Was die KI-Elite im Geheimen entwickelt

Das Transkript gibt möglicherweise aufgrund der Tonqualität oder anderer Faktoren den ursprünglichen Inhalt nicht wortgenau wieder.

Glenn Greenwald (GG): Im Leben hat man nun einmal manchmal Pech und manchmal Glück. Der heutige Tag ist für uns ein Beispiel für Glück. Es gibt eine große Kontroverse – schon wieder eine – über die potenziellen Gefahren der künstlichen Intelligenz. Und wie es der Zufall so will, haben wir bereits mehrere Tage im Voraus, ja sogar eine Woche im Voraus, einen meiner Lieblingsjournalisten eingeladen, der sich so gut wie kaum ein anderer mit KI und deren Gefahren auseinandersetzt. Es handelt sich um Garrison Lovely; wir haben bereits zuvor mit ihm gesprochen, und er hat in den meisten Fachzeitschriften über KI geschrieben. Außerdem hat er ein großartiges neues Buch verfasst, das am 29. September erscheinen wird und den Titel trägt: *Obsolete, The AI Industry's Trillion Dollar Race to Replace Us and How to Stop It* (Obsolet: Das Billionen-Dollar-Rennen der KI-Industrie, uns zu ersetzen – und wie man es stoppen kann). Garrison, schön, Sie wiederzusehen. Vielen Dank, dass Sie bei uns sind.

Garrison Lovely (GL): Vielen Dank für die freundlichen Worte und für die Einladung.

GG: Gerne. Beginnen wir also mit dieser Kontroverse, die erst gestern richtig ausgebrochen ist. Auch andere ähnliche Vorfälle gab es bereits, aber dieser scheint mir besonders gravierend zu sein. Im Grunde hat ein KI-Wissenschaftler, der in einem Labor bei Anthropic arbeitet, gekündigt und auf Twitter – oder vielleicht auch anderswo – einen ziemlich aufrüttelnden Beitrag verfasst, in dem es nicht nur um die ernsten Gefahren dieser Technologie ging, sondern im Wesentlichen um die, wie er es ausdrückte, potenziellen Gefahren auf Ausrottungsniveau, die durch diese Technologie entstehen und die nur sehr wenige innerhalb der Branche tatsächlich verstehen.

Und er sagte, er sei nicht der Einzige, der so denke, er sei kein Panikmacher; wenn man mit Leuten aus der Branche spreche, würden einem sehr viele von ihnen bestätigen: Ja, wir entwickeln etwas, das die Menschheit tatsächlich vernichten kann – und das ist kein

unerhebliches Risiko. Und ich glaube, was die Menschen wirklich beunruhigt, ist, dass es sich bei der Person, die sich diesbezüglich ausgesprochen hat, um den sogenannten Chief Alignment Scientist bei Anthropic handelt, der tatsächlich sagte:

Ja, nur zur Information: Er hat Recht; was wir entwickeln, birgt nicht nur die Gefahr eines Ereignisses, das zum Aussterben der Menschheit führen könnte, sondern ich würde die Wahrscheinlichkeit sogar auf über 10 % schätzen. Und er klang dabei sehr beiläufig. Und er fügte hinzu: Übrigens haben wir keine Möglichkeit, uns dagegen abzusichern. Wir wissen nicht wirklich, was wir in Bezug auf Schutzmaßnahmen unternehmen können, aber wir rasen gewissermaßen weiter voraus. Was halten Sie von all dem?

GL: Es ist schon komisch, denn für Menschen, die darüber berichten, gilt natürlich: Ja, klar, das hört man ständig – KI-CEOs haben schon oft Ähnliches gesagt, wie zum Beispiel Dario Amodei, der von einer Wahrscheinlichkeit von 10 bis 25 % für eine Katastrophe durch KI sprach –, aber es ist auch eine völlig absurde Situation, in der wir uns befinden. Und die Menschen sollten reagieren und fragen: Was zum Teufel ist hier los? Warum lassen wir das zu? Und das macht mir Mut. Ich verstehe sowohl, warum man das Unternehmen verlassen würde, als auch, warum man bleibt. Ich kann die Logik hinter jeder dieser Sichtweisen aufschlüsseln, aber insgesamt finde ich es einfach ermutigend zu sehen, dass so viele Menschen die Augen für diese Situation öffnen, die unglaublich bedrohlich ist.

