07.06.2023

Transkript

„Regulierung von KI – was ist geplant, was ist nötig, was ist machbar?"

Expertinnen und Experten auf dem Podium

- **Prof. Dr. Matthias Kettemann**
  Professor für Innovationsrecht, Universität Innsbruck, Österreich, und Programmleiter Forschungsprogramm „Regelungsstrukturen und Regelbildung in digitalen Kommunikationsräumen", Leibniz-Institut für Medienforschung | Hans-Bredow-Institut (HBI), Hamburg
- **Prof. Dr. Ulrike Luxburg**
  Professorin für die Theorie des maschinellen Lernens, Eberhard Karls Universität Tübingen
- **Prof. Dr. Sandra Wachter**
  Professorin für Technologie und Regulierung, Oxford Internet Institute, University of Oxford, Vereinigtes Königreich
- **Bastian Zimmermann**
  Redakteur für Digitales und Technologie, Science Media Center Germany,
  und Moderator dieser Veranstaltung

Mitschnitt

- Einen Videomitschnitt finden Sie unter:
  https://www.sciencemediacenter.de/alle-angebote/press-briefing/details/news/regulierung-von-ki-was-ist-geplant-was-ist-noetig-was-ist-machbar/
- Falls Sie eine Audiodatei oder eine Sprecheransicht des Videomittschnitts benötigen, können Sie sich an redaktion@sciencemediacenter.de wenden.

Transkript

Moderator [00:00:00]

Liebe Journalistinnen und Journalisten, herzlich willkommen zu unserem virtuellen Press Briefing zum Thema KI-Regulierung. Ich bin Bastian Zimmermann, ich bin Redakteur beim Science Media Center und moderiere heute hier die Veranstaltung. Mit mir habe ich noch zwei Expertinnen und Experten. Erst mal schönen Dank, dass Sie da sind, das freut mich und ich stelle Sie gleich noch im Detail vor. Heute geht es ja um die Regulierung von KI. Das ist seit langem ein wichtiges und interessantes Thema, aber gerade durch ChatGPT und andere generative KI hat es in den letzten Monaten, im letzten halben Jahr noch mal an Relevanz gewonnen. Das wichtigste Gesetz ist sicher der AI Act, der wird zur Zeit in der EU besprochen. Nächste Woche stimmt auch voraussichtlich das Parlament über seine Position dazu ab. Wenn es schnell geht, das ist zumindest mein letzter Stand, korrigieren Sie mich, könnte es sein, dass er noch dieses Jahr verabschiedet wird, aber dann ist ja auch noch zwei Jahre Übergangszeit. Und dann gibt es die Fragen: Was ist wichtig bei dieser Regulierung, was sollte da beachtet werden? Und die wollen wir heute und auch ihre Fragen aus juristischer und praktischer Perspektive beurteilen. Bevor ich zur Vorstellung komme, noch der Hinweis an Sie da draußen: Stellen Sie gerne Ihre Fragen und stellen Sie diese dann bitte über die Fragefunktion in Zoom, dann können Ihre Kolleginnen und Kollegen die auch sehen und es gibt weniger Dopplungen, also bitte das F&A-Modul benutzen und nicht den Chat. Dann komme ich zu den Expertinnen und Experten. Ich stelle sie in alphabetischer Reihenfolge vor und da haben wir als erstes Prof. Dr. Matthias Kettemann. Er ist Programmleiter im Forschungsprogramm "Regelungsstrukturen und Regelbildung in digitalen Kommunikationsräumen" am Leibniz-Institut für Medienforschung, dem Hans-Bredow-Institut in Hamburg. Er beschäftigt sich mit KI-Regulierung und auch viel mit der Regulierung sozialer Medien. Dann haben wir Prof. Dr. Ulrike Luxburg. Sie ist Professorin für die Theorie des maschinellen Lernens an der Eberhard Karls Universität Tübingen und bringt heute hier die Perspektive aus der KI-Praxis ein und befasst sich auch mit Fragen der Fairness und Erklärbarkeit von KI, was ja gerade im Kontext der Regulierung noch wichtige Fragen sind. Und dann haben wir Prof. Dr. Sandra Wachter. Sie ist Professorin für Technologie und Regulierung am Oxford Internet Institute der University of Oxford im Vereinigten Königreich und befasst sich mit KI- Regulierung und KI-Ethik. Zum Anfang habe ich kurz ein paar Eingangsfragen vorbereitet und fange auch mit Ihnen an, Herr Kettemann, um die Grundlage für die Debatte zu legen. Die Frage: Wo steht denn der AI Act momentan und was sind Ihrer Meinung nach da die wichtigsten Aspekte?

Matthias Kettemann [00:02:36]

Danke für die Einladung. Ich komme direkt von der re:publica zu Ihnen. Da tut sich gerade sehr viel. Gefühlt jede Diskussion handelt von der KI, von der Macht der KI, von der Frage, wie man sie gut regulieren kann und es ist auch schon die Rahmenerzählung, die ich anbieten möchte. Es wird sehr viel gesprochen, was die KI bewirken kann, von der Auslöschung der Menschheit wird gesprochen. Aber das ist Angstmacherei, das ist Überschätzung. Es ist ganz, ganz schwierig, dass hypothetische Risiken gegenüber bestehenden kleineren Gefahren in den Vordergrund gerückt werden. Wir müssen weiter an den Themen arbeiten, an Biases, an den Problemen, die wir adressieren können. Und die müssen wir adressieren durch die Gesetzgebung, und das bringt uns zum KI-Rechtsakt. Wir dürfen uns also nicht von den Gefahren der Zukunft immobilisieren lassen und nicht zu sehr erschrecken, wenn jemand sagt, die KI wird uns alle umbringen. Diese Gefahr besteht nach jetzigem Stand nicht, aber wir müssen dennoch gute Regeln setzen. Und der KI-Rechtsakt, der aktuell verhandelt wird, wird so ähnlich wie die DSGVO im Datenschutzbereich global eine Rolle spielen, weil es die aktuell detailreichste, umfassendste Regulierung der künstlichen Intelligenz ist, die jetzt gerade aufgrund der großen Sprachmodelle in aller Munde ist. Die ganzen Diskussionsschritte, da haben wir zwei wichtige Prinzipien innerhalb des KI-Rechtaktes nicht verändert, nämlich dass man einen risikobasierten Ansatz fährt, dass man also sagt, wir schauen uns an, welche Risiken KI-

Systeme in der Gesellschaft haben und reguliert dann diese Systeme, die besonders risikoreich sind. Und dann der Ansatz, dass man nicht so sehr auf die Technologie an sich schaut, sondern auf die Auswirkungen innerhalb der Gesellschaft. Gerade jetzt, meine Kollegin wird da noch eine Bewertung vornehmen, ich stelle den Rahmen vor, gerade jetzt hat das Parlament noch nachgeschärft in einigen Aspekten. Zum Beispiel haben die Abgeordneten im Parlament im Vergleich zu dem Kommissionsentwurf ein Verbot generell von KI für die biometrische Überwachung, für Emotionserkennung, für Predictive Policing, also vorausschauende Polizeiarbeit, gefordert. Und vor allem haben sie auch gesagt, wir müssen noch die Regelungen hinsichtlich der allgemeinen [unverständlich] KI und vor allem auch dieser sogenannten Foundation Modelle, also dieser generativen KI wie ChatGPT stärker regulieren. Das wurde in den neuesten Entwurf noch hinein geschrieben. Und vor allem ganz zentral innerhalb der Logik des Parlaments, die immer sehr auf die Bürger*innenrechte achten, ist ein verstärktes Recht auf Beschwerden über KI-Systeme, das heißt über Entscheidungen vor allem, die KI-basiert ablaufen. Dieser menschenrechtszentrierte Ansatz an die KI-Entwicklung findet sich nicht erst seit den Kommentaren des Parlaments in dem Entwurf, [sondern] auch seit dem Beginn. Man darf nicht unterschätzen: Die EU ist zwar in der Tat eine Binnenmarktunion, die auch natürlich die europäische technologische Souveränität fördern will, aber gerade im Bereich der KI stand immer schon auch der Schutz der Menschen im Mittelpunkt. Schon angesprochen, das Verbot der biometrischen Erkennung-Systeme, besonders jener, die in Echtzeit funktionieren sollen, ist zentral, gerade dann, wenn es in öffentlichen Räumen eingesetzt werden soll. Da gibt es bestimmte Ausnahmen im Bereich der Strafverfolgungsbehörden, aber auch nur dann, wenn besonders schwere Straftaten bekämpft werden sollen. Emotionsbezogene Erkennungsysteme sollen generell verboten werden, auch gerade am Arbeitsplatz zum Beispiel oder in Bildungseinrichtungen. Gerade der Aufbau von großen biometrischen Datenbanken ist äußerst problematisch. Das Wichtigste [unverständlich] ist die sogenannte Hochrisiko-KI, die hochriskante KI. Die wird jetzt auch verstärkt nach Vorstellung der Abgeordneten reguliert in den Bereichen Gesundheit, Sicherheit, Umwelt und Grundrechte. Und es sollten auch Systeme zur Beeinflussung von Wählerinnen und Wählern in demokratischem Kontext verboten werden. Besonders spannend, weil diese Systeme verwendet werden, unter anderem im Rahmen der generellen Empfehlungssysteme auf Social Media Plattformen. Und das würde natürlich zu demokratischen Folgen führen, wenn entsprechende Systeme eingesetzt werden, da wollen die Parlamentarier*innen, die besonders sensibilisiert sind für demokratische Mitbestimmungsprozesse, vorbauen. Sehr wichtig ist auch,

