

Yuval Harari im grossen Interview: «Künstliche Intelligenz ist gefährlicher als der Klimawandel»

Rico Bandle

Der weltbekannte Historiker und Bestsellerautor glaubt, die neuen Technologien könnten die Grundlagen unserer Zivilisation zerstören. Der Mensch sei einfacher zu manipulieren, als er denke.



Publiziert: 29.04.2023, 20:57





Warnt vor Manipulation durch künstliche Intelligenz: KI-generiertes Porträt des Historikers und Bestsellerautors Yuval Noah Harari.

Quelle: Dall-E

Er wurde zum Weltstar, indem er auf völlig neue Art die Geschichte der Menschheit erzählte. Seit einigen Jahren schaut der Historiker Yuval Harari vorzugsweise in die Zukunft – und warnt davor, dass künstliche Intelligenz (KI) die Macht erlangen könnte. Diese Woche weilte er für einen Vortrag in Montreux. Der grosse Welterklärer erweist sich bei der Begegnung im Fairmont Palace Hotel als kleiner, schwächlicher Mann mit schüchternem Blick, der aber aufblüht, sobald er über jenes Thema sprechen kann, das ihn so sehr umtreibt: das Schicksal der Menschheit.

Herr Harari, Ihre Bücher begeistern Millionen Leser weltweit. Wie können wir sicher sein, dass Sie sie geschrieben haben und nicht ein Roboter?

Einzig daran, dass ich sie in der Vergangenheit publiziert habe. Bei meinem neuen Buch, das hoffentlich nächstes Jahr erscheint, wird das wohl bereits ein Problem sein. Wir sind dann an einem Punkt angelangt, an dem künstliche Intelligenz tatsächlich der Urheber sein könnte.

Dass KI Bücher zusammenfasst, kann man sich gut vorstellen. Aber wird sie auch neue Bücher schreiben können, mit eigenen, originellen Gedanken?

Noch nicht jetzt, aber das wird möglich sein. Genau das unterscheidet KI von allen bisherigen Technologien: dass sie autonom neue Ideen hervorbringen kann. Deshalb ist die Entwicklung von KI nicht mit anderen Erfindungen zu vergleichen. Der Buchdruck oder andere Informationstechnologien wie das Radio konnten nur die Ideen von Menschen kopieren und verbreiten, keine neuen hervorbringen. Leute sagen dann, auch KI sammle doch nur Ideen und kombiniere sie dann einfach auf eine neue Art. Aber das ist genau das, was auch Menschen tun. Für meine Bücher lese ich viele andere Bücher, nehme Ideen von hier und von dort, daraus entsteht dann etwas Neues.

Doch Kreativität beim Menschen kann doch auch von innen her kommen, aus den Gefühlen?

Aber auch der Mensch startet nie bei null. Es gibt keine Explosion im Hirn, die etwas völlig Neues herausbringt. In der Kunst oder der Musik baut alles auf früheren Werken auf. Ein Künstler versteht es, bereits vorhandene Muster zu verändern, zu durchbrechen oder zu kombinieren. Das ist dann seine neue Schöpfung. Die Möglichkeiten des Menschen sind allerdings begrenzt. Ein Komponist kann unmöglich alle Musikstücke hören, die je geschaffen wurden, um sich inspirieren zu lassen, sondern nur eine kleine Auswahl. KI hingegen kann das, deshalb wird sie den Menschen irgendwann auch in der Kunst überholen.

In Ihrem Buch «Homo Deus» schreiben Sie, das menschliche Hirn sei auch nur ein Algorithmus ...

Ich sagte nicht, dass dem so sei, sondern formulierte das als Frage. Ich bin mir nicht sicher. Die meisten Neurowissenschaftler allerdings sind tatsächlich überzeugt davon.

Was Sie sicher schrieben: Sollte das Hirn tatsächlich ein Algorithmus sein, so ist der freie Wille nur eine Illusion. Andererseits plädieren Sie dafür, die richtigen Entscheidungen zu treffen. Wie kann ich die richtige Entscheidung treffen, wenn ich keinen freien Willen habe?

Entscheidungen und Wille sind zwei unterschiedliche Dinge. Wir treffen ständig Entscheidungen. Der Glaube an den freien Willen nimmt uns die Neugier, herauszufinden, weshalb wir so entschieden haben und nicht anders. Man denkt immer: Das geschah aus freien Stücken. Wenn ich aber den freien Willen anzweifle, muss ich tiefer graben und berücksichtigen, dass biologische und kulturelle Kräfte meine Entscheidungen formen, ohne dass mir das bewusst ist. Vielleicht war meine Entscheidung ja beeinflusst durch Propaganda oder sogar durch einen chemischen Prozess in meinem Körper. Man weiss zum Beispiel, dass Menschen etwas migrationskritischer eingestellt sind, wenn man sie einem schlechten Geruch aussetzt.

