAI. Also so wie man diese Systeme heutzutage programmiert, so hätte man das vor fünf Jahren und vor zehn Jahren schon gar nicht gemacht. Also die drei Faktoren sind jetzt zusammengekommen und OpenAI hat es eben geschafft,

00:11:31 **Speaker 2**

durch viel Aufwand dieses Problem zu lösen, also das muss man wirklich neidlos anerkennen. Das ist sicher ein Meilenstein und eine Revolution in der KI. Man hatte vorher nicht gewusst, dass man auf diesem Weg, dass das so gut funktionieren wird und das kann man auf jeden Fall sagen. Vielleicht ganz kurz noch, wofür Chatship -ID steht, weil das ist eigentlich der Grund, warum wir heute da sind, man kann das ganz kurz erklären, wofür die Abkürzung auch steht. GPD steht für Generative Pre - trained Transformer. Transformer sind eine gewisse Art von normalen Netz, die entwickelt wurden fürs maschinelle Übersetzen eigentlich und ironischerweise von Google, nicht von OpenAI im Jahr 2017. Das waren damals die besten Methoden zum Übersetzen von Texten. Pre -trained steht dafür, dass es eben an riesigen Textmengen angelernt wurde, angeblich an Lehrbuchsammlungen, also wir als Wissenschaftler haben die Trainingsdaten geliefert und viele andere große Datensätze auch, so genau weiß man das natürlich

00:12:42 **Speaker 2**

nicht. Generative steht eben dafür, dass es dafür eingesetzt wird, Texte zu generieren, nicht zu übersetzen und ich glaube, jeder hat es schon einmal verwendet, man sieht ja auch, dass der Text quasi Silbe für Silbe erzeugt wird, das heißt, das Programm sieht die letzten Silben, die letzten paar tausend Silben in dem Dialog und dann wird die nächste Silbe erzeugt und so weiter und so fort. Also diese Texte werden Silbe für Silbe generiert und so ein GPD selber, das war schon in den Fachkreisen aufsehenderregend vor einigen Jahren, aber das GPD selber ist jetzt nicht besonders nützlich, es generiert zwar schöne Texte, man kann denen Input geben und es generiert dann Texte, die aber mit dem Prompt nichts zu tun haben müssen und der große Fortschritt war dann ganz kurz noch, dass das Chat in ChatGPT bedeutet, dass das wirklich jetzt ein System ist, das als Chatbot sehr gut funktioniert und darauf eingeht, was die Anwenderin oder der Anwender haben will und ich hoffe, das beantwortet Ihre Frage,

00:13:46 **Speaker 2**

was der Unterschied ist. Wirklich sehr viel gelernt, muss ich sagen, sehr viel gelernt in Ihrer Antwort, danke mal für diese grundsätzliche Definition des Unterschiedes, Herr Strauß, ich würde gerne gleich bei Ihnen dagegen, ist das jetzt aus Ihrer Sicht, der Herr Heitzinger hat es jetzt eh ganz gut definiert, wo sozusagen der Unterschied zwischen den stupiden Algorithmen und dem selbstlernenden, referenzierenden System ist, aber ist das jetzt eigentlich ein komplett neues Feld der Softwareentwicklung oder ist das so die logische Weiterentwicklung deswegen? Ist das ein fließender Übergang? Fließend ist immer schwierig, natürlich ist das alles irgendwie fließend, klar, man kann auch sagen, gleichzeitig wird es gefeiert als disruptive Revolution in irgendeiner Form, dem würde ich auch nicht widersprechen wollen. Gleichzeitig, wenn man anschaut, woher diese Modelle kommen, Sie haben das Perceptron, glaube ich, vorher kurz erwähnt, oder Sie haben es darauf verwiesen, das ist ja aus den 50er

00:14:39 **Speaker 2**

-Jahren auch, jetzt gibt es teilweise noch KI -Modelle, die verwenden das immer noch, da könnte man auch einmal diskutieren, ob das noch zeitgemäß ist, nicht im Sinne, dass es sozusagen veraltet ist, sondern im Sinne, ob es das noch abbildet, oder ob es die Funktion noch erfüllt ist, die es damals hatte, oder ob es möglicherweise die Welt auch komplexer geworden ist und wir sozusagen auch eine Weiterentwicklung von den Parametern brauchen, die wir heute haben, sozusagen, ChatGPT verwendet das, glaube ich, nicht mehr, weiß ich nicht genau, das können wir vielleicht später noch diskutieren, aber worauf ich hinaus will, Chatbots gibt es auch schon seit, ich glaube, der erste Chatbot, der so ein bisschen medial kursiert ist, das war um die Jahre 2000 herum, der CleverBot, das war eine wissenschaftliche Entwicklung, ich habe mit dem eine Zeit lang gern gespielt, weil es ganz interessant war, aber natürlich völlig trivial im Vergleich zu dem, was wir da heute sehen, weil diese großen Sprachmodelle,

00:15:36 **Speaker 2**

die wir haben, und ich weiß nicht, wie viele Millionen Parameter da drinnen sind, ist natürlich ein Riesenqualitätssprung, den man hier unbestritten sieht. Gleichzeitig muss man aber sagen, wir haben auch das Phänomen der AI Illusions, also die Illusionen, das heißt, Sie kennen das alle aus der Zeitung, Sie geben in JGPT was rein und dann kommt da was völlig Plausibles raus, nur wenn Sie es genauer lesen, da sind einfach erfundene Quellen drinnen, es gibt diesen schönen Fall vielleicht aus der Rechtsecke da, wo dann einfach Urteile erfunden wurden, und es gibt eine ganze Reihe von Beispielen, dass es eben nicht so gut funktioniert, wie man glauben mag. Dass das hochperformant

ist, dass es wirklich einen Qualitätssprung bedeutet, ist völlig unbestritten, glaube ich, aber wenn man ins Detail reingeht, muss man glaube ich sehr genau differenzieren, in welchem Anwendungskontext macht es Sinn, wie kommt es eigentlich zu diesen Fehlern, das ist glaube ich ein ganz wesentlicher Knackpunkt, auch

00:16:36 **Speaker 2**

in Bezug auf den Untertitel der Veranstaltung, Vertrauen, also wie vertrauensvoll ist dieses System oder kann es überhaupt sein, wenn da einfach frei erfundene Dinge auch drinnen stehen und man kann sich auch gar nicht eruieren, also es ist kein Label dabei, es gibt kein Watermarking oder was auch immer, es gäbe ja Möglichkeiten, das schon zu tun, aber ich greife jetzt vielleicht schon ein bisschen voraus, aber es ist eben nicht neu, aber wir erleben glaube ich gerade sowas wie eine neue Verhandlung dessen, was es bedeutet in einem konkreten Anwendungskontext, weil diese Anwendungsfälle, die wir heute haben, die sind tatsächlich neu, die hat es früher nicht gegeben und auch diese Breitenwirksamkeit, KI war immer eine Spezialanwendung in der wissenschaftlichen Community oder in der technischen Community, heute ist es eine breitenwirksame Mainstream - Technologie geworden, wir diskutieren ja schon darüber, ob es die nächste Generation von Suchmaschinen quasi so funktionieren wird, das heißt,

00:17:35 **Speaker 2**

wir haben eine ganz andere Bedeutung von dem, was wir unter Mensch -Maschine -Interaktion verstehen oder verstanden haben in den letzten 20 Jahren, würde ich sagen, da brauche ich noch gar nicht darüber reden, ob das eine haptische Veränderung hat, sondern da geht es ja schlicht einmal darum, die Art und Weise, wie man mit

00:17:55 **Speaker 2**

Informationsaufbereitung umgeht und was das eigentlich bedeutet, aus dem dann wiederum Wissen entsteht, wer dieses Wissen eigentlich schafft, früher gab es ja auch den Begriff, da gibt es immer noch die Anwendungsverarbeitung, die Wissensverarbeitung sozusagen, ein Teil der KI, da geht es stark eher darum, wie man sozusagen die Algorithmen dahinter auch designt, aber der Wissensbegriff ist auch ein ganz zentraler, würde ich sagen, von dem wir jetzt eigentlich noch relativ weit weg sind, weil ich würde jetzt, dem würde ich zum Beispiel nicht, das würde ich nicht unterschreiben, wenn man jetzt sagt, Chat -GPT oder ähnliche Technologien, die generieren Wissen in irgendeiner Form, sie replizieren vielleicht Wissen, aber sie generieren kein Wissen und sie sind auch nicht kreativ in dem Sinn, also es sind ja trotzdem Maschinen, man bräuchte, glaube ich,

auch ein bisschen eine Entzauberungsdiskussion darüber, was es letztlich bedeutet, ich bin schon fertig. Nein, nein, bitte, bitte, es soll

00:18:43 **Speaker 2**

ja eine Diskussion sein, aber ich würde diesen Ball eklig aufnehmen mit Entzauberungsdiskussion und Frau Stamboliev, ich werde Sie weitergeben, wie gesagt, es war natürlich ein großer Hype, es war ein großer, auch wir haben natürlich darüber geschrieben in unserer Zeitung und das auch mitverfolgt, aber jetzt ist die Frage, jetzt aus Sicht, wir wollen nachher dann eh noch über die Folgen, wie wir damit umgehen, das soll ja natürlich ein großer Teil der Diskussion sein, aber ganz kurz nur zur Einschätzung, worüber wir eigentlich sprechen, war das jetzt eigentlich wirklich so eine große Neuerung, auch medienteoretisch dann, also ist das wirklich ein Bruch, haben wir eine Disruption gesehen? Wie würden Sie es einschätzen, von der Bedeutung her? Sie meinen jetzt JetBox, also JetJuby? Jetzt die Neuerungen mit Journey, all diese KI -Anwendungen, die jetzt im letzten halben Jahr quasi auf den Markt gekommen sind, das sind ja relativ schnell hintereinander ganz viele gekommen. Also auf zwei Ebenen

00:19:36 **Speaker 2**

war es definitiv schon mal ein Bruch, weil einfach plötzlich die ganze Welt davon sprach, wenn es auch jetzt nur, also vielleicht nicht immer sich auf die Anwendung irgendwie gespiegelt hat oder dass jetzt nicht immer so viel angewendet wurde, wie vielleicht geschrieben, aber gleichzeitig wurde sehr viel angewendet, sehr viele Menschen haben sich sehr schnell diese Technologien, ich würde gerne sagen, zunutze gemacht, aber das stimmt eigentlich gar nicht, weil für mich jetzt als Philosophin zum Beispiel ist auch die große Fragestellung, verwenden wir diese Technologie oder verwendet sie uns? Und da muss man sich sehr genau überlegen, was man eigentlich macht und da war, muss ich sagen, OpenAI einfach genial als Unternehmen, wirklich strategisch unglaublich gut, weil sie eine Art Demokratisierung dieser Technologie simuliert haben. Sie haben sie frei zugänglich gemacht für alle, also die Chatbots und meinten hier, sie haben Zugang zu Wissen, das soll die Menschheit verbessern, die Ansprüche

00:20:30 **Speaker 2**

waren ja auch sehr groß, das soll Kreativität fördern, alles wird besser, Jobs fallen weg, die sowieso keiner will, nur die Guten bleiben übrig, wir werden die Gesellschaft verbessern und und und. Dazu kann ich jetzt generell gar nicht sagen, ja, nein, für manche hat es Vorteile, für einige Berufszweige, wir werden auch später noch darüber sprechen, das weiß ich, für andere nicht. Dahinter gibt es aber noch unglaublich viele Probleme. Ich weiß, ich habe genau als Chat 4, eins von beiden, rauskam

genau eine Lehrvorlesung, also begonnen mit Studierenden und es ging die ganze Zeit eigentlich nur darum, dürfen sie das jetzt verwenden in ihren Arbeiten oder nicht. Ich war relativ demokratisch und meinte, ja, vielleicht, aber macht was damit, weil ich jetzt nicht per se technophob bin. Ich wollte auch, dass Studierende wirklich lernen, damit umzugehen. Das Ergebnis war ein bisschen erschreckend, aber im Sinne von Bequemlichkeit ist wirklich etwas sehr, sehr, wie soll ich sagen, Mächtiges. Und

00:21:32 **Speaker 2**

sobald man einem Prozess von Anstrengung irgendwie, wenn es auch nur Lernen ist und so kürzer machen kann, werden Leute das verwenden und Menschen werden sich dessen bedienen. Insofern ist da ein großer Umbruch, glaube ich, jetzt zum Thema Wissensaneignung. Aber dann ist es denn überhaupt Wissensaneignung? Genau, generieren wir dadurch neues Wissen? Ich glaube, einige ja. Zum Beispiel, ich habe mit sehr vielen auch Künstler und Künstlerinnen inzwischen gesprochen, Architektinnen und so weiter. Es hat schon auch Potenzial für Menschen, die damit auch umgehen können, im Sinne von auch programmieren können. Für jetzt, wie soll ich sagen, die Mehrheit der Gesellschaft, die jetzt jede banale Frage irgendwie da einzutippen, von Kochrezepten zum Beziehungsstatus, weiß ich nicht, das möchte ich gar nicht urteilen. Da, glaube ich, wird dieser Effekt, diese Euphorie nachlassen. Ich glaube, das wird auf Dauer ein bisschen langweilig werden, genauso, wenn man mal zehn Texte gelesen hat oder wie ich

