

«Werden die Rasierten von den Bärtigen oder die Bärtigen von den Rasierten ausgerottet?»

Einführung von Walter Thurnherr zum Thema:

Künstliche Intelligenz und Demokratie

Universität Basel, 22. Oktober 2024

«Noch sind sie unsere Knechte / Noch führen sie aus...

*...Doch bald / Werden sie weiter rechnen / Ohne uns /
Formeln finden, / die nicht mehr zu interpretieren sind»*

Friedrich Dürrenmatt, Elektronische Hirne

*«The real danger is not that computers will begin to think like men,
but that men begin to think like computers.»*

Sydney J. Harris

Anrede

Ein Besuch im Panthéon von Paris ist eine aufregende Sache. Diese nationale Ruhmeshalle Frankreichs ist faszinierend, selbst wenn, oder vielleicht gerade weil, ein Panthéon in der Schweiz unvorstellbar wäre. Für eine eidgenössische Ruhmeshalle würde es uns zwar nicht an nationalen Grössen fehlen, und schon gar nicht an Kandidatinnen und Kandidaten. Aber in der Schweiz werden bekanntlich alle nach kantonalen Vorschriften beigelegt. In Paris ist das anders. Da schreiten Sie nicht nur an feierlichen Inschriften vorbei: «Vivre libre ou mourir». Sie stehen nicht nur staunend vor dem Foucaultschen Pendel, das mittels 28 Kilogramm Messing an einem 67 Meter langen Stahlseil majestätisch schwingend die Erdrotation nachweist. Sie besichtigen nicht nur die Grabstätten von Napoleon Bonaparte, Victor

Hugo, Alexandre Dumas, Jean-Jacques Rousseau und Voltaire. Sie bleiben auch schauernd vor dem Grab von Marie Curie stehen. Es ist paradox: Über dem Eingang der nationalen Ruhmeshalle Frankreichs steht in riesigen Lettern und angeblich im generischen Maskulin: «AUX GRANDS HOMMES - LA PATRIE RECONNAISSANTE», und drinnen liegt in einem mit drei Zentimeter dickem Blei ausgelegten Sarg diese zierliche kleine Frau aus Polen, die für die Entdeckung der Radioaktivität den ersten von zwei Nobelpreisen bekommen hat, und die immer noch derart strahlt, dass man die Besucherinnen und Besucher des Panthéon vor ihrem Leichnam schützen muss.

Wer heute vor diesem radioaktiven Grab steht, muss unweigerlich an die Debatte über die künstliche Intelligenz denken: Denn die Radioaktivität wurde zwar schon 1896 entdeckt, der Atomkern aber erst 1911. Mit anderen Worten, schon die einfachste Feststellung, dass Radioaktivität ein kernphysikalisches Phänomen ist, konnte erst fünfzehn Jahre nach der Entdeckung der Radioaktivität gemacht werden. Im Prinzip hatte man keine Ahnung, woher die neue Energie kam. Die Verbindung zwischen den Bindungen im Kern und dessen Stabilität, wurde erst 1920 hergestellt. Es dauerte also rund ein Vierteljahrhundert, bis man nur schon verstand, weshalb gewisse Elemente radioaktiv sind und andere nicht. Aber auch dann lag man noch falsch: 1933 schrieb Ernest Rutherford in der Londoner Times, jene Leute, die glaubten, man könne aufgrund dieser Erkenntnisse aus den Atomen Energie gewinnen «are talking moonshine». Und Ernest Rutherford war nicht irgendjemand in der Kernforschung. Albert Einstein, auch nicht gerade ein Niemand auf dem Gebiet der modernen Physik, bestätigte 1934, dass es praktisch ausgeschlossen sei, aus Materie Energie zu gewinnen: «It is like shooting birds in the dark in a country where there are only a few birds».

Aber nur acht Jahre später wurde in Chicago der erste Atomreaktor gebaut, und weitere drei Jahre später fielen die ersten Atombomben auf Nagasaki und Hiroshima. Lange ging sehr wenig, dann ging es sehr schnell.

