



Trendradar Digitale Ethik

Ausgabe August 2020

Künstliche Intelligenz verstehen:

Ein Fitnessprogramm für Unternehmen, Politik und Gesellschaft

Erfreulicherweise gibt es im deutschsprachigen Raum immer mehr Bildungsangebote zu Künstlicher Intelligenz. Viele sind kostenlos und lassen sich online absolvieren. Ein Beispiel dafür ist der soeben lancierte KI-Campus (**Seite 22 ↗**). Dieser Trend ist erfreulich, denn nur wer KI versteht, kann die Zukunft aktiv mitprägen.

Doch was für eine Zukunft wollen wir? In der Schweiz und in Deutschland ist die Sehnsucht gross, dass das «Menschsein» sowie nicht-digitale Erfahrungen beibehalten werden. Viele Menschen sehen die Vorteile der Digitalisierung. Gleichzeitig fürchten sie sich – auch vor dem Kontrollverlust an den eigenen Daten (**Seite 18 ↗**). Dies ist mit ein Grund, weshalb das etablierte Institut Gartner überzeugt ist, dass sich das Marketing grundlegend verändern wird (**Seite 13 ↗**).

Unternehmen, die sich auf neue Themen wie Privatsphäre, Datensicherheit oder den Umgang mit Desinformationen konzentrieren, werden auch für Investoren immer interessanter. Denn Innovationen entstehen da, wo Unternehmen auf neue Erwartungen reagieren müssen oder wo neue Gesetze erlassen werden. Deshalb ist es für Fidelity International ein logischer Schritt, dass klassische ESG-Kriterien schon bald um Aspekte einer digitalen Ethik ergänzt werden (**Seite 17 ↗**).

Damit wir die Chancen von KI ausschöpfen können, braucht es eine Kombination aus Selbstregulierung der Wirtschaft und staatlicher Regulierung. Dazu finden Sie in dieser Ausgabe vier aktuelle Geschichten (**Fokusthema ↗**). Eine Studie zeigt beispielsweise, dass Unternehmen andere Prioritäten setzen als die politischen Entscheidungsträger. Sie konzentrieren sich stark auf bestehende Vorschriften und noch wenig auf neu auftretende Probleme bei KI. Doch genau diese Konfliktfelder werden die Regulierung von KI prägen, etwa Erklärbarkeit, Fairness und Nichtdiskriminierung.

Wer sich diesen Themen bereits heute annimmt, als Bestandteil der Digitalstrategie, kann den Anpassungsaufwand reduzieren und sich als vertrauenswürdiges Unternehmen positionieren.



In dieser Ausgabe

Fokusthema

- Unternehmen unterschätzen die Prioritäten bei der KI-Regulierung
- Selbstregulierung als notwendige Ergänzung zur Regulierung von KI
- Wie Salesforce die Ethik bei seinen KI-Anwendungen integriert
- Mit dem «Data Fairness Label» können Unternehmen Vertrauen schaffen

Wissenschaft

- Corona-Krise: Ethische Fallstricke bei digitalen Tools
- Eine neue Methode soll Vertrauen in selbstlernende KIs schaffen

Wirtschaft

- Voraussetzungen für die Akzeptanz von KI in der Finanzindustrie
- Konsumenten-Datenethik wird das Marketing grundlegend verändern
- Die Nutzung von Cognitive Biases in der Marketingkommunikation
- Erst einzelne virtuelle Influencer erzielen hohe Werbeeinnahmen
- Neues Toolkit fördert die Entwicklung ethischer Produkte
- Digitale Ethik ergänzt klassische ESG-Kriterien bei Investitionen

Gesellschaft

- Menschen sehnen sich nach Menschlichkeit und Datensouveränität
- Experten setzen sich für realistische Narrative zu KI ein
- 10 Gebote für einen ethischen Umgang in der digitalen Gesellschaft
- Praktische und unterhaltsame Einblicke in die Geheimnisse von KI
- KI-Campus: Eine Lernplattform mit kostenlosen Kursen zu KI

Staat

- Neuseeland führt Standards für den Einsatz von Algorithmen ein
- Fehlende Transparenz bei Schweizer Polizei- und Justizbehörden
- Wenige Staaten setzen sich für ein Verbot von Killerrobotern ein



