

Diese Ansage kennt ihr bestimmt. Ihr ruft bei einer Hotline an und soll aufgezeichnet werden.

So what? Warum nicht? Ja, weil es vielleicht um mehr geht, als ihr denkt.

Denn eure Stimme gibt viel mehr Preis, als ihr vielleicht glaubt. Besonders, wenn eine künstliche Intelligenz sie auswertet und dann vielleicht sogar völlig falsche Annahme über euch trifft.

Ihr hört 11km der Tagesschau-Podcast, ein Thema in aller Tiefe. Mein Name ist Victoria Kobmann. Heute ist Donnerstag, der 12. Oktober. Mein Gast heute ist Rebecca Ciesielski vom BR AI und Automation Lab.

Hallo Rebecca. Hallo. Du warst quasi im Silicon Valley von München unterwegs, nämlich bei einem Softwareunternehmen. Ja, Silicon Valley, das ist eine ganz interessante Analogie, weil das ist in Gilching und andere kennen das vielleicht auch diese Region unter Oberfaffenhofen. Da sind tatsächlich relativ viele unterschiedliche Unternehmen, die alle was mit Hightech machen und eben auch dieses Unternehmen, was wir uns angeguckt haben, Odearing. Ich muss sagen, ich war noch

nicht im Silicon Valley, aber auf jeden Fall nicht so, wie man sich das vorstellt, sondern das war ein sehr unscheinbare Bürokomplex und da sind wir so in die erste Etage gefahren und da war das so eine etage Büros und relativ unscheinbar außer ein großes rotes Sofa-Stand im Empfangsbereich.

Und dann sind wir von Bion Schuller, dem wissenschaftlichen Leiter, dort in einen Konferenzraum geführt worden, der schon so abgedunkelt war und dann hat er seinen Laptop aufgeklappt. Abgedunkelt, okay, was ist da passiert? Ja, vielleicht war das für die mysteriöse Stimmung, aber wahrscheinlich war es einfach nur, damit wir die Software auf dem Laptop gut erkennen können und er hat uns dann halt gezeigt, wie Software zum Beispiel funktioniert, die Emotionen

erkennt. Ich finde es übrigens immer sehr spannend, wie viel mehr Angst die Leute vor Kameras haben und

entspannt sind mit Mikrofonen. Was wir jetzt sehen, ist ein sogenanntes Arousal Valence Modell, das heißt, das zeigt wie negativ oder positiv ich klinge und wie aktiviert oder erregt ich bin.

Dann hat sich ein Mikrofon an seinen Hemd gesteckt und hat dann diese Software geöffnet und da musste er erstmal einen Text vorlesen, den hat er relativ neutral vorgelesen, damit sich die Software sozusagen an seine Stimme gewöhnt und dann hat er uns eben das gezeigt und das Klang so ein bisschen geschauspielert und dann waren wir natürlich neugierig.

Und was macht dann die Software damit, also die KI? Das ist wie so ein Vier-Eck auf dem Bildschirm und in

allen Ecken, also in den Vier-Ecken ist jeweils ein Smiley und einer davon schaut wütend und einer schaut fröhlich und einer schaut, naja, so ein bisschen neutral zufrieden und einer schaut traurig und da ist so ein roter Punkt auf diesem Koordinatensystem und der springt dann immer mit so einer kleinen Verzögerung in eine Richtung, also zum Beispiel zu dem wütenden Smiley oder zu dem zufriedenen und so weiter und ich war neugierig und wollte das selbst ausprobieren und habe mir eben dann auch dieses Mikro genommen und habe ohne die Software vorher groß an meine Stimme zu

gewöhnen und zu trainieren einfach was gesagt auf Deutsch und habe dann auch gesehen, dass die Software das einigermaßen in einer Emotion kategorisieren kann. Jetzt rede ich sehr motiviert und auch

etwas fröhlich, amused, okay. Und jetzt rede ich relativ neutral und versuche nur etwas zu erklären, was ich versuche zu beschreiben, also zum Beispiel, dass ich da jetzt gerade ein Punkt sehe auf dem

Bildschirm, der hin und her springt. Okay, da steht dann jetzt aber worried, also das war jetzt relativ negativ.

