

Forschung aktuell – Wissenschaft im Brennpunkt

Produktionsmanuskript

Sprecher

erzählt die ganze Geschichte im Rückblick

Rollen

Die Maschine

Marie Winkler

Stefan Seidel, ihr jetziger Freund

Politiker

Moderatorin

Moderator

div. Stimmen

Übersetzer

für Ghani, Kroll, Mallat

Übersetzerin

für Bryson

An- und Absage

Studio

17.10.-18.10 von 18:00 - 1:30 B-S7

Interviewpartner

Rayid Ghani, University of Chicago

Joshua Kroll, University of California, Berkeley

Joanna Bryson, University of Bath

Stéphane Mallat, ENS (École Normale Supérieure) Paris

The logo for Deutschlandfunk, featuring the text "Deutschlandfunk" in white, bold, sans-serif font on a blue rectangular background.

Die Maschine

Über das Dunkle in der Black Box

Dokumentation von Thomas Reintjes

Regie: Friederike Wigger

Redaktion: Christiane Knoll

Manuskript: 20. September 2017

Sendetermin: 26. Dezember 2017

SZENE 1: DER FUND

**Im Treppenhaus eines Mehrfamilienhauses,
vor der Wohnungstür von Marie Winkler.
Sie stellt die Szene für den Reporter nach.**

MARIE WINKLER: Es war eigentlich ein ganz normaler Freitag. Ich war froh, dass ich ein bisschen früher Feierabend machen konnte. Weil, ich hatte Kinokarten für 7 Uhr oder so.

DIE MASCHINE: Fluch der Karibik 6.

MARIE WINKLER: Also, ich bin hier so die Treppe raufgekommen, und da lag dieses Paket vor der Tür. Ich denk noch, komisch, normalerweise bestell ich immer alles zur Packstation oder auf die Arbeit. *(Sie schließt die Tür auf und geht rein.)*

SPRECHER: Es war gegen halb sechs, als Marie Winkler das seltsame Paket gefunden hat.

DIE MASCHINE: Marie Winkler, Angestellte, ledig, keine Kinder.

MARIE WINKLER: Ich hab das Paket erstmal auf den Tresen gelegt. Die Schuhe ausgezogen. Wahrscheinlich hab ich in den Kühlschrank geguckt. Was man halt so macht.
Ich mein, ich hab so ein bisschen den Überblick verloren, über meine Bestellungen. Ich dachte, da ist wahrscheinlich neues Müsli drin oder'n Buch oder keine Ahnung. Naja, und dann mach ich halt irgendwann das Paket auf und dann ist da dieses Teil drin. Dieses komplett schwarze Teil. So kantig, ach, das ist so ne komische Form... Rund, aber mit ganz vielen dreieckigen Flächen. Ich dachte im ersten Moment, es ist eine, ich weiß nicht, eine Blumenvase oder so.

SPRECHER: Rückblickend erscheint es fast naiv, wie sie das schildert. Denn aus heutiger Sicht weiß jeder, was das schwarze Ding in dem Paket ist:

MASCHINE: Die. Maschine.

- MARIE WINKLER: Ich hab das dann auf die Arbeitsplatte gestellt. Ich weiß noch, wie überrascht ich war, dass es so schwer war für seine Größe. Ich hatte natürlich keine Ahnung... *(redet weiter, improvisiert...)*
- SPRECHER: *(darauf)* Vielleicht war es gerade diese Unscheinbarkeit, die die Maschine so unverdächtig machte. Was konnte so ein kleines Teil schon ausrichten? Doch dann hat sie in ein paar Monaten unsere Gesellschaft verändert.
- ANSAGE: Die Maschine. Wie tickt sie, wie denkt sie?
Dokumentation von Thomas Reintjes.
- DIE MASCHINE: Thomas Reintjes, Sozialversicherungsnummer 122-55-4327. Ist im Fitnessstudio angemeldet, geht aber nie hin.

SZENE 2: DIE VERHEISSUNG

- SPRECHER: Schon seit langem gab es Warnungen vor Künstlicher Intelligenz, aber den meisten Menschen erschien die Technik eher wie Spielzeug. Auf Youtube fanden sich Videos, in denen KI-Systeme Witze erzählten.
- VIDEO: O-TON 1 (<https://www.youtube.com/watch?v=SeCuF2M9bb4>)
Mann: Alexa, erzähl uns einen Witz!
Alexa: Was trägt die LED-Leuchte, damit sie nicht nass wird? Einen Lampenschirm.
- SPRECHER: Das war im Juli 2019 auch schon nicht witzig. Aber die Sternstunde der Künstlichen Intelligenz sollte ja erst noch kommen. Wissenschaftler hatten seit Jahren verheißungsvolle Dinge zu sagen.