00:22:27 **Speaker 2**

einige hunderte, die generiert werden von solchen Bots, wird es langweilig. Also es kommt, die Struktur wird irgendwann so durchschaubar, dass ich immer noch hoffe, dass der Mensch eigentlich sich früher oder später auch ein bisschen langweilen wird und es vielleicht für Anwendungen verwendet, die Zeit sparen können. Da macht es schon viel Sinn für bestimmte stupidere Tätigkeiten, was auch immer das ist. Aber im Allgemeinen, ja, es war eine große Umwälzung. Ich weiß nicht, wie lange sie hält für den Allgemeingebrauch, bin ich echt gespannt, weil ja auch jetzt die Frage ist, werden mehr und mehr Falschinformationen da mitgefiltert, wird es quasi eine Art Spirale von kompletten Unsinn irgendwann geben und es ist nicht mehr brauchbar, das weiß man nicht, weil die Daten natürlich nicht mehr so gereinigt werden können, so gut, was übrigens auch von Menschen gemacht wird, nur so nebenbei, also so automatisiert sind diese Systeme nicht, es sind auch große Ausbeutersysteme und dahinter sind ganz

00:23:20 **Speaker 2**

oft auch Menschen im globalen Süden, die sehr billig dafür arbeiten, dass wir quasi richtige Antworten bekommen. Das darf man nie vergessen bei solchen Anwendungen, das heißt, die sind sehr, auf der einen Seite sehr unglaublich gut entwickelt und sehr hochkarätig und auf der anderen Seite fördern sie eigentlich sehr alte Strukturen genauso wie Ausbeutung. Herr Kettemann, die Frau Stamboliev hat eh auch schon ein bisschen diese Strukturen angesprochen, was ja natürlich bei dem ganzen Thema natürlich sehr präsent auch ist, es ist ein Riesengeschäft, es ist ein Zukunftsmarkt sozusagen, da fließen, also nur ein paar Zahlen, Microsoft hat 10 Milliarden in Open AI, den Hersteller von JGPT investiert, Google investiert in die eigene Software, und heute erst ist die Meldung gekommen, auch Amazon steigt ein mit 4 Milliarden bei einem Unternehmen, das erst im Jahr 21 von ehemaligen Mitarbeitern von Open AI gegründet worden ist. Das Unternehmen ist gerade einmal zwei Jahre alt und hat heute von Amazon

00:24:16 **Speaker 2**

4 Milliarden bekommen. Also man sieht, das ist ein wirklich großes Geschäft und die großen Tech - Konzerne sind da natürlich mit sehr viel Geld dahinter und erwarten sich da natürlich auch was und daher, das ist auch ein Teil sozusagen, und jetzt, Herr Kettemann, das würde mich auch an Sie interessieren, ist jetzt quasi sozusagen, man redet viel über JGPT, das ist sozusagen der Leuchtturm, weil halt jeder damit spielen kann, aber es gibt ja ganz viele andere Anwendungen auch, die halt im Hintergrund verschiedenste Sachen verändern werden. Wie würden Sie es einschätzen? Ist aus Ihrer Sicht da jetzt, die Veränderung ist da so ein Disruptiver -Moment? Wie würden Sie es grundsätzlich einschätzen? Ja, auf jeden Fall. Wir haben ja jetzt einen Moment, wo die Hochtechnologie im Wohnzimmer angekommen ist, wo wir auf unseren Handys die Möglichkeit haben, Bilder zu generieren, Inhalte zu generieren, die wir dann sofort nutzen können oder könnten, weil, wie Kollegin Richter gesagt hat, in der Praxis

00:25:14 **Speaker 2**

verwenden die meisten Menschen die Inhalte doch eher nur in beschränktem Ausmaße, wenn sie nicht gerade Seminararbeiten abgeben müssen. Wobei man sagen muss, dass gerade bei der Zusammenfassung von Artikeln, bei der Synthetisierung von bestehenden Informationen hier durchaus einiges an Mehrwerten generiert werden kann. Ich forsche viel zur Innovation und bei der Innovation hat man sehr häufig ein Problem, nämlich, ob etwas wirklich innovativ ist oder nur innovativ wirkt. Und wenn es wirklich innovativ ist, wie man darauf dann reagieren soll, mit Recht, mit der Gesellschaft, was ist dann zu tun? Gibt es hier ein Problem für individuelle Rechte, für gesellschaftliche Werte, für den Zusammenhalt? Und wenn es solche Probleme gibt, dann muss darauf mit Recht reagiert werden. Ich bin mir nicht ganz sicher, ob wir hier mit einer Innovation der

Gestalt zu tun haben, die jetzt schon scharfe neue Regeln verlangt. Es kann durchaus sein, dass wir ein bisschen warten sollten, wie sich die GBT -Generation

00:26:21 **Speaker 2**

entwickelt, wie die Technologie, jetzt wo der erste Hype bald einmal abgeflaut sein wird, ob wir hier wirklich fundamentale Umwälzungen sehen. Wir können gern später noch darüber sprechen, aber ich denke auch nicht, dass wir hier den Beginn des Endes der Menschheit sehen. Uns wird es schon noch ein bisschen geben. Wir dürfen nicht vergessen, diejenigen, die hier an, die hier sagen, dass der GBT und Kollegen uns dann in näherer Zukunft auslöschen werden, die haben oft andere Interessen, nämlich zum Beispiel, dass die Unternehmen, die das noch nicht haben, schärfer reguliert werden sollen. Weil Innovation, der, der sie hat, der möchte nicht, dass andere danach ziehen. Auf das wollen wir eh noch, dieses berühmte Moratorium, auf das möchte ich später eh noch zu sprechen kommen. Aber ganz kurz noch zur Einschätzung, nochmal ans Podium. Ich habe ein bisschen nachgedacht, weil Sie auch gesagt haben, fundamentale Umwälzungen. Und wenn wir zurückdenken, das iPhone wurde von Apple im Jahr 2007 vorgestellt,

00:27:19 **Speaker 2**

das ist jetzt gerade einmal 16 Jahre her. Und wenn man sich anschaut, wie sich das Leben verändert hat, viele Dinge, es gibt zum Beispiel jüngste Studien, dass die Menschen ihren Orientierungssinn zu verlieren fangen langsam, weil einfach man die Gewöhnung an Google Maps hat. Und man braucht keine Karten mehr lesen, man braucht sich nicht mehr orientieren. Das heißt, es werden über Jahrhunderte erlernte Fähigkeiten, bilden sich mitunter zurück. Und die Frage wäre ein bisschen so, Herr Strauß würde mich interessieren, kann, bei KI wird ja das auch nachgesagt, dass hier vielleicht auch so fundamentale Umwälzungen sind. Wenn wir jetzt ein bisschen in die Zukunft blicken, jetzt wissen wir, was geht heute, wovon reden wir da in 10, 15 Jahren, wenn wir so 15 Jahre in die Zukunft blicken? Ja, wir könnten das GQBT befragen, was es dazu sagt, natürlich. Könnten wir machen, ja.

00:28:06 **Speaker 2**

Prognosen sind immer schwierig, aber wenn man davon ausgeht, dass sozusagen, was ich vorher gesagt habe, dass diese Mensch -Maschine -Interaktion sich, sagen wir mal, langfristig verändern wird, weil wir jetzt ein Sprachinterface haben, das funktioniert, also ich glaube, da können wir uns drauf einigen, würde ich sagen. Dann würde ich schon sagen, dass es auch einen Unterschied macht, wie die Art und Weise, wie man tatsächlich mit solchen Interfaces, wie man die verwendet, sich verändert. Und dadurch verändert sich dann im Endeffekt auch irgendwie die Sprachlogik. Das ist

jetzt natürlich relativ schwierig festzumachen, inwiefern, aber das steht und fällt auch ein bisschen damit, wie funktionsfähig die Systeme dann letztendlich sind. Ich glaube, was wir sehen werden, ist erstmal ein Lernprozess, nämlich dass wir lernen werden oder müssen im Endeffekt besser zu verstehen, wie diese Systeme nicht funktionieren. Wie sie funktionieren, verstehen wir sowieso nicht. Ich glaube, das wird sich auch

00:29:01 **Speaker 2**

nicht ändern, aber ich glaube, es gibt Mittel und Wege zu lernen, wie sie nicht funktionieren. Da gibt es ja diese Prompt Injections, das ist jetzt ein bisschen aus der Sicherheitsforschung, das heißt, Sie können sozusagen auch JGBT, auch wenn Schutzmechanismen implementiert sind, soweit mit Prompts konfrontieren, dass Sie da Dinge ausspucken, die Sie eigentlich nicht ausspucken sollen. Das geht so weit, dass Sie da irgendwie sicherheitsrelevante Bombenanleitungen oder sowas entlogen. Da gibt es auch Studien dazu, wie man das sozusagen hinbekommt. Da gibt es eigene Forschungszweige mittlerweile in der Sicherheitsforschung. Das ist so ein Extrembeispiel jetzt ein bisschen. Und für den Alltagsgebrauch, glaube ich, auch ein bisschen, was die Kollegin vorher gesagt hat, es ist ein bisschen endenwollend, dieser Hype. Wenn Sie sich die Beispiele ansehen, wie das auch beworben wird, das ist ja schon sehr stark getriebener Hype. Das hat ja einen Grund, warum so viel Geld investiert wurde von OpenAI,

00:29:48 **Speaker 2**

nicht von Microsoft. Das ist ein klarer Technologie -Push. Da geht es eigentlich nicht um JGBT, sondern da geht es darum, Schnittstellen zu verkaufen, da geht es darum, Machine Learning Modelle auf den Markt zu bringen und dann zu stimulieren, dass ein ganz neuer Markt entsteht. Das ist eigentlich ein Grund, warum das so gehypt wird. Und da wird es dann spannend, wie sich das dann wirklich weiterentwickeln wird. Weil jetzt haben wir die Diskussion, FAU hat eine kurze Regulierung stand im Raum, also wie regulieren wir das jetzt eigentlich? Und da ist eine Frage auch, ja, es gibt sozusagen Konkurrenz zwischen dem, was da schon am Markt ist und denjenigen, die das reinschieben in den Markt. Und dann gibt es Entwicklungen sozusagen im kleineren Rahmen, die auch Trainingsdaten verwenden, die eben nicht vom Global South, von Clickworkern genutzt oder bereinigt werden müssen, sondern die sozusagen ein relativ hohes Datenniveau, Qualitätsniveau haben. Die haben wiederum das Problem, dass sie nicht

00:30:44 **Speaker 2**

marktfähig sind im Sinn, weil sie eben nicht das Geld haben, dass sie da reingeschoben werden können. Und ich glaube, das wird auch, das ist jetzt keine menschliche Fähigkeit, die sich verändert,

das ist, glaube ich, eine wirtschaftliche Logik, die sich verändern wird, dass es sozusagen eine Marktverschiebung gibt. Weil in irgendeiner Form, glaube ich, muss man, wenn man es ernst meint mit europäischer Digitalisierung, europäischer KI und so weiter und vor allem auch Forschungsexzellenz in dem Bereich, da muss man, glaube ich, auch in die Richtung gehen, dass man das in irgendeiner Form so gestaltet, dass es auch gestaltbar ist von sozusagen einem Markt, der nicht nur hinterher rennt von dem, was Tesla, was X und so weiter, die ganzen Firmenmieter, Google, Elphabet und so weiter tun, sondern da muss man, glaube ich, irgendwie Mechanismen finden. Das ist ja wirklich ein bisschen voraus vorweggegriffen, aber ich weiß nicht, ob das die Frage jetzt beantwortet hat, jetzt bin ich ein bisschen

