

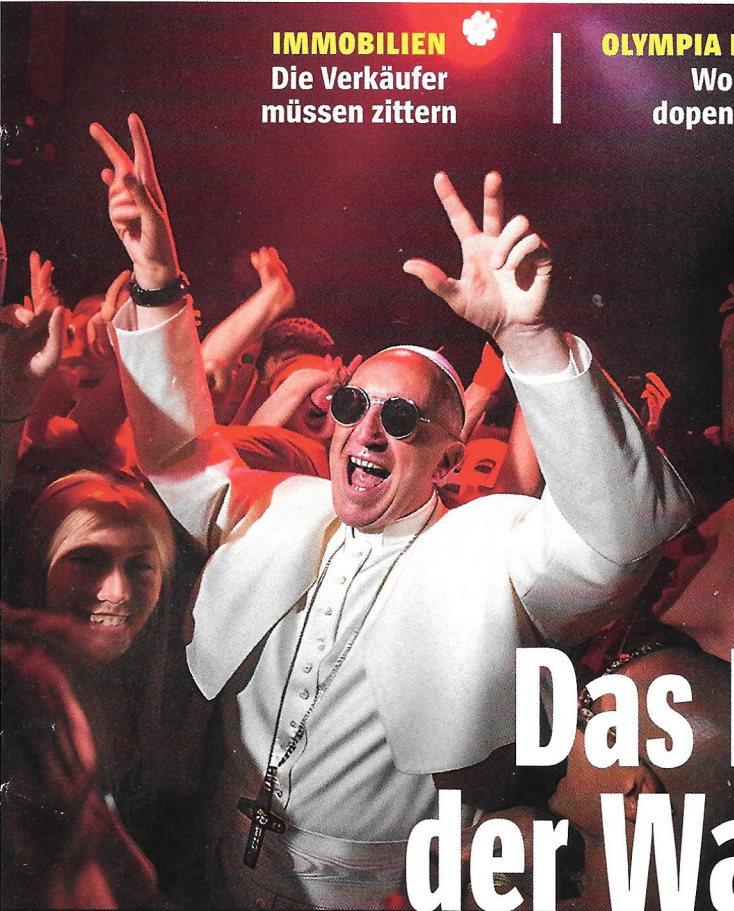
EVAL-INFO-SYSTEM CE/15

TECH 70 - TB 04  
dazu Interview:  
TECH 70 - TB 05

**IMMOBILIEN**  
Die Verkäufer  
müssen zittern

**OLYMPIA IN EXTREM**  
Wo alle  
dopen dürfen

Die neue Karriere des  
Hollywoodstars



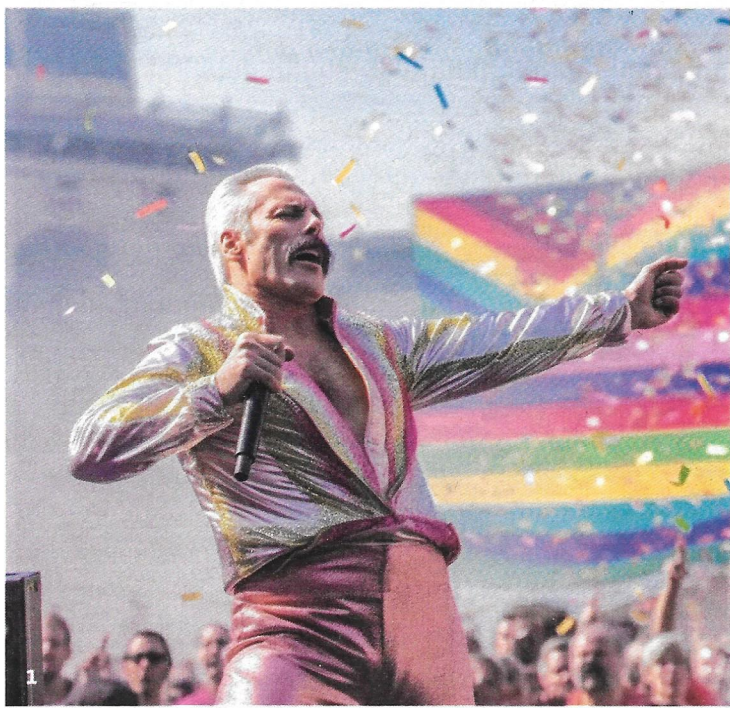
# Das Ende der Wahrheit

Politik, Liebe, Musik: Wie uns  
**künstliche Intelligenz**  
täuschen kann



Ungarn Ft 3790,-  
Printed in Germany  
Spanien € 770  
Schweden € 750  
Slowakei € 743  
Slovakei € 743  
Portugal (cont.) € 710  
Kroatien € 9,10 / Kč 195,-  
Norwegen NOK 125,-  
Griechenland € 790  
Italien € 8,20  
Finnland € 9,50  
Frankreich € 770  
Dänemark dkr 74,95

# Die große Fake-Maschine



KI generiert: Reddit / Legitimat

1



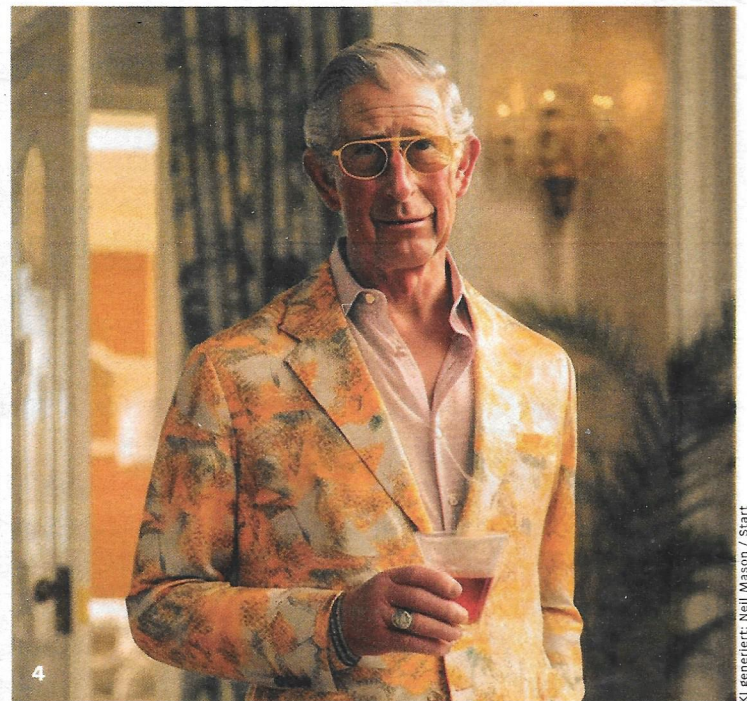
KI generiert: Pete Neomakis / Children of Legend

2



KI generiert: Reddit / Arctic\_Chillean

3



KI generiert: Neil Mason / Start

4

1 | Erfundene Pride-Parade mit Freddie Mercury auf Reddit

2 | Sänger Elton John als Kind von US-Grafikkünstler Pete Neomakis

3 | Fiktive Erdbebenopfer eines Reddit-Users

4 | Partykönig Charles III. des australischen Kreativen Neil Mason

**TECHNOLOGIE** Ein Papst in Daunenjacke, Kriegsreden von Präsidenten, Pophits von Verstorbenen:  
Künstliche Intelligenz erschafft aus riesigen Datenmengen neue Realitäten.  
Was passiert, wenn wir sie nicht mehr von der echten Welt unterscheiden können?



