

Digitalgespräch Folge 47

KI und Haftung: Wer steht ein für die Fehler Künstlicher Intelligenz?

Mit Carsten Gerner-Beuerle vom University College London, 13. Februar 2024

<https://zevedi.de/digitalgesprach-047-carsten-gerner-beuerle/>

[Der Vorspann mit Musik und Ausschnitten aus dem Gespräch beginnt]

Marlene Görger [mg]: „Herr Gerner-Beuerle, Sie sind Professor of Commercial Law am University College, London. Sie arbeiten aktuell auch zur Regulierung von KI.“

Carsten Gerner-Beuerle [Gerner-Beuerle]: „Der europäische Gesetzgeber sieht es auch ganz klar, dass es große Schwierigkeiten geben kann, für Personen, die einen Schaden erleiden, die entsprechenden Beweise darzubringen.“ – „Wann ist das System fehlerhaft? Was ist überhaupt ein Fehler? Wenn ein Fehler besteht, dann festzustellen, wer für den Fehler verantwortlich ist.“ – „Das führt dann zu tiefergehenden Fragen darüber, was Verantwortung ist, wer Verantwortung tragen kann, was Schuld ist.“

Petra Gehring [pgg]: „Und wir haben uns einfach noch nicht dran gewöhnt, dass Maschinen inzwischen auch Sprache in einer Wucht und Mächtigkeit in den Alltag hineintreiben können, die tatsächlich gesamtgesellschaftliche Effekte erzeugen kann.“

[Gerner-Beuerle]: „Wenn Risiken gemessen werden, sollen auch Risiken für zum Beispiel den demokratischen Prozess gemessen werden und reduziert werden. Was genau das in der Praxis heißt? Ich habe keine Ahnung. Ich glaube, niemand weiß das.“ – „Das ist im Augenblick alles cutting edge of science.“

[Der Vorspann endet das Gespräch beginnt]

[mg]: „Wenn ein Schaden entsteht, dann stellen wir schnell die Frage, wer daran schuld ist. Und meist hängt damit die zweite Frage zusammen, wer für den entstandenen Schaden haftet. Da muss zum Beispiel geklärt werden, wer an wen welche Summe Geld zu zahlen hat. Im Alltag hilft hier in der Regel die Haftpflichtversicherung. Sie kommt im Zweifel nicht nur für Schäden auf, die wir versehentlich verursachen – ersetzt die kaputte Fensterscheibe der Nachbarin nach dem unachtsamen Fußballspiel oder bezahlt für die Behandlung des gebrochenen Beins nach dem Skiunfall – sondern sie prüft auch, ob wirklich wir haften müssen und nicht vielleicht doch jemand anderes. Auch abstrakte Güter können beschädigt werden, etwa Grundrechte, wenn wir zum Beispiel Diskriminierung erfahren oder zu Unrecht sanktioniert werden. Und nicht nur unsere Gesundheit, sondern auch unseren guten Ruf können andere ruinieren. In der Frage nach der Haftung sind also Ideen von Wert, von Verschulden, Verantwortung und

ausgleichender Gerechtigkeit verbaut. Gerade bei großen Schadensfällen in komplexen Szenarien wird um Haftungsansprüche mitunter hart verhandelt. Man kann sich vorstellen, dass dabei nicht immer ein für alle Betroffenen befriedigendes Ergebnis zustande kommt. Und neue Technologien bringen nicht nur neue Typen von Risiken mit sich, sondern können auch die Komplexität weiter erhöhen. In der letzten Zeit diskutieren wir zunehmend, wie wir Schäden beherrschbar und kalkulierbar machen können, die mit dem Einsatz von sogenannter „Künstlicher Intelligenz“ verbunden sind. Haftungsregeln können dabei auch helfen, Unternehmensverhalten zu steuern, denn für Unternehmen sind Haftungsrisiken ein Motiv, sichere Produkte an den Markt zu bringen. Und vor allem dort, wo wichtige Werte betroffen sind, etwa in der Medizin, der Rechtspraxis oder auf dem Bildungs- und Arbeitsmarkt, wollen wir uns absichern. Was für Schäden können durch KI-Einsatz entstehen? Wie findet man heraus, wer dann haftet? Und was soll die KI-Verordnung der Europäischen Union, also der AI Act, zur Beantwortung dieser Fragen beitragen? Das ist unser Thema heute im *Digitalgespräch*. Mein Name ist Marlene Görger, ich bin Physikerin und Technikphilosophin und arbeite am *Zentrum verantwortungsbewusste Digitalisierung*."

[pgg]: „Und ich bin Petra Gehring, Professorin für Philosophie an der Technischen Universität Darmstadt. Wir haben heute Prof. Dr. Carsten Gerner-Beuerle bei uns zu Gast, er ist per Videokonferenz zugeschaltet aus London. Herzlich willkommen Herr Gerner-Beuerle, und vielen Dank für Ihre Zeit."

[Gerner-Beuerle]: „Vielen Dank für die Einladung."

[mg]: „Herr Gerner-Beuerle, Sie sind Professor of Commercial Law am University College London und haben als Rechts- und Wirtschaftswissenschaftler an etlichen renommierten Institutionen in Europa sowie den USA gelehrt und geforscht. Mit Ihrer Expertise für internationales Privat-, Gesellschafts- und Kapitalmarktrecht haben Sie auch Regulierungsvorhaben der EU-Kommission und des EU-Parlaments unterstützt. Sie sind Mitglied des European Corporate Governance Institute und sie arbeiten aktuell auch zur Regulierung von KI, die ja auf europäischer Ebene mit dem AI Act gerade stark vorangetrieben wird. Wir kommen später noch darauf zu sprechen, ob und wie der AI Act die Haftung betrifft, aber zu Beginn vielleicht noch mal grundsätzlicher: Zu klären, wer haftbar ist, das wird ja dann wichtig, wenn ein Schaden entstanden oder zumindest zu erwarten ist. Wenn man so allgemein von Haftung redet, was kann denn damit alles gemeint sein?"

[Gerner-Beuerle]: „Wenn Sie von Haftung sprechen, dann geht es grundsätzlich darum, dass jemand, der einen Schaden erlitten hat, einen Schadensersatzanspruch geltend macht gegen eine andere Partei, die für den Schaden nach dem Gesetz verantwortlich ist. Schadensersatzanspruch bezieht sich normalerweise darauf, den Zustand herzustellen, der ohne das schädigende Ereignis bestanden hätte. Das ist nicht immer möglich, und dann wird in der Praxis häufig auf Geldersatz geklagt."