Ach, das geht aber viel schneller als ich dachte, muss ich sagen. Ich dachte irgendwie, dass man so ein Text einspricht und dann erst hinterher irgendwie kommt, naja, die Person hat sich vielleicht so oder so gefühlt, aber das heißt, es wird quasi in Echtzeit angepasst. Ja, also so fast in Echtzeit und was ich da gerade erzählt habe mit worried und amused und diese ganz konkreten Emotionen,

das wird dann immer noch dazu angezeigt und das war aber, ja, das hat eher nicht so richtig gepasst oder hatte ich zumindest das Gefühl, dass es nicht so richtig passt, weil das waren dann schon sehr große Emotionen dafür, dass ich eigentlich da nur, ja, relativ neutral gesprochen habe, zumindest aus meiner Wahrnehmung. Wir wollten natürlich dann auch wissen, wie dieses Software funktioniert und da hat Björn Schuller gesagt, die misst ganz viele verschiedene Faktoren der Stimme. Intonation, Intensität, Rhythmus, dann habe ich den Klang der Stimme, also klingt die Eheise, geschrien sozusagen, geflüstert, das ist so Stimmqualitative Merkmal, dann nennen wir das und dann gibt es artikulatorische Merkmale, das heißt, wie sauber spreche ich, sozusagen. Und all das kann die KI erfassen und verarbeiten und daraus bestimmte Muster ableiten. Also dieses Programm, das mit einer künstlichen Intelligenz funktioniert, hat eigentlich das Ziel, nur aus dem, wie jemand spricht, aus der Sprache quasi, rauszuhören, welche Emotionen der oder die hat. Okay, also was würdest du jetzt sagen, wie feinfühlig, wie genau ist das Ganze? Ja, also was auffallend war, war, dass die Emotionen so ungefähr gepasst haben, also genau mit diesen

vier Richtungen, das hat schon so ungefähr gepasst, aber eben auch mit dem Schauspielern. Also wir

haben ja, oder ich habe in dem Fall ja auch versucht, meine Stimme dann so ein bisschen zu verstellen und zu sagen, jetzt spreche ich irgendwie fröhlich oder jetzt spreche ich irgendwie neutral und das war ja nicht authentisch. Also ich war in der Situation ja nicht fröhlich gerade, weil irgendwas passiert ist, sondern ich habe das ja bloß mal versucht durchzuspielen und das hat es einigermassen, wie gesagt, abgebildet. Das heißt, die Software hat noch nicht erkannt, dass du die Emotionen nur gespielt hast. Björn Schuller, der Vorsteller schon seit 20 Jahren oder über 20 Jahren in diesem Bereich, hat uns dann auch gesagt, dass die Emotionen nur so gut abgebildet werden

können, wie Menschen sie auch erkennen. Das heißt, die KI in der Zukunft wird einfach immer mehr ähnlich wie ein Mensch hören können, oh, der ist gerade total durch den Wind oder ist kognitiv gerade woanders, was auch immer es ist. Und Menschen erkennen das ja auch nicht immer richtig, aber das ist sozusagen die bestmögliche Weise, wie so ein Algorithmus die Emotionen erkennen kann, aber es geht nicht besser als Menschen. Er hat verschiedene Zahlen auch ins Feld geführt und da ist oft zum Beispiel von 80 Prozent Richtigkeit die Rede und das ist eben dann auch manchmal falsch, aber Menschen liegen halt auch manchmal falsch. Wo kommt diese Software denn

zum Einsatz? Die wird ja wahrscheinlich nicht einfach nur so produziert, ohne dass man die benutzen will. Die Software ist tatsächlich im Einsatz und zwar sogar auch in Deutschland und ein Einsatzfeld sind Core Center, was natürlich relativ nahe liegt, weil da ja ganz viel mit Sprache gearbeitet wird und Sprache das Medium ist, was da die ganze Zeit ausgetauscht wird und dementsprechend