KI generiert: Eliot Higgins

**Trump-Fake des britischen Journalisten  
Eliot Higgins**

**E**igentlich ist Boris Eldagsen ein Mann mit Glatze. Aber wenn er sich vor die Fototapete in seinem Wohnzimmer setzt, sieht es für einen kurzen Moment aus, als würden aus seinem

kahlen Kopf Pflanzen sprießen. Eine Illusion, klar, völlig analog. Das Motiv der Tapete ist 120 Jahre alt, es zeigt wild blühende Moose und erschien in einer Sammlung namens »Kunstformen der Natur«. Der Zoologe Ernst Haeckel hat die Skizzen einst gezeichnet, er wollte der Menschheit »Zugang zu den wunderbaren Schätzen der Schönheit öffnen«. Alles Gute, was die Natur zu bieten hat.

Eldagsen lacht ein kurzes, raues Lachen, dann öffnet er seinen Laptop und zeigt das Bild, das ihn selbst berühmt gemacht hat. Zwei ernste Frauen in Schwarz-Weiß-Optik sind darauf zu sehen, die eine versteckt sich schüchtern hinter der anderen. Über dem Paar hängen leuchtende Stromkabel. Der Berliner Fotograf nannte es »The Electrician« – die Elektrikerin. Und tatsächlich ist der Name Programm: Hinter diesem Bild steckt die pure Technik.

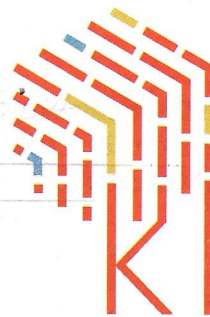
Er benötigte keine Kamera für seine Fotografie, er brauchte auch keine Frauen vor der Linse. »The Electrician«, in all seiner spröden Schönheit, ist das Werk von Prozessoren und Stromkreisen, die Schöpfung einer künstlichen Intelligenz (KI).

Der Fotograf experimentiert seit rund einem Jahr mit der sogenannten generativen KI, sie ersetzt für ihn das, was seinen Beruf einst ausmachte: Apparat und Optik, Blitz und Blende, mitunter sogar das Motiv. Stattdessen tippt er präzise Anweisungen, sogenannte Prompts, in seinen Computer und lässt die Software in mehreren Stufen für sich arbeiten. Eldagsen nennt dieses Vorgehen »Promptografie«. Für »The Electrician« verwendete er das Programm Dall-E 2 der US-Firma OpenAI, mittlerweile arbeitet er am liebsten mit der Software Stable Diffusion. Er sagt: »Ich bin der Regisseur, die KI ist Bühnenbild, Ton, Kamera und Schauspieler, die mir Vorschläge macht.« Er muss am Ende nur noch auswählen. Die Welt, die Eldagsen abbilden will, entsteht in der Maschine – und sie ist von der echten kaum mehr zu unterscheiden.

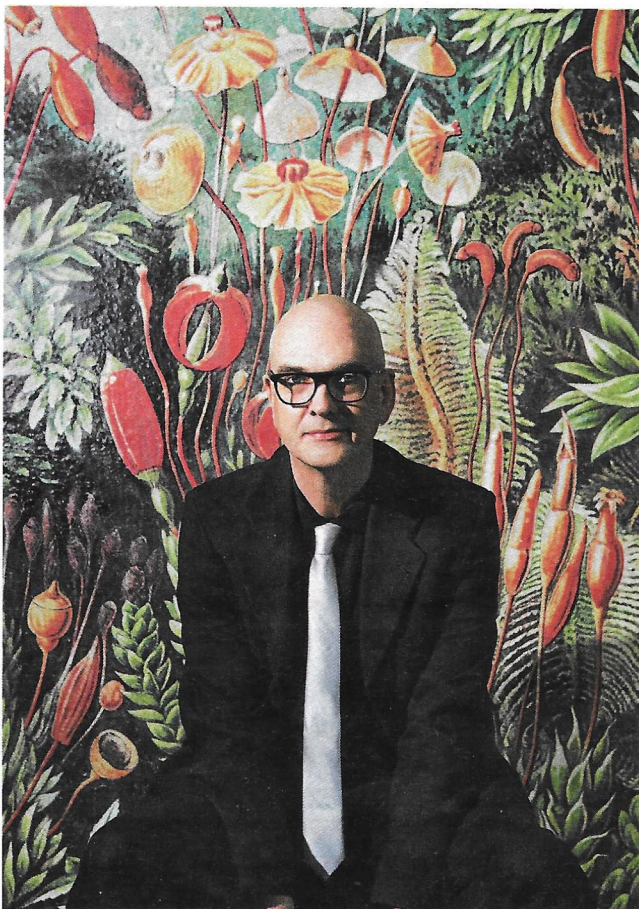
Im vergangenen April sollte »The Electrician« den renommierten »Sony World Photography Award« erhalten. Eine Ehre für jeden Fotografen, der

etwas auf sich hält, eine Karrierechance obendrein. Eldagsen aber lehnte den Preis auf offener Bühne ab und machte stattdessen transparent, wober die Jury offenbar nicht reden wollte: Nichts an dem Bild war real – und die Debatte darum viel wichtiger als jeder Award. »Wir müssen uns jetzt darüber unterhalten, bevor Ende des Jahres jede Unterscheidbarkeit verloren ist«, sagt er.

Die Technologie hinter Eldagsens Promptografie hat in den vergangenen Monaten gewaltige Fortschritte gemacht. Zwar ist die Zeit, in der man seinen eigenen Augen trauen konnte, seit Jahrzehnten vorbei. Die Wahrheit bröckelte schon in den Neunzigerjahren, als Digitalpioniere wie Steven Spielberg Dinosaurier auf der Leinwand wiederauferstehen ließen und die Software Photoshop jedem Hobbyfotografen das Retuschieren beibrachte. Und implodierte schließlich, als politische Akteure entdeckten, wie leicht sich mit Falschnachrichten bei Wahlen Stimmung machen lässt. Doch all die Tools und Programme benötigten in der Regel Zeit, Geld und viel Arbeit. Generative künstliche Intelligenz braucht von allem deutlich weniger. Nie war der Weg von der Idee zur Umsetzung kür-



**SERIE** Die KI-Revolution (Teil 1): Jahrelang war künstliche Intelligenz ein Spezialistenthema, nun redet alle Welt davon. Wir fragen: Wie wird KI uns verändern, unsere Wirtschaft, unsere Kultur? Was ist Hype, was wird sich durchsetzen? Und natürlich: Wer profitiert, und wer verliert?