Und auch dann lag man noch falsch, selbst in der Schweiz. Nach dem Zweiten Weltkrieg glaubten hierzulande viele, man könne nun in jedem Wohnquartier ein kleines Atomkraftwerk bauen, um die Häuser zu heizen. Dann glaubte man 1958, die Schweiz müsse zum Schutz der Neutralität einige hundert Atombomben beschaffen - und es ist irgendwie bezeichnend, dass Bundesrat Thomas Holenstein diese Haltung erstmals an einem Schützenfest bekanntmachte. Und selbstverständlich hielt man Atomkraftwerke für sicher: Als 1962 alt Bundesrat Hans Streuli, in seiner neuen Funktion als Präsident der «Nationalen Gesellschaft zur Förderung der industriellen Atomtechnik» den beunruhigten Vertretern des Reformierten Töchterheims von Lucens gegenüber sass, stellte er klärend fest: «Ein Werk wie das Versuchsatomkraftwerk Lucens explodiert nicht, denn es kann gar nicht explodieren». Nur sieben Jahre später explodierte es dann doch, und die schweizerische Reaktorentwicklung wurde abgebrochen. Die nachfolgenden Katastrophen von Tschernobyl und Fukushima und die dahinterstehenden Irrtümer sind Ihnen auch bekannt. Und wer heute im Panthéon vor dem Grab von Marie Curie steht, der macht automatisch einen Schritt zurück, kratzt sich am Hinterkopf und fragt sich nachdenklich: Und was erkennen wir heute nicht, von dem unsere Nachfahren sagen werden: Jesses Gott, waren die naiv!

Ich finde es deshalb nützlich, wenn die Einführung der künstlichen Intelligenz und ihre Folgen für die Demokratie diskutiert werden, würde aber die Debatte mit folgenden drei Bemerkungen – eigentlich einer

Vorbemerkung und zwei Feststellungen - gerne in einen etwas grösseren Zusammenhang stellen:

Vorbemerkung: Um die Demokratie zu zerstören oder ihr zu schaden, braucht es keine künstliche Intelligenz. Die handelsüblichen Dummheiten tun es auch.

Vielleicht wird das in Zukunft anders sein, aber zurzeit liegen die Gründe, weshalb Demokratien unter Druck sind, ganz woanders, und oft gehen sie weiter zurück, als man meint. Demokratien werden insbesondere dann in Frage gestellt, wenn ihre Voraussetzungen nicht mehr erfüllt sind: Rechtssicherheit, Gewaltenteilung, ein Minimum an Wohlfahrt und sozialer Ausgleich zwischen Reich und Arm, Solidarität mit den Hilfsbedürftigen, Bildung – und zwar auf allen Stufen, und nicht nur an den Eliteuniversitäten - , Medienvielfalt, Vertrauen in die politischen Institutionen, Zuversicht, dass man eine Zukunft schafft, die besser wird als die Gegenwart, und eine Reihe weiterer, gemeinsam getragener Werte und Normen, die nicht fortwährend neu begründet werden müssen, sondern von allen geteilt werden. Zählen Sie einmal die Länder auf, von denen Sie überzeugt sind, dass diese Voraussetzungen noch alle erfüllt sind. Selbst in Europa sind wir in einzelnen Ländern nur noch eine schwere Rezession entfernt von einem radikalen Umsturz. Und zwar nicht wegen zu viel künstlicher Intelligenz von Open AI oder Google, sondern wegen zu wenig politischer Umsicht von einzelnen Regierungen, wegen zu vielen Krisen, die laufend den Kontinent und die Welt traktieren, wegen zu viel Verzweiflung und zu viel Ohnmacht jener Menschen, die die Folgen dieser Krisen besonders stark zu spüren bekommen, und wegen zu vielen populistischen Gaunern, die diese Verzweiflung auszunützen verstehen.

Die Demokratie ist auch unter Druck, weil sie in den Augen vieler Menschen nicht mehr geeignet scheint, die dringlichsten und grössten Probleme der Welt zu lösen. «The Guardian» schrieb bereits 2018: «Democracies do seem to be doing some fairly stupid things at present». Selbst unter den Jugendlichen, die im Frühling 2020 vor dem Bundeshaus für eine griffige Klimapolitik demonstrierten, gab es solche, die Schilder aufhielten mit dem bedenklichen Satz: «Wenn Ihr Notrecht anwenden könnt gegen Covid, dann könnt Ihr auch Notrecht einsetzen für das Klima». Erstaunlich ist diese Forderung andererseits nicht, schliesslich haben sich die Regierungen dieser Generation (ähnlich wie jene früherer Generationen) nicht unbedingt durch fanatisches Verfolgen der Nachhaltigkeitsziele ausgezeichnet. Die globalen Risiken werden grösser und zahlreicher, und die Feststellung von Eliezer Yudkowsky in Bezug auf AI könnte man ohne Weiteres erweitern: «An alarmingly different kind of Moore's Law is at work today: the minimum IQ needed to destroy the world drops by one point every 18 months.»

