

Voorwoord

De naam die we onze soort gegeven hebben is *Homo sapiens* – de wijze mens. Het valt alleen wel te betwisten of we die naam hebben waargemaakt.

De laatste honderdduizend jaar hebben wij sapiens enorm veel macht vergaard, dat staat buiten kijf. Er zijn hele boekdelen te vullen met al onze ontdekkingen, uitvindingen en veroveringen. Maar macht is geen wijsheid, en na honderdduizend jaar van ontdekkingen, uitvindingen en veroveringen heeft de mensheid zichzelf in een existentiële crisis gewerkt. We verkeren op het randje van de ecologische ondergang doordat we onze eigen macht misbruiken. We zijn ook druk in de weer met het creëren van nieuwe technologieën als kunstmatige intelligentie (AI), die het potentieel hebben om aan onze controle te ontsnappen en ons te onderwerpen of uit te roeien. We zouden nu als soort bij elkaar kunnen komen om die existentiële uitdagingen het hoofd te bieden, maar in plaats daarvan stijgen de internationale spanningen, wordt mondiale samenwerking steeds moeilijker, zijn landen vernietigingswapens aan het hamsteren en lijkt een nieuwe wereldoorlog niet ondenkbaar.

Als wij sapiens zo wijs zijn, waarom zijn we dan zo zelfdestructief?

We hebben zoveel informatie vergaard over van alles en nog wat, van DNA-moleculen tot verre sterrenstelsels, maar op een dieper niveau lijkt al die informatie geen antwoord opgeleverd te hebben op de grote levensvragen: wie zijn we? Wat moet ons streven zijn? Wat is een goed leven en hoe moeten we dat aanpakken? Ondanks de kolossale hoeveelheden informatie die we tot onze beschikking hebben, zijn we net zo vatbaar voor fantasieën en waanideeën als onze verre voorouders. Het nazisme en het stalinisme zijn maar twee recente voorbeelden van de massawaanzin die bij vlagen de overhand krijgt, zelfs in moderne samenlevingen. Niemand zal betwisten dat mensen vandaag de dag veel meer informatie en macht hebben dan in de steentijd, maar het is verre van zeker of we onszelf en onze rol in het universum veel beter zijn gaan begrijpen.

Waarom zijn we zo goed in het verzamelen van steeds meer informatie en macht, maar veel minder succesvol bij het vergaren van wijsheid? Door de eeuwen heen hebben allerlei culturen het idee gehad dat er een of andere fatale karakterfout in de mens zit die ons ertoe aanzet om macht te vergaren die we niet aankunnen. De Griekse mythe over Phaëton verhaalt van een jongen die ontdekt dat hij de zoon van de zonnegod Helios is. Hij wil zijn goddelijke afkomst aan de hele wereld tonen en eist het privilege op om de strijdswagen van de zon te besturen. Helios waarschuwt hem dat geen mens de goddelijke paarden die de zonnewagen trekken kan beheersen, maar Phaëton blijft aandringen tot de zonnegod toegeeft. Phaëton stijgt vol trots op, maar verliest dan inderdaad de controle over de strijdswagen. De zon raakt uit koers, verschroeit al het plantenleven, doodt talloze wezens en dreigt de hele aarde te verbranden. Zeus komt tussenbeide en treft Phaëton met een bliksemstraal. De arrogante jongeling stort als een vallende ster naar beneden, brandend als een fakkel. De goden nemen de controle over het hemelruim weer over en redden de wereld.

Tweeduizend jaar later, toen de industriële revolutie haar eerste stapjes zette en machines allerlei menselijke taken over begonnen te nemen, publiceerde Johann Wolfgang von Goethe het gedicht 'De tovenaarsleerling', dat een vergelijkbare waarschuwing inhield. Goethe's poëem (later gepopulariseerd als een Walt Disney-animatie met Mickey Mouse in de hoofdrol) gaat over een oude tovenaarsleerling die een jonge leerling

alleen laat in zijn werkplaats en hem wat klusjes geeft, zoals water halen uit de rivier. De leerling besluit het zichzelf wat gemakkelijker te maken en gebruikt een van de bezwingen van de tovenaars om een bezem te betoveren, zodat die het water voor hem kan halen. De leerling weet alleen niet hoe hij de bezem weer kan laten ophouden, dus die blijft maar water halen, tot de hele werkplaats dreigt te overstromen. De leerling hakt in paniek de betoverde bezem doormidden met een bijl, maar de twee helften worden elk weer een hele bezem. Nu zijn er twee bezems die de werkplaats overspoelen met water. Als de oude tovenaars terugkomt, smeekt de leerling hem om hulp: ‘De geesten die ik heb opgeroepen kan ik niet verdrijven.’ De tovenaars verbreekt onmiddellijk de betovering en maakt een einde aan de overstroming. De les voor de leerling – en de mensheid – is duidelijk: roep nooit krachten op die je niet kunt beheersen.