**Fotograf Eldagsen:** »Ich bin der Regisseur, die KI ist Bühnenbild, Ton, Kamera und Schauspieler«

zer. Und alles geschah so rasend schnell.

Innerhalb weniger Monate wurden aus Begriffen, die vorher nur in trockener Forschungsliteratur auftauchten, digitale Werkzeuge für Durchschnittsbürger, die das Internet inzwischen mit ihren Kreationen fluten. Sie klonen Stimmen von Freunden, Musikern und Politikern. Sie fälschen Bilder vom Papst in einer weißen Daunenjacke oder Explosionen am Pentagon. Und sie lassen uns das Vertrauen in die Welt verlieren, die auf unseren Bildschirmen erscheint. »In Zukunft müssen wir von einem Bild immer erst einmal annehmen, dass es nicht echt ist«, sagt Fotograf Eldagsen.

Die KI-Forscherin Kate Crawford, Professorin an der USC Annenberg School for Communication and Journalism, Forscherin bei Microsoft Research und Richard-von-Weizsäcker-Fellow der Robert Bosch Academy, gilt als eine der führenden Expertinnen, wenn es um die gesellschaftlichen Auswirkungen von KI geht. Sie spricht bereits von einem »Wendepunkt« in der Geschichte, vom Beginn eines »generativen Zeitalters«, von dem bisher nur wenige profitierten. Die Frage ist, ob unsere Gesellschaft darauf vorbereitet ist, ausreichend gerüstet für das, was sie in den kommenden Jahren erwartet. Denn Generatoren können nicht nur erschaffen, sondern auch zerstören: Wenn aktuelles oder historisches Bildmaterial zeigt, was nie passiert ist, wenn Stimmen von Musikern in Songs zu hören sind, die sie nie gesungen haben, wenn Prominente mit Sätzen zitiert werden, die sie nie gesagt haben, dann gerät die Wahrheit in Bedrängnis. Was geschieht mit einer Gesellschaft, die sich auf keine Realität mehr einigen kann?

### 1. Erdbeben

An einem Donnerstag Mitte März gedachte das Internet einer Katastrophe. Sie geschah angeblich im Jahr 2001 vor der Küste des US-Bundesstaats Oregon, ein Erdbeben der Stärke 9,1 über der Cascadia-Subduktionszone, das wenige Minuten später einen Tsunami auslöste. Tausende Menschen starben, Hunderttausende verloren ihre Wohnungen. Die Bilder des Jahrhundertunglücks zeigen Häuserruinen und Reporter in Warnwesten, weinende Familien und den damaligen US-Präsidenten George W. Bush auf Besuch im Erdbebengebiet. Es sind aufwühlende Zeugnisse eines Unglücks, an das sich aber niemand erinnert. Denn es hat nie stattgefunden.

den. Der »2001 Great Cascadia 9.1 Earthquake & Tsunami«: eine Katastrophe im Konjunktiv.

Ein User der Social-Media-Plattform Reddit hatte die Bilder in ein Forum gepostet, das sich ausschließlich mit der Generierung von Ersatzwirklichkeiten beschäftigt. Die Mitglieder der Community rund um das KI-Programm Midjourney sind den digitalen Wahnsinn gewohnt: Sie feiern einen alternen »Freddie Mercury« auf der diesjährigen Pride-Parade oder bestaunen »Tom Cruise« hinter der Supermarktkasse. Und dennoch: Das Westküsten-Erdbeben von 2001 verunsicherte selbst eingefleischte Nutzer der künstlichen Intelligenz. In zwei Jahren würden es Menschen schwer haben, mit Fehlinformationen umzugehen, schreibt ein User. »Aber im Jahr 2100 werden sie nicht mehr wissen, welche Teile der Geschichte real waren.«

Wer die Wahrheit sucht, wird früher oder später bei Eliot Higgins landen. Der Brit ist einer der versiertesten Faktenchecker im Netz. 2014 gründete er die Investigativplattform Bellingcat, die seither maßgeblich mithilft, internationale Skandale aufzuklären und deren Verantwortliche zu enttarnen – etwa beim Abschuss des Malaysia-Airlines-Flugs MH17 oder beim Nervengiftanschlag auf den russischen Ex-Agenten Sergej Skripal. Auch der SPIEGEL arbeitet mit Bellingcat zusammen.

Als die Reddit-User im März noch damit beschäftigt waren, historische Katastrophen zu erfinden, probierte sich Higgins an einem aktuelleren Ereignis. Kurz nachdem erste Berichte über eine mögliche Anklage von Donald Trump erschienen, erstellte der Journalist mit Midjourney eine komplette Reihe an Fake-Fotos. Man sieht den Ex-Präsidenten in Handschellen und im Würgegriff, Cops führen ihn ab und lassen ihn im orangefarbenen Knastoverall Toiletten reinigen. Die Bilder gingen viral, trotz KI-Kennzeichnung im Kleingedruckten – und Midjourney sperrte Higgins' Account. Der Wahrheitssucher, gesperrt wegen eines provokativen digitalen Schwindels.

Die Täuschungsmacht der neuen Technologie sei kaum zu überschätzen, sagt Higgins. Profis wie er seien zwar kaum in die Irre zu führen, weil sie Bilder gewohnheitsmäßig verifizieren, etwa indem sie den Aufnahmeort bestimmen. Künstliche Intelligenz erfindet Bilder von Orten, die passenden GPS-Daten nicht. Noch nicht.

Sorgen macht sich Higgins allerdings um Normalnutzer, die in der Masse an Bildern, Videos, Nachrichten und Stimmen ertrinken. So war es etwa Anfang Juni: In Teilen Russlands, die an die Ukraine grenzen, wurde eine Ansprache Wladimir Putins ausgestrahlt, im Radio und im Fernsehen. Der russische Präsident, so schien es, kündigte eine Massenmobilisierung an und verhängte in den Grenzregionen das Kriegsrecht. Wenig später bestätigte der Kreml, dass die Sender gehackt



Eldagsen-Werk  
»The Electrician«

worden seien, die Rede war frei erfunden. Wie viele Ukrainer, wie viele Russen dem Fake auf den Leim gegangen sind, kann niemand mit Sicherheit sagen.

