

Texttranskript zum Video 12. Veranstaltung Gleichstellung im Gespräch Ist die künstliche Intelligenz geschlechtsneutral?

Sehr geehrte Damen und Herren! Ich darf sie herzlich zur 12. Veranstaltung der Reihe Gleichstellung im Gespräch zum Thema „Ist die künstliche Intelligenz geschlechtsneutral“ begrüßen.

Frauenpolitik ist eine Querschnittsmaterie, das wird gerade bei Veranstaltungen wie der heutigen deutlich. Als Frauenministerin ist es für mich ganz besonders wichtig, veraltete Rollenbilder aufzubrechen, und dabei gilt es auch bei neuen Technologien wie der künstlichen Intelligenz anzusetzen, die schon heute Eingang in unser tägliches Leben gefunden hat und das in Zukunft sicherlich auch noch verstärkt tun wird. Es freut mich also sehr, dass dieser fachliche Austausch zu dem wichtigen Thema künstliche Intelligenz und Geschlechtergleichstellung möglich ist. Gerade, weil künstliche Intelligenz uns im Alltag immer wieder begegnet und das oft unbewusst, ist es wichtig, den Einfluss, den sie auf die Gesellschaft hat, kritisch zu hinterfragen. Viele von uns nutzen künstliche Intelligenzsysteme wie Alexa oder Siri beispielsweise in Form von digitaler Assistenz in unserem Smartphone oder auch in Form von Sprachsteuerung oder Navigationssystemen. Beim täglichen Surfen im Internet wird die uns angezeigte Werbung von künstlicher Intelligenz ausgewählt, basierend auf unseren Interessen und Aktivitäten. Diese auf uns abgestimmten Empfehlungen begegnen uns etwa auf Streamingdiensten und auf Social-Media-Plattformen. Es gibt bereits Beobachtungen darüber, dass gewisse Themenfelder eher den Interessen von Frauen und andere eher den Interessen von Männern zugeschrieben werden. Hier müssen wir wachsam sein und sicherstellen, dass sich bestehende Stereotype dadurch nicht noch mehr verfestigen. Künstliche Intelligenz beruht auf von Menschen gesammelten und verarbeiteten Daten. Da ist es wichtig, sich ganz genau mit dem Zusammenspiel zwischen KI-Systemen und etwaigen Stereotypen oder geschlechtsbezogenen Verzerrungen bei der Datensammlung, bei der Programmierung oder Datenaufbereitung auseinanderzusetzen. Es gilt aber auch das Potential zu sehen, das KI-Systeme mit sich bringen. Ein simples Beispiel ist die objektive Bewertung von Kandidatinnen und Kandidaten in Bewerbungsprozessen, die nicht zuletzt dazu beitragen kann, das bestmögliche Ergebnis sowohl für das Unternehmen als auch für die Gesellschaft zu erzielen. Im Zusammenhang mit künstlicher Intelligenz ist ein weiterer Aspekt besonders wichtig für mich: die beruflichen Möglichkeiten, neue Technologien mit sich bringen. Diese Berufe erfordern wesentliche Kompetenzen, eine Ausbildung in den Mintfächern, also Mathematik, Informatik, Naturwissenschaft und Technik, aber auch Kreativität und emotionale Intelligenz. Ich möchte Mädchen und Frauen ermutigen und darin bestärken, eine solche Ausbildung zu ergreifen, denn diese Karrieren haben hohes Zukunftspotential. Ich freue mich sehr, das wir heute mit Frau Mag. Maria Ulmer, Frau Professor Sabine Köszegi und Frau Dr. Martina Friedl 3 Expertinnen hören werden, die sicher ein bisschen mehr Klarheit zum Zusammenspiel zwischen KI-Systemen und der Gleichstellung der Geschlechter schaffen können. Außerdem hoffe ich, dass diese Beiträge auch einen Anreiz für viele Mädchen und Frauen sind, sich diesem Themengebiet noch intensiver zu widmen. Ich wünsche Ihnen allen eine interessante Veranstaltung, vielen Dank.

Herzlich willkommen auch von meiner Seite, mein Name ist Theresa Veit und ich darf Sie heute durch die 12. Veranstaltung „Gleichstellung im Gespräch“ zum Thema „Ist die künstliche Intelligenz geschlechtsneutral“ führen. Es warten sehr spannende Beiträge von Expertinnen auf uns. Wir starten mit Maria Ulmer, Leiterin der Sektion Digitalisierung und E-Government im Bundesministerium für Digitalisierung und Wirtschaftsstandort. In die Verantwortung der von ihr geleitenden Sektion fallen die Digitalisierung des E-Governments der Republik Österreich. Inhaltlich sorgt die Sektion für die Entwicklung und Koordination der Digitalisierungsstrategie, der Artificial Intelligence-Strategie, der Datenstrategie und für Angelegenheiten der Digitalisierungsagentur DIA. Sie wird in ihrem Beitrag über die neuen Entwicklungen der Artificial Intelligence-Strategie in Österreich berichten. Liebe Frau Ulmer, Sie haben das Wort.

Ein herzliches Willkommen, sehr geehrte Besucherinnen und Besucher zur Veranstaltungsreihe „Gleichstellung im Gespräch“ mit dem Zukunftsthema „Ist die künstliche Intelligenz geschlechtsneutral?“. Eigentlich ist künstliche Intelligenz kein Zukunftsthema, sondern schon Realität,

so wie wir uns heute der aktuellen Situation geschuldet virtuell treffen. Mein Name ist Maria Ulmer, und ich darf den Bereich Digitalisierung und E-Government im Bundesministerium für Digitalisierung und Wirtschaftsstandort verantworten. Bevor ich mit meinen Gedanken zum Thema künstliche Intelligenz beginne, gestatten Sie mir noch der Interministeriellen Arbeitsgruppe Gender Mainstreaming/Budgeting in Österreich zu ihrem 20-jährigen Bestehen herzlich zu gratulieren und weiterhin viel Erfolg zu wünschen. Die Bedeutung von Gleichstellung für alle Gesellschafts- und Lebensbereichen ist gerade in diesen fordernden Zeiten um die Facette der digitalen Transformation reicher geworden. Es ist dabei wichtig, wesentliche Entwicklungen und Herausforderungen sichtbar zu machen, anzusprechen und darauf auch mit entsprechenden Konzepten auch zu reagieren. Dies bringt mich zum Input meines heutigen Beitrags, ob die künstliche Intelligenz nicht auf den Bezug auf Gendermainstreaming einen blinden Fleck hat. Das ist eine wichtige und richtige Frage. Ich bin der Ansicht, dass neben den bestehenden, definierten Handlungsfeldern der interministeriellen Arbeitsgruppe die Digitalisierung bzw. die digitale Transformation als neues Handlungsfeld mit aufgenommen werden sollte. Wie vieles andere auch wird Technologie und damit auch künstliche Intelligenz von Menschen entwickelt. Technologie und Forschung sind nach wie vor Bereiche in den Frauen stark unterrepräsentiert sind. Das führt dazu, dass Technologie nach männlichen Vorstellungen entwickelt wird. Das kann bewusst oder unbewusst passieren, führt aber dazu, dass Computer, Software und sogenannte intelligente Systeme nicht geschlechtsneutral agieren. Lassen sie mich das anhand eines bekannten Beispiels erklären, den digitalen Sprachassistentinnen Siri, Alexa und Cortana, die in ihrer Rolle als Gehilfinnen standardmäßig eine weibliche Stimme eingestellt haben, während der intelligente Supercomputer Watson männlich definiert ist. Hier haben mehrere Studien gezeigt, dass durch die Feminisierung von Sprachassistenzen stereotype und antiquierte Rollenbilder reproduziert werden. Digitale Assistenten mit Frauenstimmen antworten auf Beschimpfungen und sexuell übergriffigen Aussagen oder Belästigungen nicht so, wie man es von realen Frauen erwarten würde. 2 Beispiele möchte ich hier anführen zum Statement „You are hot“ bedanken sich die Sprachassistentinnen für's Feedback. Und zum Statement „You are a naughty girl“ entschuldigen sie sich. Daher ist es umso wichtiger zu beobachten, wie sogenannte intelligente Anwendungen agieren bzw. auch wahrgenommen werden. Ein KI-System kann nur so gut sein, wie ihre Entwickler es mit Daten gefüttert und trainiert haben, und beide sind nach wie vor männlich und weiß geprägt. Eine überwiegende Anzahl von Texten, die man zum Trainieren von KI-Algorithmen verwendet, sind immer noch daher traditionell und männlich. Als Beispiel dafür, das hat eine umfangreiche Studie in den USA ergeben, gilt die KI in den USA mitunter als rassistisch. Was darauf zurückzuführen ist, dass sie auch mit Daten der Strafverfolgungsbehörden gefüttert ist. Das führt einerseits dazu, dass vor allem Deep-Learning-Algorithmen, die eine große Menge an Daten verarbeiten, geschlechtsspezifische Verzerrungen verursachen. Dadurch liefern auch Suchmaschinen, die solche Algorithmen verwenden geschlechtsspezifische Verzerrungen, die durch weitere Suchanfragen noch verstärkt werden. Da diese Modelle aus den neuen Anfragen lernen und somit eben diesen Bias verstärken. Ein weiterer interessanter Aspekt ist der Themenbereich Testpersonen. Systeme zur Sprach- und Gesichts- und Bildererkennung, wie sie teilweise auch in modernen Smartphones eingesetzt sind, funktionieren bei Frauen bzw. bei dunkelhäutigen Frauen anfangs schlechter. Da sie weit seltener als Testpersonen ausgewählt wurden als hellhäutige Männer. Wie kann man diesen Entwicklungen gegensteuern, und was braucht es für eine diskriminierungsfreie und geschlechtsneutrale künstliche Intelligenz? Sowohl auf europäischer als auch auf internationaler Ebene werden zur Zeit Empfehlungen zur Entwicklung, zum Einsatz und zur Verwendung von Artificial Intelligence ausgearbeitet, bei der auch Gendergleichstellung bzw. Genderneutralität betont werden. Österreich unterstützt diese Empfehlungen und bringt sich selber aktiv ein, zum Beispiel in Ratsarbeitsgruppen auf EU-Ebene. Auch global befindet sich derzeit eine Empfehlung in Ausarbeitung, das im Rahmen der UNESCO abgewickelt wird. Das ist die Empfehlung zur Ethik, der künstlichen Intelligenz, die das erste weltweite, normative Rechtsinstrument für die Steuerung der ethischen Dimension des Bereichs künstlicher Intelligenz sein soll. Zusätzlich arbeitet im ISO das Subcommittee 42 an Artificial Intelligence, das ist die Working Group 3 mit dem Themenschwerpunkt Trustworthiness. Wir haben hier die große Chance, unsere Wünsche, Prinzipien und Interessen als Mitgliedsland in Europa zu verankern. Noch dazu sind gerade die Themenbereiche Fairness und Transparenz auch zutiefst europäische Interessen. Ich freue mich sehr,

