

**MENS
VERSUS
MACHINE**

Leesexemplaar



Geertrui Mieke De Ketelaere

**MENS
VERSUS
MACHINE**

**Artificiële intelligentie
ontrafeld**

P E L C K M A N S

INHOUD

INTRODUCTIE: EEN HANDLEIDING VOOR AI-VERTALERS 9

1	ALGORITMES	30
	Worden computers slimmer dan mensen?	31
	Algoritmes	32
	Neurale netwerken	35
	Leren in drie smaken	37
	Topmodel gezocht	42
	Zelfrijdende crash	46
	Moravec's paradox	49
	Mens versus machine	52
	Algorithmic bias	57
	Black box	60
	Zet 78	61
	INTERVIEW MET VÉRONIQUE VAN VLASSELAER	63
	INTERVIEW MET LORE SIMONS	65

2	DATA	67
	De olie van slimme computers	68
	De vijf dimensies	71
	Gezondheid	76
	Clicks en personal things	79
	GDPR	83
	Solid	84
	INTERVIEW MET RUBEN VERBORGH	87
	INTERVIEW MET FREDERIC PIVETTA	89

3	HARDWARE	92
	Wie controleert AI?	93
	AI wordt tastbaar	96
	Madonna versus Basic	96
	Basisstructuur	101
	AI-versnellers	103
	In de wolken	105
	Oplossingen aan de rand	108
	Techgiganten	111
	Dot-com bubble	114
	Halen we hen in?	115
	INTERVIEW MET PETER DEBACKER	117
	INTERVIEW MET TIJS WILBRINK	119

4	BESLISSEN	121
	Steelt AI onze jobs?	122
	Hokjesdenken	124
	Beeldherkenning	127
	Spraakherkenning	129
	Misleidende AI	132
	Context matters	134
	Failliet	137
	AI bij de corporates	138
	Computer versus mens?	141
	INTERVIEW MET JOSEPH FATTOUCH	145
	INTERVIEW MET FRANCIS WYFFELS	148

5 VERANTWOORDELIJKE AI 150

Hoe maken we AI verantwoordelijk? 151

Onverwachte gevolgen 154

Eed van Hippocrates 156

Vier grondprincipes 159

Red Queen Hypothese 164

Richtlijnen voor AI 166

Verbod 167

Asimov 168

INTERVIEW MET NATALIE BERTELS EN PEGGY VALCKE 171

INTERVIEW MET KATLEEN GABRIELS 174

CONCLUSIE VAN BREIN NAAR PLANT 177

Wat is de toekomst van AI? 178

Vier fasen 182

Breinchauvinisme 185

Andere vormen van intelligentie 188

De Celestijnse Belofte 195

EINDNOTEN 199



INTRODUCTIE

EEN HANDLEIDING VOOR AI-VERTALERS

“Dat is allemaal goed en wel, maar misschien moeten we ook iets met AI doen?” Het is een doordeweekse ochtend in 2019 wanneer die vraag weerklinkt. Samen met enkele collega’s van imec zit ik in een meeting met het management van een Belgische energieleverancier. De vraag zorgt onmiddellijk voor verwarring aan onze kant van de tafel. Het voorbije uur praatte we namelijk over hoe automatische systemen en datagedreven methodes, die steunen op alles van klantendata tot slimme meters, er mogelijk voor zorgen dat ze hun klanten persoonlijker zullen benaderen in de toekomst. We hadden het met andere woorden dus al een uur lang over artificiële intelligentie, maar dan zonder die woorden in de mond te nemen. Iets dat het management blijkbaar helemaal niet doorhad.

Vandaag lachen we natuurlijk met die vraag, gewoon een spraakverwarring tijdens een meeting. Maar tegelijk duidt het op een belangrijke en beangstigende realiteit. Een managementteam dat uiteindelijk verantwoordelijkheid moet dragen over AI-systemen, weet eigenlijk niet wat ze zijn of hoe ze in elkaar zitten. En dit specifieke managementteam was allesbehalve uniek. Er mag dan een geweldige hype bestaan rond artificiële intelligentie, kennis van deze systemen ontbreekt vaak bij beslissingsmakers.