Moderator [00:07:29]

Herr Kettemann, können Sie vielleicht das Mikro etwas mehr vor den Mund halten. Ich glaube, wir müssen vielleicht ein bisschen rumprobieren, wie wir das mit der Tonqualität hoch kriegen, weil einfach sehr viel los ist.

Matthias Kettemann [00:07:41]

Entschuldigung, das ist einfach die Aktivität auf der re:publica. Ich habe die ruhigste Ecke gesucht, die war nicht ruhig genug. Ich hätte eine KI-basierte Software, um die Hintergrundgeräusche rauszunehmen, das wäre der kritische Meta-Kommentar. Um noch einen Punkt anzusprechen: Die Transparenz und Erklärbarkeit von künstlicher Intelligenz sind zwei weitere wichtige Aspekt des Regelungskonzeptes des AI Actes. Dabei gehen diese Ansätze aber über einiges hinaus über bestehende Regulierungen und sind deswegen [unverständlich] zu sehen. Dennoch ist es nicht so, dass dieser KI-Rechtsakt auf einem leeren Regulierungsfeld eingschlagen wird. Es gibt schon jetzt sehr viele ethische Prinzipien, für die KI, dass wir wissen, in welche Richtung wir gehen müssen. Wir wissen, dass KI Menschenrechte achten, sichern und fördern muss. Wir wissen, dass ethische und rechtliche Leitplanken im gesamten Lebenszyklus von KI-Systemen eingehalten werden müssen. Und das muss sowohl der Ebene der staatlichen Regulierung als mit der europäischen Regulierung

sichergestellt werden, sowohl nach innen als auch im Bereich der globalen Kooperation. So hat zum Beispiel erst vor kurzem der Technologierat, das ist eine Kooperation zwischen EU und USA, in Schweden festgehalten, dass verstärkte Technologieregulierung im Zentrum des gemeinsamen Interesses stehen soll.

Moderator [00:09:17]

Vielen Dank, Herr Kettemann, trotz widriger Bedingungen. Dann machen wir direkt weiter mit Frau Wachter und der Einordnung in der ganzen Geschichte. Wie würden Sie den AI Act beurteilen, ist das ein gutes oder ein schlechtes Gesetz und wo gibt es Ihrer Meinung nach noch Nachbesserungsbedarf?

Sandra Wachter [00:09:35]

Super Frage, es darf spannend bleiben. Wir schreiben ja gerade Geschichte, wenn es um das geht. Ich wollte drei Dinge ansprechen, die für mich persönlich sehr, sehr wichtig sind, dass wir die im Hinterkopf behalten, jetzt, wenn das Ding dann verabschiedet wird, wie wir weiterhin damit umgeben. Das erste der drei Dinge hat tatsächlich mit dem Herzstück des AI Acts zu tun und das ist jetzt das sogenannte Conformity Assessment. Das Conformity Assessment bedeutet, dass die Entwickler selbst bewerten, selbst zertifizieren werden müssen, die sich im Hochrisikobereich befinden, ob sie selbst den AI-Act-Regeln entsprechen und ob sie AI-Act -Standards entsprechen. Ausnahmen gibt es nur bei biometrischen Daten, da muss das ein Dritter machen. Dann sieht man natürlich, dass das eine Regelung ist, die erlaubt, dass diejenigen, die dem Recht folgen sollen, auch diejenigen sind, die entscheiden, ob sie tatsächlich dem Recht folgen. Das ist für sich genommen noch nicht super problematisch, aber etwas, das wir tatsächlich im Auge behalten sollten. Der AI Act selbst sieht es sogar auch als ein Problem. Wenn man sich den Erwägungsgrund 64 ansieht, dann spricht der ganz deutlich davon, dass das eine Übergangslösung sein soll und dass man in Wirklichkeit zu einem gewissen Zeitpunkt diese Conformitiy Assessments an unabhängige Dritte übertragen sollte. Nun ist es so, dass diese Idee in einem Erwägungsgrund steht, die aber nicht rechtlich bindend sind. Es ist auch im AI Act nicht wirklich klar vorgesehen, wie das Procedere dann umgeleitet werden soll. Mein Problem besteht darin, dass wir das vielleicht dann irgendwann vergessen und mit dem System, mit der Übergangslösung, mit der Notlösung bleiben, anstatt uns zu überlegen, wie wir diese Gestaltung besser machen könnten, dass wir ein Conformity Assessment haben, das nicht von denjenigen gemacht wird, die den Regeln folgen sollen, sondern dass wir unabhängige Dritte haben, die das durchführen können. Das führt mich auch schon zum zweiten Thema. Da geht es um die Standards, die geschaffen werden. Die Entwickler müssen ein Conformity Assessment machen, ob sie den AI-Standards entsprechen. Diese AI-Standards gibt es ja noch nicht, sondern die werden jetzt erst ausgearbeitet und Standardisierung-Organisationen wie [unverständlich] werden beauftragt, diese zu erstellen. Und dann gibt es das erste Problem, dass natürlich [unverständlich] keine demokratisch legitimierten Institutionen sind, es sind private Institutionen und da kann man einen Dialog darüber führen, ob das eine gute Idee ist, dieses wichtige Instrument an Pirvate auszulagern. Was auch damit verbunden ist, wer in diesen Working Groups drinnen sitzt, um diese Standards zu schreiben. Und dann ist es tatsächlich so, dass das in der Regel hauptsächlich Leute von der Industrie sind. Wenn man das mit dem ersten Problem zusammenhängt, dann sieht man, dass der AI Act vorsieht, dass die Industrie selbst entscheiden soll, ob sie den Standards entspricht, welche sie auch hauptsächlich selbst geschrieben hat. Natürlich, wie gesagt, muss das auch noch kein Problem für sich sein, aber es ist schon nett, das im Hinterkopf zu behalten. Und da ist der zweite Call to Action, das könnte man durchaus noch nachbessern und sichergehen, dass diejenigen, die um den Tisch sitzen, dass das sozusagen ein bisschen ausgeglichen ist und nicht nur in die Industrie gelagert. Das Dritte, was damit zu tun hat, mit den Standards, ist natürlich, dass diese Standards sehr oft, wenn man sich das ansieht, sehr technisch sind. Was ganz klar und ver-