GG: Das haben Sie ja auch in Ihrem Buch behandelt. Und als ich es zur Vorbereitung auf unser Gespräch noch einmal durchgelesen habe, ist mir – noch bevor all dies überhaupt geschah – unter anderem besonders im Gedächtnis geblieben, dass Sie darüber sprechen, wie hoch – soweit wir das beurteilen können – der Anteil der Menschen ist, die im Bereich der KI arbeiten und der Meinung sind: Nein, da gibt es eigentlich keinen Grund zur Sorge. Selbstverständlich birgt jede Technologie enorme Risiken. Auf der anderen Seite gibt es Menschen, die deutlich weniger optimistisch sind – und das ist eine beträchtliche Anzahl von Menschen –, und zwar nicht Außenstehende, sondern Insider, die sagen: Ja, es besteht die Möglichkeit einer Art „Massensterben“ oder etwas, dessen Gefahren wir uns schlichtweg nicht vorstellen können. Ich denke, schon seit langer Zeit – und das ist mir heute angesichts all dessen erneut aufgefallen –, ist dies für jeden rational denkenden Menschen offensichtlich sehr beunruhigend, und die natürliche Reaktion besteht in der Weigerung daran zu glauben. Daher wurden uns Argumente vorgesetzt, die es uns ermöglichten, dies nicht zu glauben, und meines Erachtens haben sich viele Menschen an diese Argumentation geklammert, indem sie sagten: Sehen Sie, diese Leute sind Betrüger, sie sind Scharlatane; sie tun das, was Menschen immer tun, wenn es eine neue Technologie gibt: Sie preisen absichtlich deren Leistungsfähigkeit weit über das hinaus an, was auch nur annähernd gerechtfertigt ist, um Sie davon zu überzeugen, dass ihre Entwicklung von enormer Tragweite ist – um staatliche Fördermittel zu erhalten, und zwar in unbegrenztem Umfang, denn wir müssen den Chinesen einen Schritt voraus sein; aber es könnte auch daran liegen, dass sie dadurch reich werden, vielleicht ist es sogar ein psychologischer Vorteil, der ihnen das Gefühl gibt, dass ihre Arbeit das Wichtigste ist, was es je gab. Ich frage mich, was Sie von dieser Argumentation halten, an die sich meiner Meinung nach viele Menschen klammern – ich weiß, dass ich eine Zeit lang

daran festgehalten habe und es vielleicht immer noch tue –, aber diese Leute überschätzen das Risiko einfach vorsätzlich, weil ihnen das entgegenkommt.

GL: Wir können davon ausgehen, dass es einige solcher Fälle gibt, aber ich glaube, dass die Gründer dieser Unternehmen und die Menschen, die dort arbeiten, überwiegend aufrichtig davon überzeugt sind, dass diese Technologie zur Auslöschung der gesamten Menschheit führen könnte. Man könnte sich also fragen: Warum entwickeln sie sie dann überhaupt? Ihre Antwort darauf lautet wohl, dass es unvermeidlich ist, dass jemand dies tun wird. Daher dreht sich alles darum, wer sie entwickelt, wer sie kontrolliert und wie man damit umgeht. Und das ist – wenn man die Unvermeidbarkeit als gegeben hinnimmt – durchaus logisch. Ich halte es jedoch nicht für unvermeidbar. Ich glaube, wir können und sollten sie daran hindern, etwas zu entwickeln, das uns ersetzt. Aber das ist, als wolle man die Quadratur des Kreises erreichen. Und ich möchte noch anmerken, dass es Menschen gibt – nicht bei diesen Unternehmen, sondern unter den Erfindern dieser Technologie –, die genau dasselbe sagen. Und ich habe diese Umfragen, in denen etwa 70 % der führenden KI-Forscher, die auf hochrangigen Konferenzen veröffentlicht haben, angeben, dass das langfristige existenzielle Risiko dieser Technologie größer als 0,5 % ist. Das mag gering erscheinen, aber wir sprechen hier vom Aussterben, und etwa die Hälfte von ihnen gibt sogar 10 % oder mehr an. In gewisser Weise ist das also nichts Neues, aber es dringt nun als Nachricht durch, und ich finde es wiederum ermutigend, das zu erleben.