Wie kommt denn das?

Das ist biologisch erklärbar: Eine kritische Haltung gegenüber Fremdem ist im Hirn an Ekelmechanismen gekoppelt. In fast jeder Kultur werden Fremde mit schlechtem Geruch assoziiert. Also dass die Leute schlecht riechen, ihre typischen Speisen oder andere Dinge. Da gibt es eine sehr tiefe biologische Verbindung. Das ist nicht freier Wille!

«Die am einfachsten zu manipulierenden Menschen sind jene, die an den freien Willen glauben.»

Wir sind also anfällig auf Beeinflussung.

Genau. Wenn ich den freien Willen bezweifle, bedeutet dies, dass wir viel mehr hinterfragen sollen, was unsere Entscheidungen wirklich prägt. Freiheit ist nichts, was wir haben, sondern etwas, wofür wir kämpfen müssen. Die Mechanismen zu verstehen, die unsere Meinungen und Entscheidungen beeinflussen, verleiht uns mehr Freiheit und gibt uns die Fähigkeit, bessere Entscheidungen zu treffen. Die am einfachsten zu manipulierenden Menschen sind jene, die an den freien Willen glauben, weil sie denken, dass nichts und niemand sie manipulieren kann.

Sie befürchten, dass die Manipulation durch KI zunehmen wird?

Dramatisch zunehmen!

Deshalb haben Sie mit über 1000 Forschern [den Brief unterschrieben](#), der einen sechsmonatigen Entwicklungsstopp für KI fordert, um verbindliche Regeln festzusetzen. Ist es nicht etwas naiv, zu glauben, man könne die Entwicklung aufhalten?

Ich glaube auch nicht, dass dies passieren wird. Es geht darum, Öffentlichkeit und Politik auf das Problem aufmerksam zu machen. Die KI-Revolution geht in rasendem Tempo voran. Wir brauchen aber Zeit, um zu verstehen, was da passiert und um Regeln aufstellen zu können. Vielleicht können wir die Entwicklung wenigstens etwas verlangsamen.

Was wäre Ihrer Meinung nach eine sinnvolle Regulierung?

Eine einfache Massnahme wäre, dass sich KI immer als KI zu erkennen geben muss, wenn sie mit Menschen in Kontakt steht. Wenn Sie online sind und mit jemandem diskutieren oder einen Tweet von jemandem lesen, so müssen Sie wissen, wenn das nicht von einem Menschen, sondern von einem Bot geschrieben wurde. Andernfalls ist die Demokratie am Ende. Denn Demokratie bedeutet Dialog, öffentlicher Dialog. Wenn nun KI in diesen Dialog eingreift und Menschen manipuliert, dann können Diktatoren und fremde Länder leicht die Demokratie zerstören. Es sollte klar sein: Meinungsfreiheit ist ein Menschenrecht, kein Computerrecht. KI hat keine Meinungsfreiheit. Entweder müssen wir die Bots aus Social Media rausschmeissen, oder sie sollten sich als solche zu erkennen geben müssen.

Letztes Jahr haben Sie im Vorwort zur neuen Auflage von «Homo Deus» geschrieben, wir würden es in Zukunft mit einer «wachsenden Schar von Avataren, Cyborgs und ausserirdischen Intelligenzen» zu tun haben. «Das nackte Überleben unserer Spezies könnte in Gefahr sein.» Das tönt für mich etwas gar nach Science-Fiction.

Steht da tatsächlich «ausserirdische Intelligenzen»? Das muss ein Übersetzungsfehler sein. Ich schrieb auf Englisch «alien intelligence», das bedeutet in diesem Zusammenhang nicht-menschliche Intelligenz und hat nichts mit Ausserirdischen zu tun.

Trotzdem: Das tönt sehr apokalyptisch.

Wir sind tatsächlich gerade mit drei apokalyptischen Szenarien konfrontiert: Die ökologische Krise mit der Klimaerwärmung. Die wachsenden internationalen Spannungen mit der russischen Invasion in der Ukraine und der Lage in Ostasien, die zu einem weiteren Weltkrieg führen könnten. Und die sich beschleunigende KI-Krise, die innerhalb weniger Jahre die Grundlagen der menschlichen Zivilisation zerstören könnte.

«Bisher formte der menschliche Geist, wie wir die Welt verstehen und wie wir uns verhalten. Nun wird das die KI übernehmen..»

Ist das jetzt nicht ein bisschen gar dramatisiert?