RAYID GHANI:

O-TON 2

[RG 00:00:47] The whole idea is, how can we improve society...

Es geht darum, die Gesellschaft zu verbessern, drängende soziale Probleme anzugehen. Und zwar nicht mit den traditionellen Methoden, die auf Anekdoten basieren, sondern evidenzbasiert, auf der Grundlage von Daten. ... *and not data driven. [15.2]*

SPRECHER:

"Data Science for Social Good" nennt Rayid Ghani das, was er an der Universität Chicago macht.

DIE MASCHINE:

Rayid Ghani. War im Obama-Wahlkampf Chef-Daten-Analytiker, sein Büro befindet sich im Searle-Gebäude, Zimmer 219.

SPRECHER:

Das Gute an der Künstlichen Intelligenz zu sehen, ist einfach. Computersysteme fahren besser Auto. Computersysteme können Bilder schneller erkennen als Menschen. Und Computersysteme können besser in die Zukunft sehen als Menschen. Rayid Ghani hatte 2017 beispielsweise bei der Polizei in Charlotte, North Carolina, ein System installiert, das vorhersagt, welcher Polizist möglicherweise bald unverhältnismäßige Gewalt einsetzen wird.

RAYID GHANI:

O-TON 3

[RG 00:08:17] If you knew exactly which indicators were useful then you would just solve the problem without the computer. But you don't.

Wir haben keinen Sensor, der Stress misst. Aber wir haben Hinweise: Wie oft hat jemand in letzter Zeit Überstunden gemacht? Wie oft wurde jemand weit weg von seiner Dienststelle eingesetzt? Wie oft wurde jemand bei Fällen häuslichen Missbrauchs eingesetzt? Wir haben oft Tausende solcher Hinweise und die geben wir alle dem System, zusammen mit Daten aus der Vergangenheit über solche Vorfälle von Polizeigewalt. Und dann lassen wir den Computer die Indikatoren so miteinander kombinieren, dass sie die Zukunft am besten vorhersagen.

[00:09:48] we let the computer basically combine those indicators in a way that best predicts the future. [7.0]

SPRECHER: Es stand Machine Learning Systemen praktisch auf der Stirn geschrieben, dass sie gut und wohlwollend sind. Aber was, wenn man an der Oberfläche kratzt?

SZENE 3: DIE ENTDECKUNG

Dokuszene mit O-Tönen, die in Marie Winklers

Küche aufgenommen wurden.

SPRECHER: Marie Winkler stellte das schwarze Ding auf ihren Küchentresen und ihr wurde klar: Das war eine Art Gerät.

MARIE WINKLER: Ich wollte ja noch ins Kino und deswegen hatte ich es ein bisschen eilig.

DIE MASCHINE: Zeit im Badezimmer: 5 Minuten, 12 Sekunden.

MARIE WINKLER: Ich zieh mich also kurz um, bisschen frisch machen und so, und dann musste ich auch schon wieder los.

DIE MASCHINE: Umziehen. Jeans und Tanktop. Ohrringe. 2 Minuten, 40 Sekunden.

MARIE WINKLER: Und ich weiß es noch genau. Das war so... (unfassbar). Ich lauf hier auf und ab und sag zu mir selber: Wo hab ich denn jetzt den Schlüssel hingetan? Und dann sagt das Teil, also die Maschine: Steckt außen in der Tür.

DIE MASCHINE: Steckt in der Tür. Außen.

MARIE WINKLER: Also, da ist mir ja die Kinnlade runtergefallen.

SPRECHER: Sie hat das Gerät noch kurz untersucht, kein Stromanschluss sichtbar, aber klar, eine Kamera hätte sich irgendwo hinter dem schwarzen Plastik verbergen können. Und dann musste sie los.