Andere, die bis vor zwanzig Jahren geglaubt haben, dass Demokratie eine Voraussetzung für wirtschaftliches Wachstum sei, orientieren sich seither an China, das scheinbar das Gegenteil beweist.

Und wieder andere fühlen sich bei sich zu Hause nicht mehr zu Hause. In ihren Augen verändert sich alles zu schnell und in die falsche Richtung: Die Kriminalität in der eigenen Stadt steigt, die bisherigen Strukturen zerfallen, der gegenseitige Respekt, der Anstand und der Zusammenhalt schwinden, und das Vertrauen in die Politik sackt ab.

Die Demokratie ist unter Druck, weil sie eben mehr ist als die regelmässige Aufregung darüber, ob der Franz oder der Fritz gewählt wird, wie Carl Spitteler einmal meinte. Sie ist eine Errungenschaft, die empfindlich reagiert, wenn sie ausgehöhlt oder missbraucht wird, schon ohne künstliche

aber mit böswilliger Intelligenz. Umso mehr kommt es darauf an, welche Gesetze man beschliesst, um sie zu erhalten und zu verteidigen. Und das führt mich zur zweiten Bemerkung:

Zweitens: Künstliche Intelligenz ist vielleicht ein gewaltiger Segen und vielleicht ein fürchterliches Risiko und vielleicht beides. Bei solcher Ausgangslage grosser Unsicherheit wird gewöhnlich halbbatzig reguliert, obwohl gute Gründe bestehen, vorsichtig zu sein.

Vor fünf Jahren habe ich in einer Rede spekuliert, dass Friedrich Dürrenmatt, würde er heute noch leben, angesichts der Fortschritte in der künstlichen Intelligenz nicht mehr die Komödie «die Physiker» schreiben würde, sondern das Stück: «die Informatiker». Heute wüsste ich auch wie: Statt mit der Figur von Isaac Newton, mit John von Neumann; statt mit einem verwirrten Albert Einstein, mit einem warnenden Geoffrey Hinton; statt im privaten Sanatorium im Kommissionszimmer eines Parlaments.

Von den 21 Punkten, die Dürrenmatt damals als Anhang zu den Physikern veröffentlicht hatte, könnte man die meisten tel quel übernehmen: «Je planmässiger die Menschen vorgehen, desto wirksamer vermag sie der Zufall zu treffen». «Der Inhalt der Physik geht die Physiker an, die Auswirkungen alle Menschen». «Was alle angeht, können nur alle lösen». Und so weiter.

Friedrich Dürrenmatt hat auch erkannt, wie schwierig internationale Regulierung ist. 1987 war er nach zwanzig Jahren zum zweiten Mal in Moskau, dieses Mal am Friedensforum von Michail Gorbatschow, mit Tausend Intellektuellen und Wissenschaftlern aus der ganzen Welt. Vom modernen Moskau war Dürrenmatt enttäuscht: «Moskau ist eine Art Gross-Ostermundigen geworden. Man hat immer das Gefühl, das seien Vorstädte».

Und vom Kongress war er auch nicht begeistert: «Jeder hat das Gleiche gesagt, wie gefährlich die Atombombe sei, und dann hat Gorbatschow gedankt usw. Das war sehr langweilig. Da hab' ich etwas wütend gesagt, vor ungefähr 30 Millionen Jahren habe es auch so eine Sitzung gegeben und da seien die Dinosaurier zusammengekommen und hätten diskutiert, wie sie überleben können». Die Debatte ging offenbar trotzdem so weiter, wie sie begonnen hatte, und Gorbatschow hatte Dürrenmatt auch nicht darauf hingewiesen, dass die Dinosaurier schon vor 65 Millionen Jahren ausgestorben sind, aber es ging ja um Abrüstung von Atomwaffen, nicht um Saurier. Und dazu schrieb Dürrenmatt später: «Am Ende der Diskussion habe ich das dem Frisch gesagt: Das sei das Verteufelte, wir rüsten ab, und dann kommt plötzlich einer, und der hat noch eine, nicht? Da sagte er: Man darf jetzt keine Witze machen. Frisch nahm das als Witz. Aber es ist wirklich ein grosses Problem».