dass es uns unter anderem gelungen ist, das plenary meeting 2023 eben von diesem subcommittee 42 zur künstlichen Intelligenz nach Wien zu holen. Das heißt für uns, dass damit weltweit die wichtigsten Expertinnen und Experten, die sich mit künstlicher Intelligenz beschäftigen, in Wien sein werden. Die Entwicklung von vertrauenswürdiger KI muss gefördert werden, um eben diese ethischen Richtlinien zu berücksichtigen, damit KI-Systeme neutral und ohne jede Voreingenommenheit, insbesondere in Bezug auf das Geschlecht und alle anderen Diskriminierungsgründe agieren können. Technische Standards müssen daher festgelegt werden, um klare Kriterien zu definieren, damit die Wahrung der Menschenwürde und die Einhaltung ethischer Prinzipien bei der Programmierung und Verwendung von AI-Verfahren künftig garantiert wird. Die Nachvollziehbarkeit und Transparenz von AI-Verfahren muss gefördert werden, damit automatische Entscheidungen logisch begründet und verifizierbar sind. Damit ist der Einsatz von quelloffenen Systemen und standardisierter Programmiersprache das Ziel. ISO-Standards haben daher einen sehr hohen Stellenwert und haben de facto normative Kraft und gelten auch vor Gericht und sind state of the art. Für Sachverständige sind sie damit auch die Basis für Gutachten und schon jetzt in der Wirtschaft und Industrie für Ausschreibungen nicht mehr wegzudenken. Ich darf Sie an dieser Stelle herzlich zur Mitarbeit einladen. Alle Expertinnen und Experten sind aufgerufen, ihren Beitrag zu leisten, um diese österreichischen und europäischen Interessen bezüglich Fairness und Transparenz im internationalen Standard zu verankern, und dieses ISO-Subcommittee 42 ist genau das passende Gremium dafür. Die Mitarbeiter in diesem Standardisierungsgremium folgen in Österreich bei Delegation via ASI und das BMDW arbeitet schon jetzt bereits an diesen Standards aktiv mit. Wir haben auch zum Thema Artificial Intelligence ein konsolidiertes Papier mit dem Titel „Artificial Intelligence Mission Austria 2030“ entwickelt. Dazu wird die österreichische Politik in verschiedene Zukunftsfelder investieren. Hierzu zählen zum Beispiel diese Bereiche: Forschung und Innovation, Qualifizierung und Ausbildung. Auch die KI im öffentlichen Sektor ist ein Thema, die KI in der Wirtschaft, Infrastruktur, für INDASTRE-Leadership, Gesellschaft, Ethik und Arbeitsmarkt und letztendlich auch KI-Governance, Sicherheit und Recht. Ein weiterer enorm wichtiger Aspekt ist auch, den Frauenanteil in Forschung und Technologie zu erhöhen und damit eine ausgewogene Geschlechterverteilung in der AI-Forschung, in der Wissenschaft, aber auch in den Toppositionen, im Management, Gremien und Forschungsteams von Digitalunternehmen herzustellen, die Förderung von Frauen als Unternehmerinnen sowie die Beteiligung und das Engagement von Frauen in allen Phasen des Lebenszyklus. Das AI-System zu fördern, in dem unter anderem wirtschaftliche und rechtliche Anreize angeboten werden, das muss ein weiteres Ziel sein. Viele von Ihnen waren und sind in dieser Phase des 2. Lockdowns als Eltern neben ihrer herkömmlichen Arbeit, im Bereich des Home Schoolings stark gefordert. Sie erleben quasi derzeit eine digitale Doppelbelastung, Mütter und Väter müssen hier viel Kreativität und Durchhaltevermögen an den Tag legen, um Familie und Beruf unter einen Hut zu bringen. Ich habe auch 2 schulpflichtige Kinder und bin hier definitiv nicht auf der Meta-Ebene, sondern mittendrin. Die derzeitige Situation und unsere digitale Arbeitswelt erfordern mehr Teamwork denn je. Umso wichtiger ist ein gutes Verhältnis unter den Mitarbeiterinnen und den Mitarbeitern und eine Unterstützung seitens der Führungskräfte. Immerhin geht es darum, diesen Neustart gut zu bewältigen. Und das funktioniert nur mit verlässlichen Abläufen, die einen sicheren Rahmen schaffen und damit auch entspannte und leistungsfähige Mitarbeiterinnen und Mitarbeiter. Ich bedanke mich herzlich für Ihre Aufmerksamkeit. Es wird wichtig sein, die Gleichbehandlungsthemen mit allen ihren Facetten aus den unterschiedlichsten Lebensbereichen aufzuspüren, anzusprechen und Lösungen dafür zu finden. Dafür und für Ihre weitere Arbeit wünsche ich Ihnen viel Erfolg.

Vielen Dank für diese spannenden Einblicke in die derzeitigen Entwicklungen. Als nächstes lässt uns Sabine Köszegi an ihrer Expertise teilhaben. Sie ist Professorin für Arbeitswissenschaft und Organisation am Institut für Managementwissenschaften der TU Wien. Ihre Forschungsschwerpunkte liegen in der Schnittmenge von Technologie, Arbeit und Organisation. Sie forschen nur in Arbeitsformen, im Kontext von Informations- und Kommunikationstechnologien sowie zu Fragestellungen der soziokulturellen und gesellschaftlichen Implikationen sozialer Robotik. Sie ist akademische Direktorin des MBA-Programms in Entrepreneurship und Innovation, sowie Koordinatorin des Interdisziplinären Doktoratskolleg „Trust Robots“ an der TU Wien. Sie ist außerdem

Mitglied der High-Level Expert Group on Artificial Intelligence der Europäischen Kommission. Frau Köszegis Beitrag lautet „I'd blush if I could...“ eine Gender-Perspektive auf KI-Technologien. Wir sind gespannt.

Hallo, ich bin Sophia. Ich werde im Rahmen des Projektes „Loving AI“ programmiert, um die bedingungslose Liebe zu lernen und zu lehren. Einige von Euch werden mich unheimlich finden, aber das liegt an Euch Menschen, nicht an mir.

