Justice, Baby! Der Podcast zu Recht und Gerechtigkeit
Transkript

Folge #10 KI: Sind Algorithmen die besseren Richter:innen?

Szenischer Einstieg

[Tastatur-Geräusche im Hintergrund]

Podcast-Host Kathrin Schön: Hallo ChatGTP!

Computer-Stimme [weiblich konnotiert]: Hallo! Wie kann ich Ihnen heute helfen?

Schön: Schreibe mir eine Anmoderation für den Podcast „Justice, Baby!“. Das Thema: KI und Gerechtigkeit. Die Zielgruppe: Gen Y. Wir duzen unsere Hörer:innen. Haben zwei Gäste. So, Punkt. Bringe ein aktuelles Beispiel mit KI und Justiz. Und versuche kurze Sätze zu schreiben.

Computer-Stimme: Klar, kein Problem: Hi und herzlich Willkommen zur neuen Folge von „Justice, Baby!“, dem Podcast für Menschen, die sich für Recht und Gerechtigkeit interessieren. Wir haben heute zwei tolle Gästinnen dabei, die uns helfen werden, das Thema KI und Justiz zu beleuchten. Aber was hat KI eigentlich mit uns zu tun? Ganz einfach, KI wird immer mehr Teil unseres Lebens und beeinflusst auch die Gerichtsbarkeit. Wir wollen heute darüber sprechen, welche Auswirkungen das haben kann und wie wir sicherstellen können, dass wir auch in Zukunft fair behandelt werden.
Ein aktuelles Beispiel dafür ist das umstrittene Risikobewertungssystem Kompass, dass in den USA eingesetzt wird. Es wird behauptet, dass es unfaire Vorurteile verstärkt und somit die Ungleichheit in der Justiz verschärft. Aber ist das wirklich so? Das werden wir heute gemeinsam herausfinden.

[fröhliches, Upbeat Intro ertönt und läuft im Hintergrund weiter]

Anmoderation

Schön: Fragen über Fragen. Und eigentlich wäre ich jetzt ja überflüssig, aber ich hoffe, ihr habt ein bisschen mehr Spaß daran, mir zuzuhören als einer KI.
Man kommt an ChatGPT und der Diskussion über die Vor- und Nachteile, aber auch die Gefahren von Künstlicher Intelligenz ja gar nicht mehr vorbei. Deswegen nehmen wir heute mal die Möglichkeiten und den Einsatz von KI in der Justiz unter die Lupe und checken, wie die überhaupt geregelt ist. Dazu spreche ich mit Victoria Guijarro Santos von der Uni Münster und mit Lajla Fetic von der Bertelsmann Stiftung.
Mein Name ist Kathrin Schön, ihr hört „Justice, Baby! - den Podcast zu Recht und Gerechtigkeit“! Schön, dass ihr wieder rein hört.

[Intro blendet aus]

Kann Künstliche Intelligenz Gerichtsprozesse fairer machen? – Interview mit Rechtswissenschaftlerin Victoria Guijarro Santos

Schön: Vor dem Gesetz sind alle Menschen gleich. Aber in manchen Fällen ist der persönliche Eindruck vor Gericht bei einem Richterspruch das Zünglein an der Waage. Wird zum Beispiel Reue gezeigt oder war die Straftat bloß ein Ausrutscher? Bei der Bestimmung der Höhe vom Strafmaß können Richter:innen mildernde Umstände anerkennen. Dabei sind natürlich auch sie nicht frei von Vorurteilen. Also je nachdem, wo man eine Straftat in Deutschland begeht, kann es tatsächlich sein, dass das Urteil härter ausfällt als woanders. Wir haben dazu schon mal eine Folge gemacht. Wenn ihr mehr dazu wissen wollt, dann hört doch einfach mal bei Folge 2 zum Thema Strafe rein.
Aber nochmal zurück, wenn Richter:innen nicht ganz frei von Vorurteilen sind, könnten technische Innovationen dann vielleicht helfen, vergleichbarer und fairer zu urteilen? Das will ich jetzt von Victoria Guijarros Santos wissen. Sie ist wissenschaftliche Mitarbeiterin am Lehrstuhl für Internationales Öffentliches Recht und Internationalen Menschenrechtsschutz an der Uni Münster und forscht dort zum Einsatz von KI in Entscheidungsverfahren.
Hallo Victoria, herzlich willkommen und schön, dass du da bist.

Guijarros Santos: Hallo, ich freu mich! [lacht]

Schön: Victoria, was ist für dich persönlich gerecht?

Guijarros Santos: Das ist eine sehr große Frage. [lacht] Wenn ich sie in Bezug auf Künstliche Intelligenz und Daten denke, dann wäre es für mich gerecht, dass wir in demokratischen Verfahren darüber entscheiden, wie wir mit Technologien leben wollen. Und nicht nur diejenigen, die zufällig Dateninfrastrukturen in ihren Händen halten und die benutzen, um ohnehin schon marginalisierte Personen noch weiter an den gesellschaftlichen Rand zu drücken.

Schön: Damit bist du direkt im Thema unserer heutigen Folge. Künstliche Intelligenz begegnet uns inzwischen ja echt überall in unserem Alltag. Also von, keine Ahnung, zielgerichteter Werbung auf Social Media, über Spracherkennungssoftware bis hin zu Chatbots oder unserer Spotify Playlist. Und das finde ich ganz schön krass, weil einerseits sollen mithilfe von KI, Angebote und Dienstleistungen besser auf die Bedürfnisse und Präferenzen von Nutzer:innen zugeschnitten werden. Und andererseits schwingt, wenn man über den Einsatz von KI nachdenkt oder spricht, also da wo Entscheidungen getroffen werden sollen, immer auch die Erwartungen mit, dass KI für mehr Neutralität und dadurch mehr Gleichheit und Gerechtigkeit sorgen könnte. Aber geht das überhaupt? Kann eine KI, das überhaupt?

Guijarros Santos: Da würde man jetzt je nach Person ganz unterschiedliche Antworten bekommen. Google, Microsoft, Facebook würden einem natürlich erzählen, dass das geht. Und wenn man sich dann aber näher damit befasst und auch mal dahinterguckt, stellt man ziemlich schnell fest, dass das nicht funktioniert. Und ich glaube, dafür muss ich vielleicht ein paar Einschränkungen und ein bisschen mehr darüber sagen, was KI eigentlich ist und was ich jetzt hier auch im Podcast damit meine. Also zuerst einmal beschäftige ich mich mit Systemen, die Personen klassifizieren, die von Unternehmen entwickelt werden, damit andere Unternehmen oder der Staat, Personen klassifizieren können und danach ausgerichtet Güter, Dienstleistungen, Werbung oder Teilhabe allozieren. Und diese Systeme sind häufig Modelle maschinellen Lernens oder ganz einfache statistische Modelle häufig auch. Dann ist ganz