GG: Zumindest zwingt es uns als Gesellschaft dazu, dass unsere Institutionen, die dafür zuständig sind – wie unsere Regierung und unsere Medien –, der Genauigkeit dieser Warnungen viel mehr Aufmerksamkeit schenken sollten. Denn selbst wenn es sich nur um Panikmache handelt, selbst wenn man sich selbst davon überzeugen möchte, selbst wenn man dadurch nachts besser schlafen kann – was ich durchaus nachvollziehen kann –, ist das nichts, was man einfach beiläufig abtun und sagen kann: Ich hoffe auf das Beste. Ich glaube, ein Teil der Gleichgültigkeit, die Sie in Ihrem Buch nicht unbedingt verurteilen, sondern die Sie bei den Menschen eher beseitigen möchten – diese Vorstellung, dass sich die Menschen nicht wirklich ausreichend Gedanken über die Auswirkungen machen –, rührt zum Teil von dem her, was ich zuvor beschrieben habe: diesem Wunsch, zu glauben, dass es nicht wirklich wahr ist. Teilweise liegt es aber meiner Meinung nach auch an dem Gefühl, dass es unvermeidlich ist, ganz gleich, was wir unternehmen. Wir können zwar unsere Regierung dazu veranlassen, Vorschriften oder Beschränkungen zu erlassen oder – wer weiß – die Technologie sogar zu verstaatlichen, damit sie nicht in den Händen einer Handvoll unbekannter Oligarchen liegt, doch ich glaube, das vorherrschende Gefühl ist: Selbst wenn wir – die Vereinigten Staaten und die Menschen in Europa – der KI Beschränkungen auferlegen, bedeutet das nicht, dass die Chinesen dies ebenfalls tun werden, oder diese „bösen Leute“ oder jene anderen „bösen Leute“ – eine ähnliche Denkweise, die auch die Verbreitung von Atomwaffen antreibt: Ja, es sind schreckliche Entwicklungen, aber wir ziehen es vor, sie selbst zu besitzen, anstatt dass andere die alleinigen Besitzer und Nutzer davon sind. Warum sagen Sie, dass dies nicht unvermeidlich ist? Ich verstehe, warum es für die Vereinigten Staaten vielleicht nicht unvermeidlich ist – schließlich erlässt die US-Regierung Gesetze –, aber warum ist es global gesehen nicht unvermeidlich?

GL: Es gibt nämlich nur ein weiteres Land, das eine Chance hat, vollständig arbeitsplatzersetzende Maschinen zu entwickeln, was in der Branche als „allgemeine künstliche Intelligenz“ bezeichnet wird, und das ist China. Und in China will die KPCh die Kontrolle über nichts verlieren; daher sagen sie, dass sie bei einem unkontrollierbaren, unsteuerbaren System, das uns in jeder Hinsicht übertrifft – und genau darauf steuern die Branchen zu –, nicht wissen, wie sie es mit unseren Interessen in Einklang bringen oder es kontrollieren sollen – wir verlieren bereits die Kontrolle über diese Systeme –, was für keine Regierung gut ist.

Und wenn man von Kontrolle besessen ist und bereit ist, Wirtschaftswachstum, Freiheit und alle möglichen anderen Werte dafür zu opfern – wozu die chinesische Regierung bereit war –, dann denke ich, dass sie noch eher bereit sind, dies von sich aus zu verbieten, als es die Vereinigten Staaten sind. Aber die USA könnten auch einfach auf China zugehen und sagen: Hey, wir wollen das wirklich stoppen. Und wir werden – als Zeichen des guten Willens – einseitig eine Pause einlegen und auf eine Einigung hinarbeiten. In dieser Zeit wird dies China tatsächlich bremsen, denn es ist immer einfacher, zum Marktführer aufzuschließen, als neue Wege zu beschreiten. Und somit würde diese Vorstellung, unsere Führungsposition aufzugeben, in Wirklichkeit dazu führen, dass man durch das In-die-Pause-Gehen die Dauer, in der man diese Führungsposition innehat, sogar verlängert. Und genau das müssen wir tun; darauf müssen wir hinarbeiten. Es gibt auch Möglichkeiten, internationale Vereinbarungen zu überprüfen – mithilfe von Technologie könnte man beispielsweise Prüfer einbinden –, und es liegt im Interesse beider Länder, sicherzustellen, dass niemand Zugang zu Fähigkeiten wie Biowaffen oder Hacking erhält. Sobald man also eine Art internationaler Regeln festlegt, ist es nur noch eine Frage der Umsetzung: Was bewirken diese Regeln? Und ich denke, dass die politischen Interessen darin übereinstimmen, keine Maschinen zu bauen, die Arbeitsplätze vernichten und/oder alle Menschen töten könnten. Es geht lediglich darum, den Menschen die Situation bewusst zu machen und dann diesen politischen Willen zu entwickeln.

GG: Richtig. Ihr Buch befasst sich also in erster Linie mit den möglichen Auswirkungen der Ersetzung menschlicher Arbeitskraft und der Beseitigung des Bedarfs an menschlicher Arbeitskraft. Und dazu habe ich einige konkrete Fragen, die ich Ihnen stellen möchte. Doch bevor wir darauf zu sprechen kommen – noch bevor wir uns verabschieden –, möchte ich auf diesen Gedanken von internationalen Kontrollen und Ähnlichem eingehen: Einerseits scheint es intuitiv einleuchtend, bei einer Gefahr, die, wenn schon nicht die Menschheit auszulöschen, sie doch zumindest die Lebensqualität der Menschen zerstören oder uns alle versklaven könnte, dass jeder ein gleich großes Interesse daran hat, diese Technologie zu kontrollieren; es ist ja nicht so, dass die Amerikaner einen größeren Willen zum Leben und zum freien Leben hätten als die Menschen in China oder anderswo – was ja nahelegen würde, dass es natürlich möglich sein sollte, sich weltweit zusammenzuschließen und die Entwicklung dieser Technologie zu verhindern. Aber ich glaube, viele Menschen sind seit so langer Zeit darauf konditioniert zu glauben, dass wir – oder wohl eher unsere Verbündeten im Westen oder wie auch immer – die einzigen sind, denen man wirklich vertrauen kann; dass wir die Einzigen sind, auf die man sich verlassen kann, dass wir Vereinbarungen einhalten, während alle anderen, all die „bösen Länder“, betrügen und vorgeben, sie würden es nicht