Wat vertellen de stichtelijke verhalen over de tovenaarsleerling en Phaëton ons in de eenentwintigste eeuw? Wij mensen hebben duidelijk geweigerd naar die waarschuwingen te luisteren. We hebben het klimaat van de aarde al uit balans gebracht en miljarden betoverde bezems, drones, chatbots en andere algoritmische geesten opgeroepen die zich aan onze controle kunnen onttrekken en een vloedgolf van onbedoelde gevolgen kunnen ontketen.

Wat moeten we dan doen? De verhalen bieden geen antwoorden, hoogstens dat we moeten wachten tot een of andere god of tovenaars ons komt redden. Dat is natuurlijk een uiterst gevaarlijke boodschap. Dat idee moedigt mensen aan om de verantwoordelijkheid van zich af te schuiven en blind op goden en tovenaars te vertrouwen. Erger nog, het gaat voorbij aan het feit dat goden en tovenaars zelf ook een menselijke uitvinding zijn, net als strijdwapens, bezems en algoritmen. De neiging om drastische dingen uit te vinden die onbedoelde gevolgen kunnen hebben is niet begonnen met de uitvinding van de stoommachine of AI, maar met de uitvinding van religie. Profeten en theologen hebben nu en dan machtige geesten opgeroepen die liefde en geluk moesten brengen, maar de wereld uiteindelijk overstromden met bloed.

De Phaëton-mythe en Goethes gedicht geven geen nuttige adviezen, omdat ze een verkeerde voorstelling geven van de manier waarop mensen

macht verkrijgen. In beide fabels verwerft één individu enorme krachten en wordt hij gecorrumpeerd door overmoed en hebzucht. De conclusie is dat onze gebrekkige menselijke psychologie tot machtsmisbruik leidt. Wat er verloren gaat in deze primitieve analyse is dat menselijke macht nooit voortvloeit uit individuele initiatieven. Macht vloeit altijd voort uit samenwerking tussen grote groepen mensen.

Onze individuele psychologie is dus niet de factor die tot machtsmisbruik leidt. Mensen zijn immers niet alleen in staat tot hebzucht, overmoed en wreedheid, maar ook tot liefde, mededogen, bescheidenheid en vreugde. Het is zeker waar dat de akeligste leden van onze soort zich laten leiden door hebzucht en wreedheid en dat die tot machtsmisbruik kunnen leiden, maar waarom zouden mensengemeenschappen ervoor kiezen om hun akeligste leden de macht te geven? De meeste Duitsers uit 1933 waren bijvoorbeeld geen psychopaten. Waarom kozen ze dan toch voor Hitler?

Onze neiging om onbeheersbare krachten op te roepen komt niet voort uit de individuele psyche, maar uit de unieke manier waarop onze soort in grote groepen kan samenwerken. Het voornaamste argument van dit boek is dat de mensheid enorm veel macht vergaart door het opzetten van grote samenwerkingsnetwerken, maar dat de manier waarop die netwerken zijn opgezet ze vatbaar maakt voor een onverstandig gebruik van die macht. Ons probleem is dus een netwerkprobleem.

Specifieker gesteld is het een informatieprobleem. Informatie is de lijm die netwerken bij elkaar houdt, maar sapiens hebben tienduizenden jaren lang grote netwerken opgezet en onderhouden door het bedenken en verspreiden van verzinsels, fantasieën en massale waanvoorstellingen – over goden, betoverde bezems, AI en nog veel meer. Individuele mensen willen doorgaans graag de waarheid over zichzelf en de wereld te weten komen, maar grote netwerken verbinden hun leden en creëren orde met behulp van verzinsels en fantasieën. Zo hebben we bijvoorbeeld het nazisme en het stalinisme gekregen. Dat waren uitzonderlijk sterke netwerken die bijeen werden gehouden door uitzonderlijke waanideeën. Onwetendheid is kracht, zoals het beroemde citaat van George Orwell luidt.

Het feit dat het naziregime en dat van Stalin op wrede verzinsels en

schaamteloze leugens gebaseerd waren, maakte ze zeker niet tot historische uitzonderingen en het maakte hun val ook niet onvermijdelijk. Het nazisme en het stalinisme waren twee van de sterkste netwerken die mensen ooit hebben opgezet. Eind 1941, begin 1942 wonnen de asmogendheden bijna de Tweede Wereldoorlog. Uiteindelijk kwam Stalin als overwinnaar uit de bus,¹ en in de jaren vijftig en zestig van de vorige eeuw maakten hij en zijn erfgenamen ook nog een redelijke kans om de Koude Oorlog te winnen. In de jaren negentig kregen liberale democratieën de overhand, maar dat begint steeds meer op een tijdelijke overwinning te lijken. In de eenentwintigste eeuw zou een nieuw totalitair regime heel goed kunnen slagen in iets wat Hitler en Stalin niet voor elkaar kregen, namelijk het opzetten van een almachtig netwerk waarin toekomstige generaties niet eens meer kunnen proberen om de onderliggende leugens en verzinsels aan de kaak te stellen. We moeten niet klakkeloos aannemen dat netwerken op basis van waanideeën automatisch gedoemd zijn. Als we hun overwinning willen voorkomen, zullen we daar hard voor moeten werken.

