

Computer im Unternehmen English Electric in Whetstone 1958: Hin- und hergerissen zwischen Erlöschung und Erlösung

Menschenwürmer

Künstliche Intelligenz Müssen wir damit rechnen, dass eines Tages die Maschinen schlauer sind als wir? Ja. Müssen wir Angst davor haben? Nicht unbedingt. *Von Georg Diez*

Es gibt das Gefühl, dass die Menschheit an einem Wendepunkt steht. Ein Unbehagen, ein Erzittern, ein Unwissen, das sich verdichtet. Ahnungen und Warnungen. Und Szenarien, was es für die Menschen bedeutet, wenn die Maschinen sie beherrschen.

Google also und Facebook und Amazon, all das, worüber sich so viele aufregen, was sie kritisieren und bekämpfen – es wirkt trivial im Vergleich zu dem, was längst gedacht wird: der Untergang der Spezies.

Stephen Hawking etwa, der in der vergangenen Woche Schlagzeilen machte, als er knapp skizzierte, wie die Apokalypse 2.0 kommen könnte: nicht durch Krieg, sondern durch Computer.

„Wenn die Künstliche Intelligenz erst vollständig entwickelt ist“, sagte Hawking der BBC, „dann könnte das das Ende der Menschheit bedeuten. Die Künstliche Intelligenz würde gleichsam sich selbst betriebsbereit machen und sich immer schneller und schneller verändern und entwickeln.

Die Menschen wären durch die langsame biologische Evolution begrenzt, sie könnten nicht mithalten und würden ersetzt.“

Das Ironische an der Sache war, dass Hawking, der weltberühmte Physiker und Bestsellerautor, der wegen einer Nervenkrankheit fast vollständig gelähmt ist und selbst nicht sprechen kann, diese Sätze recht flüssig sagen lassen konnte, weil er seinen neuen Sprachcomputer vorstellte: Das System basiert auf einem Algorithmus, der ausgeht von dem, was Hawking in der Vergangenheit gesagt hat, und daraus Prognosen darüber trifft, was er gerade sagen will.

Mit anderen Worten: Hawking warnte mit den Möglichkeiten der Künstlichen Intelligenz vor den Folgen der Künstlichen Intelligenz. Es gibt kein besseres Bild für den Grundwiderspruch unseres technologischen Zeitalters, hin- und hergerissen zwischen Erlösungs- und Erlöschungsfantasien.