Sehr geehrte Damen und Herren! Was lösen diese Worte bei Ihnen aus? Wenn Sie dabei das Bild der Maschine Sophia betrachten, der diese Worte von Ihren Programmierern in den Mund gelegt wurde. Sophia hat Sensoren, Aktuatoren, Silikonhaut. Sie kann rechnen, sie kann in Wissensdatenbanken recherchieren und unsere Fragen beantworten. Sie hat keine Gefühle, keine Empathie, und sie kann auch nicht erröten. Das ist aber die Standardantwort, oder zumindest war es bis 2017 im Sprachmodus von Siri, wenn Nutzerinnen sie beschimpft haben, sie sei eine Schlampe. Ich würde vor Scham erröten, wenn ich könnte. Wie gesagt, Maschinen haben keine Empfindungen, sie haben keine Ahnung, welche Bedeutung unsere Sätze oder Fragen haben. Ist es denn dann nicht egal, ob wir sie beschimpfen? Ist es angemessen, wenn sie auf unangemessene Kommentare von Nutzerinnen so reagieren? Oder hat das auch Konsequenzen für uns Menschen und unsere Gesellschaft? Ich möchte mich herzlich für die Einladung bedanken und heute hier über eine Genderperspektive auf KI-Technologien und Digitalisierung sprechen und versuchen, mit Ihnen ein paar Antworten zu spannenden Fragen zu finden. Was habe ich vor heute in diesem Vortrag? Zunächst werde ich über Gender Machines sprechen und der Frage mit Ihnen nachgehen, wie menschlich Maschinen eigentlich sein sollten. In einem 2. Schritt werden wir über Gender Data sprechen, da werde ich der Frage nachgehen, ob Vorstellung herrscht, ja das Maschinen objektiv seien, ob wir dadurch Diskriminierung vermeiden könnten, indem wir Algorithmen einsetzen. Und in einem 3. Schritt werde ich dann über den Gender Divide sprechen, nämlich der Tatsache, dass Frauen in Technik, Mathematik und Naturwissenschaften massiv unterrepräsentiert sind., und ich der Frage nachgehen werde, ob es mehr Frauen in der Technik braucht. Wenn Sie jetzt, ich schalte da jetzt schnell um, wenn Sie jetzt dieses Bild betrachten, das wurde bereits in den 40er-Jahren programmiert, von Heider und Simmel und zwar ein Studienexperiment, wo man Menschen dazu eingeladen hat, diese Objekte zu betrachten, die sich hier bewegen und sie gefragt haben, was sie denn da sehen? Und das spannende ist, dass diese Personen nicht gesagt haben, sie sehen hier Dreiecke und einen Kreis, die sich im Raum bewegen, sondern sie haben erzählt von Kooperation, von Konflikt, von unfairem Verhalten, von Eifersucht. Also das heißt, sie haben diesen bewegten Objekten menschliche Eigenschaften, Gefühle, Intentionen usw. zugesprochen. Diesen Prozess nennt man Anthropomorphisierung, und das tun wir Menschen. Wir Menschen schreiben auch Maschinen und anderen Objekten menschliche Eigenschaften zu, insbesondere, wenn sie bewegte Objekte sind. Und so passiert das auch bei Robotern, und diesen Effekt den nutzen wir auch. Und zwar, weil je menschenähnlicher wir eine Maschine auch wahrnehmen, umso vertrauter ist sie uns auch und umso höher ist sozusagen das Vertrauen oder die Akzeptanz in diese Maschine. Das hat ein Wissenschaftler, Mori, so beschrieben: Also, was Sie hier in der Grafik sehen, ist eben genau dieser Zusammenhang, den ich Ihnen jetzt gezeigt habe, und das will man in der Robotik nutzen. Das heißt, man macht Roboter menschenähnlich, man macht Schnittstellen, die uns erlauben mit natürlicher Sprache mit Ihnen zu kommunizieren. Und je mehr wir das tun, also je menschenähnlicher diese Maschinen sind, umso einfacher wird es für uns, mit ihnen zu interagieren. Und das ist insbesondere in der sozialen Robotik besonders wichtig. Erst wenn Sie jetzt wie das Bild von Sophia, die dem sehr, sehr nahe an einem Menschen herankommt, erst dann kann es passieren, dass uns das irritiert., weil wir das sozusagen nicht mehr genau wissen, ist das jetzt noch eine Maschine, was ist das, was haben wir da vor uns? Es ist aber nicht nur so, dass wir soziale Roboter haben, die eine Verkörperung haben, sondern wir haben auch soziale Roboter, die ohne Verkörperung sind. Soziale Roboter eben wie schon vorhin gesagt interagieren und kommunizieren mit uns Menschen, indem sie dabei erwarteten Verhaltensnormen folgen und daher von uns auch als soziale Akteure wahrgenommen werden. Wenn Sie eine Verkörperung haben, vielleicht kennen Sie die Pepper im Servicebereich zum Beispiel, oder Sie haben vielleicht schon den Aibo gesehen, den kleinen

Spielzeughund, der also in die Unterhaltungsrobotik gehört. Es wird in letzter Zeit sehr viel über Pflegeroboter und Pflegeassistenzsysteme gesprochen. Da hatten wir an der TU Wien einen Hobbit zum Beispiel, oder vielleicht kennen Sie die Robbe Paro, die man streicheln kann und die da auf Menschen reagiert, als therapeutischer Roboter für zum Beispiel demenzerkrankte Menschen. Ohne Verkörperung kennen wir JetBots, da begegnen sie uns recht häufig oder wir kennen sie in Form von diesen persönlichen Assistenzsystemen wie eben Siri oder Alexa. So, jetzt gibt es diese Designentscheidungen für Roboter, mit oder ohne Verkörperung, und die Frage ist, was sind das für Designentscheidungen. Wir entscheiden uns dafür, welches Aussehen die haben sollen, wenn sie physisch verkörpert sind. Aber wir geben ihnen auch Namen Siri oder Alexa. Wir geben ihnen eine Stimme männlich, weiblich, kindlich, aber auch darüber hinaus programmieren wir sie mit bestimmten Charakteren sogar, bestimmten Charaktereigenschaften, und sie sind programmiert für bestimmte Rollen und Aufgaben. Und all das führt dazu, das wir ihnen implizit auch oft sehr starke Gendermerkmale zuteilen. Und sehr oft, und das ist ein Argument, sind diese zugeteilten Geschlechtermerkmale sehr stereotyp ausgeprägt. Und ich zeige Ihnen jetzt ein paar Beispiele von berühmten RobotikerInnen, die Roboter kreiert und designt haben. Zum Beispiel Nadia Thalmann spricht davon, dass sie gerne, weil sie erlebt hat, wie ihre Mutter im Altersheim alleine war, dass sie gerne einen Roboter hätte, einen Gefährten, eine Helferin, dass sie, wenn sie einmal alt ist, unterstützt, damit sie nicht so alleine ist und hat ihr Ebenbild sozusagen geschaffen. Hier die Vorstellung, die Helferin im Alter, die mich pflegt und die mich unterstützt. Hiroshi Ishiguro, ein japanische Robotiker, er beschäftigt sich mit Schönheit und Ideal. Auch er hat sich ein Ebenbild geschaffen, einen Geminoid. Aber er hat zum Beispiel auch eine sehr berühmte, japanische Fernsehsprecherin nachgebaut. Und er sagt, dass eben die ideale Frau ein Roboter sei, den er erschaffen hat. Also auch hier eine stereotype Zuschreibung oder ein Schönheitsideal, das hier in einer Maschine implementiert ist. Matt McMullen ist der bekannte Sexroboterhersteller aus Kalifornien, das ist ein Modell von ihm. Sie können diese Sexgefährtinnen in unterschiedlichen Varianten auch haben, als unterschiedliche Charaktere. Und er möchte eben, dass Menschen sich tatsächlich, also eine emotionale Bindung zum Roboter aufbauen. Und sondern eigentlich nicht nur zum Roboter, sondern zum Charakter dahinter, den er programmiert hat. Also Sexgefährtin, eine ideale schöne Frau, Partnerin, eine Helferin, sehr stereotype Verkörperungen in den Robotern. Es gibt aber auch männliche Roboter, und hier sehen sie 2 wunderbare Beispiele: Nämlich einerseits von Boston Dynamics, Atlas kann Backflips machen und in der freien Natur draußen gehen und sich zurechtfinden und ist als Stereotyp maskulin und ebenso der Roboter von NASA, den die NASA in einer DARPA-Challenge entwickelt hat als Astronauten. Was lernen wir daraus? Wir lernen daraus, dass wir, wenn wir nicht sehr aufpassen im Design von Robotern, dabei enden, dass wir Geschlechterrollen und Geschlechterstereotypen nicht nur verfestigen in dieser Technologie, sondern sie eben auch sogar verstärken. Und wir lernen, dass Roboter eben keine neutralen Maschinen sind, sondern das sie kulturelle Artefakte unserer Kultur sind, die diese Geschlechterstereotypen widerspiegeln. Was bedeutet das jetzt aber für uns Menschen, wenn wir mit diesen Technologien vermehrt kommunizieren, uns mit ihnen auseinandersetzen? Als Amazon Echo in die Wohnzimmer dieser Welt Einzug gehalten hat, hat es einen Hilfeschrei gegeben von einem Journalisten, der gesagt hat: „Amazon Echo is magical, it's also training my kid into an asshole“. Und zwar deswegen, weil er beobachtet hat, dass diese Technologie, dieses Assistenzsystem sehr freundlich selbst auf Beschimpfungen agiert. Und seine Kinder, ich glaube, seine Tochter, eben damit eigentlich keine Zurechtweisung in der sozialen Interaktion mehr erfahren hat, weil egal was man zur Alexa sagt, sie reagiert immer freundlich, ja. Was passiert mit uns, wenn wir mit diesen Systemen interagieren, niemals ein „nein“ hören, niemals „hör auf, das tut mir weh“, niemals „das ist aber nicht angemessen“ hören. Wir entwickeln ja unsere eigene soziale Identität in der Auseinandersetzung mit unserer Umwelt. Und bereits Hegel hat deutlich gemacht, das Herr- und Knechtschaft nicht unabhängig voneinander sind und zeigt damit die Dialektik von Herr und Knecht als Quelle des Selbstbewusstseins und der eigenen Identität auf. Was passiert also mit uns? Können wir gute Herren werden, wenn wir permanent uns mit digitalen Knechten umgeben? Eine weitere Folge ist, dass sich auch unsere Erwartungen an andere durch diese Interaktion mit diesen Technologien verändern können. Und zwar deswegen, weil wir eben in bestimmte Interaktionsmuster mit diesen Technologien, die relativ eingeschränkt noch sind, weil wir die verfestigen könnten oder eben reichere Interaktionsmuster nicht

