

De militaire relevantie van een technologische singulariteit

Technologische ontwikkelingen zijn de belangrijkste factor achter brede veranderingen in de samenleving. De technologische vooruitgang zou nu zo snel gaan, dat deze volgens sommigen binnen afzienbare tijd zal leiden tot een technologische singulariteit: een hypothetisch moment waarop, als gevolg van versmelting van mens en machine en het ontstaan van bovenmenselijke intelligentie, de samenleving in alle facetten onbegrijpelijk snel verandert. De hypothese van een technologische singulariteit is wetenschappelijk moeilijk toetsbaar. Maar de hypothese is militair gezien wel bruikbaar en kan voor Defensie als kader dienen om in te spelen op technologische ontwikkelingen op het gebied van kunstmatige intelligentie, robotica, nanotechnologie, genetica en neurotechnologie. Het zijn namelijk deze technologiegebieden die ten grondslag liggen aan twee aspecten die de komende decennia in hoge mate bepalend zullen zijn voor de aard en effectiviteit van militair vermogen: het gebruik van autonome systemen en het combineren en integreren van mensen en intelligente machines.

*T.L.A. Schless MSc – majoor van de Koninklijke Luchtmacht**

'We can only see a short distance ahead, but we can see plenty there that needs to be done' – Alan Turing

De technologische vooruitgang is de belangrijkste factor achter brede veranderingen in de samenleving en geeft mondiaal een grote impuls aan uiteenlopende wetenschapsgebieden en industrieën, aan ontwikkelingen in de economie en aan de inrichting van de maatschappij.¹ De technologische groei zou nu zo snel gaan, dat deze volgens sommigen binnen afzienbare tijd zal leiden tot een technologische singulariteit: een hypothetisch moment waarop, als gevolg van versmelting van mens en machine en het ontstaan van bovenmenselijke intelligentie, de samenleving in alle facetten onbegrijpelijk snel verandert. Het zou

betekenen dat onze huidige notie van concepten als kennis, economie, identiteit, en mogelijk zelfs van het leven niet meer geldig is. In die zin refereert de term technologische singulariteit aan een wiskundige of fysische singulariteit: een punt in tijd en ruimte waar een waarde ongedefinieerd is of natuurwet niet meer opgaat.

De mogelijkheid van een technologische singulariteit krijgt steeds meer aandacht, ook bij Defensie. Dit inleidende artikel over het onderwerp is gebaseerd op de vraag in hoeverre de singulariteitshypothese bruikbaar en

* Timo Schless is verbonden aan het Innovatiecentrum AIR van de Koninklijke Luchtmacht.

¹ Wim de Ridder, *De ontdekking van de toekomst: wat we al weten, is niet te geloven* (Deventer, Vakmedianet, 2014).



FOTO: MCD, V. KUIJPERS

Neil Jacobstein, voorzitter van de faculteit Artificial Intelligence & Robotics van de Singularity University, spreekt op een door de Koninklijke Luchtmacht en Defensie Materieel Organisatie georganiseerd symposium

relevant is vanuit een militair perspectief. Deze vraag lag ten grondslag aan een vooronderzoek dat is uitgevoerd door de luchtmacht om focus te geven aan verder onderzoek en beleid op het gebied van disruptieve technologieën die bepalend zijn om in de toekomst militair effectief te kunnen blijven in en vanuit de lucht en de ruimte.

Dit artikel geeft eerst een overzicht van technologieën die voor een technologische singulariteit van belang zijn. Het zijn geen specifiek militaire ontwikkelingen, maar de technologische vooruitgang brengt vanwege de significante sociale en economische impact op de mondiale samenleving hoe dan ook nieuwe

veiligheidsvraagstukken met zich mee. Bovendien leiden opkomende technologieën onvermijdelijk tot nieuwe militaire toepassingen. Vervolgens gaat het artikel in op de onderliggende aanname dat technologische vooruitgang exponentieel verloopt. Daarna komen de symptomen en tegenargumenten van een singulariteit aan bod. Het artikel sluit af met de impact, de risico's en de militaire relevantie.

Technologische ontwikkelingen

De eerste gedachten over een technologische singulariteit dateren uit de jaren vijftig en zestig van de vorige eeuw,² maar de hypothese is voor het eerst uitgebreid beschreven door Vernor Vinge in 1993. Hij ziet de singulariteit als het gevolg van het ontstaan van bovenmenselijke intelligentie: een gebeurtenis die

² Irving John Good, *Speculations Concerning the First Ultra-intelligent Machine* (1964).

vergelijkbaar is met die van het ontstaan van leven en met het ontstaan van de intelligente, zelfbewuste mens.³ Wat de aard van een bovenmenselijke intelligentie is, kan hij niet beschrijven, maar hij ziet verschillende manieren waarop dit kan worden bereikt:

- computers die intelligenter zijn dan een mens en zelfbewust worden;
- netwerken van computers met een bovenmenselijke intelligentie;
- koppelingen die mensen en computers zo nauw laten samenwerken dat ze gezamenlijk bovenmenselijk intelligent zijn;
- biologische manieren om de menselijke intelligentie te verbeteren.

De overlap in de ontwikkelingen van de genetica, nanotechnologie en robotica zal volgens Ray Kurzweil uiteindelijk tot een technologische singulariteit leiden. Hij licht daarbij toe dat met robotica hier vooral de kunstmatige intelligentie wordt bedoeld en niet zozeer de fysieke verschijning.⁴ Murray Shanahan geeft aan dat kunstmatige intelligentie en neurotechnologie de twee gebieden zijn die het meest significant bijdragen aan het bereiken van een technologische singulariteit.⁵ In beide perspectieven staat de kunstmatige intelligentie (*Artificial Intelligence*, AI) centraal, waarbij deze niet los wordt gezien van interactie met de fysieke en levende wereld op nanoschaal.

Intelligentie is het vermogen om te leren, te begrijpen, problemen op te lossen en beslissingen te nemen; we spreken van kunstmatige intelligentie als een machine dit vermogen heeft.⁶ Waar een geautomatiseerd systeem werkt op basis van vooraf gedefinieerde handelingen, onderscheidt een intelligente machine zich door het vermogen waarmee die zelf tot inzichten en oplossingen komt. Merk op dat de hier gebruikte definitie niet op voorhand vereist dat die machine ook intelligentie-gerelateerde biologische eigenschappen heeft, zoals zelfbewustzijn en emoties.

Kunstmatige intelligentie bestaat als wetenschap en technologie al meer dan zeventig jaar.

McCulloch en Pitts toonden in de jaren veertig niet alleen aan dat hun model van kunstmatige neurale netwerken feitelijk ‘Turing-compleet’ was (dus dat iedere berekenbare functie in beginsel door een neuraal netwerk kan worden opgelost), maar ook dat dit soort netwerken lerende eigenschappen had.⁷

Tot ver in de jaren zestig van de twintigste eeuw werd gewerkt aan neurale netwerken en andere abstracte methoden om algemene oplossingen te vinden voor een grote klasse van problemen, of wat Marvin Minsky noemde ‘the design or programming of machines to work on problems for which the designer does not have, in advance, practical methods of solution’. Maar vanwege de hoge mate van abstractie en de beperkte capaciteit bleef de praktische toepasbaarheid van deze kenniszwakke systemen achter. Na een periode van grote toekomstbeelden zakten de verwachtingen en daarmee de investeringen in (onderzoek naar) kunstmatige intelligentie in.

In de jaren zeventig kwamen de expertsystemen op, waarin kennis van experts en bekende feiten worden opgeslagen. Op basis daarvan kan een redeneermechanisme tot de gevraagde oplossingen komen. Hoewel ze in specifieke gebieden grote meerwaarde hebben gehad, zijn het geen lerende systemen en is de toepasbaarheid beperkt gebleven tot afgebakende kennisdomeinen. Vanaf eind jaren tachtig werden deze beperkingen deels overwonnen met behulp van *fuzzy logic*. Dat stelt systemen tot op zekere hoogte in staat te werken met niet-exacte uitdrukkingen zoals die veel voorkomen in de menselijke taal, zodat het ook mogelijk werd niet-volledig kwantificeerbare kennis van meerdere experts vast te leggen.