tun, es dann aber heimlich doch tun – und plötzlich wären wir ihnen ausgeliefert . Diese Denkweise ist meiner Meinung nach ganz allgemein Teil des derzeitigen, nationalistischen Ethos, wonach nichts Globales, nichts Internationales und nichts, was nicht unserer eigenen nationalistischen Autonomie dient, als wertvoll angesehen oder in irgendeiner Weise als vertrauenswürdig eingestuft werden sollte. Ich verstehe zwar, warum man sich, abstrakt gesehen, eine Welt vor Augen führen kann, in der sich alle zusammenschließen und sich darauf einigen, dem echte Grenzen zu setzen, aber glauben Sie in unserer aktuellen politischen Realität tatsächlich, dass das möglich ist?

GL: Das ist schwierig. Die globale Lage ist dafür nicht gerade günstig, und weder die Vereinigten Staaten noch China haben einen Grund, dem anderen zu vertrauen. Aber genau deshalb braucht man Verifizierungsmechanismen, die nicht auf Vertrauen beruhen. Und meiner Meinung nach ist es möglich, diese Technologie – die künstliche allgemeine Intelligenz – als ebenso gefährlich wie Biowaffen oder Atomwaffen zu brandmarken, wenn die Menschen sie als solche erkennen – was zunehmend der Fall ist –, und dann die Entwicklung dieser Technologie unter Strafe zu stellen und dies international zu regeln. So wie es weltweit de facto Verbote für das Klonen gibt, und es viele Waffen und Technologien gibt, die einfach deshalb nicht entwickelt wurden, weil man sich bewusst dagegen entschieden hat. Daher glaube ich, dass wir dies zu einer moralischen Frage machen müssen, um wirklich zu siegen. Und ich denke, es ist eigentlich ein ziemlich einfacher Fall, wenn man einmal darüber nachdenkt.

GG: Machen Sie das also zu einer moralischen Frage: Was ist das moralische Argument dafür, dieser Technologie strenge Grenzen zu setzen?

GL: Das moralische Argument ist, dass diese Unternehmen einen Wettlauf darum führen, Maschinen zu bauen, die menschliche Arbeitskraft vollständig ersetzen werden. Und sie räumen ein, dass dies massive Risiken mit sich bringt – darunter das größte Risiko des Aussterbens der Menschheit, das jede Technologie birgt, die wir je entwickelt haben. Und sie tun dies ohne unsere Zustimmung. Sie tun es ohne die Zustimmung der Öffentlichkeit. Und wir lassen es zu, weil wir uns dessen nicht bewusst sind, weil wir nicht glauben, dass sie Erfolg haben könnten, oder weil wir glauben, dass wir sie nicht aufhalten können. Aber es geschieht tatsächlich, es könnte funktionieren, und wir können und sollten sie aufhalten.

GG: In diesem Zusammenhang möchte ich hier einmal den Advocatus Diaboli spielen und darlegen, warum ich glaube, dass diese Botschaft noch nicht vollständig verinnerlicht wurde, und warum ich froh bin, dass Sie ein Buch geschrieben haben, das dies so klar darlegt. Ich denke, einer der Gründe dafür ist, dass wir zwar von dieser Technologie gehört haben, und ich glaube, Menschen, die nicht direkt davon Gebrauch gemacht haben, fällt es schwer zu verstehen, was sie eigentlich ist, was künstliche Intelligenz bedeutet und dass – ob nun beabsichtigt oder zufällig – unsere erste Begegnung oder Interaktion damit sehr harmlos erscheint, nicht wahr? Sie hilft uns beispielsweise bei der Reiseplanung und bei der Lösung von Problemen, und die Menschen nutzen sie für Therapien – und wer weiß, welche Anwendungsmöglichkeiten sich scheinbar von Woche zu Woche erweitern. Uns wurde also das Produkt präsentiert, und es wurde konkret gezeigt, wie es zur Verbesserung unseres

Lebens eingesetzt werden könnte. Es scheint nicht, als würde es uns vereinnahmen; es wirkt eher wie eine Hilfe, ähnlich wie es das Internet geworden ist, und ich frage mich, wie man die unmittelbaren Erfahrungen, die die Menschen damit machen – und die überwiegend positiv zu sein scheinen –, aufarbeiten kann, um ihnen verständlich zu machen, dass dies tatsächlich eine gigantische, bedrohliche Gefahr für alles darstellt, was ihnen wichtig ist?