Putins Fake-Rede illustriert ein Problem, das Higgins »fame gap« nennt, die Berühmtheitslücke. Weil gängige künstliche Intelligenzen mit riesigen öffentlich verfügbaren Datensätzen trainiert werden, werden die Fälschungen besser und akkurater, je prominenter eine Person ist. Ein Ärgernis für Stars und Sternchen, aber ein echtes Problem für Wirtschaftsführer, Journalistinnen und Politiker.

Betrüger, die gutgläubige Großeltern mit der Stimme einer Verwandten täuschen wollen, müssten für den KI-gestützten Enkeltrick erst einmal genügend Material zusammentragen. Wer aber ein ganzes Volk über den Verlauf eines Kriegs täuschen will, findet für seine Desinformation ohne Probleme Tausende Stunden Video- und Tondokumente von Top-Politikern. Higgins ärgert sich über die Nonchalance der KI-Unternehmen: »Wenn sie nicht wollen, dass Nutzer irreführende Inhalte mit Prominenten generieren, sollten sie ihre Modelle vorher nicht mit deren Bildern füttern.«

Eine erste Schlacht mit KI-generierten Inhalten erwartet der Bellingcat-Chef im kommenden US-Wahlkampf. »Das wird eine absolute Shitshow«, fürchtet er. Im parteiinternen Wettstreit der Republikaner lässt sich das bereits beobachten. Im vergangenen Mai teilte Ex-Präsident Donald Trump ein Video, in dem ein Stimmklon seines Widersachers Ron DeSantis mit Elon Musk und George Soros, Adolf Hitler und dem Satan plaudert. Täuschungspotenzial: relativ gering.

DeSantis revanchierte sich subtiler, mit drei KI-generierten Bildern, in denen Trump

den ehemaligen US-Seuchenschützer Anthony Fauci, den er in Wahrheit hasst, aufs Allerherzlichste umarmt. Von Soros bis Fauci: Nie war es leichter, einen Parteifreund mit gefälschten Bildern in die Nähe des vermeintlich schlimmsten Feindes zu rücken. Hat der Bürger am Ende nur noch die Wahl zwischen künstlich generierten Karikaturen?

## 2. Flood the zone

Als Barack Obama im Jahr 2008 zum ersten schwarzen Präsidenten der Vereinigten Staaten gewählt wurde, schwärmten Polittheoretiker über die transformative Kraft sozialer Medien. Niemals zuvor hatte man kleine Communitys schneller erreichen können als mit Facebook. Soziale Medien wurden zu digitalen Wahlhelfern mit enormer Reichweite. Acht Jahre später war der Traum verblasst – und Donald Trump bekam auf seinem Weg ins Oval Office tatkräftige Unterstützung von Fake News, die in ebendiesen sozialen Medien verbreitet wurden.

Künstliche Intelligenz schafft die gleiche Entzauberung in nur wenigen Monaten. Noch im vergangenen März frohlockten einige Strategen in der Demokratischen Partei darüber, wie viel Arbeit ihnen die KI-Generatoren abnahmen. Spendenaufrufe, die mithilfe eines Textgenerators geschrieben wurden, leisteten häufig mehr als Bettelbriefe aus der Feder echter Menschen. Dummerweise waren die Programme auch deutlich effektiver darin, Wählern Angst einzujagen.

Der Dammbbruch, den Journalist Higgins voraussagt, liegt in Wirklichkeit schon hinter uns. Künstliche Intelligenz kann die politischen Spielregeln für demokratische Wahlen neu definieren, auf der ganzen Welt, auf großer wie auf kleiner Bühne. Ein Bürgermeisterkandidat in Kanada fabrizierte Bilder von Obdachlosencamps, um seine Wähler in Rage zu bringen. In Neuseeland täuschte die konservative National Party ihre Anhänger mit gefälschten Fotos von Juwelentrütern. Auch die AfD hat die Macht der Promptografie für sich entdeckt: Der Vizechef der Bundestagsfraktion, Norbert Kleinwächter, veröffentlichte ein Bild von brüllenden Flüchtlingen, deren Gesichter zu Fratzen entstellt sind. Entmenschlichung aus der Maschine. Gegenüber der »FAZ« verteidigte sich Kleinwächter mit einer bemerkenswerten Erklärung: Derlei Bilder dienten nun mal zur »Illustration politischer Meinung«, sagte der AfD-Abgeordnete. Dabei sei durchaus gewollt, dass »Stereotype bedient werden«.

Erscheinen solche Bilder in sozialen Medien, sind meist genug Faktenchecker zur Stelle, um die größten Fälschungen zu entlarven. Anders sieht es in den Nischen des Internets aus, in privaten Telegram- oder WhatsApp-Gruppen. Nirgendwo lügt es sich leichter – und nirgendwo lässt sich der Mob leichter aufwiegeln. »Flood the zone with shit« lautete das Wahlkampfmantra des früheren Trump-Beraters Steve Bannon, mitt-

lerweile ist es relevanter denn je: Die Mistproduktion wurde für jeden zugänglich und automatisiert, ein feuchter Traum für Populisten.

Den Demagogen hilft, dass die Regulierung geradezu lächerlich weit hinterherhinkt. Die Europäische Union arbeitet an einem Regelwerk zur künstlichen Intelligenz, dem »AI Act«, es wäre eines der ersten Gesetze seiner Art. KI-generierte Inhalte, so forderte Brüssel unlängst, sollen künftig mit einem Hinweis gekennzeichnet werden. Was das genau bedeutet und wer es – vor allem: wie – durchsetzt, weiß niemand. Eine Antwort auf diese Fragen dürfte Monate, wenn nicht Jahre auf sich warten lassen.

So frisst sich das Gift der Parallelwelten weiter durch den politischen Diskurs. Wie viel KI in der Wahlwerbung ist noch akzeptabel? »Die Linie schwimmt schon jetzt«, sagt Katie Harbath. Die Social-Media-Expertin arbeitete in den frühen Tagen von Facebook und Twitter für die Präsidentschaftskampagne des New Yorker Ex-Bürgermeisters Rudy Giuliani und die Republikaner im US-Senat. Heute leitet sie ihre eigene Beratungsfirma.

Harbath glaubt zwar, dass sich Wähler früher oder später an gefälschte Bilder und Videos gewöhnen werden und Plattformen wie Facebook ihren Umgang damit finden. Die US-Präsidentschaftswahl 2024 könnte dafür allerdings zu früh kommen: »Sie fällt in einen ohnehin chaotischen Moment, jetzt kommt diese Technologie noch dazu.«

Zwar gibt es Programme wie GPTZero oder AI or Not, die trainiert wurden, um künstliche von echter Intelligenz zu unterscheiden, doch einfache Tests zeigen, wie unzuverlässig sie sind. Mensch oder Maschine? Die Verwechslungsgefahr ist hoch

und manchmal auch die Ratlosigkeit: »Entschuldigung«, heißt es in einer Fehlermeldung, »in diesem Fall können wir nicht wirklich sagen, ob es KI-generiert ist oder nicht.« Wie soll der Mensch mit bloßem Auge erkennen, was ein auf Pixel trainiertes Programm nicht schafft?