Dürrenmatt hatte natürlich recht. Und was für die Atomtechnologie galt, gilt erst recht für die künstliche Intelligenz: So viele Vorteile, wenn man sie hat. Nicht nur militärische. Auch im Bereich der Medizin, denken Sie nur an die Diagnostik, an das Design von Molekülen in der Pharma-Forschung oder an das Proteinfolding. Im Energiebereich, vom Reaktorbau bis zur möglichst effizienten Nutzung von Stromnetzen. Von der Chemie über die Materialwissenschaften bis zur Mathematik. Von der Kriminalitätsbekämpfung über die Kommunikation bis zur Meteorologie. Überall wird bereits KI eingesetzt. Und an vielen Orten mit einem unerhörten Gewinn. So viele Vorteile, wenn man sie selber nutzen kann.

Aber auch erhebliche Risiken, wenn man andere einfach machen lässt. Nicht nur militärische. Gewinnen wir das Wettrüsten der eigenen KI, die uns vor den Angriffen der anderen KI schützen soll? Bei Cyberangriffen zum

Beispiel, oder wenn es um die Echtheit von Bildern und Videos geht? Wie verläuft in Zukunft die politische Meinungsbildung in Demokratien, wenn man nicht mehr weiss, welche Aussagen von Menschen und welche von Bots gemacht worden sind? Wenn Parlamentsmitglieder von Mails überschwemmt werden, von denen sie nicht wissen, ob diese von besorgten Bürgerinnen und Bürgern oder von einer KI redigiert worden sind. Wo ziehen wir in Zukunft die Grenzen der Privatsphäre und des Datenschutzes, wenn Dritte sie mittels KI auf beispiellose Art überschreiten können? Und wie verteidigen wir diese Grenzen?

Sind wir vorbereitet auf die Auswirkungen der KI auf den Arbeitsmarkt? Allenfalls sind Letztere nicht so gross, wie befürchtet. Aber vielleicht eben doch? Und welche Folgen hätte das längerfristig? Wer kontrolliert die Unternehmen, die die KI kontrollieren? Sie können bei Large Language Models (LLM) ja nicht einfach «Open Source» einfordern, wie bei anderen Algorithmen, wenn Sie Schlimmeres verhindern wollen. Also, nicht nur wer, sondern auch wie kontrolliert man die Unternehmen, die KI einsetzen? Oder nur schon: Wie misst man, was die KI-Unternehmen mit ihrer KI anrichten? Überhaupt: Wie reguliert man vorgängig eine Technologie, deren Wirkung man erst dann richtig einschätzen kann, wenn man sie bereits eingesetzt hat? Schliesslich sind insbesondere LLM, wie Demis Hassabis jeweils erklärt, «so unreasonably effective». Wie schafft man und wie reguliert man «Alignment»? Kommt dazu: Sie können Ihre KI im Nachbarland oder irgendwo auf der Welt einsetzen, und sind also auch vom Einsatz der KI anderer Staaten betroffen. Also: Wie regulieren Sie KI auf internationaler Ebene? Wir schaffen es ja nicht einmal, den CO₂-Ausstoss international zu regulieren, obwohl wir beim Klima eine viel klarere Idee haben, was zu tun wäre als bei der KI? Und selbst wenn Sie ein internationales Abkommen aushandeln und

die schlimmsten KI verbieten könnten: «Das Verteufelte ist, wir rüsten ab, und dann kommt plötzlich einer, und der hat noch eine, nicht?» .

Dass wir nicht besonders gut sind bei der Regulierung der Digitalisierung, haben wir schon mit den sozialen Medien bewiesen. Unsere Reaktion nach 2007: staunen, loben, einführen, anwenden, überrascht sein, überfordert sein, zu spät sein und dann halbhatzig und halbherzig legiferieren. Aber die sozialen Medien haben viel verändert und verschoben, und dies bekanntlich nicht nur in die positive Richtung: wie wir uns informieren, wie wir unsere politische Meinung bilden, wie wir debattieren und uns beschimpfen, wie wir kränken, provozieren und zetern im Netz, wie wir uns politisch engagieren, wie wir Kampagnen führen und sie finanzieren, wie wir wählen und abstimmen und welche Schlussfolgerungen wir daraus ziehen. All das hat sich völlig geändert, spätestens seit 2009, als Google begann, die Suchergebnisse zu personalisieren, Twitter den «Retweet-Button» erfand und Facebook den «Like-Button» einführte. Auf kaum einem Gebiet sieht man, dass jene Erfahrung, die die Briten «the law of unintended consequences» nennen, derart zutrifft wie bei der Digitalisierung. Oder mit anderen Worten: Vieles in unserem Leben ist angenehmer, kurzweiliger und einfacher geworden, seit das iPhone auf den Markt kam. Aber die Digitalisierung hat vieles verändert, was wir gar nicht ändern wollten, und einiges ging komplett schief.