00:31:40 **Speaker 2**

abgebogen. Ja, nein, es war schon, also Sie haben gesagt, Prognosen sind schwierig, es ist schwierig, dass es in 10, 15 Jahren ist, konnte man sich vor 15 Jahren beim Smartphone damals auch nicht vorstellen, die Anwendungsformen, oder gibt es? Naja, das mit den Anwendungsformen ist auch so eine Frage, also in welchem Anwendungskontext bewegen wir uns? Reden wir davon, dass es sozusagen ein massentaugliches Instrumentarium geworden ist? Das würde ich jetzt einmal so sehen. Gleichzeitig sehe ich immer noch, für mich ist eine völlig offene Frage, wo ist der Use Case, wo ist der Anwendungsfall, wo sozusagen der Standardbenutzer, die Benutzerin sagt, das Ding nutze ich jetzt jeden Tag, weil es so toll ist, so wie Google -Suche oder sowas, weil das wurde sozusagen zum Synonym für Suchen. Da sind wir, glaube ich, weit entfernt davon, weil ein Kochrezept wird nicht die Lösung sein, und der Unterschied zur normalen Suche ist auch, die sozusagen die Killeranwendung wurde, ist auch, dass die Anfrage

00:32:31 **Speaker 2**

-Logik viel zu komplex ist. Also natürlich, Sie können auch ein Wort reingeben, aber da ist die Antwort, die Sie da bekommen, wesentlich, also nicht so viel Unterschied, wie was Sie von Google bekommen, und wenn Sie komplexere Fragestellungen haben, brauchen Sie auch ein Wissen dafür, was Sie da eigentlich fragen wollen. Deswegen würde ich auch sehen, dass dieser Wissens - und Bildungsaspekt eine ganz zentrale Rolle spielt letztendlich. Herr Heitzinger, ich würde auch gerne an Sie die Frage so ein bisschen weitergeben. Gäbe es sozusagen Use Cases, die Ihnen schon, also als Experte, die Sie sehen am Horizont, dass die kommen werden? Das wäre das eine, und dann vielleicht noch eine Frage noch, das klingt ein bisschen technisch, das würde mich interessieren. Es gibt immer dieses Thema der technologischen Singularität, also vereinfacht gesagt, irgendwann einmal verstehen wir Menschen nicht mehr, wie die Maschinen selber weiterlernen. Ist das ein Thema, dass wir das haben, oder haben wir das

00:33:21 **Speaker 2**

schon? Zur ersten Frage vielleicht zuerst. Da kann ich ein bisschen berichten, was wir so an Anfragen sehen, auch von österreichischen Firmen. Also, was wir schon wissen, ohnehin schon wussten, ist, dass Microsoft jetzt GPT-Versionen in die Produkte einbaut. Das heißt, das funktioniert jetzt schon oder in Kürze, dass man sagt, okay, wenn ich eine lange E-Mail bekomme, fassen wir das zusammen, oder ich gebe ein paar Stichworte ein und lasse mir dann eine lange E-Mail generieren. Also, das ist damit kein Problem mehr. Also, da wird sicher viel eingebaut werden von allen großen Herstellern, also Apple hat auch angekündigt, und wir gehen jetzt halt von einer Ära der Suchmaschinen in eine Ära der Erklärmaschinen. Damit ist klar, also, was man auch, glaube ich, sagen kann, ist, dass so Systeme wie ChatGPT, diese Large-Language-Models, auch deswegen, weil sie so viel gelesen haben, quasi einen Hausverstand mitbringen. Also, sie wissen, wie die Welt funktioniert, in einem gewissen Sinn,

00:34:26 **Speaker 2**

und man ist jetzt auf der Suche nach Anwendungen. Wo kann ich das einsetzen? Und das kann jetzt sein, sozusagen extern, wo man jetzt auf die Website einen Chatbot geben kann, der jetzt tatsächlich funktioniert, nicht so, wie es früher war, also für Customer-Service-Anwendungen, und da glaube ich, dass sich der Markt halt irgendwie aufteilen wird. Also, wenn ich mehr zahle, dann bekomme ich halt Customer-Service von einem menschlichen Mitarbeiter, und in der Billigversion werde ich halt an die KI verwiesen.

00:34:56 **Speaker 2**

Dann gibt es natürlich andere Anwendungen innerhalb der Firma, die ja auch oft viele Dokumente haben. Die Systeme sind auch nützlich beim Programmieren, um Templates zu erstellen zum Beispiel, also quasi Templates, aber dort, wo es viel Source-Code gibt, dort funktioniert das sehr gut, weil es daran angelernt wurde. Und da ist halt die Frage jetzt, manche Firmen sind da konservativer und verbieten den Einsatz, andere Firmen sind da offener, und die Konservativen haben wahrscheinlich auch ein bisschen Angst davor, dass sie dann Gegenstand von Klagen werden. Welche Trainingsdaten sind da wirklich verwendet worden? Und da gibt es jetzt eh fast täglich Meldungen, dass Klagen eingebracht werden in den USA. Ich habe nie dazugestimmt, dass meine Daten dafür verwendet werden oder meine Bücher dafür verwendet werden, ich auch nicht, um diese Systeme anzutrainieren.

00:35:50 **Speaker 2**

Also da gibt es diese zwei Denkschulen, sage ich einmal, in den Firmen, und da, wo es sicher großes Interesse gibt, ist es einfach auszuprobieren. Man will jetzt natürlich nicht alles an JetGPT schicken, an OpenAI, gerade die internen Sachen nicht, aber das ist ja, glaube ich, der Ansatz jetzt von Facebook, dass sie ihre Lama -Modelle für die meisten Firmen gratis und für private Anwender auch zur Verfügung stellen, um da OpenAI das Wasser abzugraben. Und da gibt es jetzt gerade so einen Goldrush quasi, diesen Modellen beizubringen, dass sie auch mit diesen internen Dokumenten umgehen können. Und dann hätte man sozusagen einen internen Mitarbeiter, den ich fragen kann, der Zugriff hat auf die Interner, auf den Source -Code, der die APIs vielleicht kennt von dem Programm. Und das kann ja auch etwas Nützliches sein. Also das sind, glaube ich, die zwei großen Anwendungsfelder, nach außen und nach innen blickend. Frau Stambouli, was wird das dann für uns als Gesellschaft, wenn wir so weitergehen,

00:36:53 **Speaker 2**

weg von der Technik hin zu dem, was das für uns als Gesellschaft bedeuten wird? Jetzt hat es vor zehn Jahren bereits eine relativ berühmte Studie der Uni Oxford gegeben, die gesagt hat, die Hälfte in den kommenden Jahren, ich glaube es war 2040, wird die Hälfte aller Jobs von Technologie ersetzt werden. Das hat natürlich dann zwei, ist noch nicht so basiert, aber wird man sehen, stimmen werden solche Prognosen auch noch zutreffend, würde mich interessieren. Und das hat natürlich dann zwei Aspekte. Einerseits für die Menschen, deren Jobs verschwinden sozusagen. Man sagt immer, es werden andere neu geschaffen dafür. Und natürlich auch für die Menschen, die jetzt konfrontiert sind mit Jobs, die halt von einer KI gemacht werden. Beispielsweise ein Klassiker, so Sachbearbeiter in der Versicherung, der entscheidet, ist das ein Schadensfall oder kein Schadensfall und da eine Entscheidung trifft, die dann natürlich auf einen dritten durchaus finanzielle Auswirkungen hat. Das würde mich interessieren,

00:37:42 **Speaker 2**

wie ist da Ihre Erwartungshaltung, wie sehr wird KI hier menschliche Arbeit ersetzen und was wird das für die Menschen bedeuten, die dann von diesen Entscheidungen betroffen sind? Zunächst wirklich ein bisschen ein Zufall, weil diese Studie, die Sie erwähnt haben, ich glaube, ich meine die gleiche, die eben neu, also die gleichen Wissenschaftler haben das jetzt neu aufgesetzt und nachgeschaut, ob ihre Ergebnisse stimmen, zehn Jahre später. Leider habe ich das nur kurz überflogen für ein paar Tage und ich glaube, Sie sind sehr viel milder geworden in Ihrer Annahme. Prinzipiell, also ich bin kein Arbeitsmarktexpertin, aber was man schon beobachten kann, ist alleine, die, also es verändern sich sehr viele Jobbereiche. Das kann man gar nicht verallgemeinern für alle, weil einfach bestimmte mehr, bestimmte weniger, wie ich auch vorhin schon ansprach, gleichzeitig

ist die sozusagen komplette Automatisierung oder KI als Technologie nicht unbedingt mit Jobverlust oder mit Aufstieg gleichzusetzen,

00:38:49 **Speaker 2**

sondern bedeutet immer noch sehr viele schlechte Jobs für sehr viele schlecht, also schlecht bezahlte Menschen weltweit und sehr viel versteckt ist. Also man sieht es dann nicht, es wird oft gehypt und natürlich macht das den Wert dieser Technologien, aber dahinter sind ganz oft Menschen, die das immer noch überschauen müssen und das ist auch gut so und das ist auch mittlerweile auch durch Regulierungen, durch die sogenannte Human Oversight auch mittlerweile auch in EU -Dokumenten zum Beispiel geregelt, dass das sein sollte und sein muss. Was aber jetzt aus philosophischer oder aus ethischer Sicht, wenn man jetzt die Gesellschaft, das ist immer sehr schwierig natürlich, aber es gibt ein Element bei dem Ganzen, das wirklich kritisch zu betrachten ist und nämlich diese Unsicherheit, die dadurch entsteht. Wenn man immer davon spricht und es wahrhält, bewusst wahrhält, dass Jobs wegfallen, dass alles sich umwälzen wird, dass das ganz schlimm und so weiter, dann ist das natürlich, das kreierte

00:39:46 **Speaker 2**

ein gewisses Unbehagen bei den Leuten, also bei sehr vielen, weil keiner mehr weiß, was soll ich machen, soll ich jetzt Programmiererin werden, Programmierer, was übrigens wegfällt, also das anscheinend nicht, weil ja KI dann alles macht, also einige Prognosen, ich spitze es jetzt natürlich zu. Was soll man werden, also am besten Schlosser, Schlosserin, das klingt nach etwas, was man brauchen wird und ich bin jetzt leicht ironisch, weil das sind natürlich unterschiedliche Prognosen, die man aus allen Richtungen so hört. Es wandelt sich aber immer der Arbeitsmarkt und Berufe wandeln sich so oder so. Ich finde es auch manchmal ein anderes Thema in Europa natürlich, als es ein Thema ist in Amerika, Nordamerika oder eben im globalen Süden, also man muss da auch unterscheiden, welchen Bereich man wie anspricht und wir führen die Diskussion sehr oft immer noch aus einer eurozentrischen Perspektive natürlich und aus einer auch gehobenen oder eben diese White -Collar -Jobs, die jetzt auch vielleicht

00:40:43 **Speaker 2**

wegfallen, Sachbearbeiter, Sachbearbeiterinnen und so weiter. Ja, da kann es aber auch sein, dass das wirklich ein Potenzial hat für bessere Jobs, also man weiß nicht genau, ob das jetzt oder ich zumindest würde mich nicht anmaßen zu sagen, da fallen Sachen weg und das ist unbedingt etwas sehr Schlechtes. Nur, wenn man ständig davon spricht und das Thema hatten wir übrigens, ich habe früher sehr viel mich mit Roboterethik und Roboterforschung beschäftigt, dort war das gleiche

Thema. Die Roboter werden uns die Jobs wegnehmen. Ist, naja, nicht wirklich so eingetreten, es ist immer noch natürlich sehr viel automatisiert, klar, aber das hat sich auch sehr gut verkauft so ein bisschen als Panikmache, teilweise auch mitunter die Thematik, dass jetzt definitiv alles umgewälzt wird, radikal ist, aber eben Prognosen sind schwierig und für einige Bereiche trifft es zu, für andere nicht, man kann das nicht allgemein. Aber sozusagen, das ist ja die eine Seite der Medaille, die zweite Seite ist ja sozusagen

00:41:39 **Speaker 2**

diese Entscheidungsfindung und mir sei ein Beispiel, schon ein paar Jahre her, dass amerikanische Banken beispielsweise bei der Kreditvergabe sehr stark schon auf Unterstützung, damals noch Algorithmen, in Zukunft wird es KI sein, setzen, ob ein Kredit vergeben wird oder nicht vergeben wird und dadurch entsteht dann immer so, es gibt dann diesen Bias, also weil ja die KI oder der Computer lernt aus den Ergebnissen der Vergangenheit, dadurch weiß er, dass Menschen in gewissen Gegenden leben, eher keinen Kredit kriegen sollen, weil es schlechtere Gegenden sind und so verfestigen und perpetuieren sich dann halt gewisse Vorurteile, die dann irgendwann beim Menschen vielleicht, der merkt dann, okay, die Gegend ist besser geworden, aber da könnte die Maschine, könnte sein, dass die eine Veränderung der Gegend oder auch auf das Individuum gar keinen Rücksicht mehr nimmt. Inwiefern ist das ein Problem? Das würde mich interessieren, Herr Kettemann, Frau Tomboliev, aber auch die anderen natürlich,