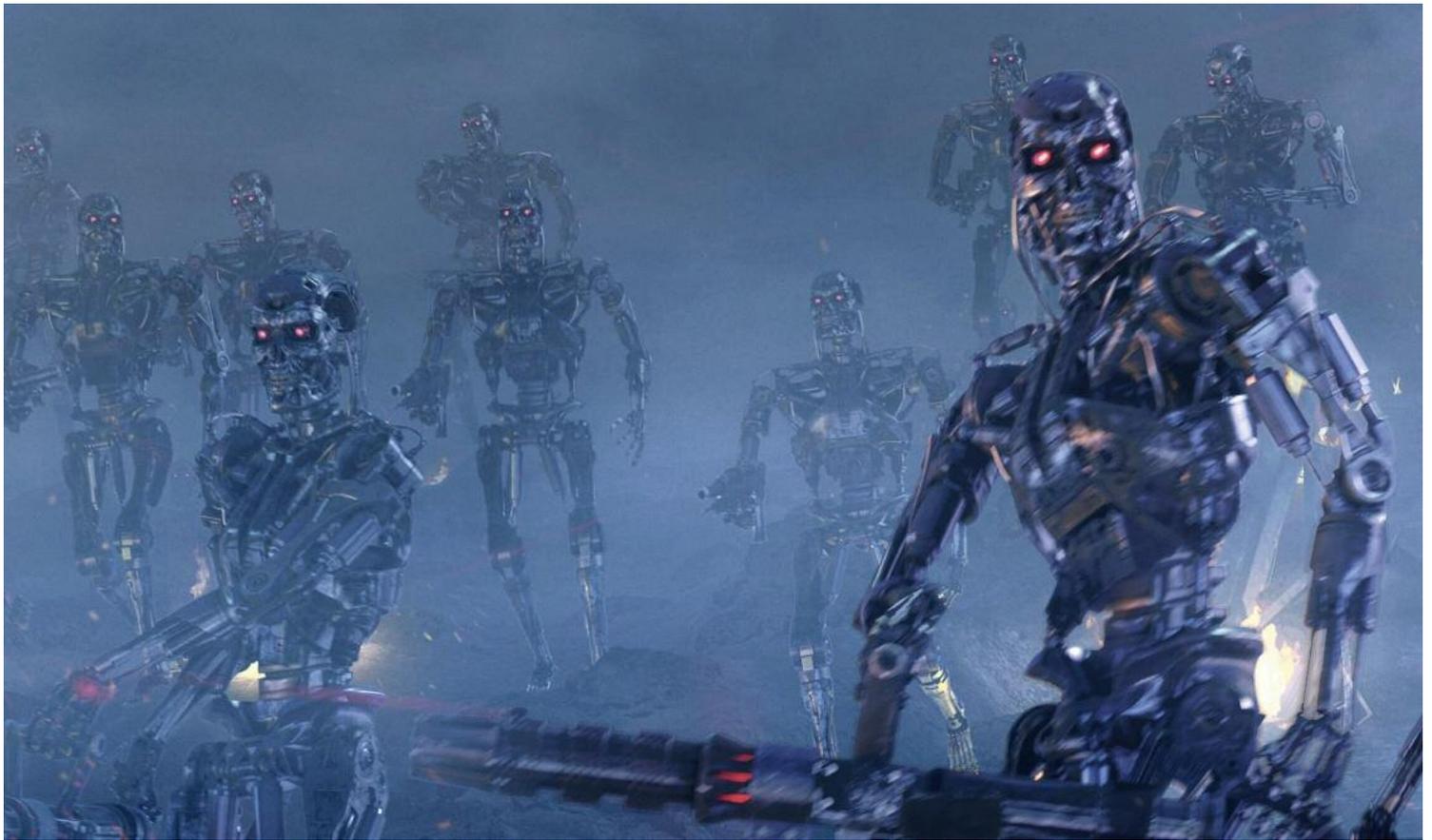
Was Hawking heute hilft, könnte ihn und uns morgen, übermorgen zerstören: Aber weil wir uns längst auf den Weg ge-

macht haben in diese Zukunft, weil an den Robotern und Algorithmen gearbeitet wird, die die Welt beherrschen könnten, ist es nun wohl zu spät umzukehren. Wir hängen fest zwischen der Gegenwart und der Zukunft, daraus entsteht die Angst.

Was dieses Heute, Morgen, Übermorgen ist oder sein könnte, darüber machen sich gerade viele Leute, Journalisten, Experten, Wissenschaftler, ihre Gedanken. Fast im Wochenrhythmus erscheint ein neues Buch, das das Internet, die Computerwelt, die digitale Gesellschaft erklären soll.

Es gibt auch viele Fragen: Was ist der digitale Mensch, was sind seine Rechte, was ist Freiheit und Überwachung, wie manipulierbar ist die Wirklichkeit, was ist mit dem Ende der Arbeit, was für ein Kapitalismus soll es sein, wie funktioniert Demokratie unter digitalen Bedingungen?

Wir wachen gerade auf und merken, wie sehr sich alles verändert. Und das scheint Angst zu produzieren.



Szene aus „Terminator 3 – Rebellion der Maschinen“, 2003: „Gefährlicher als die Atombombe“

Die eine Angst, die dabei entsteht, kombiniert schwer verständliche Zusammenhänge und leicht verständliche Antworten – es ist eine Angst, die eine relativ einfache Gratifikation bereithält: Man kann gegen etwas sein, gegen Google etwa, und fühlt sich dadurch besser.

Die andere, die entscheidende Angst ist in Wahrheit gar keine Angst, dazu ist sie viel zu existenziell, dazu macht sie einen viel zu ratlos. Es ist wie eine riesige Wand aus Eis und Schnee, die auf einen zurast, man hört schon das Dröhnen, während sich nebenan noch jemand darüber beschwert, dass der Skilift zu teuer ist.