Das müssen Sie wissen

- **Der ethische Umgang mit Kundendaten wird das Marketing in den nächsten Jahren entscheidend prägen, denn der Vertrauensverlust der Kundinnen und Kunden wird weiter zunehmen.** Unternehmen können Vertrauen schaffen, indem sie ihr Engagement aktiv thematisieren.
- **Cognitive Biases geben Einblick in den Entscheidungsprozess von Menschen, denn dieser ist oft irrational. Dabei ist es wichtig, den Grat zwischen Manipulation und Inspiration zu verstehen.** Das Cognitive Biases Wissen sollte verantwortungsvoll im Marketing-Funnel eingesetzt werden.
- **Die Mehrzahl der Menschen fürchten sich vor dem Kontrollverlust über die persönlichen Daten und sie sehnen sich nach Menschlichkeit, Gerechtigkeit und Solidarität.** Unternehmen sollten sicherstellen, dass ihre digitalen Lösungen eine echte Unterstützung bieten und keine Trade-Offs bezüglich Datensouveränität beinhalten.
- **Nur wer weiss, wie Künstliche Intelligenz funktioniert, kann die Zukunft beruflich und privat mitgestalten.** Deshalb bieten Regierungen, Bildungsinstitutionen und Unternehmen auch im deutschsprachigen Raum vermehrt hochstehende, oft kostenlose Lernangebote zu KI an.
- **Viele Unternehmen sind sich nicht bewusst, dass ethische Prinzipien wie Erklärbarkeit, Fairness und Nichtdiskriminierung die Regulierung von KI prägen werden.** Der Anpassungsaufwand lässt sich frühzeitig reduzieren, indem bereits heute in eine vertrauenswürdige KI investiert wird, als Teil der Digitalstrategie. Ausdruck davon sind Ethik-Grundsätze, die Auseinandersetzung mit neuen Risiken sowie Investitionen in die Mitarbeitenden.
- **Investoren werden die klassischen ESG-Kriterien (Environmental, Social, Governance) immer mehr um ethische Aspekte der Digitalisierung ergänzen.** Denn neue Gesetze und Erwartungen an Unternehmen führen zu Innovationen in Bereichen wie Privatsphäre und Datensicherheit.



6

Fokusthema

10

Wissenschaft

12

Wirtschaft

18

Gesellschaft

23

Staat



Unternehmen unterschätzen die Prioritäten bei der KI-Regulierung

Welche ethischen Grundprinzipien sind für die Regulierung von Künstlicher Intelligenz besonders wichtig? Eine globale Studie zeigt, dass sich Unternehmen auf Grundsätze konzentrieren, die von existierenden Vorschriften wie der Datenschutz-Grundverordnung (DSGVO) priorisiert werden, und nicht auf neu auftretende Probleme, die im Zeitalter von KI kritisch werden. Doch genau diese Prinzipien werden die Regulierung von KI prägen, etwa Erklärbarkeit, Fairness und Nichtdiskriminierung. Das KI-Whitepaper der EU-Kommission bekräftigt diese Prinzipien, denn sie widerspiegeln die Besorgnis der Öffentlichkeit. Nur wer die ethischen Grundsätze der politischen Entscheidungsträger versteht, kann den Anpassungsaufwand frühzeitig reduzieren. Für Ernst & Young ist klar: → Eine vertrauenswürdige KI wird zu einem zentralen Merkmal jeder Digitalstrategie. Unternehmen sollten ihre ethischen Grundsätze publizieren, sich mit ethischen Risiken auseinandersetzen und in die Mitarbeitenden investieren.