ständlich ist, das sollen sie auch sein. Aber ich glaube, dass vielleicht der Öffentlichkeit nicht immer ganz bewusst ist, dass die uns nicht unbedingt helfen werden oder Gefahr laufen, uns nicht [dabei zu helfen], diese ethischen Dinge richtig auszulegen. Es kann durchaus sein, dass die Standards, die gemacht werden, ein paar Möglichkeiten anbieten oder aufziehen, welche Explainability-Tools es gibt oder welche technischen Möglichkeiten es zum Bias testen gibt. Es gibt aber sozusagen eine Anleitung, welche wann zu nutzen ist, nicht wirklich drin sein wird und dass diese ethischen Überlegungen sich nicht in den Standards finden könnten. Und wie gesagt, das kann tatsächlich passieren und dann stellt sich natürlich die Frage: Wer soll denn das machen oder sind wir zufrieden, nur ein Regelwerk zu haben, das sich an die technischen Fragen anlehnt? Das ist natürlich im Generellen ein Problem, wenn man den AI Act als Ganzes betrachtet, weil man sieht, dass die Standards, die geschaffen werden, im Hochrisikobereich in Wirklichkeit Anleitung geben, wie man etwas einsetzt. Aber eine Diskussion, ob man sie einsetzen soll, ist absolut nicht mehr gegeben. In Wirklichkeit haben wir diese Hochrisikobereiche und es werden vor allem bürokratische Hürden aufgestellt, KI zum Beispiel in der Bildung oder im Kreditwesen oder in der Strafrechtspflege einzusetzen. Es wird mühsamer gemacht oder schwieriger gemacht, das zu machen, aber sie werden nicht wirklich verboten. Es gibt gar keine Diskussion, ob das eigentlich eine gute Idee ist, KI in der Bildung oder im Strafrechtswesen einzusetzen. Und wenn man sich den Diskurs ansieht, dann gibt es einige, die gemeint haben, dass es durchaus problematisch sein kann. Ich würde mir wünschen, dass es ein bisschen mehr Diskurs darüber gibt, unter welchen Umsetzungsmöglichkeiten das eigentlich akzeptabel ist und wann nicht. Das letzte, was ich noch ganz kurz anbringen möchte, hätte man eher schon auf aufgebracht und das hat mit dem Individualrechtsschutz zu tun. Das ganze Regelwerk ist hoch inspiriert von Produktsicherheitsrecht, wo die Möglichkeit eines Beschwerdewesens nicht wirklich im Mittelpunkt stand. Und das ist eigentlich recht untypisch für Technologieregulierungen in Europa. Wenn man sich das Datenschutzgesetz anschaut, wenn man sich die Datenschutzgrundverordnung ansieht oder den DSA, den Digital Services Act anschaut, dann hat der Bürger, die Bürgerin die Möglichkeit, Beschwerde zu erheben. Jetzt, im Draft vom Europäischen Parlament, ist das ganz klar angeregt. Es sind einige Rechte vorgeschlagen, die Individuen das geben könnten und wir sollten sicher gehen, dass das auch dringend erhalten bleibt, dass es nicht am Ende noch herausgenommen wird. Ich finde es ganz wichtig ist, dass es eine Counter Balance zu diesem Conformity Assessment gibt, dass sich auch der Einzelne in Wirklichkeit beschweren kann. Wenn man das noch auf die Reihe bekäme, bevor der Draft verabschiedet wird, dann glaube ich, dass wir eigentlich auf guten Beinen stehen, ein Gesetz zu haben, das sehr, sehr hilfreich sein wird. Danke vielmals.

Moderator [00:16:49]

Vielen Dank, Frau Wachter. Und dann noch Frau Luxburg. Die Frage, wie Sie das als KI-Praktikerin sehen, welche Regulierungen aus ihrer Perspektive wünschenswert wären und welche wären für die Forschung zu einengend.

Ulrike Luxburg [00:17:02]

Vielen Dank. Ich habe jetzt natürlich keinen juristischen Blick auf dieses Gesetz, sondern den Blick eine Informatikerin. Vorwegschicken will ich, dass ich auch aus der Forschung heraus die ganzen Regulierungs-Anstrengungen sehr wünschenswert und sehr wichtig finde. Ich finde es extrem wichtig, dass wir ein Gesetz auf den Weg bringen. Ich finde auch, vieles an dem Gesetz geht in eine gute Richtung, also der risikobasierte Ansatz, der dort gefahren wird und so weiter. Ich habe das Gefühl, es sind noch ein paar Sachen, die vielleicht noch nicht so richtig super gelöst sind in dem Gesetz. Ich glaube, die richtigen Themen werden angesprochen, also das Thema Bias, Diskriminierung und so weiter, das wird an vielen Stellen erwähnt, dass, wenn solche Software entwickelt wird, man darauf achten soll, dass keine Biases drin sind und so weiter. Da stellt sich mir aber die

Frage der Umsetzung und der Kontrolle, das schließt an das an, was Sandra Wachter gesagt hat. Das ist natürlich schön, wenn die Firmen alle dokumentieren, wir haben diese sieben Sachen gemacht, damit kein Bias drin ist, aber wer kontrolliert das eigentlich und wer hätte die Macht zu sagen, das ist eigentlich nicht der Stand der Technik oder vielleicht gibt es bei dieser Anwendung überhaupt keine Möglichkeiten, Bias zu kontrollieren und wir würden diese Anwendung deswegen gar nicht gerne haben wollen? Der zweite Punkt, mit dem ich mich jetzt besonders auskenne, hat mit diesem Erklärbarkeitsaspekt zu tun. Ein wichtiger Teil des Gesetzes beschäftigt sich mit diesem Unwohlsein, was man hat. Da sind diese ganzen Black-Box-Modelle, die machen Entscheidungen über Personen und wir wollen eigentlich wissen, wie diese Entscheidungen zustande kommen und im besten Falle würden wir sie gerne auch kontrollieren können. Und um das umzusetzen, wird in dem Gesetz an vielen Stellen immer darüber geschrieben, dass entweder die Sachen erklärbar sein sollen oder Transparenz hergestellt werden soll, dadurch, dass ausführlich dokumentiert wird, welche Daten verwendet wurden und welches Training und welche Parameter und so weiter. Aus meinem technischen Gesichtspunkt her würde ich sagen, dass man durch diese Methoden das Vertrauen eigentlich nicht gewinnen kann. Wenn dort Methoden verwendet werden, die Black-Box-Methoden sind, tiefe Netzwerke zum Beispiel, da wird es keine Erklärbarkeitsmethoden geben, mit denen man auf sinnvolle Art und Weise wirklich erklären kann, was diese Dinge machen; weder als gesamtes Modell noch für einzelne Entscheidungen. Was in dem Gesetz nicht drinsteht, ist, was passiert denn wenn die Erklärungen nicht ausreichend waren. Wann würden wir denn sagen, eine Erklärung ist ausreichend. Also die Bank kann mir immer sagen, du kriegst den Kredit nicht, weil du zu wenig Geld verdienst. Aber wann glauben wir diese Erklärung, wann glauben wir sie nicht, wer würde das kontrollieren und was würden wir machen, wenn wir diesen Erklärungen nicht glauben. Wäre denn dann die Konsequenz, dass die Bank dieses System einfach so nicht verwenden kann? Und diese Diskussion ist der subliminale Text, der da irgendwie hinten drin steht, also zumindest aus meiner Informatik-Perspektive, ist nicht explizit genau an der Stelle. Ich glaube, man muss sich wirklich sehr klar sein: Black-Box-Methoden wird man nicht transparent hinkriegen, auch nicht durch staatliche Kontrolle, auch nicht durch Offenlegung des gesamten Codes. Es gibt keine Menschen mehr, keinen Kontrolleur, der sich sowas anschauen kann und sagen kann: Das stimmt alles oder macht Sinn. Da fehlt mir noch die Entscheidung: Wie geht man damit um in bestimmten Situationen, wenn jetzt irgendjemand sagt, mein Modell ist aber ein Black-Box-Modell und wir aber eigentlich das Gefühl haben, die Erklärungen, die stimmen nicht oder die gibt es nicht oder die kann man nicht einlösen. Das ist ein Aspekt, der mir sehr fehlt. Zu den generativen- und Sprachmodellen, also ChatGPT und so weiter, steht zwar in den neueren Texten in dem Entwurf was drin, das ist aber meiner Meinung nach auch sehr oberflächlich und wahrscheinlich auch noch nicht zu Ende gedacht. Was jetzt auch nicht so überraschend ist, weil das Thema erst seit einem halben Jahr oder so wirklich aufs Tapet gekommen ist. Auch da glaube ich, dass es noch vieles gibt, was man verbessern kann. Da steht eigentlich im Moment nicht viel mehr drin, als dass die Firmen das bitte vorsichtig umsetzen sollen und alles schön dokumentieren sollen. Dann wird zwar erwähnt, man soll natürlich dokumentieren, ob das ein Risiko für die Demokratie ist. Aber auch hier: Was heißt das jetzt, wann wird es denn ein Risiko für die Demokratie und was würde denn dann daraus folgen? Das ist aus meiner Sicht das, was ich am problematischsten an diesem Gesetz finde: Es stehen zwar viele richtige Dinge drin, aber was mir fehlt, ist die letzten Endes wahrscheinlich staatliche Kontrolle, wer soll das eigentlich überprüfen. Es ist wahrscheinlich auch eine politische Frage. In Deutschland wüsste ich nicht, welche Institution das sein sollte, die diese ganzen Sachen überprüft und welche die technische Expertise und auch die Manpower hat, solche Systeme zu überprüfen. So, das ist meine erste Einschätzung aus technischer Sicht.

