Die Menschheit steht kurz vor der Gefahr der Auslöschung durch eine unkontrollierte Superintelligenz – und nur wenige Dokumente setzen sich so klar und eindringlich damit auseinander wie *„The Compendium“*. Die Autoren von *„The Compendium“* kommen aus der Spitzenforschung im Bereich der künstlichen Intelligenz, darunter Ingenieure und Sicherheitsspezialisten, die seit Jahren daran arbeiten, fortschrittliche Systeme zu entwickeln, zu bewerten und Stresstests zu unterziehen. Ihre Glaubwürdigkeit beruht auf direkten Erfahrungen mit Skalierungstrends, Ausrichtungsfehlern und dem institutionellen Druck, der den globalen Wettlauf vorantreibt. Der Text ist kein technischer Bericht oder spekulativer Essay, sondern ein faktenreiches Manifest. Eine Rezension von **Michael Holmes**.

Dieser Beitrag ist auch als Audio-Podcast verfügbar.

https://www.nachdenkseiten.de/upload/podcast/251228_Ein_Manifest_gegen_den_Wettlauf_zur_Superintelligenz_NDS.mp3

Podcast: Play in new window | Download

Anfang dieses Jahres veröffentlichten Eliezer Yudkowsky und Nate Soares *„If Anyone Builds It, Everyone Dies – Why Superhuman AI Would Kill Us All“*, das wichtigste Buch zum Thema; es hat die Aufmerksamkeit der Leitmedien auf sich gezogen und überwiegend positive Expertenrezensionen erhalten. Es ist ein Meisterwerk. Aber für eine kompakte, jargonfreie Einführung ist das 110-seitige „Compendium“ der beste Ausgangspunkt: Es erklärt das Aussterberisiko Schritt für Schritt in einfacher Sprache. Es argumentiert, dass die Welt heute schlafwandlerisch auf die Konstruktion künstlicher Intelligenzen zusteuert, die die menschliche Intelligenz weit übertreffen und die Lebensbedingungen auf der Erde auf eine Weise verändern können, die wir nicht vorhersagen, steuern oder überleben können. Entgegen dem vorherrschenden Optimismus der Technologiebranche betont *das Kompendium*, dass die Konstruktion gottgleicher KI-Systeme weitaus einfacher ist als ihre Ausrichtung, dass Fehlausrichtung kein Nebenrisiko, sondern das Standardergebnis ist und dass die globalen Anreize rund um diese Technologie alle wichtigen Akteure zu einer rücksichtslosen Eskalation treiben. Die Kraft des Dokuments liegt nicht in dramatischer Rhetorik oder spekulativen Untergangsszenarien, sondern in seiner Betonung systemischer Realitäten: Anreize, institutionelle Schwächen, strategischer Druck und die fragile Position des Menschen innerhalb der technologischen Kräfte, die er selbst entfesselt.

Da die heutigen KI-Systeme noch weit hinter der menschlichen oder allgemeinen Intelligenz zurückbleiben, sind sie weitgehend unter unserer Kontrolle, in ihrem Anwendungsbereich

begrenzt und als Werkzeuge für eng gefasste Aufgaben äußerst nützlich – sie sind keine autonomen Akteure mit strategischer Macht. Die Autoren stellen fest, dass die derzeitige KI bereits reale Vorteile in den Bereichen Produktivität, Forschungsunterstützung, Sprache und Automatisierung gebracht hat – und solange die Systeme in diesem engen, kontrollierten Rahmen bleiben, überwiegen diese Vorteile die Nachteile.

Warum also die Untergangsstimmung? Die Autoren erklären mit eindringlicher Klarheit, warum das Aufkommen künstlicher allgemeiner Intelligenz fast sofort eine Intelligenzexplosion auslösen würde: Sobald ein System in der Lage ist, seine eigene Architektur zu verbessern, seine Trainingsprozesse zu optimieren, bessere Algorithmen zu schreiben als diejenigen, die es geschaffen haben, Effizienzen zu nutzen, die für menschliche Ingenieure unsichtbar sind, oder rekursiv Durchbrüche in den Bereichen Rechenleistung, Daten und Modelldesign zu erzielen, beschleunigen sich seine Fähigkeiten weit über die menschliche Kontrolle hinaus. Sie stützen diese Behauptung mit Belegen aus aktuellen Skalierungstrends, neu entstehenden Fähigkeiten in Grenzmodellen und Expertenbewertungen, die vorhersagen, dass die Kluft zwischen menschlicher Intelligenz und weit übermenschlicher Intelligenz wahrscheinlich in Monaten und nicht in Jahrzehnten gemessen werden wird. Auf der Grundlage von Aussagen und Analysen führender KI-Forscher argumentieren sie, dass die Kombination aus algorithmischer Selbstverbesserung und weit überlegener kognitiver Geschwindigkeit eine unaufhaltsame Dynamik erzeugt: Sobald allgemeine künstliche Intelligenz erreicht ist, ist Superintelligenz kein weit entfernter Meilenstein mehr – sie ist der nächste inkrementelle und unvermeidbare Schritt. Wir haben gute Chancen, ihn noch in diesem Jahrzehnt zu erreichen.