DE NAÏEVE KIJK OP INFORMATIE

Het is moeilijk om de kracht van waannetwerken goed in te schatten, en dat komt door een vrij algemeen gebrek aan kennis over de werking van grote informatienetwerken, of die nu waanideeën inzetten of niet. Dit gebrek aan kennis is ingekapseld in iets wat ik ‘de naïeve kijk op informatie’ noem. Fabels als de Phaëton-mythe en ‘De tovenaarsleerling’ geven een al te pessimistisch beeld van de individuele menselijke psychologie, maar de naïeve kijk op informatie verspreidt een overdreven optimistisch beeld van grootschalige menselijke netwerken.

De naïeve kijk stelt dat grote netwerken meer kennis vergaren over geneeskunde, natuurkunde, economie en talloze andere onderwerpen door veel meer informatie te verzamelen en verwerken dan individuen aankunnen, wat het netwerk niet alleen sterk maakt, maar ook wijs. Farmaceutische bedrijven en gezondheidszorginstellingen kunnen bij-

voorbeeld informatie over ziekteverwekkers verzamelen en daarmee de onderliggende oorzaken van veel ziekten bepalen, waardoor ze effectievere medicijnen kunnen ontwikkelen en slimmere beslissingen kunnen nemen over het gebruik daarvan. Volgens deze visie leidt voldoende informatie tot de waarheid, en waarheid leidt weer tot macht en wijsheid. Onwetendheid lijkt juist nergens toe te leiden. Netwerken die waanideeën of onwaarheden inzetten, kunnen nu en dan opkomen ten tijde van historische crises, maar op de langere termijn zullen ze het stevast verliezen van rivalen die scherpzinniger en eerlijker te werk gaan. Een zorginstelling die informatie over ziekteverwekkers negeert of een farmaceutische gigant die opzettelijk desinformatie verspreidt zal uiteindelijk verliezen van concurrenten die verstandiger met informatie omspringen. De naïeve kijk impliceert dus dat waannetwerken afwijkingen zijn en dat we er meestal wel op kunnen vertrouwen dat grote netwerken verstandig met hun macht omgaan.



De naïeve kijk op informatie

Uiteraard houdt de naïeve kijk er rekening mee dat er heel wat mis kan gaan op de weg van informatie naar waarheid. We kunnen altijd eerlijke vergissingen maken bij het verzamelen en verwerken van informatie. Kwaadwillige actoren kunnen vanuit hebzucht of haat belangrijke feiten verbergen of pogingen doen om ons te misleiden. Daardoor leidt informatie soms tot fouten en niet tot de waarheid. Door halve informatie, gebrekkige analyses of desinformatiecampagnes kunnen bijvoorbeeld zelfs experts verkeerde conclusies trekken over de ware oorzaak van een ziekte.

De naïeve kijk gaat er echter van uit dat de remedie voor de meeste problemen die we bij het verzamelen en verwerken van informatie te-

genkomen het verzamelen en verwerken van nog meer informatie is. Vergissingen zijn natuurlijk nooit helemaal uit te sluiten, maar in de meeste gevallen betekent meer informatie meer nauwkeurigheid. Eén arts die de oorzaak van een epidemie wil vaststellen en daarvoor één patiënt onderzoekt, zal minder snel in zijn opzet slagen dan duizenden artsen die gegevens over miljoenen patiënten verzamelen. En als de artsen samenspannen om de waarheid verborgen te houden, zal hun bedrog uiteindelijk uitkomen als medische informatie toegankelijker gemaakt wordt voor het grote publiek en voor onderzoeksjournalisten. Hoe groter het informatienetwerk, des te dichter het bij de waarheid moet liggen, zo luidt het standpunt.

Zelfs als we informatie correct analyseren en belangrijke waarheden ontdekken, garandeert dat natuurlijk nog niet dat we verstandig met die kennis zullen omspringen. Wijsheid wordt doorgaans gelijkgesteld aan het nemen van de juiste beslissingen, maar wat 'juist' betekent, hangt af van waardeoordelen die sterk variëren bij diverse volkeren, culturen of ideologieën. Wetenschappers die een nieuwe ziektekiem vinden, kunnen een vaccin ontwikkelen om de bevolking te beschermen. Maar als deze wetenschappers – of hun politieke bovenbazen – een racistische ideologie aanhangen die stelt dat sommige bevolkingsgroepen inferieur zijn en uitgeroeid moeten worden, dan kan die nieuwe medische kennis gebruikt worden om een biologisch wapen te ontwikkelen dat miljoenen slachtoffers maakt.