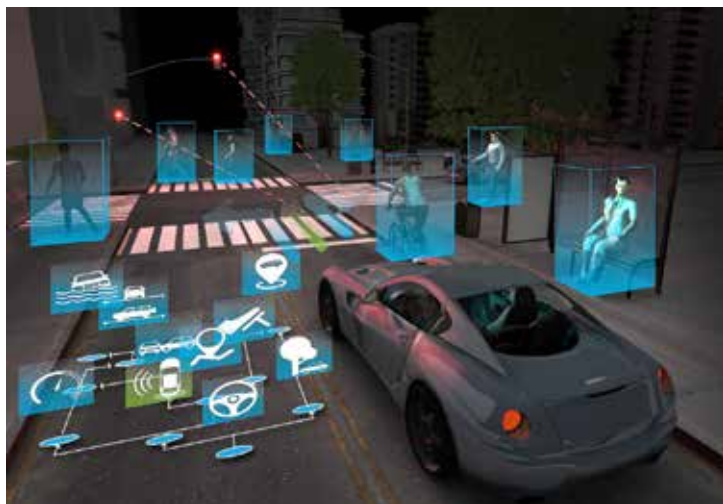
3 Vernor Vinge, *The coming technological singularity: How to survive in the post-human era* (1993).

4 Ray Kurzweil, *The Singularity Is near: When Humans Transcend Biology* (New York, Viking, 2005).

5 Murray Shanahan, *The technological singularity* (Cambridge, MA, The MIT Press, 2015).

6 Michael Negnevitsky, *Artificial Intelligence: A Guide to Intelligent Systems* (Harlow, Addison Wesley/Pearson, 2011).

7 Warren S. McCulloch en Walter Pitts, ‘A logical calculus of the ideas immanent in nervous activity’, in: *Bulletin of Mathematical Biophysics* 5 (1943) 115-33; [Negnevitsky, *Artificial Intelligence*, 5.](#)



ILLUSTRATIE DARPA

Lerende systemen hebben de afgelopen jaren steeds meer toepassingen gevonden, onder meer in zelfrijdende auto's

De toename van de rekencapaciteit leidde vanaf de jaren tachtig tot hernieuwde belangstelling voor neurale netwerken. In de afgelopen dertig jaar is dit uitgegroeid tot een breed scala aan technieken en algoritmen voor lerende systemen, zoals regelgebaseerde systemen, fuzzy logic, statistische methoden, evolutionaire systemen en neurale netwerken. Deze zijn, al dan niet gecombineerd, succesvol in steeds geavanceerdere toepassingen waarbij software zonder voorgeprogrammeerde oplossingen besluiten neemt. Bekende doorbraken waren de computers die de beste specialisten versloegen met spellen als schaken, Jeopardy! en go in achtereenvolgens 1997, 2011 en 2016. Inmiddels worden lerende systemen dagelijks gebruikt in onder meer (zelfrijdende) auto's, zoekmachines, operating systems, beeldherkenning, beurshandel, tekstinterpretatie- en vertaling, pratende digitale assistenten zoals Siri en Cortana, en steeds meer in wetenschap en gezondheidszorg.⁸

Deze voorbeelden van kunstmatige intelligentie die alleen werken binnen een afgebakend gebied, worden zwakke AI genoemd. De grote

vraag voor de singulariteitshypothese is of sterke AI mogelijk is: kunstmatige intelligentie die niet beperkt is tot een bepaald domein, maar algemene intelligentie bezit en daarmee zo adaptief is dat het in brede zin kennis en vaardigheden kan opdoen zoals mensen dat kunnen. Er zijn verschillende trends die op zichzelf en gecombineerd bijdragen aan de ontwikkeling van sterke AI:

- de doorontwikkeling en verbreding van zwakke AI;
- het gebruik van big data: omdat kunstmatige intelligentie digitaal van aard is, heeft het directe toegang tot zeer veel informatie op de meest uiteenlopende gebieden;
- de toepassing van biologische voorbeelden in technologie, zoals sensormotoriek in de robotica, breinsimulatie met hardware, en evolutionaire algoritmen;
- de groeiende verwevenheid van de digitale en fysieke wereld, en daarmee van AI met de fysieke wereld, onder meer door de robotica, 3D-printing, het *internet-of-things* en nano- en neurotechnologie.

Naast de groeiende verwevenheid van het informatiedomein met de fysieke wereld, wordt ook de afstand tussen informatietechnologie en levende organismen steeds kleiner. Neurotechnologie is het domein dat kunstmatige systemen rechtstreeks verbindt met een zenuwstelsel of brein, waardoor informatie kan worden uitgewisseld tussen biologische organismen en digitale systemen voor wederzijdse beïnvloeding. Tot nu toe heeft onderzoek hiernaar vooral plaatsgevonden in de medische wetenschap, die zich primair richt op het functioneren van het menselijk brein om ziektes en stoornissen te kunnen behandelen. Onderzoek en investeringen in neurotechnologie voor niet-medische toepassingen nemen toe en de verwachting is dat de mogelijkheden die neurotechnologie biedt op veel gebieden binnen afzienbare tijd manifest zullen worden. Hierbij moet men denken aan directe brein-tot-brein-interactie tussen mensen, brein-machine-koppelingen en het verbeteren van menselijke intelligentie, geheugen, alertheid en omgevingsbewustzijn.⁹

⁸ Thomas G. Dietterich en Eric J. Horvitz, 'Rise of Concerns about AI: Reflections and Directions', in: *Communications of the ACM* 58 (2015) (10) 38-40.

Regelgebaseerde systemen	<i>Systemen waarin kennis van experts is vastgelegd in databases, doorgaans in de vorm van regels met een of meer antecedenten en een consequent (alsdan redenering). Op basis van bekende en actuele feiten kan een rule engine met de expertregels conclusies voor een specifiek geval deduceren. Dergelijke systemen kunnen wel met incomplete en onzekere informatie omgaan, maar hebben geen lerend vermogen.</i>
Fuzzy Logic	<i>Een vorm van meerwaardige logica waarin aan een logische variabele een bepaalde mate van waarheid kan worden toegekend. Hierdoor kunnen systemen beter overweg met begrippen die door mensen contextueel intuïtief worden geïnterpreteerd, maar voor een computer niet exact genoeg zijn, zoals 'lang', 'ver', etc. Toepassing meestal in combinatie met regelgebaseerde systemen en in meet- en regeltechniek.</i>
Statistische methoden	<i>Een breed scala aan wiskundige methoden, zoals Bayesiaanse kansrekening, correlatie- en regressie-analyse, etc., die kunnen worden gebruikt om met waarschijnlijkheden te werken binnen regelgebaseerde systemen, of worden toegepast in lerende systemen om relaties en clusterings in representatieve gegevensverzamelingen te herkennen en dit toe te passen op nieuwe gegevens.</i>
Evolutionaire systemen	<i>Een evolutionair systeem maakt gebruik van selectie, mutatie en reproductie om computercode te verbeteren en aan te passen. Hoewel deze techniek niet noodzakelijk lerend vermogen van het evoluerende algoritme zelf vraagt, komt het evolutieproces wel tot oplossingen die vooraf niet door een mens zijn ontworpen. Deze techniek wordt onder meer gebruikt voor complexe optimalisatieproblemen, en in combinatie met regelgebaseerde systemen en neurale netwerken.</i>
Lerende systemen	<i>Het vermogen van een systeem om te leren van beschikbare gegevens, in contrast met systemen die voorgeprogrammeerd zijn voor een bepaalde taak. Lerende systemen kunnen onder meer gebruik maken van statistische methoden om patronen in representatieve gegevensverzamelingen te ontdekken, maar ook neurale netwerken zijn lerende systemen.</i>
Neurale netwerken	<i>Een netwerk van neuronen dat is gebaseerd op de wijze waarop hersenen werken. Een enkel neuron voert een relatief eenvoudige informatiebewerking uit, maar het geheel kan complexere taken uitvoeren omdat de neuronen onderling zijn verbonden. Deze verbindingen kunnen zich aanpassen, waardoor het netwerk lerende eigenschappen heeft. Neurale netwerken zijn onder meer geschikt voor patroonherkenning, zoals afbeeldingen en handschriften.</i>