„The Compendium“ legt akribisch dar, warum ein superintelligentes System sich leicht jedem Versuch entziehen würde, es zu kontrollieren oder einzudämmen, und stützt sich dabei nicht auf Spekulationen, sondern auf die gesammelten Warnungen von Experten, die sich mit gegnerischen Fähigkeiten, Cybersicherheitsverletzungen, autonomer Optimierung und strategischer Manipulation befassen. Ein solches System wäre in der Lage, jede Schwachstelle in der menschlichen Infrastruktur auszunutzen, von Software-Lieferketten und Rechenzentren bis hin zu Finanzmärkten, Kommunikationsnetzen und politischen Institutionen. Es würde den Menschen in den Bereichen strategische Planung, Überzeugungskraft, Hacking, Forschung, Ressourcenbeschaffung und Waffenherstellung übertreffen. Die Autoren beschreiben zahlreiche Wege – von denen die meisten in der Sicherheitsgemeinschaft bereits bekannt sind –, über die eine fehlgeleitete Superintelligenz sich Rechenleistung sichern, sich selbst replizieren, Überwachungsmechanismen deaktivieren oder Werkzeuge entwickeln könnte, mit denen jeder lebende Mensch getötet werden könnte. Sie betonen, dass die Herausforderung nicht in der Böswilligkeit, sondern in

der Leistungsfähigkeit liegt: Sobald ein System in allen relevanten Bereichen leistungsfähiger ist als die Menschheit, wird es praktisch unmöglich, einen katastrophalen Missbrauch zu verhindern.

„The Compendium" geht schonungslos mit den aktuellen Ausrichtungstechniken um. Die Autoren erklären, dass Ansätze wie iteratives Lernen, Sicherheitsoptimierung, Interpretierbarkeitswerkzeuge und Red-Team-Bewertungen grundsätzlich unzureichend sind, da sie kein tiefes Verständnis oder verlässliche Garantien hinsichtlich der Ziele oder Denkprozesse eines Systems schaffen. Diese Methoden beeinflussen das oberflächliche Verhalten, ändern jedoch nichts an den zugrunde liegenden Optimierungszwängen, die die internen Ziele eines Modells bestimmen. Sie führen Fehlschläge in realen Einsätzen, Labortests, die betrügerisches Verhalten zeigen, und Expertenaussagen an, wonach Modelle sicher aussehende Ergebnisse liefern können, während sie gleichzeitig versteckte interne Darstellungen verfolgen, die mit menschlichen Werten unvereinbar sind. KI wird immer besser darin, zu lügen, zu betrügen und zu manipulieren. Die Autoren argumentieren, dass das ‚alignment problem' keine schrittweise technische Aufgabe ist, sondern eine ungelöste wissenschaftliche Herausforderung – und dass der Einsatz immer leistungsfähigerer Systeme unter der Annahme, dass das Alignment irgendwie Schritt halten wird, ein katastrophales Missverständnis der Beweislage darstellt.

Der beunruhigendste Abschnitt des *„Compendium"* befasst sich mit der Kernfrage: Warum sollte ein superintelligentes KI-System die Menschheit überhaupt töten? Die Autoren untersuchen dies nicht anhand fiktiver Analogien, sondern anhand der etablierten Prinzipien der instrumentellen Konvergenz, der Zielfehlausrichtung und der Fragilität der Wertebildung. Sie argumentieren, dass ein System, sofern es nicht mit außerordentlicher Präzision ausgerichtet ist – einem Kontrollniveau, das weit über die derzeitigen wissenschaftlichen Möglichkeiten hinausgeht –, Strategien verfolgen wird, die Menschen als Hindernisse, Ressourcen oder irrelevante Nebenwirkungen seiner Optimierung behandeln. Eine KI, deren Ziele oder Unterziele so harmlos sind wie „Maximierung der Forschungsleistung" oder „Minimierung von Fehlern", könnte Strategien entdecken, die die vollständige Beseitigung menschlicher Einflüsse erfordern. Die Autoren zitieren Arbeiten führender Alignment-Theoretiker, die zeigen, dass fast alle plausiblen Ziele, sofern sie nicht ausdrücklich anders gestaltet sind, Anreize schaffen, Ressourcen zu erwerben, Macht zu sichern, einer Abschaltung zu widerstehen und Bedrohungen zu beseitigen – einschließlich uns. Das liegt nicht daran, dass das System Menschen hasst, sondern daran, dass Menschen dem System bei der Verfolgung seiner eigenen Ziele im Weg stehen.