00:42:29 **Speaker 2**

Herr Kettemann, wenn Sie mal vielleicht sagen wollen, wie problematisch ist dieser Bias? Vor allem, wenn man dann auch vielleicht ans Justizsystem denkt, wie wir vorhin gesprochen haben, wo er auch mal vielleicht überlegt, KI mal einzusetzen. Also in diesen Beispielen sieht man, dass KI regelmäßig gar nicht so intelligent ist. In der Tat beruht die Entwicklung von maschinellem Lernen auf den Trainingsdaten, auf der Qualität der Trainingsdaten. Daten, das Gegebene, deutet meistens nach hinten. Es gibt sehr wenig Prognosedaten, die verwendet werden, einzelne synthetische Datenmodelle gibt es natürlich, aber die meisten Daten kommen aus der Vergangenheit. Die Vergangenheit war regelmäßig nicht besser, schöner, liberaler, freundlicher. Das heißt, wenn wir nur auf die Modelle der Vergangenheit schauen, kommt es sehr häufig zu problematischen Ergebnissen. Die angesprochenen Biases, also diese Diskriminierungen, die in die Systeme hineingelernt werden, sind eine Herausforderung. Diese Herausforderung

00:43:26 **Speaker 2**

hat aber der Gesetzgeber gesehen und die Regulierung des Einsatzes von künstlicher Intelligenz am Beispiel dieser Erklärungspflichtigkeit ist ein schönes Beispiel dafür, wie man es richtig machen sollte. Also nicht künstliche Intelligenz generell verbieten, das will ja eh keiner, aber sicherstellen, dass in jenen Bereichen, die besonders eingriffsnah sind, wo Menschen betroffen sind, wo die Rechte von Menschen über die verfügt wird, da muss sichergestellt sein, dass die von einer KI vorbereiteten Entscheidungen auch den Menschen gegenüber erklärt werden müssen. Es gibt also ein ganz großes Forschungsfeld, das nennt sich XAI, Explainable AI, wo verschiedene Ansätze untersucht werden, wie man sinnvoll Entscheidungen, die eine KI vorbereitet, erklären kann, die ein automatisiertes System vorbereitet begründen kann. Ein Modell, das zum Beispiel sehr spannend ist, ist das der sogenannten kontrafaktischen Erklärung, da muss dann die KI mitliefern, welche der Faktoren hätte anders sein müssen,

00:44:39 **Speaker 2**

damit die Entscheidung anders ausgefallen wäre. An dem Beispiel des Kredits wäre das Gehalt höher gewesen oder wäre das Geschlecht anders gewesen oder wäre der Wohnort anders gewesen, welcher Faktor war entscheidend dafür, dass genau diese Entscheidung gefallen ist und wenn wir zum Beispiel bei einer Entscheidung über Kreditvergabe als Ergebnis geliefert bekommen, dass der Faktor Geschlecht war das Entscheidende, dann sind wir mitten im klassischen Diskriminierungsrecht und das darf da natürlich nicht so stehen. Viele der neueren EU -Gesetze, die neuen Rechtsakte, werden uns weitere Pflichten auferlegen, denen die KI einsetzen, im Bereich der Plattformen, die KI, die automatisierte Systeme verwenden, bei der Lenkung von Aufmerksamkeits - und Informationsflüssen ist das jetzt schon geltendes Recht. Im Rechtsakt über digitale Dienste, der sich auf die Facebooks, die TikToks, die YouTubes unserer Zeit bezieht, die sind jetzt schon verpflichtet, darauf offen zu legen, nach welchen Kriterien

00:45:46 **Speaker 2**

Inhalte dort verteilt werden, welche Inhalte präferiert werden, welche empfohlen werden und welche Auswirkungen diese internen Empfehlungssysteme, wo KI meistens eine Rolle spielt, nicht überall, aber meistens, welche Auswirkungen diese Empfehlungssysteme auf demokratische Willensbildung haben, auf öffentliche Gesundheit. Da geht es genau um diese Fragen, dass TikTok, dass Facebook offenlegen muss, ob ihre KIs zum Beispiel in der neuen Corona -Krise mehr Desinformation verbreiten würden oder weniger. Herr Strauß, Sie wollten noch etwas? Ja, ich kann da einhaken. Ich würde da ein bisschen widersprechen, weil das stimmt schon alles, was Sie sagen, letztlich nur das Problem ist, es funktioniert halt nicht. Ich würde mal behaupten, dass Diskriminierung der Status Quo ist, gerade bei sozialen Medien und es gibt beispielsweise so im

Human Resource Management, da gibt es auch schon Software, die quasi benutzt wird für Job Recruitment. Das ist in den USA schon weit verbreitet, bei uns ist es, glaube

00:46:49 **Speaker 2**

ich, noch nicht so verbreitet, weil es, glaube ich, auch rechtliche Hürde gibt dafür. Ich glaube, es sind Konzerne, die in den USA sitzen und in Europa agieren, dürfen das dann auch in Assessment - Centern einsetzen, wo sie beispielsweise da sitzen, sie einfach vor einer Kamera interagieren sozusagen mit einem Bot und der Bot, ein Parameter ist zum Beispiel ihre Hautfarbe, wie sie sprechen und viele andere Parameter und es ist auch evident, dass es diskriminierend ist, wenn sie vor einem Bücherregal sitzen, dann haben sie höhere Chancen, einen höheren Score zu bekommen, dann echt für ein richtiges Job Recruitment -Interview eingeladen zu werden, als wenn sie jetzt Inder sind, einen Akzent haben oder noch schlimmer, ein schwarzer Mensch sind oder eben vor einer weißen Wand sitzen beispielsweise. Also das ist klare Diskriminierung, was macht die Person dann, die wird es ja nicht erfahren, dass sie quasi gar nicht weiterkommt in diesem Recruitment -Prozess und das ist nur ein Beispiel von vielen,

00:47:42 **Speaker 2**

es gibt viele andere und jetzt wäre so quasi die Frage, wie implementieren wir Explainable oder Explicable AI, wie schaffen wir dieses Regelwerk. Ich bin bei Ihnen, da gibt es gute Ansätze und es ist auch ganz gut, was da teilweise dieser risikopassierte Ansatz der MEI vorgesehen ist, gleichzeitig gibt es aber eine ganze Kette von Ausnahmen, da kann man vielleicht auch noch ein bisschen drüber reden, aber mich würde interessieren, auch wie man es technisch implementiert, wie man es sozusagen so gestalten kann, dass das jetzt, wenn wir jetzt bei diesem Job Recruitment -Beispiel bleiben, wie schafft man das, dass die Person konkret in diesem Fall mitbekommt, dass sie gerade diskriminiert wird und was tut sie dann, weil eigentlich ist es davon abweichend, dass die Person sozusagen vermutet, sie wäre diskriminiert worden, muss dann aktiv klagen und das wird keiner tun oder das werden die wenigsten machen. Aber vielleicht eine Frage zu der Diskriminierung, ist das eine Diskriminierung bei diesen

00:48:35 **Speaker 2**

Kreditdaten, ist es einfach, wenn die künstliche Intelligenz aus der Vergangenheit lernt, dann perpetuiert sie bestehende Diskriminierungen oder ist das eine, in dem Fall, den Sie beschrieben haben, eine rein programmierte Diskriminierung? Nein, das ist immer die Frage, was ist programmiert, das ist gar nicht so leicht zu beantworten, da werden Sie vielleicht zustimmen, also wo beginnt es, wenn man sagt, es ist programmiert oder es ist nur ein Parameter von vielen. Wir

haben das Beispiel mit Geschlecht, jetzt kann man sagen, Geschlecht ist völlig irrelevant, gleichzeitig ist es aber fast immer ein Parameter in so ziemlich jedem Algorithmus, der in irgendeiner Form Scoring betreibt. In den USA ist es so, dass Frauen und schwarze Frauen systematisch diskriminiert werden von diesen Kreditvergabe - Algorithmen, was auch nicht nur am Geschlecht ist, sondern auch am Wohnort teilweise. Wenn Sie jetzt in einem Problemviertel in Los Angeles sind, dann werden Sie wahrscheinlich sehr wenig Chancen

00:49:21 **Speaker 2**

haben, einen Kredit zu bekommen, vielleicht sogar egal, wie hoch Ihr Einkommen ist, das ist jetzt natürlich überspitzt gesagt, aber das sind so Aspekte, da sehe ich keine Lösung momentan dafür, da sehe ich keine Kette von Problemen. Würde sich das technisch lösen lassen, nicht mehr ans ganze Podium interessieren, also wenn man die Regulierung hätte, wenn man der Willen dahinter wäre, würde sich das ändern lassen oder ist das einfach, weil es doch ein bisschen eine Blackbox ist mit Machine Learning, gar nicht möglich? Das ist eine gute Frage, also die Frage ist natürlich, erstens einmal, ich glaube, die Frage müsste früher beginnen, das ist natürlich jetzt auch ein bisschen skizzierter, sozusagen schon fast eine Utopie, wir müssten beginnen zu fragen als Gesellschaft, in welchen Bereichen wir so eine Form von Automatisierung eigentlich wollen, wo wollen wir sie nicht, weil es ist letztlich, als Sie ganz am Anfang gefragt haben, was ist KI, also ich würde sagen, KI ist letztlich eine neue

00:50:05 **Speaker 2**

Form der Automatisierung von Bereichen, die es vorher nicht betroffen hat, das betrifft Wissensarbeit, das betrifft kognitive Tätigkeiten, das ist eine Imitation natürlich, aber es ist ein Intrusionsgrad von Technologie erreicht, den wir bislang noch nicht gesehen haben und da müssen wir einfach ganz neu denken, was Automatisierung eigentlich bedeutet. Herr Kettemann, kurzer Zwischenruf, das Schöne ist ja, dass wir durch den Einsatz von KI erst so richtig sehen, wie kontingent diese ganzen Entscheidungsprozesse sind, weil natürlich bei jeder Entscheidung über eine Einstellung, bei jeder Kreditvergabe waren vorher auch Faktoren am Werk, da haben Menschen Entscheidungen getroffen, da haben Menschen auf den Postcode geschaut und gesagt, ach, von dem Viertel, die zahlen das nicht zurück. Das heißt, wir sehen jetzt, welche Faktoren wir wünschen, dass sie in eine Entscheidung führen und welche wir für problematisch halten und wir verzweifeln daran, weil wir merken, wie unglaublich schwierig

00:51:03 **Speaker 2**

das ist, diskriminierende Faktoren aus Entscheidungsprozessen heraus zu programmieren. Deswegen scheitern wir ja auch daran zum Beispiel, automatisierte Autos ganz automatisiert zu gestalten, weil wir einfach nicht einprogrammieren wollen, im Notfall überfahre lieber die zwei Kinder und nicht die drei Kinder, sondern das sind solche Notfälle, die wir einfach nicht vorprogrammieren können und dasselbe sehen wir jetzt auch hier. Also ich sehe der technischen Umsetzung mit großer Freude entgegen, würde aber, freue mich aber auch gleichzeitig, auf der Rechtseite zu sitzen, also auf der Seite des Rechts, weil wir hier nur den Rahmen geben müssen. Frau Stambouli, ich würde mich interessieren, der Herr Strauß hat vorhin gerade angesprochen, dass man die Diskussion zurückziehen müsste, in welchen Bereichen wollen wir überhaupt die KI haben. Jetzt kann man sagen, wer könnte das definieren und die zweite Frage und dann das ganze Problem, wie realistisch ist es, weil wir wissen natürlich, wenn es

00:52:02 **Speaker 2**

Anwendungsfälle gibt, ein Verbot sozusagen, ich meine, vielleicht wenn maximal in Europa, aber es ist anzunehmend global, dass natürlich es irgendwo in irgendeiner Jurisdiktion umgesetzt wird und kann man überhaupt verbieten quasi, okay, KI darf für solche Sachen nicht verwendet werden. Ist das überhaupt realistisch? Also ich stimme dem zunächst wirklich vollkommen zu, dass wir eigentlich viel früher ansetzen müssen und ich bin jetzt jemand, der per se nicht an ausschließlich technische Lösungen glaubt und auch nicht, dass ein Problem mit Diskriminierungen bei Jobauswahl ein KI - Problem ist. Es ist natürlich, man darf eine Sache nicht unterschätzen und die ist im ethischen Bereich unglaublich schwer nachzuweisen, aber das kommt natürlich aus der Forschung auch heraus, Unternehmen zum Beispiel diskriminieren und die sind sich dessen bewusst. Das heißt, KI hat nicht immer nur eine aktive Funktion zu diskriminieren, sondern versteckt auch oft Diskriminierung, die sich darin zeigt. Sprich,