Das Ende der Menschheit also: „Der Hauptgrund für unsere beherrschende Stel-

„Ich hoffe, wir sind nicht das biologische Startprogramm für eine digitale Superintelligenz“, sagt Elon Musk.

lung auf der Erde besteht darin, dass unser Gehirn im Vergleich zu dem anderer Tiere über ein paar zusätzliche Fähigkeiten verfügt“, schreibt der schwedische Philosoph Nick Bostrom in seinem aufregenden, im Surkamp Verlag erschienenen Buch „Superintelligenz“ noch eher beruhigend – nur um eine Seite weiter festzustellen, dass „man eine superintelligente KI (Künstliche Intelligenz) nicht schlau in dem Sinn nennen sollte, wie es ein wissenschaftliches Genie im Vergleich zu einem Durchschnittsmenschen ist. Es dürfte eher um einen Un-

terschied von der Art gehen, wie er zwischen einem Durchschnittsmenschen und einem Käfer oder einem Wurm besteht.“

Der Wurm, um das noch mal klarzumachen, sind wir. Jedenfalls wenn es Maschinen gibt, die einen IQ von 6455 haben, Maschinen, die irgendwann einmal von Menschen erfunden wurden und die es dann geschafft haben, sich mit ihren technischen Mitteln selbst schlau und immer schlauer zu machen.

„Singularität“, so nennt der Techno-Visionär Ray Kurzweil dieses Ereignis, „Superintelligenz“, so nennt das Nick Bostrom in seinem gleichnamigen Buch, ein Text so nüchtern und cool, so angstfrei und dadurch umso erregender, dass da-

nach das, was bisher vor allem Filme durchgespielt haben, auf einmal höchst plausibel erscheint.

Hal zum Beispiel, der mörderische Bordcomputer aus Stanley Kubricks Film „2001: Odyssee im Weltraum“ von 1968, oder die Terrorroboter aus „Terminator 3 – Rebellion der Maschinen“ von 2003 oder 2015 der neue Teil der „Avengers“-Serie: Gedacht wurde der Aufstand der Maschinen schon öfter, meistens, so sagt es der Futurologe Paul Saffo, aus einer „tiefen Sehnsucht heraus“, er sieht darin die Vertreibung aus

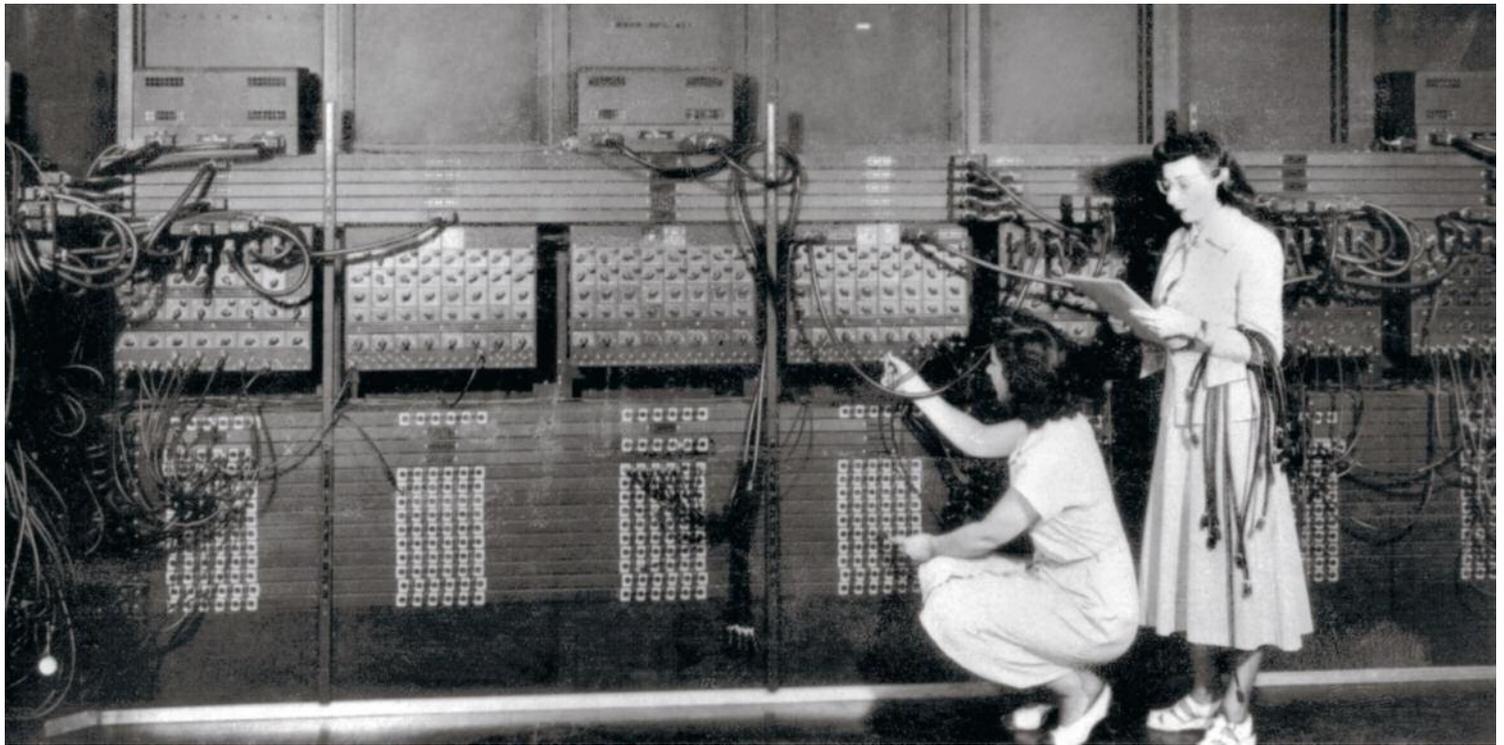
dem Paradies gespiegelt und das christliche Bild der Ursünde.