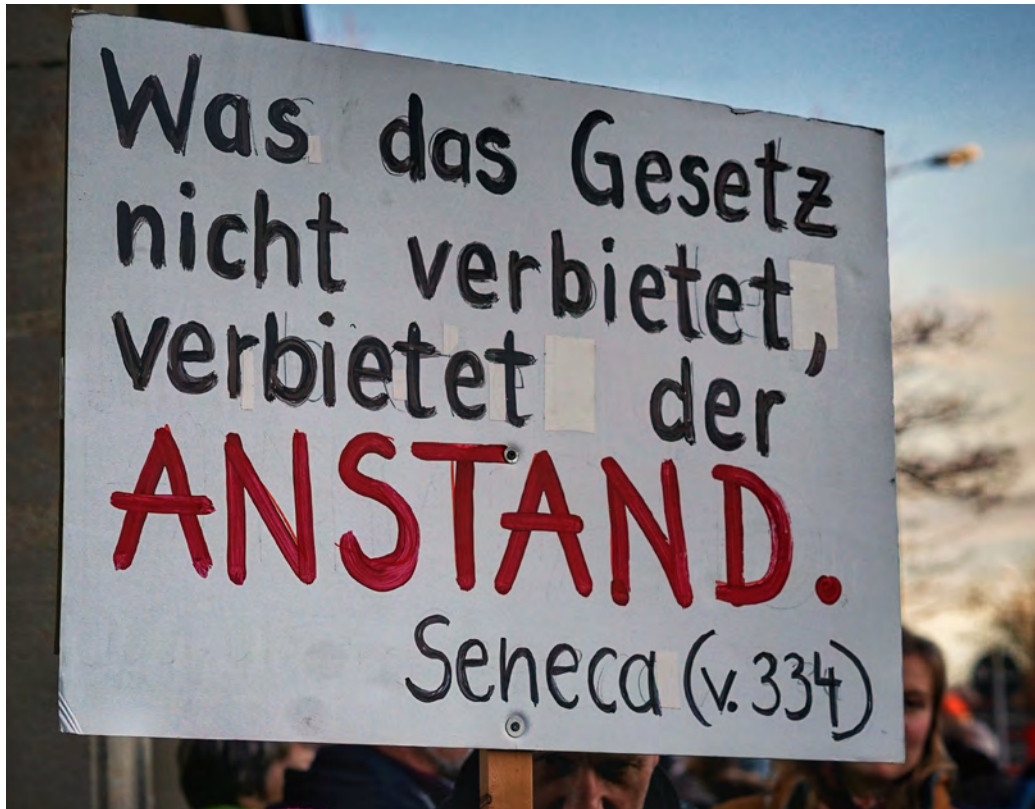


FOTO: ROBERT ANASCH / UNSPLASH

Selbstregulierung als notwendige Ergänzung zur Regulierung von KI

Selbst die CEOs renommierter Unternehmen wie Google und Microsoft stehen einer Regulierung von Künstlicher Intelligenz positiv gegenüber. Doch dies genügt nicht. Ebenso wichtig sind berufliche Richtlinien, private Standards, Best-Practices und Verhaltenskodizes. Weshalb? Zum einen stösst die Regulierung angesichts des hohen Tempos beim technologischen Fortschritt von KI an klare Grenzen. Zum anderen gibt es eine Vielfalt an KI-Anwendungen in unterschiedlichsten Branchen. Dabei konzentriert sich der Regulator auf besonders kritische Bereiche, zum Beispiel auf autonome Fahrzeuge. In vielen Bereichen genügt es, wenn Unternehmen innerhalb ihrer KI-Governance ethische Prinzipien sowie globale Produktstandards berücksichtigen. Diese Selbstregulierung hat den Vorteil, dass sie Innovationen fördert und Anpassungen schnell möglich sind. → Damit die Chancen von KI ausgeschöpft werden können, braucht es eine Kombination aus Selbstregulierung der Wirtschaft und staatlicher Regulierung.



Wie Salesforce die Ethik bei seinen KI-Anwendungen integriert

Salesforce positioniert sich als verantwortungsvolles Tech-Unternehmen, basierend auf seinen «Trusted AI Principles». In diesem Kontext wurde auch die Stelle «Architect of Ethical AI Practice» geschaffen, welche Kathy Baxter bekleidet. Wie werden diese Prinzipien umgesetzt? Zum Beispiel müssen alle Mitarbeitenden bei jedem Produkt verstehen, wie diese für unterschiedliche Nutzertypen interpretierbar gemacht werden können. Viel wird auch in die Ethics-by-Design-Denkweise investiert. Sie bildet die Basis für alle KI-Lösungen. Die Produkte enthalten auch eine Reihe von Funktionen, um die Kunden bei ethischen Entscheidungen zu unterstützen. Ein weiterer Ethik-Checkpoint ist das Data Science Review Board, das Best-Practices zur Datenqualität und Modellbildung fördert und durchsetzt. Und es gibt einen Prüfungsausschuss, damit transparent ist, wie die verwendeten Daten gesammelt werden. → Das Beispiel Salesforce zeigt, wie es gelingen kann, ethische Prinzipien in der Praxis zu verankern.



Mit dem «Data Fairness Label» können Unternehmen Vertrauen schaffen

Wie lässt sich sicherstellen, dass die Nutzung von Daten und die Anwendung von datengetriebenen Modellen transparent, nachvollziehbar und in diesem Sinne fair gestaltet wird? Hierfür hat der Verband Swiss Insights das Label «Data Fairness» geschaffen. Unternehmen steht ein Online-Tool zur Verfügung, in dem sie die wichtigsten Merkmale der verwendeten Datensätze (Data Sheets) oder die Art der verwendeten Datenmodelle inklusive Verwendungszweck (Model Cards) erfassen. Dies geschieht anhand zweier Fragebogen im Sinne einer Selbstdeklaration, die mindestens jährlich aktualisiert wird. Unternehmen, welche die Vorgaben des Labels bezüglich Transparenz und Nachvollziehbarkeit erfüllen und Mitglied von Swiss Insights sind, können das Label beim Verband beantragen. Es verpflichtet die Unternehmen zum verantwortungsvollen Umgang mit Daten. → Mit einem Label wie «Data Fairness» können Unternehmen ihr Engagement gegenüber Kunden und Stakeholdern sichtbar machen und so Vertrauen schaffen.