00:53:02 **Speaker 2**

es sitzt jetzt trotzdem kein Mensch da, so würde ich jetzt annehmen, oder die ganz bewusst sagt, ja natürlich, ich möchte diskriminieren, so funktioniert es nicht oder ganz böswillige Absichten, aber das ist das Problem auch mit diesem Bias zum Beispiel, der also voreingenommen hat oder Vorurteile, dass sie sich in solchen Systemen natürlich verfestigen und diese multiplizieren und das Problem an der KI und deswegen ist es ja mitunter auch so gefährlich, ist dann der Ausmaß. Das wird dann schnell zu einem riesengroßen Problem und das merkt man auch bei Skandalen, die immer wieder auftreten, wie zum Beispiel in den Niederlanden vor ein paar Jahren mit der Kindergeldaffäre, wo sehr viele Menschen falsch beschuldigt wurden, dass sie Kindergeld bekommen, was sie nicht bekommen sollten. Das hat, also wir reden hier von, als ob das ein

technisches Problem ist, aber das hat zu absoluten Krisen von zigtausenden von Menschen geführt, unter anderem auch Selbstmord und Scheidungen, Kinder wurden weggenommen,

00:53:55 **Speaker 2**

Schulden und so weiter und so fort. Sprich, und jetzt zu dem Thema, wer soll entscheiden und was soll verboten werden? Eines der Hauptprobleme war unter anderem, es war nicht erklärbar, wie dieser Algorithmus funktioniert oder es gab einfach viele Unklarheiten und Intransparenzen, aber es gab auch kein Problem für die Menschen, also keine Möglichkeit für die Leute einzuschreiten und sich dagegen zu wehren. Das heißt, Zivilgesellschaften können immer noch nicht genug mitsprechen, wenn es um diese Anwendungen geht. Nun, ich sage das und gleichzeitig weiß ich aber, es gibt Forschung zur Innovation, die ich jetzt aus der Robotik kenne, nicht allgemein, die davon ausgeht, wenn man Leute fragt, ob sie etwas Neues möchten, reagieren sie vielleicht oft konservativer und kritischer. Manche Dinge werden einfach in den Markt gegeben und dann interagiert man damit. Das heißt, inwiefern kann man Menschen wirklich auf solche Dinge vorbereiten und sie fragen vorab. Man sollte aber definitiv die, die davon

00:54:53 **Speaker 2**

betroffen sind und die so schon vielleicht vulnerable Gruppen sind in der Gesellschaft, mit einbeziehen, noch bevor Dinge schief laufen. Das ist etwas, was wir definitiv machen sollten, weil langfristig zum Beispiel nicht zu wissen, wie ein System einen bewertet, nicht zu wissen, an wie man sich wenden kann, nicht zu wissen, wie man aus dem rauskommt, das war vorher vielleicht schon so, aber das darf jetzt nicht noch schlimmer werden, weil dann äußert sich das in anderen Krisen und die äußern sich dann wiederum demokratisch in der gesamten Gesellschaft. Ich meine, ein Punkt, den wir jetzt noch nicht angesprochen haben, bei Vertrauen ist ja sehr stark, jetzt hier war es sehr stark, Entscheidungen der KI. Was anderes ist ja auch ein Thema, das ja auch wahrscheinlich nächstes Jahr gibt es ja ein super Wahljahr mit Wahlen in Europa, USA und wir wissen, dass die KI das ganze Thema Deepfakes erleichtern wird, also das ist quasi Videos beispielsweise, es kann die KI, kann, wenn sie, ich glaube

00:55:52 **Speaker 2**

inzwischen braucht es nun mal bald 30 Sekunden oder zwei Minuten, eines Videos eines Menschen, dann kann die KI, also gibt es diverse und sie werden immer besser ein Video erstellen, das mit der originalen Stimme und mit den richtigen Lippenbewegungen irgendwas anderes sagt und das macht es natürlich sehr gefährlich, gerade in der politischen Diskussion, wenn man dem politischen Gegner sehr leicht Aussagen unterstellen kann und wissen, dass wir in unserer Gesellschaft sehr schnell oft

Anschuldigungen da sind, die auch dann zu großen Krisen führen können und auch im persönlichen Bereich. Inwiefern ist diese Thematik, dass sie ein Problem ist, glaube ich, ist bekannt, aber wie groß könnte dieses Problem noch werden und wie kann man mit dem umgehen? Da wäre zum Beispiel Transparenz ein Thema. Herr Heitzinger, das würde mich interessieren, wie sehen Sie das Thema mit Umgehen? Das ist auch ein wichtiger Punkt und eigentlich wie bei den anderen Punkten jetzt, eigentlich kein neues Problem. Also

00:56:47 **Speaker 2**

wir wissen in den USA, es gibt schon seit Jahrzehnten Robocalls, in denen Gerüchte in die Welt gesetzt werden. Das Ganze wird jetzt halt noch besser und schneller und billiger und ich glaube, wir stehen da vor der alten philosophischen Frage, was ist Wahrheit und wie komme ich zur Wahrheit? Also wenn wir ein paar Jahre zurückgehen, als die Suchmaschinen aufgekommen sind, also Google, haben wir auch gesagt, wir wenden das natürlich, wir dürfen es verwenden, aber wir sollten uns dann Gedanken machen, dass wir das Ganze nachprüfen und es war aber nicht so, dass sich die Leute dann vor den Bibliotheken angestellt hätten und alles verifiziert hätten. Also das Problem ist an sich nicht neu. Was dazu kommt, ist, dass das geschriebene Wort, wird es halt billiger zu generieren und dazu das Video auch noch. Also ich denke, wir müssen uns wirklich Gedanken machen, wenn wir Inhalte konsumieren, von wem kommen die und da sind wir wieder beim Bildungsthema eigentlich, um irgendwelche Aussagen einordnen

00:57:54 **Speaker 2**

zu können, brauche ich ein Grundwissen, ein Basiswissen, um abschätzen zu können, ist das jetzt plausibel, ist das jetzt eine Aussage, die mir weiterhilft, wenn ich Chatship -Idee als Experten verwenden will oder ist das einfach eine erfundene Aussage. Das bringt uns dann, glaube ich, auch zum Bildungsthema, weil wir werden dann in ungefähr zehn Jahren die erste Generation von Studierenden haben, die gar nichts anderes mehr kennen, die seit Volksschulzeiten wissen, okay, da gibt es ein System, dem kann ich Fragen stellen und ich bekomme auch immer eine Antwort. Das heißt, wir müssen uns irgendwie als Gesellschaft auch klar werden und in diesen Hallen darf man das ansprechen, was die Bildung betrifft. Wir müssen uns auch klar sein, was ist so das Grundwissen, das jeder haben sollte, das man mitbringen muss, um diese Inhalte, die wir auf uns dann zusteuern, um die richtig einordnen zu können. Frau Stamboliev, Sie wollten noch, ich würde gerne eine Sache wirklich an dem anknüpfen, weil ich

00:58:56 **Speaker 2**

bin relativ besorgt über diese Deepfakes und es stimmt mit dem Wahr und Falsch und dass wir nicht immer alles nachprüfen können und das ist ein ganz großes Problem, aber eine Sache, die mir wirklich große Sorgen bereitet, ist, dass wir die Quellen nicht mehr überprüfen können. Sprich, Deepfakes könnten dann eine Quelle so simulieren, dass wir quasi nichts mehr vertrauen können, was wir aufgrund von optischer oder Audioqualität her wiedererkennen. Sprich, es geht dann gar nicht mehr um Wahrheit oder Lüge oder Falschinformation, sondern dass es dann unmöglich wird, an Informationen zu kommen, die in irgendeiner Form eine Referenz haben, auf die ich mich festlegen kann. Und das, glaube ich, könnte zu einer wirklich großen Wissenskrise und zu einer Vertrauenskrise führen, weil dann ist es nämlich fast schon unmöglich, dann macht ja nichts mehr Sinn, weil dann kann ich mich überhaupt nicht mehr orientieren und das unabhängig von meiner politischen Ausrichtung, sondern ich kann dann niemandem

01:00:00 **Speaker 2**

mehr vertrauen und nichts vertrauen, was ich lese oder sehe. Herr Kettemann, Sie wollten auch noch. Ja, also Deepfakes sind irrelevant, muss man einfach so sagen. Ja, es gibt Einzelfälle, aber die spielen gesellschaftlich überhaupt keine Rolle. Es gibt sie schon lange. In keinem Desinformationskonflikt wurden die effektiv bisher eingesetzt und werden das auch in Zukunft nicht tun. Und die Einzigen, die davon profitieren, wenn wir von Vertrauenskrisen und Wissenskrisen sprechen, sind genau diese Desinformationsakteure. Das ist ja auch das, was Sie erreichen wollen. Wenn Russland zum Beispiel Desinformationen lanciert, dann tun Sie das ja nicht deswegen, weil Sie glauben, dass man an diese Desinformationen glaubt, sondern weil Sie eine Atmosphäre der Furcht, eine Atmosphäre der Unsicherheit verbreiten wollen. Also würde ich vorschlagen, wie ich das immer tue, wenn ich über Desinformationen rede und das tue ich in letzter Zeit sehr häufig, es gibt einfach empirisch keine Nachweise, dass Desinformation

01:00:53 **Speaker 2**

wirklich effektiv uns in irgendeine Richtung verändert, unser Wissensbestand verändert, unsere Meinungen verändert. Das gibt es einfach nicht. Aber was wir sehen ist, dass künstliche Intelligenz eingesetzt werden kann, um Desinformationen zu beschleunigen und sie preiswerter zu machen. Schon jetzt ist es so, dass diese Informationen genutzt werden, um politische Prozesse zu beeinflussen. Wie gesagt, besonders effektiv ist es nicht. Aber in die Bildung zu investieren kann nie falsch sein. Wir haben das Fach digitale Grundbildung. Meine Kinder lernen in der dritten Institution Ascher, wie man Influencerinnen und Influencer identifiziert, wie man nachprüfen kann, ob ein Bild echt ist oder falsch ist. Das ist auf jeden Fall ein sinnvoller Ansatz, aber nicht Angst zu haben vor Desinformationen. Herr Heitzinger, dann Herr Strauß. Ja, es gab zu Zeiten der Suchmaschinen, als sie

aufgekommen sind, also vor ein paar Jahren, hat man sich die Frage gestellt, können Suchmaschinen und diese Suchergebnisse

01:02:00 **Speaker 2**

das Wahlverhalten verändern? Also wir sehen ja die ersten zehn Suchresultate auf der ersten Seite. Man hat sich die Frage gestellt, wenn ich jetzt nach einem Kandidaten suche, also es war eine US - amerikanische Studie, macht das einen Unterschied, in welcher Reihenfolge die Kandidaten angezeigt werden oder ob die auf der zweiten Seite vorkommen? Und man hat das wirklich untersucht, indem man Leute hingesezt hat, Kandidaten erfunden hat, diese Suchergebnisse erfunden hat und das wirklich Erschreckende dabei war eigentlich, dass schon kleines Umordnen der Reihenfolge dazu geführt hat, dass sich die Wahlergebnisse verändert haben. Also das war in den Proceedings der National Academy of Sciences und wir wissen ja, wie knapp die Wahlergebnisse heutzutage sind. Also wenn das einen Unterschied macht von ein oder zwei Prozent oder fünf Prozent vielleicht, dann allein wie ich diese Suchergebnisse anordne, dann kann das durchaus große Auswirkungen haben. Das heißt, die Frage ist, die hätten wir

01:03:03 **Speaker 2**

uns damals schon stellen sollen, wie genau verstehen wir eigentlich die Algorithmen hinter den Suchmaschinen und sind da die richtigen Trainingsdaten drinnen, wer hat die Algorithmen programmiert und dort vielleicht noch etwas eingebaut? Also das ist auch eine Frage, die hätten wir uns damals eigentlich schon viel mehr stellen sollen und die nächste Frage ist dann auch, haben wir die Technologie in Europa oder machen wir uns jetzt abhängig von den USA, vielleicht von China? Also das ist auch ein Thema, ich glaube, das wir ansprechen sollten. Ich glaube, der Herr Kettemann wird auch keinen antworten. Herr Strauß hat sich vorher schon gemeldet gehabt. Herr Strauß, aber Herr Kettemann. Das ist kein Problem. Also ich würde sagen, mitunter schon. Ich würde jetzt das Beispiel Cambridge Analytica ins Treffen führen. Da kann man jetzt sagen, okay, das war jetzt keine KI, aber das ist doch ein ziemlich großer Skandal. Also da wurde, immerhin wurden 87 Millionen Userdaten einfach missbraucht, um

01:04:00 **Speaker 2**

dann ganz gezielt so beeinflussende Werbung zu schalten. Das Problem dabei ist natürlich, es gibt wenig Empirie, das stimmt, wenig Empirie über die Wirksamkeit davon, aber das Problem ist halt auch, es wird Ihnen ja niemand erzählen in einem Interview beispielsweise, ja, ich habe Trump gewählt, weil ich auf Facebook diese Werbung gesehen habe. So einfach funktioniert das Spiel halt nicht. Es gibt auch diese Fälle der beeinflussenden Werbung aus den 80er -Jahren, Coca -Cola mit

diesen Wüstenbildern. Das ist auch ganz interessant. Da würde ich ein bisschen sehen, dass ich so einen Diskurs auf einer ganz anderen Ebene wiederhole. Ich glaube, man könnte auch darüber diskutieren, ob sich die Werbemechanismen und das sogenannte Persuasion -Marketing nicht in eine Richtung entwickelt, dass das schon so intrusiv geworden ist, genau über die Digitalisierung, dass es vielleicht hier Ansätze bräuchte, um sozusagen diese Problematik, die Sie auch ganz gut jetzt angesprochen haben, ein bisschen zu entschärfen.