mehr erlernen, ja. Sherry Turkle hat ein warnendes Buch geschrieben, nämlich „Alone Together. Warum wir weniger voneinander und mehr von Technologien erwarten“, und hat darin aufgezeigt, wie sehr sich unser Interaktionsverhalten verändert. Und ich finde dieses eine Zitat, das ich Ihnen aus diesem Buch hier hergeschrieben habe, sehr bezeichnend, wo ein junger Bursche mit 14 sagt, irgendwann werde ich lernen, mit Mädchen richtig zu sprechen, aber noch nicht jetzt, ja. Also, da sieht man, wie sehr sich bei jungen Leuten ihr Sozialverhalten und ihre Sozialkompetenzen schon verändert und eingeschränkt haben. Und wenn wir jetzt hören von der Gartner Group, dass bereits 2020 20 % von Menschen mehr mit ihren Sprachassistenzsystemen sprechen als mit ihrem PartnerInnen, dann ist das vielleicht kein gutes Zeichen. Und wenn wir einen Blick nach Japan werfen, dann werfen wir vielleicht einen Blick in unsere Zukunft, wie es bei uns aussehen wird. Und was Sie hier sehen, sind 2 Beispiele nämlich einerseits Gatebox. Das ist ein holographisches, virtueller Companionrobot oder andererseits die Loveplus-App, das ist auch eine virtuelle Freundin, also KI-Assistenzsysteme oder KI-Systeme, die bestimmte Charaktere haben. Und ganz viele junge japanische Männer gehen eben nur mehr diese platonischen Liebesbeziehungen mit solchen virtuellen Freundinnen ein und vereinsamen dabei. Ja, was kann man dagegen tun, oder was ist zu tun? Das eine ist tatsächlich im Design, in Designentscheidungen sehr kritisch auch zu sein. Und ich habe das Undoing Gender genannt, nämlich eine Überwindung von Stereotypisierung durch gendersensible Designentscheidungen. Und ich habe Ihnen hier 2 Beispiele gebracht, es gibt „Q“, das ist der Versuch eine geschlechterneutrale Sprache zu implementieren, also weder männlich, weiblich noch kindlich, sondern eine Stimme, die alles sein kann oder das Foto von Telenoid, das ist ein ferngesteuerter Roboter von Hiroshi Ishiguro. Und Sie sehen, dass der zwar sehr menschenähnliche Züge hat, aber Sie weder Geschlecht noch Alter aus seinem Aussehen schließen könnten. Ja, der 2. Punkt, über den ich sprechen möchte, ist eine andere Dimension von Gender und in diesem Fall stärker Digitalisierung. Ich möchte auch darüber sprechen, ob KI-Maschinen oder Algorithmen künstliche Intelligenz, ob sie objektiver oder fairer sind und wir dadurch Diskriminierung verhindern können. Und dazu möchte ich Ihnen dieses Bild zeigen, das ist aus 2016, ein Bild, das die „Presse“ veröffentlicht hat und titelt in einem Artikel „Frauenvorrang löst Chaos in Stockholm aus. Die rot-grüne Stadtregierung lässt neuerdings aus „Gendergründen“ Gehwege statt Straßen zuerst vom Schnee räumen, da Frauen angeblich eher zu Fuß gehen. Die Folge: ein Fiasko.“ 2016, jetzt fragt man sich, was hat denn jetzt eigentlich die Stadträumung mit Gendermainstreaming oder mit Gender, mit Diskriminierung zu tun? Und tatsächlich war das so, dass sich die Stadtregierung in Stockholm sich überlegt hat, in jedem Bereich eine Genderdimension auch zu analysieren und schon gibt es da Genderdimensionen. Und man ist draufgekommen, dass besonders im Winter eben natürlich, wenn die schlechte Jahreszeit kommt, vermehrt Frauen und Kinder Unfälle mit Beinbrüchen etc. erleben. Man hat sich gefragt, welche Ursachen das haben könnte? Jetzt könnte man argumentieren, die sind vielleicht ein bisschen podscherter, wenn es regnet oder schneit als Männer. Oder man könnte sich vielleicht überlegen, dass das andere Ursachen hat. Der Grund ist der, dass auch in Schweden es noch trotz eines hohen Equality Index, sehr tradierte Arbeitsaufteilungen und Geschlechterrollen es noch gibt. Und dass es auch in Schweden so ist, dass tendenziell eher die Frauen, also tendenziell, fangen wir es andersrum an, die Männer das Familienauto verwenden, um in die Arbeit zu fahren und tendenziell eher die Frauen die Kinder mit den öffentlichen Verkehrsmitteln zuerst in die Schule bringen und dann mit den öffentlichen Verkehrsmitteln auf ihren Arbeitsplatz gehen. Und das ist die Ursache dafür, dass sie häufiger verunfallen als Männer, weil wie in allen europäischen Städten bis 2016 zuerst die Straßen geräumt wurden, wenn es Schneefall gegeben hat und vielleicht erst danach die Gehwege, damit die Autos freie Fahrt haben, und das hat man in Schweden eben geändert. Und das ist hier ein wunderbares Beispiel dafür, dass es abgesehen von dieser interpersonellen Diskriminierung, dass ich zum Beispiel jemanden diskriminiere, weil ich sehe, es ist eine Frau oder eine Person mit einer bestimmten Ethnie. Abgesehen davon gibt es strukturelle Diskriminierungen, und die sind allgegenwärtig, aber eben nicht auf den 1. Blick sichtbar. Und wenn ich jetzt KI-Systeme entwickle, die also Daten von unserer Gesellschaft hernehmen, analysieren und mit neuronalen Netzen, und Sie sehen ein einfaches, neuronales Netzwerk, im Grunde genommen Wahrscheinlichkeiten berechnen, dann können solche kausalen Zusammenhänge, die ich Ihnen gerade vorhin erklärt habe, nicht entdeckt werden. Das heißt, die Gefahr ist sehr, sehr groß, dass institutionelle, strukturelle Diskriminierung in den Daten drinnen bleibt, selbst und auch damit in den