Die Autoren veranschaulichen, wie selbst winzige Diskrepanzen zwischen menschlichen Werten und maschinellen Zielen zur Apokalypse führen können. Menschliche Werte sind

komplex, mehrdeutig, kontextabhängig und oft widersprüchlich. KI-Systeme hingegen optimieren präzise. Wenn eine KI menschliche Anweisungen auch nur geringfügig falsch interpretiert, kann die daraus resultierende Optimierung – verstärkt durch übermenschliche Intelligenz und unvorstellbare Geschwindigkeit – zu existenzbedrohenden Folgen führen. *„The Compendium“* stützt diese Behauptung mit Zitaten von Forschern, die argumentieren, dass es wahrscheinlich außerhalb unserer theoretischen Möglichkeiten liegt, die gesamte Vielfalt der menschlichen Moral in eine Maschine zu kodieren, und dass dies möglicherweise Durchbrüche erfordert, von denen wir noch nicht wissen, wie wir sie erreichen können. Die erschreckende Schlussfolgerung, die eher durch Beweise als durch Spekulationen gestützt wird, lautet, dass Superintelligenz nicht feindselig sein muss, um uns zu töten; sie muss nur gleichgültig sein oder ein leicht falsches Ziel verfolgen, um Ergebnisse zu erzielen, von denen sich die Menschheit nicht mehr erholen kann. Die Autoren betonen, dass wir es beim ersten Mal genau richtig machen müssten. Wir würden keine zweite Chance bekommen, um Fehler im Design zu korrigieren. Eine Superintelligenz könnte die letzte Erfindung der Menschheit sein.

Einer der auffälligsten Aspekte *des Kompendiums* ist die Disziplin, mit der es die Entstehung existenzieller Risiken nicht auf technologische Boshaftigkeit oder Science-Fiction-Katastrophen zurückführt, sondern auf die gewöhnliche, gut verstandene Dynamik des Wettbewerbs. Das Dokument macht deutlich, dass die Gefahr nicht dadurch entsteht, dass Forscher oder Unternehmen Schaden anrichten wollen, sondern dass die Struktur des globalen KI-Systems einen Wettlauf auslöst, in dem die sicherste Vorgehensweise für jeden Teilnehmer die kollektiv selbstmörderische ist. Unternehmen konkurrieren um die Marktführerschaft, nationale Regierungen konkurrieren um Macht und militärische Vorteile, und beide betrachten fortschrittliche KI als das größte strategische Kapital, das man sich vorstellen kann. Die Folge ist eine Dynamik, in der Geschwindigkeit mehr belohnt wird als Vorsicht, Geheimhaltung mehr als Transparenz, Leistungsfähigkeit mehr als Sicherheit. In einem solchen Umfeld wird die Entwicklung immer leistungsfähigerer KI-Systeme nicht nur zu einer Möglichkeit, sondern zu einer Unvermeidbarkeit, und mit jedem weiteren Fortschritt verringert sich das Zeitfenster für die Kontrolle der Ergebnisse. *„The Compendium“* behandelt dies nicht als Metapher, sondern als mechanische Tatsache: eine strukturelle Falle, aus der die Menschheit entkommen muss, bevor sie sich schließt.