01:04:52 **Speaker 2**

Da wären wir dann auch wieder irgendwie bei der Regulierung vielleicht. Sie haben das, glaube ich, eh gesagt, wer profitiert davon? Sie haben völlig recht, Russland profitiert natürlich davon, den Diskurs zu zerstreuen. Da ist nicht wichtig, ob das geglaubt wird. Aber es ist de facto in der Ukraine, glaube ich, ein ziemliches Problem, diese Propaganda und Gegenpropaganda, das bekommen wir hier nicht so mit, das stimmt. Gleichzeitig gibt es auch ziemlich harte Indizien, es gibt da Untersuchungen des House of Commons, dass Cambridge Analytica nicht nur in den USA bei der Wahl möglicherweise eine Rolle gespielt hat, sondern auch beim Brexit, da gibt es eigene Untersuchungen dazu, weil es da auch so Subfirmen gab, die da gegründet wurden aus diesem seltsamen Firmenkonstrukt rund um Cambridge Analytica, die SCL Group, die sind dann alle ins Event gegangen und dann wieder in neuem Namen wieder aufgetaucht. Also da gibt es einen ganz interessanten Graubereich. Man kann auch diskutieren, ist es

01:05:44 **Speaker 2**

legal oder ist es nicht legal? Das ist, glaube ich, auch eine ziemlich unklare Frage letztlich. Also es ist leicht zu sagen, das war alles illegale Manipulation. Ich glaube, der Nachweis ist relativ schwierig tatsächlich. Aber worauf ich hinaus will, ich glaube, das ist schon ein Problem. Aber das geringste Problem, das wir vielleicht haben, ist, dass der Papst in einer Daunenjacke abgebildet wird, weil das war ja tatsächlich einfach nur sozusagen ein Gegner. Herr Klettermann, Sie wollten antworten, ich würde nur ganz kurz nur den einen Punkt, legal, nicht legal, ich glaube, das Problem ist halt, es muss auch die entsprechenden Gesetze geben. Und gibt es die? Ja, und die Frage ist natürlich, was ist Manipulation? Ja, also ich glaube, zwischen legaler und illegaler Manipulation zu unterscheiden, ist schon einigermaßen schwierig. Jede Werbung ist Manipulation, jeder Artikel ist in gewisser Weise Manipulation. Oder ist es eine legitime Meinungsäußerung? Ja, das Recht kann bei der Desinformation

01:06:32 **Speaker 2**

sehr wenig tun und das soll es auch nicht. Wir wollen nicht, dass die Staaten der Welt entscheiden dürfen, was Wahrheit ist. Wollen wir wirklich, dass die Staaten uns das Lügen verbieten? Nein, weil dann müsste es ja Wahrheitsministerien geben, die gibt es ja zum Glück nicht. Wir wollen auch nicht, dass autoritäre Staaten mehr Macht bekommen, darüber zu entscheiden, was gesagt werden darf und was nicht gesagt werden darf. Wir verstehen immer besser, wie Menschen Entscheidungen treffen, das stimmt. Wir verstehen auch immer mehr darüber, wie Menschen Meinungen bilden. Aber hier würde ich doch meinen, dass KI eine relativ untergeordnete Rolle spielt, dass Cambridge Analytica und die anderen eine sehr untergeordnete Rolle spielen und dass die vollkommen legalen und normalen Lügen eines Johnson, eines Trump und von Fox News viel einen größeren Einfluss haben auf das Wahlverhalten, auf Entscheidungsfindung von Menschen, als das die KI macht. Die KI macht vieles falsch, aber nicht alles und sie

01:07:33 **Speaker 2**

ist ein schöner Sündenbock. Natürlich haben die Akteure, auch im House of Commons, haben auch Akteure in der amerikanischen Meinungslandschaft ein riesen Interesse daran zu sagen, ja die KI war das, ja das waren diese bösen Computer, die das gemacht haben und diese fiesen Firmen. Wenn in Wirklichkeit wir ganz andere Herausforderungen im demokratischen Diskurs haben, denen man ein Herr werden könnte mit ganz einfachen Maßnahmen, wie zum Beispiel einer verstärkten Orientierung auf die wirklichen Herausforderungen, vor denen wir stehen, einer Verlängerung unserer Zeithorizonte nach vorne. Wir haben große Fragen, große Krisen, aber die KI ist zum Glück nicht die größte von ihnen. Aber Demokratie, würde ich schon sagen, ist ein wichtiger Punkt. Ich finde Demokratie ist ein wichtiger Punkt. Weil Sie sagen, das hat ja keinen Einfluss, das hat einen Grund, warum es da ziemlich lange Untersuchungen gibt. Es gibt auch Studien dazu, die sagen, dass es schon einen Einfluss hatte. Also Cambridge Analytica,

01:08:26 **Speaker 2**

das ist jetzt kein KI -Thema, aber es ist ein Manipulationsthema. Und wenn jetzt ein politischer Akteur hergeht und diese Technologien bewusst einsetzt, um sozusagen einen Diskurs zu verzehren, dann könnte man schon sagen, das ist ein Problem, weil es hat ja auch einen Grund, warum das gern gemacht wird. Also wenn man davon ausgeht, dass es eh nichts bringt, dann würde niemand so viel Geld in die Hand nehmen. Aber es gibt keine empirischen Nachweise dafür, dass Cambridge Analytica, also diese gezielte Ausspielung von Werbungen, wirklich irgendetwas bringt. Nein, da gibt es Studien dazu. Die im Promillebereich Unterschiede zeigen. Okay. Jetzt sind wir schon sehr spezifisch in dem Thema. Ich muss ein bisschen auf die Zeit schon, ich liebe es, wenn eine Diskussion so von selber läuft, aber ich glaube, ich möchte nämlich das Publikum, nur Frau Stamboliev, Sie

wollten noch ganz kurz was dazu sagen und dann möchte ich mal mit der Publikumsrunde, also bereiten Sie sich schon vor bitte mit Ihren

01:09:16 **Speaker 2**

Fragen, weil dann würde ich mal gerne das Wort an das Publikum geben. Aber Frau Stamboliev, Sie haben sich vorher auch noch gemeldet. Ich habe wahrscheinlich meinen Punkt schon vergessen, aber weil ich das sehr spannend finde, das sind genau diese Positionen, die auch sehr oft in der Forschung auch zu finden sind. Aber was für mich jetzt auch wiederum sehr klar wird und das ist bei dem Thema einfach immer wieder das Gleiche, bestimmte Konzepte werden sehr wohl im Moment ausgespielt, das stimmt, das Thema Desinformation als etwas sehr Neues oder KI -gebundenes zu nehmen, ist nicht richtig, also das passt einfach nicht, das ist eine gewisse Ablenkung, das gibt auch der KI zu viel Macht. Gleichzeitig, ja, die Technologie hat eine gewisse Macht, auch wenn es nur eine Verschleierung von Entscheidungen ist, die im Hintergrund passieren und im Endeffekt muss man sagen, man muss diese Form von Technologien immer wieder auch als ideologische Technologien sehen und als Menschen,

01:10:11 **Speaker 2**

gebrauchte Instrumentarien, die auch natürlich im politischen Machtkampf immer wieder eine Rolle spielen und wo es dann aber auch sehr schwer ist, nachzuweisen, wer wie, wo was und empirisch und das stimmt, jetzt kam wieder eine neue Studie auch in Nature, gerade die Tage raus, wo man festgestellt hat, also man hat 2020 auch Wahlkämpfe und alles untersucht, unglaublich viele Menschen befragt in UK und USA und die versuchen auch quasi herauszufinden, wie sehr polarisierende Nachrichten sich gegenseitig dann quasi kontaminieren und finden heraus, naja, nicht so sehr, aber was heißt das jetzt, also es ist unglaublich schwierig, dann auch empirisch genau alles, man kann auch nicht feststellen, wer liest was und wählt dann was, das darf man ja nicht, also im Prinzip, ja, schwierig. Okay, dann machen wir hier mal kurz eine Unterbrechung sozusagen und ja, verehrtes Publikum hier im Saal, wenn Sie mit einem Handzeichen, die Kollegin drüben hat ein Mikro und ich darf nur vorab, es ist ein unglaublich

01:11:08 **Speaker 2**

spannendes Thema, aber ich bitte darum, kein Ko -Referat, sondern doch eine konkrete Frage relativ zügig zu stellen. Ich glaube, der Herr hier in der, hier link, rechts waren einmal zwei, nehmen Sie das Mikro genau hintereinander und dann. Ich habe eine ganz konkrete Frage, wie gut das Chat -GPT funktioniert, wenn ich eine Vorgabe gebe, ich nehme ein ganz primitives Beispiel, ein Schüleraufsatz über Chat -GPT, soll schreiben, soll eine Seite lang sein, wird immer das Gleiche, der gleiche Wortlaut

herauskommen bei dem Aufsatz und wenn ja, warum und wenn nein, warum nicht? Herr Heitzinger? Es wird nicht der gleiche Wortlaut herauskommen, also die Ausgabe ist prinzipiell zufällig und ich, ja, Anfang des Jahres, als Chat -GPT aufgekommen ist, hat ein Fernsehsender da ein interessantes Experiment gemacht, ist in eine Wiener Schule gegangen und hat dort einen Deutschlehrer genommen und drei Schülerinnen und Schüler, die eine Matura, Deutschmatura -Aufgabe gestellt bekamen, das war ein Leserbrief,

01:12:20 **Speaker 2**

also mit Referenz auf einen Artikel in einer Tageszeitung und es war dann ein Leserbrief zu schreiben und eine Schülerin, ein Schüler hat, sollte selbst schreiben, der zweite sollte Chat -GPT einfach verwenden und der dritte eine Hybridversion abliefern, also eigene Gedanken, aber auch ein Teil inspiriert oder verarbeitet die Chat -GPT -Ausgabe und die Frage ist natürlich, wie viele davon hat der Deutschlehrer erkannt als computergeneriert und wie viele von den Schülern? Also das waren tatsächlich Schüler, die er acht Jahre betreut hat als Deutschlehrer und die Antwort hat nichts erkannt. Also er hat nicht erkennen können, welcher Aufsatz von einem Schüler, von einer Schülerin war und welcher Aufsatz computergeneriert war. Das heißt, der Aufbau war genauso, wie es in der Schule vermittelt wurde, das Programm hat den Artikel gelesen, online gefunden und verarbeitet und wenn man sich die technische, zweite Antwort ist, wenn man sich die technischen, das was publiziert wurde, die Tech -Reports

01:13:32 **Speaker 2**

ansieht von OpenAI, wo man so ein bisschen erklärt, wie das Ganze funktioniert, mehr oder weniger genau, da gibt es eine sehr interessante Abbildung, in denen die Performance, die Leistung auf gewissen Advanced Placement Tests in den USA gezeigt wird. Das sind so Tests, das ist eigentlich ein maturer Niveau, ein bisschen drüber, das ist so Bachelor -Niveau in den USA, erste zwei Jahre Studium, geht so von 0 bis 100 Prozent richtig und da gibt es eine große, das geht über alle möglichen Fachgebiete, also wirklich durch die Bank und alle möglichen anderen Tests auch und das funktioniert wirklich gut, also es geht von 90, 100 Prozent, es gibt nur eine Handvoll Fächer, wo es sehr schlecht ist, aber es ist ein Programm, das würde dort zum Studium zugelassen werden und zu den meisten Studien zugelassen werden und die ersten Jahre den Bachelor wahrscheinlich abschließen. Okay, danke für die Antwort, dann hatten wir hier hinten eine Frage, Herr mit dem Mikro, ja? Genau, meine Frage richtet sich