Later die dag zit ik opnieuw in mijn kantoor. Een van mijn onderzoekers komt binnen met de laatste onderzoeksresultaten. Onze nieuwe aanpak toont theoretisch een verbetering van de accuraatheid van een AI-systeem. Een klein eureka-moment voor mijn team dat AI-systemen ontwikkelt binnen imec, vanwaar we die inzichten terug doorspelen aan de industrie. Maar dan gaan de radertjes draaien in mijn hoofd. De wiskundige formule achter die nieuwe resultaten is misschien erg elegant, maar de implicaties ervan zijn veel complexer. Hoe relevant is dit voor de industrie die op onze algoritmes steunt om hun operaties efficiënter te maken? En zelfs als het nieuwe systeem relevant is, zal het de meerkost dan waard zijn? Hoe leg je dit onderzoek in godsnaam uit in mensentaal terwijl je transparant blijft? En lost dit überhaupt een probleem uit de echte wereld op?

Maar de vragen reiken ook verder dan puur de economische invalshoek. Wat als dit algoritme op het internet terechtkomt, en voor duistere doeleinden wordt ingezet? Het zou namelijk niet de eerste keer zijn dat een technologie van enthousiaste, enigszins naïeve, onderzoekers plots de basis vormt voor software die de privacy of basisrechten van mensen schendt. Een mentaliteit waarop ik mezelf regelmatig betrap. Ik ben vaak enthousiast over technologische vooruitgang, zonder me te realiseren dat er ook een maatschappelijke invalshoek aan is. Hoe zorgen we ervoor dat dit niet gebeurt met de formule die voor me ligt?

Wanneer ik thuiskom, rinkelt de telefoon, een bezorgd familielid belt me. Ze zag een debat over AI en dat maakte haar bang. Dus belde ze maar naar mij. Nemen robots ons werk over? Blijven autonome wagens ongevallen veroorzaken? Wie is daar schuldig aan? Hoe lossen we ethische dilemma's op, zoals de vooroordelen die we als mensen in algoritmes stoppen? Ik probeer alles zo goed mogelijk te beantwoorden, maar er komt vooral angst uit die telefoon. Angst die ik maar moeilijk wegneem.

AI-VERTALER

Spijtig genoeg zijn dit soort dagen al lang niet meer de uitzondering. AI of artificiële intelligentie neemt zeer snel aan belang toe in onze maatschappij. Doorbraken stapelen zich op, en AI-systemen dringen overal door in onze leefwereld: van de series die Netflix ons aanraadt op basis van ons eerder kijkgedrag tot de poortjes op het vliegveld die onze gezichten scannen.

Zelfs komkommers zijn niet meer veilig, dat bewees de Japanner Makoto Koike, een systeemingénieur uit de Japanse auto-industrie, die in 2016 begon mee te helpen op de komkommerboerderij van zijn ouders. Die teelt was op dat moment nog een erg manueel proces. Vooral het sorteren van de komkommers in categorieën kostte veel werk. Tijdens het oogstseizoen spendeerde Makoto's moeder maar liefst acht uur per dag aan het onderverdelen van de oogst in negen categorieën. Afhankelijk van de vorm, grootte en kleur verdeelde ze de komkommers volgens een systeem dat te moeilijk was om aan deeltijdse arbeiders te leren. Niettemin vond Makoto dat het beter kon, en daarom liet hij er AI op los. Hij gebruikte een zogenaamd neuraal netwerk, een AI-systeem dat onze eigen hersenen imiteert en dat we verder in het boek uitleggen, en trainde dat op een hoop foto's van komkommers. Dat deed hij met open-source tools van Google, die iedereen gewoon gratis

online kan terugvinden. Het systeem leerde al snel om de methode van Makoto's moeder over te nemen en de komkommers onder te verdelen in negen categorieën.