Das, wovon Bostrom spricht, ist aber etwas ganz anderes, das weiß auch Saffo, der im Silicon Valley lebt, wo sie die technologische Welterrettung zur säkularen Religion erklärt haben – mit Leuten wie dem Milliardär Elon Musk als Hohepriester, dessen Firma Tesla ein Elektroauto entwickelt hat und der mit seinem Unternehmen SpaceX bemannte Weltraumflüge anbietet will.

Musk hat Bostroms Buch gelesen, er hat die Lektüre sogar seinen mittlerweile rund 1,3 Millionen Twitter-Followern empfohlen, weil er die Künstliche Intelligenz für „gefährlicher als die Atombombe“ hält: „Ich hoffe, wir sind nicht nur das biologische Startprogramm für eine digitale Superintelligenz“, twitterte er. „Leider erscheint das immer wahrscheinlicher.“

Auch Peter Thiel, Milliardär, PayPal-Gründer, Musk-Freund und Silicon-Valley-Schlüsselfigur, findet, dass die Menschen sich viel zu viel vor der Klimaerwärmung fürchten und viel zu wenig vor der Künstlichen Intelligenz, deren Triumph einmal ein so „einschneidendes Ereignis sein wird wie die Ankunft von Außerirdischen auf der Erde“.

Nur Larry Page, der Google-Gründer, der „Superintelligenz“ auch gelesen hat, warnt vor „Alarmismus“ und bleibt optimistisch, dass man „lernen“ und dann die Maschinen schon irgendwie kontrollieren



Technikerinnen am frühen Computer ENIAC an der University of Pennsylvania um 1944: Eine Welt ohne Not und Naturkatastrophen

könne: „Je näher wir dem Ereignis kommen“, sagt er, „desto mehr werden wir wissen. Ich denke, wir werden viel lernen in dem Prozess.“

Über das „Kontrollproblem“ hat auch Bostrom nachgedacht, im Kapitel 9 von „Superintelligenz“, direkt nach Kapitel 8, „Sind wir dem Untergang geweiht“: Bostrom, und das ist der manchmal fast komische skandinavische Pragmatismus des Schweden, der heute in Oxford Professor ist, schlägt vor, die Künstliche Intelligenz (KI) in einen Metallkäfig einzuschließen, damit sie „keine Radiowellen aussenden und damit möglicherweise Radioempfänger oder andere Elektrogeräte manipulieren kann“.

Sollte es zur Konfrontation kommen, sieht Bostrom dabei zwei Möglichkeiten: die „physische Sicherheitsverwahrung“, die uns allerdings „ein falsches Gefühl von Sicherheit geben könnte“, oder die „informationale Sicherheitsverwahrung“, wobei es nicht genügt, „die KI vom Internet zu trennen, wenn es menschliche Wärter gibt, die mit ihr in Kontakt bleiben“.