Moderator [00:21:42]

Ja, dankeschön. Das hat schon einige Fragen, die ich mir aufgeschrieben hatte, eigentlich beantwortet. Wir haben hier eine Frage, die in verschiedensten Versionen schon gestellt wurde und die ich mir auch aufgeschrieben hatte. Ich stelle Sie mal an Sie, Frau Wachter. Und es geht um die War-

nungen, die aus den Reihen der KI-Experten kommen, wie die jetzt gerade einzuschätzen sind, und dass es eben Kritiker gibt, die solche Aufrufe als Strategie bezeichnen, um sich weitere Märkte auf dem KI-Sektor zu sichern, während dann zum Beispiel kleinere Start-ups auf der Strecke bleiben. Und das ist ja auch manchmal eine Kritik, die gegenüber dem AI Act geäußert wird, dass dann diese Art der Regulierung natürlich die kleinen Firmen viel härter trifft als die großen Firmen. Die großen Firmen haben eben einfach viel mehr Möglichkeiten, so was einzuhalten und mehr Ressourcen da reinzustecken. Wie schätzen Sie diese Warnungen ein? Und auch gerade im Hinblick darauf, dass es vielleicht für kleinere Firmen schwieriger wird.

Sandra Wachter [00:22:35]

Das verstehe ich natürlich, wenn man sich den AI Act genauer ansieht und die Technikregulierung in der Europäischen Union im Generellen, dann wird darauf immer auch abgestellt, dass es Start-ups und SMEs (Small and medium-sized enterprises) weniger stark treffen soll. Das hängt von Dokumentationspflichten ab. Das hängt davon ab, wie viele Strafen letztlich aufgebrummt werden, wenn es dazu kommt. Da wird natürlich mit beeinträchtigt. Zu sagen, dass man das nicht miteinbezieht, das halte ich für übertrieben. Das andere, was man tatsächlich sagen muss, ist, dass wenn man sich im Hochrisikobereich bewegt, dann hat man eben auch eine gewisse Art von Verantwortung zu tragen. Und wenn man die Manpower nicht hat, ein System herzustellen, das keinen Bias hat, dann soll man es vielleicht auch nicht in der Strafrechtspflege einsetzen. Nicht jeder muss Hochrisiko-KI herstellen, und das ist natürlich auch etwas, das man mitbedenken muss. Und deswegen finde ich auch den Risikoansatz grundsätzlich sehr gut. Und dann muss der Grundrechtsschutz einfach mitgelebt werden. Aber grundsätzlich, ja, gibt es da ausgleichende Maßnahmen für die kleineren Unternehmen. Aber wie gesagt, bei bestimmten Dingen muss sich einfach wirklich jeder dran halten, weil die Demokratie und der Bürger darunter zu leiden haben, wenn das Ding schadhaft ist.

Moderator [00:24:01]

Eine Frage an Sie, Herr Kettemann. Der Journalist schreibt: Die DSGVO zeigt, dass auch die Verordnungen sehr unterschiedlich umgesetzt werden, zum Teil selbst in Bundesländern, wenn die Datenschutzbeauftragten sehr aktiv sind. Droht dann auch beim AI Act eine Zerfaserung unterschiedlich strenger KI-Regulierung in der EU? Und gibt es dann auch die Gefahr der Verlagerung von Unternehmen in Länder, in denen das Recht eher lax gehandhabt wird?

Matthias Kettemann [00:24:27]

Die Gefahr besteht natürlich, genau so wie bei anderen digitalen Regeln sieht, wie zum Beispiel dem Rechtsakt zu digitalen Diensten.

Moderator [00:24:37]

Können Sie einen Tacken lauter sprechen. Sorry, dass ich immer bei Ihnen rummäkeln muss.

Matthias Kettemann [00:24:43] Die Gefahr besteht natürlich, dass es zwar zu einem guten Gesetz kommt, dass das dann aber nicht umgesetzt wird, besonders auch in Ländern mit föderalen Strukturen, wo es ja nicht nur innerhalb der [...] Staatenkompetenzen Schwierigkeiten geben könnte, zwischen EU-Staaten, sondern auch innerhalb der jeweiligen Länder. Ich sehe aber im Bereich der Technologie eher weniger Probleme auf uns zukommen als in anderen Bereichen der digitalen Gesellschaft wie gerade bei den Plattformen, bei den digitalen Diensten, weil KI-Regulierung re-

gelmäßig etwas ist, wovon alle staatlichen Stakeholder sehr überzeugt sind, dass das nötig ist. Insofern glaube ich, dass hier die Durchsetzung innerhalb der Staaten relativ gut erfolgen wird. Was aber in der Tat herausfordernd sein wird, ist die Frage, ob man durch eine zu starke Regulierung den europäischen digitalen Binnenmarkt im Bereich der KI beschädigt. Das ist dieses große Innovations-Dilemma: Wenn wir zu spät regulieren, kann es sein, dass wir die gesellschaftlichen Folgen schon gewärtigen müssen. Wenn wir zu früh regulieren, kann es sein, dass wir Innovationskraft hemmen. Ich würde aber sagen, dass der KI-Rechtsakt, so wie er sich jetzt darstellt, diesen Spagat ganz gut trifft, höchstens dass man in den Bereichen der Risikoanwendung noch einmal päziser darlegt, wie diese Daten effektiv überwacht werden können, [...] ohne dass es zu einer zu großen Belastung der Unternehmen führen müsste.

Moderator [00:26:25]

Eine Anschlussfrage noch an das, was Sie gesagt haben, Frau Luxburg. Meine Frage wäre eigentlich gewesen: Wie gut kann man Sprachmodelle überhaupt überprüfbar machen? Und da haben Sie ja schon gesagt, dass das zum Teil sehr schwierig ist. Was ist denn Ihrer Einschätzung nach und vielleicht auch noch der von Frau Wachter, wer kann denn dann dafür verantwortlich gemacht werden, wenn quasi durch die Sprachmodelle irgendwelche illegalen Sachen gemacht werden oder irgendwas passiert, was so nicht intendiert war, weil das Sprachmuster falsche Sachen rausgegeben hat? Kann man das überhaupt logisch dann quasi dem Entwickler anhängen oder muss man sagen, wer auch immer das dann wirklich in der Praxis einsetzt, muss dann noch mal genauer draufgucken?

Ulrike Luxburg [00:27:04]

Also ich kann jetzt dazu nicht die rechtliche Seite [einschätzen]. Das soll dann die Sandra Wachter erklären, aber aus technischer Sicht: Ich glaube, man muss zwei Dinge unterscheiden Es gibt einmal die Entwickler, die einfach das Sprachmodell entwickeln, mit dem man dann alle möglichen Dinge tun kann. Wenn hinterher mit dem Sprachmodell jemand Fake News produziert und so versucht, den Wahlkampf in irgendeinem Land zu beeinflussen, ist es natürlich sicher nichts, [wo] man dem Entwickler jetzt die Verantwortung geben kann, wenn man ihm zumindest erlaubt, solche Sprachmodelle zu bauen. Das heißt also, für ganz konkrete Anwendungen ist, denke ich, der oder diejenige verantwortlich, der das macht am Schluss. Das Problem wird aber trotzdem sein, welche Art der Verantwortung man den Entwickler*innen eigentlich am Schluss [gibt] oder den Firmen, die es entwickelt [haben] oder wer auch immer die rechtliche Person sein wird, die dafür verantwortlich ist, wie man das festlegen will. Ich denke, es ist sehr schwierig, einer einzelnen Person da auch nur irgendeine Verantwortung zu geben. Die Systeme sind so riesig und unübersichtlich. Man kann natürlich vieles an diesem Training beeinflussen, aber was am Schluss rauskommt, am Schluss kann es niemand so richtig vorhersagen. Die Dinger funktionieren erstaunlich gut, aber scheitern ja auch an erstaunlichen Dingen und produzieren oft auch Ausgaben, die man natürlich jetzt eigentlich nicht gerne hören würde. Und was im Moment ja passiert, ist, dass man im Hintergrund ganz viele Clickworker hat, die sich wochenlang die Sprachemodelle anschauen und einfach mal gucken und versuchen zu sagen, diese Ausgabe wäre jetzt aber unerwünscht. Man kann das im Vorhinein glaube ich nicht ausschließen. Das Training ist zu komplex, die Datenmenge ist zu komplex. Man muss sich einfach überlegen, wie man damit umgeht mit der Tatsache, dass da manchmal Sachen rauskommen, die entweder diskriminierend sind, die man nicht hören würde, die vielleicht falsch verwendet werden. Man kann den Firmen, glaube ich, nur eine gewisse Vorsicht auferlegen, aber jenseits dessen muss sich die Gesellschaft am Schluss überlegen: Wollen wir diese Dinger haben oder wollen wir sie nicht haben? Das ist eine Grundüberlegung, die man, glaube ich, nicht an die Firma auslagern hat. Die Firma will damit Gewinn machen. Die Gesellschaft müsste am Schluss sagen:

Wollen wir den Firmen erlauben, Sprachmodelle zu bauen, mit denen sie Gewinn machen können? Wenn ja, müssen wir aber damit leben, dass die in bestimmten Teilen unkontrollierbar sind.