Dat systeem sloot Makoto aan op een automatische verdeler. Met een hoop eigen elektronica zoals een Raspberry Pi, een kleine computer voor knutselprojecten, stak hij een robot in elkaar die volgens de categorisering van het AI-model de komkommers onderverdeelde. Daarmee automatiseerde hij de taak die zijn moeder hele dagen bezighield.¹

Kennis van AI democratiseert zich dus razendsnel. Een niet-expert zoals Makoto stak zelf een systeem in elkaar dat enkele decennia geleden ondenkbaar was. Ondertussen dringt AI stelselmatig door in alle uithoeken van onze maatschappij. In 2017 stelde Andrew Ng, een belangrijke AI-onderzoeker die onder meer bij Google en het Chinese Baidu werkte: "Net zoals elektriciteit bijna alles transformeerde 100 jaar geleden, vind ik het vandaag moeilijk om me een industrie in te beelden die AI niet zal transformeren in de komende jaren."²

AI is dus de nieuwe elektriciteit volgens Ng. Elektriciteit transformeerde namelijk onze hele wereld, straten en huizen werden verlicht, apparaten namen huishoudelijke taken over en elektrische machines automatiseerden werk in fabrieken. Vandaag zorgen AI-algoritmes er dan weer voor dat Amazon ons efficiënter producten aanraadt op hun platform, dat dokters ons betere diagnoses geven, dat menselijk werk geautomatiseerd wordt en dat we er online in slagen om bijna elke taal te vertalen.

Tegelijk duiken er ook heel wat dreigingen op. We zijn bang dat AI onze jobs zal innemen, dat het onze privacy zal schenden, dat er killerrobots op ons afkomen en zelfs dat AI een superintelligent wezen zal bouwen dat sciencefictiongewijs onze wereld overneemt.

Daarom baren mijn dagelijkse meetings me zorgen. Managers die de eindverantwoordelijkheid dragen, of zullen dragen over dit soort systemen, snappen de werking en de implicaties ervan niet. AI-ingenieurs hanteren daarnaast nog te vaak een erg nauwe, technische bril. De maatschappelijke gevolgen zijn voor hen nog al te vaak externaliteiten die ze negeren. Er zit een zekere ironie in het feit dat we, alhoewel AI ervoor zorgt dat online vertalers zoals Google Translate vandaag vrij goed werken, eigenlijk personen ontbreken

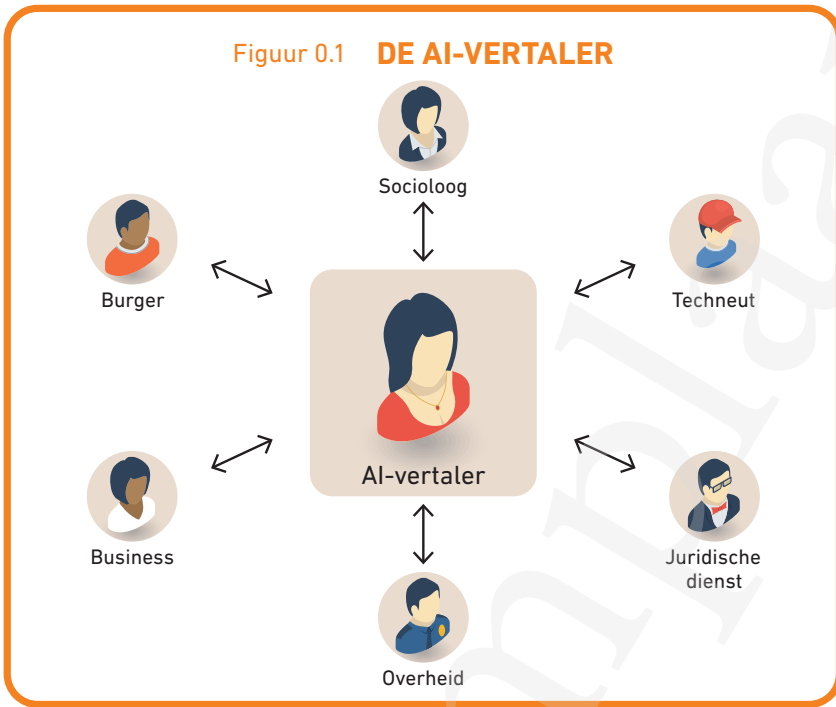
die de specificiteiten van AI vertalen tussen verschillende groepen. Volgens mij ligt de grote uitdaging van AI niet in de technische sfeer, maar wel in die vertaalslag. Natuurlijk bestaan er nog een reeks technische uitdagingen voor AI, die we niet mogen onderschatten. Maar waar de implementatie van deze systemen nog vaak faalt, is in de vertaling van het technische naar de echte wereld.