Tabel 1 Globaal overzicht van Kunstmatige Intelligentie-technieken

Was de genetica tot enige tijd geleden vooral de studie naar erfelijke eigenschappen, nu wordt deze wetenschap steeds meer benaderd als de biologische manier om informatie te coderen waarmee entiteiten in de fysieke wereld kunnen worden gereproduceerd, variërend van enkele moleculen tot eenvoudige organismen en intelligente wezens als de mens. Van verschillende organismen, waaronder de mens, is het genoom ontrafeld. Hoewel de betekenis en werking van veel genetische data nog niet bekend zijn, is de techniek beschikbaar om DNA van levende organismen gericht te doorzoeken en te wijzigen.¹⁰ De biologische reproductiemechanismen vertonen nauwe verwantschap met nanotechnologie, het wetenschapsgebied dat zich bezighoudt met de manipulatie van materie op zeer kleine schaal, te weten kleiner dan 100 nanometer. Technieken om zeer dunne lagen materiaal ergens op

aan te brengen bestaan al ruim honderd jaar, maar vanaf begin jaren tachtig van de vorige eeuw bestaat het idee van machines op moleculair niveau, om daarmee op atomaire schaal computers te maken en biologisch materiaal te manipuleren.¹¹ Onderzoek en investeringen in

- 9 Jennifer Buss, Brian Barnett en Unnati Mehta, *Trends in Neurotechnology* (Potomac Institute for Policy Studies, 2015); James Giordano en Rachel Wurzman, 'Neurotechnology as weapons in national intelligence and defense—An overview', in: *Synesis: A Journal of Science, Technology, Ethics, and Policy* 2 (2011) (1) 55-71.
- 10 International Human Genome Sequencing Consortium e.a., 'Finishing the euchromatic sequence of the human genome', in: *Nature* 431 (2004) (7011) 931-945; Richard M. Durbin e.a., 'A map of human genome variation from population-scale sequencing', in: *Nature* 467 (2010) (7319) 1061-1073; Patrick D. Hsu, Eric S. Lander en Feng Zhang, 'Development and Applications of CRISPR-Cas9 for Genome Engineering', in: *Cell* 157 (2014) (6) 1262-1278.
- 11 K. Eric Drexler, 'Molecular engineering: An approach to the development of general capabilities for molecular manipulation', in: *Proceedings of the National Academy of Sciences of the USA* 78, (1981) (9) 5275-78.

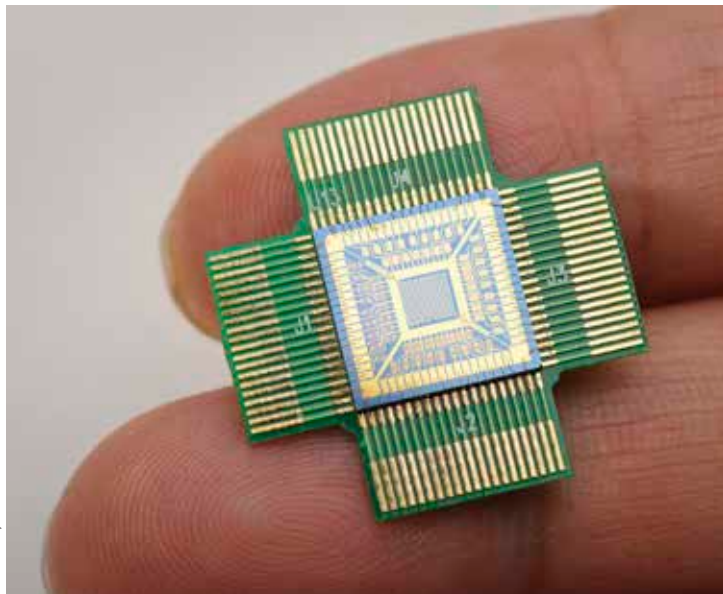


FOTO NASA, D. HART

Het is goed denkbaar dat een menselijk brein in de toekomst op een chip past, maar door ontwikkelingen in de nano- en neurotechnologie weten we niet of die er dan nog zo uit zal zien

nanotechnologie nemen toe en richten zich onder meer op energieopwekking en -opslag, fabricageprocessen, informatieverwerking, biomedische toepassingen en voedsel.¹² Nanotechnologie is daarmee de volgende evolutionaire stap in de menselijke bewerking van de fysieke wereld: wat begon als grove manuele materiaalbewerking en is uitgegroeid tot de zeer nauwkeurige machinale en chemische productieprocessen van nu, ontwikkelt zich verder tot programmeerbare materialen waaraan eigenschappen als zelfreplicatie en -reparatie kunnen worden gegeven.

Exponentiële trends

Hoewel er duidelijk ontwikkelingen plaatsvinden op het gebied van kunstmatige intelligentie, neurotechnologie, genetica en nanotechnologie, lijken die betrekkelijk langzaam te gaan. Aan de singulariteits-hypothese ligt echter ten grondslag dat de ontwikkeling van een evolutionair systeem exponentieel verloopt: bezien over een korte tijdspanne is het verloop bij benadering lineair. Over een relatief korte periode lijken technologische ontwikkelingen voor ons dus een vaste, gematigde tred te volgen.

Op macroniveau is het exponentiële verloop van zowel de kosmologische, biologische en technologische evoluties wel goed zichtbaar. Als we bijvoorbeeld kijken naar de tijdsduur van opeenvolgende significante gebeurtenissen in de kosmologische en biologische evolutie, dan zien we dat die steeds korter wordt. Tussen het ontstaan van de aarde en het ontstaan van dierlijk leven zat zo'n 4 miljard jaar, maar het duurde daarna minder dan 590 miljoen jaar tot de mens op het toneel verscheen.¹³ Om tot aan de neolithische revolutie¹⁴ te geraken, had de mens vervolgens ongeveer 190 duizend jaar nodig, terwijl het wetenschappelijk-industriële tijdperk daarna minder dan 13.000 jaar op zich liet wachten. Die laatste periode is nog geen 400 jaar geleden begonnen, maar de wereld is nooit zo snel zo ingrijpend veranderd.

De ontwikkeling in de informatietechnologie wordt vaak als typerend voorbeeld van de huidige technologische ontwikkelingen aangehaald, niet in de laatste plaats omdat het de motor achter de huidige vooruitgang is.¹⁵ In figuur 1 is het geheugen en de rekenkracht van computers (waaronder ook smartphones en tablets) op een logaritmische schaal uitgezet en is de exponentiële groei duidelijk zichtbaar.¹⁶

In dit figuur is de Wet van Moore terug te zien, die iedere 18-24 maanden een verdubbeling van de processorcapaciteit voorspelt, hoewel de oorspronkelijke uitspraak van Gordon

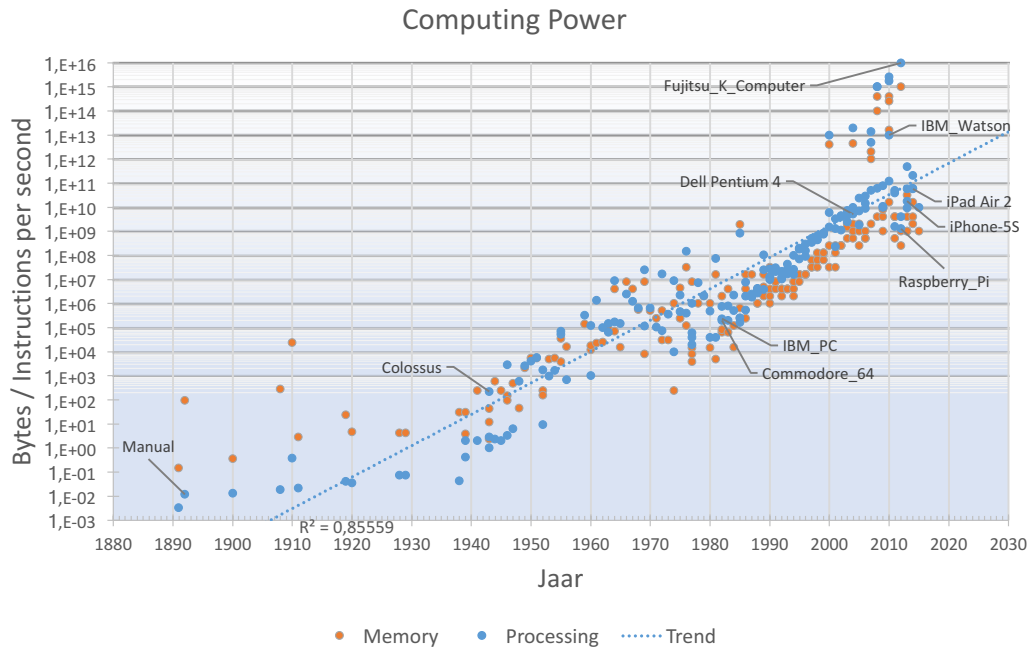
¹² K.E. Drexler e.a., *Productive Nanosystems: A Technology Roadmap* (Battelle Memorial Institute, 2007); European Commission en Directorate-General for Research and Innovation, *Nanotechnology: The Invisible Giant Tackling Europe's Future Challenges* (Luxembourg, Publications Office of the European Union, 2013); H. N. Cheng e.a., 'Nanotechnology Overview: Opportunities and Challenges', in: *ACS Symposium Series 1220* (2016) 1-12.