Und was wir unterschätzt haben: Wir leben in einer Umwelt, die uns permanent überfordert. Mein Grossvater wurde 1880 geboren. Er hatte nicht studiert, sondern war Bauer. Aber alles, was er anfasste, verstand er. Das, womit er hantierte, musste er selber reparieren können, und er konnte und tat es auch. Viele hier drin haben studiert und sind viel klüger als mein Grossvater. Aber wir verstehen nichts mehr von dem, was wir anfassen.

Keiner hier schraubt am Wochenende sein Handy auseinander, wenn es nicht mehr funktioniert. Keine Waschmaschine, kein Auto, kein Toaster ohne Halbleiter, aber nur noch die wenigsten könnten erklären, was ein Transistor überhaupt macht. Das führt zu einer eigenartigen Entfremdung und Entwurzelung, die Botho Strauss mit den Worten beschrieb: «Die Menschen sind in ihrem Bewusstsein weit unter das Niveau ihres Wissens gesunken. Ihre kunstvollen Einrichtungen, Produkte der Technik und der Wissenschaft überragen turmhoch, was sie darüber denken». Oder wie es kürzlich jemand aus dem Open AI Alignment Team sagte, als er über die Dinge sprach, die Large Language Models können, ohne dass man das erwartet hätte: «If you can do amazing things, but you can't really control it, then it is not so amazing». Und nur weil wir diesen Dingen dann Namen geben, wie «grokking», «double descent» usw., haben wir es ja noch nicht verstanden. Passiert in einem solchen Umfeld eine Panne, und dann noch eine zweite Panne, dann verlieren Sie das Vertrauen. Und ohne Vertrauen, verlieren Sie die Demokratie.

Und ich habe noch nicht einmal von der AGI, der «Artificial General Intelligence» gesprochen, oder der «Artificial Super Intelligence», und was diese für die Demokratie bedeuten könnten. Was heisst es für eine Demokratie, wenn eine Maschine statt ein Mensch entscheidet, oder vorentscheidet? Und ich spreche nicht von Entscheiden bei einem Rotlicht im Strassenverkehr, sondern, wenn es um die Steuerung von kritischen Infrastrukturen geht, oder um die Kontrolle über ein Raketenabwehrsystem, oder um die Beurteilung von angeklagten oder verdächtigten oder unverdächtigten Personen, oder um die Überwachung eines Patienten auf der Intensivstation. Ich spreche von Entscheiden, die vielleicht zu 99.5% richtig sind und schnell. Aber in 0.5 % auch völliger Unsinn.

Spätestens an dieser Stelle wird Sie der Hinweis nicht erstaunen, dass nicht nur Stanislaw Ulam aus Lwiw, John von Neumann aus Budapest und Ray Kurzweil aus New York sich mit der Singularität beschäftigten, also mit dem Zeitpunkt, an dem der technologische Fortschritt eben zum Beispiel auf dem Gebiet der künstlichen Intelligenz unkontrollierbar und unumkehrbar wird - sondern auch Friedrich Dürrenmatt aus Konolfingen, und zwar schon 1969 in seinem «Monstervortrag über Gerechtigkeit und Recht.»¹ Es ist ein hellsichtiger Text voller Dürrenmattscher Visionen, weshalb ich gerne aus diesem «KI-Demokratie-Papier» zitiere:

«Ohne Technik ist der Mensch nicht mehr lebensfähig. Die Maschinen vermögen, je perfekter wir sie bauen, all unsere Arbeiten auszuführen, auch unser Regieren über Menschen. Der Mensch kann nur durch Maschinen objektiv regiert werden, scheint es, es bleibt ihm nichts anderes übrig, als sich ihnen anzuvertrauen - [...] - Er wird der totale Konsument, der Konsument, der für sein Konsumieren nicht zu arbeiten braucht. Was wird für diesen Menschen wichtig? Kunstwerke? Wird sich eine ungeheure Sonntagsmalerei ausbreiten? Wird es noch eine Politik geben, oder wird die Politik eine Farce sein? Eine Zeremonie? Feierliche Protestmärsche gegen unumstössliche Computerbeschlüsse, deren Sinn nur noch den Computern einleuchtet? Werden Sekten entstehen, neue Religionen, weil sich der materiell gesicherte Mensch nur noch für das Metaphysische interessiert?