auch analysiert werden kann. Was für Core Center sind das? Wir wissen nicht, welche Core Center das alle einsetzen, aber wir haben einen Core Center gehabt, was uns auch Einblicke gegeben hat und das ist das Core Center 1188.0. 1188.0, wenn ich das höre, da denke ich doch an diese Werbung aus den 90ern. 1188.0, die Auskunft für Deutschland, Telegate. 1188.0, da werden Sie geholfen. Das ist meine Assoziation mit 1188.0, diese Werbung natürlich mit Verona Pot. Ich wusste gar nicht, dass es die Hotline überhaupt noch gibt heute. Ja, das war überraschend, also die hatten ihre große Hochphase, das haben sie uns auch erzählt, eben in den 90ern mit dieser Auskunft, die sie waren. Wie ich mir die Auskunft merke, ganz einfach. Mit 11 wurde ich eingeschult. Für jede Sendung brauche ich mindestens 88 Garten. Und dann kam natürlich das Internet und Telefonauskünfte wurden weniger wichtig und das Businessmodell von 1188.0 hat sich auch verschoben hinzu einem Core Center Betreiber, also quasi Core Center für andere Unternehmen. Das machen die jetzt hauptsächlich, aber sie sind auch immer noch ein bisschen Auskunft, das haben sie uns auch gesagt. Wenn die einen Core Center für andere sind, dann ist das so, wie wenn ich jetzt als Kunde, als Kundin, wo anrufe und irgendeine Frage habe, dann rufe ich vermeintlich bei irgendeiner Firma an. Ich lande aber bei der 1188.0, oder wie? Genau, das haben wir in unserer Recherche auch an sehr vielen Stellen gehört, dass es immer mehr Unternehmen gibt, die ihre Core Center Outsourcen an andere Unternehmen. Das würde man aber als anrufende Person nicht mitbekommen. Also man würde bei einem Unternehmen XY anrufen und dann würde die Person am anderen Ende sagen, dass man dort gelandet ist bei diesem Unternehmen und würde jetzt nicht sagen, ich bin eigentlich bei einem Core Center Betreiber angestellt und gar nicht bei dem Unternehmen selbst. Wo landet man da? Wo sitzen die? Ich suche sie jetzt raus, ne? Ein Moment, wir haben uns dort das Core Center angeschaut, also das Unternehmen selbst, das ist in Essen und die haben aber ihre Core Center in unterschiedlichen deutschen Städten, eben zum Beispiel in Rostock und dort sitzen relativ viele Personen in einem Großhausbüro und alle haben so Headset auf und telefonieren den ganzen Tag, also wie man sich so ein Core Center vorstellt. Und was hat die 1188.0 jetzt mit dieser StimmKI zu tun? Also es gab da auch Smilies und zwar auf dem Bildschirm der Agentinnen und Agenten 2, also einmal einen, der gezeigt hat, wie die Anrufenden klingen und einmal einen, der gezeigt hat, wie die Agentinnen und Agenten selbst klingen. Und die konnten ihre Farbe wechseln, also das war auch eine sehr grobe Einschätzung, also von Grün, was eher so freundlich fröhlich ist, über gelb, so neutrale Eindruck der Stimme, bisschen zu rot, aggressiv, fütend, aufgeregt und das ging sowohl bei den Smilies für die Anrufenden als auch für die Mitarbeitenden, also schon eine sehr grobe Einteilung, aber man konnte sich irgendwie ein Bild machen und das war so ganz interessant, dass wir, obwohl wir ja nicht gehört haben, was die Anrufenden gesagt haben, weil das als über Headset war, gesehen haben, wie diese Smilies zum Beispiel ihre Farbe verändert haben über das Gespräch und dann erst rot eingestiegen sind, mit wütend oder aufgeregt

wahrscheinlich, über irgendwie neutral, bisschen zu fröhlich zum Teil manchmal. Ja, man kennt's, oder? Man ruft an und ist erstmal irgendwie ein bisschen sauer, weil was nicht funktioniert und dann denkt man aber, okay, vielleicht lag's auch ein bisschen an mir. Ja, oder die Person am anderen Ende der Leitung sagt, was ändert, doch wieder beruhigt oder gibt's eine Hilfestellung und so weiter? Ja. Wie war denn dein Eindruck da vor Ort? Wie finden die Mitarbeiter das? Also die Mitarbeiter

mit denen wir gesprochen haben, die haben gesagt, das ist eine Unterstützung für sie, aber was sie auch gesagt haben und das war ganz interessant, dass sie versuchen die Anrufenden wieder auf Grün,

also auf freundlich zu switchen, war der Ausdruck und das hat bei mir den Eindruck verstärkt oder hervorgerufen, dass es da auch ein bisschen um Spiel geht. Also dass das wie so ein Gamification Ansatz ist, dass die Mitarbeitenden schon versuchen, bestimmt auch die Anliegen der Anrufenden zu lösen.