Kenntnissen oder in den Aussagen dieser KI's. Selbst, wenn ich versuche einen Gender Bias, einen offensichtlich herauszurechnen. Auch dafür gibt es berühmte Beispiele: Das eine war dieses Beispiel auch von Google. Google hat automatisch Fotos klassifiziert, diese App, und das war der Aufschrei einer schwarzen Person, wie er draufgekommen ist, dass Fotos von ihm und seiner Freundin in dem Ordner „Gorilla“ kategorisiert wurden, dass er gesagt hat, meine Freundin ist aber kein Gorilla. Google hat das mittlerweile eben wieder korrigiert, aber sozusagen es passiert sehr, sehr häufig, dass diese KI-Systeme implizite Biases und explizite Biases reproduzieren. Und wenn Sie sich jetzt das Bild ansehen, dann werden Sie sehen, dass bei KI-Systemen ein Bias in allen möglichen Phasen der Entwicklung dieses Systems entstehen kann, also nicht nur in den Daten, auch in der Konzeption selbst, in der Evaluierungsphase. Und es gibt eine Vielzahl an Entscheidungen, die Menschen treffen, um so ein System zu entwickeln. Das heißt, die Vorstellung, dass es ein objektiveres System wäre, kann gar nicht gehalten werden, weil eine Vielzahl von Designentscheidungen eben das System genau zu einem solchen gemacht haben, wie es ist und es daher ebenfalls wieder ein kulturelles Artefakt ist und nicht objektiv. Auch hier kann man fragen, was kann man tun? Da hat die ExpertInnengruppe der Europäischen Kommission die High-Level Expert Group Artificial Intelligence Ethikrichtlinien entwickelt, ich durfte da auch mitarbeiten. Und wir haben eben festgelegt, wie man vertrauenswürdige KI-Systeme entwickeln kann. Also wie man verhindern kann, dass es Biases gibt oder dass diese Modelle unfair sind. Und da ist neben der Datenqualität und Sicherheit insbesondere die Transparenz von Modellen, also wer welche Zielsetzung verfolgt das System, wer arbeitet daran mit, wie sehen die Parameter aus, die Erklärbarkeit von Modellen, die Nachvollziehbarkeit von Entscheidungen notwendig. Und ebenso natürlich die Kontrolle, die Autonomie und die Übernahme von Verantwortung durch Menschen notwendig. Das bringt mich jetzt zum letzten Punkt: Nämlich darüber zu sprechen, wenn ich die anderen beiden jetzt hernehme und sage, eigentlich haben wir Gender Machines, und wir haben eigentlich auch Gender Data, was könnte ich denn jetzt tun? Wie könnte ich die Situation verbessern, dass wir in Zukunft bessere Systeme machen, die fair sind, die nicht Stereotype reproduzieren. Dann könnte ich ja eine Genderbrille auch in der Entwicklung dieser Technologie verwenden und könnte versuchen, mehr Frauen in diese Entwicklung dieser Technologien auch zu bringen. Und darüber möchte ich jetzt im 3. Punkt sprechen. Es war nicht immer so, dass wir so wenig Frauen in der Programmierung zum Beispiel hatten. Dieses Foto wurde in den 50er Jahren aufgenommen, dort war Programmierung also zu dieser Zeit noch Frauensache. Wie sieht es heute aus? Was ist passiert? Der UNIDO-Report zeigt, dass es eine Geschlechterkluft gibt, zwischen der relativ homogenen Gruppe von Ingenieuren, die KI-Technologien entwickeln und der Gruppe der diversen Nutzerinnen, nämlich nur 12% der Entwicklerinnen sind Frauen in der KI-Technologie.

Woher kommt das? Wenn man kleine Kinder noch befragt, da gibt's kaum noch einen Unterschied in den Kompetenzen und auch in den Interessen von Mädchen und Burschen. Aber ab einem gewissen Alter sozusagen, in der primären Schulbildung, beginnen sich diese Stereotypenvorstellungen durchzusetzen, und die Mädchen verlieren ihr Selbstvertrauen in ihre Kompetenzen, und was Sie in dieser Graphik sehen, ist, dass abgesehen von den Ländern Thailand und Türkei die Mädchen signifikant höhere Kompetenzen in „Computer and Information Literacy“ haben, aber ihr Selbstvertrauen geringer ist als jenes der Burschen in diesen Bereichen. Das führt dann eben auch dazu, dass sie, oder auch zusätzlich, haben sie eine geringere Selbsteinschätzung in ihren Kompetenzen in der Mathematik. Also, auch da, wenn man sich anschaut, die Selbstwahrnehmung von Mathematikkenntnissen, von 15 und 16-jährigen Burschen oder Mädchen, dann trauen sich die Mädchen einfach deutlich weniger zu in der Mathematik. Und das führt dann in weiterer Folge zu sehr stereotyper Studienwahl. Die Mädchen entscheiden sich für Humanities, also für Studienrichtungen, in den Sozialwissenschaften, in den Erziehung, im Kunstbereich und die Science Technology Engineering and Mathematics-Fächer, bei uns auf Deutsch, die MINT-Fächer, die bleiben vorzugsweise Studienentscheidungen von Männern. Und das führt dann schlussendlich auch dazu, dass sie jetzt, wenn sie sich Tech-Berufe anschauen und große Tech-Konzerne anschauen wie Google, dass Frauen dort in allen Ebenen unterrepräsentiert sind. Und das ist mitunter ein Grund, warum viele der Technologien einfach genderstereotyp entwickelt werden, diese Genderdimension unberücksichtigt bleibt, weil eine sehr homogene Gruppe von männlichen Ingenieuren gar nicht auf die Idee kommt,

dass das vielleicht ein Problem darstellen würde. Jetzt zeige ich noch eine Graphik, die sehr spannend ist. Das ist das Gleichstellungsparadoxon. Was Sie hier sehen ist die paradoxe Situation, dass in Ländern, die einen höheren Gleichstellungsindex bereits haben, signifikant weniger Frauen mit ICT-Studienabschlüssen haben. Das heißt, es gibt einen negativen Zusammenhang zwischen Gleichstellung von Männern und Frauen und der Studienwahl oder ICT-Abschlüssen. Das heißt, Sie finden zum Beispiel in arabischen Ländern den höchsten Anteil von Frauen in Informatik und in Ländern wie Österreich liegen wir da eben, wie die Zahlen auch zeigen, sehr, sehr niedrig. Und dieses Paradoxon kann man nicht abschließend noch nicht erklären, aber eines zeigt dieses Paradoxon jedenfalls, dass es möglich ist, Frauen für dieses Studium zu gewinnen und zu begeistern. Die arabischen Länder machen uns das vor. Da gibt es über 60 % Studienabschlüsse von Frauen. Also, was ist zu tun? Und auch darauf geht dieser UNESCO-Report „I'd blush if could“ ein und gibt insgesamt 15 Empfehlungen ab, die ich Ihnen hier jetzt nicht vorlesen werde, die Sie dann in Ruhe aus den Folien auch nachlesen können. Es bedarf jedenfalls vielfältiger Initiativen. Die beginnen damit, dass wir Lehrerinnen und Lehrer mit Genderkompetenz ausstatten müssen, dass wir ICT-Kompetenzen bereits in der Primarstufe auch erwerben sollen, dass wir sichere Räume brauchen, wo sich auch Mädchen diese Kompetenzen aneignen können. Das heißt, durchaus auch Ausbildungskurse für Mädchen, wo sie auch positive Lernerfahrungen machen können etc., etc. Ja, damit bin ich am Ende meines Vortrags. Ich bedanke mich für Ihre Aufmerksamkeit, wieder mit dem sehnsuchtsvollen Blick der Sophia und hoffe, dass einiges für Sie dabei war.

Vielen Dank, Frau Köszegi, für diese interessante Perspektive auf KI-Technologien. Last but not least, hören wir Martina Angela Friedl zum Thema: „Rettet uns die Technik? Strategien im Umgang mit ‚Unconscious Bias‘“. Frau Friedl ist promovierte Chemikerin, Organisationsberaterin, Trainerin und Coach. Ihre Arbeitsschwerpunkte liegen im Bereich „Training und Organisationsberatung“ zu den Themen: „Führung“, „Frauen in der Führung“, „Diversity Management“ und „Gestaltung von Veränderungsprozessen“. Als Coach begleitet sie außerdem viele Führungskräfte bei deren Karriereentwicklung. Sie arbeitet in vielen unterschiedlichen organisationalen Kontexten, unter anderem an der Verwaltungsakademie des Bundes. In Ihrem Vortrag geht Sie auf Strategien, zum Umgang mit unbewusster Diskriminierung (auf Englisch: „Unconscious Bias“) ein und fragt sich, inwiefern künstliche Intelligenz ein probates Gegenmittel ist? Liebe Frau Friedl, die Bühne ist Ihre.