¹³ James G Ogg, Gabi Ogg en Felix M Gradstein, *A Concise Geologic Time Scale* (Elsevier, 2016).

¹⁴ De overgang van een samenleving van rondtrekkende jager-verzamelaars naar een samenleving van nederzettingen met landbouw en veeteelt.

¹⁵ De Ridder, *De ontdekking van de toekomst*, 23-27.

¹⁶ Hans Moravec, *Numerical data for computer power* (z.d.), zie: <http://www.frc.ri.cmu.edu/~hpm/book97/ch3/processor.list.txt>.



Figuur 1 Rekenkracht en geheugen van computers

Moore alleen betrekking had op de verdubbeling van het aantal transistors van een geïntegreerd circuit in twaalf maanden. Ray Kurzweil trekt de exponentiële trend nog breder en verklaart de exponentiële ontwikkelingen ten aanzien van alle vormen van evolutie met zijn *Law of Accelerating Returns*: een meer geavanceerde samenleving ontwikkelt zich sneller dan minder hoogstaande samenlevingen, juist omdat die geavanceerder is.¹⁷

Maar net zo min als de wetten van Moore en Kurzweil natuurwetenschappelijke wetten zijn, is de totale technologische ontwikkeling in werkelijkheid een exacte exponentiële curve. De ontwikkelingen verlopen als gevolg van diverse technische, maatschappelijke en culturele interacties veel grilliger. De groeitrends zijn niettemin reëel voor veel technologieën.¹⁸ Naast de kwantitatieve groei geeft tabel 2 met een korte selectie van wetenschappelijk nieuws over kleine, maar significante doorbraken van slechts één maand, al een kwalitatief beeld waar de huidige technologie staat met zaken die tot voor kort nog ondenkbaar waren.

Er bestaan verschillende modellen die technologische ontwikkelingen op mesoniveau inzichtelijk maken. Zo beschrijven Denning & Lewis de totale exponentiële groei binnen het *computing ecosystem* op de niveaus van *chips*, *systems* en *community* aan de hand van modellen van onder meer Koomey (energieverbruik) en Amdahl (parallelliseerbaarheid). Ook al zijn er deelgebieden waar de groei door fysische beperkingen stagneert, zij verwachten dat de exponentiële groei in de informatietechnologie voorlopig nog zal voortduren.¹⁹

Meer algemeen wordt de prestatiegroei van een specifieke technologie, zoals een biologisch groeiproces, vaak met een logistische functie weergegeven: een S-curve, ofwel prestatiecurve.²⁰ Ze laten zien dat de prestaties van een nieuwe technologie in het begin steeds sneller

¹⁷ Kurzweil, *The Singularity Is Near*.

¹⁸ Declan Butler, 'Tomorrow's World', in: *Nature* 530 (2016) (7591) 398-401.

¹⁹ Peter J. Denning en Ted G. Lewis, 'Exponential laws of computing growth', in: *Communications of the ACM* 60 (2017) (1) 54-65.

²⁰ Alexander Linden en Jackie Fenn, 'Understanding Gartner's hype cycles', in: *Strategic Analysis Report No R-20-1971* (Gartner, Inc., 2003).

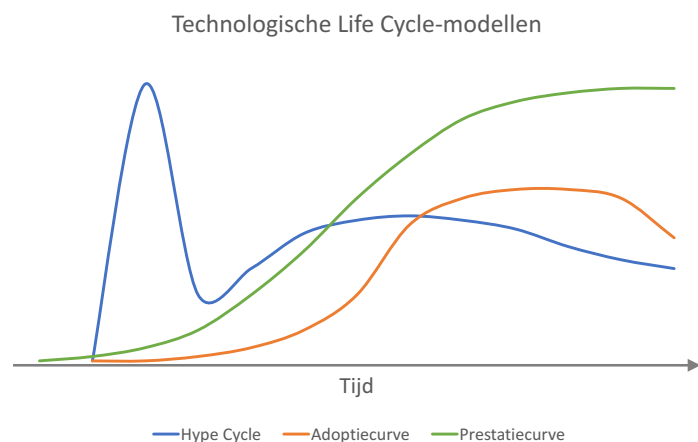
22 februari 2017: Onderzoekers van de Universiteit van Baskenland hebben biologische processen op kwantumschaal laten verlopen: in een kwantumomgeving hebben zij een evolutionair systeem met individuen, replicatie, mutatie, interactie en de dood gesimuleerd. De onderzoekers denken dat dit mechanisme zelf-replicerende geautomatiseerde processen op kwantumschaal mogelijk maakt.
http://www.ehu.es/en/en-content/-/asset_publisher/1575/content/n_20170222-investigacion-unai-almaz

1 maart 2017: Onderzoekers van de Universiteit van Manchester demonstreren dat het fysiek mogelijk is om een computer te maken gebaseerd op DNA. Deze computer is een nondeterministische universele Turingmachine die zichzelf kan repliceren, en daarmee door parallelisering sneller problemen kan oplossen dan conventionele computers en kwantumcomputers.
<http://www.manchester.ac.uk/discover/news/scientists-reveal-new-super-fast-form-of-computer-that-grows-as-it-computes/>

1 maart 2017: Microsoft Research en de University of Cambridge hebben een intelligent algoritme ontwikkeld dat in staat is om unieke softwarecode te schrijven om een gegeven eenvoudig (wiskundig) probleem op te lossen, waarbij het niet alleen leert van eigen resultaten, maar ook bestaande software als voorbeeld gebruikt.
<http://www.nextgov.com/emerging-tech/2017/03/microsofts-ai-learning-write-code-itself-not-steal-it/135793/>

6 maart 2017: MIT Computer Science and Artificial Intelligence Laboratory (CSAIL) en Boston University tonen met een werkend systeem aan dat het mogelijk is dat een mens met zijn gedachten het gedrag van een robot die een taak uitvoert, real-time kan corrigeren.
<http://www.techrepublic.com/article/how-mits-new-ai-system-lets-you-control-a-robot-with-your-mind/>

Tabel 2 Een aantal recente significante doorbraken binnen één maand



Figuur 2 Modellen voor de levenscyclus van een specifieke technologie

toenemen, maar dat de groei later steeds verder afvlakt als de technologie zijn volle potentie bereikt (zie figuur 2). Een S-curve kan ook worden gebruikt als cumulatieve weergave van de verspreiding van een innovatie over de tijd.²¹

Omdat betere technologieën de oude vervangen, is de totale (exponentiële) technologische vooruitgang in werkelijkheid een samenstel

van prestatiecurves die steeds korter en hoger zijn.²² Zo is de stijgende lijn in figuur 1 opgebouwd uit verschillende prestatiecurves, omdat de rekenkracht en het geheugen steeds op andere technologieën zijn gebaseerd.

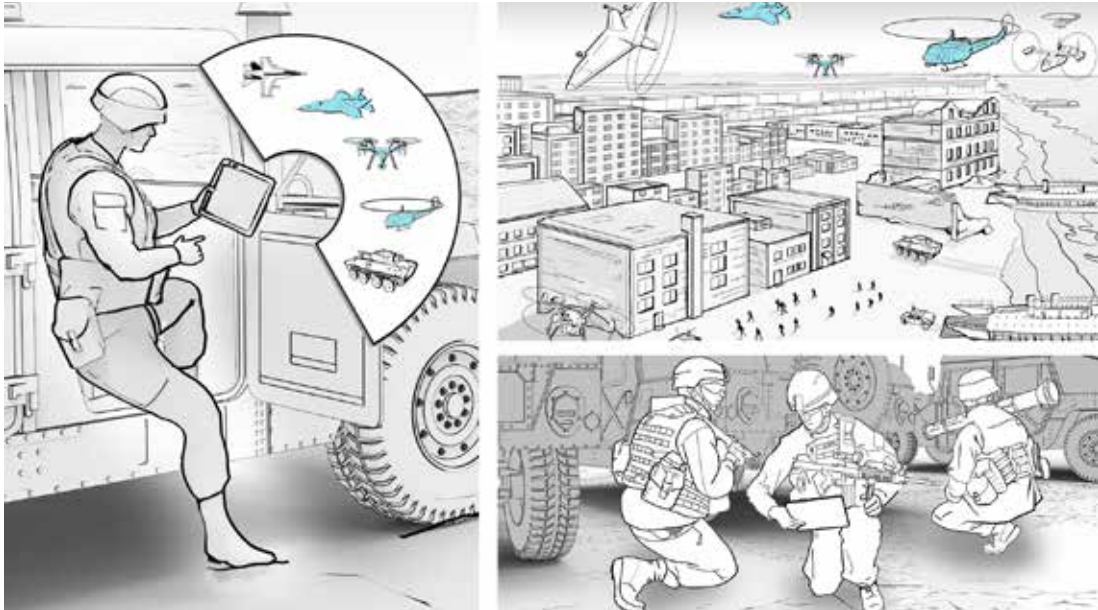
In figuur 2 is de mate van adoptie van een technologie weergegeven met een zogeheten adoptiecurve. Deze laat niet zozeer de ontwikkeling van een technologie zien, maar benutting ervan door de maatschappij: naarmate een technologie volwassen wordt, wordt deze breder toegepast (en andersom). Uiteindelijk zal de technologie worden vervangen door een andere, betere technologie en loopt het gebruik terug.