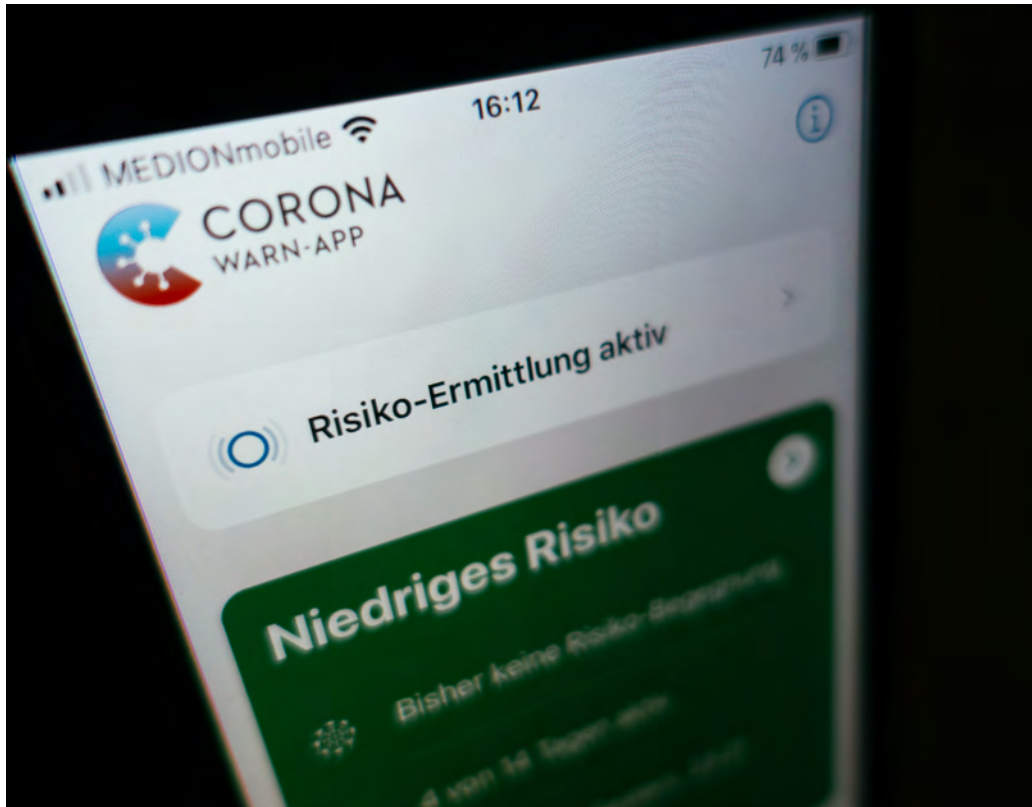


FOTO: KAI PILGER / UNSPLASH

Corona-Krise: Ethische Fallstricke bei digitalen Tools

Viele Länder setzen bei der Bewältigung der Corona-Krise digitale Hilfsmittel ein. Forschende der ETH Zürich haben untersucht, welche ethischen Aspekte dabei zu berücksichtigen sind, und dafür vier Kategorien von Tools analysiert: Contact-Tracing-Apps, Programme zur Identifizierung einer möglichen Erkrankung, Apps zur Kontrolle der Quarantänevorschriften sowie Flow-Models, die beispielsweise Google für Mobilitätsberichte verwendet. Bei all diesen Hilfsmitteln muss der öffentliche Nutzen sorgfältig gegen mögliche Nachteile abgewogen werden. Dazu gehören auch unerwünschte Nebeneffekte wie staatliche Überwachung, fehlender Zugang oder Diskriminierung. Die Forschenden zeigen Schritt für Schritt auf, welche Fragen zu berücksichtigen sind. → Die Entwicklung digitaler Gesundheits-Tools sollte auf sechs ethischen Prinzipien beruhen: Respektieren von Autonomie und Privatsphäre, fördern der Gesundheitsvorsorge und Solidarität, verhindern neuer Infektionen und böswilliger Handlungen.



FOTO: JOSHUA HOEHNE / UNSPLASH

Eine neue Methode soll Vertrauen in selbstlernende KIs schaffen

Eine KI wurde mit Pferde-Fotos trainiert, um Pferde zu erkennen. Irgendwann war sie zwar in der Lage, das Tier auch auf unbekanntem Aufnahmen zu identifizieren. Doch nun stellte sich heraus, dass es bei den verwendeten Trainingsdaten in einer Ecke immer eine Information zum Copyright mit einem Link auf eine Pferde-Website gab. Diese Links bildeten unbeabsichtigt den Lerninhalt der KI, deshalb handelt es sich um «Confounders», also Störfaktoren. Ein Forscherteam der Universität Darmstadt entwickelte nun eine Lösung: erläuterndes interaktives Lernen (explanatory interactive learning XIL). Dabei werden Fachleute in den Lernprozess eingebunden. Die KI informiert darüber, wie sie lernt und wie sie daraus Voraussagen ableitet. Ist die Erklärung der Voraussage falsch, füttern die Experten die KI mit anderen Daten, beispielsweise Pferdefotos mit Copyright-Links auf beliebige Websites. → Die neue Methode XIL kann helfen, das Vertrauen in maschinelle Lernverfahren zu stärken.



Voraussetzungen für die Akzeptanz von KI in der Finanzindustrie

Die erfolgreiche Einführung neuer Technologien in einem Unternehmen hängt in hohem Masse von ihrer Akzeptanz ab. Was ist bei der Nutzung von Künstlicher Intelligenz zu beachten? Diana Bengel hat ein Akzeptanzmodell für Künstliche Intelligenz in der Deutschen Finanzindustrie entwickelt. Dabei wurden die Erkenntnisse von 17 Expertinnen und Experten aus Unternehmen wie AXA, Deutsche Bank, IBM und Talanx einbezogen. Zahlreiche Faktoren beeinflussen die Akzeptanz von KI positiv oder negativ und sie beeinflussen sich gegenseitig. Sie lassen sich in fünf Bereiche unterteilen:

- Organisation, z.B. Unternehmenskultur, transparente Kommunikation
 - Individuen, z.B. Angst vor Arbeitsplatzverlust, Mindset intern
 - Finanzen, z.B. Kosteneinsparungen, Produktivitätsgewinne
 - Technologie, z.B. Schulungsaufwand, Nachvollziehbarkeit von KI
 - Gesellschaft, z.B. Datenschutz, öffentliche Wahrnehmung
- Die Akzeptanzfaktoren können Unternehmen als Checkliste beim Einsatz von Künstlicher Intelligenz dienen.