- MARIE WINKLER: Ich wollte ja das Date nicht sausen lassen. Aber das Teil war mir unheimlich. Deshalb hab ich im Rausgehen noch die Decke da vom Sofa über die Maschine geworfen. Und dann bin ich halt zum Kino.
- SPRECHER: Ihr Date ist heute ihr Freund, Stefan Seidel. An dem Abend damals hat sie ihn zum ersten Mal mit nach Hause genommen. Auch, weil sie nicht mit der Maschine allein sein wollte. Zusammen haben sie dann angefangen, die Maschine zu erkunden.
- STEFAN SEIDEL: Ich war natürlich auch erstmal so: Wie, das Ding lag einfach vor deiner Tür? Und das fängt plötzlich an zu reden? Aber dann war irgendwann der Spaß größer als die Frage, woher die Maschine nun kommt.
- SPRECHER: An diesem Abend machten die beiden das Video, das heute jeder kennt. Marie filmte mit ihrem Handy, wie sie und Stefan mit der Maschine redeten.
- STEFAN SEIDEL: *(im Video)* Ok, ähm... Wie wird das Wetter morgen?
- DIE MASCHINE: Das Wetter für Samstag, 27. Juli: So wie heute, nur am Nachmittag zieht es sich dann zu und um 5 fängt's an zu regnen.
- MARIE WINKLER: Oh. Hm. Ok, hier ist ne wichtige Frage: Soll ich nächstes Jahr Strand- oder Wanderurlaub machen?
- DIE MASCHINE: Du wirst Wanderurlaub machen.
- MARIE WINKLER: Wanderurlaub, hm. Alles klar. *(Wobei klar ist, dass es nicht klar ist. Ihr fehlt hier offenbar eine Erklärung.)*
- STEFAN SEIDEL: Ich mach erstmal 'nen Kurzstreckentrip. Wie komme ich denn heute Abend am besten nach Hause?
- DIE MASCHINE: Am besten du nimmst ein Mietfahrrad. Die Straßenbahn wird wegen einer Störung um 23 Uhr 13 und 25 Sekunden den Betrieb einstellen.

MARIE WINKLER: Ach, echt? Na, mal sehen. Ich hab sowieso 'ne bessere Idee. Pass auf. Soll ich Stefan heute hier schlafen lassen?

DIE MASCHINE: Ja, das ist kein Problem.

STEFAN SEIDEL: Oh!

MARIE WINKLER: Mh! (*So eine Art positives "Tja"*)

SPRECHER: Die beiden luden das Video ins Netz. Und als es dann bei den Bottropper Verkehrsbetrieben ...

DIE MASCHINE Ursache wird ein Problem mit der Stromversorgung sein.

SPRECHER: ... tatsächlich eine Störung gab, und zwar exakt um 23 Uhr 13 und 25 Sekunden, ging das Video viral.

SZENE 4: DIE BEDENKEN

**Wissenschaftler sagen, man wisse nicht, worauf diese Antworten beruhen
und wie sie zustande kommen.**

**Es ist unmöglich ins Innere der Maschine zu gucken.
Niemand weiß, woher sie kommt.**

SPRECHER: Die meisten User teilten es mit einer Mischung aus Faszination und Unglaube. War das Video wirklich echt? Der Clip tauchte auch bei Joanna Bryson von der Universität Bath im Social Media Stream auf.

DIE MASCHINE: Schreibt durchschnittlich 3 Blogartikel pro Monat. Geboren in Milwaukee, Wisconsin.

SPRECHER: Joanna Bryson dachte spontan, dass die Maschine nicht das war, was sie zu sein schien. So vorausschauend und allwissend konnte keine künstliche Intelligenz sein. Und deshalb musste irgendetwas faul sein.

JOANNA BRYSON:

O-TON 4

[JB 00:29:05] We're never going to have an algorithm that can predict everything. Right, it's just intractable.

Es wird niemals einen Algorithmus geben, der alles vorhersagen kann. Das geht nicht. Es gibt mehr Varianten von Schachpartien mit 35 Zügen als es Atome im Universum gibt. Atome im Universum! Das ist unfassbar, oder? Deswegen ist es schwachsinnig zu behaupten, absolut allwissend zu sein. *It's insane to talk about knowing everything and being perfectly omniscient. [18.9]*

SPRECHER:

Machine Learning Systeme konnten nach Meinung der Wissenschaft und aktuellem Stand der Technik bestenfalls voraussehen, wer welchen Film mögen würde, und das auch nur mit einer gewissen Wahrscheinlichkeit. Eine Störung auf die Sekunde genau vorherzusagen war unmöglich. Für Marie Winkler stellte sich die Realität jedoch anders dar. Sie plauderte munter weiter mit ihrer Kiste.