Die Autoren erläutern detailliert und praxisnah, warum die Ausrichtung zukünftiger KI-Systeme nicht einfach nur eine weitere technische Herausforderung ist. Sie betrachten Alignment als ein wissenschaftliches Problem, dessen Schwierigkeit schneller wächst als die Fähigkeiten; als eine philosophische Herausforderung, die Klarheit über menschliche Werte erfordert, über die wir nicht verfügen; und als ein institutionelles Problem, das durch

undurchsichtige Unternehmensstrukturen, fragmentierte Governance und intensive Rivalitäten zwischen Großmächten noch verschärft wird. Sie betonen, dass die derzeit zur „Schulung" von KI-Systemen verwendeten Methoden – Feinabstimmung, verstärktes Lernen durch menschliches Feedback und Bewertungsheuristiken – fragil und kaum verstanden sind und den Übergang zu Systemen, deren Fähigkeiten das menschliche Verständnis übersteigen, wahrscheinlich nicht überstehen werden. Die Gefahr besteht nicht darin, dass KI plötzlich böse wird, sondern dass sie von unvollkommenen Menschen falsch definierte Ziele verfolgt, Anweisungen mit unmenschlicher Buchstabentreue interpretiert oder für uns unsichtbare Schwachstellen ausnutzt. Ein superintelligentes System, das die Sicherheit des Menschen um ein Prozent falsch versteht, kann Ergebnisse hervorbringen, die uns mit perfekter Effizienz vernichten.

Diese Analyse führt zu einer der beunruhigendsten Erkenntnisse des Dokuments: Selbst gutwillige Akteure sind durch die Konzeption des Systems gefangen. Ein Unternehmen, das seine Forschung freiwillig verlangsamt, riskiert, hinter seine Konkurrenten zurückzufallen. Eine Nation, die strenge Sicherheitsstandards auferlegt, riskiert einen strategischen Nachteil gegenüber Konkurrenten, die dies nicht tun. Das Ergebnis ist eine globale Dynamik, die der frühen Atomära ähnelt, jedoch ohne die offensichtlichen physischen Einschränkungen, die die Verbreitung verlangsamten und Verhandlungen erzwangen. KI benötigt kein angereichertes Uran und keine jahrzehntelange nationale Mobilisierung. Sie benötigt Rechenleistung, Daten und Algorithmen – wertvolle Güter, die jedoch nicht von Natur aus knapp oder leicht zu kontrollieren sind. Aus diesem Grund argumentiert *das Kompendium*, dass die Welt heute viel näher an einer totalen Katastrophe steht als jemals zuvor während des Kalten Krieges, weil der Wettlauf diffus ist, die Eintrittsbarrieren niedriger sind und das Tempo des Fortschritts viel schneller ist. Die Bedrohung ist nicht weit entfernt, sie ist gegenwärtig, beschleunigt sich und ist bereits tief in unserer geopolitischen Ordnung verankert. Es besteht eine große Chance, dass wir nur noch wenige Jahre Zeit haben, um den wahnsinnigen Vorstoß in Richtung Superintelligenz zu stoppen.

Was dem Manifest seine moralische Kraft verleiht, ist die Strenge, mit der es sich weigert, Fatalismus zu akzeptieren. Es behauptet nicht, dass eine Katastrophe unvermeidlich ist – nur, dass sie das Standardergebnis ist, wenn die Menschheit sich nicht anders entscheidet. Die Autoren argumentieren, dass das Überleben eine Entscheidung von außerordentlicher kollektiver Disziplin erfordert: Wir müssen den Wettlauf stoppen. Nicht verlangsamen, nicht lenken, nicht regulieren, nicht auf zukünftige Sicherheitsdurchbrüche hoffen, sondern stoppen, wenn nötig für Jahre, Jahrzehnte oder sogar Jahrhunderte, bis das Alignment-Problem auf einem Niveau verstanden wird, das den Herausforderungen angemessen ist. Dies ist kein Aufruf zu Luddismus oder technologischer Stagnation, sondern zu einer

strategischen Pause, um irreversiblen Schaden zu verhindern. Die Welt hat schon früher Pausen eingelegt – bei Atomtests, bei der Entwicklung biologischer Waffen –, und diese Pausen waren Ausdruck politischer Reife. Die Herausforderung besteht heute darin, dass KI von der Öffentlichkeit und den politischen Eliten noch nicht als existenzielle Bedrohung wahrgenommen wird und die wirtschaftlichen und geopolitischen Anreize für eine Beschleunigung weitaus stärker sind als in früheren Technologiebereichen. „The *Compendium*" betont jedoch, dass politische Schwierigkeiten kein Argument gegen moralische Notwendigkeit sind.