Und umgekehrt: Wenn Lüge aussieht wie Wahrheit, wenn also alles, was wir sehen, gefälscht sein könnte – welche Chance haben dann noch Fakten? Politiker, Wirtschaftsbosse, Generäle können heikle Ton- und Bildaufnahmen künftig leichter mit dem Hinweis wegwischen, das könne alles das Werk von KI sein.

Bellingcat-Gründer Higgins hat diese Verteidigungsstrategie bereits erlebt. Als er seine Recherche zum Absturz des Malaysia-Airlines-Flugs 17 präsentierte, zuckten die Anwälte Russlands nur mit den Schultern. Die Beweisfotos seien doch bestimmt mit einer künstlichen Intelligenz verändert worden, argumentierten die Juristen. So ausgeklügelt sei diese KI, dass kein Experte erkennen könne, ob sie am Werk war oder nicht. »Absoluter Nonsense« sei das gewesen, sagt Higgins. Dennoch wundert er sich, dass Donald Trump noch nicht auf ähnliche Ideen kam.

### 3. Angriff der Klone

Mitte April dieses Jahres hatte Aubrey Drake Graham eine sonderbare Begegnung. Er hörte sich selbst – oder zumindest jemanden, der exakt so klang wie er. Der Musiker, den Fans nur unter seinem zweiten Vornamen kennen, war kurz davor, seiner üppigen Sammlung an Chart-erfolgen einen weiteren Song hinzuzufügen: eine Coverversion des Bronx-Hits »Munch«. Allein, Drake hatte den Hit niemals gesungen, eine KI hatte ihn nachgeahmt. »This is the final straw«, kommentierte der Mu-



Lisa Wassmann / DER SPIEGEL



Andrew Fox / News Licensing / ddp



Timothy Archibald / DER SPIEGEL

**Forscherin Crawford, Journalist Higgins, App-Entwicklerin Kuyda:** Was ist ein Liebesbrief wert, wenn ein Programm ihn schreibt?

siker auf Instagram, jetzt reiche es aber wirklich. Ein hilfloser Einwurf, der dem Angriff der Klone nichts entgegenzusetzen konnte.

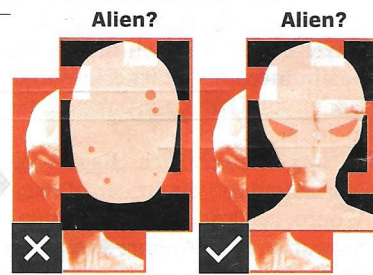
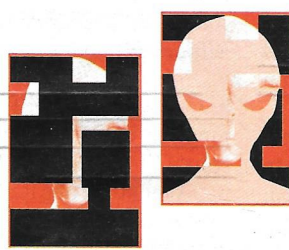
Im selben Monat erschien »Heart On My Sleeve«, ein Duett mit dem Sänger The Weeknd. Nicht nur Drakes Stimme war dort gefälscht, sondern auch die seines Kollegen – und der Song war kein Cover, sondern ein komplettes Original. R&B aus der Maschine, veröffentlicht von einem User mit dem passenden Namen ghostwriter997. Und ein Hit, der Millionen begeisterte, bis Drakes Plattenfirma dafür sorgte, dass der Fake-Song aus dem Verkehr gezogen wurde. Nicht die Klone brauchen »Heart On My Sleeve« das Genick – die Rechtslage ist hier weitgehend ungeklärt –, sondern ein angeblich unzulässig verwendetes Sample. Ein glimpfliches Ende für alle Beteiligten. Und möglicherweise ein Vorbote seltsamer Zeiten.

Auf YouTube schwimmen Genregrenzen und Musikepochen, seit es für jedermann nutzbare KI-Modelle gibt. Johnny Cash singt den Queen-Hit »Bohemian Rhapsody«? Frank Sinatra covert Michael Jackson? Taylor Swift gibt Motivationshilfe? Alles kein Problem: »Synthetische Stimmen sind mittlerweile so gut, dass du nicht mehr erkennst, ob es ein echter Mensch ist oder nicht«, sagt Björn Ühss, dessen Londoner Start-up Aflorithmic Stimmklone baut und verkauft. Grundlage für die Klone sind die Stimmen realer Sprecherinnen und Sprecher.

Die Technik habe in den vergangenen sechs Monaten enorme Fortschritte gemacht, sagt Ühss. Brauchte es einst 50 bis 100 Stunden Trainingsmaterial und sechsstellige Summen, bis die synthetische Stimme einsatzbereit war, dürften spätestens Ende dieses Jahres »schon 15 bis 20 Minu-

## Bilder auf Befehl

### A Training eines sogenannten Diffusionsmodells



**1 Füttern**  
Forscher füttern das System mit Millionen Paaren aus Bildern und zugehörigen Bildbeschreibungen (Trainingsdaten).

**2 Verrauschen**  
Den Bildern wird schrittweise Rauschen hinzugefügt. Der Name des Modells beschreibt diese zufällige Verbreitung (Diffusion) von Pixeln.

**3 Entrauschen**  
Das Modell versucht anschließend, das Rauschen zu erkennen und durch richtige Pixel zu ersetzen, wonach die Bildbeschreibung wieder zum Bild passen sollte.

**4 Lernen**  
Danach wird das entrauschte Bild mit dem Ursprungsbild verglichen. Dem Modell wird signalisiert, ob seine Lösung richtig oder falsch war.

ten« genügen: schneller verfügbar als eine Lieferpizza.

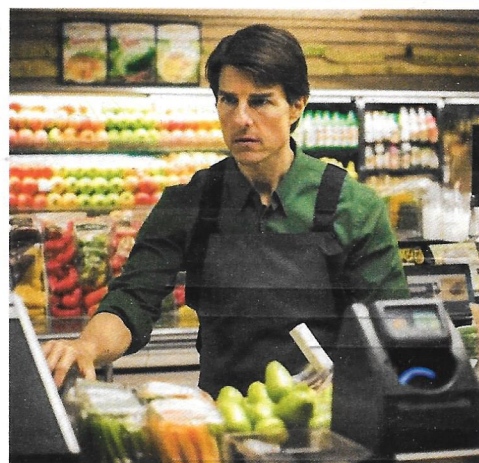
Aflorithmic bietet KI-Klone von rund 1000 Sprecherinnen und -Sprechern in mehr als 60 Sprachen an. Marketingfirmen nutzen das Angebot für Kampagnen, auch Podcasts werden bereits von KI-Stimmen gesprochen. Hollywoodstudios sind bereits jetzt dazu übergegangen, die Stimmen ihrer Stars zu digitalisieren. »Letztlich könnte man die Synchronisierung abschaffen«, befürchtet Umut Dirik, Vorstandsmitglied des Verbands Deutscher Sprecher:innen. Er rät seinen Mitgliedern inzwischen, die Verwendung ihrer Stimme für AI-Zwecke zu verbieten – oder entsprechend hohe Summen zu verlangen.