Ook in dit geval blijft de naïeve kijk op informatie erbij dat extra informatie het probleem minstens voor een deel oplost. De naïeve kijk denkt dat meningsverschillen over normen en waarden bij nader inzien ofwel te wijten zijn aan een gebrek aan informatie of aan opzettelijke desinformatie. Volgens deze visie zijn racisten slecht geïnformeerde mensen die de feiten over biologie en geschiedenis gewoon niet kennen. Ze denken dat 'ras' een valide biologische categorie is en ze zijn gehersenspoeld door valse complottheorieën. Racisme is dus te verhelpen door die mensen meer biologische en historische feiten te geven. Het kan wat tijd kosten, maar in een vrije informatiemarkt zal de waarheid vroeg of laat komen bovendrijven.

De naïeve kijk is natuurlijk genuanceerder en zinniger dan in een paar alinea's uitgelegd kan worden, maar het kernprincipe is dat informatie in essentie iets goeds is en hoe meer we ervan hebben, hoe beter. Met genoeg informatie en genoeg tijd, zullen we vanzelf de waarheid ontdekken over dingen die variëren van virusinfecties tot racistische vooroordelen, waarmee we niet alleen onze macht vergroten, maar ook de wijsheid die nodig is om die macht goed te gebruiken.

Deze naïeve kijk rechtvaardigt het najagen van steeds krachtiger informatietechnologieën en het is sinds jaar en dag de semiofficiële ideologie van het computertijdperk en het internet. In juni 1989, een paar maanden voor de val van de Berlijnse Muur en het IJzeren Gordijn, verklaarde Ronald Reagan dat 'de Goliath van de totalitaire controle razendsnel ten val gebracht zal worden door de David van de microchip' en dat 'de grootste Big Brother steeds hulpelozer tegenover communicatietechnologie staat. [...] Informatie is de zuurstof van de moderne tijd. [...] Ze dringt door muren die bezet zijn met prikkeldraad. Ze dwarrelt over grenzen met stroomhekken en boobytraps. Briesjes van elektronische stralen blazen door het IJzeren Gordijn alsof het van gaas is.'² In november 2009 sprak Barack Obama in dezelfde trant tijdens een bezoek aan Sjanghai, waar hij tegen zijn Chinese gastheren zei: 'Ik ben een groot gelover in technologie en ik ben een groot gelover in openheid als het om informatiestromen gaat. Ik denk dat de samenleving alleen maar sterker wordt als de informatie vrijer stroomt.'³

Ondernemers en grote bedrijven hebben vaak al even rooskleurige visies op informatietechnologie gegeven. In 1858 stond al in een redactioneel artikel van *The New Englander* over de uitvinding van de telegraaf: 'Het is onmogelijk dat oude vooroordelen en vijandigheden nog langer kunnen bestaan nu zo'n instrument is geschapen voor het uitwisselen van gedachten tussen alle naties op aarde.'⁴ Bijna twee eeuwen en twee wereldoorlogen later zei Mark Zuckerberg dat het doel van Facebook is 'mensen te helpen meer te delen om de wereld opener te maken en het begrip tussen mensen te helpen bevorderen'.⁵

De eminente futuroloog en ondernemer Ray Kurzweil geeft in zijn boek *The Singularity is Nearer* uit 2024 een overzicht van de geschiedenis

van informatietechnologie en concludeert dat ‘de realiteit is dat bijna alle aspecten van ons leven steeds beter worden als gevolg van de exponentiële verbeteringen binnen de technologie’. Terugkijkend op de lange, rijke menselijke geschiedenis gebruikt hij voorbeelden als de uitvinding van de drukpers om aan te voeren dat informatietechnologie van nature neigt tot het voortbrengen van ‘een cirkel van deugdzaamheid die bijna alle aspecten van het menselijk welzijn bevordert, onder meer geletterdheid, onderwijs, welvaart, volksgezondheid, democratisering en een afname van geweld’.⁶

De naïeve kijk op informatie wordt misschien het bondigst samengevat in Googles doelstelling om ‘de informatie op aarde te organiseren en universeel toegankelijk en nuttig te maken’. Googles antwoord op Goethes waarschuwingen is dat één leerling die het geheime toverboek van zijn meester jat zeker tot rampen zal leiden, maar een hele hoop leerlingen die vrije toegang hebben tot alle informatie op aarde niet alleen handige betoverde bezems zullen creëren, maar ook zullen leren hoe ze daar verstandig mee moeten omgaan.

GOOGLE VERSUS GOETHE

Het moet gezegd worden dat er talloze gevallen zijn waarin meer informatie mensen inderdaad de kans gaf om de wereld beter te begrijpen en hun macht verstandiger te gebruiken. Denk bijvoorbeeld aan de drastische afname van de kindersterfte. Johann Wolfgang von Goethe was de oudste van zeven kinderen, maar alleen hij en zijn zus Cornelia haalden hun zevende verjaardag. Hun broer Hermann Jacob stierf op zijn zesde door ziekte, hun zus Catharina Elisabeth op haar vierde, hun zus Johanna Maria op tweejarige leeftijd, hun broer Georg Adolf met acht maanden en een vijfde, naamloos gebleven broertje was dood geboren. Cornelia stierf op haar zesentwintigste door ziekte, zodat Johann Wolfgang het enige overlevende kind uit het hele gezin was.⁷