entscheidend, dass diese Systeme auf Daten entwickelt werden. Und diese Daten kommen nicht aus dem Nichts, sondern irgendjemand anders, irgendein Unternehmen in den meisten Fällen, hat entschieden, dass wir diese bestimmten Daten generieren, dass wir die speichern, dass wir die labeln. Und entscheidet dann darüber, wie wir die benutzen, für wen und gegen wen wir die benutzen.
Und dieses Verständnis, dass Daten aus einem sozioökonomischen Prozess geleitet werden und dass die nicht aus dem Nichts kommen, dass ist total wichtig zu verstehen, weil man dann auch viel besser verstehen kann, dass allein schon die Frage, ob wir nur Daten über arme Personen haben, zum Beispiel jetzt im Bereich von Sozialleistungen, wo auch automatisierte Entscheidungsverfahren genutzt werden, dass das auch schon, den Grundstein dafür legt, dass man dann, wenn man Sozialbetrug aufdecken will, dass es dann häufig darum geht, Sozialbetrug unter armen Personen aufzudecken und nicht darum geht, Steuerbetrug unter reichen Personen aufzudecken.
Und jetzt in Bezug auf Amazon, zum Beispiel jetzt die Werbung, die uns da angezeigt wird, dass in den Fällen sind das eben die Daten aus unserer Realität und wenn unsere Realität von Geschlecht, race, Behinderung, von diesen Machtverhältnissen durchzogen ist und so wirkmächtig ist, dass sie auch unsere Präferenzen steuert, dann wird das in diesen Daten wiedergegeben und wenn wir dann die Modelle darauf trainieren, dann geben diese Modelle auch diese Herrschaftsverhältnisse wieder. Und deshalb ist diese ganze Idee, dass KI neutral wäre, dass sie fair wäre, dass sie objektiv wäre, von vornherein unsinnig. Und ein ganz großer PR-Gag eigentlich von diesen großen Datenunternehmen, die uns diese Modelle verkaufen wollen.

Schön: Die Gerechtigkeit, die Fairness, das Potenzial der KI steht und fällt mit den Daten, mit denen sie trainiert wird.

Guijarros Santos: Genau, aber nicht nur das. Und es ist nicht nur die Frage, welche Daten ich benutze. Sondern auch die Frage, was für Fragen stelle ich diesen Daten eigentlich. Also ich finde wirklich dieses Beispiel, von welche Straftaten wollen wir aufdecken, indem wir KI benutzen, glaube ich ganz anschaulich.
Und es ist auch ein Beispiel, was in den Niederlanden benutzt wurde oder auch genutzt wird, da gibt es ein algorithmisches System, das heißt SyRi, wie das iPhone nur mit einem Y, und da soll vorhergesagt werden, wer Sozialbetrug begeht. Also da sind schon ganz viele Fragen, die man sich stellen kann. Kann man überhaupt aus Daten und mit dem, was dieses Modell maschinellen Lernens erfassen kann aus den Daten, kann man damit überhaupt herausfinden, ob eine einzelne Person Sozialbetrug begeht? Weil das, was diese Modelle machen und das, was sie in den Daten finden wollen, sind Muster, und zwar Muster in diesen vergangenen Daten über eine große Masse an Personen. Und das wird dann verwendet, um über Individuen in der Gegenwart eine Einzelentscheidung zu treffen. Wo schon mal eine relativ große Lücke ist, zwischen dem also, was wir von diesem System wollen, und dem, was dieses System eigentlich kann.
Und gerade bei diesen Fragen mit Sozialbetrug kommt dann auch noch hinzu, dass es dann um Sachen geht, wie, man wird dann als Risiko geflaggt, wenn man diese tausenden Formulare, die man von der Behörde ausfüllen soll, wenn man da mal irgendwie so ein paar Fehler gemacht hat, dann ist es schon verdächtig. Und das ist natürlich noch mal viel prekärer, wenn man überlegt…- Also SyRi wurde in den Niederlanden dann auch noch vor allem in Vierteln

verwendet, in denen migrantische Personen leben, also Personen, die auch Niederländisch vielleicht nicht unbedingt perfekt beherrschen und es zu mehreren Fehlern kommen kann.

Schön: Jetzt sprichst du ja schon von einem ganz konkreten Anwendungsszenario. Ich würde gerne wissen, in welchem Entscheidungsverfahren maschinelle Lernsysteme denn sonst schon eingesetzt werden und wie genau die funktionieren, vielleicht auch im deutschsprachigen Raum?

Guijarros Santos: Mhm, also diese Frage ist immer ganz bisschen tricky und ich finde es auch wichtig, dass hervorzuheben, dass viele dieser Systeme von privaten Unternehmen entwickelt werden. Und auch von Unternehmen oder vom Staat eingesetzt werden und niemand in dieser Kette hat ein sonderlich großes Interesse daran, dass preiszugeben. Deshalb ist es häufig relativ schwierig herauszufinden, wer diese Systeme eigentlich schon benutzt und um die Frage aber trotzdem zu beantworten, [lacht] es gibt nämlich…- und das ist möglich, man kann diese Frage beantworten, weil es Investigativjournalist:innen gibt, weil es NGOs gibt, weil es Forscher:innen gibt, die sich die Mühe machen, trotzdem irgendwie irgendwas herauszufinden. Und darauf basierend weiß ich, dass in Deutschland zum Beispiel beim Bundesamt für Migration und Flüchtlinge, wird ein Spracherkennungstool benutzt, dass den Dialekt von Geflüchteten erkennen soll und danach dann bestimmt wird, vorher eine Person kommt. Und dabei ist dann das große Problem, dass dieses Spracherkennungssystem funktioniert, nur…- also hat eine statistische Genauigkeit von 80%. Und es ist auch nicht so ganz klar, was diese statistische Genauigkeit von 80% eigentlich bedeutet.
Also was die Datengrundlage war, was für Dialekte waren da drin, waren das unterschiedliche syrische Dialekte, unterschiedliche ägyptische Dialekte oder war das jetzt irgendwie einmal Hocharabisch und dann Syrisch und dann 80% lag es richtig? Und selbst wenn es irgendwie nur Hocharabisch und ein syrischer Dialekt war, sind 20% immer noch falsch. Und dann kommen viele Geflüchtete auch nicht nur aus Ländern, wo man Hocharabisch spricht oder einen anderen Dialekt hat. Also es funktioniert ganz häufig nicht und das ist ganz dramatisch, weil es in Asylverfahren häufig auf die Glaubwürdigkeit von Geflüchteten ankommt. Und dass wenn dann irgendwie so eine KI sagt, die Person sagt zwar, sie ist aus Syrien, ist sie aber gar nicht, dabei ist sie aus Syrien, aber war halt einer dieser negativen Fälle, dann bestimmt das das ganze Verfahren aufgrund von so einer Software. Und sonst weiß man auch unter Privatem, das eine ganz bekannte Beispiel sind Bewerbungsalgorithmen, also Systeme die genutzt werden, um Bewerbungen zu screenen und auszusortieren. Und auch da weiß man aber ziemlich wenig, wer das jetzt eigentlich schon benutzt. Es gibt unfassbar viele Angebote, aber, ob die Unternehmen das dann auch wirklich machen, ist noch relativ unklar, man weiß aber, dass sie sich auch schon darauf vorbereiten.

Schön: Im Grunde geht es da um Erleichterung und um Effizienz. Aber wem nutzen diese Algorithmen tatsächlich?

Guijarros Santos: Ja, das finde ich eine total wichtige Frage. Also gerade wie sie benutzt werden. Nutzen sie denjenigen, die ohnehin schon viel Macht in ihren Händen halten und werden genutzt, um Migrant:innen, um rassifizierte Personen, um Frauen, queere Personen noch weiter an den gesellschaftlichen Rand zu drücken? Und warum passiert das? Ist das BAMF inhärent böse? Wahrscheinlich nicht. Das BAMF will, dass diese Asylverfahren möglichst

effizient vonstattengehen und daran haben ja natürlich auch Asylbewerber:innen ein Interesse. Niemand will für mehrere Jahre in so einer Unterkunft leben. Und nicht wissen, was in der Zukunft passiert. Das ist eine total prekäre Situation und alle wollen diese Situation ändern. Es gibt also ein hohes Interesse, dass diese Verfahren möglichst schnell ablaufen. Aber dafür muss man nicht unbedingt diese Technologie benutzen. Man kann auch mehr Übersetzer:innen einstellen. Es gibt auf jeden Fall viele andere Wege, anstatt einer Software, die in 20% der Fällen falsch liegt.