¹ Stanislaw Ulam hatte in seinem Nachruf auf John von Neumann nicht nur dessen Leistungen bei der Architektur und dem Bau der ersten programmierbaren Computer sowie von Neumanns Beitrag zum Bau der Atombombe beschrieben, sondern er berichtete auch von den Gesprächen über die Beschleunigung der technologischen Fortschritte und den Bedenken über die Auswirkungen eben dieser Fortschritte auf die Gesellschaft: «One conversation centered on the ever accelerating process of technology and changes in the mode of human life, which gives the appearance of approaching some essential singularity in the history of the race beyond which human affairs, as we know them, could not continue». Seit rund zwanzig Jahren wird deshalb gern mit dem Ausdruck der Singularität hantiert, wenn es um künstliche Intelligenz geht. Mit der Atombombe hätten wir bis anhin «Schwein gehabt», mit Betonung auf «bis anhin». Mit der künstlichen Intelligenz werde das allenfalls anders sein.

Fegen blutige Glaubenskämpfe über die von Computern verwaltete Erde? Werden die Rasierten von den Bärtigen oder die Bärtigen von den Rasierten ausgerottet? Wird der Fussball so lebensentscheidend, dass sich die Anhänger der verschiedenen Mannschaften zerfetzen? - [...] - Wird die von Computern beherrschte Menschheit eine Revolution gegen die Maschinen entfesseln, wird der Mensch wieder ein Pfahlbauer?»² Soweit Friedrich Dürrenmatt.

Persönlich bin ich nicht sicher, ob - und wenn ja - wie schnell es so schlimm kommen wird, aber mein Punkt ist: Die Risiken für die Demokratie sind bereits vor dem Erreichen der Singularität gross genug, um vorsichtig zu sein. Sie müssen nicht gleich von der Herrschaft der Maschinen über die Menschen ausgehen. Aber nehmen Sie zum Beispiel den einen Punkt der politischen Information, der für uns in der Schweiz eine erhebliche Rolle spielt, schliesslich stimmen wir alle drei Monate ab.

Sie wissen: Der Bundesrat ist verpflichtet, die Stimmbürgerinnen und Stimmbürger kontinuierlich über die Vorlagen zu informieren. Dabei muss er gemäss Gesetz die Grundsätze der Vollständigkeit, der Sachlichkeit, der Transparenz und der Verhältnismässigkeit beachten. Das gelingt nicht immer, ist aber auch nicht so einfach. Was genau ist «sachlich», was «vollständig» und was «verhältnismässig»? Die Medien, die Parteien, die Verbände, die Gewerkschaften und die Initiativ- oder Referendumskomitees haben diese gesetzliche Vorgabe des Bundesrates nicht. Sie können dazu dichten was sie wollen, verdrehen und verkürzen, was das Zeug hält. Und das ist auch richtig so, denn wir trauen den Stimmenden zu, dass sie kritisch hinterfragen, was ihnen von anderen vorgesetzt wird. Dabei helfen ihnen zum

² Friedrich Dürrenmatt, «Monstervortrag über Gerechtigkeit und Recht, nebst einem helvetischen Zwischenspiel, eine kleine Dramaturgie der Politik», 1969.

Teil die Medien, denn Letztere verfolgen dasselbe Ziel: Sie hinterfragen und ordnen ein (gewisse Medien verfolgen auch andere Ziele). Das kritische, selbstständige Nachdenken ist eine Bürgerpflicht. Aber es ist heute auch eine eigentliche Zumutung geworden, denn es wird immer schwieriger zu erkennen, wenn man angeschwindelt wird.

Und mit künstlicher Intelligenz dürfte das noch schwieriger werden. Früher oder später wird es möglich sein, dass Unternehmen, Parteien oder Verbände mittels KI jede Bürgerin und jeden Bürger laufend mit einer individuell zugeschnittenen Information bedienen können - möglicherweise einer zutreffenden Information, allenfalls einer nur leicht verfälschten oder nicht vollständigen Information, und vielleicht auch mit einer krassen Lüge. In Ansätzen wurde das ja schon von Cambridge Analytica vorgemacht, und bekanntlich stellte bereits 2018 das «Senate Intelligence Committee» fest, dass fremde Regierungen mit den neuen technologischen Mitteln in den US-Wahlkampf von 2016 eingegriffen haben: «Masquerading as Americans, these (Russian) operatives used targeted advertisements, intentionally falsified news articles, self generated content, and social media platform tools to interact with and attempt to deceive tens of millions of social media users in the United States.»