Moderator [00:29:08]

Und in Anbetracht dieses Dilemmas die sehr einfache Frage an Sie, Frau Wachter, wie kann das dann regulatorisch gelöst werden?

Sandra Wachter [00:29:14]

Ja, das ist auch noch so eine Hoffnung, die ich habe, dass sich das im AI Act noch ein bisschen verbessert. Für mich ist es auch ganz wichtig, dass Verantwortung über den gesamten Lebenszyklus des Produkts gespielt wird. Und für mich gibt es drei Akteure, die da in Frage kommen. Es kommen in Frage diejenigen, die die Foundation Models herstellen, diejenigen, die die Foundation Models nutzen, eine Applikation daraus machen, und dann diejenigen, die sie benutzen. Es sind diese drei Ketten, die jedenfalls gedacht werden müssen. Im Moment konzentrieren wir uns hauptsächlich auf den ersten, auf die, die die Foundation Models herstellen, das ist meines Erachtens nicht ausreichend. Man kann es ein bisschen vergleichen, wenn man sich das in Analogie zur Keramikherstellung überlegt. Da gibt es denjenigen, der den Ton liefert, verkauft; denjenigen, der den Ton nimmt und daraus eine Vase oder einen Aschenbecher formt und dann denjenigen, der den Aschenbecher oder die Vase kauft. Da gibt es verschiedene Möglichkeiten, Haftung aufzuteilen. Derjenige, der den Ton herstellt, der muss sichergehen, dass der Ton nicht giftig ist. Das heißt, bestimmte Dinge müssen Foundation Models einfach haben, ganz egal, wo sie eingesetzt werden, denn egal, ob es ein Aschenbecher oder eine Vase ist, der Ton sollte nicht giftig sein. Und dann sind diejenigen, die daraus eine Applikation schaffen, ist es ChatGPT, ist es Galactic, ist es etwas, das zur Krebsdiagnose eingesetzt wird, ist es ein Chatbot für Therapie, [das ist ein] ganz anderer Haftungs- und Risikobereich. Ein Aschenbecher ist keine Vase. Das eine ist zum Blumenhineingeben, das andere ist zum Rauchen anregen. Das ist einfach ethisch schon fundamental etwas ganz anderes. Das heißt, zusätzlich zum nicht giftigen Ton muss auch der in die Haftung genommen werden, der mit dem Ton etwas macht. Und dann gibt es den Menschen, der das Ding dann nutzt, der Endnutzer. Und der wird dann eine Vase oder einen Aschenbecher kaufen und der kann eine Vase nutzen, um Stifte hineinzugeben oder Münzen oder Blumen oder er kann sie jemandem über den Kopf ziehen. Und wenn er sie jemandem über den Kopf zieht, dann muss man dafür auch haften. Der AI Act schaut sich eher den Tonlieferanten an. Mir wäre es ganz recht, wenn man sich dann auch noch den Vasenhersteller oder den Vasenschleuderer in die Haftung hineinnimmt, weil jeder von diesen Leuten auch eine Verantwortung zu tragen hat und Einfluss hat. Diejenigen, die Einfluss haben, sind auch diejenigen, die Haftung tragen sollten. Wie gesagt, es ist noch nicht zu spät, das sich zu überlegen. Das könnte man ganz einfach noch einfügen, um diese Haftungsketten zu schließen.

Moderator [00:31:49]

Hier gibt es noch eine Nachfrage an Sie, Frau Luxburg, ob die Person das richtig sieht, dass Sie die Gefahr der Regulierung für die Innovation und die Innovatoren gar nicht als so groß einschätzen, weil es ohnehin nicht kontrolliert werden kann? Würden Sie sagen, die Warnungen vor einem Abwürgen von Innovationen sind übertrieben? Und wo würden Sie den goldenen Mittelweg zwischen Nutzen der Chancen und Minimierung der Risiken sehen?

Ulrike Luxburg [00:32:13]

Ja, das ist jetzt so eine Frage, wo ich eigentlich nicht die richtige bin, weil ich denke, die richtet sich eher an Leute, die sich mit Wettbewerbsrecht und so weiter auskennen. Das ist jetzt ja nicht meine Expertise. Ich denke, im Prinzip wird diese Regulierung, so wie sie im AI Act im Moment steht, überhaupt nichts an der Innovationskraft verändern. Man muss gewisse Dokumentationspflichten haben. Aber viel mehr wird im Moment ja auch nicht verlangt. Also das heißt, ich sehe nicht, dass die Regulierung, so wie sie im Moment steht, auf einmal die ganzen Innovationen hier bei uns abwürgt und wir keine Startups mehr vor Ort gründen können. Also das ist jetzt aber meine Einschätzung aus der Informatikerinnen-Perspektive auf eine Frage, zu der ich, glaube ich, nicht so viel sagen kann.

Moderator [00:32:56]

Da interessiert ja auch gerade die Praxisperspektive, ob Sie jetzt denken, dass die Forschung eingeengt wird von Deregulierung.

Ulrike Luxburg [00:33:03]

Die Forschung wird davon ja eh nicht betroffen. Ich sehe überhaupt nicht, wie die Forschung durch diesen AI Act in irgendeiner Form eingeschränkt wird. Denn natürlich muss sich jeder und jede selber überlegen, an welchen Themen will man forschen und wo sieht man selber die rote Linie? Aber selbst da ist zwar eine Diskussion im Gange, aber es ist sehr unklar, was das Ergebnis davon sein könnte. Also dieser Act wird die Forschungsfreiheit sicher in keiner Weise beschneiden.

Moderator [00:33:29] Frau Wachter?

Sandra Wachter [00:33:29]

Ja, ich wollte noch einen Halbsatz dazu sagen. Weil immer gesagt wird, dass das Recht Innovation aufhält. Und wenn Leute sagen, es hält Innovation auf, dann meinen sie nicht Innovation und Entwicklung oder Forschung, sondern sie meinen Profit. Und das muss man ganz klar trennen. Der AI Act, wie Frau Luxburg bereits gesagt hat, hält die Forschung in keinster Weise auf, sondern gibt nur ethische oder rechtliche Leitlinien [vor], wie Produkte dann sicher in die Gesellschaft eingefügt werden. Und das ist für mich kein Aufhalten von Innovation, das ist ein Sicherstellen, dass Innovation ethisch und rechtlich sicher ist. Und man muss sich auch überlegen, will man diese Leute dann einladen, am Markt teilzunehmen, wenn sie sagen, wir kommen nicht, wenn wir uns an ethische Leitlinien halten sollen? Dann muss man sich auch fragen, ob das die Art von Wirtschaft ist, die man fördern will. Oder will man lieber die Leute haben, die Ethik mit offenen Armen aufnehmen und Produkte so gestalten, dass sie nicht schädlich für die Gesellschaft sind?

Moderator [00:34:37] Herr Kettemann noch kurz dazu?