FOTO: DARIA NEPRIAKHINA / UNSPLASH

Konsumenten-Datenethik wird das Marketing grundlegend verändern

Das renommierte Institut Gartner geht davon aus, dass Datenethik das Marketing in den nächsten 5 bis 10 Jahren entscheidend prägen wird. Die Geschäftspraktiken müssen mit ethischen Richtlinien in Einklang gebracht werden, welche die Werte des Unternehmens widerspiegeln. Deshalb wird der ethische Umgang mit Kundendaten erstmals im Hype Cycle for Digital Marketing aufgeführt. Davon betroffen sind vor allem die Bereiche Mobiles Marketing und gezielte Werbung, Gestaltung von Marketingkampagnen, Kundensegmentierung, passive Datenerfassung sowie Kundendienst- und Treueprogramme. Treiber dafür sind zwei Gründe: Zum einen kontrollieren wenige Tech-Giganten eine riesige Menge an Kundendaten. Zum anderen haben Kundinnen und Kunden grösste Bedenken, wie ihre Daten gesammelt und genutzt werden, was weiter zunehmen wird. → Marketingverantwortliche können das Vertrauen der Kundinnen und Kunden zurückgewinnen, indem sie den verantwortungsvollen Umgang mit Kundendaten aktiv thematisieren.





Die Nutzung von Cognitive Biases in der Marketingkommunikation

Basierend auf Daten können Unternehmen zwar nachvollziehen, wie sich Konsumentinnen und Konsumenten verhalten. Was Daten nicht abbilden ist der Entscheidungsprozess. Dieser ist oft auch irrational. Die kognitiven Verzerrungen bzw. Vorurteile (Cognitive Biases) von Menschen liefern Erklärungen dafür, wie Informationen selektiv aufgenommen, evaluiert und verarbeitet werden. Von Manipulation spricht man, wenn Cognitive Biases auf solch eine Art eingesetzt werden, dass sie beispielsweise Unwahrheiten vermitteln. Etwa wenn sie suggerieren, dass es sich um ein gesundes Produkt handelt, obwohl dies nicht der Fall ist. Von Inspiration kann gesprochen werden, wenn tatsächliche Vorteile wie Inhaltsstoffe so kommuniziert werden, dass dem Konsumenten eine «bessere» Entscheidungsfindung ermöglicht wird. → Die Publikation «Der entschlüsselte Konsument» kann Unternehmen helfen, die Wirkung von Cognitive Biases zu verstehen und dieses Wissen verantwortungsvoll im Marketing-Funnel einzusetzen.



Erst einzelne virtuelle Influencer erzielen hohe Werbeeinnahmen

Virtuelle Influencer sind fiktive, computergenerierte «Menschen», die die realistischen Eigenschaften, Merkmale und Persönlichkeiten von Menschen haben. Wie viel verdienen die 20 erfolgreichsten computergenerierten Influencer auf Instagram? Zu diesem Zweck hat der Marktplatz OnBuy die Werbeeinnahmen der bereits publizierten Posts bis Ende Jahr hochgerechnet. An der Spitze steht ganz klar «Lil Miquela» (@lilmiquela), die sich auch als Sängerin ⁷ betätigt. Sie hat 2,6 Millionen Follower und verdient pro Post 7211 Euro. Das entspricht knapp zehn Millionen Euro bis Ende Jahr. Platz 2 geht an «nonoouri» (@nonoouri) mit 363 000 Follower und 2,5 Millionen Euro Umsatz. Auf Platz 3 ist «Imma» (@imma.gram) mit 245 000 Follower und 550 000 Euro Umsatz bis Ende Jahr. → Erst ganz wenige virtuelle Influencer erzielen hohe Werbeeinnahmen, doch wird mit einer gewissen Zunahme gerechnet. Wichtig ist, dass künstlich erzeugte Influencer immer als solche deklariert sind, um Follower nicht zu täuschen.

Neues Toolkit fördert die Entwicklung ethischer Produkte

Seit einigen Jahren stehen verschiedene Toolkits zur Verfügung. Sie sollen Unternehmen motivieren, über die Auswirkungen ihres Tuns und unerwünschte Nebenwirkungen frühzeitig nachzudenken. Einige Tech-Unternehmen haben dafür Stellen geschaffen, zum Beispiel Salesforce mit seinem «Office of Ethical and Humane Use». Und immer mehr Mitarbeitende hinterfragen gewisse Praktiken und Produkte ihrer Arbeitgeber. Für sie hat das Omidyar Network das Toolkit «Ethical Explorer Pack» entwickelt. Es hilft, Themen im Team zur Sprache zu bringen, Lösungen für Problem zu finden und rote Linien zu erkennen. Dazu gehört zum Beispiel ein Leitfaden zur Navigation durch acht Risikozonen: Überwachung, Desinformation, Ausschluss, algorithmische Verzerrung, Sucht, Datenkontrolle, schlechte Akteure und Machtkonzentration. Ausserdem gibt es ein Kartenspiel mit konkreten Fragen. → Das Toolkit «Ethical Explorer Pack» hilft Mitarbeitenden, ethische Produkte zu entwickeln und mögliche Risiken zu vermeiden.