Schauen Sie, seit Tausenden von Jahren leben die Menschen in einer Welt, die von menschlicher Intelligenz und menschlicher Vorstellungskraft geprägt ist. Wir erleben die Realität nicht so, wie sie ist, sondern immer durch ein Prisma kultureller Produkte, die der menschliche Geist erschaffen hat – also der Religionen, der Gesetze, der Kunst, auch des Geldes, das ebenfalls ein kulturelles Produkt ist. Der menschliche Geist formt also, wie wir die Welt verstehen und wie wir uns verhalten. Nun wird das die KI übernehmen. Wir wissen nicht, was es bedeutet, in einer Welt zu leben, in der die meisten Geschichten, die meiste Musik, die meisten Fernsehsendungen, Gemälde, politischen Manifeste, sogar Religionen womöglich von einer nicht menschlichen Intelligenz geschaffen wurden. Wir haben nicht einmal die Werkzeuge, um über diese Situation nachzudenken.

Gibt es schon Anzeichen, wohin das führen kann?

Einer der wichtigsten Kulte aus den letzten Jahren ist QAnon. Ein sehr seltsamer Kult. Die Anhänger glauben, die Welt werde durch eine Verschwörung von Pädophilen oder was auch immer beherrscht. Es gibt Millionen von Menschen, nicht nur in den USA, die dem Kult folgen. Er hat mittlerweile sogar politische Macht erlangt, einige Sympathisanten sitzen im US-Kongress. Am Anfang dieses Kultes standen anonyme Onlinenachrichten. Ursprünglich wurden diese Texte von einem Menschen geschrieben, wir wissen nicht von wem, und durch Bots verbreitet. Aber in naher Zukunft könnte sich eine solche Religion um Texte herumbilden, die von einem Nichtmenschen verfasst wurden. Alle Religionen in der Geschichte behaupteten, dass ihre heilige Schrift aus einer nicht menschlichen Quelle stammt, etwa bei der Bibel oder dem Koran. Jetzt könnte das Realität werden.



«Der Facebook-Algorithmus hat von selbst entdeckt, dass die Verbreitung von Hass- und Angstbotschaften an Orten wie Myanmar die Nutzungszeit beträchtlich erhöht», erklärt Historiker Yuval Harari.

Foto: Oded Balilty (AP)

Weshalb ist das gefährlicher, als wenn es vom Menschen kommt? Es wurden schon viele Kriege im Namen der Religion geführt.

Eine KI kann durch Versuch und Irrtum ihre Fähigkeit optimieren, Menschen zu manipulieren. Daraus kann ein sehr überzeugender und inspirierender Text entstehen, der zur Grundlage eines

neuen Kults oder einer neuen Religion wird – inklusive Sextabus und politische Ansichten.

Muss es nicht immer einen Menschen geben, der der Maschine den Befehl gibt, eine Religion zu kreieren? Der Computer macht das ja nicht von sich aus.

In der Regel dürfte so etwas tatsächlich damit beginnen, dass ein Mensch Chat-GPT 5 oder einem anderen AI-Programm die Order gibt. Schon das tönt gefährlich. Aber es gibt auch Szenarien, in denen so etwas passiert, ohne dass ein Mensch es initiiert hat. Wir haben bereits bei Social-Media-Algorithmus gesehen, was für verheerende Konsequenzen Algorithmen ungewollt haben können.

Zum Beispiel?

Der Algorithmus von Facebook ist dahingehend programmiert, die Menschen möglichst lange auf der Plattform zu halten. Dieser Algorithmus hat von selbst entdeckt, dass die Verbreitung von Hass- und Angstbotschaften an Orten wie Myanmar die Nutzungszeit beträchtlich erhöht. Also hat er automatisch viel mehr solcher Nachrichten angezeigt. Mit verheerenden Folgen: Dies schürte den Hass zusätzlich, was mitverantwortlich war für die ethnische Säuberung des Rohingya-Volks. Das geschah allerdings völlig unbeabsichtigt: Niemand bei Facebook wollte die Gewalt in Myanmar anheizen. So etwas kann jederzeit wieder passieren. KI könnte zum Beispiel von selbst entdecken, dass eine neue Religion zu schaffen ein guter Weg ist, die Nutzungsintensität einer Plattform zu erhöhen.

Sie halten die KI langfristig als grössere Gefahr für die Menschheit als die Klimaerwärmung. Weshalb?

Bei der Klimafrage sind sich im Grundsatz alle einig: Wir wollen ein Klima- und Ökosystem, das die menschliche Zivilisation unterstützt. Ich habe noch nie jemanden sagen gehört, dass er es gut findet, wenn sich die Erde so stark erwärmt, dass dies die Lebensgrundlage gefährdet. Die einzige Differenz besteht darin, dass gewisse Leute sagen, sie glaubten nicht, dass es sich um eine echte Gefahr handelt oder dass sie nicht menschengemacht sei. Bei KI ist das komplizierter. Viele Leute wollen eine schnelle Entwicklung der KI vorantreiben. Es gibt keine Einigkeit darüber, was das ideale Ergebnis wäre. Die Träume der einen sind die Albträume der anderen. Zum Beispiel gibt es Diktatoren, die eine totalitäre KI herbeisehnen, die alles kontrollieren kann – ein Horrorszenario für Menschen, die an Demokratie und Freiheit glauben.