Schön: Es gibt ja auch schon einige Einsatzbereiche von KI in Rechtsverfahren, also ob im Steuerrecht oder in so digitalen Vertragsbaukastensystemen oder eben bei legal Chatbots. Die Algorithmisierung ist also voll im Gange, auch in der Justiz. Wir haben nämlich in Folge 2 gehört, dass es in Deutschland regionale Unterschiede gibt bei der Festlegung des Strafmaßes. Und wäre es da nicht fairer, wenn eine KI die Urteilsfindung zumindest begleitet und Richter:innen in Frankfurt, wie Richter:innen im gleichen Fall, in ähnlichen Fall in München oder Schleswig-Holstein geurteilt haben? Weil auch da spielen persönliche Biases gegenüber Angeklagten eine große Rolle.

Guijarros Santos: Ja, auf jeden Fall. Es gibt also das Bewusstsein, dass Entscheidungen unterschiedlich sind. Dass sie den Eindruck erwecken, dass irgendwie willkürlich entschieden wird, dass es total drauf ankommt, welche Richterin jetzt vor einem sitzt, also woher sie kommt, was für Biases hat sie. All diese Dinge. Und dann setzen wir jetzt wieder KI dazwischen, und die löst dann das Problem für uns. Und ich glaube, wir müssen einfach wirklich sehr, sehr darauf achten, was für ein System man einsetzt und wofür man das einsetzt. Und wie sehr auch dann die Richter:innen trainiert werden, diese Systeme zu verstehen.
Ich finde es toll, wenn algorithmische Systeme oder KI-Systeme genutzt werden, um Entscheidungsverfahren zu analysieren und dadurch zu begleiten. Und damit meine ich, dass zum Beispiel auch einfach mal diese ganzen Gerichtsurteile überhaupt in Daten übersetzt werden, also Daten lesbar gemacht werden, dass sie dann ausgewertet werden können. Und dass wir dann auch überhaupt eine gute Grundlage haben, um zu sagen, pass auf, Kammer X im Gericht X, ihr entscheidet häufig so und so. Kammer Y in Gericht Y entscheidet aber häufig so und so. Warum passiert das eigentlich? Warum entscheiden wir so unterschiedlich? Was sind unsere Entscheidungsparameter für denselben Fall?
Das fände ich super, wenn man algorithmische Systeme für so was verwendet. Stattdessen, was jetzt passiert ist aber, auch angeleitet von diesem Effizienz-Paradigma, was wir schon angesprochen haben, auch Strafverfahren dauern Ewigkeiten, auch da müssen Personen Ewigkeiten drauf warten, was jetzt in ihrer Zukunft passiert, werden sie in Freiheit leben oder nicht, wieviel Geld müssen sie bezahlen oder nicht? Werden sie bestraft oder nicht? All diese Dinge, all diese Verfahren dauern Ewigkeiten, und auch da haben natürlich die Angeklagten ein Interesse, dass das schnell abläuft. Die Gerichte sind überlastet, also auch das soll jetzt erleichtert worden. Und dann wird jetzt einfach eine Software implementiert, die dann eine Entscheidungshilfe macht in dem Sinne, dass sie den Richter:innen anzeigt ja oder nein. Oder Freiheitsstrafe, so oder so hoch oder Schuld, so oder so. Und mir fällt es schwer, es anders zu sagen, aber das ist es einfach Quatsch. Was sie machen, so wie sie jetzt gerade verwendet werden, Muster in den Daten über die Vergangenheit erkennen, über eine große Personengruppe. Und das wird dann auf dem Individuum in der Gegenwart, in der

Einzelfallentscheidung genutzt. Und dadurch werden eben Stereotype und Muster auch auf dieses Individuum übergestülpt.

Schön: Wie kann man sich denn jetzt dagegen wehren, wenn man von einer falschen Entscheidung durch eine Künstliche Intelligenz betroffen ist? Wie ist die aktuelle Gesetzeslage in Deutschland, aber auch in der EU?

Guijarros Santos: Es gibt in der Datenschutzgrundverordnung eine Norm, die grundsätzlich verbietet, dass automatisierte Entscheidungen getroffen werden, wenn sie eine hohe persönliche Relevanz haben. Ich formuliere das jetzt ein bisschen anders als es im Gesetz steht, aber mehr oder weniger ist es da drin. Strafverfahren, Asylverfahren, Bewerbungsverfahren haben eine hohe persönliche Relevanz. Aber dann müssen die Verfahren auch voll automatisiert sein, also eine Software muss komplett entscheiden, ja oder nein und nicht nur eine Empfehlung abgeben. Und daran scheitert dann häufig schon diese Anwendung dieser Norm, weil sie eben nur auf vollautomatisierte Verfahren zutrifft und nicht auf teilautomatisierte Verfahren, wenn also zwischen der Ausgabe einer Software und der letztendgültigen Entscheidung noch eine Caseworkerin, eine Richterin, eine Personalerin sitzt. Und wenn es aber voll automatisiert wäre, dann wäre die Regelung, dass man ein Recht auf den eigenen Standpunkt hat. Was ich auch sehr kritisch finde, weil dann eine einzelne Person darlegen muss, ich bin ganz anders als die Norm. Ich bin ganz anders als das Muster, was dieses algorithmische System in den Daten erkannt hat. Was einmal, glaube ich, kaum jemanden überzeugen wird, also zu sagen, ich bin ganz anders als das, was diese superfancy AI gerade ausgespuckt hat. Deshalb finde ich diese DSGVO-Norm, selbst wenn sie auf diese Systeme anwendbar wäre, sehr kritisch.
Und dann gibt es noch das Allgemeine Gleichbehandlungsgesetz, das EU Antidiskriminierungsrichtlinien umsetzt und es immer dann greift, wenn eine betroffene Person vermutet, dass sie diskriminiert worden ist, also dass sie im Einzelfall individuell benachteiligt worden ist. Aufgrund benannter Antidiskriminierungskategorien wie Geschlecht, race, Alter, Behinderung, Religion. In den Fällen ist es dann so, ich muss überhaupt erstmal wissen, dass ich von einem algorithmischen System bewertet wurde, dann muss ich die Vermutung haben, dass ich diskriminiert worden bin, also dass ich aufgrund dieser benannten Kategorien individuell benachteiligt worden bin, aber es reicht nicht nur, dass ich dieses Gefühl dafür habe, sondern ich muss auch ein Gericht davon überzeugen, dass ich vermute, diskriminiert worden zu sein. Also der Zugang zum Recht, ist schon eine unfassbar hohe Hürde. Und wenn man dann diese Hürde aber geschafft hat, wenn man also einmal in diesem Gerichtsverfahren drin ist, dann haben die Gerichte in Deutschland die Tendenz, immer nur diesen Einzelfall zu behandeln und immer nur darüber zu sprechen, ob jetzt in dieser Entscheidung, wo die Personalerin, also denkt man an einen Bewerbungsalgorithmus, wo die Personalerin entscheiden musste Bewerberin ja oder nein, ob sie in dem Moment auch von dieser Software angeleitet war und auch deswegen dann eine Entscheidung getroffen hat gegen die Bewerberin und dann kommt es wieder darauf an, ob diese Software, wie es in der *Machine Learning* Forschung heißt, ab die ein Bias hatte oder nicht, dann wird das ganze Ding mega technisch und kompliziert. Und es geht dann nicht darum, ob also so wie die Gerichte das häufig auslegen, ob nicht überhaupt schon der Einsatz dieses Algorithmischen Systems diskriminierend ist. Und ob man deshalb das unterlassen sollte.
Und dann gibt es jetzt auch noch den KI-Verordnungsentwurf der EU. Es gibt zum Beispiel