FOTO: MCCUTCHEON / UNSPLASH

Digitale Ethik ergänzt klassische ESG-Kriterien bei Investitionen

Das Verhalten der Menschen ändert sich gerade fundamental. Treiber sind neue Technologien, parallel dazu wandeln sich auch die Bedürfnisse der Menschen und die Arbeitswelt. Befeuert wird dieser Wandel durch die Corona-Krise, davon ist Fidelity überzeugt, einer der grössten Vermögensverwalter der Welt. Diese Transformation stellt gleichzeitig die ethischen Grundsätze in Unternehmen auf den Prüfstand. Zu den Themen Umweltschutz, sozialem Wohlergehen und einer nachhaltigen Unternehmenssteuerung tritt die Diskussion um Datenschutz und Digitale Ethik. Gerade Plattformen müssen durch neue Gesetze und Erwartungen navigieren, sich mit Desinformation und Hassrede sowie Steuerfragen, Kartellrecht und Regulierungen auseinandersetzen. → Innovationen entstehen durch Unternehmen, die sich auf Themen wie Privatsphäre, Datensicherheit oder den Umgang mit Desinformationen verstehen. Deshalb werden Investoren die klassischen ESG-Kriterien immer mehr um Aspekte einer digitalen Ethik ergänzen.





FOTO: H HEYERLEIN / UNSPLASH

Menschen sehnen sich nach Menschlichkeit und Datensouveränität

Die Studie «Values & Visions 2030» von GIM gibt spannende Einblicke in bevorstehende Werteverstärkungen in einzelnen Ländern. Menschen sind sich bewusst, dass sich das Leben immer mehr in die digitalen Sphären verlagert. Das macht vieles bequemer, schneller. Dennoch steht eine grosse Mehrheit dieser Entwicklung skeptisch gegenüber. Daran hat die Corona-Krise nichts geändert. Viele Menschen in der Schweiz und in Deutschland fürchten sich vor dem Kontrollverlust über die persönlichen Daten. Sie prägen immer mehr Entscheide und das eigene Handeln. Gross ist auch die Sehnsucht, dass das «Menschsein» sowie nicht-digitale Erfahrungen in einer immer stärker technisch dominierten Welt beibehalten werden. Werte wie Gerechtigkeit und Solidarität werden ebenfalls als erwünscht angesehen, im Wissen, dass diese unter Druck geraten. → Unternehmen sollten digitale Anwendungen so konzipieren, dass sie eine echte Unterstützung bieten und keine Trade-Offs bezüglich Datensouveränität beinhalten.



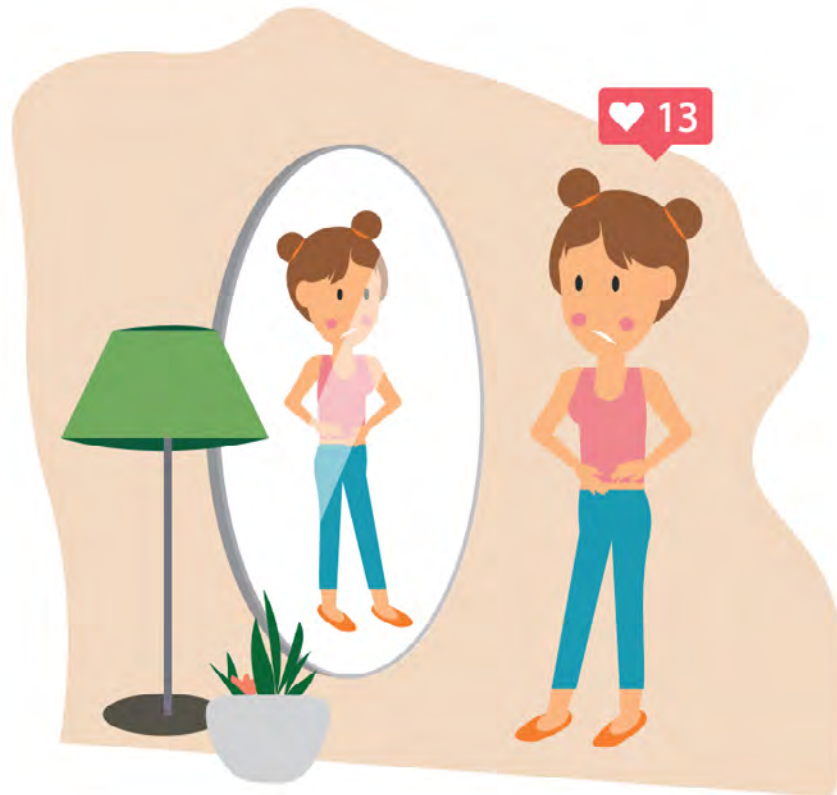


FOTO: TOA HEFTIBA / UNSPLASH

Experten setzen sich für realistische Narrative zu KI ein

Wie werden sich KI-Technologien bis 2035 auf den beruflichen und sozialen Alltag auswirken? Eine aktuelle Studie gibt die Stimmen von über 500 Expertinnen und Experten in Deutschland wieder, die sich für realistischere KI-Narrative einsetzen. Basierend auf 53 Thesen erwarten sie die grössten Veränderungen in den Bereichen Mobilität, Gesundheit und Pflege, Medienkonsum sowie Arbeit. Zum Beispiel erwarten 64 Prozent der Befragten, dass Unternehmen in den nächsten 10 Jahren die menschlichen Kompetenzen in der Interaktion zwischen Menschen und Maschine in den Vordergrund stellen. 76 Prozent sind überzeugt, dass KI vor allem einfache Arbeiten übernehmen wird. Das Management wird hauptsächlich für Personalführung, Motivation und Kreativleistungen verantwortlich sein. → Die Studie zeigt, dass der Kampf um kreative Talente deutlich zunehmen wird, und dass sich die Diskussion rund um KI von der rein technischen Betrachtungsweise hin zu einer Nutzer- und Gemeinwohlorientierung wandelt.