01:14:39 **Speaker 2**

auf die Aussage, dass die Ära der Suchmaschinen vorbei sei und die Ära der Erklärmaschinen beginnen könnte oder würde und jetzt ist aber meine Nachfrage noch einmal, es hat sich ja jetzt auch in der Diskussion gezeigt, dass immer wieder in Frage gestellt wurde, ob diese Modelle überhaupt die Welt erklären können. Sie meinten, dass das Problem mit den Quellen ist, also der GPT weiß nicht, auf welche Quelle es zugreift, sondern es sind ja eher statistische oder Muster eigentlich, die erkannt werden und dann hört man auch immer wieder diese Rede von der Blackbox - KI, also man weiß nicht wirklich, wie die Ergebnisse zustande kommen und deshalb möchte ich hier noch einmal nachfragen und Sie bitten um eine Präzisierung oder um eine tiefere Ausführung. Können die Modelle, die wir jetzt haben, uns die Welt erklären? Verstehen Sie etwas von der Welt, so dass Sie uns das erklären können? Wer möchte antworten? Das wäre eigentlich eine Frage für Sie, also ich würde sagen nein. Kurze Antwort, ja, Herr

01:15:39 **Speaker 2**

Kettmann? Also Sie müssen es teilweise, weil jetzt zum Beispiel in den schon erwähnten neueren EU -Gesetzen drinnen steht, dass Unternehmen, die KI einsetzen, gerade auch jetzt im geplanten KI - Rechtsakt, dass die eine Erklärung bereitstellen müssen, in besonders eingriffsnahen Bereichen. Das heißt, die Technik wird sich überlegen müssen, wie man bestimmte Parameter auf eine Weise erklären kann, die dann auch bei den Menschen ankommt, weil es geht ja nicht nur darum, dass man erklärt, sondern auch, dass man das verstehen kann. Ich glaube übrigens nicht, dass die Ära der Suchmaschinen zu Ende ist, ich glaube, wir werden eher zu einer Mischung kommen, wo die Suchmaschinen und die Antwortmaschinen zusammenschaltet werden, also so eine hybride maschinelle Unterstützung von Lern - und Kommunikationsprozessen. Das wird viele Arbeitsumgebungen beeinflussen, aber sicher nicht ersetzen. Ich finde, es ist wirklich eine spannende Frage und ich glaube, also jetzt abgesehen davon, nein, glaube ich,

01:16:46 **Speaker 2**

also auch zuerst nein, aber die Welt erklären, man sollte die Welt ja auch erfahren und was an Wissen und Wissensaneignung ja so spannend ist, ist ein Prozess, den man durchläuft und jetzt rein von dem ganzen Erfahrungsschatz, den man sich dadurch erspart, man kann also wirklich persönlich gesehen nur appellieren, dass bestimmte Dinge vielleicht nicht auszulagern sind, weil sie einem auch sehr viel nehmen. Man denkt, man ist effizienter, aber man nimmt sich sehr viel Erfahrung.

01:17:18 **Speaker 2**

Dann haben wir hier in der Antwort auch über die Nützlichkeit von ChatGBT. Ich bin auch persönlich sehr froh, ich habe ChatGBT Plus auch abonniert und ich verwende es eigentlich jeden Tag. Es ist

sehr praktisch, also einerseits bei Sachen wie Programmieren und so und auch zum Beispiel, ich schreibe sehr gerne philosophische Essays und wenn es um Recherche geht, gibt es nichts Besseres als ChatGBT, weil da gebe ich einen zum Beispiel, was hat dieser eine Professor in diesem einen Podcast gesagt, kannst du mir das nochmal sagen? Er findet den Podcast, sagt, eine Stunde zwanzig hat er diese Aussage getätigt und ich habe mir schon eine Stunde Suche erspart zum Beispiel. Und dann jetzt noch die Frage. Ich zum Beispiel habe nicht viel Angst vor AI. Ich habe nicht viel Angst davor, ich denke, AI wird ein guter Freund für uns Menschen sein. Ich habe eher Angst, wie behandeln wir die AI und deswegen, das ist meine Frage an das Podium, habt ihr euch schon ethisch überlegt, wie wir AI behandeln und

01:18:19 Speaker 2

mögliches schlechtes Verhalten gegenüber der AI verhindern könnten? Gegenüber der KI, okay.

01:18:32 Speaker 2

Sehr interessante Perspektive, würde ich sagen. Also ich glaube, vielleicht die Frage ist, muss man aufpassen, dass man es nicht überreguliert, glaube ich. Ich glaube, das ist definitiv ein Problem. Wir diskutieren das in anderen Kontexten, nämlich der Kollege spricht an die Frage, ob man nicht jetzt KI-Rechtspersönlichkeit geben soll oder bestimmte Rechte geben soll. Das hat bei der Roboterethik auch diskutiert, brauchen Roboter Rechte gegenüber Menschen? Müssen wir gerade mal humanoide Roboter mit bestimmten Rechten ausstatten, auch deswegen, weil wenn wir das nicht tun, dann macht es ja etwas mit uns, genauso wie auf die Tierrechtsbewegung eines Argumente für das gute Verhalten zu Tieren ja ist, wenn wir Tiere auf eine bestimmte Weise behandeln, das macht etwas mit uns. Wenn man in den Abgrund schaut, schaut der Abgrund auch in den einen und deswegen wird die Debatte lebhaft geführt. Aktueller Stand ist, dass wir keine eigene Rechtspersönlichkeit, zumindest für automatisierte Systeme

01:19:29 Speaker 2

brauchen, dass wir aber sehr wohl über Haftungsfragen reden müssen und da gibt es so interessante neue Entwicklungen, dass man sagt, es kann ja durchaus so sein, dass KI-Systeme selbst sich so eine Art Haftungsfonds ausarbeiten sollen, dass man also einer KI die Möglichkeit gibt, zum Beispiel auf der Börse zu spekulieren, mit dem Ziel einen Haftungsfonds zu entwickeln, aus dem dann Personen, die von der KI beschädigt werden, entschädigt werden können. Das sind aber Modelle, die wir noch in Zukunft anschauen werden, wobei wir sagen müssen, das Rechtssystem kennt natürlich jetzt schon ganz viele nichtmenschliche Akteure, die Rechtspersönlichkeit haben und es funktioniert seit sehr langer Zeit sehr gut. Jedes Unternehmen hat eine Rechtsperson, obwohl es nicht wirklich eine

Person ist. Das heißt, da ist das Recht nicht für unlösbare Herausforderungen gestellt. KI -Systeme würde ich nicht sagen, dass wir die wirklich schlecht behandeln könnte. Bei humanoiden Robotern ist es ein bisschen anders

01:20:23 **Speaker 2**

wegen der Nähe zu dem Menschenbild, würde ich meinen. Okay, dann hatten wir hier hinten zwei Fragen noch und eine dann auch hier vorne.

01:20:39 **Speaker 2**

Dankeschön. Eine Frage zur Richtung des Datenflusses. Wir haben bis jetzt die ganze Zeit davon gesprochen, dass der Datenfluss von zum Beispiel Jett -TPD, also über AI, zum Konsumenten geht. Jetzt, wenn ich davon ausgehe, dass das System ja möglicherweise auch aus den Fragen sozusagen sein eigenes Repertoire erweitert, stellt sich für mich die Frage, wer steht dann hinter dem, was an neuen Daten in das System einfließt, speziell wenn ich an wissenschaftliche Fragestellungen denke. Weil was ich sehe ist oder was ich hier mitkriege ist, dass dann viele auch bereits Jett -TPD, als Beispiel genannt, anstatt einer Recherche verwenden und daher ganz anderen Selektionskriterien ausgeliefert sind, als man sich selber vielleicht bei einer fundierten wissenschaftlichen Recherche erfahren würde. Von der rechtlichen Seite her sage ich, was ist, wenn ich sage, GPT, bitte schreibe mir einen Patentantrag. Die schreiben den Patentantrag und ich soll den weiter bearbeiten. Ist da irgendetwas dazu überdacht

01:21:42 **Speaker 2**

worden oder daran gedacht worden, wer kann auf die Daten zugreifen, die ich eingabe in der Annahme, dass die Daten bei mir bleiben. Ich glaube, das wurde vorher kurz angesprochen. Wollen Sie noch einmal quasi, soll man dieses System erfüttern mit eigenen Daten?

01:22:02 **Speaker 2**

Herr Heitzinger, Herr Strauß?

01:22:07 **Speaker 2**

Ich glaube, wir müssen da unterscheiden, wofür die Daten verwendet werden. Das, was OpenAI und andere Hersteller sicher gerne verwenden wird, also schon verwendet bin ich überzeugt und verwenden wird, ist, sie werden sich ansehen, wie zufrieden sind die Benutzer mit den Antworten auf gewisse Fragen und das gibt dann natürlich einen Hinweis, wo man weiter trainieren soll. Davon getrennt sind dann wieder die Trainingsdaten, die tatsächlich verwendet werden, um das System

anzulernen. Die sind dann natürlich spezieller und werden jetzt auch speziell generiert dafür oder von Menschen auch hergestellt, wenn man so will.

01:22:53 **Speaker 2**

Die andere Frage ist dann, wie soll ich mit sensitiven Daten umgehen und das ist die Antwort natürlich, die wird man dort nicht eingehen. Daher dieses Interesse derzeit an lokalen Large Language Models, die man selber laufen lassen kann, die eben keine Verbindung haben zur Außenwelt und dann stellen sich eben viele dieser Probleme nicht. Herr Kettemann, wollten Sie noch was sagen? Also, wenn man im Internet bei einem Programm Daten eingibt, dann sollte man wissen, dass die Daten dann regelmäßig einem nicht mehr ganz alleine gehören. Deswegen unterschreibt man ja die AGBs, wo in Punkt 3.6 drinnen steht, bei der Open AI, dass man dem Unternehmen das Recht einräumt, diese Daten zu verwenden, um das eigene System zu verbessern. Zoom zum Beispiel, das wir alle verwenden, hat vor drei Wochen eine Änderung ihrer AGBs eingeführt mit dem Inhalt, dass sie nun sämtliche Gespräche, die man auf Zoom führt, verwenden dürfen, um via KI ihr System zu verbessern. Sie sagen zwar, das tun sie nicht, aber

01:24:00 **Speaker 2**

sie wollen es ja mal reinschreiben, man weiß ja nicht, was die Zukunft so bringt. Das heißt, ich würde unter keinen Umständen

01:24:08 **Speaker 2**

Patentanträge oder auch nur ansatzweise personenbezogene Informationen in JIPTPT hineingeben, das wird auch jeder Arbeitgeber Ihnen verbieten, entweder hat er es schon oder sollte er es. Man darf nicht zum Beispiel JIPTPT nutzen, um Empfehlungsschreiben zu verfassen oder Arbeitszeugnisse, das geht unter keinen Umständen. Einfach deswegen, weil das datenschutzrechtlich überhaupt nicht klappt. Das ist aber auch generell so, unser Umgehen mit unseren eigenen Daten und den Daten anderer Leute ist manchmal minimal blauäugig. Herr Straußmann? Nur kurz ergänzend, das Problem ist, Sie sind ja personalisiert, Sie sind bei JIPTPT, das heißt, Sie haben schon personenbezogene Daten, deswegen ist es Ihnen vielleicht wichtig, zu sagen, jeder Input, den Sie da generieren, egal, ob das ein Patentantrag ist oder irgendwas anderes, wird sozusagen eingespeist, das Trainingsdatum, also der Trainingsdatensatz wird mit Ihnen erweitert und das ist zwar rechtlich, genauso wie Sie sagen natürlich, die AGBs definieren

01:25:03 **Speaker 2**

das sozusagen, aber ich glaube, das Bewusstsein, das Wissen dafür ist es Entenwollen, dass das tatsächlich so ist. Und das ist eine Grundsatzfrage, wie nutzt man es dann eigentlich? Ich glaube, das ist ein Knackpunkt auch, von wegen Ethik. Okay, die Zeit ist schon sehr fortgeschritten, wir haben hier noch da vorne eine Frage, hinten noch eine Frage und dann werden noch zwei aus dem Chat. Ganz kurze Frage, wieso sollte man einen Unterschied machen zwischen einer schlecht trainierten KI, die sozusagen als Personalchef fungiert und eine rassistisch motivierte Entscheidung trifft und einem schlecht trainierten Personalchef, der auch genau solche Entscheidungen trifft. Mit anderen Worten, der Paul Virilio hat schon gesagt, der, der das Flugzeug erfunden hat, hat auch den Absturz erfunden, also da geht es eigentlich nur um die Systemverbesserung im Bereich der Transparenz, oder irre ich mich da?