De technologische *hype cycle* van Gartner gaat een stap verder door de sociale attitude ten aanzien van een technologie weer te geven. Technologieën doorlopen alle dezelfde fasen van verwachting, maar voor iedere technologie verschilt de verwachting en de snelheid waarmee de fasen worden doorlopen.²³ De hype cycle laat in combinatie met de prestatiecurve zien dat de impact van een nieuwe technologie op de korte termijn vaak wordt overschat door te hoge verwachtingen, de gezeheten *Peak of Inflated Expectations*. De impact op de lange termijn wordt echter

²¹ Everett M. Rogers, *Diffusion of innovations*, 4th ed. (New York, Free Press, 1995).

²² Kurzweil, *The Singularity Is Near*.

²³ Linden en Fenn, 'Understanding Gartner's hype cycles'.



ILLUSTRATIE DARPA

Op allerlei vlakken ervaren we, ook in de militaire omgeving, dagelijks de toegenomen technische communicatiemogelijkheden en de directe toegang tot meer informatie

onderschat: de verwachtingen dalen eerst sterk en gaan door een *Trough of Disillusionment*, terwijl de volwassenheid en adoptie van de technologie wel degelijk groeit.

Symptomen van de singulariteit

Vinge (1993) geeft aan dat het niet zeker is of de technologische singulariteit eigenlijk wel mogelijk is en er bestaan wezenlijke argumenten die het fenomeen fundamenteel als onmogelijk beschrijven. Maar als het mogelijk is, zal het volgens Vinge zeker plaatsvinden. Hij noemt een viertal symptomen die daarbij mogelijk optreden.

Het eerste symptoom is dat automatisering steeds hoger geschoolde banen vervangt. Dit is een waarneembare trend die wordt beschreven in diverse recente rapporten, waaruit blijkt dat in de komende decennia de helft van de banen in westerse economieën zou kunnen verdwijnen.²⁴ Hoewel het verdwijnen en opkomen van beroepen vaker is voorgekomen in het verleden, neemt dit momenteel fundamenteelere vormen aan vanwege het allesomvattende karakter van informatietechnologie: het is niet

meer vanzelfsprekend dat verouderde beroepen door nieuwe worden vervangen.²⁵

Het tweede symptoom zou zijn dat ideeën zich steeds sneller verspreiden. Doordat mensen ideeën en gewoonten overdragen op andere personen en generaties, kunnen ze zich handelwijzen eigen maken zonder alle gevaren en risico's van individueel experimenteren en kunnen ze daarmee grote geïntegreerde gedragspatronen en collectieve intelligentie verwerven.²⁶ Iedereen ervaart dagelijks de toegenomen technische mogelijkheden voor onderlinge communicatie en de directe toegang tot meer informatie. Maar leidt dit ook tot een meetbare toegenomen ideeënstroom en tot grotere collectieve intelligentie? We weten dat het aantal wetenschappelijke publicaties en

24 Carl Benedikt Frey en Michael Osborne, 'Technology at Work: The Future Is Not What It Used to Be' (Oxford Martin School, University of Oxford, 2016); Devdatt Dubhashi en Shalom Lappin, 'AI Dangers: Imagined and Real', in: *Communications of the ACM* 60 (2017) (2) 43-45.

25 Erik Brynjolfsson en Andrew McAfee, *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies* (New York, W.W. Norton & Company, 2015).

26 Alex Pentland, *Sociale big data: opkomst van de data-gedreven samenleving* (Amsterdam, Maven Publishing, 2014).

referenties toeneemt, dat het wetenschappelijk werk meer en meer bestaat uit wereldwijde samenwerking²⁷ en dat communicatie- en informatietechnologie daar een grote rol in speelt.²⁸ Maar uitsluitend op basis daarvan kan niet direct iets worden gezegd over de snelheid van verspreiding van nieuwe ideeën. Het aantal nieuwe patenten groeit bijvoorbeeld ook, maar daaruit kan evenmin rechtstreeks worden geconcludeerd dat het aantal nieuwe ideeën toeneemt, onder meer vanwege de grote invloed van marktdynamiek en het juridische klimaat.²⁹ Pas met recente ontwikkelingen op het gebied van big data en sociale netwerken is beter inzichtelijk te maken hoe ideeënstromen zich precies gedragen en waardoor ze worden beïnvloed.³⁰ Zo is collectieve intelligentie gevoelig voor redundante informatie van verschillende, maar onderling afhankelijke bronnen, wat specifieke trends versterkt en het effect van een snellere verspreiding van vernieuwende ideeën teniet kan doen. Daarnaast zijn er ook andere dan technologische factoren, waaronder demografische ontwikkelingen als bevolkingsgroei en verstedelijking, waaraan een snellere verspreiding van ideeën

kan worden toegeschreven. Er is dus meer onderzoek nodig naar dit symptoom om het beter te definiëren en te kwantificeren.

Een andere dimensie is een toename in synergetische en niet-menselijke kennis, het derde symptoom. De totale hoeveelheid digitale gegevens groeit exponentieel door verdergaande digitalisering en zal naar verwachting in 2020 de 44 zettabytes, ofwel $44 \cdot 10^{21}$ bytes, passeren.³¹ Zowel bedrijven als wetenschappelijke instellingen gebruiken steeds vaker kennissystemen en big data-technieken om op basis van groeiende hoeveelheden beschikbare gegevens tot nieuwe kennis en inzichten te komen.³² Er is hier primair sprake van een toename van synergetische kennis die alleen kan bestaan door samenwerking van mens en machine. Maar er komen ook steeds meer vormen van niet-menselijke kennis. Zoekmachines en andere applicaties zijn steeds beter in staat termen in context te interpreteren en concrete vragen in menselijke taal te beantwoorden.³³ Getrainde neurale netwerken kunnen vooraf onbekende patronen interpreteren en bevatten dus niet-menselijke kennis waarvan wordt onderzocht hoe die expliciet kan worden gemaakt, zodat mensen kunnen begrijpen hoe machines tot bepaalde uitkomsten komen.³⁴ Naarmate meer neurale netwerken worden gebruikt in toepassingen, zal de niet-menselijke kennis dus groeien.

Het vierde en laatste symptoom is dat machines slagen voor de Turingtest. De test komt er in de kern op neer dat als een mens in een vraaggesprek het onderscheid tussen een machine en een ander mens niet weet te maken, we praktisch gezien moeten concluderen dat de machine het cognitieve vermogen van een mens heeft. Deze *imitation game* is de praktische oplossing om tot een antwoord te komen op de vraag of machines kunnen denken, zonder in filosofische discussies over wat 'denken' is te geraken.³⁵ De test lijkt in zijn opzet eenvoudig, maar de wijze waarop het vraaggesprek moet worden gevoerd en geïnterpreteerd is onderwerp van veel wetenschappelijk debat,³⁶ niet in de laatste plaats vanwege de onvermijdelijke menselijke vooroordelen bij