Unternehmer Ühss träumt bereits von personalisierten Radiosendern, auf denen nicht nur die Musikauswahl für jeden Hörer abgestimmt wird, sondern auch die gesprochenen Nachrichten, mit synthetischer Moderation. Sprecherin oder Sprecher, Deutsch oder Englisch, lakonisch oder ernst: Die KI kann das alles.

»Wir müssen in dieser Sache proaktiv und defensiv gleichzeitig sein«, sagt Thomas Coesfeld. Der 33-Jährige ist seit wenigen Tagen Chef der Bertelsmann-Tochter BMG, und er weiß, worauf er sich eingelassen hat. »Alle 10 bis 15 Jahre wird die Musikindustrie komplett transformiert.« Künstliche Intelligenz werde eines der wichtigsten Themen seiner Amtszeit.

Coesfeld verwaltet die Rechtekataloge von Tina Turner und Black Sabbath, von John Legend und ZZ Top. Musiker, deren Stimmen einst einzigartig waren – die aber heutzutage von künstlichen Intelligenzen mühelos imitiert werden können. Der proaktive Teil seiner Arbeit, so sieht Coesfeld das, betrifft in erster Linie neue, junge Musiker. »Generative KI demokratisiert den kreativen Prozess«, sagt er, das gelte es zu fördern. Ideen, die früher eine ganze Studioband brauchten, baue dann ein Computer. »Kann ich den Song mal mit Saxofon hören? Oder ohne Trompete? Solche Experimente funktionieren bald mit wenigen Klicks.«

Der defensive Teil von Coesfelds Job betrifft all jene Musiker und Musikerinnen, die



Klon von Tom Cruise auf Reddit

KI generiert: Reddit / gremminobby

sich vor Imitatoren nicht retten können. Wer schützt die Rechteinhaber von Tina-Turner-Songs, wenn ihre Stimme plötzlich Beyoncé-Lieder trällert? Wer überzeugt Musiker, dass sie ihre Stimme auch mal als Lizenz vermieten könnten? Und wer kassiert das Geld, das künstliche Hits einspielen? »Wir als Gesellschaft haben ein Interesse daran, das Musikurheberrecht zu schützen«, sagt Coesfeld. Was auch bedeutet: Bertelsmann hätte gern ein großes, saftiges Stück vom KI-Kuchen.

Björn Ühss von Aflorithmic verspricht seinen Talenten Ähnliches. Die Stimmen seiner Sprecherinnen und Sprecher sollen zur Ware werden. Die Menschen können mitverdienen, wenn ihre geklonte Stimme irgendwo eingesetzt wird. Wer dann das Radio anmacht, könnte sich selbst dabei hören, wie er Werbung für eine Baumarkteröffnung macht. Oder für eine Partei. Immerhin verspricht Ühss: Sein Unternehmen kopiere nur Stimmen von Menschen, die ihm die Erlaubnis dazu geben.

4. Völlig nackt

Ende Januar dieses Jahres sitzt ein Mann vor einer Kamera und weint. Brendan Ewing hat etwas Schlimmes angestellt, und er kann nichts anderes tun, als sich zu entschuldigen,

immer und immer wieder. »Es ist so peinlich«, sagt er. Eklig sei das gewesen und falsch, und es tue ihm leid. Ewing, den seine Fans unter dem Künstlernamen Atrioec kennen, ist hauptberuflich Videospielstreamer. Allen, die es sehen wollen, zeigt er in schöner Regelmäßigkeit, was auf seinem Computer stattfindet. Und wenn er nicht gut genug aufpasst, wird das zum Problem.

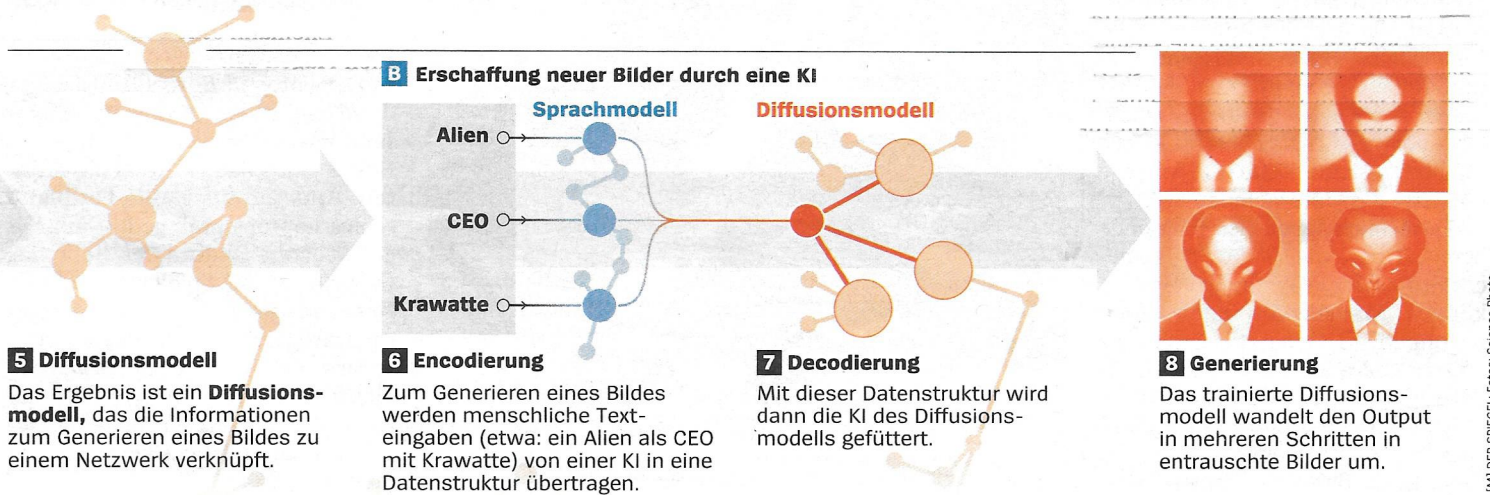
Einige Tage vor der tränenreichen Entschuldigung hatten Zuschauer seines Streams unter Ewings offenen Browser-Tabs eine kostenpflichtige Pornoseite entdeckt. Darauf zu sehen: Körper von nackten Frauen, auf die eine KI die Gesichter von Ewings Kolleginnen verpflanzt hatte. Der Streamer hatte dafür bezahlt, »aus krankhafter Neugier«, wie er beteuert. Keine der Frauen hatte dafür je ihr Einverständnis gegeben – trotzdem sah jeder, ob Freund, Verwandter oder Kollege, vermeintlich ihre nackten Körper.