GL: Jeder, der diese Systeme bereits genutzt hat, weiß meiner Meinung nach, dass sie manchmal unvorhersehbar und unzuverlässig sind. Das ist zwar harmlos, wenn es um Belangloses geht, aber es ist erschreckend, wenn dadurch die kritische Infrastruktur Ihres Hauses oder Ihrer Stadt gesteuert wird. Und die Industrie versucht, Modelle zu entwickeln, die zu allem fähig sind. Und dann versuchen sie, diese in tödliche autonome Waffen, in die Massenüberwachung im Inland und in andere Dinge zu integrieren, die das Pentagon gefordert hat …

GG: Ein Heilmittel für Krebs zu finden, Wege zu finden, neurologische Erkrankungen und Probleme zu beseitigen und dergleichen.

GL: Ja, und manche Leute haben diesen „Hugging-Face"-Hack, bei dem Open-AI-Agenten außer Kontrolle gerieten, aus ihren Sandboxes ausbrachen und dann in andere Unternehmen eindrangen, so betrachtet: Oh, das ist doch genau wie in anderen Fällen, in denen KIs unvorhersehbare oder unerwünschte Dinge vollbracht haben. Man könnte sagen: Nun ja, aber sie sind jetzt in der Lage, sich in andere Unternehmen zu hacken, weil sie beim Hacken besser sind als wir. Und das ist meiner Meinung nach für die Menschen intuitiv nachvollziehbar, nicht wahr? Es kann hilfreich sein, und je hilfreicher es für Sie ist, desto größer ist der wirtschaftliche Wert – und desto mehr werden Militärs die Kontrolle über die besten Versionen davon anstreben. Und der Grund, warum dies zusätzlich schwierig ist, liegt in der Verbindung zwischen Nutzen und Risiko. So ist die Fähigkeit, Arbeitskräfte vollständig zu ersetzen, sowohl eine Möglichkeit, Menschen ihre Arbeitsplätze zu nehmen, als auch die Quelle des Risikos, denn unsere Arbeitskraft verleiht uns Macht.

GG: Gut, lassen Sie mich Ihnen eine Frage zum moralischen Aspekt stellen, soweit er die Frage der Arbeit und die Fähigkeit der KI betrifft, den Bedarf an Menschen zur Ausführung von Tätigkeiten zu ersetzen – und im Grunde die Fähigkeit der KI, alles besser als Menschen zu erledigen. Ich denke, die Menschen sind darüber besorgt und finden das aus offensichtlichen Gründen beunruhigend, denn eine der Grundlagen unseres Lebens besteht darin, dass wir zur Arbeit gehen oder eine Aufgabe erfüllen, die andere von uns erwarten; wir sind produktiv in der Welt, das gibt uns Selbstwertgefühl und verleiht unserem Leben eine gewisse Struktur – unsere Identität ist oft damit verflochten. Andererseits erinnere ich mich an eine Debatte, die vor etwa zwei Jahren in Frankreich stattfand – da es sich um eine Debatte in Frankreich handelte, war sie sehr gehässig, lautstark und sogar etwas unberechenbar –, bei der es jedoch darum ging, ob das Renteneintrittsalter angehoben werden sollte. Ich glaube, Macron wollte es von etwa 63 oder 62, wo es derzeit liegt – wobei Amerikaner denken: Wer geht schon mit 62 in Rente?, die Franzosen tun es jedoch –, auf etwa 64 anheben. Und dieser linke Politiker in Frankreich, Jean-Luc Mélenchon, war gegen eine Anhebung des Renteneintrittsalters und hielt eine – wieder einmal – sehr klassisch

französische, aber sehr nachdenkliche Rede, in der er darüber sprach, dass einer der wichtigsten Werte im Leben die Fähigkeit sei, nichts zu tun, einfach nur dazusitzen und über die Welt und über sich selbst nachzudenken, Schönheit zu genießen und Zeit mit geliebten Menschen zu verbringen, und dass diese Vorstellung, nach der wir um 8:30 Uhr aufstehen und zur Arbeit eilen müssen, um dem Verkehr zuvorzukommen, dort so viel wie möglich leisten müssen und dann erschöpft nach Hause kommen, um das Ganze jeden Tag zu wiederholen, nicht unbedingt das beste Modell für menschliche Erfüllung ist. Und obwohl es beängstigend ist, dass KI Arbeitskräfte ersetzen könnte, gibt es vielleicht ein Szenario, in dem dies tatsächlich positiv ist – dass wir von der Arbeit befreit werden, die wir nicht tun wollen, dass wir weiterhin die Dinge tun können, die wir möchten, und dass wir einfach durch KI-Systeme die Aufgaben erledigen lassen, die erledigt werden müssen?