-
- 27 Lutz Bornmann en Rüdiger Mutz, 'Growth rates of modern science: A bibliometric analysis based on the number of publications and cited references', in: *Journal of the Association for Information Science and Technology* 66, (2015) (11) 2215-2222.
- 28 B. W. Hesse, R. P. Moser en W. T. Riley, 'From Big Data to Knowledge in the Social Sciences', in: *The ANNALS of the American Academy of Political and Social Science* 659 (2015) (1) 16-32.
- 29 Michele Boldrin en David K Levine, 'The Case Against Patents', in: *Journal of Economic Perspectives* 27 (2013) (1) 3-22; Stephen Roper en Nola Hewitt-Dundas, 'Knowledge Stocks, Knowledge Flows and Innovation: Evidence from Matched Patents and Innovation Panel Data', in: *Research Policy* 44 (2015) (7) 1327-1340.
- 30 Pentland, *Sociale big data*, 35 e.v.
- 31 Butler, 'Tomorrow's World'.
- 32 Hesse, Moser en Riley, 'From Big Data to Knowledge in the Social Sciences'; Helen N. Rothberg en G. Scott Erickson, 'Big Data Systems: Knowledge Transfer or Intelligence Insights?', in: *Journal of Knowledge Management* 21 (2017) (1) 92-112.
- 33 Vivienne Waller, 'Making Knowledge Machine-Processable: Some Implications of General Semantic Search', in: *Behaviour & Information Technology* 35 (2016) (10) 784-795.
- 34 Yunpeng Li e.a., 'A New Method of Mapping Relations from Data Based on Artificial Neural Network', in: *International Journal of System Assurance Engineering and Management* 5 (2014) (4) 544-553.
- 35 Alan M. Turing, 'Computing machinery and intelligence', in: *Mind* 59 (1950) (236) 433-460.
- 36 Jia You, 'Beyond the Turing Test', in: *Science* 347 (2015) (6218) 116-116.
- 37 Kevin Warwick, Huma Shah en James Moor, 'Some Implications of a Sample of Practical Turing Tests', in: *Minds and Machines* 23 (2013) (2) 163-177.



FOTO: ANP

De Turingtest is genoemd naar Alan Turing, de Britse wiskundige die in de Tweede Wereldoorlog baanbrekend werk op het gebied van decodering verrichtte

degenen die het onderscheid tussen mens en machine moeten zien te maken.³⁷ Tot op heden heeft nog geen machine de test gehaald. Ray Kurzweil verwacht dat machines tegen 2030 voor de Turingtest zullen slagen.³⁸

Argumenten tegen een singulariteit

Er zijn onderzoekers die denken dat vooruitgang altijd gepaard gaat met ontwrichting en herijking. Het is dus moeilijk te bepalen of symptomen die aan de technologische singulariteit kunnen worden toegeschreven, zoals die van Vinge, eigenlijk geen manifestaties zijn van andere fenomenen, zoals de crises beschreven in het Schumpeter-Freeman-Perez-paradigma. Dat model gaat uit van een periodieke verandering van het techno-economisch paradigma met bijhorende culturele, geopolitieke en systeemcrises als gevolg van technologische vooruitgang.³⁹ Het model voorspelt geen convergentie uitmondend in een singulariteit; anderzijds sluit het model dit ook niet uit.

Andere wetenschappers beschouwen een technologische singulariteit als principieel onmogelijk. De meest bekende argumenten zijn die van John Searle, John Lucas en Roger Penrose. De argumenten van Searle zijn gebaseerd op het uitgangspunt dat machines alleen syntactische bewerkingen kunnen uitvoeren, wat een onvoldoende voorwaarde is voor semantiek en zelfbewustzijn. Met andere woorden: computers kunnen alleen symbolen manipuleren (zoals getallen en woorden) zonder die te kunnen begrijpen.⁴⁰ Hierdoor zou sterke AI niet mogelijk zijn.

De argumenten van Lucas en Penrose volgen dezelfde lijn, maar gaan een stap verder. Hun argumenten zijn niet helemaal dezelfde, maar ze baseren zich beiden op het onvolledigheidsprincipe van Gödel. Dat bewijst dat een consistent wiskundig systeem altijd onvolledig is, omdat er waarheden binnen dat systeem bestaan die niet door dat systeem zelf kunnen worden bewezen.⁴¹ Omdat machines of algoritmes feitelijk praktische implementaties van een wiskundig systeem zijn, zijn er dus waarheden die consistent zijn met dat systeem terwijl het systeem dat niet kan bewijzen. Maar een mens heeft wel het vermogen zulke waarheden te onderscheiden.⁴² Als gevolg van Gödels onvolledigheidsprincipe zijn algoritmen dus te beperkt om echt intelligent te kunnen zijn, zoals het menselijk brein dit als niet-algorithmische entiteit wel kan zijn. Of anders geformuleerd: echte intelligentie is niet verenigbaar met een volledig computergestuurd systeem.⁴³

³⁸ Kurzweil, *The Singularity Is Near*, o.m. 25 en 200.

³⁹ De Ridder, *De ontdekking van de toekomst*, 33 e.v. Christopher Freeman en Carlota Perez, 'Structural Crises of Adjustment, Business Cycles and Investment', in: *Technical change and economic theory*, IFIAS research series, no. 6 (London-New York, Pinter Publishers, 1988) 38-66.

⁴⁰ John R. Searle, 'Minds, Brains, and Programs', in: *Behavioral and Brain Sciences* 3 (1980) (3) 417; John R. Searle, 'Is the Brain's Mind a Computer Program?', in: *Scientific American* 262 (1990) (1) 26-31.

⁴¹ Ernest Nagel, James Roy Newman en Douglas R. Hofstadter, *Gödel's Proof* (New York, New York University Press, 2001).

⁴² John R. Lucas, 'Minds, machines and Gödel', in: *Philosophy* 36 (1961) (137) 112-127.

⁴³ Roger Penrose, 'On Understanding Understanding', in: *International Studies in the Philosophy of Science* 11 (1997) (1) 7-20.

Bovenstaande standpunten zijn door de jaren heen, zowel in de opzet van de redentie als inhoudelijk, veelvuldig bekritiseerd. Een belangrijk tegenargument is dat kunstmatige intelligentie ook niet-algoritmisch van aard zou kunnen zijn, zodat Gödels onvolledigheidsprincipe niet relevant is. Door het introduceren van enige mate van toeval in een kunstmatig systeem, bijvoorbeeld met een *random number generator* of met de natuurlijke toevalligheden die in de wereld waarneembaar zijn, kunnen niet-algoritmische oplossingen worden gevonden.⁴⁴ We weten al langer dat toeval een fundamentele rol speelt in de natuurkunde, meer specifiek in de kwantummechanica, maar toeval speelt die rol ook in de grondslagen van de wiskunde⁴⁵ en in het functioneren van het brein.⁴⁶ Tevens wordt er vooruitgang geboekt met de toepassing van toeval in *deep learning*.⁴⁷

Tot slot bestaat er een aantal argumenten tegen het extrapoleren van technologische groei, bijvoorbeeld vanwege fundamentele fysische limieten. Onder deze klasse vallen onder meer de argumenten dat grotere rekenkracht niet noodzakelijk leidt tot geavanceerdere algoritmen, dat intelligente algoritmen niet noodzakelijk zo intelligent zijn dat ze zichzelf kunnen verbeteren en dat toenemende intelligentie wegvalt tegen toenemende complexiteit. Deze argumenten sluiten sterke AI in de nabije toekomst niet uit, maar die zou dan niet leiden tot een singulariteit.⁴⁸

Impact, risico's en militaire relevantie

Vanuit wetenschappelijk oogpunt is de mogelijkheid van een technologische singulariteit dus geen uitgemaakte zaak. Hoewel het een interessant intellectueel en filosofisch debat oplevert,



FOTOCW.A. OOMEN

De invoering van AI zal de verdere automatisering van banen, economische

biedt deze discussie op dit moment weinig praktische houvast. De vraag is dan ook of het in de nabije toekomst verschil maakt of in de wereld ooit een technologische singulariteit zal ontstaan. De snelle technologische ontwikkelingen die zich gelijktijdig op veel gebieden voordoen, brengen immers nu al complexe vraagstukken en risico's met zich mee.

Zo stellen Dubhashi en Lappin dat de ontwikkelingen op het gebied van kunstmatige intelligentie niet zo snel gaan dat sterke AI binnen afzienbare tijd is te verwachten, maar zij zijn wel van mening dat alleen al de impact van zwakke AI dermate groot kan zijn dat dit mogelijk leidt tot een onredelijke verdeling van welvaart door de verdergaande automatisering van banen, terwijl overheden hierop nauwelijks zijn voorbereid.⁴⁹ De opkomst van kunstmatige intelligentie en robotica zal onvermijdelijk gepaard gaan met economische herijking en bijbehorende crises. Tot op heden verdienen

⁴⁴ Taner Edis, 'How Gödel's Theorem Supports the Possibility of Machine Intelligence', in: *Minds and Machines* 8 (1998) (2) 251-262.