We mogen ons niet blindstaren op de technologie. In grote delen van de AI-gemeenschap heerst er nog steeds de visie dat enkel technische regels nodig zijn om de slechte gevolgen van AI te voorkomen. Dat is fout volgens mij. We moeten natuurlijk ook technische fixes najagen, maar uiteindelijk draait het erom dat we kennis rond AI-systemen democratiseren en er een breder debat over voeren. Ingenieurs zullen er nooit in slagen om alle mogelijke toekomstscenario's te voorzien of de complexiteit van onze wereld volledig te begrijpen. De AI-systemen die zij bouwen zullen dus nooit perfect zijn. We moeten daarom zoveel mogelijk disciplines aan het debat laten deelnemen. Tegelijk mogen we als er iets fout loopt niet enkel naar de ingenieur wijzen als de schuldige, want vaak is die persoon gewoon één radertje in een complexer geheel.

Juristen, managers, burgers en ingenieurs kijken allemaal op een andere manier naar AI, en vragen andere dingen van deze systemen. Daarom hebben we personen nodig die daartussen staan. Personen die de noden en de eisen van ingenieurs vertalen naar managers, of die de bezorgdheden van burgers aan technenuten overbrengen. Personen die analyseren of een algoritme mogelijk schadelijk is als we het online plaatsen. Of die een management-team weten uit te leggen dat AI niet de beste oplossing is voor hun noden (of net wel). Personen die de werking van technische AI-systemen uitleggen aan juristen, zodat zij daar een correcte regelgeving rond opbouwen.

We hebben met andere woorden bruggenbouwers nodig die tussen werelden springen (zie figuur 0.1). AI is vandaag per definitie multidisciplinair, maar ondertussen behandelen we het vanuit aparte silo's. We zitten met te veel AI-specialisten, en te weinig generalisten. Te veel mensen die veel weten over één aspect, en te weinig over verschillende disciplines tegelijk.

Figuur 0.1 **DE AI-VERTALER**



SURFEN OP HET ALGORITME

Mijn ervaring met AI gaat al even terug. Ik schreef er als jonge ingenieur mijn masterthesis over. Dat was in het begin van de jaren '90 in Duitsland. Vandaag daarentegen ben ik de programmadirecteur artificial intelligence bij imec, het grootste onafhankelijke onderzoeksinstituut van België dat zich specialiseert in digitale en nanotechnologie.

Toen ik met AI begon, zaten we nog lang niet in de 'zomer' van grote aandacht en hype waarin we nu zitten. De computers waren trager, de data beperkter en de algoritmes minder indrukwekkend. Als je toen een neuraal netwerk van acht lagen en een dertigtal variabelen wilde opbouwen (wat dat betekent leggen we verder in het boek uit), dan duurde zoiets drie dagen, zelfs met de nieuwste en snelste computers. Ik woonde in Stuttgart, ongeveer vijf uur rijden van het Italiaanse Gardameer. Ik kon dus perfect met mijn medestudenten in een Volkswagenbus naar Italië rijden om te gaan windsurfen terwijl mijn algoritme zich trainde. Toen ik terugkwam, was mijn model klaar. Sindsdien maakte de exponentiële groei van AI dit soort tripjes

spijtig genoeg voor vele studenten onmogelijk, het trainen van standaard algoritmes gaat vandaag super snel dankzij veel snellere computers.

AI onderging al pieken en dalen van hype en technologische ontwikkeling, iets dat eigen is aan de technologie-industrie. Sinds die vroege dagen van windsurfen zag ik hoe de Amerikaanse dotcom-bubbel barstte aan het begin van de 21ste eeuw, maakte ik van binnenuit mee hoe techbedrijven zoals Microsoft en SAS zich moesten heruitvinden door de komst van nieuwe spelers en help ik nu een toekomstgerichte Europese aanpak voor AI uit te bouwen bij imec.