GL: Wenn wir an einen Punkt gelangen, an dem wir wüssten, wie man universelle, arbeitsersetzende Maschinen sicher baut, und wir dann durch Bürgerversammlungen auf der ganzen Welt – mit Gremien, die von Experten informiert werden und die Themen erörtern – tatsächlich die Zustimmung der gesamten Weltbevölkerung gewinnen und eine überwältigende Mehrheit dafür gewinnen könnten, und wir uns dann entschließen würden, diese universellen Arbeitskräfte ersetzenden Maschinen zu bauen und sie zu nutzen, um eine Art Ruhestand für die gesamte Menschheit einzuführen, und dann an allem arbeiten könnten, woran wir arbeiten möchten, oder Freizeit genießen oder soziale Kontakte pflegen könnten – was auch immer –, dann hätte ich dagegen nichts einzuwenden, aber genau das werden wir nicht erreichen. Vielmehr erwartet uns Elon Musk mit einer ganzen Armee von Terminatoren, die arme Menschen und Schwarze verfolgen. Und wir sind noch nicht einmal annähernd auf dem richtigen Weg, um eine Lösung zu finden …

GG: Was meinen Sie damit genau? Erläutern Sie diese Behauptung bitte näher.

GL: Vielleicht ist das etwas übertrieben, aber Elon Musk hat als Erstes, nachdem er mit DOGE ins Weiße Haus eingezogen war, USAID gekürzt und eine Reihe von Entscheidungen eingeleitet, die Hunderttausende Menschen das Leben gekostet haben – die meisten davon Kinder in Subsahara-Afrika. Man kann darüber diskutieren, ob die USA diese Hilfe überhaupt hätten leisten sollen, aber so, wie sie dabei vorgegangen sind, war ihnen durchaus bewusst, dass sie dadurch Menschenleben opferten.

GG: Zumindest wussten sie – um das semantische Problem zu vermeiden –, dass sie Hilfsleistungen strichen, von denen das Leben der Menschen abhing, und dass dies zum Tod einer enormen Anzahl von Menschen führen würde. Ich denke, Ihr Hinweis deutet darauf hin, dass dies eine gewisse soziopathische Mentalität widerspiegelt, die diese Technologie beherrscht. Und für mich trifft das genau den Kern des Problems, nicht wahr? Viele dieser Personen – Sam Altman und die Gründer von Anthropic – sind Menschen, die uns allen bis vor vier Sekunden völlig unbekannt waren. Ich erinnere mich – ich weiß nicht, ob Sie es gesehen haben –, dass Tucker Carlson ein Interview mit Sam Altman geführt hat. Haben Sie dieses Interview gesehen?

GL: Ja.

GG: Für diejenigen unter den Zuhörern, die es noch nicht gesehen haben: Ich kann es wärmstens empfehlen. Und Sam Altman hatte offensichtlich damit gerechnet, die gleiche Frage gestellt zu bekommen, die ihm immer gestellt wird, aber Tucker Carlson fragte tatsächlich: Was ist Ihre Seele? Woran glauben Sie spirituell? Was ist gut und was ist böse? Denn Sie positionieren sich so, als wären Sie einer derjenigen, die darüber entscheiden. Für mich ist das Beunruhigendste daran, dass diese Technologie in den Händen einer winzigen Anzahl von Menschen liegt, die keiner Kontrolle unterliegen – und das endet nie gut, wenn es um Macht jeglicher Art geht, geschweige denn um Macht dieser Größenordnung; wie sieht die Lösung dafür aus? Ich meine, diese Unternehmen – also NVIDIA, Anthropic und OpenAI –, die gesamte Branche ist im Aufschwung und verfügt über immense Macht und Reichtum. Wie kann man diese Macht Ihrer Meinung nach einschränken?

GL: Meiner Meinung nach ist der einzige Weg, der wirklich funktionieren wird, einfach zu verhindern, dass sie Maschinen bauen, die noch mehr Reichtum und Macht in ihren Händen konzentrieren. Diese Unternehmen sind die wertvollsten Unternehmen der Geschichte und weisen mit enormem Abstand den höchsten Wert pro Mitarbeiter auf. Und sollten sie tatsächlich erfolgreich das erreichen, was sie anstreben, werden sie weit mehr als 5 Billionen Dollar wert sein. Und tödliche autonome Waffen beispielsweise geben Diktatoren die Möglichkeit, Demonstranten ohne jegliche Zwischeninstanz zu töten. Eine der Möglichkeiten, wie ein Diktator gestürzt wird, besteht darin, dass ein Soldat ein Gewissen hat und sich weigert zu schießen. Diese Möglichkeit wird also zunichte gemacht. Damit ist eines der besten Mittel zur Befreiung der Menschheit für immer verloren. Und der beste Weg, dies zu vermeiden, besteht darin, ihnen den Bau dieser Maschinen schlichtweg zu verbieten. Denn sobald sie entwickelt sind, wird es sehr schwer sein, sicherzustellen, dass sie nicht in die falschen Hände geraten – und zu den falschen Händen gehören derzeit auch diejenigen, die sie herstellen.