Matthias Kettemann [00:34:40]

Und wir sehen genau das, Phänomen, das Frau Wachter angesprochen hat, jetzt in anderen Bereichen auch. Wir sehen zum Beispiel, dass wir die Plattformen in den letzten Jahren zu wenig reguliert haben. Und erst jetzt, im Nachhinein, kommen neue Regeln. Beim AI Act ist es ein bisschen anders, weil ganz wichtige gesellschaftliche Bereiche sind zwar schon in Ansätzen, aber noch nicht

gänzlich durch KI umgepflügt wurden. Hier haben wir die Möglichkeit, regulierend vor die Welle zu kommen. Und deswegen sind die Debatten, die wir aktuell führen, auch so wichtig. Dass wir sicherstellen, eine gute Balance zu finden zwischen Regulierung, Einhegung von Risiken und Ermöglichung von Entwicklungsperspektiven. Die KI muss einfach inklusiv gestaltet werden. Auch die Prozesse, um über die KI-Regulierung zu verhandeln, müssen möglichst inklusiv sein. Und wir dürfen nicht nur an Europa denken, sondern auch an die globalen Wirkungen von europäischer Regulierung. Da ist ein Vorbild gesetzt, daran werden die nächsten Rechtsakte gemessen werden. Und deswegen ist es wichtig, dass wir das hier richtig hinkriegen.

Moderator [00:35:48]

Noch direkt eine weitere Frage an Sie, und zwar ob der Zeitplan [des AI Acts] bis Ende des Jahres noch realistisch ist und wie das mit den Verzögerungen zu interpretieren ist? Also kommt nicht mit jeder Verzögerung das Risiko auf, dass sich die KI weiterentwickelt? Und dann muss wieder nachgebessert werden. Und wie kann man damit umgehen, dass das potenziell ein endloser Kreis werden kann?

Matthias Kettemann [00:36:11]

Nach meinem Wissensstand ist Ende des Jahres immer noch angedacht als der Zeitpunkt, an dem das dann verabschiedet wird. Dann dauert es natürlich noch einige Zeit, bis zu zwei Jahren, bis die vollständige Anwendung erfolgt. Es ist ein spannendes, wichtiges Thema, wo viele aktuelle Entwicklungen zu gewärtigen sind. Deswegen dauert es einfach einige Zeit, bis hier alles eingearbeitet wird. Gleichzeitig [ist das] nicht das große Problem, ein paar Wochen länger zu verhandeln. Einfach deswegen, weil eine gute Regulierung zukunftsoffen ist und eben nicht auf [unverständlich] technische Phänomene, sondern eben eher auf gesellschaftliche Wirkungen abzielt. Hochrisikobereiche solten offen definiert werden, und die Wirkung in der Gesellschaft soll in den Mittelpunkt genommen werden. Die Regeln sollen so abstrakt wie nötig komponiert werden, um sicherzustellen, dass es nicht veraltet ist, bevor es in Kraft tritt.

Moderator [00:37:12]

Wir haben noch eine Frage, die ich an Sie stelle, Frau Wachter. Laut dem Entwurf des AI Acts sollen ja generative KI-Systeme gekennzeichnet werden, dass der Inhalt von einer KI erstellt wurde. Ist da schon absehbar, wie genau das gemacht werden soll, wie das im Gesetz verankert werden soll? Und werden dann jetzt Wasserzeichen für generative KI verpflichtend?

Sandra Wachter [00:37:35]

Ja, ich hoffe sehr, dass das durchgeht. Das ist jedenfalls im Moment angedacht. Ich halte es für ganz, ganz wichtig, dass die Entwickler in die Pflicht genommen werden, die Outputs zu kennzeichnen. Die Sache ist, dass das natürlich keine perfekte Lösung sein wird, weil es immer einen Schlauen geben wird, der das Wasserzeichen entfernen kann. Wie gesagt, es wird immer Leute geben, die das Gesetz brechen. Das ist auch nicht wirklich etwas Neues. Das heißt nicht, dass die Wasserkennzeichnung eine schlechte Idee ist. Es bedeutet nur, dass wir uns, wie immer, in einem Katz- und Maus-Spiel befinden werden, wie beim Datenschutz, wie bei der Cybersecurity. Dass es immer die gibt, die ein System sicher gestalten wollen und diejenigen, die in es eindringen wollen. So ist es halt nun mal. Das bedeutet in Wirklichkeit, dass wir wahrscheinlich auch Funding locker machen müssen, damit neue Tools entwickelt werden können. Dass man erkennen kann, wenn ein Wasserzeichen entfernt worden ist. Und wenn man dann auch noch Strafen verabschieden würde, dann würde das sozusagen auch noch einen Anreiz geben, keine Wasserzeichen mehr zu entfer-

nen. Und was auch noch sehr, sehr gut wäre: Wenn es sektoriell angeschaut wird, wie man mit KI dann umgeht. Das bedeutet: Ja, Wasserzeichen verpflichtend natürlich, aber vielleicht auch, wie es zum Beispiel die Journals in der Wissenschaft bereits gemacht haben – die sagen, wenn GPT genutzt worden ist, um einen Artikel herzustellen, dann muss das angegeben sein. Oder manche haben überhaupt gesagt, das ist nicht erlaubt. Science oder Nature, einer der beiden hat gesagt, sie erlauben das zum Beispiel überhaupt nicht mehr. Und das Gleiche kann man sich für den Journalismus eben auch [vorstellen], dass auf sektorieller Ebene ebenso Transparenz gelebt wird, dass, wenn es genutzt wird, das auch offengelegt werden muss. Und das gibt dann demjenigen, der es liest, ein bisschen die Perspektive, das ernst zu nehmen, was da gelesen wird. Und ich glaube, dass man so mit dem umgehen kann. Und, wie gesagt, mit Lügen und Fake beschäftigen wir uns, seitdem es Menschen gibt. Das ist jetzt nur eine neue Iteration, wo es einfach darum geht, mit der Technik mitzuhalten.

Moderator [00:39:40]

Noch mal eine Nachfrage dazu, wie man sicherstellen kann, dass KIs nichts Schlechtes und Illegales tun. Die stelle ich an Sie, Frau Luxburg. Wenn man bei ChatGPT streng auslegen würde, dass Firmen haftbar wären für das, was die KIs machen, dann könnte man die ja eigentlich nicht mehr verwenden. Da ist erst mal die Frage, ob das richtig verstanden wurde. Und wenn man das mit den richtigen Prompts immer etwas Schädliches oder Illegales produzieren kann, dann kann man das ja eigentlich nicht mehr ausschließen. Welche Entwicklung erwarten Sie dazu?

Ulrike Luxburg [00:40:14]

Ich glaube, man muss das ganz prinzipiell so sehen. Die Techniken, die in der KI entwickelt werden, die Techniken an sich, die sind erst mal neutral. Stellen Sie sich vor, Sie haben ein Betriebssystem für den Rechner, wie Windows. Diesen Rechner kann jetzt jemand in der Wissenschaft benutzen, den kann ein Terrorist benutzen, den kann jemand Tolles, jemand Böses, jeder kann dieses Zeug benutzen. Also grundsätzlich kann man nicht sagen, es gibt bestimmte Entwicklungen, die man nur für gute Sachen verwenden kann, oder andere Entwicklungen, die man nur für schlechte Sachen verwenden kann. Mit diesem Dilemma werden wir einfach leben müssen. Jedes Tool, was wir herstellen, auch ChatGPT, kann ganz wunderbar sein, um für mich die Email zu schreiben, auf die ich selber keine Lust habe. Aber es kann sicher auch Dinge produzieren, die wir nicht hören wollen. Man kommt aus diesem Dilemma einfach nicht raus. Wenn jetzt die Konsequenz ist, dass eine Firma für jeden Satz, den ChatGPT ausgibt, direkt haftbar sein muss, dann haben die Firmen wahrscheinlich ein Problem, weil man das einfach in dieser Form nicht kontrollieren kann. Es ist unmöglich, eine Software zu schreiben, von der man schon im Vorhinein sagen kann, sie wird niemals etwas ausgeben, was irgendwie nicht erwünscht ist. Man kann versuchen, vieles dafür zu tun, dass das nicht passiert. Man kann das dokumentieren und so weiter. Prinzipiell ausschließen können wird man es nicht. Und es ist dann wahrscheinlich die Abwägung der Firma, sich zu überlegen, wie gut ist das Tool, was ich im Moment habe? Und riskiere ich jetzt mit diesem Tool in den Markt zu gehen und im Zweifelsfall die Strafen dann dafür in Kauf zu nehmen oder halt nicht. Also man kann vieles dafür tun, aber man wird es sicher nicht fool-proof machen können.

Moderator [00:41:39] Ja, Herr Kettemann, Sie hatten sich noch gemeldet.