[pgg]: „Wenn da jetzt einfach nur Menschen einander gegenüberstehen, kann man sich das relativ gut vorstellen. Also, jemand tut was und schädigt damit jemand anderes. Aber wie ist das, wenn Produkte ins Spiel kommen? Also wenn Hersteller Produkte an den Markt bringen und jetzt das Produkt irgendwie in einen Schadensfall verwickelt ist?“

[Gerner-Beuerle]: „Das Komplizierte an der Fallgestaltung ist ja, denke ich, dass verschiedene Personen beteiligt sind und es nicht immer klar ist, wer für welche Handlung und für welchen Schaden letztendlich verantwortlich ist. Ähm, dass ein Produkt den Schaden verursacht als solches, ist nicht weiter problematisch und vom Recht auch nicht schwierig zu erfassen. Es ist eben dann eine Handlung letztlich – äh, einer Person, etwa des Herstellers des Produktes, ähm, die für den Schaden verantwortlich ist. Also man kann den, die Kausalität kann man immer zurückführen auf eine Person – das mag eine natürliche Person sein oder eine juristische Person – die nach dem Gesetz dann verantwortlich ist und Schadensersatz unter Umständen zu leisten hat. Wie der Schaden entstanden ist, eben zum Beispiel durch ein fehlerhaftes Produkt oder vielleicht durch eine Handlung, ein, ein Schlag oder sowas, wenn sich zwei Menschen gegenüberstehen, das macht dem Prinzipnach keinen Unterschied. Die Probleme entstehen dadurch, dass, wenn ein Produkt etwa fehlerhaft ist, dass es dann meistens schwieriger wird, die Kausalität nachzuweisen, also für denjenigen, der den Schaden einklagt, genau darzulegen vor Gericht, wenn es dazu kommt, dass eine andere Person verantwortlich ist. Wenn Sie von Produkthaftung sprechen, dann kann man sofort daran denken, dass es eben nicht nur den Hersteller des Produktes gibt. Der Hersteller, der bezieht sich dann vielleicht auf weitere Lieferanten, äh, die bestimmte Teile zum Produkt bereitstellen, die dann eingebaut werden und vielleicht dann auch nicht so gut überprüft werden oder nicht so gut überprüft werden können von dem Hersteller, weil dem Hersteller dann die technische – das technische Wissen da auch fehlt. Und der Hersteller sendet das Produkt dann weiter an jemanden, der es verkauft letztlich, sodass, dass wir eine Kette von Personen haben und es nicht immer einfach festzustellen ist, wo in dieser Kette jetzt der Fehler entstanden ist und wer dementsprechend für den Fehler verantwortlich ist.“

[mg]: „Jetzt kommt ja durch den Einsatz von KI dann noch mal erschwerend hinzu, dass unter Umständen auch Prozesse passieren, wo jetzt gar kein Mensch mehr so richtig kausal drinsteckt. Ist das richtig so, oder stimmt das eigentlich noch nicht?“

[Gerner-Beuerle]: „Das kann man vielleicht so sehen [lacht], aber letztlich sind Menschen ja doch beteiligt am in den Umlauf bringen eines KI-Produktes oder eines KI-Systems. Also ich, ich denke hier stellen sich zwei Probleme, die rechtlich zu unterschiedlichen Fragen führen und in der Praxis dann auch unterschiedliche Probleme stellen. Das erste Problem oder die erste Frage ist, inwieweit das KI-System selber verantwortlich gemacht werden sollte. Da gab es eine Diskussion – oder es gibt immer noch eine Diskussion darüber, ob ein KI-System Rechtspersönlichkeit erhalten soll, sodass es verklagt werden kann. Die Idee dahinter ist, dass das KI-System selbstständig Entscheidungen trifft. Das tut es vielleicht so im gewissen Rahmen, aber letztlich, äh, handelt das KI-System ja so, wie es programmiert ist. Also die Prämisse ist

auch – kann man schon hinterfragen, aber auf jeden Fall: Die Idee hier ist, dass das KI-System selbstständig Entscheidungen trifft und dass eben das System dann der richtige Gegner in einer Klage sein soll. Das – ist nicht unbedingt eine überzeugende – ähm – Sichtweise, denke ich, und es ist auch vom Gesetzgeber nicht weiterverfolgt worden. Aber das ist die erste Frage, die sich da stellt. Und die zweite Frage, die auch mit Ihrer Frage zusammenhängt, Frau Görger, ist, dass, wenn das KI-System besonders kompliziert ist, was häufig KI-Systeme ja sind, und vielleicht sich dann auch selber weiterentwickelt, also wenn es – wenn wir über ein Machine-Learning-System sprechen, dann ist es häufig sehr schwierig nachzuvollziehen, was der Grund ist für eine, sagen wir, fehlerhafte Klassifizierung, oder was auch immer das System macht. Es ist auch schwierig, überhaupt erst mal festzustellen, was eine fehlerhafte Klassifizierung ist, weil ein gewisser Error-Margin immer dabei ist, wenn ein Algorithmus eine Prediction trifft. Das Problem hier ist, festzustellen: Wann ist das System fehlerhaft. Was ist überhaupt ein Fehler? Und wenn ein Fehler besteht, dann festzustellen, wer für den Fehler verantwortlich ist. Also wenn wir nicht von einem KI-System als einer Rechtsperson sprechen und KI-Systeme sind nach derzeitiger Rechtslage keine Rechtsperson, dann stellt sich die Frage, welche Person, natürliche oder juristische Person, hinter dem KI-System verantwortlich ist. Und das kann noch schwieriger sein als bei Produkten, anderen Produkten, mit denen wir – mit denen wir vertrauter sind vielleicht, einem Fernseher oder so zum Beispiel."

[pgg]: „Kann ich noch mal kurz zu der Frage, ob KI selbst haften können sollte? So, auf so einer ganz abstrakten Ebene gibt es ja viel so Menschähnlichkeitsdiskussionen über KI. Schon das Wort legt das so ein bisschen nahe. Aber sobald wir uns vorstellen, dass wir mit KI-Produkten oder KI-Systemen zu tun haben, klänge das wie ne, äh, ein kompletter Umbau unseres Rechtssystems, wenn man plötzlich ein nichtmenschliches System, weil es so was Ähnliches wie Intelligenz an den Tag legt, haftbar machen würde.“

[Gerner-Beuerle]: „Ja, das führt dann zu tiefergehenden Fragen darüber, was Verantwortung ist, wer Verantwortung tragen kann und was Schuld ist. Viele Haftungsvorschriften verlangen ja einen Nachweis von Verschulden, und nicht alle können schuldhaft handeln. Auch nicht alle Menschen können schuldhaft handeln, Babys zum Beispiel sind nicht schulfähig. Die Diskussion überlasse ich aber lieber Ihnen [lacht], was Verantwortung ist und was Schuld bedeutet und wer schuldhaft handeln kann. Aber die, die praktischen Probleme sind ja auch ganz klar im Vordergrund hier. Wenn ein KI-System haftbar wäre, dann müsste sichergestellt werden, dass das KI-System als Rechtsperson vernünftig kapitalisiert ist, um den Haftungsanspruch auch bedienen zu können, sonst nützt es wenig [lacht], das KI-System verklagen zu können. Das heißt, letztlich müssen dann doch wieder Personen im Hintergrund, die hinter dem KI-System stehen und die wahrscheinlich einen Gewinn machen dann von dem KI-System, müssen die eben dann in Verantwortung genommen werden.“

[mg]: „Ich glaube es lohnt sich, mal an konkrete Beispiele zu gehen, um zu zeigen, was es vielleicht auch schwierig macht, dann im Feld der Produkthaftung auch weiterzukommen. Das scheint ja schon mal die deutlich einfachere oder

naheliegendere Diskussion zu sein und ich kann mir aber vorstellen, dass es da ganz unterschiedliche Fälle gibt und vielleicht auch in unterschiedlichen Bereichen unterschiedliche Eigenschaften von KI zu Problemen führen. Kennen Sie denn Beispiele, wo Produkthaftung zu KI unklar ist und deswegen vielleicht auch Unsicherheiten beim Einsatz bestehen?"