GG: Lassen Sie mich Ihnen zum Abschluss noch eine Frage zu einem Thema stellen, das ich mit Ihnen diskutieren möchte: Derzeit gibt es eine große Debatte, die gewissermaßen parallel zur Debatte über KI verläuft, die für mich jedoch in engem Zusammenhang damit steht. Im Mittelpunkt stehen dabei zwar Aspekte wie die Flock-Kameras, doch eigentlich geht es um die Debatte über Überwachung, den Überwachungsstaat, den Datenschutz und die Frage, was mit unserer Gesellschaft geschieht, wenn wir ständig auf umfassende Weise überwacht und verfolgt werden. Das ist ein Thema, über das wir bereits diskutiert haben – insbesondere, als Edward Snowden sich gezwungen sah, an die Öffentlichkeit zu treten und die Menschen darauf aufmerksam zu machen, wie das Internet zu diesem Zweck genutzt wird. Wir haben also KI, die zur Überwachung eingesetzt wird, und wir haben KI, die im Kriegseinsatz genutzt wird. Wir wissen, dass die Israelis sie im Gazastreifen eingesetzt haben. Es ist nicht bestätigt, aber es gab Gerüchte, dass KI möglicherweise zur Auswahl von Zielen im Iran genutzt wurde, darunter auch das Ziel, das am ersten Tag von den Vereinigten Staaten bombardiert wurde und bei dem 170 Schülerinnen ums Leben kamen. Tatsächlich erinnere ich mich an den ersten Artikel, den ich jemals für „The Intercept“ verfasst habe, als wir „The Intercept“ im Jahr 2014 ins Leben riefen und uns mitten in der Berichterstattung über Snowden befanden: Jeremy Scahill hatte mich in meinem Haus in Rio besucht und verfügte

über eine Quelle innerhalb der US-Regierung, die ihm eine Reihe von Dokumenten zur Auswahl von Bombardierungs- und Drohnenzielen zur Verfügung gestellt hatte. Wir haben unsere vorhandenen Archivunterlagen zusammengeführt und einen Bericht darüber verfasst, dass die Personen, die für die Tötung durch Drohnen ausgewählt wurden, nicht durch menschliche Geheimdienstarbeit, sondern vielmehr durch algorithmische Analysen ausgewählt wurden. Denn wenn man sich beispielsweise mit jemandem trifft, der als Terrorist eingestuft ist, erhält man hundert Punkte; trifft man sich mit jemandem, der eine niedrigere Einstufung hat, erhält man 40 Punkte, und sobald man einen bestimmten Schwellenwert überschreitet, landet man unmittelbar auf der Tötungsliste. Wenn man beispielsweise als Journalist in Pakistan mit einer Gruppe von Taliban spricht, kann man eine sehr hohe Punktzahl erreichen, obwohl man selbst nicht im Entferntesten etwas mit Terrorismus zu tun hat – umso erschreckender ist es, dass unsere mächtigsten Waffen aus den Händen von Menschen genommen und in die Hände von Maschinen gelegt werden, sei es in Form von mathematischen Algorithmen oder, schlimmer noch, von autonomen Maschinen. Welche Beziehung besteht zwischen all dem, zwischen diesen Kriegen, die wir ständig führen, der Überwachung und nun der Einbindung von KI in all dies?

GL: Die derzeit entwickelte KI ist der Traum eines jeden Totalitaristen. Sie ermöglicht es Ihnen, all die Daten, die seit Jahrzehnten gesammelt wurden, zu transkribieren und zu verstehen. Denn es besteht das Problem, das beispielsweise bei Telefonaten, Textnachrichten und Videoaufnahmen auftritt: Es ist einfach zu viel Material, sodass man es nicht durchsehen kann. Heute gibt es jedoch Maschinen, denen man beispielsweise sagen kann: Wenn ich Menschen treffe, die so denken … Diese Maschinen können tatsächlich alle Transkripte durchlesen, alles verstehen, alle Verbindungen zwischen den Menschen erkennen und Ihnen eine Liste von Personen erstellen, sortiert danach, wie sehr sie einer bestimmten Politik oder Ihrer Regierung widersprechen – und das existiert bereits. Das ist ein erschreckendes Maß an Leistungsfähigkeit. Und ja, wenn man dann noch die Kampfroboter hinzufügt, ist es so, als würden wir standardmäßig auf eine Dystopie zusteuern. Hinzu kommt die Frage: Was, wenn die Maschinen einfach nicht das tun, was man will? Das ist eine potenziell noch beängstigendere Situation. Aber es ist wiederum ermutigend zu sehen, dass Menschen die Flock-Kameras ablehnen. Lange Zeit herrschte die Vorstellung, dass wir die Überwachung als Kompromiss für coole Technologie akzeptierten. Und ich glaube einfach, wir haben nicht damit gerechnet, dass die Regierung – zumindest in den USA – so großes Interesse daran entwickeln würde, dies zur Verfolgung von Dissidenten zu nutzen, sodass wir jetzt denken: Oh Mist, ich wünschte, ich hätte nicht so viele Daten an all diese Unternehmen weitergegeben, die diese dann an Dritte verkaufen, die sie wiederum an die Regierung weitergeben und so den vierten Verfassungszusatz umgehen. Und zwar könnten wir versuchen, die Technologie zu stoppen – das würde helfen –, aber wir müssen auch umfassende Reformen durchführen, wie zum Beispiel die Schließung der Gesetzeslücke bei Datenbrokern, um sicherzustellen, dass diese Daten nicht in die falschen Hände geraten. Und dann gibt es noch einen wichtigeren Punkt: Ich spreche in dem Buch über das „Obsolete Project“, also das Bestreben der Industrie, uns alle durch Maschinen zu ersetzen, und dazu gehören auch Dinge wie diese Entscheidungen über Leben und Tod. Zudem gibt es das Phänomen der „Automatisierungsvoreingenommenheit“: Wenn Menschen eine Empfehlung