Die ervaring bundel ik in dit boek, als handleiding voor toekomstige AI-vertalers. We zullen de technische aspecten van AI dus bespreken, maar ook de maatschappelijke implicaties ervan. We bekijken wat de wet van Moore is, en wat General Adversarial Networks zijn, maar ook wat dat betekent voor onze privacy en de toekomst van ons werk. We kijken hoe mens en machine dus beter kunnen samenwerken.

Dat is volgens mij nu meer dan ooit nodig. AI kent natuurlijk een gigantische hype, maar dat blijft niet noodzakelijk duren. Zoals we verderop in het boek zullen zien, ging AI door periodes van zomer en winter, fases dus van grote hype en bijhorende onderzoeksbudgetten, en decennia's van teleurstelling en opdrogende investeringen. Vandaag leven we in een zomer, maar als we de tegenstellingen die AI momenteel tegenkomt, op ethisch maar ook technisch gebied, niet behandelen, dan lopen we de kans dat het opnieuw winter wordt. Iets wat enorm problematisch zou zijn. AI bezit een gigantische capaciteit om onze wereld beter te maken en een vroegtijdige winter laat die toepassingen misschien afsterven. Daarom zijn AI-vertalers meer dan ooit nodig, vandaar dit boek.

Maar om dat te doen moeten we eerst beginnen met ons onderwerp af te lijnen. Wat is dat dus, AI?

WAT IS AI?

De term artificiële intelligentie of AI bestaat maar uit twee woorden, maar zorgt tegelijk voor veel verwarring. Academics en onderzoekers debatteren namelijk al decennia over de exacte betekenis van het concept. Tijdens de voorbije jaren zagen we ook een vermenigvuldiging van het gebruik van

de term, waarbij alle computersystemen die in staat zijn om automatische taken en beslissingen uit te voeren het label toegewezen kregen. AI bleek zelfs een marketingterm, en bedrijven begonnen het te gebruiken om hun producten een geavanceerd, technologisch imago te bezorgen.

Laten we de twee woorden dus uit elkaar halen. Over het woord ‘artificieel’ bestaat niet zo veel discussie. Iets is artificieel als het niet op een natuurlijke manier ontstaat of gebeurt, en is een synoniem van kunstmatig. Je mag dus evengoed kunstmatige intelligentie zeggen.

‘Intelligentie’ blijkt een pak moeilijker. Filosofen kunnen het concept anno 2020 nog steeds niet eenduidig definiëren. Psychologen daarentegen beschrijven intelligentie als een mentale eigenschap die bestaat uit heel wat componenten, zoals analytische intelligentie, praktische intelligentie, sociale intelligentie, emotionele intelligentie, creativiteit en wijsheid.³

Intelligentie definiëren is dus moeilijk. Verschillende disciplines en onderzoekers zijn het niet met elkaar eens over wat het zou moeten zijn. Bijgevolg is het ook moeilijk om een duidelijke, allesomvattende definitie te maken van artificiële intelligentie. Daarom kiezen we in dit boek voor een pragmatische weg, en nemen we een gangbare en vooral praktische definitie. Mits wat simplificatie kun je vandaag AI-systemen aflijnen als:

**“Computersystemen die zelfstandig leren,
beslissingen nemen en ze ook uitvoeren.”**

Die definitie bevat zowat alle belangrijke kenmerken van hedendaagse AI-systemen. AI werkt zo op een computer. Het leert zelfstandig iets doen, een menselijke programmeur hoeft niet exact, stap voor stap, te vertellen wat er moet gebeuren, zoals wel het geval is wanneer je normale software schrijft. AI leert dingen, het leert bijvoorbeeld een kat op een foto herkennen of een bordspel spelen. En ten slotte is het ook naadloos aangesloten op een beslissingssysteem en onderneemt het ook actie, het herkent bijvoorbeeld een bepaald type komkommer uit een reeks foto's en geeft een commando aan een robotarm om die komkommer weg te nemen.