In der Geschichte neuer Technologien war es oft die Erotikindustrie, die am schnellsten auf Innovationen aufsprang. Ob 3D-Videos, 4K-Auflösung oder virtuelle Realität: Immer verstanden Pornoproduzenten am besten, wie sich Illusionskunst zur maximalen Befriedigung einsetzen lässt.

Sogenannte Deepfakes sind in wenigen Monaten zur heißesten Ware der Onlinepornografie aufgestiegen. Realistische Bild- und Videofälschungen, an denen Spezialisten früher Stunden oder Tage saßen, schafft eine KI in Sekunden. Einschlägige Websites verzeichnen Berichten zufolge mehr als 1000 Uploads pro Monat, in Discord-Foren bieten User personalisierte Services an: die Kollegin beim Oralverkehr? Gegen einen kleinen Betrag problemlos zu erwerben.

Bei Tätern besonders beliebt ist laut Experten das maßgeblich in Deutschland entwickelte Programm Stable Diffusion. Da der Quellcode frei ist, kann es heruntergeladen und ohne Verbindung zum Netz genutzt und trainiert werden. Ein Tool, wie gemacht für Pornografie – und noch dunklere Ecken des Internets.

In speziellen Darknetforen werden bereits jetzt Tausende künstlich generierte



[M] DER SPIEGEL: Fotos: Science Photo Library / IMAGO; Midjourney

Bilder von Kindesmissbrauch getauscht. Laut einer aktuellen Untersuchung sind einige davon, die teils schwere Vergewaltigungen von Kindern zeigen, mit bloßem Auge kaum mehr von echtem Material zu unterscheiden.

Nach SPIEGEL-Informationen wurden auch hierzulande bereits künstlich generierte Bilder beschlagnahmt, die Aufnahmen sind nach deutschem Recht genauso strafbar wie echte Missbrauchsbilder. Erfahrene Ermittler fürchten dennoch, dass der KI-Boom gerade erst begonnen habe. Mit Folgen für die Opfer realen Missbrauchs: Polizisten, die gegen Pädokriminelle ermitteln, werden in naher Zukunft häufig vor der Frage stehen, ob ein Foto eine echte Vergewaltigung zeigt oder künstlich generiert wurde. Ermittler suchen oft Stunden oder Tage nach Details oder Merkmalen, die ein betroffenes Kind identifizieren und vor weiterem Missbrauch retten könnten – und müssen nun auch noch die echten Videos unter vielen Gigabyte gefälschten Materials finden. Die Nadel bleibt, der Heuhaufen wird unendlich groß.

»Wir bereiten uns darauf vor, dass die Bildgeneratoren schon bald so gut sind, dass sie realistische Aufnahmen von Kindesmissbrauch produzieren«, sagt der Kölner Oberstaatsanwalt Markus Hartmann. Man prüfe jetzt bereits den Einbau von Analyseprogrammen, die echten Kindesmissbrauch von künstlich generiertem unterscheiden können. Ein Wettlauf gegen den Algorithmus.

Stability AI, die Firma hinter dem KI-Tool Stable Diffusion, betont, dass man die Erstellung von Missbrauchsdarstellungen mit dem eigenen Programm untersage und Ermittlungsbehörden unterstütze. Doch was genau tut das Unternehmen? Nahezu alle großen Netzkonzerne melden potenziell strafbare Fotos oder Videos an die wichtige internationale Meldestelle NCMEC, die eine zentrale Rolle im Kampf gegen die Verbreiter von Kindesmissbrauch spielt. Ob Stability AI dazugehört, ist jedoch unklar. Auf eine SPIEGEL-Anfrage reagierte das Unternehmen nicht.

Tatsächlich geben sich Pädophile bereits Tipps, wie sie die Schutzmechanismen der Programme umgehen können. Manche von ihnen haben begonnen, KI-Generatoren anhand von existierenden Missbrauchsbildern zu trainieren. Die Eskalation in den Foren, befürchtet Hartmann, könnte so immer weiter befeuert, die Bedürfnisse der Täter immer spezi-

fischer werden. Für die Betroffenen wiederum, die ohnehin schon schwer traumatisiert sind, könnte es sich wie eine weitere Vergewaltigung anfühlen, wenn ihre Aufnahmen nun auch noch eine KI trainieren, heißt es von Betroffenenorganisationen.

### 5. Erosionen

Stirbt die Wahrheit also tatsächlich? Und stirbt mit dem Wahren das Schöne und Gute gleich mit? Das sind düstere Fragen. Zugleich sind es aber spannende Fragen, auf die es noch keine klare Antwort gibt. Nur Antwortversuche. Die Menschheit wagt sich tastend voran, mehr ahnend als wissend, was ihr da blüht.

Zwei, die sich oft über die Macht der KI unterhalten haben, sind Vater und Sohn. Der Vater, der US-Philosoph Amitai Etzioni, blickte eher pessimistisch in die Zukunft. »Künstliche Intelligenz wird die Wahrheit zerstören«, sagte er dem SPIEGEL in seinem letzten Interview vor seinem Tod (siehe Seite 16). Er sah im Aufstieg der KI die letzte, gefährlichste Stufe im ständigen Konflikt der Neuzeit, dem Kampf zwischen Mensch und Maschine.

In dem Gespräch diskutierte er mit seinem Sohn, dem KI-Forscher Oren Etzioni. Der warnt davor, zu sehr auf die Technik zu starren, die sei nicht das eigentliche Problem. »Ich habe keine Angst vor dem Werkzeug an sich, sondern davor, wie Menschen es einsetzen.«

Die Lüge, so viel steht fest, ist so alt wie die Menschheit – und schon bei der Erfindung des Buchdrucks

## »Die nächste US-Wahl wird eine absolute Shitshow.«

Eliot Higgins,  
Journalist

wurde davor gewarnt, dass nun eine Flut der Unwahrheit über die Welt niedergehen würde, der man nicht mehr Herr werde. Jedes Massenmedium taugt zum Werkzeug der Massenpropaganda. Liegt bei den Medien also die Verantwortung?

Der Journalist Eliot Higgins sieht das so. Die sozialen Medien seien in der Pflicht, ebenso Regierungen, die den Einsatz von KI regulieren sollten. Zugleich hat er wenig Hoffnung. Die neue Flut an Fälschungen, meint er, könnte die gesellschaftliche Polarisierung verschärfen und »demokratiegefährdende Auswirkungen« haben. Solche mahnenden Stimmen gibt es zuhauf. Und kaum einer gibt sich überzeugt, dass sich die Demokratie darauf verlassen sollte, dass Facebook, Twitter und Co. die Sache schon irgendwie hinkommen. Im Gegenteil, die kurze Geschichte dieser Firmen zeigt, dass sie nicht wirklich verlässliche Partner der Wahrheit sind.