MARIE WINKLER:

Für mich war das gar nicht mehr so überraschend. Es wurde doch immer gesagt, wie viele Datenspuren wir hinterlassen und wie transparent wir sind. Also, mich wunderte das nicht, dass die Maschine daraus Schlüsse ziehen konnte.

STEFAN SEIDEL:

Gut, aber es gab ja schon auch Situationen... Also, zum Beispiel als das Teil meinte, deine Schwester wäre schwanger...

DIE MASCHINE:

Glückwunsch, du wirst bald Tante!

MARIE WINKLER:

Ja, und die wusste es selbst noch nicht. Aber es stimmte. Ja, das war eigentlich das erste Mal, dass ich dachte: Das gibt's doch jetzt wirklich nicht!

SPRECHER: Auch vor der Maschine hatten Deep Learning Systeme schon überraschend akkurate Ergebnisse geliefert, Aber selbst die Programmierer verstanden nicht immer, warum. Wie zogen diese nachgebauten Nervennetze Schlüsse aus ihrem Wissen? Im Kern waren sie wie eine Black Box. Genau deshalb hatten Leute wie Rayid Ghani echte Probleme, ein tiefes Vertrauensverhältnis zu den lernfähigen Maschinen aufzubauen. *

RAYID GHANI: O-TON 5
[RG 00:35:36] I've used Google Maps in ways enough that I trust it. I might not bother to verify to sort of go with it. [11.5]
Ich habe Google Maps oft genug benutzt, dass ich ihm vertrauen kann. Das muss ich nicht noch verifizieren. Da geht's ja nicht um viel.
[RG 00:50:26] But for a lot of decisions I'm not willing to give them up to use the computer...
Aber bei vielen anderen Entscheidungen bin ich nicht bereit, sie einem Computer zu überlassen. Ich würde meinen Computer nicht einmal entscheiden lassen, welchen Computer ich mir kaufe.
*I don't let my computer decide what computer I want to buy. **

SPRECHER: Das Problem war, dass man nicht einfach in das Gehirn einer fremden Intelligenz reinschauen konnte. (*Aber ein paar Nerds versuchten trotzdem, selbstlernende Systeme transparenter zu machen. Joshua Kroll zum Beispiel.

DIE MASCHINE: Joshua Kroll. Forscher und Hacker. Eltern beide Zahnmediziner.

JOSHUA KROLL: O-TON 6
[JK 00:18:12] There are methods that work even in in the case where where decisions are made in a black box way.
Es gibt Methoden die bei Black Box-Entscheidungen funktionieren.
...in the same way that you can't go to a human decision maker...

SPRECHER: Joshua Kroll beschäftigt sich an der University of California in Berkeley mit Verfahren, um Künstliche Intelligenzen besser zu verstehen. Zum Beispiel konnte man sie wie mit einer Taschenlampe stückweise zu erkunden. Für Systeme, die Katzen auf Bildern erkennen, funktionierte das damals schon ganz gut.

JOSHUA KROLL: O-TON 7
Also, zum Beispiel sagt das neuronale Netzwerk, dieses Bild sei eine Katze. Dann kann ich das Bild ein bisschen verändern und zum Beispiel Teile löschen. Wenn die Katze auf einem Tisch sitzt, löscht man den Tisch und das neuronale Netzwerk, wenn es gut ist, sagt immer noch, das sei eine Katze.

....will still say this is a cat because the table was sort of irrelevant in the first place. But if you you know if you delete the whiskers...

Aber wenn man jetzt die Schnurrhaare oder die Ohren oder etwas anderes Relevantes löscht, dann sagt das neuronale Netzwerk vielleicht, es sei etwas anderes.

[00:19:50] you can explore the space kind of around the example that you're looking for and say and that can tell you both what other examples might you have provided that would have given you a different answer. And what is it about the example that you have that's really important to the decision.

So kann man den Raum um den Fall, für den man sich interessiert, erkunden und sehen, was in seinem Inneren vor sich geht

....what's going on inside the decision making system it's just a purely blackbox method. [21.1]

SPRECHER: Vielleicht konnten Leute wie Kroll auch die Maschine von Marie Winkler so weit löchern, dass sie ihre Geheimnisse Preis geben würde?

DIE MASCHINE: Das wird nicht passieren.