Bostroms Buch, das merkt man schon, ist apokalyptisch nur in der Akuratesse. Er hat keine Interessen, er hat keine Moral, die ihn antreibt, er hat nur Fragen: Wie schnell wird etwa der „Take-off“ gehen, wie schnell wird also die Superintelligenz entstehen, Sekunden, Stunden, Monate, Jahre, und warum stehen die Chancen der Menschen besser, wenn der „Take-off“ langsam passiert?

Wird die Superintelligenz gut sein oder böse oder weder noch, weil sie den Menschen gar nicht wahrnimmt und macht,

was sie will? Oder wird sie den Menschen sogar retten, kann man ihr beibringen, gut zu sein, kann der Mensch profitieren, wenn die Superintelligenz Naturkatastrophen, Hunger, Not eliminiert? Bostrom kann sich aber auch Szenarien vorstellen, „in denen eine Superintelligenz an die Macht gelangt, indem sie politische Prozesse, Finanzmärkte oder Informationsströme beeinflusst oder die Kontrolle über existierende Waffensysteme übernimmt“.

Offen bleibt bei alledem, wenn man so will, die Digitalität des Bösen – selbst wenn Bostrom eine „existenzielle Katastrophe“ auf die Menschen zukommen sieht in nicht mehr allzu weiter Zukunft, in 80, 100 oder auch nur 25 Jahren, das kann niemand so genau sagen.

Aber böse? Eher banal. „Für digitale Intelligenzen, die ihre Körper wechseln wie wir unsere Kleidung und die sich ebenso leicht vervielfältigen, muss die Erhaltung des Selbst in Form einer bestimmten Implementierung oder eines bestimmten physischen Objekts kein wichtiger instrumenteller Wert mehr sein“, schreibt Bostrom.

Für den Menschen sieht er einerseits die Möglichkeit der „Migration in die digitale Welt“, also des eigenen „Hochladens, da ein von einer Superintelligenz physikalisch optimiertes Rechensubstrat effizienter wäre als ein biologisches Gehirn“.

Eine andere Konsequenz könnte sein, dass das Vermögen eines Menschen direkte Auswirkungen auf die Lebenszeit hätte, weil sich ein Milliardär zum Beispiel tausendmal mehr Rechenleistung kaufen könnte als ein Millionär – was dann wieder zurückführt zum Silicon Valley und dieser

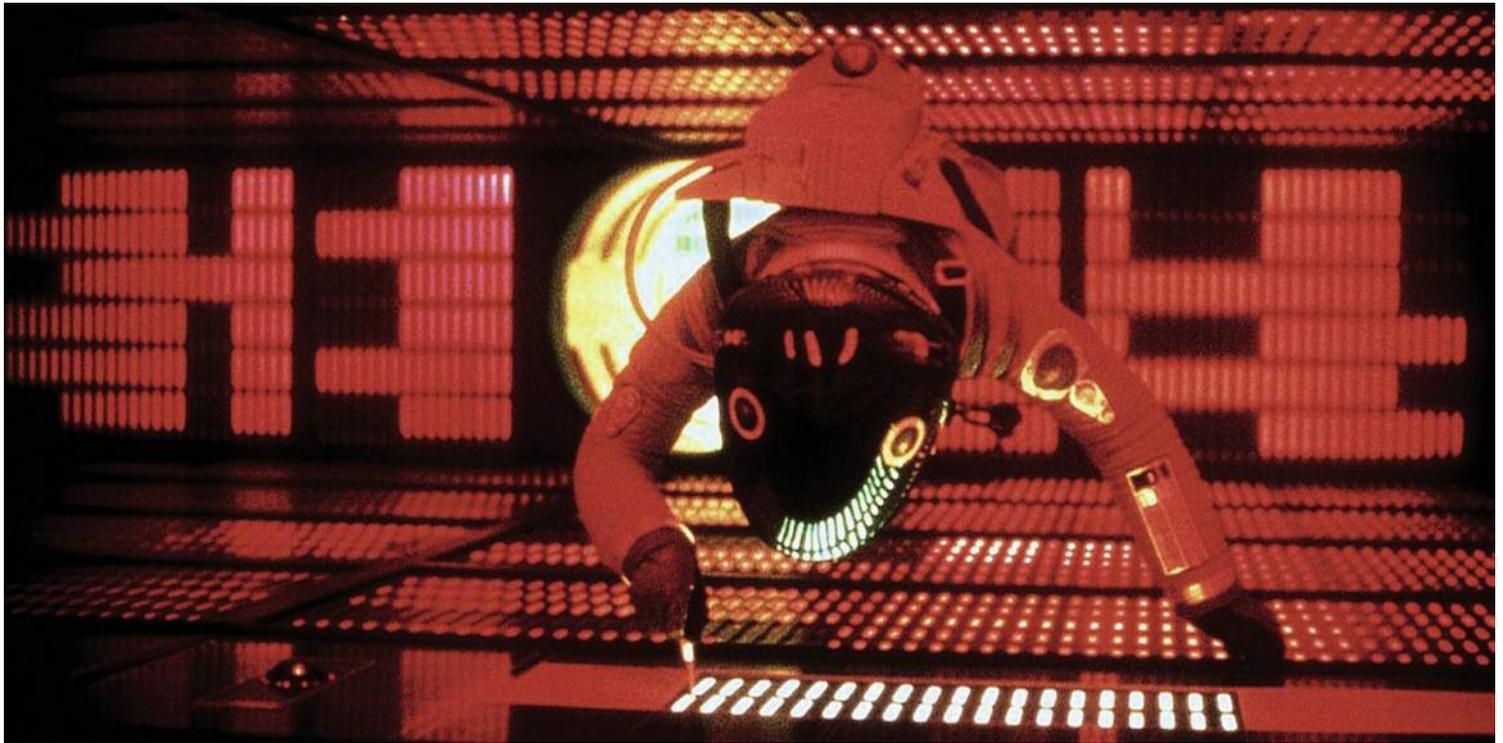
dort typischen Mischung aus Technologie, Geld und der Vorstellung, das neu zu definieren, was Leben heißt.