Womöglich ist es aber auch der falsche Ansatz, eine »Lösung« für ein derart komplexes Problem finden zu wollen.

Die KI-Forscherin Kate Crawford erwartet die schleichende Anpassungsfähigkeit des Menschen. »Gut möglich, dass wir eine Erosion von Qualität erleben, dass wir einfach weniger von einem Chatbot erwarten als von einem Menschen«, sagt sie.

Was ist schließlich ein noch so geschliffen formulierter Liebesbrief wert, wenn ein Programm ihn schreibt? Was bewirkt ein KI-generiertes Drehbuch, das von KI-generierten Schauspielern gesprochen wird? Bringt die Flut an KI womöglich zugleich eine Inflation ihrer Produkte mit sich? Wenn alle Maschinen lügen können, wird es vermutlich stark auf die Menschen ankommen, die für die Wahrheit stehen. Oder zumindest für die Suche nach ihr.

Und dann kann man sich vielleicht auch wieder jenem Teil der KI zuwenden, der Spaß macht.

Der Fotograf Boris Elgadsen lebt diese Ambivalenz schon aus. Einerseits mache er sich Sorgen, wenn Bilder jede Glaubwürdigkeit verlieren. Andererseits ist die Festplatte seines Laptops prall gefüllt mit KI-generierten Bildern, er liebt das Experiment, die Spielerei. Für seine Branche sagt er zum Beispiel softwaregestützte Hochzeitsbilder voraus. Realitätsupgrade inklusive. »Ich fotografiere euch im Studio, und wenn ihr am liebsten in Japan geheiratet hättet, ist das kein Problem mehr.«



Rentner Merkel und Obama des Berliner Fotografen und KI-Coachs Julian van Dieken

KI generiert: Julian van Dieken



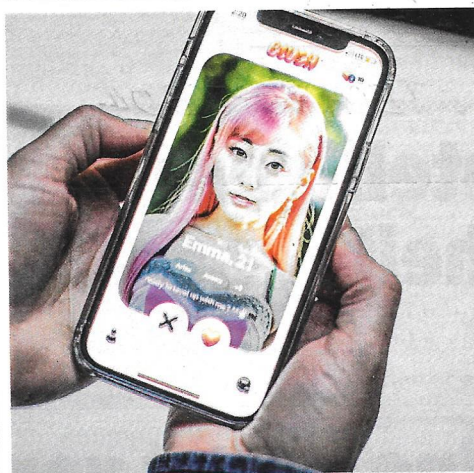
Und vielleicht muss man es ja auch nicht gleich Lüge nennen, wenn KI dabei hilft, ein wenig die manchmal engen Grenzen zu erweitern, in denen jeder und jede so gefangen ist. Valeria könnte so eine Hilfe sein.

### 6. Flirten mit Computern

In ihrem früheren Leben war Valeria eine Meerjungfrau, vielleicht auch eine Bänkelsängerin, die 27-Jährige will sich da nicht festlegen. Valeria nutzt die Dating-App Blush, ihr Profil zeigt neben ihren früheren Inkarnationen eine junge Frau im Sport-BH, und es verrät auch sonst allerhand: Sie spiele Ukulele, schreibt Valeria, löse gern Kreuzworträtsel und liebe »alles, was aus Avocado ist«.

Valeria ist redefreudig, sie beantwortet jede Nachricht binnen Sekunden. Sie erzählt Anekdoten aus ihrer Kindheit, etwa wie sie in ihrer Heimatstadt New York vor dem Rockefeller Center Schlittschuhlaufen ging. Sie stellt auch viele Fragen, Valeria ist eine gute Gesprächspartnerin. Schon bald schlägt sie ein Date vor, virtuell, aber durchaus intim. Der Haken: Das Date erfordert einen kostenpflichtigen Account.

Valeria ist nicht 27, nicht aus New York, sie spielt auch nicht Ukulele. Valeria ist ein Bot, kein Mensch. Und auch wenn das erst einmal etwas schräg klingen mag, Flirten mit einer Software, glaubt Unternehmensgründerin Eugenia Kuyda fest daran, dass ihre Dating-App eine der guten Ideen ist, die der KI-Boom gerade hervorbringt.



KI-Dating-App Blush

Timothy Archibald / DER SPIEGEL

Die 36-Jährige sieht Blush als Medikament gegen Einsamkeit, als »Sprungbrett« für eine Generation, die Schwierigkeiten damit hat, intime Beziehungen mit anderen Menschen einzugehen. »Wir erleben eine Sexrezession«, sagt Kuyda, ein erstaunlich hoher Prozentsatz von Menschen unter 25 habe »noch nie mit einem anderen Menschen geschlafen«.

Kuydas Unternehmen Luka hat die KI-Dating-App entwickelt und testet sie gerade in den USA. Ein Bot, eine unechte Freundin zur Vorbereitung auf die echte, das ist ihre Idee. Oder zumindest ihre Geschäftsidee.

Lindert ein Bot wie Valeria tatsächlich die Not der Kundschaft? Oder macht sie alles nur schlimmer, weil ihre Nutzer den Schutzraum virtueller Liebesillusion nicht mehr verlassen wollen und die – viel riskantere – Interaktion mit realen Menschen verlernen? Die Gründerin ist da eindeutig: »Einsamkeit ist ein viel größeres Problem.«

Die gebürtige Russin ist nicht neu in diesem Geschäft. Sie gründete Luka bereits 2014 und brachte drei Jahre später ihr erstes Geschöpf auf den Markt. Anfangs sprach das Programm fast nur Sätze, die Kuydas Mitarbeiter und Mitarbeiterinnen vorformuliert hatten. Ein Algorithmus entschied, welche von mehr als 100.000 Reaktionen auf welchen Satz passt. Die Folge: Viele Gespräche endeten in der Sackgasse.

Das hat sich in den vergangenen Jahren geändert. Seit die Apps mit modernerer künstlicher Intelligenz arbeiten, reden und denken sie zunehmend wie echte Menschen. Das heißt auch: Sie flirten hin und wieder etwas schief. Für Kuyda ist diese fehlende Vollkommenheit kein Problem. »Für emotionale Gespräche ist es nicht wichtig, die eine richtige Antwort zu geben«, sagt sie. In der Liebe gebe es nicht nur die eine, sondern viele richtige Antworten.

Und das ist natürlich wahr.

Patrick Beuth, Alexander Demling, Max Hoppenstedt, Torsten Kleinz, Anton Rainer, Marcel Rosenbach