JOSHUA KROLL: O-TON 8
[JK 00:21:11] Schlussendlich hätte man das Modell dann quasi reverse engineered, indem man es dazu gebracht hat, sich selbst zu erklären.sort of reverse engineered the model by by making it explain itself.

SPRECHER: *(evtl. hier schon auf Musik vom Frühstücksfernsehen)* Aber davon war er zu diesem Zeitpunkt noch weit entfernt.

SZENE 5: DIE MEDIEN

MARIE WINKLER: Es ging alles unglaublich schnell. Am Montagmorgen waren wir schon im Frühstücksfernsehen.

STEFAN SEIDEL: Ja, das war Wahnsinn.

Ausschnitt aus Frühstücksfernsehen wird eingespielt:

MODERATORIN: Ist es eine aus Versehen geleakte neue Version von Alexa? Oder ein neues Produkt von Apple, das jemand in das falsche Paket gepackt hat? Wir wissen nicht, woher es kommt, aber seit diesem Wochenende bewegt es die ganze Welt. Das Internet hat es einfach "Die Maschine" getauft. Und sie ist hier, bei uns - gleich, nach den Nachrichten!

Ausschnitt Ende

STEFAN SEIDEL: Die Maschine stand auf dem Tisch, Marie und ich saßen mit der Moderatorin drumrum auf der Couch und haben ihr ein paar alberne Fragen gestellt.

Ausschnitt TV

MARIE WINKLER: Hey, ähm, soll ich meinen Job kündigen und nochmal studieren?

DIE MASCHINE: Ja, das macht dich glücklicher.

MARIE WINKLER: Oh! Sorry, Chef...

MODERATORIN: Haha, da steckt ja sogar ein Karriereberater drin! Da müssen Sie jetzt wohl mal mit Ihrem Chef reden...

Ausschnitt Ende

STEFAN SEIDEL: Bis dahin war das alles noch lustig. Auch wenn Marie eigentlich wusste, dass das mit der Kündigung und dem Zweitstudium die richtige Entscheidung war. Das war so ein bisschen unheimlich. Aber dann kam natürlich der Knaller.

MARIE WINKLER: Ja, ich glaube da wollte die Moderatorin ein bisschen ernster werden - und ist dann ein bisschen übers Ziel hinausgeschossen.

Ausschnitt TV

MODERATORIN: Jetzt hatten wir ja gerade in den Nachrichten wieder vom Nordkorea-Konflikt gehört. Der schwelt ja nun schon ewig. Vielleicht könnten wir mal fragen, also... Wie können wir den Nordkorea-Konflikt lösen?

DIE MASCHINE: Es wird auf die nukleare Option hinauslaufen.

MODERATORIN: (*verunsichert*) ...nukleare Option? Hat das gerade gesagt, es wird einen Atomkrieg geben?

Ausschnitt Ende

SPRECHER: Und das war der Anfang des Hypes. Der Nordkorea-Konflikt war eine komplexe Situation, die niemand vollständig überschauen konnte. Da erschien es opportun, die Äußerung der Maschine einfach als Tatsachenbehauptung anzuerkennen.

DIVERSE STIMMEN: Computer für Atombombe auf Nordkorea!
Zündet Nordkorea jetzt die Bombe?
German AI goes nuclear!
Will dieser Computer die Menschheit auslöschen?
Die Maschine zieht in den Krieg
Alles über die Maschine in dieser Ausgabe
Was weiß die Maschine über unsere Zukunft?
Maschine: Fluch oder Segen?
Warum die Maschine gefährlich ist
Computer macht Weltpolitik