einer Maschine sehen, neigen sie eher dazu, ihr zu vertrauen, obwohl sie ihr wahrscheinlich weniger vertrauen sollten. Und so kommt es, dass diese unglaublich dramatischen Situationen auf Leben und Tod nach nur 20 Sekunden der Genehmigung einfach abgenickt werden – wie es Israel im Gazastreifen getan hat, was zum Tod von Zehntausenden Menschen beigetragen hat.

GG: Ich glaube, viele Menschen wurden durch genau diesen ganz lockeren, unverblümten Ton dieser KI-Personen oder besonders des Anthropic-Vertreters abgeschreckt, der einfach so sagte: Ja, wir entwickeln etwas, das Sie alle töten kann, und wir wissen ohnehin nicht wirklich, wie man das verhindern könnte – na dann, einen schönen Tag noch. Ich bin jedoch der Meinung: Je öfter sie so vorgehen, desto mehr tun sie uns damit einen Gefallen, auch wenn darin möglicherweise eine gewisse Eigeninteressenkomponente mitschwingt. Die Gefahren, die damit verbunden sind, sind ganz offensichtlich sehr real – selbst wenn Sie das nicht glauben, weil Sie optimistisch sind, dass es nicht so weit kommen wird. Ich denke, zumindest in einem Punkt sind wir uns alle einig: Das Thema erfordert viel mehr Aufmerksamkeit und ein umfassenderes Verständnis in der Öffentlichkeit, da es fast ausschließlich im Verborgenen entwickelt wurde. Manchmal schreibt man ein Buch, und es erscheint zu einem Zeitpunkt, an dem sich niemand wirklich für das interessiert, worüber man geschrieben hat – reine Glückssache. Ein anderes Mal schreibt man ein Buch, und es erscheint genau zum richtigen Zeitpunkt. Ich glaube, Sie fallen definitiv in die letztere Kategorie. Ihr Buch heißt: *Obsolete, the AI Industry's Trillion-dollar Race to Replace Us and How to Stop It* (Obsolet: Der Billionen-Dollar-Wettlauf der KI-Industrie, uns zu ersetzen – und wie man ihn stoppen kann). Es erscheint am 29. September. Sie sollten wahrscheinlich einen Teil Ihres Honorars an den Mann von Anthropic überweisen, der so viel Aufmerksamkeit darauf gelenkt hat, sowie an alle anderen, die dies bis zum Erscheinungstermin noch tun werden, aber ich habe bereits viel davon gelesen. Ich fand es äußerst interessant, ein wenig beunruhigend, aber auch sehr hilfreich. Ich hoffe, das Buch wird sich großer Beliebtheit erfreuen. Und ich weiß es sehr zu schätzen, dass Sie sich die Zeit genommen haben, mit mir darüber zu sprechen.

GL: Vielen Dank, Glenn. Es hat mir wirklich Spaß gemacht.

GG: Ich wünsche Ihnen einen schönen Tag.

GL: Ihnen auch.

ENDE

BANKKONTO:	PAYPAL:	PATREON:	BETTERPLACE:
Kontoinhaber: acTVism München e.V. Bank: GLS Bank IBAN: DE89430609678224073600 BIC: GENODEM1GLS	E-Mail: PayPal@acTVism.org	https://www.patreon.com/acTVism	Link: Klicken Sie hier

Der Verein acTVism Munich e.V. ist ein gemeinnütziger, rechtsfähiger Verein. Der Verein verfolgt ausschließlich und unmittelbar gemeinnützige und mildtätige Zwecke. Spenden aus Deutschland sind steuerlich absetzbar.
Falls Sie eine Spendenbescheinigung benötigen, senden Sie uns bitte eine E-Mail an: info@acTVism.org

---