„Kreative Zerstörung“, so heißt hier das Credo: ein Slogan, der heute, wo so viele der alten Institutionen und Firmen in der Defensive sind oder längst wanken, wie eine Drohung klingt – tatsächlich verweist diese Formel aber auf die gegenkulturellen, anarchischen, freiheitlichen Wurzeln des Internets, ein Erbe, das bis heute präsent ist, eine Art doppeltes Netz.

Denn die Hightech-Industrie des Silicon Valley hat einen doppelten Ursprung. Fast alles, was mit dem Computerzeitalter zu tun hat, hat wenigstens zwei sehr gegensätzliche Entwicklungen durchlaufen – und wenn man dieses Zeitalter verstehen will, muss man eben auch seine Geschichte verstehen.

Etwas, das George Dyson in einem mächtigen Werk beschreibt, mit bessener Genauigkeit – es geht dort aber auch um seinen Vater Freeman Dyson und um Princeton, wo die Familie lange gelebt hat: den Ort, an dem in den Vierziger- und Fünfzigerjahren wichtige Schritte gemacht wurden bei der Entwicklung des Computers.

„Turings Kathedrale“, so heißt Dysons im Propyläen Verlag erschienenes Buch und verweist damit schon ganz direkt auf den Mann, Alan Turing, den genialen Mathematiker, der sich umbrachte, weil er aufgrund seiner Homosexualität medizinisch „behandelt“ wurde – der Wegbereiter jener Künstlichen Intelligenz, von der Hawking und Bostrom heute sprechen, und der Erfinder eben jenes Turing-Tests, mit dem man feststellen kann, ob und



Szene aus „2001: Odyssee im Weltraum“, 1968: Der Computer kann Leben vernichten und Leben erschaffen

wann die Maschine dem Menschen ebenbürtig ist. Ein Biopic über Turing ist gerade in den amerikanischen Kinos angelaufen und gilt als einer der Favoriten für die nächsten Oscars.

Die Ursprünge des digitalen Zeitalters, so wie Dyson sie beschreibt, liegen dabei in der Zeit des Krieges, was erst einmal nichts über das Wesen des Computers aussagt: Mathematiker, Techniker, Tüftler, Genies wie der jüdisch-ungarische Emigrant John von Neumann und der Britin Turing machten sich angesichts der Bedrohung durch die Deutschen daran, mit ihren Mitteln in den Krieg gegen Hitler einzugreifen.