[Gerner-Beuerle]: „Also zurzeit ist es unklar, ob Produkthaftung überhaupt Anwendung findet auf KI. Da gibt es Meinungen zu und auch Urteile, die nicht ganz eindeutig sind. Deswegen haben die europäischen Institutionen auch vor ungefähr zwei Jahren einen Gesetzesvorschlag auf den Weg gebracht, der die Produkthaftungsrichtlinie – ich glaube, so heißt die auf Deutsch, ich bin mir nicht ganz sicher, Product Liability Directive – der die Produkthaftungsrichtlinie reformieren wird. Das Gesetzesvorhaben ist noch nicht verabschiedet, wird aber wahrscheinlich verabschiedet werden, denke ich. Im Augenblick sind Diskussionen in den europäischen Institutionen da zu den Einzelheiten – ähm – des Gesetzesentwurfs und wie er an den KI, an das KI-Gesetz dann auch angeglichen werden kann, sodass wir dann auf europäischer Ebene ein harmonisches Gesetzeswerk haben. Also bislang war es unklar, ob Software, ein KI-System eben, überhaupt unter den Begriff „Produkt“ fällt. Wenn ein KI-System unter den Begriff „Produkt“ fällt, dann ist es möglich, die Produkthaftungsrichtlinie oder das entsprechende Umsetzungsgesetz in den EU-Mitgliedsstaaten anzuwenden, und der Hersteller des Systems würde dann haften für Fehler, die eintreten. Mir sind keine Fälle bekannt, konkret zumindest in, äh, aus Großbritannien, in denen Produkthaftung eingeklagt worden wäre. Das mag in EU-Mitgliedsstaaten anders sein – ich kenne nicht unbedingt die Rechtslage in der gesamten EU natürlich – aber mir sind keine Fälle bekannt, ich glaube, die Gerichtsurteile sind hier relativ selten, eben weil es bislang unklar ist, ob KI-Systeme unter den Begriff „Produkt“ fallen. Und das, das Gesetzesvorhaben, das auf europäischer Ebene jetzt diskutiert wird, das würde das klarstellen, dass sie eben darunterfallen.“

[mg]: „Woran liegt das? Es ist ja eigentlich jetzt mal so aus einer Laiensicht irgendwie klar, dass das ein Produkt ist, ne? Ich kann da was kaufen, ich kann mir nen Profi-Zugang zu *ChatGPT* kaufen, so, dann erwarte ich eigentlich: das ist ein Produkt. Wie ist das begründet, dass das nicht klar ist, dass das ein Produkt ist?“

[Gerner-Beuerle]: „Bislang die Produkthaftungsrichtlinie bezog sich auf Objekte, die – also handfeste Objekte im Prinzip – die verkauft werden können. Und Software wäre dann unter Umständen ein Bestandteil. Aber es war eben strittig, ob Software der gesetzlichen Definition von Produkt entspricht. Aber das wurde vertreten und ich denke, es ist auch nicht fernliegend zu sagen: „Ja, man kauft Software, genauso wie man andere Produkte kauft, und dann ist das KI-System eben ein Produkt.“ Also wenn Software irgendwo integriert ist und man kauft dann zum Beispiel einen Computer mit Software, dann ist auch keine Frage, dass die Produkthaftung Anwendung findet. Die schwierige Frage ist, ob Software alleine, äh, als Produkt zählt.“

[mg]: „Hmhm. Das heißt, wenn ich ein Gerät kaufe, auf dem ist schon irgendeine Software vorinstalliert, dann ist da dieses Gesamtpaket ein Produkt.“

[Gerner-Beuerle]: „Ja.“

[mg]: „Aber wenn ich auf mein eigenes Gerät Software installiere, die ich anders vorher bezogen habe, dann ist schon sozusagen beides auch irgendwie aufgebrochen. Dann ist mein Computer jetzt ja auch nicht mehr in seinem Originalzustand [lacht].“

[Gerner-Beuerle]: „Genau.“

[pgg]: „Vielleicht da noch mal nachgefragt: dieses nochmal Veränderbare, Bewegte, das Softwarelösungen unter Umständen sowieso haben, aber KI in ganz besonderer Weise, sofern gelernt wird: Das Programm adaptiert sich...“

[Gerner-Beuerle]: „Hmhm.“

[pgg]: „... und es adaptiert sich an Eingabedaten oder an Daten, die es sonst wie nutzt, auf die es zugreift. Und insofern ist es vielleicht zum Punkte Null der Inbetriebnahme in einer bestimmten Weise konfiguriert, so dass der Hersteller sagen kann: „Ja klar, das ist mein Produkt, aber im Laufe der Verwendung verändert es sich und verändert es möglicherweise ja auch sein Verhalten und seine Schädigungspotenziale.“ Wie fasst man denn diese dynamische Seite von KI-Produkten in den Haftungsgedanken rein?“

[Gerner-Beuerle]: „Ja, das ist noch nicht ganz klar. Im ersten Gesetzesentwurf der Kommission aus dem Jahr 22 stand da, glaube ich, noch nichts zu drin. Aber das ist eben eine der Fragen, die nach der alten Produkthaftungsrichtlinie nicht wirklich zu beantworten waren. Also: Die Produkthaftungsrichtlinie in ihrer Originalfassung verlangt, dass der Fehler, der Defekt, zum Zeitpunkt des Inverkehrbringens vorliegt. Das ist nicht unbedingt ja der Fall bei einem KI-System, und trotzdem könnte man sich Situationen vorstellen, wo es sachgerecht ist, den Hersteller dann in Haftung zu nehmen. Also ein, wie Sie sagen, ein KI-System kann ja lernen und bekommt neue Daten zugespeist. Und auf der Basis werden dann die Voraussagen, die das KI-System trifft, eben genauer oder vielleicht auch nicht genauer, je nachdem, was es für Daten bekommt und je nachdem, in welchem Zusammenhang es eingesetzt wird. Man könnte sich jetzt vorstellen, dass etwa der Benutzer des KI-Systems keine klaren Anweisungen bekommen hat, was die Parameter eines Datensatzes sein müssen, so dass der dann geeignet ist für das bestimmte KI-System und für den, den Use-Case, in dem es eingesetzt werden soll. Und wenn der, der Benutzer, die – diese Informationen nicht hat, und es falsch einsetzt, dann ist der Fehler vielleicht auch dem Hersteller anzulasten. Das ist nach der derzeitigen Fassung der Richtlinie ist das schwer zu begründen, weil der Fehler eben zum Zeitpunkt des Inverkehrbringens noch nicht vorhanden war. Das soll jetzt aber auch in die Richtlinie rein, zumindest besprechen die europäischen Institutionen, inwieweit es dann – oder unter welchen Voraussetzungen es vertretbar ist, den Begriff des Fehlers weiter auszulegen.“

[pgg]: „Sodass es vielleicht, um noch mal wieder so eine, vielleicht ein bisschen platte Analogie zu benutzen, ähm – sowas wie so ein allmähliches giftiger Werden irgendeiner KI-Lösung geben kann, die am Anfang schon harmlos war, aber dann sich durch Daten oder durch irgendwelche Maschine-Maschine-Prozesse, die möglicherweise der Nutzer überhaupt nicht mitgekriegt hat oder schon gar nicht vorhersehen konnte, verändert zum Unguten.“

[Gerner-Beuerle]: „Ja. Das muss natürlich zurückzuführen sein auf den Hersteller, das Problem, und damit sind wir bei anderen Herausforderungen, die wir ja schon angesprochen haben, dass eben in diesem Beispiel mehrere Personen beteiligt sind an dem Inverkehrbringen des – und dem Benutzen des KI-Systems. Und bei einem Schadensfall müsste dann, müsste dann festzustellen sein oder festgestellt werden, ob der Schaden darauf zurückzuführen ist, dass der Hersteller ein ungeeignetes System in Verkehr gebracht hat, das dann, als es sich weiterentwickelt hat, eben vielleicht an Präzision verloren hat, anstelle an Präzision zuzugewinnen, oder ob es der Benutzer in einem ungeeigneten Zusammenhang verwendet hat.“