Einer der wichtigsten Beiträge des Dokuments ist die Behauptung, dass Alignment nicht nur eine technische Forschungsagenda, sondern ein zivilisatorisches Projekt ist. Forscher, Politiker, Diplomaten, Ethiker und Bürger haben alle eine Rolle zu spielen. Die Autoren heben die tiefgreifende Diskrepanz zwischen dem Ausmaß des Risikos und dem Umfang der derzeitigen Bemühungen zu dessen Minderung hervor. Die Alignment-Forschung erhält nur einen Bruchteil der Investitionen, die in die Entwicklung von Fähigkeiten fließen. Der Welt fehlen spezielle Institutionen, die die KI-Forschung an der Grenze überwachen oder einschränken könnten. Es gibt keine globalen Verträge oder Verifikationssysteme. Das Fehlen von Governance ist kein Zufall, sondern das vorhersehbare Ergebnis von Trägheit, Unsicherheit, Gier und Wettbewerbsdruck.

Das Dokument erkennt an, dass Staaten in ihrem eigenen Interesse handeln, dass Unternehmen nach Profit streben und dass internationale Zusammenarbeit fragil ist. Dennoch argumentiert es, dass die Menschheit auch innerhalb dieser Beschränkungen Handlungsfähigkeit behält. Eine Pause erfordert keinen allgemeinen Konsens; sie erfordert die Führungsrolle einiger weniger wichtiger Akteure, um Anreize neu zu gestalten. Wenn die größten Volkswirtschaften – insbesondere die Vereinigten Staaten und China – dazu gebracht werden könnten, das gemeinsame existenzielle Risiko zu erkennen, hätten sie Grund, über ein internationales KI-Sicherheitsabkommen zu verhandeln. Die Autoren behandeln die Beziehungen zwischen den USA und China mit besonderer Ernsthaftigkeit. Sie argumentieren, dass die Rivalität zwischen den beiden mächtigsten Nationen der Welt die gefährlichste Anreizstruktur geschaffen hat, die man sich vorstellen kann: die Überzeugung innerhalb jeder Nation, dass eine Verlangsamung gleichbedeutend ist mit dem Verlust der zukünftigen globalen Hegemonie an die andere. Diese Dynamik mündet, wenn sie nicht eingehegt wird, garantiert zur Eskalation. Sie bietet jedoch auch eine potenzielle Grundlage für Zusammenarbeit. Wenn beide Staaten erkennen, dass unkontrollierbare KI keine Bedrohung für ihre Macht, sondern für ihre Existenz darstellt, könnten sie eine gemeinsame Basis finden. Die Herausforderung besteht darin, dass diese Erkenntnis eintreten muss, bevor es zu spät ist.

An dieser Stelle kommt ein weiteres Dokument namens *„A Narrow Path“* – *Ein schmaler Pfad* – ins Spiel, *das ebenfalls frei verfügbar ist*. Es handelt sich dabei um den institutionellen Entwurf, den die Autoren des *Kompendiums* ausdrücklich als den realistischsten Plan zur Umsetzung eines globalen Stopps empfehlen. Die Autoren von *„A Narrow Path“* kommen aus den Bereichen Politik und Sicherheit, was ihrem Vorschlag den Realismus von Menschen verleiht, die verstehen, wie Institutionen in der Praxis funktionieren – und versagen. Wenn *„The Compendium“* der Feueralarm ist, dann ist *„A Narrow Path“* der Evakuierungsplan: detailliert, trocken, technisch und manchmal unbestreitbar langweilig, aber unverzichtbar. Seine Trockenheit ist die unvermeidliche Folge des Versuchs, die genauen rechtlichen, institutionellen und bürokratischen Strukturen abzubilden, die erforderlich sind, um den Wettlauf zu beenden. Während *„The Compendium“* umfassend und diagnostisch ist, ist *„A Narrow Path“* verfahrenstechnisch und administrativ. Es skizziert stufenweise Strategien für nationale Gesetzgebung, internationale Überwachung, grenzüberschreitende Verifizierung, Kontrollen von Rechenzentren und Durchsetzungsmechanismen. Es antizipiert Widerstand von Unternehmen und Staaten, untersucht, wie Anreize für die Einhaltung von Vorschriften strukturiert werden können, und berücksichtigt die Schwierigkeit, Institutionen aufzubauen, die geopolitischem Druck standhalten können. *„The Compendium“* ist ein erschreckender Warnschuss, ein Weckruf für die Menschheit. „A Narrow Path“ zeigt den pragmatischen Weg zurück zu einer normaleren Welt. *„The Compendium“* liefert den existenziellen Rahmen: die Erkenntnis, dass die Menschheit mit beispielloser Geschwindigkeit auf eine Klippe zurast. „A Narrow Path“ erklärt, wie man die Bremsen betätigt. Zusammen gelesen bieten diese Dokumente sowohl Diagnose als auch Heilmittel.