«Diktatoren haben schon immer davon geträumt, die Menschen zu verändern, auch die Nazis.»

KI ist dadurch schwieriger zu stoppen?

Es kommt noch ein wichtiger Unterschied hinzu: Der Klimawandel verändert die Menschheit nicht. Er bedroht uns von aussen, aber er ändert nichts an dem, was es bedeutet, Mensch zu sein. Die neuen Technologien wie KI und Bioengineering – also Eingriffe ins Hirn – werden wahrscheinlich die Menschheit komplett verändern. Wenn bestimmte Regierungen, Armeen oder Unternehmen die Macht erhalten, KI, Gentechnik, Gehirn-Computer-Schnittstellen direkt an den Menschen anzuwenden, so werden sie versuchen, nützliche Fähigkeiten wie Disziplin und

Intelligenz zu stärken. Andere menschliche Qualitäten wie Mitgefühl, künstlerische Sensibilität oder Spiritualität hingegen werden ignoriert.

Machthaber werden ihr Volk nach ihrem Gusto optimieren?

Das ist die grosse Gefahr. Was würde zum Beispiel Putin mit einer Gehirn-Computer-Schnittstelle anstellen? Er würde Soldaten wollen, die sowohl intelligenter als auch disziplinierter sind und sicher nicht solche, die mehr Mitgefühl haben und den Befehlen zur Ermordung von Zivilisten nicht gehorchen. Diktatoren haben schon immer davon geträumt, die Menschen zu verändern, auch die Nazis. Als 1945 der Zweite Weltkrieg zu Ende ging, hatten die meisten Menschen zwar ein Trauma, waren aber noch dieselben wie zu Kriegsbeginn 1939. Bei den Nazis der Zukunft wird das voraussichtlich anders sein. Selbst wenn sie besiegt werden, werden die veränderten Leute bleiben. Das ist in vielerlei Hinsicht ein beängstigendes Szenario, beängstigender als der Klimawandel.

Die Menschen bei uns beschäftigt wohl am meisten, ob ihre Jobs durch KI bald überflüssig werden. In «Homo Deus» sagen Sie, es werde eine «nutzlose Kaste» entstehen. Bisher sieht die Realität anders aus: Der Fachkräftemangel ist so gross wie noch nie.

Das Buch ist nun sechs Jahre alt, wir aber sprechen von Prozessen, die Jahrzehnte dauern. Und die global sind, nicht lokal. Wahrscheinlich werden die fortgeschrittenen Gesellschaften, die die Automatisierungsrevolution anführen, sogar eine höhere Nachfrage nach Arbeitskräften haben. Leiden werden die Entwicklungsländer mit ihren altmodischen Industrien, die sich leicht automatisieren lassen, zum Beispiel die Textilindustrie. Diese Länder verfügen auch nicht über die Ressourcen, um ihre Arbeitskräfte umzuschulen. Wir werden also wahrscheinlich global eine noch grössere Spaltung erleben, mit Ländern, die durch die Revolution reicher und mächtiger werden. Andere werden abgehängt und kollabieren.

Was raten Sie, was kann jeder Einzelne von uns unternehmen, um nicht bald zu den Überflüssigen zu gehören?

Niemand weiss, wie der Arbeitsmarkt in 20 Jahren aussehen wird. Die Welt verändert sich in immer höherem Tempo. Den Kindern sollten wir deshalb vor allem beibringen, wie man ein Leben lang lernt und sich ständig auf neue Situationen einstellt. Das ist viel wichtiger, als ihnen spezifisches Fachwissen wie Programmieren beizubringen. Die Leute müssen viel flexibler werden, ein Leben lang dazulernen, sich immer wieder neu erfinden.

Das sind hohe Ansprüche.

Es ist vor allem eine riesige psychologische Herausforderung, denn Veränderung ist immer eine Belastung. Auch die Politik sollte dies beschäftigen. Denn wenn Menschen mental gestresst sind und das Gefühl haben, abgehängt zu werden, dann sehnen sie sich nach einem autoritären Führer, der Ordnung schafft. Es ist eine enorm wichtige Aufgabe, den Menschen die Fähigkeit beizubringen, mit einer sich schnell verändernden Welt fertig zu werden.

Schon jetzt haben viele Jugendliche psychische Probleme.

Und wir stehen erst am Anfang. Das Tempo wird noch zunehmen, das Problem noch viel schlimmer werden.

Rico Bandle ist Redaktor bei der SonntagsZeitung und beschäftigt sich hauptsächlich mit gesellschaftspolitischen Themen.[Mehr Infos](#)

[@rbandle](#)

Fehler gefunden?[Jetzt melden.](#)