[mg]: „Hmhm. Gibt es denn schon Beispiele für Schäden, die durch KI-Systeme verursacht wurden?“

[Gerner-Beuerle]: „Also ich weiß nicht, ob es Beispiele gibt, die dann tatsächlich eingeklagt worden sind. KI-Systeme werden eingesetzt ja bereits, aber selten ohne menschliche Kontrolle. Also im medizinischen Bereich ist es häufig so, dass KI-Systeme eingesetzt werden, um eine Diagnose zu erstellen etwa, soviel ich weiß, passiert das nicht völlig ohne Aufsicht eines Doktors, sodass dann automatisch immer eine natürliche Person Verantwortung trägt und die Haftungsfragen dann einfacher zu beantworten sind. Ein Beispiel aus den USA, wo KI-Systeme ohne weitere menschliche Zwischenstufe Entscheidungen getroffen haben, die signifikant sind, sind KI-Systeme, die bei einer Jobsuche verwendet werden. Also die, die verwendet werden von einem Arbeitgeber, um Kandidaten auszuwählen für ein Interview, für ein Bewerbungsgespräch. Und da ist, wenn ich das richtig verstehe, in den USA: Manche Unternehmen haben bei dieser ersten Stufe, wo die die Vorselektion getroffen wird, keine Person im Spiel. Das KI-System trifft also die letztlich relevante Entscheidung. Und wenn das KI-System dann vernünftig funktioniert und es wurden, das da wurden Fälle gebracht, wo die Arbeitssuchenden, die keine Einladung zum Bewerbungsgespräch erhalten haben, dann sagten, dass – dass gegen sie diskriminiert worden sei. Das ist ja, wie Sie sagten, am Anfang auch eine Art von Schaden. Wenn mir wegen eines KI-Systems ein Job entgeht, äh, dann stellt sich in der Tat die Frage der Haftung. Und in den USA ist das auch weit, äh, besprochen worden und das ist dann auch ein überzeugendes Argument gemacht worden, dass manche KI-Systeme eben in der Tat diskriminiert haben, weil sie nicht vernünftig kalibriert waren oder weil das Training-Dataset schon einen, äh, inhärenten Bias hatte. Und dann, soviel ich weiß, sind da auch tatsächlich, äh, ist da Schadensersatz zugesprochen worden.“

[pgg]: „Man muss wahrscheinlich aber klagen, um dann wirklich in die Details dieser softwarebasierten Entscheidung auch Einblick zu bekommen. Denn, äh, so aus der – aus der Ferne wird man nicht genau sagen können, wie der Algorithmus jetzt funktioniert hat und warum man nicht eingeladen worden ist. Also vermutlich muss man den ganzen Weg eines Gerichtsverfahrens gehen, um rauszukriegen, ob da wirklich genau so verfahren wurde und ob die Software angreifbar ist.“

[Gerner-Beuerle]: „Ja, in der Tat. Dann muss dann das System offengelegt werden, und Algorithmen sind ja – solche Systeme sind, äh, sind ja Eigentum und häufig sind, sind sie Business Secrets, die die Unternehmen, die sie benutzen, nicht gerne offenlegen. Also das müsste dann unter Umständen eingeklagt werden. Der europäische Gesetzgeber – da Sie auf diesen Punkt ansprechen – der europäische Gesetzgeber sieht es auch ganz klar, dass es große Schwierigkeiten geben kann für Personen, die einen Schaden erleiden, die entsprechenden Beweise darzubringen. Das technische Wissen wird häufig fehlen, aber dann wird auch schlicht der Zugang zu den notwendigen Daten und zu dem System selber fehlen und die Gesetze, die jetzt in der Pipeline sind in Europa, die sehen vor, dass, wenn derjenige, der einen Schaden erlitten hat, anderweitig nicht die Beweislast bringen kann, dass das Gericht dann verlangen kann, dass tatsächlich Dokumentation und der Algorithmus selber auch offengelegt wird.“

[pgg]: „Noch mal eine Frage – äh, inwieweit können sich Hersteller oder vielleicht auch in dem Fall so Arbeitgeber, die so Lösungen einsetzen, äh, durch so was wie Allgemeine Geschäftsbedingungen, irgendwelche Dinge, die man vorher unterschreiben muss, da absichern und quasi den Haftungsfall ausschließen? So nach dem Motto: „Wir setzen KI ein, wir informieren Sie darüber, wir können letztlich nicht ausschließen, dass Folgendes passiert. Ähm – Wenn Sie unterschreiben, haben Sie da eingewilligt, dass das in Ordnung ist und dass Sie damit dann eben leben.““

[Gerner-Beuerle]: „Ja, in gewissen Grenzen geht das. Die Produkthaftungsrichtlinie und ein anderer Gesetzesentwurf in Europa, die sogenannte IA Liability Directive, also eine weitere Richtlinie, die außerhalb des Bereichs Produkthaftung die Haftungsregeln in den Mitgliedsstaaten dann umändern würde, damit sie, damit sie anwendbar sind in diesem Kontext. Ähm – diese Gesetze würden AGB überschreiben. Also das ist bindendes Recht dann und davon kann man nicht vertraglich abweichen. Wo diese Gesetze nicht Anwendung finden, können Hersteller und diejenigen, die KI-Systeme benutzen, dann versuchen, ihre Haftung einzuschränken, und das, soweit ich das beurteilen kann, geschieht das auch in der Praxis. Da versucht natürlich jeder, der ein KI-System benutzt, und inzwischen sind sie ja – werden sie ja häufig benutzt, also nicht ohne menschliche Überwachung, aber sie tragen bei zu vielen Funktionen – ähm – da wird dann versucht, vertraglich eben die Haftung zu beschränken.“

[mg]: „Wissen Sie, ob es Versicherungen gibt, die da selbst reagieren, die zum Beispiel bestimmte Haftungsrisiken, die auf KI dann basieren, ausschließen aus ihrem – ihren Policen. Ist Ihnen da was bekannt?“

[Gerner-Beuerle]: „Ich weiß nicht, ob es dazu speziell Versicherungsklauseln gibt schon, allgemein gesprochen: Versicherer werden bestimmtes Handeln ausschließen. Bestimmtes Handeln wird nicht versichert sein, zum Beispiel was mit Vorsatz erfolgt, unter Umständen auch mit grober Fahrlässigkeit. Und normalerweise bezieht sich der Versicherungsschutz dann auf Handlungen, die, die fahrlässig vorgenommen werden.“

[pgg]: „Jetzt haben Sie eben schon angedeutet: Das ist ja in der Praxis oft so eine ganze Kette von Akteuren, die irgendwie digitale Lösungen nutzt und wenn jetzt so ganz normale Leute die gut zugänglichen KI-Lösungen, die man im Netz so bekommen kann, nutzen, dann – ähm – nutzen sie einerseits – ja – eine Lösung, die vielleicht bald ein Produkt sein wird, also sagen wir mal *OpenAI's*, *ChatGPT* und Systeme, die darauf aufsetzen. Wenn ich jetzt als Nutzerin eines solchen Systems, sagen wir mal generative Text-KI eingesetzt habe und dann irgendein Text selber in einer Situation verwende, in der ich Verantwortung übernehme, dann kann es ja sein, dass Fehler, die aus der generativen KI stammen, dann von mir weitergereicht werden, sage ich jetzt mal so, und umgekehrt: Wenn jetzt jemand betroffen ist davon, dass diese Fehler drin saßen in dem Text, den ich abgeliefert habe...“