Johann Wolfgang von Goethe kreeg zelf vijf kinderen, die op zijn

oudste zoon August na allemaal binnen twee weken na hun geboorte sterven. Naar alle waarschijnlijkheid kwam dat doordat de bloedgroepen van Goethe en zijn vrouw Christiane onverenigbaar waren, waardoor de moeder na de eerste succesvolle zwangerschap antilichamen tegen het bloed van de foetussen ontwikkelde. Deze aandoening, die rhesusziekte genoemd wordt, is tegenwoordig zo effectief te behandelen dat het sterftecijfer onder de twee procent ligt, maar rond 1790 was de gemiddelde sterfte vijftig procent en voor Goethes vier jongste kinderen was het een doodvonnis.⁸

Al met al lag de overlevingskans van de kinderen in Goethes familie – een welvarende Duitse familie aan het eind van de achttiende eeuw – op een schrikbarende vijftienvijftig procent. Slechts drie van de twaalf kinderen bereikten de volwassen leeftijd. Dit gruwelijke statistische feit was niet uitzonderlijk. Rond de tijd dat Goethe ‘De tovenaarsleerling’ schreef, in 1797, bereikte naar schatting maar vijftig procent van de Duitse kinderen de leeftijd van vijftien⁹ en hetzelfde gold waarschijnlijk voor de meeste andere delen van de wereld.¹⁰ In 2020 overleefde 95,6 procent van de kinderen op de hele wereld hun vijftiende verjaardag¹¹ en in Duitsland was dat cijfer 99,5 procent.¹² Dit doorslaande succes was niet mogelijk geweest zonder het verzamelen, analyseren en delen van gigantische hoeveelheden medische gegevens over dingen als bloedgroepen. In dit geval bleek de naïeve kijk op informatie dus juist.

Toch ziet de naïeve kijk op informatie maar een deel van het hele plaatje en de geschiedenis van de moderne tijd draaide niet alleen om het terugdringen van de kindersterfte. In recente generaties heeft de mensheid de grootste toename aller tijden in zowel de hoeveelheid als de snelheid van onze informatieproductie meegemaakt. Elke smartphone bevat meer informatie dan de bibliotheek van Alexandrië in de klassieke tijd¹³ en geeft zijn eigenaar instant verbinding met miljarden andere mensen over de hele wereld. Maar ondanks al die informatie die met adembemende snelheid circuleert, is de mensheid dichter bij de totale zelfvernietiging dan ooit.

Ondanks – of misschien dankzij – onze schat aan data blijven we broeikasgassen de atmosfeer in pompen, rivieren en oceanen vervuilen, bossen kappen, complete leefgebieden vernietigen, talloze soorten tot

uitsterven drijven en de ecologische basis van onze eigen soort in gevaar brengen. We produceren ook steeds sterkere massavernietigingswapens, van thermonucleaire bommen tot dodelijke virussen. Het ontbreekt onze leiders niet aan informatie over die gevaren, maar in plaats van samen te werken aan oplossingen, begeven ze zich steeds verder in de richting van een wereldoorlog.

Zou alles beter worden van meer informatie – of slechter? Daar zullen we spoedig achter komen. Talloze bedrijven en overheden zijn verwikkeld in een race om de krachtigste informatietechnologie uit de geschiedenis te ontwikkelen – AI. Sommige toonaangevende ondernemers, zoals de Amerikaanse investeerder Marc Andreessen, geloven dat AI uiteindelijk alle problemen van de mensheid zal oplossen. Op 6 juni 2023 publiceerde Andreessen een artikel met de titel ‘Why AI will save the world’, dat vol staat met boude beweringen als ‘Hier is het goede nieuws: AI zal de wereld niet verwoesten, maar die misschien juist redden’ en ‘AI kan alles wat ons lief is beter maken’. Zijn conclusie: ‘De ontwikkeling en uitbreiding van AI is geen risico dat we moeten vrezen, maar een morele plicht die we onszelf, onze kinderen en onze toekomst verschuldigd zijn.’¹⁴

Ray Kurzweil is het met hem eens en voert in *The Singularity is Nearer* aan dat ‘AI de centrale technologie is waarmee we de prangende problemen van nu het hoofd kunnen bieden, waaronder het overwinnen van ziekte, armoede, milieuschade en al onze menselijke zwakheden. We hebben de morele verplichting om de belofte van deze nieuwe technologieën te realiseren’. Kurzweil is zich sterk bewust van de potentiële gevaren van de technologie en hij analyseert ze uitgebreid, maar hij gelooft dat ze succesvol ingeperkt kunnen worden.¹⁵