Na klar, aber auf der anderen Seite auch einfach versuchen, dass dieser Smiley irgendwie wieder Grün wird, wenn er vorher eben ärger ausdrückt und rot war und das war schon echt interessant, also dass das auch ein Spiel ist vielleicht. Also es gibt nur 0 und 1, es gibt nur sauer oder fröhlich quasi. Ja, wobei man da auch noch so eine, wie so eine Verlaufskurve darunter gesehen hat, wo man dann sehen konnte, okay, da hat sich quasi der Anteil der Fröhlichkeit gerade wieder gesteigert oder gesenkt und so weiter. Also es war schon so ein bisschen differenzierter als nur wütend oder glücklich. Wenige Meter entfernt von der Agentin saß die Teamleiterin Yvonne Eklof Martin und die hatte auch so eine ähnliche Ansicht. Also die hat auf ihrem Bildschirm vor sich auch Smilies gesehen, aber eben nicht nur die zwei, also nicht nur für ein Gespräch, sondern für alle Gespräche in ihrem Team. Also die Teamleiterin, die schaut jetzt aber nicht auf die Kunden, sondern auf die Mitarbeiter oder wie, wie die ihre Gespräche so führen. Da haben wir dann sowas gesehen wie 50 Prozent Freundlichkeit neben einem Smiley oder 27 Prozent Ärger neben einem anderen

Smiley und das war schon ja auch relativ eindrücklich, dass sie sogar so eine Prozentangabe daneben hatte und sie konnte auf diese einzelnen Gespräche in Echtzeit klicken und sich dann genau angucken, wie ist der Verlauf des Gesprächs und wie hat sich die Stimmung des Gesprächs entwickelt.

Wir sind natürlich bestrebt, gute Kundengespräche zu führen, freundliche Kundengespräche zu führen

und so können wir natürlich reagieren, sollte der Agent nicht so schöne Gespräche führen, dass wir dagegen steuern können. Also ich stelle mir das irgendwie schon auch wie eine Drucksituation

für denjenigen, der da in der Leitung sitzt, oder? Man möchte ja, dass der Kunde zufrieden ist. Das war gar nicht meine erste Assoziation. Meine erste Assoziation war, das wissen die Menschen, die anrufen mir gar nicht. Also die Menschen, die anrufen, die hören quasi in der Wandansage, die wir wahrscheinlich alle kennen. Zur Prüfung und Verbesserung unserer Servicequalität würden wir den Inhalt dieses Gesprächs mit ihrem Einverständnis gerne aufzeichnen. Wenn sie damit

einverstanden sind, drücken sie bitte die Taste 1. Also bei 1188.0 ist kein Hinweis darauf, dass auch die Emotionen ausgewertet werden und dann gibt es quasi diese zweite Ebene, wo jemand auf der anderen Seite des Telefons sitzt und die ganze Zeit auf einem Bildschirm sieht,

wie die Stimme der Anrufenden gerade klingt. Ja, stimmt, das ist ja eigentlich schon interessant. Also wenn ich gefragt werde, ob was aufgezeichnet wird, dann rechne ich damit, dass man sich das vielleicht hinterher nochmal anhört, aber ja nicht, dass meine Emotionen von einer künstlichen Intelligenz analysiert werden. Das ist ja schon noch mal was anderes, ne? Und das ist auch eine große Debatte, die darüber geführt wird, ob man Menschen darüber informieren sollte und in wiefern das gemacht werden sollte. Und das haben wir natürlich dann auch den Leiter des Kurs-Hinterbereichs

von 1188.0, hier in Hausmann gefragt. Also die Kunden werden nicht vorab informiert, sie werden darüber

informiert, wenn was aufgezeichnet wird, das tun wir auch, aber da hier nichts gespeichert wird und keine Kundenprofile daraus abgeleitet werden. Und ich sage mal, in dem System nicht hinterlegt wird,