eine Norm, die darin anschließt, dass ja in der DSGVO, die Norm nur für vollautomatisierte Entscheidungen trifft und in der KI-Verordnung wird vorgeschlagen, dass auch bei, also bei allen Systemen, die KI gestützt sind, dass bei jeder dieser Entscheidungen immer irgendwie eine menschliche Aufsicht da ist und deshalb dann irgendwie auch erlaubt zu intervenieren. Wenn also eine Betroffene meint, ich bin anders, dann könnte sie das sagen und dann könnte aufgrund der KI-Verordnung auch die Nutzerin dieser KI dann intervenieren. Ich finde diese DSGVO-Norm insgesamt nicht sonderlich hilfreich, weil ich nicht glaube, dass das überzeugend ist und ich das auch problematisch finde, dass das Entscheidungsmuster an sich dadurch legitimierbar gemacht wird.

Schön: Findest du KIs prinzipiell deswegen super kritisch oder findest du eigentlich spannend, dass es diese Technologie gibt? Und was wünscht du dir für die Zukunft im Umgang mit KI?

Guijarros Santos: Ich habe auch eine positive Einstellung gegenüber Technologien, auch wenn das nicht so anklang. Ich glaube, dass Technologien auch total viel Kreativität loslösen können, also es gibt jetzt auch gerade diese ganzen Text-zu-Bild-KI-Systeme wie DALL-E oder Stable Diffusion und es macht auch total Spaß, das zu machen, was einzugeben und dann kriegt man witzige Bilder, diese ganzen Sachen auszuprobieren. Ich finde auch, dass KI-Systeme super sinnvoll eingesetzt werden können. Wie ich das vorhin erwähnt hatte, bei den Strafverfahren zum Beispiel, dass man mal die Entscheidungsmuster analysiert und dass man dann auch mal richtig auf einer evidenzbasierten Grundlage darüber spricht, was läuft hier eigentlich gerade falsch, finden wir das falsch oder nicht und das dann anfängt zu korrigieren und die Institutionen zu reformieren. Ich glaube, dass total viel Potenzial darin steckt.
Aber das passiert gerade nicht. So, wie diese Daten jetzt gerade genutzt werden, werden sie eben viel häufiger dazu genutzt, um marginalisierte Personen weiter an den gesellschaftlichen Rand zu drücken, gar nicht beabsichtigt. Aber dadurch, dass unsere Herrschaftsverhältnisse reproduziert werden. Ich finde es auch problematisch, dass diese Infrastruktur, die benötigt wird, um Daten zu generieren, um Daten zu speichern, um Daten zu labeln, dass die Infrastruktur in den Händen wenigen privaten Unternehmen liegt, dass die Öffentlichkeit darauf gar keinen Zugriff hat, dass wir gar nicht mitentscheiden, in was für Beziehungen wir mit Technologie eigentlich leben. Die einzige kleine Entscheidungsmöglichkeit, die wir haben, ist Cookies, ja oder nein. Und ich glaube, dass man das alles anders gestalten kann. Ich glaube, dass wir in einen Prozess eintreten sollten, in dem wir in der Öffentlichkeit viel offener darüber diskutieren, in was für Beziehungen wir mit welchen Technologien leben wollen, ob wir das überhaupt wollen.
Genau, also, wenn ich frage, was mein Ausblick ist, dann hoffe ich, dass wir offener darüber reden. Und dass wir einen Weg finden, Technologien wie KI-Systeme so zu nutzen, dass sie sozial marginalisierte Personen vom gesellschaftlichen Rand loslösen, dass sie uns dadurch allen nützen und dass wir demokratisch darüber entscheiden.

Schön: Vielen Dank für den Einblick in die rechtliche Dimension von KI in Entscheidungsverfahren und danke, dass du dir die Zeit genommen hast!

Guijarros Santos: Ja sehr gerne! Vielen Dank für die Einladung.

[kurze musikalische Zwischensequenz]

Kann Künstliche Intelligenz moralisch handeln? – Interview mit Tech Governance Expertin Lejla Fetic

Schön: Puh, ich muss das erstmal sacken lassen. Ich finde es nämlich schon erstaunlich, dass bisher nicht transparent gemacht werden muss, wo zum Beispiel in öffentlichen Verwaltungen oder eben auch auf dem Arbeitsmarkt, Entscheidungen mit Hilfe von Künstlicher Intelligenz vorbereitet werden. Wie leicht es ist, einer Empfehlung zu folgen oder ihr nicht zu folgen, das merken wir selbst in unserem Alltag. Social Media Portale oder Unternehmen, die Klamotten, Bücher oder Elektronik verkaufen, die schlagen ja gezielt Produkte vor, die uns interessieren können. Und seien wir mal ehrlich, es ist ganz schön schwer, sich von der einen oder anderen Information nicht beeinflussen zu lassen. Das Versprechen dahinter, das verstehe ich schon individuell maßgeschneiderte Lösungen für uns, das kann aber auch in einem ganz anderen Kontext verwendet werden, zum Beispiel wird das gerade bei einem Nachrichtenportal in Kuwait ausprobiert. Dort gibt es seit kurzem eine virtuelle Nachrichtensprecherin, Fedha, so heißt die Frau, soll mit Hilfe von Künstlicher Intelligenz geschaffen worden sein und ein kleines Vorstellungsvideo von ihr gibt es sogar schon im Internet, indem sie klassisches Arabisch spricht. Die Kuwait News testen sie gerade und hoffen dann, dass sie den Nutzerinnen und Nutzern zukünftig personalisierte Nachrichten präsentieren kann, z.B. auch mit einem kuwaitischen Akzent.
Aber wir müssen eigentlich gar nicht ins Ausland schauen, in Deutschland gibt es auch jede Menge KI generierten Content, zum Beispiel Radio Helgoland. Das ist ein online Radiosender, der von sich selbst sagt, dass er wohl der erste Radiosender weltweit ist, der ausschließlich von Künstlicher Intelligenz gesteuert und moderiert wird. Auch die Moderator:innen existieren nur virtuell. Irgendwie ist das ja witzig und gleichzeitig hat Victoria Guijarros Santos ja auch einen guten Punkt. Wir sollten uns nicht blind auf Algorithmen verlassen und sollten vor allem die Einsatzmöglichkeiten von Künstlicher Intelligenz in unserem Alltag mitbestimmen. Weil in dem Bereich im Moment so viel passiert, hinken unsere Gesetze wie immer bei Innovationen, das ist ganz normal, ein bisschen hinterher und haben eigentlich auch noch keine gute Lösung für die ethischen Probleme mit KI. Aber genau über die möchte ich jetzt mit Lajla Fetic sprechen. Sie ist Co-Leiterin des Projekts *Ethik der Algorithmen* bei der Bertelsmann Stiftung und wurde 2021 für ihre Arbeit als eine der One Hundred Billion Women in AI ausgezeichnet. Ich freue mich sehr, dass du da bist! Herzlich willkommen, liebe Lajla.

Fetic: Vielen Dank! Vielen Dank für die Einladung.