[Gerner-Beuerle]: „Hmhm.“

[pgg]: „...stellt sich auch schon wieder eine Haftungsfrage. Bin ich da im Zweifel immer dran, wenn ich so gut zugängliche Lösungen aus dem Netz nutze, weil ich quasi alle Risiken übernehmen muss, die da unbekannterweise in diesen riesigen Sprachmodellen drinstecken und den – unfassbar komplizierten Prozessen, die mir da meinen Text geliefert haben, oder kann ich das auch irgendwie abwälzen und sagen: „Ja gut, ich bin ja jetzt nun auch einfach nur eine ganz normale Nutzerin?““

[Gerner-Beuerle]: „Hmhm. Ja. Also im Prinzip sind Sie dran [lacht], wenn Sie das machen. [mg lacht] Die, die genauen Voraussetzungen, die zur Haftung führen können, die müssten dann natürlich noch genauer überprüft werden. Also Produkthaftung: Vor Produkthaftung müssen Sie keine Angst haben, wenn Sie als Nutzer *ChatGPT* benutzen und das zu einem Schaden führt bei irgendjemand anderem aus irgendwelchen Gründen. Die Produkthaftung bezieht sich auf Hersteller, also diejenigen, die das professionell machen, Unternehmen, die, die ein Produkt, um Gewinn zu machen, in den Verkehr bringen. Allgemeine Haftungs Vorschriften finden aber Anwendung, also das ganz normale Schadensrecht, etwa in Deutschland nach dem Bürgerlichen Gesetzbuch, das findet Anwendung. Und dann stellt sich die Frage, was Sie hätten wissen müssen. Also, wenn es eine Verschuldenshaftung ist – meistens ist es nach deutschem Recht ja eine Verschuldenshaftung im Deliktsrecht – dann stellt sich die Frage, was Sie hätten wissen müssen, und es ist allgemein bekannt, dass *ChatGPT* nicht fehlerfrei ist, sondern halluziniert, wie alle – alle Large-Language-Models, ja? Ähm – wenn Sie sich darauf verlassen also, dann geschieht das aus eigenem Risiko. Und wenn Sie es verwenden, um andere Personen, anderen Personen Informationen zu übermitteln oder so, und dann entsteht da ein Schaden, dann würden Sie also in der Tat haftbar sein, wenn Sie fahrlässig handeln. Eine andere wichtige Frage, die zwischen dem Europäischen

Parlament und der Kommission und den Mitgliedstaaten jetzt im Rahmen der Verabschiedung des KI-Gesetzes ja auch ziemlich heiß debattiert wurde, ist, inwieweit Hersteller wie zum Beispiel *OpenAI* haftbar sein sollen. Im Augenblick ist es schwer zu begründen, denke ich, die haftbar zu machen nach allgemeinem Deliktsrecht. Sagen wir mal, sie benutzen *OpenAI*, um einen Text zu fabrizieren und dann verlässt sich jemand anderes auf den Text. Um ein konkretes Beispiel zu nennen, sagen wir mal: Sie benutzen *OpenAI*, um zu analysieren, was gute Investmentoptionen sind und dann benutzt es jemand, um darauf basierend eine Investition zu treffen im, sagen wir mal, am Aktienmarkt oder so und verliert Geld, ist es sehr schwer zu begründen, dass *OpenAI* für so etwas haftbar sein sollte nach allgemeinem Deliktsrecht. Interessanter wird es, wenn da das KI-Gesetz in Kraft tritt, da das KI-Gesetz sich ja an Hersteller nicht nur generell von KI-Systemen, sondern speziell von generativen KI-Systemen – also was im englischen Text *Foundation Models* heißt – wendet und alle möglichen Compliance-Requirements diesen, diesen Herstellern aufgibt. Also interessante Fragen werden sich stellen, wenn das Gesetz in Kraft tritt und dann jemand das Argument bringt, dass zum Beispiel *ChatGPT* nicht ausreichend akkurat ist, so akkurat wie es das KI-Gesetz verlangt, dann würde im Prinzip eine Haftung von *OpenAI* möglich sein."

[mg]: „Was heißt dann in dem Fall akkurat? Also gerade die Ergebnisse jetzt von so nem generativen Modell hängen ja maßgeblich auch davon ab, was für eine Anfrage ich reingebe. Die müsste dann ja auch akkurat genug sein, sage ich mal."

[Gerner-Beuerle]: „Ja, das frage ich mich auch, das ist eine gute Frage. Das Gesetz selber, also das KI-Gesetz, das definiert es nicht weiter, das stellt keine, das stellt keine quantifizierten Thresholds auf hier, kann es auch nicht, denn was ausreichend akkurat ist, ich denke, man kann überzeugend sagen: Das hängt von dem Zusammenhang ab, von der Situation ab und von dem potenziellen Schaden. Also es ist – hängt vom Risiko ab und vom Nutzen, den, den man bekommt, indem man ein, so ein System einsetzt. Also es kann nicht allgemein beantwortet werden, was ausreichend akkurat ist, das KI-Gesetz, das formuliert eben so, wie ich es gesagt habe. Es muss, es muss ausreichend akkurat sein. Also das ist jetzt meine Übersetzung, der deutsche Text mag das etwas anders formulieren. Aber, äh, also solche vagen Voraussetzungen stellt das KI-Gesetz auf und die müssen dann implementiert werden. Ich denke, was als nächstes kommen wird, ist, dass Standard-Setting Bodies, also zum Beispiel in Deutschland TÜV oder so – ähm – genauere Richtlinien aufstellt, herausgeben wird. TÜV arbeitet auch daran, die haben so eine Division, die sich mit solchen AI-Haftungsfragen beschäftigt oder mit Performance Fragen von AI. Solche, solche Standard-Setting Bodies werden dann genauere Richtlinien aufstellen und auf diese Weise werden sich Industriestandards herausbilden."

[mg]: „Und, äh, Sie haben ja auch eine internationale Perspektive, auch auf unterschiedliche Rechtssysteme. Ist Ihr Eindruck, dass verschiedene Rechtssysteme unterschiedlich gut auf diese KI-Problematik oder diese KI-Besonderheiten vorbereitet sind? Also lässt sich zum Beispiel im britischen Recht das Aufkommen solcher Systeme leichter integrieren als etwa im Deutschen?"

[Gerner-Beuerle]: „Also, wenn Sie fragen, ob Gesetzgeber in verschiedenen Ländern unterschiedlich gut vorbereitet sind, würde ich sagen: Ja. Wenn Sie fragen, ob manche Rechtssysteme besser dazu geeignet sind, auf KI zu reagieren als andere, das ist eine schwierige Frage, ich denke nein. Die Antwort ist nein. Es gibt nichts im englischen Recht, was das englische Recht jetzt grundsätzlich geeigneter machen würde, um mit diesen Herausforderungen umzugehen als etwa das deutsche Recht. Was man sagen kann, ist, dass Gesetzgeber und Politiker, wenn man sich das anschaut, was passiert in der Welt, doch schon sehr unterschiedliche Ansätze verfolgen, um KI zu regulieren. Europa, denke ich, ist weit vorne im Hinblick darauf, KI detailliert zu regulieren. Es gibt sonst nirgendwo in der Welt etwas, was genauso umfassend und detailliert wäre wie das KI-Gesetz in Europa und mit dem KI-Gesetz zusammen eben etwa die, die Revision der Produkthaftungsrichtlinie, die ich erwähnte, die aber, wie gesagt, noch nicht verabschiedet sind, aber verabschiedet werden, denke ich, in naher Zukunft. Ob das gut ist oder nicht, ist eine andere Frage. Gut, von der Perspektive der Industrie und gut von der Broader Social Perspective, also ob es gut ist für die Gesellschaft oder so, das ist eine andere Frage. Aber es ist auf jeden Fall ein klar anderer Ansatz als der, der in Großbritannien verfolgt wird oder in den USA. In Großbritannien gibt es keine, bislang keine konkreten Gesetze und ich denke, es wird auch keine konkreten Gesetze in naher Zukunft geben, sondern der Ansatz hier ist, das Wissen und das Verständnis erstmal von KI-Systemen zu unterstützen und zu entwickeln. Dazu hat der Premierminister vorgeschlagen, als wir dieses AI-Summit, Safety-Summit, hier hatten, letztes Jahr, ein, ein AI-Safety-Institute zu errichten, was jetzt errichtet wird in Großbritannien und grundsätzlich sind auch andere, andere Institute und Universitäten sind auch sehr aktiv hier, sich darüber Gedanken zu machen, was die Risiken sind von KI und was die richtige Antwort ist, um diese Risiken zu reduzieren, aber es gibt keinen umfassenden Gesetzesentwurf. Und abgesehen von diesen Programmen der Förderung von Wissen und Verständnis wird die konkrete Regulierungsaufgabe den existierenden äh, Regulatory Bodies übertragen, also je nach Sektor dann zum Beispiel der Finanzmarkt-Regulierungsbehörde.“