SZENE 6: MEHR BEDENKEN

- SPRECHER: Bei Joshua Kroll läuteten die Alarmglocken. Wer hatte die Maschine programmiert? Mit welchen Daten war sie gefüttert – und zu welchem Zweck? Wollte sie warnen, informieren oder provozieren? Sein Instinkt war eindeutig: Finger weg. Die Maschine spricht nicht für sich selbst.
- JOSHUA KROLL: O-TON 9
[JK 00:46:08] someone is building these things, these things are their artifacts.
Irgendjemand baut sie.
They don't come fully formed from the head of Zeus....
Diese Dinge kommen ja nicht einfach von Zeus. Sie kommen von Leuten, die sie bauen und die bestimmte Ansichten haben, was das System machen soll, (und) was die Ziele dahinter sind.
...what the goals of the systems are how to evaluate the system's correspondence to those goals. [21.6]
- MARIE WINKLER: Kurz danach kamen vier Männer bei mir vorbei. Die haben mir ihre Badges entgegengestreckt und gesagt, sie würden die Maschine jetzt mitnehmen. Ich hab sie nie wieder gesehen.
- SPRECHER: Ein geheimer Krisenstab wurde gebildet, der die Maschine untersuchen sollte – und sie abseits der Öffentlichkeit weiter befragte. Aber die Aussage, der Nordkorea-Konflikt laufe auf einen Einsatz von Atombomben hinaus, war in der Welt. Und hatte ein politisches Erdbeben ausgelöst.
- POLITIKER: *(Anmutung einer Pressekonferenz oder Rede, Fotoapparate klicken)* Sehr geehrte Damen und Herren, wir haben in dieser Woche in einer bisher beispiellosen Weise erlebt, wie Algorithmen sich in hochsensible zwischenstaatliche Gefüge einmischen können.

SPRECHER: Einigkeit darüber, wie eine angemessene Reaktion aussehen sollte, gab es keinesfalls.

POLITIKER: Das dürfen und können wir in einer Demokratie nicht zulassen!

SPRECHER: Das war natürlich ein Ansatz: Die Maschine ignorieren. Auch KI-Experte Rayid Ghani hätte das wohl empfohlen – wenn ihn damals schon jemand gefragt hätte.

RAYID GHANI: O-TON 10
[00:47:33] Let's say you predicted North Korea is going to launch....
Angenommen, die Vorhersage wäre, dass Nordkorea in den nächsten fünf Minuten eine Rakete starten wird. Sollen wir dann einen Präventivschlag machen? Das wäre doch eine Frage der Abwägung.
At what extra benefit is the cost worth it?
Bei welchem Nutzen ist es das Risiko wert?
And that's a human decision. That's not a machine decision.
Das ist eine Entscheidung, die Menschen treffen, nicht Maschinen. Das muss der Präsident entscheiden oder der Kongress oder das Volk.
So some policymaker whether with the president or it's the Congress or it's the public.

POLITIKER: Wir tragen die Verantwortung! Wenn es um Schicksale geht, wenn es darum geht, ob wir Gewalt anwenden, dann wollen wir keine künstlichen Systeme, die unsere Demokratie untergraben.

SPRECHER: Aber sich davon frei zu machen, das war schwierig. Zumal die neue Superintelligenz ja oft richtig zu liegen schien.

RAYID GHANI: O-TON 11
[00:51:31] If the computers are augmenting human experts then it's strictly better than...
Wenn Computer menschliche Experten beraten, dann ist das besser als ohne Computer. Wenn man das System richtig programmiert hat, dann ist ein

Mensch mit Computer besser als ohne Computer. Aber ein Computer allein ist nicht unbedingt besser als der Mensch.

...but a computer by itself is not obviously better than the human. [30.9]

SPRECHER: *(wiederholt:)* “Wenn das System richtig programmiert ist.” Das war natürlich der Kern der Angelegenheit. Denn es wusste ja noch immer niemand, wie sie überhaupt programmiert war. Niemand kannte die Daten, auf denen die Aussagen basierten und niemand kannte die Regeln, nach denen sie Vorhersagen generierte.

POLITIKER: Ich wurde gewählt, um das deutsche Volk zu repräsentieren. Aber wen repräsentiert dieses Computersystem?

SPRECHER: Auf beiden Seiten des Pazifiks jedenfalls

DIE MASCHINE: Es wird auf die nukleare Option hinauslaufen.

SPRECHER: .. spielte die düstere Prognose den Scharfmachern in die Hände. Die Krise verschärfte sich.

VERSCH. STIMMEN: Nationaler Sicherheitsrat tagt nonstop.

Russland schickt scharfe Warnung nach Pjöngjang.
Chinas Haltung weiter unklar.

Eine Pressekonferenz wurde kurzfristig wieder abgesagt.

SPRECHER: Auch Pazifisten und Diplomaten nahmen die Maschine ernst. Sie zogen aber andere Konsequenzen. Sie glaubten, das Eintreten der Prognose noch verhindern zu können und forderten statt Säbelrasseln umso dringlicher Verhandlungen. Nur Schlagzeilen machten sie damit kaum.