Schön: Was ist für dich persönlich eigentlich gerecht?

Fetic: Ja, wie viel Zeit haben wir? [lacht] Ich hoffe, ein paar Tage. Nee, das ist eine total gute und richtige Frage und ich denke häufig darüber nach, weil im Grunde genommen das Projekt, in dem ich arbeite, was ich leiten darf, beschäftigt sich indirekt implizit immer mit Gerechtigkeit und deshalb ist die Frage sehr relevant, was ich darunter verstehe. Und ich kann darunter wahrscheinlich keine zufriedenstellende Antwort geben. Und wieso? Das liegt wahrscheinlich daran, dass ich vor allem den Gedanken von Amartya Sen ganz spannend finde an dieser Stelle der, die Idee der Gerechtigkeit in einem Buch festgehalten hat und dort eine Sache beschrieben hat, die mich seitdem sehr, sehr prägt. Und zwar, dass wir im Grunde

genommen gar nicht so richtig beschreiben können, was Gerechtigkeit bedeutet, sondern vielmehr beschreiben können, was ungerecht ist. Und dass das etwas ist, was im Zweifel auch der Schlüssel zu vielen Fragen ist. Denn ganz am Anfang, so beschreibt er in seinem Buch, entwickeln selbst Kindern ein Gefühl von Ungerechtigkeit und das hält dann auch bis in das Erwachsenenalter und das ist das, was unser Tun leiten sollte am Ende des Tages, nämlich diese Ungerechtigkeit anzugehen und die Welt ein Stück weit gerechter zu machen. Und insofern prägt mich das auch in meinem täglichen Leben und auch in meiner Arbeitswelt, dass ich nicht unbedingt ein Bild von einer absolut gerechten Gesellschaft im Kopf habe, aber mit meinem Tun gerne Ungerechtigkeit ein Stückweit weniger stattfinden lassen will und ein Stückchen dazu beitragen möchte, dass die Welt gerechter ist. Also insofern habe ich ein ganz dynamisches Verständnis von Gerechtigkeit an dieser Stelle und das prägt mich eben entsprechend auch im Kontext der Algorithmen.

Schön: Wir haben im Gespräch gerade eben mit Victoria Guijarros Santos gehört, dass maschinelle Lernsysteme nicht die besseren Entscheidungen treffen, aber dafür strukturelle Ungleichheiten im System aufdecken können. Welche Rolle spielen Ethik und Moral beim Einsatz von KI denn deiner Meinung nach?

Fetic: Also zunächst einmal zum Ethik Begriff. Das ist für mich die Lehre vom guten Handeln und hilft uns analytisch über die normativen Weltvorstellungen an der Stelle nachzudenken und die auch ein Stück weit in Frage zu stellen. Moral wiederum beschreibt das gute Handeln als solches und ist an der Stelle eben normativ, wenn man so möchte. Und im Kontext von algorithmischen Systemen ist Ethik ja nicht wegzudenken. Weil Algorithmische Entscheidungssysteme, so wie ich es bezeichne, ich spreche seltener von Künstlicher Intelligenz, Algorithmische Entscheidungssysteme sind Werkzeuge, die wir nutzen, um unser Leben im besten Falle besser, leichter, schneller, vergnüglicher zu gestalten, und dabei treffen wir implizit und explizit Annahmen darüber, was denn unser Leben besser, schöner, schneller macht, und über diese Annahmen sollten wir sprechen. Denn die sind mitunter nicht für alle gleich. Also insofern kann ich der Vorrednerin, kann ich Victoria nur zustimmen, ja Algorithmische Entscheidungssysteme können Entscheidungsmuster und Wertvorstellungen aufdecken, aber immer nur dann, wenn wir diese Systeme, diese Werkzeuge auch ganz bewusst entwickeln und einsetzen und in der Praxis ist das manchmal eben nicht der Fall, und das wiederum führt zu der Reproduktion von sozialer Ungleichheit, von Vorurteilen. Und damit können mitunter Menschen systematisch diskriminiert oder auch ausgebeutet werden.

Schön: Wie müssen ethische Regeln denn gebaut sein oder konstruiert sein für den Einsatz von Algorithmische Entscheidungssystemen, wie du es jetzt genannt hast, um genau so etwas zu vermeiden?

Fetic: Im besten Falle entwickeln wir Regeln, die nicht für die Systeme, sondern vor allem für die Menschen gelten. Denn was ich häufig in der Diskussion erlebe, und deswegen sprechen wir auch nicht mehr von Ethik der Algorithmen, was ich häufig erlebe, ist, dass diesen Systemen ein Stück weit Verantwortung unterstellt wird, die diese Systeme nicht übernehmen können. Denn am Ende kann nur der Mensch für sein Handeln oder für das Handeln insgesamt verantwortlich sein und dazu gehört auch das Entwickeln technologischer Innovationen. Folglich müssten aus meiner Sicht vor allem die Regeln für Menschen gelten, die die die Systeme entwickeln, aber auch einsetzen oder auch evaluieren. Und da hilft mir im Kopf, das

Bild von soziotechnischen Systemen. Also wir sprechen nicht nur von rein technologischen Systemen, wenn wir von Algorithmen und Künstlicher Intelligenz sprechen, sondern von soziotechnischen Systemen. Das bedeutet, dass sind Werkzeuge, die im sozialen Kontext auch eingebettet sind. Und wenn wir den größeren sozialen Kontext von Algorithmen miteinbeziehen, dann müssen wir Regeln gestalten, die von Anfang an die Entwicklung mit begleiten und auch den sozialen Kontext dabei mit vor Augen halten.
Um vom Abstrakten ein bisschen wegzukommen und konkreter zu werden, weil auch das treibt meine Arbeit oder unsere Arbeit bei der Bertelsmann Stiftung an, es gibt Algorithmen, die beispielsweise helfen, Verkehrsströme zu analysieren und entsprechend Empfehlungen auszugeben, wie man mit diesen Verkehrsströmen umgehen könnte. Also um beispielsweise den Nachmittagsstau in der Stadt an der Stelle zu entlasten. Das hat einen großen sozialen Impact. Wenn ich im Stau bin mit dem Auto, dann bin ich genervt, wenn der Stau gerade da ist, wo eine Spielstraße ist oder eine Fahrradstraße ist, dann hat das auch einen großen Einfluss darauf, wie Menschen zusammenleben. Und wenn wir solche Systeme entwickeln, dann müssen wir ganz am Anfang, uns Gedanken darüber machen, wofür wollen wir denn dieses System eigentlich nutzen? Und insofern ist mit dem ersten Gedanken für die Entwicklung eines solchen Systems eine Regel wichtig, wie beispielsweise, dass man transparent darüber redet, wieso, weshalb, warum entwickeln wir an dieser Stelle. Also mein Ziel ist es im Grunde genommen, dass wenn wir technologische Innovationen entwickeln, wir uns von Anfang an darüber Gedanken machen, was hat es für gesellschaftliche Implikation, wo könnten wir bestimmte ethische Fragestellungen aufkommen lassen und wie müssten diese dann auch breit diskutiert werden?

Schön: Jetzt bist du schon mittendrin im Thema Regeln und Regulierungen und das bringt mich zu dem Punkt, den ich unbedingt mit dir besprechen möchte, denn du hast ja im Grunde ein Regelwerk für den Einsatz von Algorithmen entwickelt, die sogenannten Algo.Rules. Das sind ethische Gestaltungsprinzipien für KIs in Unternehmen beziehungsweise maschinelle Lernsysteme in Unternehmen. Was gehört alles zu diesen Regeln dazu? Und welche Regeln müssen beachtet werden, wenn man eben maschinelle Lernsysteme entwickelt?