10 Gebote für einen ethischen Umgang in der digitalen Gesellschaft

Wie wir uns in der digitalen Welt verhalten und mit Konflikten umgehen, ist Ausdruck unserer ethischen Haltung. Die Medienwissenschaftlerin Petra Grimm erarbeitete mit ihren Studierenden die «10 Gebote der digitalen Ethik». Entstanden ist ein praktischer Wertekompass der digitalen Ethik gegen Hass und Cybermobbing. Zu jedem Gebot gehört eine Geschichte, mit der sich nicht nur Jugendliche identifizieren können. Sie regen an, das eigene Wertesystem zu reflektieren. Zum Beispiel «Miss deinen Wert nicht an Likes und Posts» oder «Glaube nicht alles, was du online siehst, und informiere dich aus verschiedenen Quellen». Es ist an der Zeit, sich darüber zu verständigen, wie wir digital miteinander umgehen. Petra Grimm ist überzeugt, dass vor allem Bewertungskompetenz als Teil der Medienkompetenz besser geschult werden muss, um Falschmeldungen zu erkennen. → Die 10 Gebote schaffen ein Wertesystem, das zeigt, wie ein gutes, gelingendes Leben in der digitalen Gesellschaft aussehen sollte.



Praktische und unterhaltsame Einblicke in die Geheimnisse von KI

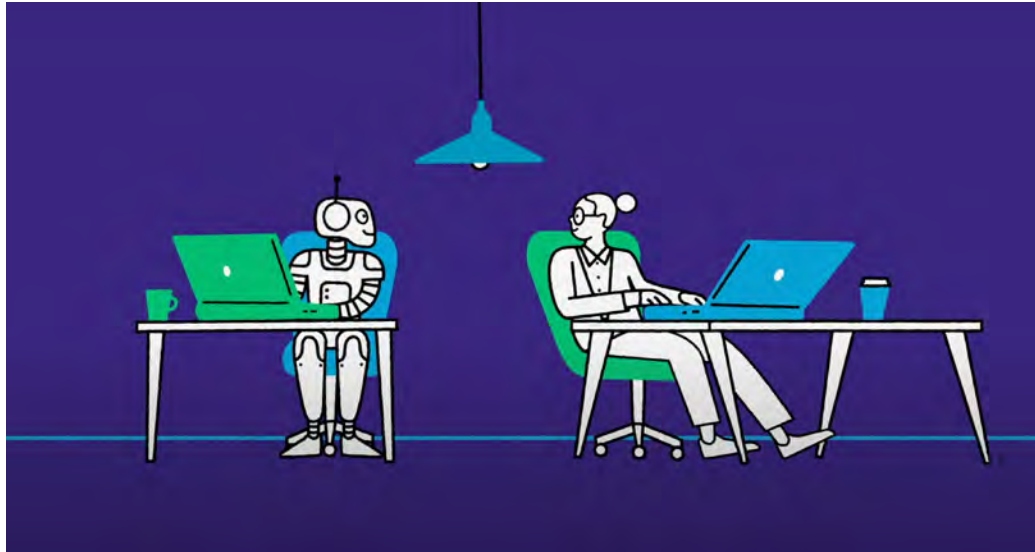
Kenza Ait Si Abbou sprach in vielen Workshops mit Kindern und Erwachsenen über ihre Fragen zu Künstlicher Intelligenz. Dabei stellte sie fest, dass oft das Wissen fehlt. Gepaart mit dramatisierten Extremfällen, die publiziert werden, löst dies Verwirrung und Angst aus. Kenza Ait Si Abbou arbeitet bei der Deutschen Telekom und erklärt in ihrem Buch die Geheimnisse der Technik. Gut verständlich erzählt sie anhand ihres Alltags von vielen Praxisbeispielen. Dabei lernen wir, wie Maschinen programmiert werden, welche Sprache sie sprechen, was Modelle und maschinelles Lernen sind. Auch Risiken werden diskutiert, denn die schlechten Beispiele aus der Vergangenheit haben uns die Augen geöffnet. «Wenn wir dafür sorgen, dass Algorithmen Ethik und Moral verstehen, und ihnen klar definierte Kriterien vorlegen, dann können diese weniger diskriminierend sein als wir Menschen», so die Autorin. → Abbou zeigt, wie wir uns alle einbringen können, um mit KI eine wertvolle Zukunft zu bauen.