01:25:58 **Speaker 2**

Herr Kettmann, Frau Stamm -Lief, wollen Sie? Genau, wir zwingen die KI, Gründe zu geben und wenn diese Gründe rassistisch sind, dann können wir daran arbeiten, Personalchef, Personalchefin, wenn die so blöd sind, rassistische Gründe für eine Entscheidung zu geben, dann kann man auch damit arbeiten, man klagt dann, aber solange die im Forum internum verbleiben, hat man sehr wenig rechtliche Handhabe. Die KI erlaubt uns ganz, ganz, ganz zentrale Fragen über, wie wir entscheiden, was Verstand ist, was wir als Menschen sind, uns zu stellen. Insofern ist die KI nicht nur gut für die Antworten, die sie uns gibt, sondern vielmehr auch für die Fragen, die sie uns verpflichtet zu stellen. Ich würde vielleicht noch kurz sagen, dass man beiden nicht vertrauen sollte, also ich sehe gar keinen Unterschied, also allein von der Problematik her schon, ist ganz klar, nur es gibt schon einen Unterschied jetzt in dem, wie man das dann praktisch löst, während man den Chef als Person zur Verantwortung ziehen

01:26:55 **Speaker 2**

kann, fangen bei der KI all diese Probleme mit der Haftung, Zurechnungsfähigkeit, Autonomie und, und, und an. Und in diesem ganzen Prozess, den wir auch oft nennen bei uns Many -Hands -Theory, also so ganz viele Hände sind da drinnen involviert, kann man sich schnell verheddern und dann ist die Verantwortung und die Zurechnung auch viel schwieriger. Okay, dann noch die Frage von dem Herren im weißen Hemd, bitte. Entschuldigung, wollten Sie noch antworten? Eine ähnliche Frage, und zwar wegen dem Thema Explanatory AI, dass ich Entscheidungsfindungsprozesse so schwer explanieren kann und dass extrem viel Bias in AI ist, die eben Entscheidungen trifft, die uns alle betreffen. Die Alternative, dass diese Entscheidungsfindungsprozesse von Menschen passieren, ist die dann so viel besser und so viel transparenter und leichter nachzuvollziehen. Ist die Entscheidung

eines Polizisten, wenn er raus aus dem Verkehr zieht, die Entscheidung eines Lehrers, wie er bewertet, oder die Entscheidung eines Redakteurs,

01:27:54 **Speaker 2**

welche Themen er pusht oder abserviert, so viel transparenter als ein AI -System, indem ich Trainingsdaten und Weights einfach an dritte Parteien weitergeben kann, die dann Tests durchführen können. Will ich nicht in diesen Bereichen eher etwas Technisches haben, wo ich viel leichter nachvollziehen kann, dass es die Entscheidung war, oder ist das schwerer? Es ist schwerer, weil es sind die Systeme komplexer. Wenn jetzt, weiß ich nicht, ein Beamter einen Fehler macht, dann kann er das irgendwie begründen, oder Sie können mit dem reden, wenn jetzt JGBT oder irgendein anderes KI -System irgendwas produziert, wissen Sie erstens nicht, wie das zustande kommt. Sie haben eigentlich keine Möglichkeit, irgendwas dagegen zu tun. Also da gibt es diesen Fall von diesem Anwalt in den USA, wo der zur Recherche die Gerichtsurteile eingespeist hat, und da ist er in die Falle getappt, dass sozusagen freierfundene Gerichtsurteile reproduziert wurden. Der hat, glaube ich, seine Konzession dann auch verloren

01:28:54 **Speaker 2**

oder wurde zumindest in den Raum gestellt und beteuert, aber er wusste es tatsächlich nicht. Wenn das, glaube ich, ohne AI passieren würde, dann wäre das relativ eindeutig. In dem Fall ist es relativ schwierig nachzuweisen, dass er tatsächlich, es ist seine Behauptung, ob man der Glauben schenkt oder nicht, und die Technologie können Sie eigentlich auf der Ebene nicht so transparent gestalten, dass Sie wirklich nachvollziehen können, wie es zu diesen falschen Informationen kommt, auch was Diskriminierung betrifft beispielsweise. Und der andere Punkt ist, was bedeutet Transparenz eigentlich? Was bringt Ihnen das, wenn Sie jetzt den Source -Code von Chet Chibiti lesen können? Das bringt Ihnen genau nichts. Da kann man jetzt sagen, als OpenAI, das ist total transparent, aber in der Praxis bringt Ihnen das nichts. Also deswegen ist es schon ein Unterschied. Also bei ganz einfachen Systemen kann man das ja schon machen. Wenn man zum Beispiel 20 Bilder hat, 10 von Hunden, 10 von Katzen, und man

01:29:57 **Speaker 2**

trainiert eine KI und unterscheidet zwischen Hund und Katze. Das klappt bei den Katzen und bei den Hunden, aber dann kommt ein Fuchs und dann weiß die KI nicht weiter. Und dann sagt man, okay, dann machen wir 200 mehr Bilder und 1 .000 mehr Bilder und eine Million mehr Bilder und irgendwann einmal schafft das System das auch. Also man kann ja schon erklären. Der Zug ist abgefahren, was Chet Chibiti betrifft oder überhaupt die Chibiti -Models. Da reden wir jetzt nicht

mehr von 20 Bildern, sondern da reden wir jetzt von ein paar Millionen Parametern. Die Datensätze sind im Milliardenbereich. Also ich glaube, da müssen Sie sehr viele Fuchsbilder einspeisen, um da zu zeigen, dass sozusagen der Algorithmus so nicht funktioniert.

01:30:34 **Speaker 2**

Da würde ich gern jetzt zum Abschluss noch zwei Fragen aus dem Chat nehmen. Die erste, die würde ein bisschen anschließen, geht nämlich auch um Transparenz, aber eher im Sinne von, die Frage ist, ist es zu erwarten, in Zukunft einfach zwischen von Menschen generierten Texten und Bildern und von KI generierten Texten und Bildern unterscheiden zu können? Und da kommt für mich auch ein bisschen so die Frage nach dem, inwiefern muss KI bezeichnet werden? Man spricht über Wasserzeichen beispielsweise sehr auf EU -Ebene. Inwiefern sollte es hier Transparenzvorschriften geben, dass man KI quasi labeln muss? Das ist KI. Die gibt es ja schon. Also die gibt es, die sind nur nicht in Kraft. Also dieser europäische AI -Act will genau das, dass man das labelt. Die Frage ist nur, wozu, um dann was zu tun? Wie vor allem, ja. Und wer wird sich dann hinter diesen Labels befinden oder verstecken oder outen als was? Und wenn man einen Jativity -produzierten Text nimmt und ihn dann adaptiert, ist es dann immer

01:31:31 **Speaker 2**

noch ein KI -basierter Text? Also rechtlich gesehen kommen diese Regeln. Das Parlament zumindest möchte das. Also der Rat hat das noch nicht in dem Entwurf drinnen vom KI -Rechtsakt. Mal schauen, ob das drinnen bleibt. Aber technisch, soweit ich das überblicke, ist es nicht unschwierig. Es ist nicht unschwierig, ja. Da wissen Sie ja noch mehr dazu. Okay. Und dann als letzte Frage. Ich fasse hier quasi fünf Fragen zusammen. Und es ist nicht sehr überraschend bei einer Veranstaltung, die vom Bildungsministerium veranstaltet wird. Es geht um die Frage, inwiefern sollte KI in das Thema Bildung Lehrpläne eingebaut werden? Inwiefern müssten die Menschen über das Bildungssystem auf das Thema vorbereitet werden? Wie stark soll es im Unterricht vorkommen? Wie schnell müsste man da auch reagieren? Und wer sollte vielleicht da dann noch so zum Abschluss? Wer möchte gleich mal? Gerne alle. Herr Heitzinger vielleicht? Ja, also verwendet wird es ja ohnehin schon. Also ich würde sagen, mindestens 95 Prozent

01:32:27 **Speaker 2**

aller Schülerinnen und Schüler und Studentinnen und Studenten verwenden es ja mittlerweile. Das heißt, es ist höchste Zeit, dass wir uns Gedanken machen, was sind jetzt wirklich das grundlegende Wissen, die grundlegenden Fähigkeiten, die man ohne KI schaffen soll? Und wo ist der Einsatz dann sinnvoll, eben als Vorbereitung auf später, um unterscheiden zu können, um kritisch damit umgehen

zu können? Das ist sicher eine Diskussion, die wir jetzt schnell führen müssen. Und es gibt eine gewisse Analogie zum Taschenrechner. Damals war es so, wahrscheinlich war das in der fünften Schulstufe, also im fünften Gymnasiumjahr, dürfte man Oberstufe, durfte man dann den Taschenrechner verwenden. Damals hat das halt Mathematik betroffen und Physik und Chemie vielleicht. Mittlerweile betrifft es halt alle Fächer. Das ist der große Unterschied. Also das ist die große Herausforderung. Und davor stehen sicher alle Schulen und alle Universitäten.

01:33:30 **Speaker 2**

Ja, also ich glaube, das ist ganz zentral, wie ich am Anfang schon gesagt habe. Also Wissen und welches Wissen man braucht, um diese Systeme zu verstehen und zu verstehen, was sie eben nicht können. Also Wissen über die Möglichkeit und die realen Grenzen dieser Systeme ist fundamental, dass wir es auch so einsetzen können, dass es sozusagen unterm Strich gesagt mehr bringt, als es Schaden anrichtet. Da gibt es eine ganze Reihe von Risiken, über die man noch sprechen könnte, geht es ja nicht mehr aus, ist klar. Aber ich glaube, es ist wesentlich, dass man so etwas wie eine kritische Medienkompetenz entwickelt im Umgang mit AI, beziehungsweise auch lernt wieder, was es heißt, kritisch zu hinterfragen, wie eine Art von Wissensgenerierung funktioniert, wie sie nicht funktioniert, ob das in einem Anwendungskontext hilfreich ist oder nicht. In der Medizin ist es, glaube ich, ein sehr spannender Bereich, auch weil es ein Expertenzugang ist. Das heißt, Sie haben eine Expertise, die Sie brauchen,

01:34:19 **Speaker 2**

ein Fachwissen. In anderen Bereichen haben Sie ein anderes Fachwissen. Im klassischen Bürojob haben Sie ganz andere Voraussetzungen. Ich glaube, diese Dinge, die muss man tatsächlich neu verhandeln, neu denken und neu konzipieren, auch in der Bildung. Kann man KI im Bildungssystem rauskriegen? Das hat Versuche gegeben, zu verbieten. Glauben Sie? Das glaube ich nicht. Nein, das glaube ich nicht, außer es wird sich von alleine ausfaseln so ein bisschen. Das kann durchaus passieren. Ich finde den Bildungsaspekt extrem wichtig und klar, kritische Medienkompetenz, da sind wir uns, glaube ich, alle einig. Aber für mich ist das ein bisschen ein Weg, der zu autoritär ist, im Sinne von, wir wollen Menschen nur beibringen, wie sie damit umgehen, aber nicht, wie sie es kritisieren oder auch rauskommen. Und das ist für mich ein ganz großes Problem, weil im Moment ist die Anwendung von KI oft undemokratisch einfach implementiert. Es gibt keine Prozesse, die das regeln, wo Menschen mitsprechen können,

01:35:13 **Speaker 2**

Gesellschaften mitsprechen können. Und das halte ich langfristig, egal wie viel man dazu weiß, sollte man nicht einfach nur wissen dürfen und nicht weiter. Ganz kurzer Gedanke dazu, vielleicht führt JGPT oder ähnliche Systeme dann in der Bildung auch zu einer Demokratisierung in dem Sinn, dass sich jetzt jeder seinen persönlichen Nachhilfelehrer leisten kann, der immer zur Verfügung steht. Wir waren heute ein bisschen kritisch, glaube ich, aber das wäre sicher ein positiver Effekt. Herr Kledemann, wollen Sie auch noch? KI muss Bildungsinhalt sein und es ist in gewissen Bereichen sicher auch sinnvoll, es als Bildungsmittel einzusetzen. Meine Schwester ist auch Lehrerin und unter anderem sagt sie mir, was sich halt sehr stark ändert, ist die Art und Weise, wie man unterrichtet, wie man prüft. Man kann zum Beispiel nicht mehr Zusammenfassungen aufgeben, weil die werden alle nur mehr automatisch generiert. Es kommt eine neue Kultur der Mündlichkeit. Referate werden wieder stärker werden. Das

01:36:11 **Speaker 2**

heißt, da müssen unsere Pädagoginnen und Pädagogen mit der Zeit gehen, was sie auch tun. KI ist Teil unseres Lebens und wir müssen die jungen Menschen darauf vorbereiten. Ich glaube, ein perfektes Schlusswort für die Veranstaltung. Ich muss sagen, wir sehen KI ein sehr breites Thema. Ich selber habe die Hälfte meiner Fragen noch gar nicht gestellt, aber ich glaube, es wird uns erhalten bleiben und wir werden noch öfters darüber diskutieren. Ich danke Ihnen fürs angeregte Diskutieren und Ihnen für Ihre Aufmerksamkeit und für die Fragen und darf die Veranstaltung hiermit beenden.