Fetic: Ich war nicht alleine bei dieser Entwicklung, es waren über 500 Expert:innen beteiligt an der Entwicklung der Algo.Rules und mit den Algo.Rules sind wir damit dann auch nicht alleine, weil selbst die Europäische Kommission sich in einer High Level Expert Group Gedanken dazu gemacht hat, welche ethischen Prinzipien wichtig wären für die Entwicklung solcher Systeme. Und die Algo.Rules sind trotzdem ein Stück weit anders als die meisten Ethikregelwerke an der Stelle, weil im Grunde genommen kommen wir ohne Ethik aus und das verwirrt oder irritiert erstmal. Diese neuen Gestaltungsprinzipien sind aber ganz bewusst nicht ethisch oder geben moralische Normen vor. Wir haben nämlich in dem Entwicklungsprozess gemerkt, uns geht es nicht darum zu sagen, wann ist ein algorithmisches System gut oder schlecht, das können wir von außen gar nicht beurteilen. Wir möchten aber, dass im Entwicklungsprozess und im Einsatzprozess die beteiligten Personen die richtigen Fragen stellen, um über ethische und gesellschaftliche Dimension ihrer Arbeit nachzudenken.
Das klingt auch wieder abstrakt. Ich mache es mal konkret. Die erste Algo.Rule ist das Thema Kompetenzen aufbauen. Diese erste Regel klingt erst mal ganz easy peasy. In der Praxis ist es aber, glaube ich, eines der wichtigsten und schwierigsten zu erreichenden Ziele, die wir uns an dieser Stelle vorstellen können. Denn was wir meinen mit Kompetenzaufbau ist nicht,

dass die IT-Nerds noch besser werden im Coden, sondern wir stellen uns vor, dass die ITler, die Entwickler:innen ein Gefühl dafür bekommen sollen, welche sozialen Implikationen ihr Handeln hat, also Kompetenzen dahingehend aufbauen, dass sie wissen, für welchen Kontext entwickle ich da eigentlich. Und diejenigen, die Algorithmische Entscheidungssysteme einsetzen, die brauchen ein Stück weit Kompetenzen um technische Entwicklungsseiten der Algorithmische Entscheidungssysteme und sie müssen auch Kompetenzen aufbauen, um die Entscheidungssysteme gut zu kommunizieren, einordnen zu können und entsprechend die Prognosen, die beispielsweise so eine Entscheidungssystem mitgibt, gut interpretieren zu können, und das ist eine der Regeln.

Schön: Gleichzeitig werden die meisten algorithmischen Werkzeuge von Privatunternehmen entwickelt, die ja auch erstmal ein monetäres Interesse daran haben, dass ihre Werkzeuge gekauft, genutzt werden und da ethische Fragestellungen oder soziale Fragestellungen nicht unbedingt im Vordergrund stehen. Wie geht man mit dieser Diskrepanz um?

Fetic: Natürlich ist die Fragestellung, was ist der Anreiz für Unternehmen, solche zusätzlichen Regeln zu befolgen, total wichtig. Und teilweise ist diese Frage noch nicht ganz beantwortet. Aber vielleicht einen Einblick in die aktuelle Lage. Ich glaube für uns Nutzer:innen, Konsument:innen wird immer deutlicher, dass solche Systeme nicht neutral sind, dass solche Systeme auch mit Gefahren und Risiken einhergehen und deshalb wächst ein Stück weit auch der Druck für Unternehmen, entsprechend nachweisen zu können, wir nehmen diese Sorgen ernst und wir entwickeln aktuell unsere Systeme unter besonderen Qualitätsansprüchen. Und insofern glaube ich, dass es durchaus ein Druck der Konsument:innen und Nutzer:innen gibt, der wiederum dann überzeugend auf Unternehmen wirkt. Und gleichzeitig sehe ich auch, dass viele Unternehmen das Thema sehr ernst nehmen, um sich ja besser behaupten zu können im Markt der vielen Anwendungen und auch in einem Markt, der von Monopolunternehmen geprägt ist. Viele Unternehmen, die Algorithmische Entscheidungssysteme entwickeln, die wir häufig nutzen, also gerade im Kontext sozialer Plattformen, kommen nicht aus Deutschland oder nicht aus Europa. Und insofern ist es fast so ein Verkaufsargument für europäische Unternehmen, hier mit besonders großem Wert, also intensiven Wertvorstellungen, auch voranzuschreiten.
Nichtsdestotrotz gibt es auch Unternehmen, die das nicht ganz so ernst nehmen und die glücklicherweise jetzt aber einen Anreiz bekommen, der nicht ganz freiwillig daherkommt, nämlich durch die KI-Verordnung, die auf europäischer Ebene gerade diskutiert wird, die im Grunde genommen unwilligen Unternehmen auch ein Stück weit die Verantwortung aufdrängt, indem sie regulativ entsprechend Maßnahmen vorgibt, aber am Ende kann auch das selbst nicht ausreichen, denn wenn ein Unternehmen so ein System entwickelt und es dann aber von Organisationen oder Menschen eingesetzt wird, die sich nicht der Risiken und der verschiedenen Aspekte eines richtigen Einsatz bewusst sind, kann auch dann viel schieflaufen. Also insofern hier nochmal mein Appell, ich nenne das immer KI-Lebenszyklus, den gesamten Lebenszyklus im Kopf zu behalten. Und die Entwicklung eines solchen Systems, die ist eine total wichtige Komponente. Die Anwendung eines solchen KI-Systems ist aber fast genauso wichtig und da wissen wir, dass wir Menschen oftmals richtig große Fehler machen, also weil wir eben nicht die rational denkenden Wesen sind, von denen wir gerne hätten, dass wir es wären, sondern Entscheidungen auf Basis von Biases, also Verzerrungen im Kopf treffen und selbst bei der Unterstützung durch Algorithmen werden wir ein Stück weit von

diesen Biases gelenkt und treffen teilweise auch trotz automatischer Unterstützung die falschen Entscheidungen. Also insofern, es gibt viele Momente, wo es schief laufen kann. Gleichzeitig bleibe ich da optimistisch, weil die Sensibilisierung die Aufmerksamkeit gerade so groß ist für die potentiellen Gefahren, dass ich glaube, dass wir da zu einem guten Ende kommen.

Schön: Was sind die inhaltlichen Eckpfeiler von der Europäischen KI-Verordnung? Könntest du da nochmal ins Detail gehen?