Die Absicht, damit das Abstimmungsverhalten zu beeinflussen, kann durchaus aufgehen. Denn das Schlimme ist ja, selbst wenn die Lüge als solche enttarnt wird, bleibt etwas hängen. Donald Trump ist nicht der erste, der das erkannt hat. Jeder Intrigant auf dem Schulhausplatz weiss das. Künstliche Intelligenz kann helfen, Desinformation zu identifizieren. Aber sie kann sie eben auch erzeugen und verstärken.

Wie also regulieren Sie das?

Das Dilemma wird noch grösser, wenn Sie berücksichtigen, wie die direkte Demokratie funktioniert. Denn wir bestimmen mittels Abstimmung nicht die beste Regulierung, sondern jene, deren Folgen die Mehrheit der Betroffenen zu tragen bereit ist. Beziehungsweise, wir definieren die «beste» Regulierung nicht technokratisch, sondern als jenen Ansatz, der zurzeit von den meisten als beste Regulierung akzeptiert wird. In diesem Sinn funktionieren zumindest wir in der Schweiz politisch wie eine «backpropagation» im Machinelearning: Wir versuchen es mit einem Ansatz, und wenn er nicht gut genug ist, verschieben sich die Gewichte in den unteren Ebenen, es kommt zu einer neuen Abstimmung und dann hoffentlich zu einem Gesetz, das besser akzeptiert wird. Und so weiter. Dieses iterative, stabilisierende und in diesem Sinn erfolgreiche Durchwursteln setzt jedoch voraus, dass wir die Möglichkeit vieler Versuche haben, um nachzujustieren. Bei den meisten Sachgebieten, von den Sozialversicherungen bis zur Landwirtschaftspolitik, ist das auch so: Da wird ständig revidiert und ergänzt und angepasst. Aber wer sagt, wie viele Anläufe wir bei der künstlichen Intelligenz zur Verfügung haben, bevor wir einen Zustand erreichen, den wir nicht mehr rückgängig machen können?

Und überhaupt, wer garantiert, dass wir deshalb das Problem nicht einfach verdrängen - so wie wir das immer wieder tun, wenn das Problem von beruhigend globaler Dimension ist. Als ich in der UNO arbeitete, hatte ich einen Freund, ein Franzose mit trockenem Humor. Nach gewissen Sitzungen, bei denen keiner einen konkreten Antrag machte, aber alle an die globale Bedeutung und globale Verantwortung erinnerten, stand er jeweils auf und sagte: «Um es mit Balzac zu sagen: 'Je weniger Gedanken man hat, desto leichter erhebt man sich'».

Meine Damen und Herren, die Zwischenbilanz lautet also: Wir wissen nicht genau, was künstliche Intelligenz ist. Wir wissen nicht genau, wie KI funktioniert. Wir wissen nicht genau, wie sich die KI weiterentwickeln wird und welche Auswirkungen KI auf die Demokratie haben wird, höchstens, dass wir die Demokratie auch ohne KI kaputtmachen können. Wir wissen nicht genau, wie man KI regulieren soll. Und wir wissen auch nicht, ob wir nicht schon zu spät sind.

Vielleicht sollte ich hier schliessen, und wir gehen alle an die Bar. Ich habe aber noch eine dritte Bemerkung:

Drittens: Es ist nichts in Stein gemeisselt. Es gibt Vorsichtsmassnahmen, die man treffen kann. Aber es braucht den politischen Willen, sich mit KI auseinanderzusetzen, bevor die Abhängigkeiten zu gross sind. Und Abhängigkeiten entstehen sehr schnell.