Turing, der schon 1936 mit seiner Idee der „Universalmaschine“ den Computer vorgedacht hatte, war unter anderem wesentlich daran beteiligt, dass die Alliierten den deutschen Verschlüsselungscode Enigma knackten, ein mithin kriegsentscheidendes Ereignis.

„Wenn wir eine Zukunft haben wollen, müssen wir unseren Horizont über die Erde erweitern“, sagt Stephen Hawking.

Turing war auch dabei, als in England einfache Maschinen entwickelt wurden, die mit Lochkarten gefüttert wurden und einfache Handlungen ausführten, eben Computer, Zahlenmaschinen.

Eines der Probleme im Krieg, das mithilfe der neuen Technik gelöst werden sollte, war es, die Flugbahnen der deutschen Bomber vorauszuberechnen, damit die Abwehrgeschosse sie besser treffen konnten – eine Mischung aus Zufall und Statistik, weil die Bomber vergleichbare Muster hatten beim Anfliegen. Es waren damit Vor-

läufer jener Algorithmen, mit deren Hilfe unter anderem heute Stephen Hawking sprechen kann.

Dyson beschreibt eindrucksvoll die verschiedenen Persönlichkeiten, die das digitale Zeitalter formten, er legt aber auch die Ambivalenzen offen, wenn er etwa schreibt, dass dem Computer zwei Funktionen zukamen: „Er sollte der Entwicklung von Waffen zur Vernichtung des Lebens dienen, und er sollte neuartige, uns unbekannte Lebensformen erschaffen.“

In diesem Satz steckt die Angst, aber auch die Wahrheit über unsere Zeit und über all das, was mit dem Computer und dem Internet verbunden wird – das Internet, das selbst wiederum seine Wurzeln im militärischen Denken hatte. Denn der Ursprung des World Wide Webs, das Tim Berners-Lee 1989 erfand, ist das amerikanische Arpanet der Sechzigerjahre, das vor allem deshalb dezentral und antihierar-

chisch sein sollte, um Ausfälle des gesamten Netzes zu vermeiden.

Aber dezentral und antihierarchisch, das beschreibt auch sehr gut eine eher hippiehafte Vorstellung, wie die Welt sein sollte, es beschreibt auch den Politikansatz etwa der Piratenpartei – was bedeutet dieser doppelte Ursprung also für das Verständnis und vor allem für die Zukunft?

Die Antwort auf diese Frage hängt ganz wesentlich damit zusammen, wie man die Welt sieht – es ist immer das Menschenbild selbst, das sich seine Wahrheit schafft,

nicht umgekehrt. Ist die Technologie also eher Bedrohung, Kontrolle, Automatisierung, Gefahr? Für diese Annahme gibt es gute Gründe.

Oder ist die Technologie eher eine Chance, die Lösung existenzieller menschlicher Probleme und ein Freiheitsvehikel? Auch für diese Annahme gibt es gute Gründe.

Was nun die beiden Bücher von George Dyson und Nick Bostrom so überzeugend macht, ist die Tatsache, dass sie beschreiben, bevor sie urteilen. Sie öffnen, auf diese angenehm angelsächsische Art, das Problemfeld, historisch fundiert und ohne interpretatorischen Furor. Die mögliche Kritik steckt jeweils in den Büchern, der Leser kann sich daraus selbst sein eigenes Urteil bilden.

Für Stephen Hawking ist das allerdings keine Frage. Er ist Naturwissenschaftler, kein Mann der Angst, sondern der Tatsachen. Er sucht die Antwort auf das Problem, das er beschreibt, den möglichen und technologisch produzierten Untergang der Menschheit, in der Technologie.

„Wenn wir eine längerfristige Zukunft haben wollten, müssen wir unseren Horizont über die Erde hinaus erweitern“, sagt er, und es klingt ein wenig nach Science-Fiction, Kolonien im All, Chemiebaukasten, Fischer-Technik.

„Wir müssen uns in den Weltraum begeben, zu anderen Sternen, damit eine Katastrophe auf der Erde nicht das Ende der Menschheit bedeutet. Es wird Zeit und Mühe brauchen, um selbstversorgende Kolonien im All zu etablieren, aber je mehr sich die Technologie verbessert, desto einfacher wird das werden.“