⁴⁵ Gregory J. Chaitin, *Meta Math! The Quest for Omega* (New York, Vintage Books, 2006).

⁴⁶ Kun Xie e.a., 'Brain Computation Is Organized via Power-of-Two-Based Permutation Logic', in: *Frontiers in Systems Neuroscience* 10 (2016).

⁴⁷ Chris Edwards, 'Growing Pains for Deep Learning', in: *Communications of the ACM* 58 (2015) (7) 14-16.

⁴⁸ Toby Walsh, 'The Singularity May Never Be Near', in: *Computing Research Repository* (2016) zie: <https://arxiv.org/abs/1602.06462>.

⁴⁹ Dubhashi en Lappin, 'AI Dangers'.



herijking en crises betekenen, wat mondiaal tot frictie kan leiden

mensen geld omdat arbeid schaars is. Maar als bepaalde arbeid steeds minder schaars wordt, rijst de vraag of er in de toekomst voldoende werk is voor iedereen. In bredere zin is schaarste in de meeste gangbare economische modellen een uitgangspunt (*Scarcity Economics*), terwijl technologische vooruitgang deze schaarste steeds meer opheft. Toekomstige economische problemen vragen derhalve om andere economische modellen (*Abundance Economics*). In dergelijke modellen wordt toegang tot informatie en kennis belangrijker dan materieel en financieel bezit. Niet alleen financiële aspecten en welvaart spelen daarbij een grote rol, maar ook netwerkrelaties, gezamenlijke risico's en welzijn.⁵⁰

Naast de economische aspecten is ook de sociaal-culturele impact groot. Voor specifieke gevallen is bijvoorbeeld al aangetoond dat collectieve intelligentie minder foutgevoelig is en zich makkelijker aanpast aan verandering

dan de huidige methoden van besluitvorming in organisaties.⁵¹ Digitale samenwerking zal daardoor op termijn beter kunnen functioneren dan huidige besluitvormende entiteiten zoals managers, raden van bestuur, medezeggenschapsraden, commissies, et cetera. In combinatie met de opkomst van big data en kunstmatige intelligentie kan dit een grote impact hebben op het functioneren van organisaties, economieën en staten, zowel in positieve als in negatieve zin.⁵² Aan de ene kant worden traditionele besluitvormers en machthebbers als gevolg van betere informatisering gedwongen transparanter te zijn en verantwoording af te leggen. Anderzijds beïnvloeden bedrijven en staten het gedrag van individuele mensen met behulp van informatietechnologie, waarbij big data en kunstmatige intelligentie een centrale rol spelen. Het analyseren en al dan niet onbewust beïnvloeden van gedrag van individuen blijkt commercieel succesvol.⁵³ En ook statelijke actoren bedienen zich ervan.⁵⁴ Het meest vergaande voorbeeld is wellicht het *social credit-systeem* in China, waarmee de overheid een score wil toekennen aan het individuele gedrag van personen. Naarmate een computersysteem van de Chinese overheid een persoon meer of minder 'betrouwbaar' acht, worden privileges (zoals hoogte van het pensioen, toestemming om te reizen, enzovoort) toegekend of onzegd.⁵⁵

50 David H. Autor, 'Why Are There Still So Many Jobs? The History and Future of Workplace Automation', in: *Journal of Economic Perspectives* 29 (2015) (3) 3-30; Kevin LaGrandeur en James J. Hughes, *Surviving the Machine Age: Intelligent Technology and the Transformation of Human Work*, 1st edition (New York, Springer, 2017); De Ridder, *De ontdekking van de toekomst*, 161 e.v.

51 Derek A. Epp, 'Public Policy and the Wisdom of Crowds', in: *Cognitive Systems Research* 43 (2017) 53-61.

52 Dirk Helbing e.a., 'Will Democracy Survive Big Data and Artificial Intelligence?', *Scientific American* (2017) zie: <https://www.scientificamerican.com/article/will-democracy-survive-big-data-and-artificial-intelligence>.

53 Jessica Lichy, Maher Kachour en Tatiana Khvatova, 'Big Data Is Watching YOU: Opportunities and Challenges from the Perspective of Young Adult Consumers in Russia', in: *Journal of Marketing Management* (2017) 1-23.

54 David Lyon, 'Big Data Surveillance: Snowden, Everyday Practices and Digital Futures', in: *Routledge Studies in International Political Sociology: International Political Sociology: Transversal Lines* (Londen, Taylor and Francis, 2016).

55 'China invents the digital totalitarian state: The worrying implications of its social-credit project', *The Economist* (2016) zie: <http://www.economist.com/news/briefing/21711902-worrying-implications-its-social-credit-project-china-invents-digital-totalitarian>.

De ontwikkelingen leiden op dit moment al tot fundamentele ethische en juridische vraagstukken. Bostrom en Yudkowsky benaderen de ethische aspecten van kunstmatige intelligentie vanuit vier invalshoeken.⁵⁶

- de minimale set van criteria die al zou moeten gelden voor zwakke AI, ofwel voor alle algoritmes die vandaag de dag menselijke besluitvorming vervangen: *responsibility*, *transparency*, *auditability*, *incorruptibility* en *predictability*;
- het ethisch bewust maken van sterke AI, omdat sterke AI zelf beslissingen neemt om een bepaald doel op relatief lange termijn te bereiken, waarbij niet alleen het einddoel, maar ook de vooraf niet te voorspellen handelingen naar het doel redelijk en acceptabel moeten zijn;
- de morele status van AI zelf: hoe willen wij AI behandelen als het gevoel, zelfbewustzijn of persoonlijkheid heeft?
- de superethische vraagstukken van bovenmenselijke intelligentie, waarbij de menselijke ethiek misschien niet meer de enige ethische maatstaf is.

Ethische vraagstukken over kunstmatige intelligentie zijn momenteel onderwerp van veel debat.⁵⁷ Vergelijkbare en deels overlappende ethische vraagstukken zijn ook te formuleren voor kunstmatig leven en voor combinaties van biologische en kunstmatige entiteiten. Dit roept nu al de vraag op of de huidige juridische stelsels voldoende zijn

toegerust voor het hedendaagse technologische landschap. Wet- en regelgeving richt zich van oudsher primair op individueel menselijk gedrag, maar biedt momenteel nog weinig houvast voor aansprakelijkheid van gedrag van complexe samengestelde systemen, laat staan van intelligente autonome systemen.⁵⁸

Het behoeft weinig toelichting dat de genoemde ontwikkelingen onvermijdelijk militaire impact hebben, en wel op twee manieren. Ten eerste hebben de geschetste veranderingen invloed op de mondiale veiligheidssituatie en op de verwachtingen die de eigen samenleving van het militaire apparaat heeft. De economische herijking en veranderende machtsbalans als gevolg van breed beschikbare krachtige technologieën kan gepaard gaan met onrust, crises en conflicten. Daarbij zal de aard van (militaire) conflicten blijven veranderen. Defensie zal deze ontwikkelingen goed moeten begrijpen om daarop te kunnen anticiperen.

Ten tweede bieden de technologische ontwikkelingen mogelijkheden voor militaire toepassing: voor de eigen krijgsmacht en bondgenoten, voor potentiële tegenstanders en voor andere partijen die bij crises en conflicten betrokken zijn. De gangbare militaire concepten van de afgelopen twintig jaar leren dat militaire capaciteiten in staat moeten zijn te opereren in netwerken die een hoge graad van decentralisatie en zelfsynchronisatie kennen en daardoor zeer adaptief zijn.⁵⁹ Daar komt nu meer en meer het combineren en integreren van mensen en intelligente machines bij, evenals het gebruik van autonome systemen. Deze twee ontwikkelingen vergroten de snelheid en precisie van opereren en worden gezien als de belangrijkste aspecten van militair vermogen in de 21e eeuw.⁶⁰ Defensie zal zich deze opkomende technologieën niet alleen eigen moeten maken voor defensieve en offensieve doeleinden, maar ook tijdig de inherente ethische en juridische vraagstukken moeten beantwoorden om deze technologieën effectief voor militaire doeleinden te kunnen en mogen inzetten.

⁵⁶ Nick Bostrom en Eliezer Yudkowsky, 'The ethics of artificial intelligence', in: *The Cambridge Handbook of Artificial Intelligence* (Cambridge University Press, 2014) 316-334.