Im Gegensatz zur Radioaktivität ist künstliche Intelligenz kein Naturphänomen. Das heisst, wer fatalistisch zuschaut und behauptet, da sei nichts mehr zu machen, braucht eine Ausrede. Stephen Hawking bemerkte einmal auf die ihm eigene Weise, es sei ihm aufgefallen, dass selbst jene Leute, die davon ausgehen, alles sei vom Schicksal vorherbestimmt, nach links und nach rechts schauen, bevor sie die Strasse überqueren. Es ist eben nicht alles vorherbestimmt. Wenn man schon weiss, das KI auch immense politische Folgen hat oder haben könnte, dann sollte sich die Politik damit beschäftigen - sich kundig und vertraut machen mit den absehbaren, beabsichtigten Folgen und mit den möglichen unbeabsichtigten Folgen. Sich dafür interessieren und nicht denselben Fehler machen, der bei der Einführung der sozialen Medien passiert ist. Natürlich, die Politik sollte nicht

zu früh und nicht zu viel legiferieren, aber das heisst nicht, dass man sich damit nicht auseinandersetzen soll.

Meines Erachtens schadet es auch nicht, sich an den Erfahrungen der Nukleartechnologie zu orientieren, und zum Beispiel unabhängige Aufsichtsbehörden einzurichten, die die gesetzlichen Möglichkeiten haben, KI zu prüfen und zu kontrollieren.

Es spricht auch nichts dagegen, die Zusammenarbeit mit unseren Universitäten und Hochschulen zu suchen, nicht nur, um die neue Technologie besser zu verstehen, sondern um sich auch international besser einbringen zu können. Im Gegensatz zu dem, was zuweilen behauptet wird, hat die Schweiz nicht eigentlich ein Problem, angehört zu werden. Aber sie hat zuweilen (nicht immer) ein Problem, Inhalte zu formulieren, die andere nicht schon zehn Mal gehört haben. Vor allem, wenn wir nicht einfach die Regulierungen von Dritten übernehmen wollen, dann müssen wir international auf spezifischen Gebieten kompetenter sein als andere.

Schliesslich könnte es sich lohnen, frühzeitig an mögliche Gegenmassnahmen zu denken, um die Demokratie zu schützen.

Das Interessante an der gegenwärtigen Diskussion über die SRG ist ja, dass die hier behandelte Thematik nur wenig zur Sprache kommt. Desinformation ist zwar auch bei der SRG möglich. Und über die Qualität einzelner Sendungen lässt sich trefflich streiten. Ich habe mich auch schon aufgeregt. Doch kann man gegen die SRG immer eine Beschwerde einreichen, zuerst bei der Ombudsstelle, dann bei der UBI, der unabhängigen Beschwerdeinstanz. Die SRG ist verpflichtet, «Informationsangebote für eine umfassende, vielfältige und sachgerechte Berichterstattung» zur Verfügung zu stellen. Sie muss Hintergrundinformationen liefern und «Beiträge zur

Vertiefung, Einordnungen und Analysen des Geschehens» wie es in der SRG-Konzession heisst. Und diese Funktion, diese Aufgabe wird im Zeitalter der künstlichen Intelligenz wichtiger. Private Medien haben nicht dieselbe Verpflichtung. Kaum ein Zufall, dass ausgerechnet der Chief Commentator der Financial Times schreibt: «Countries that are lucky enough to have inherited high-quality public service broadcasters...from a better past, should defend them to the death.» Es gebe immer Fragen zum Angebot, zur Qualität und zur Finanzierung: «Yet these institutions are still among the most important sources of a body of facts, national sensibility, and national conversation.»

Meine Damen und Herren, damals, als die Radioaktivität entdeckt wurde, war man begeistert. Wie Sie wissen, schaffte es kaum eine Entdeckung so schnell vom Labor ins Spital zur konkreten Anwendung. Die Faszination und der Rummel hielten an. Bald wurde eine radioaktive Zahnpasta verkauft: «Doramad, für ein strahlendes Lächeln», wie es in der Werbeanzeige hiess. Dann folgte die «Dorothy Gray Cleansing Creme», ein radioaktiver Make-up-Entferner, für den bis in die 1950er-Jahre Werbung gemacht wurde, oder «Iradia», die radioaktive Unterwäsche, die eine wohlige und sanfte Wärme im Unterleib versprach. Mit anderen Worten: Technologische Anwendungen, selbst wenn sie gut gemeint sind, können auch in die Hosen gehen. Marie Curie sagte damals: «C'est maintenant le moment de comprendre davantage, afin de craindre moins.» Es sei jetzt der Moment, die Dinge zu studieren, um sich weniger vor ihnen zu fürchten. Ich würde sagen : «C'est maintenant le moment de comprendre davantage, afin de mieux légiférer.» Denn wenn die Demokratie nicht früh genug der KI den Stempel aufdrückt, wird früher oder später die KI der Demokratie den Stempel aufdrücken.