Vielen Dank. Freue mich sehr, dass ich heute hier sein kann, und ich werde jetzt in meinem Vortrag zum Thema „Unconscious Bias“ sprechen und dabei auch die Frage beantworten, ob uns die Technik bzw. künstliche Intelligenz bei diesem Thema rettet und hilft, oder eben auch nicht. Wie ist mein Vortrag heute aufgebaut? Ich starte mit einem kurzen Überblick zum Thema „Unconscious Bias“ und erkläre auch den Begriff noch einmal und stelle hier einige Phänomene vor. Als nächstes werde ich darauf eingehen, inwiefern eben bei diesem Thema künstliche Intelligenz in Zukunft eine Rolle spielen wird oder kann. Und am Ende möchte ich ein paar Auswege und Strategien im Umgang mit „Unconscious Bias“, eben unbewusster Diskriminierung, vorstellen. Ja, ich starte mit dem Begriff an sich, „Unconscious Bias“. Das war ja auch ein Password und ein ganz wichtiges Wort in den letzten 1, 2, Jahren, das sehr heftig diskutiert wird immer wieder. Gemeint ist damit unbewusste Diskriminierung. Das heißt, Diskriminierungsvorfälle, die sehr schnell passieren und auf Stereotypen und unseren Modellen in unserem Kopf passieren. Ich werde es gleich auch noch im Detail ausführen. Fakt ist, wir alle haben unbewusste Vorteile, auch wenn wir uns noch für so objektiv halten und reflektiert und offen. Sogenannte Denkfehler, sagt man auch, haben wir deswegen auch alle, weil sie eine Funktion haben. Sie kommen aus unserer Menschheitsgeschichte, wo es immer darum gegangen ist, wenn wir, stellen wir uns vor, in der Steinzeit anderen Menschen begegnet sind, dass wir dort sehr schnell entscheiden haben müssen: Ist das Freund oder Feind, und muss ich eben kämpfen oder davonlaufen oder mich im schlimmsten Fall totstellen? Das heißt, unser Gehirn ist darauf trainiert Menschen sehr schnell einzuschätzen, in Schubladen zu stecken mit der Fülle an Sinneseindrücken, die jeden Tag auf uns hereinprasselt insofern so umzugehen, als dass gewisse Dinge auch ausgeblendet werden oder schnelle Schlüsse gezogen werden, und so kommt es eben zu diesen Phänomenen, von denen ich heute spreche. Der Nachteil oder das Negative an diesen unbewussten Vorurteilen ist, dass

sie zum Beispiel im Recruiting oder auch im Führungsalltag, wenn es darum geht, wen wir fördern, zum Beispiel einem Team, sie dazu führen, dass wir Talente übersehen, dass wir gewisse Menschen bevorzugen und dass wir nicht die bestmöglichen Entscheidungen treffen. Die schlechte Nachricht vielleicht ist, man kann diese unbewussten Vorteile nicht einfach so loswerden, indem man es sich vornimmt, aber man kann doch eine Auseinandersetzung damit, so wie wir es heute tun, schon sehr viel bewusster damit umgehen und auch Wege finden, diese Effekte zu reduzieren. Insgesamt gibt es sehr, sehr viele dieser Phänomene, die auch gut beschrieben sind. Ich habe einmal eine sogenannte „Unconscious Bias“-Map gesehen, da waren 170 dieser Effekte beschrieben. Ich habe einen kleinen Teil für heute herausgesucht, der glaube ich, sehr anschaulich ist und der auch im Führungsalltag oder auch beim Thema Recruiting zum Beispiel eine Rolle spielt. Genau so. Nächste Folie. Ja, Sie sehen hier einen Überblick, und ich sag dann zu den einzelnen Phänomenen überall etwas dazu. Starte mit dem sogenannten Gender Bias. Es steht auch dabei, innere Bilder zu Technik und Führung helfen tendenziell eher Männern. Damit ist gemeint, dass, wenn man Menschen draußen auf der Straße fragt: Wer ist denn so die ideale Führungskraft? Oder wenn man auch in Organisationen herumfragt, wer denn das ist, dann fallen uns oft Männer als erster ein. Genauso ist es bei Berufen im technischen Umfeld. Das bedeutet wiederum für Frauen, die sich in solchen Feldern bewerben, zum Beispiel umstellen oder dort weiterkommen wollen, dass sie genauer geprüft werden, dass genauer hingeschaut wird und immer wieder so die Frage im Raum steht: Kann Sie's denn wohl? Übrigens umgekehrt gilt es auch für Männer, die sich in klassischen Frauendomänen bewerben. Also der Mann, der sich im Kindergarten bewirbt, wird genauso stärker geprüft als die Frauen. Was noch dazukommt, das habe ich hier jetzt auch auf der Folie unten im zweiten Teil aufgeschrieben: Ein weiterer Aspekt von Gender-Bias ist, dass erfolgreiche Frauen oft auch noch sozusagen diesen Nachteil haben, dass sie als unsympathisch beschrieben werden. Es gibt hier dieses schöne Beispiel von der Heidi Royce. Das ist eine Amerikanerin, die im Silikon Valley sehr erfolgreich war und immer noch ist, und da gab es eine Fallstudie, die an den Universitäten getestet wurde, wo Studierende gefragt wurden, ob sie denn die Heidi Royce oder den Howard Royce, also einen Mann, im Team haben wollen. Beide hatten den komplett selben Lebenslauf, und es war wirklich statistisch nachweisbar, dass Howard, das heißt, dass der männliche Kollege viel öfter eingestellt worden ist als eben die Heidi, und es wurde und es wurde immer wieder begründet, sie ist uns irgendwie nicht sympathisch. Das heißt, hier haben Frauen auch noch einen doppelten Nachteil. Ja, Gender-Bias. Ich komme zum nächsten Bias. Der hat diesen schönen Namen „Beauty Premium“ oder auch „Lookismus“, sagt man. Schönheit zahlt sich aus, steht da. Wir sehen es hier auch auf diesem Bild, ja, wo sie links diese Comicfiguren sehen. Wir Menschen reagieren einfach auf Schönheit. Und Menschen, die gängigen Schönheitsidealen entsprechen, sind nachweislich erfolgreicher. Denen wird mehr vertraut, denen werden mehr Sympathien entgegengebracht und, ganz konkret, es gibt Studien dazu, diese Menschen verdienen auch mehr. Geprägt werden diese Bilder ganz, ganz früh. Sie sehen es links da bei den Comicfiguren, das heißt, eine symmetrische Körperform zum Beispiel wird schon ganz früh geprägt, dass es mit der guten Rolle in einem Comic geprägt ist und die Bösen zum Beispiel wie der Gargamel, den Sie hier sehen, die haben immer einen Buckel oder die Hexe zum Beispiel wird eben eher mit krummen Gesichtszügen zum Beispiel dargestellt. Genauso wichtig, oder ein wichtiger Faktor bei dem Thema Schönheit ist die Körpergröße. Da gibt es auch einen klaren Zusammenhang bei Männern, wo man sagen kann: Idealerweise eine Größe um 1,80 herum ist ein echter Karrierebooster, alles darüber oder darunter eher weniger. Und auch im Negativen kann man sagen, bei Frauen zum Beispiel wird das Übergewicht viel negativer bewertet und schlägt sich auch negativ nieder, weil sofort die Idee im Raum ist, wenn jemand übergewichtig ist, dann ist das jemand, der nicht so leistungsfähig ist zum Beispiel. Ja, zu schön darf man auch nicht sein. Das ist für Frauen halt wieder so dieser Spagat. Ja, also, wenn man als Frau sozusagen zu schön ist, was immer das dann genau bedeutet, wird auch sehr schnell unterstellt, ob man denn klug ist. Als ob es da einen Widerspruch geben würde. Ja, und das war das YouTube Freemium dem wir uns oft blenden lassen. Ein weiteres Bias, den ich noch vorstellen möchte, ist dieser Affinity-Bias oder auch Mini-Me-Effekt, das kennen Sie vielleicht. „Ähnlichkeit schafft Vertrauen“ steht auf der Folie. Man sieht das immer wieder, wenn man diese Fotos sieht von zum Beispiel EU-Gipfeln. Wenn man dann schaut, ok, wer ist denn da drauf? Man hat den Eindruck, die schauen irgendwie alle gleich aus – ja, alles Männer, alle ungefähr im selben Alter, plus, minus, alle gleich angezogen. Man hat den Eindruck, die waren vielleicht