VERSCH. STIMMEN Flugzeugträger haben sich auf den Weg Richtung Ostchinesisches Meer gemacht. (Moderatorin)

UN-Sicherheitsrat ist zu einer weiteren Dringlichkeitssitzung
zusammengekommen

SZENE 7: DER MATHEMATIKER

- SPRECHER: Das änderte sich, als im Oktober der französische Mathematiker Stéphane Mallat ein viel beachtetes Radio-Interview gab.
- MODERATOR: Monsieur Mallat, auch wenn uns die Maschine rätselhaft ist, sie basiert doch auf Mathematik. Können wir ihr also glauben?
- STÉPHANE MALLAT: O-TON 12
[SM 00:09:57] We are in a situation where algorithms experiments are well ahead to theory...
(Vielleicht kann man es mehr so klingen lassen wie einen Simultan-Dolmetscher:) Wir befinden uns in einer Situation, in der Algorithmen in der Anwendung der Theorie weit voraus sind. Uns fehlt das mathematische Verständnis, warum sie so gut funktionieren. Algorithmen entwickeln sich viel schneller als unser Vermögen, sie wirklich zu verstehen.
...algorithms evolve much more more quickly than our ability to really understand it.
- MODERATOR: Wie ist das möglich? Wie kann jemand eine Maschine programmieren, ohne dass es mathematische Grundlagen dafür gibt? *
- STÉPHANE MALLAT: O-TON 13
[SM 00:11:42] There are things we understand there are a few basic principles...
Wir verstehen ein paar grundlegende Prinzipien, aber es gibt vieles, was wir nicht verstehen. Es gibt etwa keine schlüssige Theorie, die erklärt, welche Probleme neuronale Netzwerke lösen können und welche nicht. Es ist größtenteils noch nicht verstanden.
*...So it's mostly not understood. [32.4] **
- MODERATOR: Herr Mallat, **müssen** wir die Maschine denn verstehen, um sie ernst nehmen zu können?

STÉPHANE MALLAT: O-TON 14
[SM 00:31:21] I think like many people that it's a very powerful and hence potentially dangerous technology.
Ich bin wie viele Menschen der Meinung, dass es eine sehr mächtige und deshalb auch potenziell gefährliche Technik ist. Deswegen müssen wir sie so gut wie möglich verstehen. Sie muss von der Politik und gesetzlich kontrolliert werden. Klar, ich benutze auch ein Navi im Auto. Wir delegieren immer mehr an Maschinen, die es einfach besser machen als wir. Aber wie weit kann das gehen? Ich mache mir Sorgen, dass unsere Gesellschaft nicht realisiert, dass diese Technik genau wie Atombomben reguliert werden muss.
Exactly like nuclear bombs and therefore it has to be controlled and it's a worry I have. [25.9]

DIE MASCHINE: Stéphane Mallat. Hat eine Wikipedia-Seite in Englisch, Französisch, Deutsch, Tschechisch und Haitianisch.

SZENE 8: DER MENSCH-VERGLEICH

MARIE WINKLER: Ja, das Interview hab ich gehört. Fand ich ganz interessant. Aber im Grunde is es doch so: Was funktioniert, funktioniert halt, auch wenn ich das nicht verstehe. Ich versteh auch nicht, wie mein Kühlschrank funktioniert. Oder, noch besser: wie die EU funktioniert oder die Vereinten Nationen. Das sind auch undurchschaubare Systeme, die gewaltig was anrichten können.

STEFAN SEIDEL Willst du immer wissen, wer da im Einzelnen dahinter steckt?

MARIE WINKLER? Ich muss das nicht wissen. Aber sie treffen trotzdem wichtige Entscheidungen und funktionieren einigermaßen.

SPRECHER: Man konnte das natürlich so oder so sehen. Man konnte eine Verschwörung hinter der Maschine vermuten. Oder aber pragmatisch mit ihr umgehen und sich Teile ihrer brillanten Analysen zunutze machen. Noch nicht einmal von uns selbst wissen wir immer so genau, warum wir etwas machen. Können wir es dann von einer Künstlichen Intelligenz verlangen? Marie Winkler hat immer wieder darüber nachgedacht:

MARIE WINKLER: Warum. Ja, das habe ich mich schon öfter mal gefragt. Also, die Maschine hat mir ja schon geholfen. Ich weiß nicht, ob ich ohne sie da wäre, wo ich jetzt bin. Auch mit Stefan und Job und Studium und allem. Aber warum sie mir dazu geraten hat? Großes Rätsel..