war er ärgerlich oder freundlich, ist das nicht gut für mich. Das war so die erste Assoziation, die zweite Assoziation ist natürlich, ja okay, da werden die Menschen, die dort arbeiten, auch die ganze Zeit gemonitort. Ich finde, das hat irgendwie so ein bisschen Vibes von totaler Überwachung am Arbeitsplatz, oder? Ja, also sie selbst sagen darüber, dass es eine Hilfestellung ist und das ist natürlich auch für die Agentinnen und Agenten eine gute Hilfestellung sein kann, wenn eben die Teamleiterin direkt sieht, okay, dieses Gespräch ist gerade am Eskalieren, vielleicht geht sie dann mal hin. Ich sehe das Gespräch in Echtzeit, hier vor Ort habe ich die Möglichkeit, dann die entsprechenden Zumagenten zu gehen, ihnen aufzumunden, aufzubauen, zu fragen, was war im Gespräch, ob es eine bestimmte Situation gab, was diesen Ausschlag bei mir hier über Starsport ausgemacht hat. Und Jörn Hausmann, der Manager

von 1188.0, hat auch gesagt, das ist mit dem Betriebsrat abgesprochen und das ist auch mit einem externen Datenschützer abgesprochen, dieser Anwendung. Also das heißt, die haben sich da schon abgesichert. Nichtsdestotrotz bleibt natürlich der Eindruck und den hatten wir vor Ort auch, dass es eine permanente Monitoring ermöglicht. Ist das eigentlich erlaubt? Also eigentlich muss

man ja zustimmen, wenn man aufgezeichnet wird. Aber ist das denn legal, dann ohne Erlaubnis zu sagen, ich speise jetzt die Stimme in eine Software ein, sodass die in Echtzeit von einer KI analysiert wird und wir wissen, wie es dir geht. Also das finde ich irgendwie schon was anderes. Also wir haben mit mehreren Expertinnen und Experten in diesem Bereich geredet und der Grundkonsens ist

schon, dass man das im Einzelfall prüfen muss, weil es kein allgemeines, kein pauschales Gesetz gibt, was Emotionserkennung am Arbeitsplatz verbietet. Aber eigentlich waren sich trotzdem alle relativ einig, dass eine Bandansage, die das nicht konkret thematisiert und nicht sagt, hier werden ihre Emotionen ausgewertet, sind sie damit einverstanden, dass das rechtlich unzulässig ist.

Also dass die Personen, die dort anrufen, darüber aufgeklärt werden müssen, wenn ihre Stimme auf diese Art und Weise verarbeitet wird. Und der Rechtsexperte Peter Wette, der auch in dem Bereich schon seit sehr vielen Jahren aktiv ist und lange als Professor in dem Bereich gearbeitet hat, also mit beschäftigten Datenschutz sich sehr gut auskennt, der hat gesagt.

Emotionsanalyse in Callcentern, sowohl zu Lasten von Beschäftigten wie von Kunden, ist datenschutzrechtlich unzulässig. Wow, was sagt die 1188 Null dazu? Habt ihr denen das mal gesagt?

Ja, also die haben gesagt, dass sie davon ausgehen, dass sie auf der richtigen rechtlichen Seite stehen.

Hm, okay, das sieht der Jurist anscheinend anders, wenn das doch dann anscheinend zumindest drittig ist.

Warum passiert das denn dann trotzdem?

Also Peter Wetter hat uns zum Beispiel auch gesagt, dass er in den 20 Jahren, die er Callcenter schon beobachtet, rechtlich und auch zum Beispiel Gutachten erstellt, dass er immer wieder erlebt hat, dass Callcenter diese rechtlichen Grenzen austasten. Und das geht, haben wir auch in dieser Recherche miterlebt, weil eben diese Einzelfallprüfung so aufwendig ist.

Also das heißt, es muss erst mal im Einzelfall geprüft werden, es muss auch erst mal herausgefunden werden, was genau ist jetzt rechtlich unzulässig und was ist vielleicht noch okay oder eben in dem Graubereich,

der gerade noch akzeptabel ist und dann müssen daraus Konsequenzen gezogen werden.

Und passiert das?

Ja, also wir waren in unserer Recherche zum Beispiel in Mecklenburg-Vorpommern bei der Landesdatenschutzbehörde dort, also in jedem Bundesland gibt es eine Datenschutzbehörde und die sind dafür zuständig zu prüfen, ob es zum Beispiel an Arbeitsplätzen Datenschutzverstöße gibt.