Anderen zijn sceptischer. Niet alleen filosofen en sociale wetenschappers, maar ook veel toonaangevende AI-deskundigen en ondernemers als Yoshua Bengio, Geoffrey Hinton, Sam Altman, Elon Musk en Mustafa Suleyman hebben het grote publiek gewaarschuwd dat AI onze beschaving kan vernietigen.¹⁶ Een artikel uit 2024 van Bengio, Hinton en heel wat andere deskundigen stelde dat ‘de onbelemmerde voortgang van AI zou kunnen culmineren in een grootschalig verlies van leven en de biosfeer, en marginalisering of zelfs het uitsterven van de mensheid’.¹⁷ In een

enquête uit 2023 onder 2778 AI-onderzoekers schatte meer dan een derde de kans dat geavanceerde AI kan leiden tot zoiets ergs als het uitsterven van de mens op minstens tien procent.¹⁸ In 2023 hebben bijna dertig overheden – waaronder die van China, de Verenigde Staten en het Verenigd Koninkrijk – de Bletchley-verklaring over AI ondertekend, waarin onderkend wordt dat ‘er potentieel ernstige, zelfs catastrofale schade, opzettelijk dan wel onbedoeld, kan voortvloeien uit de meest significante capaciteiten van deze AI-modellen’.¹⁹ Het gebruik van zulke apocalyptische termen door experts en overheden is niet ingegeven door de wens om een Hollywood-beeld te scheppen van opstandige robots die door de straten racen en mensen afknallen. Dat is niet zo’n aannemelijk scenario en het leidt mensen alleen maar af van de echte gevaren. Deskundigen waarschuwen juist voor twee andere scenario’s.

Ten eerste kan AI bestaande menselijke conflicten op scherp zetten, waardoor de mensheid verdeeld raakt. Zoals het IJzeren Gordijn in de twintigste eeuw de rivaliserende grootmachten verdeelde in de Koude Oorlog, zo kan in de eenentwintigste eeuw het Digitale Gordijn – gemaakt van computerchips en -programma’s in plaats van prikkeldraad – rivaliserende machten verdelen in een nieuw wereldwijd conflict. De AI-wedloop zal steeds destructievere wapens produceren, waardoor zelfs een klein vonkje een catastrofale vlammenzee kan veroorzaken.

Ten tweede zal het Digitale Gordijn misschien niet de ene groep mensen tegen de andere opzetten, maar eerder alle mensen tegen de nieuwe AI-heerschappij. Waar we ook wonen, het zou kunnen dat we verstrikt raken in een web van ondoorgrondelijke algoritmen die ons leven beheersen, onze politiek en cultuur omvormen en zelfs ons lichaam en onze geest modificeren, terwijl wij de machten die ons beheersen niet meer kunnen begrijpen, laat staan dat we ze kunnen tegenhouden. Als een eenentwintigste-eeuws totalitair netwerk de wereld weet te veroveren, zou dat gerund kunnen worden door een niet-menselijke intelligentie in plaats van door een menselijke dictator. Mensen die China, Rusland of een postdemocratische vs aanwijzen als hun belangrijkste bron van totalitaire nachtmerries hebben het gevaar niet helder voor ogen. Chinezen, Russen, Amerikanen en alle andere mensen worden evenzeer bedreigd

door het totalitaire potentieel van niet-menselijke intelligentie.

Het gevaar is zo groot dat AI alle mensen op aarde zou moeten interesseren. Niet iedereen kan een AI-expert worden, maar we moeten met zijn allen voor ogen houden dat AI de eerste technologie aller tijden is die zelfstandig beslissingen kan nemen en met nieuwe ideeën kan komen. Alle voorgaande menselijke uitvindingen hebben mensen meer macht verschaft, want hoe krachtig het nieuwe werktuig ook was, we hadden de beslissingen over het gebruik ervan zelf in de hand. Messen en bommen beslissen niet zelf wie ze vermoorden. Het zijn domme werktuigen, zonder de intelligentie die nodig is om informatie te verwerken en zelfstandige beslissingen te nemen. AI kan daarentegen wel op eigen houtje informatie verwerken en kan mensen dus vervangen in het beslisproces. AI is geen werktuig, het is een actor.

Door die mate van informatiebeheersing kan AI ook zelfstandig nieuwe ideeën genereren op de meest uiteenlopende terreinen, van muziek tot geneeskunde. Van oudsher speelden grammofoons onze platen af en microscopen onthulden de geheimen van onze cellen, maar grammofoons konden geen nieuwe symfonieën componeren en microscopen geen nieuwe medicijnen creëren. AI kan inmiddels al kunst voortbrengen en zelfstandig wetenschappelijke ontdekkingen doen. De komende decennia zal het waarschijnlijk zelfs het vermogen ontwikkelen om nieuwe levensvormen te creëren, ofwel door het schrijven van genetische codes of door het uitvinden van anorganische computercodes die anorganische entiteiten tot leven brengen.

Zelfs nu de AI-revolutie nog in een embryonaal stadium verkeert, kunnen computers al beslissingen over ons nemen – om ons al of niet een hypotheek te geven, om ons aan te nemen voor een baan of om ons naar de gevangenis te sturen. Die trend zal alleen maar groeien en versnellen, zodat het steeds moeilijker wordt om nog iets van ons eigen leven te begrijpen. Kunnen we erop vertrouwen dat computeralgoritmen verstandige beslissingen nemen en een betere wereld creëren? Dat is een veel grotere gok dan een betoverde bezem water te laten halen. En we spelen niet alleen met mensenlevens. AI zou niet alleen de loop van onze menselijke geschiedenis kunnen veranderen, maar de evolutie van alle levende wezens.