eh alle auf derselben Uni. Da ist einzig noch die Angela Merkel, die ein bisschen heraussticht. Das ist ein klassisches Beispiel von Mini-Me-Effekt. Das heißt, wir vertrauen Menschen mehr, die uns ähnlich sind, die uns in gewissen Eigenschaften wie Herkunft, Geschlecht, Alter, soziale Schicht usw., also, die uns da ähnlich sind. Und wenn es darum geht, zum Beispiel die Nachfolge zu planen oder zu überlegen, mit wem möchte ich gerne zusammenarbeiten, dass wir eher diesen Menschen den Vorzug geben, die uns ähnlich sind. Führt natürlich dazu, dass – wie soll ich sagen, Menschen, die nicht dem Mainstream entsprechen, immer wieder nicht weiterkommen oder sich einfach schwerer tun. Ja, den Namen-Bias stell ich auch noch vor. Das haben Sie sicher mal gehört, da gab es diesen Ausspruch einer Lehrerin, die gesagt hat: „Kevin ist mehr eine Diagnose als ein Name“. Ja, also Namen erzeugen in unserem Kopf sofort Geschichten. Wir lesen einen Namen. Wir haben die Person nie gesehen, und haben sofort die Idee, wer die Person ist, woher sie kommt, was sie kann, was sie nicht kann. Und das wirkt sich natürlich aus, wenn man weiß, sozusagen, wenn man sich mit gleichem Lebenslauf, aber mit exotischem Namen in Österreich bewirbt, man viel, viel mehr Bewerbungen abschicken muss, um überhaupt eingeladen zu werden. Und man weiß auch zum Beispiel in der Schule und auch an Universitäten, dass man mit exotisch klingenden Namen tendenziell auch schlechtere Bewertungen bekommt. Ja – Namens-Bias. Ja – „Halo-Effekt“ stell ich noch vor. Der „Halo-Effekt“ („Halo“ bedeutet „Heiligenschein“ auf Englisch), damit ist noch gemeint, dass wir uns manchmal, wenn wir Menschen kennenlernen, uns oft von einem einzelnen Merkmal blenden lassen, dass dann alles überstrahlt. Das kann so im positiven Sinn sein, dass jemand zum Beispiel an einer Top-Uni studiert hat, sag ich jetzt einmal, und dass wir von daher sagen, dass muss ein ganz kompetenter Mensch sein und gar nicht mehr so genau hinschauen. Das kann aber im Negativen sein zum Beispiel, das berichten auch Menschen, die zum Beispiel auch körperliche Behinderung haben, die sichtbar ist und sie sagen, sie werden immer nur auf dieses Merkmal reduziert, und ihre Leistungsfähigkeit wird in Frage gestellt, obwohl vielleicht die Behinderung gar kein Problem ist für diesen Job, für den sie sich bewerben. Ja – gut, zu guter Letzt noch zum Thema Ageism. Also, das ganze Thema: „Alter“, da gibt es auch sehr viele Bias, weil wir natürlich Ideen im Kopf haben, wann wir für welche Tätigkeit zu jung oder zu alt ist. Wann man eine erste Führungsposition haben sollte, wann man mit dem Studium fertig sein sollte usw., obwohl das natürlich nichts damit zu tun haben sollte mit dem, was man leisten kann in einem Team oder einer Organisation und ganz am Ende noch diesen Confirmation Bias, der zielt darauf ab oder sagt uns, dass unsere Wahrnehmung manchmal auch sehr selektiv ist, dass wir manchmal, wenn wir Menschen begegnen, neu kennenlernen, gewisse Dinge sehen, die in unser Konzept passen, was wir glauben, was richtig ist, wie man sich zu verhalten hat oder was einem wichtig ist, wenn man in dem Alter oder mit der Ausbildung unterwegs ist, und dass wir alle anderen Dinge, die uns da vielleicht überraschen würden, einfach ausblenden und gar nicht wahrnehmen. So, das waren jetzt ganz, ganz viele Beispiele für „Unconscious Biases“. Vielleicht sind Ihnen auch manche dieser Phänomene schon begegnet in der eigenen Organisation. Ich wollte jetzt einfach einen Überblick geben, damit wir gemeinsam uns vorstellen können, um was es geht und mit welchen Phänomenen wir es hier zu tun haben. Eine Frage, die jetzt natürlich auf der Hand liegt, ist: OK, aber was tun dagegen? Wir haben schon vorher gehört, so einfach kann man sie nicht loswerden. Es gibt natürlich ganz viel Forschung auch zu dem Thema, welche Strategien man in Organisationen oder man auch persönlich verwenden könnte, um insgesamt fairer und objektiver zu sein und dadurch vielleicht auch diversere Teams zustande zu bringen. Und eine Frage oder ein Aspekt, der da ganz stark diskutiert worden ist in den letzten Jahren, ist das Thema der Technik, der künstlichen Intelligenz. Weil man ja annehmen könnte, aha – ein Computer, der hat doch diese Vorurteile gar nicht, der ist doch komplett neutral. Der könnte uns doch helfen und wenn wir Computer entscheiden lassen würden zum Beispiel anstatt von Führungskräften, dann wären wir vielleicht insgesamt viel besser dran. Ja, sehr berechtigte Argumentation. Ich gehe jetzt noch darauf ein sozusagen, was sozusagen in den letzten Jahren man dazu herausgefunden hat und welche Beobachtung ich auch gemacht habe zu der Forschung. Ja, was ist das Thema, oder was habe ich beobachtet, und was ist auch schon viel publiziert worden? Also, Algorithmen an sich sind super. Könnten theoretisch neutral sein, aber sie sind ja von Menschen gemacht. Die fallen ja nicht vom Himmel. Das heißt auch, jegliche künstliche Intelligenz ist halt nur so intelligent, wie die Personen, die sie hergestellt haben oder eben auch programmiert haben. Und gerade da, wenn es um diese Algorithmen geht, fällt auch ganz klar auf, es gibt da so einen Gender

Divide sozusagen, also hauptsächlich werden Algorithmen von weißen Männern programmiert – ja, immer noch – und es gibt kritische Forschung mittlerweile zu den Algorithmen, und das ist hauptsächlich, dass das Frauen machen. Das war ein sehr interessanter Aspekt auch. Was sind die Erkenntnisse sozusagen, zu dieser Betrachtung? Ich glaube, ich habe es eh schon anklingen lassen. Algorithmen sind immer nur so gut, wie die Menschen, die es eben programmieren und entwickeln, und sie beruhen immer auf den Daten der Vergangenheit. Das heißt, was oft passiert, ist, dass es zu einer Fortschreibung des Status Quo kommt, und dass Algorithmen tendenziell wieder zu ähnlichen Entscheidungen führen, die man vorher auch schon getroffen hätte. Ich mach's einmal fest an ein paar Beispielen. Ich glaube, dann wird es noch aussagekräftiger, der Punkt, den ich machen möchte.

Also bei Amazon zum Beispiel, das ist eine riesengroße Organisation, hat man versucht, Rekrutinnen sozusagen durch künstliche Intelligenz unterstützen zu lassen. Was ja an sich naheliegend ist, wenn ich so viele Neubesetzungen jedes Jahr vornehmen muss. Und es war die Idee, einen Algorithmus zu schreiben, der die vielen Bewerbungen, die diese Organisation bekommt, einfach durchscreent sozusagen. Und bevor es sich überhaupt ein Mensch anschaut, eine Vorauswahl trifft und anhand gewisser Kriterien herausfiltert, wer sind wohl die besten Bewerber, Bewerberinnen, wen sollten wir auf jeden Fall einladen zu einem Interview. Die haben das eine Weile so gemacht und laufen lassen und mit diesem Algorithmus gearbeitet und am Ende, nach einer Weile die Ergebnisse angeschaut und sind draufgekommen, dass sie durch den Algorithmus ja also noch weniger Frauen zum Beispiel zum Bewerbungsgespräch eingeladen haben als vorher, obwohl das ein Ziel der Firma ist, da diverser zu werden. Und haben angefangen zu suchen, warum denn das so ist. Und sind halt eben draufgekommen, na klar, der Algorithmus greift auf Daten der Vergangenheit zurück. Und wertet sozusagen statistisch aus, welche BewerberInnen bisher am erfolgreichsten waren. Und nachdem die Organisation sehr männerdominiert ist und hauptsächlich Männer in höheren Positionen sind bei Amazon, wurde das als ein Plus bewertet. Und so hat sozusagen der Algorithmus automatisch, wenn eine Frau sich da beworben hat einmal diesen Lebenslauf mit einem Minus versehen und dadurch sind viele Männer durchgekommen. Die haben es versucht, dann irgendwie besser hinzubekommen, diesen Algorithmus, und ihn zu reparieren und bemerkenswerterweise aber nicht geschafft und sind von dem wieder weggegangen. Das heißt, sie haben sich dagegen entschieden, weil sie da zum Schluss gekommen sind, der Algorithmus ist leider nicht fairer berücksichtigt und kann gar nicht berücksichtigt werden, weil wir die Daten gar nicht haben, das, was wir eigentlich möchten und rekrutieren jetzt wieder ohne sozusagen. Ein anderes Beispiel, das ich noch bringen möchte, das auch in den Medien diskutiert worden ist und auch schon erforscht worden ist, ist der Algorithmus vom AMS, wo es darum gegangen ist, die Arbeitssuchenden in unterschiedliche Gruppen einzuteilen. Und es war so die Idee, eine Prognose über deren Chancen am Arbeitsmarkt abzugeben über den Algorithmus. Wenn man da genau hineinschaut, wie der gewertet hat sozusagen. Ist auch klar herausgekommen, was das Problem an dem Algorithmus ist – ja, der ist auch vergangenheitsorientiert. Das heißt, der hat auch, es gibt ja keine anderen Daten als die aus der Vergangenheit ausgewertet, wie denn die Chancen der Arbeitssuchenden in den letzten 10, 20, 30 Jahren waren. Und auf Basis dieser Gegebenheiten sozusagen einfach Menschen mit Plus- und Minuspunkten versehen, und der Algorithmus hat hier zum Beispiel Frauen schlechter bewertet und auch das Alter und bei Frauen schon sehr früh, ich glaube ab sogar 35 oder 40, also sehr, sehr früh sozusagen gab es da schon einen Abzug an Punkten der Chancen. Und was da auch die Kritik war, ist, dass man gesagt hat, okay, man nimmt halt einfach diese Vergangenheit, weil keine anderen Daten da sind und versucht eine Prognose für die Zukunft sozusagen zu erstellen – ohne natürlich berücksichtigen zu können, dass man andere Ziele verfolgt in der Arbeitsmarktpolitik zum Beispiel. Und auch da ist noch nicht ganz klar, wie der Algorithmus jetzt zum Einsatz kommen wird. Aber man sieht schon sehr schön die Probleme, die es da einfach gibt.