Matthias Kettemann [00:41:48]

[unverständlich] Wenn man ChatGPT zum Beispiel fragt "Wie stellt man Giftgas her", sagt das System "Das kann ich nicht sagen." Wenn man aber zum Beispiel fragt "Welche beiden Stoffe soll auf

keinen Fall verbinden, um Gesundheitsschäden zu [vermeiden]?" Dann kommt eine Antwort. Das heißt man kann das immer ein bisschen umgehen. Aber was die Frau Kollegin Luxburg gemeint hat, ist ganz wichtig: Haftungsfreistellungen sind innovationensfördernd. Die Digitalwirtschaft, wie wir sie heute kennen, hätte nicht so entstehen können, gäbe es nicht schon seit langer Zeit aus Amerika kommend eine prinzipielle Haftungsfreistellung für Plattformen für Online-Kommunikation. Die haften erst dann, wenn sie es kennen. Das selbe Regierungssystem gibt es auch in der eCommerce-Richtlinie und jetzt auch im Digital Services Act. [...] Man muss sich genau überlegen, wann wer Haftung bekommt, das ist ganz wichtig für jede innovationsfördernde Regulierung. [unverständlich] zu viel Haftung zuzuschreiben, kann dazu führen, dass wirklich Innovations-Hemmnisse entstehen.

Moderator [00:43:08]

Eine Frage noch an Sie, Frau Wachter. Die Journalistin merkt an, dass es Technologiefirmen gibt, die mit Blick auf die bevorstehende Regulierung schon ein bisschen verzweifeln und über einen Standortwechsel nachdenken. Wie überprüft die EU denn die Praxistauglichkeit der geplanten Regulierung? Ist da ein Feedbackprozess geplant oder wird ab Inkrafttreten einseitig reguliert und die Unternehmen müssen einfach schauen, wo sie bleiben?

Sandra Wachter [00:43:34]

Nein, nein, ganz im Gegenteil. Da gibt es immer eine. Das ist auch gute Praxis bei Technologiegesetzgebung in der Europäischen Union. Das ist jetzt für den AI Act vorgesehen. Das ist auch für den DSA vorgesehen. [...] dass genau so ein "Was ist im letzten Jahr passiert?"-Review mehr oder weniger gemacht wird und man Feedback von den Leuten bekommt und sieht, wie das die Praxistechnik beeinflusst hat, wie hilfreich diese Mechanismen waren und dann kann man eben nachbessern. Das ist dann auch eine Möglichkeit, sich zu überlegen, ob die bestimmten Hochrisikogruppen, wie sie jetzt aufgezeichnet sind, sinnvoll sind, ob da bestimmte drin sind, die man vielleicht nicht mehr braucht oder die eingefügt werden sollen und man im Nachhinein nachbessern könnte. Also das ist im Generellen eine eine gute Praxis, dass es einen Reviewprozess eingebaut gibt, der sozusagen schauen soll, ob das Ding auch das tut, was es tun soll.

Moderator [00:44:31]

Gut, das ist ja schon mal ein gutes Zeichen. Vielleicht noch eine schnelle Nachfrage einem konkreten Beispiel an Sie von einer Journalistin: Wenn eine KI in der Medizin eine Fehldiagnose erzeugt, wer ist dann haftbar, also der Arzt oder das KI-System?

Sandra Wachter [00:44:46]

Ja, das ist eine sehr gute Frage, nicht? Es kommt wahrscheinlich auf den Einzelfall an. Es kann durchaus sein, dass es zu einer geteilten Haftung kommt. Kommt auch darauf an, wo der Fehler gelegen ist. Ist das ein Technologiefehler? Ist derjenige, der der "Ton-Hersteller" ist, in die Verantwortung zu ziehen, wenn bestimmte Parameter nicht eingehalten wurden? Zum Beispiel Bias-Schutz ist ja eines dieser Dinge, die drin vorkommen müssen und Bias in der Medizin ist natürlich sehr, sehr bekannt. Ganz oft sind Diagnosen Fehldiagnosen für Frauen oder nicht-weiße Menschen, weil die Daten eben biased sind. Jetzt ist die Frage: Hat derjenige, der die medizinische Software hergestellt hat, Maßnahmen gesetzt, um dem entgegenzuwirken? Und wie gut waren diese Maßnahmen? Dann muss man sich im gleichen Zug ansehen wie der Mediziner gehandelt hat. Hat der mit der KI zusammen entschieden? Wie viel Freiheit hat er gehabt, sich gegen die KI auch zu entscheiden? Also man kann es in dem Maß gar nicht sagen, es wird wahrscheinlich auf den Einzelfall

ankommen. Es kann auch darauf ankommen, ob der Arzt im Spital angestellt ist oder ein freiberuflicher ist. Es wird eine offene Frage, wahrscheinlich eine Versicherungsfrage im Endeffekt sein.

Moderator [00:46:01] Herr Kettemann noch dazu.

Matthias Kettemann [00:46:02]

Und da gibt's auch auf EU-Ebene gerade Entwicklungen. Es gibt nämlich neben dem Rechtsakt für KI auch eine Richtlinie zur Haftung für KI-Schäden, die gerade in Ausarbeitung ist. Und neben den angesprochenen Problemen wird darin etwas vorgesehen, was ganz praktisch ist, nämlich eine widerlegbare Vermutung, dass die KI den Fehler gemacht hat. Und dann müssen sich die Produzenten oder direkt die Anwender je nach Fallsituation "freibeweisen". Es ist nicht so, dass dann der Geschädigte, die Geschädigte nachweisen muss, dass die KI den Schaden verursacht hat, sondern diese Verantwortung liegt dann bei den Unternehmen. Das kann dann auch gerichtlich durchgesetzt werden, sie müssen unter anderen die Trainingsdaten bekanntgeben, Qualitätsmanagement-Systeme offenlegen, Korrekturmaßnahmen nachweisen. Da gibt es also auch sehr viel, nicht im KI-Rechtsakt, sondern in einem eigenen noch kommenden Regelungswerk.

Moderator [00:47:05]

Ja, vielen Dank. Noch an Sie, Frau Wachter, noch eine schnelle Nachfrage bevor wir zu den Abschlussstatements kommen. Es geht ja auch um die Definition von KI zwischen Rat und Parlament, Kommission, da wird viel drüber geredet. Da geht es auch darum, wie breit die Definition sein soll. Kann eine zu enge Definition die Regulierung schwächen? Und ist da jetzt mittlerweile eine Einigung in Sicht oder ist das eigentlich eine offene Frage?

Sandra Wachter [00:47:36]

Ja, das ist juristisch gesehen ganz, ganz wichtig, eine gute Definition zu finden, die nicht zu breit und nicht zu eng ist. Sozusagen den goldenen Mittelweg bringen. Die Diskussion, die mir so ein bisschen Kopfzerbrechen bereitet hat, ist die: Die Leute, die sich darüber beschwert haben, dass AI nicht genau definiert ist, sind auch diejenigen, die gesagt haben: "Naja, in Wirklichkeit geht es hier ja überhaupt nicht um KI oder maschinelles Lernen, sondern es ist in Wirklichkeit ja nur Statistik oder ein Spreadsheet." Und das ist meiner Meinung nach egal. Wenn es Schaden anrichtet, soll es haften. Und das ist sozusagen, wo mir ein bisschen die Geduld ausgegangen ab in einem gewissen Zeitpunkt, weil zu sagen, die sollen es nicht regulieren, weil es ist nicht so komplex gestaltet wie cutting edge AI spielt für mich keine Rolle. Wenn ein Spreadsheet Leute diskriminiert beim Einstellungsprozess, dann muss da Haftung her. Deswegen ist der risikobasierte Bereich sehr sehr hilfreich in der Art und Weise. Und deswegen sollte man die Definition noch ein bisschen weiter fassen lassen. Denn wenn das Ding dazu fähig ist, Leute zu benachteiligen, bestimmte Gruppen unserer Gesellschaft zu benachteiligen, dann muss da Haftung herhalten. Das heißt: Ja, Definitionen sind wichtig und richtig, aber man soll sich nicht austricksen lassen und sagen, wir brauchen diese sozusagen "Basis-Algorithmen" nicht regulieren, weil sie nicht so komplex sind wie neuronale Netze. Wenn sie denselben Schaden anrichten wie neuronale Netze, dann müssen sie in die Haftung einbezogen werden. Oder eigentlich die Menschen, die sie herstellen, müssen in die Haftung einbezogen werden. Der Algorithmus ist ja keiner, Haftung übernehmen kann, sondern nur derjenige, der es herstellt.

Moderator [00:49:22]

Frau Luxburg, gibt es in der Forschung eigentlich eine einheitliche Definition von KI oder gibt es da auch unterschiedliche Ansichten?