INFORMATIE ALS WAPEN

In 2016 publiceerde ik *Homo deus*, een boek over enkele gevaren waarvoor de nieuwe informatietechnologieën de mensheid konden stellen. Daarin voerde ik aan dat de echte held van de geschiedenis altijd informatie is geweest in plaats van *Homo sapiens* en dat wetenschappers niet alleen de geschiedenis, maar ook biologie, politiek en economie steeds meer bekijken in termen van informatiestromen. Dieren, staten en markten zijn allemaal informatienetwerken die gegevens opnemen vanuit hun omgeving, beslissingen nemen en vervolgens weer gegevens de wereld in sturen. Mijn boek bevatte de waarschuwing dat betere informatietechnologie mogelijk macht van ons kan afpakken en onze fysieke en mentale gezondheid kan schaden, ondanks onze hoop dat we er gezondheid, geluk en macht mee kunnen verwerven, en de hypothese dat de mens, als we niet oppassen, wel eens kan oplossen in de stortvloed van informatie, als een kluit aarde in een snelstromende rivier, en dat de mensheid achteraf hoogstens een rimpeling in de kosmische datastroom geweest zal zijn.

In de jaren na de publicatie van *Homo deus* is het tempo van die veranderingen alleen maar toegenomen en is er inderdaad macht van mensen op algoritmen overgegaan. Veel scenario's die in 2016 als sciencefiction klonken – zoals algoritmen die kunstwerken kunnen maken, zich als mensen kunnen voordoen, cruciale beslissingen over ons leven nemen en meer van ons weten dan wijzelf – zijn in 2024 de alledaagse werkelijkheid geworden.

Er zijn sinds 2016 nog veel andere dingen veranderd. De klimaatcrisis is verergerd, internationale spanningen zijn geëscaleerd en een golf van populisme heeft de samenhang van zelfs de meest robuuste democratieën ondermijnd. Het populisme is ook een radicale klap in het gezicht gebleken voor de naïeve kijk op informatie. Populistische leiders als Donald Trump en Jair Bolsonaro, en populistische bewegingen en complotdenkers als QAnon en de antivaxxers, beweren dat alle traditionele instellingen die gezag ontleen aan de claim dat ze informatie vergaren en waarheden ontdekken keihard liegen. Bureaucraten, rechters, artsen, mainstreamjournalisten en wetenschappers behoren tot een samen-

spannende elite die niet geïnteresseerd is in de waarheid en opzettelijk desinformatie verspreidt om macht en privileges te vergaren over de rug van ‘het volk’ heen. De opkomst van politici als Trump en bewegingen als QAnon heeft een specifieke politieke context die uniek is voor de omstandigheden in de Verenigde Staten aan het eind van de jaren tien, maar populisme als anti-establishmentdenken bestond al lang voor Trump en is relevant voor talloze andere historische omstandigheden, nu en in de toekomst. Kortom, het populisme beschouwt informatie als wapen.²⁰

Informatie → Macht

De populistische kijk op informatie

In zijn extremere verschijningsvormen stelt het populisme dat er überhaupt niet zoiets bestaat als objectieve waarheid en dat iedereen ‘zijn eigen waarheid’ heeft, die aangewend wordt om rivalen te bedwingen. Volgens dit wereldbeeld is macht de enige realiteit. Alle sociale interacties zijn een machtsstrijd, omdat mensen uitsluitend in macht geïnteresseerd zijn. Als iemand beweert dat hij geïnteresseerd is in iets anders – zoals waarheid of rechtvaardigheid – is dat alleen maar een trucje om macht te verwerven. Overal en altijd waar het populisme erin slaagt om dat idee van informatie als wapen te verspreiden, wordt de taal zelf ondermijnd. Zelfstandige naamwoorden als ‘feiten’ en adjectieven als ‘correct’ en ‘waarheidsgetrouw’ worden ongrijpbaar. Zulke woorden worden niet meer in verband gebracht met een gemeenschappelijke objectieve werkelijkheid. Alle uitspraken over ‘feiten’ of ‘waarheden’ stuiten dan onherroepelijk op vragen als: ‘Over wiens feiten en wiens waarheid heb je het?’