Und diese Behörden, also jetzt nicht im speziellen Mecklenburg-Vorpommern, sondern generell haben uns gesagt,

dass sie im Jahr manchmal bis zu über 10.000 Beschwerden bekommen und sie müssen halt allen diesen Beschwerden nachgehen.

Und da kommt dann keiner mehr hinterher, verstehe ich.

In Deutschland gibt es ja eigentlich die Datenschutzgrundverordnung, die müsste sowas regeln.

Da steht diese künstliche Intelligenz allerdings, glaube ich, noch nicht so drin.

Aber auf EU-Ebene, da tut sich gerade was, nämlich der AI-Act, den hatten wir ja auch schon mal in einer anderen FKM-Folge zur KI,

der soll in Zukunft doch einiges dazu regeln.

Haben die auch auf dem Schirm, dass die KI unsere Stimmen analysieren kann?

Wenn die AI-Act-Verordnung in einer Weise verabschiedet werden sollte, wie sie gerade diskutiert wird

oder wie zumindest die Parlamentsposition gerade ist, dann könnte es sein, dass Emotionserkennung am Arbeitsplatz pauschal verboten wird.

Das ist alles noch sehr, sehr viel im Wagen, weil das gerade noch diskutiert wird.

Wohin führt das, glaubst du? Wohin könnte sich diese neue Technik entwickeln?

Also was wir bei Beyond Schulle, wir waren zweimal dort, was wir auch mitbekommen haben, ist, dass er nicht nur von deutschen Unternehmen angefragt wird oder seinen Unternehmen, sondern eher als Wissenschaftler und eben Hersteller oder zumindest Vordenkel von solchen Softwareanwendungen,

wird zum Beispiel anscheinend auch von Diktaturen, das hat er uns zum Beispiel erzählt, gefragt, ob er für die Lügendetektoren oder auch Software zur Homosexualitätserkennung bauen kann.

Homosexualitätserkennung, aber an der Stimme dann?

Das hat er uns gesagt, ja.

Das war übrigens früher etwas, was wir immer wieder angefragt wurden aus Regimen, wie soll ich sagen,

die man nicht unterstützen möchte, auch man zum Beispiel Homosexualität erkennen kann automatisch.

Und das war eines der Themen, die ich immer abgelehnt habe.

Und sind das Länder, in denen Homosexualität strafrechtlich relevant wäre?

Ja, also wir mussten ihm versprechen, dass wir nicht sagen, welche Länder das waren, aber genau, das wäre wirklich problematisch für die Menschen.

Also er sagt, er hat das abgelehnt, das zu entwickeln, aber es zeigt natürlich, wenn er es sogar schon bei ihm aufschlägt,

dass diese Gedanken, dass das vielleicht möglich ist anhand der Stimme, dass das in der Welt ist.

Und wenn man mal ein bisschen weiter noch wegschaut, zum Beispiel nach China,

dann sieht man auch, dass zum Beispiel im Bildungsbereich Emotionserkennung dort schon eingesetzt wird.

Also dass es Schulen gibt, die zum Beispiel Gesichtserkennung, also da geht es dann jetzt nicht um Stimme,

sondern um Emotionen, die aus Gesichtern herausgelesen werden sollen, in Schulen aufhängen und da die ganze Zeit die Kinder beobachten und gucken, ob die gerade konzentriert sind und was vermeintlich in deren Köpfen abgeht.

Und das ist ja schon, ja, das ist justopisch.

Also zum Glück wohnen wir jetzt nicht in einem Überwachungsstaat, aber wenn wir jetzt mal Deutschland angucken,

kann das ja auch irgendwie gefährlich sein für uns?

Ja, das ist die Frage. Also wir haben bei 1188 Nuiya relativ deutlich gesehen,

dass die Software dort am Rahmen geknüpft war, also eben, dass keine Profile gebildet werden,

dass die Mitarbeitenden jetzt nicht deshalb bestraft werden, wenn sie unglücklicher oder ärgerlicher und so weiter klingen.

Also es wurde uns zumindest so gesagt, sobald das aber eingeführt wird, also sobald man sagt, okay, diese Software misst wirklich Leistung und je nachdem, wie freundlich oder wie fröhlich jemand klingt,

wird er besser bewertet oder negativer bewertet, wenn die Person weniger freundlich klingt, dann könnte das natürlich sehr negative Auswirkungen auch haben, also am Arbeitsplatz zu einer großen Kontrolle,

zu einer großen Überwachung führen.