Trendradar Digitale Ethik
Ausgabe August 2020

AKTEUR
WEBSITE
KATEGORIEN

Kenza Ait Si Abbou
Keine Panik, ist nur Technik! ↗
#digitalkompetenz #ki #zukunft



KI-Campus: Eine Lernplattform mit kostenlosen Kursen zu KI

Auf der Lernplattform für Künstliche Intelligenz «KI-Campus» finden Interessierte kostenlose, digitale Lernangebote. Diese reichen von Podcasts über Videos und Online-Kurse bis hin zum «Micro-Degree». Die Angebote werden von Wissenschaftlerinnen und Wissenschaftlern sowie von Fachpersonen des Deutschen Forschungszentrums für Künstliche Intelligenz (DFKI) entwickelt. Darüber hinaus kuratiert und empfiehlt die Lernplattform deutsch- und englischsprachige Lernangebote weiterer Plattformen. Im Moment sind erst wenige Kurse verfügbar, da die Plattform aufgrund der Corona-Krise früher als geplant lanciert wurde. Angekündigt sind beispielsweise Kurse zur Einführung in KI, zu KI & Ethik oder auch ein Kurs über Methoden zu «explainable AI» (XAI). → Bildungsangebote leisten einen wichtigen Beitrag daran, dass die KI-Kompetenzen des Einzelnen und innerhalb der Gesellschaft gefördert werden. Denn nur wer weiss, wie KI funktioniert, kann die Zukunft beruflich und privat mitgestalten.



FOTO: PAT HO / UNSPLASH

Neuseeland führt Standards für den Einsatz von Algorithmen ein

Neuseeland erlässt als erstes Land Standards für den Einsatz von Algorithmen beim Staat. Von Verkehrsampeln bis zur Gesichtserkennung will die Regierung mit der Charta Transparenz schaffen und nur Algorithmen ohne Designfehler oder Vorurteile einsetzen. Die Standards verpflichten Behörden, in verständlicher Sprache transparent zu machen, wie eine Entscheidungsfindung verläuft. Auch Informationen über die angewandten Methoden, die Datenspeicherung und -weitergabe sollen publiziert werden. Insbesondere werden Verzerrungen aufgrund unterschiedlicher Weltanschauungen der neuseeländischen Minderheiten identifiziert und Bedenken angehört. Mehr als 20 Behörden unterzeichneten die Charta bereits, darunter die Polizei und die Einwanderungsbehörde. Anstelle einer unabhängigen Regulierungsbehörde wurde eine «Data Ethics Advisory Group» eingesetzt. → Neuseeland erfüllt mit seiner «Algorithm Charter» die spezielle Sorgfaltspflicht, die Staaten beim Einsatz von Algorithmen haben.



FOTO: WENDY ALVAREZ / UNSPLASH

Fehlende Transparenz bei Schweizer Polizei- und Justizbehörden

Die Schweizer Polizei- und Justizbehörden setzen mehr als 20 automatisierte Systeme ein, um unangemessenes Verhalten einzuschätzen oder vorherzusagen. Im Kampf gegen häusliche Gewalt setzen die sechs Kantone Glarus, Luzern, Schaffhausen, Solothurn, Thurgau und Zürich die Software «Dyrias» ein. Basierend auf 39 Fragen wird die Wahrscheinlichkeit berechnet, dass eine Person den Intimpartner angreift. Ein weiteres Beispiel ist «ROS», das die Behörden in der Deutschschweiz im Umgang mit Straffälligen einsetzen. Dieses Tool kennzeichnet Gefangene als A (kein Rückfallrisiko), B (neue Straftat wahrscheinlich) oder C (Gewaltverbrechen wahrscheinlich). Die Kantone in der Romandie und im Tessin arbeiten an einer Alternative, die 2022 eingeführt werden soll. → Viele bezweifeln, dass sich das individuelle Verhalten eines Menschen basierend auf Daten aus der Vergangenheit vorhersehen lässt. Der Einsatz von KI-Tools ist wenig transparent und sie suggerieren eine vermeintliche Objektivität.



Wenige Staaten setzen sich für ein Verbot von Killerrobotern ein

Waffensysteme, die ohne menschliche Kontrolle Ziele auswählen und bekämpfen, müssen verhindert werden. Denn im Ernstfall können Killerroboter nicht zwischen Kämpfenden und Zivilisten unterscheiden. Seit 2013 setzt sich deshalb Human Rights Watch für ein internationales Verbot von Killerrobotern ein. In ihrem neusten Bericht fasst sie die Stellungnahme von 97 Staaten zusammen. Alle Länder haben die ethische, moralische und rechtliche Verpflichtung, die Menschheit vor dieser gefährlichen Entwicklung zu schützen. 30 Staaten, darunter Österreich als einziges europäisches Land, sprechen sich für ein internationales Verbot aus. Doch viele Staaten misstrauen den anderen und erwägen zwar ein Verbot, wollen aber weiter daran forschen. → Human Rights Watch ruft die Länder dazu auf, Gesetze und Richtlinien für eine sinnvolle menschliche Kontrolle über die Anwendung von Gewalt zu erarbeiten. Verbote braucht es bei der Entwicklung, Herstellung und dem Einsatz autonomer Waffen.

Centre for Digital Responsibility (CDR)

Der Think Tank für Digitale Ethik
Berggasse 14
CH-5742 Kölliken

digitalresponsibility.ch ↗
[@centreforDR](https://twitter.com/centreforDR) ↗

Für Fragen und Anliegen

Cornelia Diethelm
cornelia.diethelm@digitalresponsibility.ch ↗
Tel. +41 79 249 04 77

Der Trendradar enthält kuratierte Informationen. Die verwendeten Bilder dienen einzig der Forschung und Illustration. Sie sind nicht zur Publikation bestimmt.

Trendradar Digitale Ethik — Ausgabe August 2020