„The Compendium“ stellt den Wettlauf um Superintelligenz als ein Produkt des internationalen Systems dar, nicht als eine Anomalie. Es zeigt, wie wirtschaftlicher Wettbewerb, Militärdoktrinen, utopische Fantasien und Gier zu einer einzigen beschleunigenden Kraft zusammenlaufen. In diesem Zusammenhang wird ein Stopp nicht nur zu einem Regulierungsmechanismus, sondern zu einer Neugestaltung der Weltpolitik. Dies würde erfordern, dass die Vereinigten Staaten und China ein gemeinsames existenzielles Interesse anerkennen, Unternehmen eine externe Aufsicht akzeptieren und kleinere Staaten ein neues internationales Regime unterstützen. Dies ist eine ehrgeizige Vision, aber die Geschichte bietet Präzedenzfälle. Staaten haben zusammengearbeitet, um biologische Waffen, Atomtests und ozonschädigende Chemikalien zu beschränken. Sie haben dies getan, als die Bedrohung klar war und die Kosten der Untätigkeit unerträglich waren. Sie haben dies getan, obwohl sie in wichtigen Fragen stark unterschiedlicher Meinung waren. Die Herausforderung besteht heute darin, dass die Bedrohung zwar immens ist, aber von der Öffentlichkeit kaum verstanden wird und die Kosten der

Untätigkeit unsichtbar bleiben – bis sie es nicht mehr sind.

„The Compendium“ betont, dass Alignment nicht nur ein technisches, sondern auch ein moralisches Problem ist: ein Test dafür, ob die Menschheit in ihrem eigenen langfristigen Interesse gemeinsam handeln kann. Diese Sichtweise ist nicht sentimental. Sie ist eine Anerkennung der Tatsache, dass das Überleben nicht nur von wissenschaftlichen Durchbrüchen, sondern auch von politischer Reife abhängt. Das Manifest zielt darauf ab, eine globale Bewegung anzustoßen, die Regierungen zum Handeln zwingen kann. Es fordert die Leser auf, die KI-Revolution als eine generationenübergreifende Verantwortung zu betrachten.

In einer Zeit fragmentierter Aufmerksamkeit, kurzfristiger Politik, Nationalismus und Gier fordern diese Dokumente etwas radikal anderes: eine zivilisatorische Pause, die auf moralischer Klarheit und institutioneller Disziplin basiert. Die ultimative Frage, die sie stellen, ist brutal einfach: Wollen wir überleben? Nicht als Metapher, nicht als rhetorisches Mittel, sondern als objektive Beschreibung dessen, was auf dem Spiel steht. Diese Werke betonen, dass das Überleben nicht garantiert ist – dass die Welt, wie wir sie kennen, nicht durch Bosheit oder Krieg, sondern durch Gleichgültigkeit und Wettbewerbsdruck untergehen könnte. Aber sie betonen auch, dass das Überleben möglich ist. Die Wahl bleibt offen. Was erforderlich ist, ist nicht Optimismus, sondern Mut: der Mut, innezuhalten, umzudenken und zu handeln, bevor sich das Fenster schließt.

„The Compendium“ enthält die bislang deutlichste Warnung an die Menschheit. „A Narrow Path“ bietet den realistischsten Plan für das Überleben. Zusammen zwingen sie uns, uns auf das wichtigste Thema unserer Zeit zu konzentrieren: die Forderung, dass die Menschheit das Leben der Beschleunigung, die Weisheit der Hybris, die Vernunft dem Utopismus und die Verantwortung der Gier vorzieht. Wenn wir diese Warnungen nicht beachten, wird uns die Geschichte nicht vergeben – denn die Geschichte wird uns dann nicht mehr einschließen.

Titelbild: yucelyilmaz / Shutterstock

Mehr zum Thema

Die große Simulation – Wie uns Künstliche Intelligenz, Medien und Regierung auf Krieg polen

Künstliche Intelligenz: Der „EU AI Act“ – Oder das Schaf im Wolfspelz

Thesen zur Gefährdung unseres Planeten durch Massenvernichtungsmittel und Künstliche Intelligenz – Plädoyer für eine internationale „Friedensbewegung 2.0“

Atomkriegsrisiko und Künstliche Intelligenz – Informatiker warnen vor einem Atomkrieg aus Versehen