[pgg]: „Es wird aber dann doch geregelt. Ich hatte jetzt im Moment gedacht Sie sagen: „Man wartet einfach ab und die Gerichte werden das dann im Einzelnen ausbuchstabieren.“ Aber Sie sagen, es gibt dann doch so was wie feldbezogene Verordnungen und so eine Art Kontrollregime?“

[Gerner-Beuerle]: „[Lacht] Also der Plan ist, kein sektorübergreifendes Gesetz zu erlassen, wie das in Europa jetzt geschehen ist mit dem KI-Gesetz, sondern den entsprechenden Regulierern der, der jeweiligen Sektoren es zu überlassen, Maßnahmen zu treffen. Die wissen aber nicht unbedingt, was die beste Herangehensweise ist. Also da gibt es, da gibt es mehrere Probleme mit diesem Ansatz. Die Regulierer wissen erstens nicht unbedingt, was sie machen sollen. [lacht] Es ist also noch nicht viel geschehen. Soweit ich das sehen kann, noch gar nichts. Zweitens fehlt den Regulierungsbehörden die Kompetenz unter Umständen oder zumindest die Ressourcen fehlen denen, um effektiv zu regulieren. Und drittens: Wenn es tatsächlich

zur Regulierung kommen sollte irgendwann, stellt sich das Problem, dass wir dann ein sehr fragmentiertes System haben, das nicht unbedingt gut zusammenarbeitet. Also das ist auch kein idealer Ansatz [lacht]. Ist es ein besserer Ansatz als das, was in der EU passiert? Vielleicht. Zumindest sind die Kosten für AI-Companies geringer."

[mg]: „Wenn wir gerade bei den AI-Companies sind: Die Diskussionen, die Sie zu Regularien oder Regulierungsbemühungen verfolgen – inwieweit spielen da Überlegungen zu Vorsatz oder auch Fahrlässigkeit eine Rolle? Also, wir kennen ja Äußerungen berühmter KI-Entwickler, offene Briefe, in denen das eigene – ich nenn's jetzt mal Produkt – als besonders risikobehaftet dargestellt wird, als maximal menschenheitsgefährdend. Ähm – Nimmt man das dann ernst bei solchen Regulierungsbemühungen, redet man da dann auch irgendwie von Vorsatz oder zumindest von Fahrlässigkeit, wenn solche Produkte entwickelt werden? Weil sie werden dann ja entwickelt in dem Bewusstsein, da was – oder in der Behauptung, was besonders Gefährliches zu machen."

[Gerner-Beuerle]: „Ich glaube nicht, dass die Unternehmen, die sie dann auf den Markt bringen, zum Beispiel *OpenAI*, tatsächlich die Intention haben, jemanden zu gefährden. Zumindest wäre das grob unternehmensschädigend, wenn sie das sagen würden. Was Sam Altman sagt und andere, ist, dass es das Potenzial gibt, zu großen Schäden zu führen, dass diese KI-Systeme unter Umständen zu großen Schäden führen können, wenn sie ohne weitere Vorkehrungen, ohne Regulierung dann auf die Menschheit losgelassen werden. Aber ich glaube nicht, dass das irgendjemand will und die, die Produkte, die KI-Systeme, die zurzeit auf dem Markt sind: Ich glaube nicht, dass irgendjemand behauptet, dass die tatsächlich solche Risiken bereits hervorrufen. Also, wir reden hier über zwei verschiedene Sachen, denke ich. Was das Potenzial ist, das mag gewaltig sein, da mag das Risiko sehr hoch sein, ich kann das nicht wirklich beurteilen, und auch, und auch Computer Scientists sind da nicht einer Meinung, ja? Und was im Augenblick auf dem Markt ist, da gibt es auch Risiken, die *ChatGPT* hervorruft. Existenzielle Risiken sind es nicht, glaube ich. So sehe ich das zumindest. Und inwieweit tatsächlich Regulierung hier notwendig ist und inwieweit Haftungsvorschriften hier zum Einsatz gebracht werden müssen, das ist eine Frage, die man debattieren kann."

[mg]: „Was wären denn, wenn Sie das jetzt so andeuten, so typische Risiken, die Sie vor Augen haben?"

[Gerner-Beuerle]: „Zum Beispiel mit *ChatGPT* jetzt?"

[mg]: „Hmhm."

[Gerner-Beuerle]: „Also, was ein großes Problem ist, gesellschaftlich ein großes Problem ist, was auch von dem KI-Gesetz angesprochen wird, ist, dass *ChatGPT* und andere Large Language Models dazu verwendet werden können, um Misinformation, äh, zu kreieren, mit sehr geringen Kosten zu kreieren und in Umlauf zu bringen, sodass der Diskurs – und hier denke ich, haben viele den politischen Diskurs vor allem im Kopf – geschädigt

werden kann. Und es ist ja ausgiebig besprochen worden, was das zum Beispiel für demokratische Wahlen bedeuten kann. Das KI-Gesetz versucht das zu regulieren, indem ein paar Vorschriften sagen, dass, wenn Risiken gemessen werden – Risiken eines Systems müssen gemessen werden und müssen dann reduziert werden, soweit machbar, das verlangt das Gesetz – wenn das geschieht, sollen auch Risiken für zum Beispiel den demokratischen Prozess gemessen werden und reduziert werden. Was genau das in der Praxis heißt? Ich habe keine Ahnung. Ich glaube, niemand weiß das. Und wie man diese Risiken reduziert, ist auch eine gute Frage. Also, das ist ein konkretes Beispiel von Risiken, die existieren. Das führt uns ein bisschen weg von unserem Thema: Muss jemand dafür haften? Also, es ist ein gesellschaftliches Problem, das ist klar. Da, denke ich, stimmen alle überein. Was man dazu machen soll, ist eine, ist eine andere Frage. Soll *OpenAI* dafür haften, wenn Misinformationen verbreitet werden? Ich weiß nicht, ob das die richtige Antwort auf das Problem wäre."

[mg]: „Es ist ja auch schwierig, sich vorzustellen, Sie sagten ja am Anfang: Idealerweise soll ein Zustand vor dem Schaden hergestellt werden, und wenn das nicht möglich ist, gibt es einen finanziellen Ausgleich. Das ist natürlich bei einem ruinierten gesellschaftlichen Diskurs schwierig [lacht] zu sagen, was das heißt."