Fetic: Also die KI-Verordnung beschäftigt sich allgemein grundsätzlich mit der Entwicklung und dem Einsatz von KI-Systemen, weil wir gesehen haben, dass bestimmte Gefahren und Risiken immer wieder auftauchen, unabhängig der konkreten Anwendung. Wir sehen, dass Algorithmische Entscheidungssysteme ja technisch fehlerhaft sein können und in der konkreten Anwendung dadurch zu Gefahren führen. Wir sehen, dass also auch ethische Fehler oder Diskriminierung auftauchen können. Beispielsweise werden bestimmte Nutzer:innen-Gruppen diskriminiert oder deren Belange nicht ausreichend in der Entwicklung der Systeme miteinbezogen. Also beispielsweise bei Gesichtserkennungstechnologien. Also solche Systeme sind nicht fehlerfrei, wir wissen vor allem von People of Color aus den Staaten, die wissenschaftlich dazu gearbeitet haben, dass solche Systeme die Gesichter von schwarzen Frauen nicht gut oder nicht gleich gut erkennen und insofern sind diese Gesichtserkennungstechnologien diskriminierend, gegenüber bestimmten Bevölkerungsgruppen oder bestimmten Merkmalen ganz besonders.
Und das kann mir als Einzelperson teilweise dann nicht auffallen, weshalb es wichtig ist, dass wir eine Regulierung haben, die grundsätzlich dafür Sorge trägt, dass solche großen grundsätzlichen Fehler auch adressiert werden, und dafür ist die KI-Verordnung da. Und was sie macht ist, die Unternehmen zu verpflichten, bestimmte Anforderungen zu treffen. Klingt total abstrakt. Ich mach‘s mal deutlich. Also wenn ein System ein besonders großes Risiko mit sich bringt, beispielsweise weil es in dem Personal Recruiting Kontext eingesetzt wird, dann muss das Unternehmen, was diese Systeme entwickelt und das Unternehmen, das solche Systeme einsetzt, dafür Sorge tragen, dass die Anforderungen, die wir in der KI-Verordnung haben, auch erfüllt werden. Das sind dann Anforderungen, beispielsweise der Transparenz da, das heißt, man muss darüber informieren, dass das System eingesetzt wird, dass diese Systeme unter menschlicher Aufsicht sozusagen ihr Werk tun können und im Grunde genommen soll das dann entsprechend absichern, dass in besonders risikohaften Bereichen solche Systeme Entscheidungen treffen oder Entscheidungen vorbereiten, die wir einfach nicht ungesehen haben wollen.
Trotzdem gibt die aktuelle KI-Verordnung auch schon…- weißt sie schon ein paar Lücken auf, denn die große Frage wird natürlich sein, was verstehen wir eigentlich unter KI, was wird denn da eigentlich reguliert und auch wie man eben das Risiko genau bemisst? Und das wird gerade auf europäischer Ebene heiß diskutiert. Also wie breit, wie eng ist die Definition, sind darunter nur Systeme zu verstehen, die lernen oder sind es vielleicht auch Systeme, die eher statisch sind, also denen man Regeln vorgibt und die diese dann einfach nur umsetzen und im Grunde genommen, was bedeutet Risiko, wer muss das eigentlich anzeigen, wer darf eigentlich darüber bestimmen, welche Systeme Hochrisikosysteme sind und welche eben kein so großes Risiko darstellen?
Und dann der dritte Schritt ist, wenn beispielsweise Anforderungen wie menschliche Aufsicht

an der Stelle stehen. Was bedeutet denn menschliche Aufsicht? Sitzt da jemand daneben und schaut den maschinellen Systemen beim Lernen zu? Ist das menschliche Aufsicht? Oder ist es eher die Person, die das am Ende des Tages anwendet? Hat die eine Schulung bekommen? Und ist dann sozusagen die Aufseherin des Entscheidungssystems? Da sind noch viele Sachen unklar.

Schön: Wie werden aber die tatsächlich Betroffenen in die Entwicklung von diesen Regulierungen einbezogen? Also gibt es da eine Art Bürger:innen Rat oder eine Art Betroffenen Rat?

Fetic: Beispielsweise gibt es bei europäischer Regulierung immer die Möglichkeit, sich in den Konsultationsprozessen zu beteiligen. Wir wissen aber auch aus anderen Politikfeldern, dass sich am Ende vor allem Unternehmensvertreter:innen beteiligen und nicht jetzt meine Mutter und mein Vater, die irgendwie Lust haben europäische Regulierungen zu feedbacken. Dafür fehlt meistens die Zeit und im Zweifel vielleicht auch die Kompetenzen. Und auch zivilgesellschaftliche Organisationen, die die Kompetenzen haben, haben leider oftmals nicht genug Zeit, sich im Detail mit allen Verordnungen und mit allen Direktiven zu beschäftigen. Weshalb diese Konsultationsprozesse, was Regulierung angeht, oftmals sehr unternehmerisch im Fazit ausfallen. Das ist schade, weil Zivilgesellschaft sich oftmals für die Belange von betroffenen Gruppen, marginalisierten Gruppen einsetzt und insofern wünsche ich mir da so ein Stück weit einen Ausgleich, Mechanismus, der da auch Sorge trägt, dass eben auch zivilgesellschaftliche Stimmen ausreichend gehört werden. Aktuell finden KI-Systeme in so vielen Lebensbereichen statt und ich merke immer wieder, dass die meisten Menschen sich dessen gar nicht bewusst sind. Meine neueste Zahnbürste kommt laut Hersteller mit KI daher, die mir dann eben also erlernter maßen sagt, wann ich den Zahnbürstenkopf zu wechseln habe. Dann mache ich die Social Media Plattform auf, um mich zu informieren. Auch die basieren auf *recommender systems* und empfehlen mir personalisiert, welche Artikel, welche Tweets ich zu lesen habe. Dann gebe ich meistens ein, wie der schnellste Weg zur Arbeit ist und nutze Google Maps, also Navigationssysteme. Auch die basierend auf Algorithmischen Entscheidungssystemen und errechnen mir die schnellste Route. Da könnte man beispielsweise auch Algorithmen einsetzen, die einem die schönste Route ausrechnen oder eine, die mit besonders viel Sonnenstrahlen daherkommt, aber das zeigt einfach nochmal, auch da sind Algorithmische Entscheidungssysteme da, um unser Leben im besten Falle zu verbessern. Und deshalb ist es eigentlich so wichtig, dass diese Anwendungen nicht an den Anwender:innen vorbei entwickelt werden. Und in der Regel investieren Unternehmen auch in User Experience Tests und so weiter. Was dabei aber häufig passiert ist, dass diese Unternehmen oftmals mysterio-typische Vorstellungen von Menschen haben. Und wenn ich jetzt gerade an den Global North denke, also an Europa und Amerika, dann sind es oftmals so Vorstellungen von einem weißen cis Mann, ab 40+, der dann die Anwendung nutzt und also als junge Frau mit Migrationshintergrund ist das jetzt nicht der beste Ankerpunkt für mein Nutzungsverhalten und das wiederum verpassen einige Unternehmen leider Gottes und deshalb sind diese Anwendungen teilweise an den Nutzer:innen-Interesse vorbei.

Schön: Seit dem 1. Januar 2023 dürfen selbstfahrende Autos bis zu 130 Kilometer pro Stunde auf deutschen Autobahnen fahren. Das Gesetz ist schon da, bisher ist aber noch kein einziges Auto dafür zugelassen. Aber eine dringende Frage gibt es hier schon noch, und zwar angenommen, es passiert ein Unfall, wer übernimmt dann die Verantwortung?