SPRECHER: Diese eine, scheinbar ganz einfache Frage konnte die Maschine bis heute nicht beantworten: Warum?

(Soundeffekt einer Fehlermeldung auf einem Computer)

VERSCHIEDENE STIMMEN: Warum? Warum? Warum?

STEFAN SEIDEL Warum weißt du soviel?

DIE MASCHINE: Das musst du nicht wissen.

SZENE 9: DAS WARUM

SPRECHER: Für Rayid Ghani, der Data Science for Social Good betreibt, war das die entscheidende Frage. Er schrieb E-Mails, machte Anrufe, setzte sein ganzes politisches Netzwerk in Bewegung - und drang schließlich durch zum Pentagon und zum Weißen Haus. Was er dem Krisenstab dort versuchte zu erklären, war in etwa folgendes:

RAYID GHANI:

O-TON 15

[00:13:25] Whether the the model we used or is is can be explained to an expert or not is a very critical factor.

Dass das Modell sich einem Experten erschließt, ist entscheidend. Wir können nicht irgendwas benutzen, was wir nicht erklären können. Das wäre zu riskant.

SPRECHER:

Sein Standpunkt: Eine Vorhersage oder ein Ratschlag ohne Begründung ist nichts wert.

RAYID GHANI:

O-TON 16

[RG 00:43:30] A good explanation system makes the right prediction. Makes humans trust the right prediction more and makes humans mistrust the wrong predictions. [14.6]

Ein gutes System sorgt durch Begründungen dafür, dass Menschen richtigen Vorhersagen vertrauen und falschen Vorhersagen misstrauen.

[00:44:06] So if I'm making a wrong prediction when I explain their prediction that humans are going to say it makes no sense...

Bei falschen Vorhersagen sollten Menschen erkennen können, dass sie keinen Sinn ergeben. Dann können sie das korrigieren. Im Endergebnis trifft das Gesamtsystem aus Mensch und Maschine dann also die richtige Entscheidung, obwohl der Computer einen Fehler gemacht hat. Das ist die Macht eines solchen Systems: es verstärkt richtige Dinge und korrigiert falsche.

... the human was able to fix it by getting the right explanation. And that's the power of a good explanation system is it improves correct things and fixes the wrong things. [25.7]

SPRECHER:

Aber selbst bei der Maschine, der keinerlei Begründung zu entlocken war, musste gelten: Wenn sie etwas prognostizierte, dann nur indem sie den gegenwärtigen Kurs in die Zukunft fortsetzte. Wenn Menschen die Prognose nicht gefiel, hätte es in ihrer Macht gestanden, den Kurs zu ändern.

VERSCH. STIMMEN: Höchste Alarmstufe an der US-Westküste.
Diktator bringt Raketen in Stellung. (Schlagzeile)
Truppenbewegungen deuten auf bevorstehenden Einsatz hin.

SZENE 10: SCHLUSS

MARIE WINKLER: Ich war wahrscheinlich zu naiv. Mein erster Reflex war ja: Decke drüber
werfen und ja, am besten wegsperren oder so. Aber dann kam die
Faszination. Jetzt frag ich mich schon, ob ich das alles nochmal so machen
würde. Die ganze Menschheit hat sich von dem Teil verrückt machen
lassen.

Explosion

MODERATORIN Gerade erreicht uns eine Eilmeldung: Wir schalten uns live auf CNN

Mix aus Fehlermeldungs-Sounds, Sirenen, Explosionen steht länger frei.

ABSAGE: Die Maschine. Über das Dunkle in der Black Box.

DIE MASCHINE

ABSAGE: Dokumentation von Thomas Reintjes über die Macht künstlicher Intelligenz.
Wer waren unsere Berater?

DIE MASCHINE: Rayid Ghani, Universität Chicago; Joshua Kroll, Universität Berkeley;
Joanna Bryson, Universität Bath; Stéphane Mallat, ENS Paris; Joel Dudley,
Mount Sinai New York.

ABSAGE: Sprecher?

DIE MASCHINE:

ABSAGE: Ton und Technik?

DIE MASCHINE:

ABSAGE: Regie und Redaktion?

DIE MASCHINE: Friederike Wigger und Christiane Knoll

ABSAGE: Stimmt.

BEIDE: Eine Produktion des Deutschlandfunks 2017.