[Gerner-Beuerle]: „Also, wenn man sich solche Rechtsgüter anschaut demokratischen Diskurs, aber auch generell Menschenrechte werden ja auch erwähnt vom KI-Gesetz, das Risiko für Menschenrechte muss auch im gewissen Rahmen bleiben, gering bleiben, reduziert werden, soweit möglich. Also wenn man sich solche Rechtsgüter anschaut, stoßen Haftungsvorschriften natürlich an ihre Grenzen. Es ist schwer zu quantifizieren, was der Schaden jetzt ist, wenn Misinformation verbreitet wird. In manchen Situationen ist es zu quantifizieren, denke ich. Wenn Sie jetzt wieder an das Beispiel denken des Arbeitssuchenden, der keinen Arbeitsplatz bekommt, weil das KI-System diskriminiert oder weil die Daten, die das KI-System zum Training benutzte, nicht ohne Bias waren. Wenn man daran denkt, kann man das quantifizieren: den Gewinn oder Gehalt, den der Arbeitssuchende dann eben nicht bekommt. Ähm – in anderen Fällen ist es sehr schwierig, das zu quantifizieren. Also das KI-Gesetz, das erwähnt ja auch Menschenwürde. Was genau das heißt und wenn Menschenwürde verletzt ist, weil zum Beispiel toxic messages verbreitet werden, die, die herabwürdigend sind gegenüber bestimmten, bestimmten Bevölkerungsgruppen, wie da der Schaden zu quantifizieren ist, das ist schwierig. Ich denke, da stoßen Haftungsvorschriften an ihre Grenzen."

[pgg]: „Eine komplexe Eigenschaft von so, von mit Dialogsystemen verknüpften großen Sprachmodellen, wie *ChatGPT*, ist ja auch, dass sie in der Lage sind, sozusagen Als-ob-Antworten zu geben, also sich in eine Rolle hineinzuimagineren und dann aus dieser Rolle heraus eine bestimmte Logik oder eine bestimmte Sichtweise oder einen Ratschlag zu entwerfen. Also, äh: „Was würde jemand sagen, der versuchen wollte...“ und dann kommt irgendwas Verbotenes oder Dramatisches, äh, und das kann man ja dann im deskriptiven Modus gewissermaßen rein logisch darlegen, und das machen diese Systeme auch. Da kann man natürlich sagen: „Ja gut, es ist ja nur eine Auskunft, was jemand machen würde, der so wäre.“ Aber das sind natürlich – ähm – in der Praxis recht,

äh, blauäugige Formen, sich herauszureden. Wenn wir jetzt mal uns vorstellen, ein Mensch würde das sagen: „Naja, ich habe jetzt ja nur erklärt, was ein Geld- und Wertzeichenfälscher tun würde, der Geld- und Wertzeichen fälschen wollte und da die allerbeste Methode dafür nutzen wollte, der würde es so machen“, ja? Also da ist ja schon ziemlich klar, dass wir, wenn Menschen so vorgehen würden, mal mindestens strafrechtlich gesehen sowas wie irgendeine Tatbeteiligung hätten oder eine Anstiftung oder sowas in der Art...“

[Gerner-Beuerle]: „Hmhm.“

[pgg]: „...das können die Maschinen und man könnte es ihnen womöglich auch irgendwie abtrainieren oder da Sperren einbauen. Da kommen wir vielleicht wieder an so einen Punkt, wo man doch noch einmal näher an der Haftungsfrage ist. Also wenn ein Hersteller sich jetzt vielleicht doch nicht einfach rausreden kann und sagen: „Es kommt halt darauf an, wie man die Maschine benutzt.“ Nein, wenn ich im Modus des „Als ob“ gewissermaßen alles Mögliche von der Maschine bekommen kann, was als ganz normale Auskunft gegen Gesetze verstoßen würde, dann stellt sich doch die Frage: Geht das nicht zulasten des Herstellers?“

[Gerner-Beuerle]: „Ja. Im Augenblick weiß ich nicht, ob man den Hersteller dafür haftbar machen kann, das versucht das KI-Gesetz ja auch zu erreichen. Das KI-Gesetz – ähm – verlangt eben, dass solche Dinge gemessen werden, dass Klarheit darüber besteht, inwieweit da das System, also ein Large Language Model zum Beispiel, sagen wir, bestimmte Vorteile unterstützt oder so. Es kann, es kann ja ganz subtil sein, dass Leute bestimmte Vorurteile haben gegenüber bestimmten Bevölkerungsgruppen und das KI-System, wie Sie sagen, reagiert, ist so trainiert, dass es in einer affirmierenden Weise reagiert auf Prompts und dann diese Vorurteile unter Umständen verstärkt oder dazu verwendet werden kann, um illegale Aktivitäten zu ermöglichen. Das versucht *OpenAI*, glaube ich, schon jetzt zu limitieren. Ich weiß nicht, wie erfolgreich die sind. Ich habe es nicht ausprobiert, aus *ChatGPT* irgendwelche Informationen zu bekommen, um eine Bombe zu bauen oder sowas. Aber, äh, *OpenAI* hat glaube ich erklärt, äh, dass sie da Schutzvorrichtungen ein- eingebaut haben. Also auf jeden Fall: das KI-Gesetz versucht dieses Problem, spricht dieses Problem an, indem es verlangt, dass Hersteller von solchen Systemen messen, inwieweit eben illegale Aktivität unterstützt wird oder Vorurteile – ähm – bestätigt werden. Der Hersteller, wenn das KI-Gesetz in Kraft tritt, der Hersteller muss dann Vorkehrungen dafür treffen, dass das so weit wie möglich eben nicht geschieht. Und wenn er das nicht tut im Rahmen des wissenschaftlich Möglichen, des technisch Möglichen, dann würde der Hersteller haften, in der Tat, für Non-Compliance mit den gesetzlichen Voraussetzungen. Dass das in der Praxis möglich ist, setzt natürlich voraus, dass man tatsächlich vernünftig messen kann, inwieweit etwa *ChatGPT* Vorurteile bestätigt oder so was. Und da gibt es sehr interessante Forschung in Computer Science oder an der Schnittstelle von Computer Science – Regulation – Human Values – äh, die Vorschläge machen dazu, wie man das messen kann. Das ist im Augenblick alles so Cutting Edge of Science, also da ist noch nichts in Industriestandards oder so implementiert, sondern das sind Wissenschaftler, die sich darüber Gedanken

machen, was da getan werden kann. Und wir sind in der Situation, im Augenblick, wo vieles noch unklar ist und Fundamentalforschung geleistet werden muss, um sowas wie das KI-Gesetz – ähm – praktikabel zu machen."

[mgj]: „Sie sagten ja schon, oder wir haben's ja immer mal wieder angesprochen, dass Haftungsfragen auch natürlich nach Verantwortung fragen und nach Verschulden fragen. Sehen Sie jetzt in dieser KI-Debatte eine Verschiebung der Verortung von Verantwortung – also gerade in diesem, in diesem Beispiel, das wir gerade diskutiert haben? Das System gibt, wenn ich es geschickt manipulierte, Informationen preis, die es eigentlich nicht preisgeben soll, oder produziert Anleitungen, die ansonsten sehr, sehr aufwendig nur zu bekommen wären. – Irgendwie ist es ja kontraintuitiv, die Verantwortung von der Person wegzuschieben, die sich da derart betätigt [lacht] hin zu dem Hersteller eines Systems, das dann dazu in der Lage ist. Aber ist das so neu, wie ich das jetzt vermute, oder gibt es analoge Konstruktionen auch schon im, sage ich mal, Vor-KI-Rechtsbild?"