Wir haben über Ethik unter anderem mit Kate Crawford gesprochen.

Sie ist KI-Ethikerin aus Australien und beschäftigt sich schon sehr lange mit Emotionserkennung und dass man sich immer stärker selbst darauf hin konditioniert, so zu klingen, wie der Algorithmus das möchte.

This idea of discrete emotions, that we, you know, just feel this clear thing called joy or fear or anger,

I think has been very strongly disputed in the more recent scientific literature.

Und das haben wir ja zum Beispiel eben auch bei Björn Schulle gemerkt, der relativ stark

geschauspielert hat,  
damit der Algorithmus ihn richtig versteht.  
Das heißt also, wenn ich schauspielern muss, damit die KI meine Emotionen eindeutig zuordnen kann,  
dann ist das natürlich fehleranfällig.  
Schwierig, wenn dann daraus falsche Schlüsse gezogen werden.  
Also, was kann das denn bedeuten für Leute, die dort anrufen und dann falsch eingeschätzt werden?  
Also, was uns zum Beispiel die Call-Agents in dem Corsair-Literaturen,  
die die Call-Agentin in dem Corsair-Literaturen von 1188.0 auch gesagt hat, ist, dass sie schon wahrnimmt,  
dass Menschen, die, sie hat gesagt, nicht Deutsch klingen, dass sie auch schlechter bewertet werden von dem Algorithmus.  
Das hat ja eben keine Konsequenzen in dem Fall, aber sie sagt, sie blendet das dann aus, weil sie hört ja, dass die Person freundlich klingt.  
Aber das heißt trotzdem, dass der Algorithmus das nicht richtig abbildet.  
Und die KI-Ethik-Experte, Nikate Crawford, was ist ihr Fazit zu dieser Software?  
Wie schätzt die das ein?  
Also, sie spricht sich auch für ein Verbot aus.  
Sie schaut sehr gespannt auf den AI-Akt, weil sie findet, dass diese Software, dass Emotionserkennung sehr stark reglementiert werden, beziehungsweise auch verboten werden sollte.  
So, yet again, I think these are red flags to why we should be deeply concerned about the spread of these systems.  
Und Rebecca, wie beunruhigt bist du denn nach dieser Recherche jetzt?  
Ja, das ist schon sehr fragwürdig zum Teil.  
Und auch die Frage, ob man Lügendetektoren bauen kann, ob man Homosexualitätserkennung damit bauen kann,  
das ist alles sehr höchst problematisch.  
Und ich glaube, wir sollten generell über Emotionserkennung, aber über alle Daten und über alle Algorithmen,  
die über solche Körperdaten, also über solche biometrischen Daten, Entscheidungen über Menschen treffen,  
viel stärker reden und das viel mehr im Fokus der Öffentlichkeit haben.  
Rebecca, danke, dass du uns davon erzählt hast.  
Ciao, tschüss!  
Rebecca Cisielski vom BR AI und Automation Lab hat sich angeschaut, wo und wie künstliche Intelligenz  
heute schon überall eingesetzt wird, um unsere Stimme zu analysieren.  
Und was Unternehmen damit machen können, gemeinsam mit ihren Kollegen Sami Kammes und Simon Würz.  
Ihr findet ihr Radiofeature abgehört in der AID-Audiothek, wir verlinken euch das in unseren Show-Notes.  
Folgenautor ist Lukas Waschbisch.  
Mitgearbeitet hat Katharina Hübel, Produktion Ursula Kierstein, Jonas Lasse-Teichmann, Konrad

Winkler und Christine Dreyer.

Redaktionsleitung Lena Gürtler und Fumiko Lipp.

FKM ist eine Produktion von BR24 und NDR Info.

Mein Name ist Victoria Kopmann, wir hören uns morgen wieder.

Und hier noch ein Hör-Tipp. Wer noch mehr wissen will über künstliche Intelligenz, die Entwicklung geht rasend schnell, viele sprechen davon, dass die KI dabei ist, die Welt zu verändern.

Damit ihr da nicht den Überblick verliert, empfehle ich euch den KI-Podcast von der AID.

Wir verlinken ihn auch in unseren Show-Notes, genau wie weitere FKM-Folgen über die Möglichkeiten und Gefahren von KI.

Tschüss!