Fetic: Und das ist eine Frage, die sehr spannend ist und wahrscheinlich auch sehr häufig gestellt wird, weil das autonome Fahren so schön greifbar macht, was eben maschinelles Lernen ermöglicht und trotzdem damit kommt im Grunde genommen ein her, dass wir oftmals dem Hype rund um KI vielleicht ein Stück weit zu sehr verfallen, weil gerade beim autonomen Fahren, also darüber wird zwar häufig gesprochen, aber wir sind gar nicht so weit. Und deswegen, das ist auch so ein interessanter Aspekt in dieser Diskussion, dass wir oftmals den Algorithmischen Entscheidungssystemen mehr Magie unterstellen, als es eigentlich ist am Ende.
Es ist Statistik und wir unterstellen oftmals uns selber, dass wir da viel weiter sind als wir im technischen Sinne eigentlich an der Stelle wären. Beim autonomen Fahren ist es so, dass man deshalb fünf Stufen definiert hat, die beschreiben sollen, wo wir im technologischen Fortschritt sind, und die fünfte Stufe stellt sozusagen voll autonome Fahrzeuge dar, und da sind wir einfach noch nicht. Ob wir da jemals sein werden vor dem Hintergrund der schlechten Dateninfrastruktur, der Sensorik und so weiter, das würde ich auch noch mal kurz in Frage stellen. Also zumindest mit Blick auf jetzt die absehbare Zeit. Und ich würde sagen, wir sind aktuell bei Stufe 3, sagen einige Hersteller. Hoch automatisiertes Fahren ist das. Das bedeutet also, dass unter bestimmten Bedingungen eine Art Automatisierung stattfinden kann. Und da kommt jetzt die Frage ins Spiel, die du gestellt hast, was passiert denn, wenn ebenso ein hoch automatisiertes System einen Fehler macht? Würde ich sagen, dass die aktuelle Rechtslage noch nicht ganz eindeutig ist. Wir haben aktuell die Situation für bestimmte Dinge, die Fahrzeughersteller:innen haften, für bestimmte Aspekte, die Fahrzeughalter oder eben auch diejenigen, die das Fahrzeug nutzen. Und das würde auch in diesem Fall zur Anwendung finden. Das bedeutet also, sobald ein System einen Fehler macht, würde ein Stück weit überprüft werden müssen, ob die Hersteller:innen alle Sicherheitsvorstellungen umgesetzt haben, die sie aktuell umzusetzen haben und dann würde man sich eben anschauen, was der Fahrzeughalter oder die Halterin und die Nutzenden im Grunde genommen unternommen haben, um den Unfall zu verhindern, oder was entsprechend die kausalen Zusammenhänge waren des Unfalls. Das heißt am Ende auch hier nicht so richtig viel Neues. Am Ende haften Menschen und diejenigen, die die Systeme nutzen, sind nicht völlig von ihrer Verantwortung befreit, das heißt, sie müssen sich entsprechend auch mit den Grenzen und Möglichkeiten eines technischen Systems vertraut machen. Das heißt nicht, dass alle Fahrzeughalter:innen jetzt verstehen müssen, wie der Motor funktioniert und auch nicht wie das KI-System funktioniert, aber sie müssen wissen, wie sie diese Systeme steuern, wann sie eingreifen müssen und wann es eben auch bestimmte Grenzen gibt, die man nicht überschreiten sollte. Und insofern, also die Antwort ist, am Ende, haftet der Mensch. Ein Stück weit werden es wahrscheinlich die Hersteller:innen tun müssen. Das wird aber eben auch, gerade weil das Thema so neu ist, auf europäischer Ebene diskutiert.

Schön: Werfen wir doch mal gemeinsam einen Blick in die Zukunft. Wie stellst du dir den Einsatz von KI für und mit den Menschen, für die sie gemacht wird, im Idealfall vor? Was wünschst du dir?

Fetic: Im Idealfall hätte ich wahrscheinlich meinen Job gar nicht mehr, weil dann müsste ich nämlich nicht mehr über die gesellschaftlichen Implikationen von KI oder Algorithmischen Entscheidungssystemen sprechen, weil wir alle verstanden haben, dass das einfach nur technologische Werkzeuge sind und technologische Werkzeuge eben ganz viel Potenzial auch

haben, uns unser Leben einfacher zu machen, schöner zu gestalten, aber eben auch nicht ganz umsonst sind. Das heißt also, die Entwicklung kostet Geld, kostet Kraft und Energie so zu voranzutreiben, dass sie unseren Wünschen und unseren Bedürfnissen am besten und am ehesten entspricht. Dann braucht es Menschen wie mich nicht mehr, die darauf hinweisen, dass eben diese technologischen Mittel nur so gut sind wie die Menschen, die sie entwickeln und einsetzen, das wäre mein Traum.
Und wenn ich dann noch zusätzlich was zu träumen dürfte, dann nicht mehr, das wir so grundsätzlich auf technologische Hypes aufspringen. ChatGPT ist ein großes Stichwort gerade. Selbst meine Mutter spricht mich darauf an, die sonst mit Technik nicht so viel am Hut hat. Und da wünschte ich mir so ein bisschen eine aufgeklärtere Haltung zu. Also dass wir nicht jedes Mal ganz doll erschrocken sind, dass da etwas technologisch Neues kommt. Oftmals sind es richtig tolle lebenshelfende Systeme, die man einfach nur richtig gut zu nutzen wissen muss. Und das ist mein Wunsch, dass wir die KI-Systeme als intelligentere Hammer nutzen und wenn wir uns auf die Finger hauen, nicht den Hammer anschreien, sondern im Zweifel besser darüber nachdenken, wie wir den Hammer künftig nutzen wollen.

Schön: Vielen Dank für deinen Einblick und das Gespräch, liebe Lajla.
Jetzt haben wir vieles über Künstliche Intelligenz, Algorithmen und Maschinelles Lernen gehört. Es macht total viel Spaß damit rumzuspielen, aber KI kann eben auch gefährlich werden, wenn wir uns komplett auf sie verlassen. Was ist wahr und was ist echt? Das ist gar nicht mehr so leicht zu beantworten. Zum Beispiel, wenn durch den Einsatz von KI unechte Fotos mit echten Menschen generiert werden. Ihr habt sicher das gefakte Bild von Papst Franziskus mit der dicken Balenciaga Daunenjacke auf dem Schirm oder andere Deep Fakes wie das Video von Bundeskanzler Olaf Scholz dem Programmierer:innen Worte aus einem russischen Actionfilm in den Mund gelegt haben. Bei Algorithmischen Entscheidungssystemen geht es eben immer auch um Verantwortung, also dass wir ganz bewusst mit den Bildern und Infos, die uns Algorithmen im Alltag vorschlagen, umgehen und auseinandersetzen, aber auch mitbestimmen, wie KI als Werkzeug im Alltag zukünftig verwendet wird.

[Einsatz Upbeat Podcast-Outro im Hintergrund]

<u>Abmoderation & Credits</u>

Schön: Was sagt ihr denn zum Einsatz von KI? Gerade nach dieser Folge. Habt ihr mit Hilfe von Künstlichen Intelligenzen schon mal eine Entscheidung getroffen oder anders herumgespielt? Schreibt uns bei Insta, Facebook, per Mail oder lasst euer Feedback einfach hier bei Spotify. Ich freue mich auf jeden Fall sehr über eure Kommentare.
Nächstes Mal gehts bei „Justice, Baby!“ dann um Kohle, Cash und Moneten. Ja, wir sprechen über Geld und wie gerecht unser Steuerrecht eigentlich ist. Steuerrecht but make it cool, mal sehen, ob wir das nächstes Mal hinbekommen. Ich freue mich jedenfalls, wenn ihr dann wieder rein hört. Mein Name ist Kathrin Schön. Ich sage vielen Dank auch an meine Kolleginnen im Podcast Team und bis bald.

„Justice, Baby!“ ist ein Podcast der Stiftung Forum Recht.
Redaktion: Andrea Wojtkowiak, Vanessa Mittmann und ich, Kathrin Schön.
Juristische Beratung: Karolina Hanisch.
Produktion: Axel Seyboth und Anna Kunzmann von L‘agence.

Kommunikation und Distribution: Silke Janßen, Franziska Walter, Romy Klemm, Sabine Faller und Hannah Schelly.

[Outro blendet aus]

Transkript: Felicia Stahnke