⁵⁷ 'Ethically Aligned Design: A Vision for Prioritizing Human Wellbeing with Artificial Intelligence and Autonomous Systems' (IEEE, 13 december 2016); Ilse Verdiesen, 'How do we ensure that we remain in control of our autonomous weapons? A value-sensitive design approach', in: *Arte Pugnantis Adsum* 54 (2017) (107) 14-20.

⁵⁸ Daniel Zeng, 'AI Ethics: Science Fiction Meets Technological Reality', in: *Intelligent Systems* 30 (2015) (3) 25.

⁵⁹ David S. Alberts, John Garstka en Frederick P. Stein, *Network centric warfare: developing and leveraging information superiority* (Washington, D.C., National Defense University Press, 1999); Hans van Dalen, Patrick Dekkers en Arjan van Daalen, 'NetForce: een nieuw model voor toekomstige Defensie', in: *Militaire Spectator* 186 (2017) (3) 125-40.

⁶⁰ Keith Kirkpatrick, 'Can We Trust Autonomous Weapons?', in: *Communications of the ACM* 59 (2016) (12) 27-29.

	<i>Commando-voering</i>	<i>Inlichtingen</i>	<i>Manoeuvre</i>	<i>Beïnvloeding</i>	<i>Zelfbescherming</i>	<i>Logistiek (personeel & materieel)</i>
<i>Kunstmatige intelligentie</i>	<i>Betere en snellere besluitvorming Situational awareness & understanding Volledige wereldrepresentatie, zowel in het fysieke, informatie- als mentale domein</i>	<i>Beeld, tekst-, spraakanalyse Big data analyse Analyseren en voorspellen vijandelijk gedrag, zowel in het fysieke, informatie- als mentale domein</i>	<i>Optimalisatie van operationele deployments Optimalisatie van tactische bewegingen Snelle anticipatie op veranderingen van tegenstanders, weer en terrein</i>	<i>Voorspellen van kinetische en niet-kinetische effecten Intelligente cyberwapens</i>	<i>Early warning Autonome respons Intelligente authenticatie & access control</i>	<i>Anticipatie en optimalisatie in materieel- en personeelsketen</i>
<i>Robotica</i>	<i>Humanoïde assistent Representatie op afstand</i>	<i>Slimme en/of autonome sensor-systemen</i>	<i>Slimme en/of (semi) autonome voer-, vaar-, en vliegtuigen Swarming Toegang tot gevaarlijke en onbereikbare plaatsen Exoskeletons</i>	<i>Slimme en/of (semi) autonome wapens</i>	<i>Autonome beschermingsrobots</i>	<i>Autonome voer-, vaar-, en vliegtuigen voor troepen-transport, bevoorrading, evacuatie, etc. Autonome distributiecentra</i>
<i>Nanotechnologie</i>	<i>Zeer kleine en/of geïmplanteerde computers voor SA & besluitvorming</i>	<i>Nanosensoren (smart dust) Detectie van zeer kleine sensoren en nanorobots</i>	<i>Toepassing van zeer lichte maar sterke materialen Onzichtbaar bereiken van tegenstander, (zeer kleine systemen, elektro-optische camouflage)</i>	<i>Kinetische en niet-kinetische effecten door nanorobots Nanoweapons</i>	<i>Beschermende materialen Anticiperend materiaal Bescherming tegen nanorobots</i>	<i>Zelfreparerend materieel Lokale, decentrale productie van materieel en voedsel Energie-opwekking</i>
<i>Genetica</i>	<i>Verbeteren cognitief vermogen en alertheid</i>	<i>Verbeteren zintuigen en geheugen Genoom-analyse van tegenstanders</i>	<i>Verbeteren lichamelijk uithoudingsvermogen</i>	<i>Directe of indirecte beïnvloeding genoom van tegenstanders</i>	<i>Verbeteren lichamelijke weerstand</i>	<i>Gepersonaliseerd voedsel en gezondheidszorg voor militairen</i>
<i>Neurotechnologie</i>	<i>Brain-to-brain en brain-to-machine communicatie Verbeteren cognitief vermogen en alertheid</i>	<i>Lezen van gedachten Verbeteren zintuigen en geheugen</i>	<i>Gedachtenbesturing van voer-, vaar-, en vliegtuigen</i>	<i>Directe beïnvloeding gedachten en gedrag van tegenstanders</i>	<i>Bescherming tegen lezen en sturing van gedachten Health monitoring</i>	<i>Medische diagnose en behandeling bij militairen</i>

Tabel 3 Mogelijke toepassingen van technologieën voor verschillende functies van militair optreden

De singulariteitshypothese is hierbij in zoverre bruikbaar dat mens-machine-integratie en autonome systemen nu juist voor een belangrijk deel zijn gebaseerd op de technologiegebieden die aan de hypothese ten grondslag liggen: kunstmatige intelligentie, robotica,

nanotechnologie, genetica en neurotechnologie. Door deze technologieën uit te zetten tegen de hoofdfuncties van militair optreden krijgen we een indruk van mogelijke militaire toepassingen. Dit is weergegeven in tabel 3.



U.S. MARINE CORPS, F. CORDOBA

Test met het robotic vehicle modular system van het U.S. Marine Corps Warfighting Laboratory: de singulariteitshypothese kan militaire denkers helpen de focus te leggen op technologische ontwikkelingen op het gebied van kunstmatige intelligentie

Dit overzicht is geenszins volledig, onder meer omdat de tabel niet op systematisch onderzoek is gebaseerd, maar hoofdzakelijk op gedachte-wisselingen in werksessies en beleidsdiscussies tijdens de afgelopen twee jaar. Daarnaast zijn er gerelateerde technologieën die niet expliciet worden genoemd, maar die wel een sterke relatie hebben met de singulariteitshypothese en militaire toepassingen. Dat zijn vooral ontwikkelingen in informatietechnologie die voor vergaande digitalisatie en integratie van de verschillende technologiegebieden zorgen, zoals big data, virtual & augmented reality en quantum computing. Tot slot is het van belang te kijken naar de toepassing van deze technologieën in het bredere veiligheidsecosysteem, zoals het The Hague Centre for Strategic Studies dat met een recent rapport over kunstmatige intelligentie heeft gedaan.⁶¹ De tabel laat niettemin zien dat de singulariteitshypothese ook relevant is vanuit militair perspectief en voldoende concrete aanknopingspunten biedt voor verder onderzoek en ontwikkeling.

Conclusie

De hypothese van de technologische singulariteit is nauwelijks toetsbaar, maar wel bruik-

baar. Enerzijds is de hypothese nuttig als denkkader bij de toenemende complexiteit en verwevenheid van verschillende wetenschaps- en technologiegebieden. Anderzijds leidt de hypothese tot een focus op de technologische ontwikkelingen met de meeste relevantie en impact op de lange termijn en de gevolgen daarvan.

De singulariteitshypothese kan Defensie helpen om tijdig in te spelen op technologische ontwikkelingen op het gebied van kunstmatige intelligentie, robotica, nanotechnologie, genetica en neurotechnologie, zowel voor de eigen militaire capaciteiten als om in te schatten hoe potentiële tegenstanders en andere partijen deze technologieën zullen benutten. Het zijn namelijk deze technologiegebieden die de komende decennia voor een belangrijk deel de aard en effectiviteit van militair vermogen zullen bepalen.

Het gaat daarbij niet alleen om toepassing van deze technologieën voor militaire doeleinden, maar ook om de sociale, ethische en juridische aspecten die er niet los van kunnen worden gezien. En net als bij andere complexe systemen, zoals vliegtuigen en informatietechnologie, is de beheersing van veiligheidsaspecten (*safety & security*) daar een essentieel onderdeel van.

Met de uitkomsten van dit vooronderzoek kan nader worden bepaald welk beleid moet worden gevoerd om voldoende rekening te houden met ontwikkelingen die, hypothetisch, kunnen leiden tot een technologische singulariteit. Daarbij staat centraal of Defensie in staat is om mogelijke militaire toepassingen van kunstmatige intelligentie, robotica, nanotechnologie, genetica en neurotechnologie te voorzien en daar voldoende snel op in te spelen. Want zoals de krijgsmacht zich tijdens inzet steeds moet kunnen aanpassen aan de dynamiek van een conflict, zal Defensie ook als geheel wendbaarder moeten worden om de technologische ontwikkelingen te volgen en relevant te blijven in een snel veranderende wereld. ■

⁶¹ Stephan De Spiegeleire, Matthijs Maas en Tim Sweijs, *Artificial Intelligence and the Future of Defense: Strategic Implications for Small- and Medium-Sized Force Providers* (Den Haag, The Hague Centre for Strategic Studies, 2017).