[Gerner-Beuerle]: „Über die Frage müsste ich genauer nachdenken [lacht]. Da müsste ich mir ein bisschen Gedanken drüber machen. Das ist eine gute Frage. Wenn es eine Verschiebung von der Verortung von Verantwortlichkeit geben sollte (ich weiß es nicht, ob das tatsächlich passiert) dann vielleicht weg von – zumindest nach KI-Gesetz, das ist ein Kritikpunkt ja, an dem KI-Gesetz – weg von individueller Verantwortlichkeit hin zu Verantwortung von – Verantwortung von denen, die solche KI-Systeme in Umlauf bringen und dann unter Umständen auch Haftung. Also weniger Caveat emptor und eher etwas, ich weiß nicht, ob paternalistisch das richtige Wort ist, aber regulierend von oben herab, während das KI-Gesetz verlangt, eben dass Risiken akzeptabel sein müssen. Ich weiß nicht, ob ein Gesetzgeber das sektorübergreifend vernünftig verlangen kann. Das hängt ja sehr von der Situation ab und von denen, die ein KI-System verwenden und die – diejenigen, die dem Output des KI-Systems dann ausgesetzt sind."

[pgg]: „Eine gegenläufige Überlegung könnte lauten: Wir haben uns vielleicht einfach noch nicht dran gewöhnt, dass auch von Maschinen, die jetzt im Fall von generativer Text-KI im Wesentlichen Sprache produzieren, ähnlich viel sehr generelle Gefahren ausgehen können, die vielleicht – was uns viel vertrauter ist – von Chemikalien in anderen Bereichen oder von Suchtstoffen oder von Dingen diesen Typs, wo wir ja auch paternalistisch sind, sofern wir sagen: Das ist zum Beispiel auf Dauer ungesund. Das merkt man vielleicht sogar gar nicht unbedingt – äh – dass irgendwie ein Schadstoff im Boden oder in der Luft ist, trotzdem gibt es Grenzwerte, und da, wo es gefährlich wird für die Bevölkerung, für die Gesellschaft, da greift der Gesetzgeber selbstverständlich ein und da wird auch Industrie in die Pflicht genommen usw. Also das sind wir im Bereich, nehmen wir mal der Chemie gewöhnt und wir haben uns einfach noch nicht dran gewöhnt, dass Maschinen inzwischen auch Sprache in einer Wucht und Mächtigkeit in den Alltag hineintreiben können, die Kausalketten oder fast kausalarartige Ketten auslösen, die tatsächlich gesamtgesellschaftliche Effekte erzeugen können."

[Gerner-Beuerle]: „Ja, das, äh, ist durchaus überzeugend, das so zu sehen. Ich weiß nicht, ob ich wirklich letztlich schon überzeugt bin davon [lacht], aber ich finde es durchaus überzeugend, diesen Gedankengang. Der Unterschied zu Chemikalien zum Beispiel ist natürlich, dass – dass KI so verbreitet ist und verbreitet sein wird, dass es in alle Bereiche, in alle Sektoren, alle Industrien oder, ähm, und wirklich jeden Bereich des Lebens eingreifen wird. Wie weit man da regulieren sollte und überhaupt regulieren kann, vernünftig regulieren kann, ist eine offene Frage. Ob Chemikalien gefährlich sind, wie gefährlich sie sind, lässt sich, glaube ich, ja relativ objektiv messen, denke ich. Ich bin kein Chemiker, aber stelle ich mir vor [lacht]. Inwieweit ein KI-System gefährlich ist, ist viel schwieriger zu messen. Äh, und dann kann ich Ihnen noch ein Gegenbeispiel bringen von meinem Gebiet, in dem ich viel gearbeitet habe, der Finanzmarktregulierung. Also es gibt bestimmte Vorschriften in der europäischen Finanzmarktregulierung und auch in anderen Ländern, die sich an KI-Systeme wenden schon. Das hat jetzt nichts mit dem KI-Gesetz zu tun oder anderen Gesetzesvorhaben. Also: Es gibt in normalen Kapitalmarktgesetzen Vorschriften, die zum Beispiel High-Frequency Trading betreffen. High-Frequency Trading basiert auch auf KI. Und diese Vorschriften, die verlangen dann etwa, dass die Unternehmen, die Investmentfirms, die solche Algorithmen einsetzen, dafür Sorge zu tragen haben, dass die Märkte nicht zu sehr gestört werden, dass nicht übermäßige, exzessive Volatilität eintritt als Folge der Verwendung von solchen Algorithmen. Und hierbei frage ich mich, wieso solche Voraussetzungen auf Unternehmen, die KI-Systeme einsetzen, übertragen werden, aber nicht auf Menschen, die genau solche Folgen haben können mit ihren Handlungen. Und das ist ähnlich bei Misinformation natürlich, Menschen können viel Blödsinn sagen [lacht] und können den demokratischen Prozess stören. Der Unterschied ist nur, dass es wesentlich schneller geht mit KI-Systemen und vielleicht die Reichweite auch wesentlich größer ist. Rechtfertigt das eine fundamental andere Herangehensweise an die Regulierung? Ich weiß es nicht. Vielleicht, ja. Vielleicht haben Sie recht und es tut das. Aber das ist, das sind schwierige Fragen, denke ich, die auch weiter debattiert werden und die wahrscheinlich dann Politiker auch unterschiedlich beantworten werden in unterschiedlichen Ländern.“

[pgg]: „Ja, genau. Also die Frage der Öffentlichkeit, das wäre auch eine klassische Debatte zwischen Presserecht auf der einen Seite: Nicht jeder hat voraussetzungslos Zugang zu beliebig großen Öffentlichkeiten...“

[Gerner-Beuerle]: „Ja.“

Petra Gehring [pgg]: „...oder auf der anderen Seite der Idee: Naja, klar. Also, seit es das Netz gibt, kann auch ein Individuum im Grunde wie ein Massenmedium agieren.“

[Gerner-Beuerle]: „Ja. Ohne KI wird ja auch viel Misinformation jetzt verbreitet und kann auch verbreitet werden über das Internet. Und das stört auch demokratische Prozesse.“

[pgg]: „Das ist richtig. Aber dass eine Sache schlecht ist, spricht nicht dafür, dass die nächste auch noch schlecht sein sollte.“

[Gerner-Beuerle]: „Und hier wird ja auch reguliert [lacht]. “

[pgg]: „Auch das ist der Fall, ja, Bots sind ja nicht unbedingt legal, also gerade aggressive.“

[Der Abspann mit Musik beginnt.]

[Gerner-Beuerle]: „Ja, es wird reguliert, oder zumindest überlegt sich die EU, wie sie regulieren kann. Und da gibt es ja auch jetzt weitere Gesetzesvorhaben, die in der Pipeline sind.“

[mg]: „Und damit ist dieses *Digitalgespräch* zu Ende und wir bedanken uns bei Carsten Gerner-Beuerle vom University College London für dieses spannende Gespräch und die interessante Diskussion. Viele Grüße nach Großbritannien. Und natürlich auch vielen Dank an Sie, liebe Zuhörerinnen und Zuhörer, für das Interesse und die Aufmerksamkeit, und wenn Sie mögen, hören wir uns wieder, wie immer in drei Wochen zur nächsten Folge des *Digitalgesprächs*, einem Podcast von ZEVEDI, dem *Zentrum verantwortungsbewusste Digitalisierung*.“



This work is licensed under CC BY-NC-ND 4.0. To view a copy of this license, visit <https://creativecommons.org/licenses/by-nc-nd/4.0/>