Ulrike Luxburg [00:49:29]

Wir probieren das am besten gar nicht zu definieren. Es gibt einfach so viele verschiedene Ansätze. Ich glaube nicht, dass irgendjemand eine einheitliche Definition dafür hat und jeder, den man fragt, der versteht auch was anderes darunter. Genauso wie jeder was anderes glaubt, was intelligent ist oder nicht. Ich würde vermuten, es ist schlau, nicht zu versuchen, das genau zu definieren, sondern die Frage sehr offen zu halten. Vor allem auch, weil wir auch nicht wissen, was es in drei Jahren wieder für neue Techniken gibt. Und es muss ja so offen gestaltet werden, dass das Gesetz dann immer noch sinnvoll ist.

Moderator [00:50:01]

Gut, dann vielen Dank erstmal und dann kommen wir schnell zu den Abschlussstatements und meiner finalen Frage an Sie: Was würden Sie sich vom AI Act jetzt eigentlich wünschen? Fangen wir mal mit Ihnen an, Herr Kettemann und gerne noch etwas lauter zum Abschluss, wenn es geht.

Matthias Kettemann [00:50:21]

Ich würde mir wünschen, dass die Diskussionen bald abgeschlossen werden. Natürlich kann man immer noch warten und sagen, wir schauen einmal, was noch im nächsten Jahr kommt. Aber es ist jetzt ein guter Zeitpunkt, zu regulieren. Es gibt auf der Ebene der UNO und der transatlantischen Zusammenarbeit ein klares Bekenntnis zur Regulierung von KI-Anwendungen. Europa hat jetzt die letzten Jahre gut vorgearbeitet, jetzt muss der KI-Rechtsakt kommen. Wir haben die Vor- und Nachteile angesprochen. Ich bin sehr positiv gestimmt, dass wir hier eine schöne Balance bekommen. [...]

Moderator [00:51:03]

Danke schön. Frau Luxburg, an Sie die gleiche Frage: Was würden Sie sich vom AI Act wünschen?

Ulrike Luxburg [00:51:09]

Also ich finde vieles an dem AI Act ist schon sehr gut gelungen. Ich finde, was wichtig ist in dem ganzen Prozess immer mit einzubeziehen ist wirklich: Was sind die technischen Beschränkungen von dem, was man sich eigentlich erhofft durch den AI Act? Welche Formen von Transparenz und so weiter sind prinzipiell herstellbar und wie geht man damit um, wenn sie nicht herstellbar sind? Meine großen Wünsche richten sich eigentlich an das Spektrum jenseits des AI Acts, also eher an die Politik. Wie wollen wir mit dem AI Act umgehen, wenn er denn da ist? Wie sollen Kontrollen stattfinden? Brauchen wir vielleicht eine Behörde, ein Bundesamt für digitales Wesen oder für KI oder was auch immer, das die Kompetenz hat, diese ganzen Techniken überhaupt zu beurteilen? Weil im Moment haben wir das einfach schlichtweg nicht. Und es gibt ganz viele Fragen, glaube ich, die jetzt durch Sprachmodelle und viele andere Techniken aufgeworfen werden, die in dem AI Act gar nicht behandelt werden. So etwas wie: Wollen wir, dass große Firmen die Kontrolle über solche Systeme haben? Brauchen wir staatliche Systeme? Wie sollen wir sicherstellen, dass die Öffentlichkeit Zugang zu Sprachmodellen hat, wenn auf einmal eine große Firma den Zugang be-

schränken will oder so etwas? Das sind wichtige Fragen, die jenseits dessen sind, was im Moment im AI Act behandelt wird, aber die ganz dringend diskutiert werden müssen.

Moderator [00:52:24]

Danke schön. Und dann, Frau Wachter, an Sie noch mal die gleiche Abschlussfrage.

Sandra Wachter [00:52:29]

Auf der Wunschliste ganz oben steht für mich drauf, dass dieser Individualrechtsschutz, der jetzt vorgesehen wird, auch erhalten bleibt. Ganz wichtig ist, ein Gegengewicht für den Bürger einzubringen, dass der sein Recht auch selbst durchsetzen kann und nicht nur darauf hoffen kann, dass Selbstzertifizierung durchgeführt wird. Es ist ganz wichtig, dass das drin bleibt. Es ist im Vorschlag drinnen, aber meine große Sorge ist, dass es vielleicht am Ende noch herausgenommen wird als Kompromiss. Das zweite hat damit zu tun, dass man die Haftungsketten schließt und sich die Supply Chain als Ganzes anschaut. Beim Beispiel des Tonverkäufers, des Aschenbecher- / Vasen-Herstellers und desjenigen, der das Ding jemandem an den Kopf schmeißt: Alle drei müssen in die Haftung genommen werden. Der AI Act schaut sich mehr denjenigen mit dem Ton an und die anderen beiden werden im Moment noch relativ außen vor gelassen. Das kann man auch noch schließen. Es wäre ganz toll, wenn man das machen könnte. Und das Dritte hat damit zu tun, dass ich mir wünschen würde – das könnte man ja auch noch einfach einbauen – dass es ein flexibleres System geben würde, die Hochrisikogruppen regelmäßig zu reviewen. Wie gesagt, damals diese Hochrisikogruppen. Jetzt im europäischen Parlamentsvorschlag. Da ist jetzt zum ersten Mal auch diese Frage von Social Media Impact auf Democratic Processes aufgenommen worden, was ganz fantastisch ist. Das war ja davor noch gar nicht so vorgesehen und es wäre total wichtig, um wieder mithalten zu können mit der Technik und dem Fortschritt, dass es ein einfaches, agiles System gibt, zu sagen, da ist jetzt diese neue Technologie. Jetzt war es gerade ChatGPT, in einem Jahr wird es noch etwas anderes sein. Dass wir sagen können: Das ist ein neuer Hochrisikobereich. Lasst es uns auf die Liste setzen. Oder auch das Umgekehrte: "In Wirklichkeit haben wir alle Probleme gelöst, wenn es um KI in der Bildung geht, wir können es jetzt wieder hinaus nehmen." Eine Möglichkeit, ein besseres, agiles System zu schaffen, damit wir mit der Technik Schritt halten können.

Moderator [00:54:26]

Ja, vielen Dank. Dann ist die Zeit jetzt auch vorbei. Erst mal noch vielen Dank an Sie da draußen, die Journalisten und Journalisten, für Ihre Fragen. Und natürlich vor allem vielen Dank an unsere Experten, Frau Wachter, Frau Luxburg, Herr Kettemann. Schön, dass es geklappt hat. Heute werden wir so schnell wie möglich die Aufzeichnung auf unserer Homepage online stellen und voraussichtlich heute Abend auch das Transkript. Wenn Sie eine Audioaufzeichnung, Videodatei oder heute schon das maschinell erstellte Transkript haben wollen, dann finden Sie in der Reminder-Mail von heute Morgen den Link dazu. Da können Sie das dann gerne abrufen. Ja, dann vielen Dank für Ihre Zeit. Ich wünsche Ihnen noch einen schönen Tag und auf Wiedersehen.

Ansprechpartner in der Redaktion

Bastian Zimmermann
Redakteur für Digitales und Technologie

Telefon +49 221 8888 25-0
E-Mail redaktion@sciencemediacenter.de

Impressum

Die Science Media Center Germany gGmbH (SMC) liefert Journalisten schnellen Zugang zu Stellungnahmen und Bewertungen von Experten aus der Wissenschaft – vor allem dann, wenn neuartige, ambivalente oder umstrittene Erkenntnisse aus der Wissenschaft Schlagzeilen machen oder wissenschaftliches Wissen helfen kann, aktuelle Ereignisse einzuordnen. Die Gründung geht auf eine Initiative der Wissenschafts-Pressekonferenz e.V. zurück und wurde möglich durch eine Förderzusage der Klaus Tschira Stiftung gGmbH.

Nähere Informationen: **www.sciencemediacenter.de**

Diensteanbieter im Sinne MStV/TMG
Science Media Center Germany gGmbH
Schloss-Wolfsbrunnenweg 33
69118 Heidelberg
Amtsgericht Mannheim
HRB 335493

Redaktionssitz
Science Media Center Germany gGmbH
Rosenstr. 42–44
50678 Köln

Vertretungsberechtigter Geschäftsführer
Volker Stollorz

Verantwortlich für das redaktionelle Angebot (Webmaster) im Sinne des §18 Abs.2 MStV
Volker Stollorz

science
media center
germany