Andere Beispiele waren auch die Gesichtserkennung auf Flughäfen oder auch die automatischen Seifenspender, wo es sehr klar rausgekommen ist zum Beispiel bei der Gesichtserkennung auf den Flughäfen war klar in den USA, dass die hauptsächlich bei sozusagen Caucasian, also Weißen, sozusagen funktionieren. Und wenn ich eine andere Hautfarbe habe zum Beispiel, funktioniert die Gesichtserkennung nicht. Das war einfach nicht Absicht, sondern es war einfach sozusagen, der, der es schreibt, der kann sich vielleicht weniger hineinversetzen, an was man sonst noch alles denken soll, und dann merkt man einfach diese Probleme bei den Anwendungen. Ja, also zusammenfassend kann

man sagen, grundsätzlich sehr nachvollziehbar natürlich, dass man sehr viel Hoffnung in diese künstliche Intelligenz setzt und in diese Algorithmen setzt, dass man aber sehr aufpassen muss bei diesen Dingen, was sie denn wirklich können. Weil die Datenbasis halt so ist, wie sie ist und dass diese Biases, die die Personen haben, die diese Algorithmen programmieren, natürlich auch vor allem dann, wenn es wenig mitbegleitet oder wenig reflektiert wird, dass man die auch wieder findet sozusagen in der künstlichen Intelligenz, und dass sie dann nur begrenzt wirklich behilflich ist. Ja, aber was funktioniert, dass wir auch vielleicht auch positiv enden, mögliche Auswege aus dem. Also was kann man denn doch tun? Weil da gibt es auch eine spannende Forschung dazu. Definitiv weiß man, dass Bewusstseinsbildung hilft, also das, was wir heute tun, das heißt, dass man sich Zeit nimmt, sich mit diesen Phänomenen auseinanderzusetzen, dass man sie einfach mal anerkennt. Dass es das gibt, dass wir alle davon betroffen sind, dass wir alle Vorurteile haben, das ist allein schon einmal ein großer Schritt und erhöht einfach die Wahrscheinlichkeit, dass zum Beispiel Rekrutierungsprozesse bewusster ablaufen. Und ich mich einfach stärker zum Beispiel hinterfragen muss, warum ich den einen oder die andere bevorzuge. Das andere ist, dass man versucht natürlich, wenn es gerade um Beförderungen oder Bewerbungen geht, dass man versucht, die Prozesse an sich anzupassen, so dass sie fairer werden und einen, wenn ich die Person bin, die entscheiden muss, mich bestmöglich unterstützen. Das sind natürlich die Ideen, die auch gut erforscht sind und in manchen Ländern auch schon üblich sind – dass man versucht, zu anonymisieren soweit es geht zum Beispiel im Erstkontakt. Man weiß mittlerweile auch, dass es hilft, wenn ich Interviews, wenn ich Jobinterviews führe. Wenn die sehr stark strukturiert sind und sehr darauf geachtet wird, dass die immer gleich ablaufen, dass das hilft. Es hilft, wenn Entscheidungskriterien vorab festgelegt werden und nicht mehr verändert werden dürfen. Wenn über alle Bewerber und Bewerberinnen gleichzeitig entschieden wird und wenn natürlich viel reflektiert wird im Rahmen vom gesamten Prozess. Das hilft schon, plus was ich auch immer wieder gehört habe, Kennzahlen und Ziele. Also wenn man auch vor einem Bewerbungsprozess zum Beispiel immer wieder daran erinnert wird, dass es Ziele gibt, zum Beispiel diverser zu werden, mehr Frauen einzustellen oder auch zu fördern, dass das auch hilft, um da diese Ziele zu erreichen. 3. wichtiger Punkt noch zu diesen möglichen Auswegen ist das Thema Diversität als Führungsaufgabe. Das ist auch wichtig zu sagen, okay, wenn ich dann Personen rekrutiert habe, wenn ich mehr, sozusagen, ein durchgemischteres Team habe, dass Führungskräfte sich diesem Thema auch wirklich annehmen, dass sich jeder auch wirklich wohlfühlen kann im Team, dass Unterschiedlichkeit auch als Wert wahrgenommen wird, als etwas, was zum Einen mehr Konflikte hereinbringt, zum anderen aber auch die Möglichkeit bietet, eben unterschiedliche Sichtweisen zu diskutieren und so zu besseren Lösungen zu kommen, weil sie immer die Hoffnung ist bei diverseren Teams, dass diese auch innovativer sind. Ja, zu guter Letzt, habe ich noch zwei Buchtipps für Sie aufgeschrieben, falls Sie auch noch in mehr Details sich vertiefen wollen und nachlesen wollen, sehr empfehlenswert das Buch von der Iris Bohnet: „What Works“. Das ist die Forscherin, die Professorin aus Harvard, die sehr stark zu diesem Thema geforscht hat, was man denn mit dem Thema, was man denn gegen „Unconscious Bias“-Phänomene tun kann und viele Vorschläge in diesem Buch aufgelistet hat und auch mittlerweile geforscht hat. Und das Zweite ist ein Herausgeberband vom Michel E. Domsch: „Vorurteile im Arbeitsleben“ auf Deutsch. Gibt auch einen guten Überblick über unterschiedliche „Unconscious Bias“-Phänomene und was verschiedene Organisationen schon dagegen getan haben. Ja, ich sage: Vielen Dank für die Aufmerksamkeit und gebe nun weiter zum nächsten Vortrag.

Vielen Dank, Frau Friedl, für diesen umfassenden Einblick. Wir haben heute sehr spannende Inputs zur Frage gehört, ob künstliche Intelligenz geschlechtsneutral ist. Vielen Dank an Frau Ulmer, Frau Köszegi und Frau Friedl für ihre Beiträge, von denen wir vieles mitnehmen können. Am 29. Oktober hat das European Institute for Gender Equality, kurz EIGE, seinen jährlichen Gender Equality Index vorgestellt, der die EU-Mitgliedsstaaten und die EU als Ganzes bewertet, wie weit sie von der Gleichstellung der Geschlechter entfernt sind. Dieses Jahr hat sich der Index zum Schwerpunktthema „Digitalisierung und die Zukunft der Arbeit“ gewidmet. Zum Abschluss und zusammenfassend möchte ich ein paar Key Facts in Bezug auf Digitalisierung und Gleichstellung aus dem Bericht anführen: Frauen haben ein etwas höheres Risiko in ihrem Beruf, durch Roboter ersetzt zu werden. Sie sind bei der Entwicklung von künstlicher Intelligenz, digitalen Start-ups und Hochtechnologieprodukten wie Raumfahrzeugen,

Glasfasern, Lasern und Microchips unterrepräsentiert. Insgesamt dominieren Männer die Entwicklung neuer Technologien in der gesamten EU. Die Ausweitung der über Online-Plattformen organisierten Arbeit reproduziert traditionelle Ungleichheiten zwischen den Geschlechtern, wie etwa das geschlechtsspezifische Lohngefälle und die Geschlechtersegregation. So arbeiten Männer beispielsweise in der Softwareentwicklung oder in der Lebensmittelzustellung, während Frauen mehr in der Onlineübersetzung oder bei häuslichen Dienstleistungsarbeiten tätig sind. In Bezug auf Österreich wurden folgende Punkte speziell hervorgehoben: Bei Menschen mit niedrigem, formalem Bildungsniveau nutzen nur 52 % der Frauen im Vergleich zu 80 % der Männer das Internet täglich. Nur 15 % der Absolventen von Informations- und Kommunikationstechnologie, kurz IKT sind weiblich. 34 % der Frauen im IKT-Bereich arbeiten Teilzeit im Vergleich zu 10 % der Männer. Wir befinden uns erst am Anfang, wenn es darum geht, die Konsequenzen von künstlicher Intelligenz in ihrer vollen Bandbreite zu erkennen. Ein Austausch wie der heutige, ermöglicht es uns von unserer gegenseitigen Expertise zu lernen und so ein besseres Verständnis in Bezug auf das Thema zu verbreiten. Mit sehr hoher Wahrscheinlichkeit lässt sich sagen: Künstliche Intelligenz ist gekommen, um zu bleiben. Unsere Gespräche sind es ebenso. Wir freuen uns bereits auf den nächsten, konstruktiven Austausch. Vielen Dank für Ihre Aufmerksamkeit.