Het is belangrijk om te bedenken dat deze op macht gerichte en uiterst sceptische kijk op informatie geen nieuw fenomeen is en niet is uitgevonden door antivaxxers, flat-earthers, bolsonarista’s of Trump-aanhangers. Vergelijkbare ideeën zijn al lang voor 2016 gepropageerd, soms door bijzonder vooraanstaande denkers.²¹ Eind twintigste eeuw beweerden radi-

caal-linkse intellectuelen als Michel Foucault en Edward Said bijvoorbeeld dat wetenschappelijke instellingen als ziekenhuizen en universiteiten geen tijdloze, objectieve waarheden zoeken, maar hun macht gebruiken om te bepalen wat voor waarheid doorgaat, ten dienste van kapitalistische en kolonialistische elites. Hun radicale kritieken mondden soms zelfs uit in beweringen dat ‘wetenschappelijke feiten’ niets meer of minder zijn dan een kapitalistisch of kolonialistisch ‘discours’, dat machthebbers niet werkelijk geïnteresseerd kunnen zijn in waarheid en dat we er nooit op moeten vertrouwen dat ze hun eigen fouten zullen inzien en rechtzetten.²²

Deze specifieke tak van links-radicaal denken gaat terug op Karl Marx, die halverwege de negentiende eeuw aanvoerde dat macht de enige realiteit is, dat informatie een wapen is en dat elites die naar eigen zeggen waarheid en gerechtigheid dienen in feite hun eigen kleine klassenprivileges najagen. In de woorden van het *Communistisch manifest* uit 1848: ‘De geschiedenis van alle tot nu toe bestaande samenlevingen is de geschiedenis van de klassenstrijd. Vrij man en slaaf, patriciër en plebejer, heer en lijfeigene, gildemeester en gezel, in één woord: onderdrukker en onderdrukte, stonden voortdurend tegenover elkaar in een onophoudelijke strijd, soms verborgen, soms openlijk.’ Deze binaire interpretatie van de geschiedenis impliceert dat elke menselijke interactie een machtsstrijd is tussen onderdrukkers en onderdrukten. Elke keer dat iemand iets zegt, is de vraag dan dus niet: ‘Wat wordt er gezegd? Is het waar?’ maar: ‘Wie zegt dit? Wiens privileges worden hiermee gediend?’

Het is natuurlijk niet erg waarschijnlijk dat rechtse populistten als Trump en Bolsonaro ooit iets van Foucault of Marx hebben gelezen en ze presenteren zichzelf juist als felle antimarxisten. Ze verschillen ook enorm van marxisten in hun voorgestelde beleidsmaatregelen op het gebied van belastingen en uitkeringen. Maar hun basale kijk op de samenleving en op informatie is verrassend marxistisch, want ze zien alle menselijke interacties als een machtsstrijd tussen onderdrukkers en onderdrukten. Trump verklaarde in zijn inauguratietoespraak van 2017 bijvoorbeeld dat ‘een kleine groep mensen in onze nationale hoofdstad heeft geprofiteerd van zijn regeringspositie, terwijl het volk de lasten droeg’.²³ Dat soort retoriek is schering en inslag bij het populisme, dat de politicoloog Cas

Mudde omschreef als een ‘ideologie volgens welke de maatschappij uiteindelijk verdeeld wordt in twee homogene en vijandige kampen – “het zuivere volk” versus “de corrupte elite”²⁴. Zoals marxisten claimden dat de media fungeren als spreekbuis voor de kapitalistische klasse en dat wetenschappelijke instellingen als universiteiten desinformatie verspreiden om de kapitalistische overheersing te bestendigen, beschuldigen populistten diezelfde instellingen ervan dat ze de belangen van de ‘corrupte elites’ dienen, ten koste van ‘het volk’.

De populistten van nu lijden aan hetzelfde euvel als de radicale anti-establishmentbewegingen uit voorgaande generaties. Als macht de enige realiteit is en als informatie alleen een wapen is, wat zegt dat dan over de populistten zelf? Zijn zij ook uitsluitend geïnteresseerd in macht en liegen zij ons ook alleen maar voor om aan de macht te komen?

Populisten hebben op twee manieren geprobeerd zich uit dit dilemma te redden. Sommige populistische bewegingen scharen zich naar eigen zeggen achter de idealen van de moderne wetenschap en de traditie van sceptisch empirisme. Ze zeggen dat je inderdaad nooit moet vertrouwen op instellingen of gezagsdragers – ook niet op zelfverklaarde populistische partijen en politici. Je moet ‘je eigen research doen’ en alleen vertrouwen wat je met eigen ogen kunt zien.²⁵ Die radicaal-empirische houding impliceert dat grootschalige instanties als politieke partijen, rechtbanken, kranten en universiteiten pertinent niet te vertrouwen zijn, maar dat individuen die er even de moeite voor doen nog steeds zelf de waarheid kunnen vinden.

Die aanpak klinkt misschien wetenschappelijk en zal vrije geesten misschien aanspreken, maar de vraag blijft dan wel hoe mensengemeenschappen kunnen samenwerken om een systeem voor gezondheidszorg op te zetten of milieuwetgeving door te voeren, want dat vereist grootschalige institutionele organisatie. Kan één individu alle benodigde research plegen om te bepalen of de aarde opwarmt en wat daaraan gedaan moet worden? Hoe zou één persoon klimaatgegevens van over de hele wereld moeten verzamelen, laat staan betrouwbare cijfers over voorbije eeuwen? ‘Op je eigen research vertrouwen’ klinkt misschien wetenschappelijk, maar in de praktijk komt het neer op het geloof dat er geen objectieve waarheid is.