DE ORIGINELE DROOM

De originele droom van AI was om computersystemen te bouwen die menselijke intelligentie nabootsen om zo gevaarlijke, saaie of vuile taken over te nemen. In het Engels zijn dat de drie D's: dangerous, dull en dirty. De huidige grotere droom achter die autonome en zelflerende systemen is om alle vormen van menselijke intelligentie na te bootsen, en daarbij ook 'moeilijke' taken te realiseren. Een vierde D werd daardoor toegevoegd aan het lijstje: difficult. Dat is een gigantische uitdaging want menselijke intelligentie neemt enorm veel vormen aan. Denk maar aan verbaal-linguïstische intelligentie, logisch-mathematische intelligentie, visueel-ruimtelijke intelligentie, muzikale intelligentie, creativiteit, lichamelijke intelligentie, intrapersonlijke intelligentie (intelligentie die gaat over de collaboratie tussen personen) of interpersoonlijke intelligentie (inzicht in hoe je jezelf mentaal en lichamelijk in balans houdt).

Als we dus menselijke intelligentie willen nabootsen, dan moet AI minstens de volgende stappen ondersteunen:

- Verwerken: een AI-systeem moet input tot een bruikbaar formaat omvormen.
- Begrijpen: een AI-systeem moet verbanden of patronen zien in de verwerkte input.
- Redeneren: een AI-systeem moet logica toepassen op de inzichten en zo tot een beslissing komen.
- Ageren: een AI-systeem moet op basis van deze beslissing de juiste actie ondernemen.

Dat zijn allemaal stappen die mensen ook nemen. We halen via onze zintuigen indrukken binnen en die verwerken we zodat we onze omgeving begrijpen. Daarna denken we na over de situatie waarin we zitten, en maken we een keuze over wat we willen doen. Die actie ondernemen we vervolgens.

Het is een cyclus die we elke dag duizenden keren herhalen, heel vaak zelfs onbewust. Elke keer wanneer we de deur opendoen, wanneer we groenten snijden of een bal heen en weer gooien, doorlopen onze hersenen deze cyclus. Voor ons is het makkelijk, we doen bijna niets anders. Maar voor computers is dit aartsmoeilijk, en beginnen we nu pas stapvoets systemen te bouwen die zo'n cyclus doorlopen voor heel specifieke taken.

Om die cyclus te doorlopen vereist een AI-systeem drie grote bouwblokken:

- 1 Ten eerste is er het algoritme. Wiskundige formules dus die computers in staat stellen om te leren.
- 2 Ten tweede verwacht een AI-systeem data. Die dienen als (ruwe) input voor het algoritme, en doorgaans hebben we eerst grote hoeveelheden ervan nodig om een systeem intelligent te maken.
- 3 Ten derde heeft AI ook hardware nodig. AI vereist veel van computers, en verwacht naast grote dataopslag ook snelle chips om berekeningen uit te voeren.

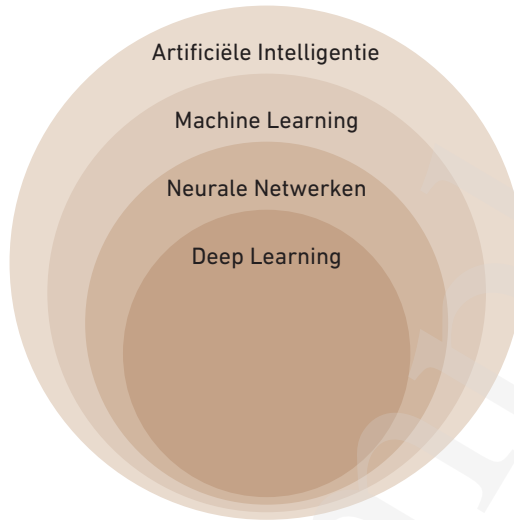
Met die drie bouwblokken wordt een computer vrij rudimentair intelligent.

MACHINE LEARNING EN NEURALE NETWERKEN

Los van die drie bouwblokken bestaat er vandaag nog een waterval aan termen die we gebruiken in de context van AI, denk maar aan machine learning, deep learning of neurale netwerken (zie figuur 0.2). Om die termen te verklaren moeten we kort terugkeren in de geschiedenis van AI. De term ontstond in 1956 toen de Amerikaanse computerwetenschapper John McCarthy hem voor het eerst gebruikte op een academische conferentie. AI gebruikten we aanvankelijk als overkoepelende term voor machines die taken intelligent uitvoeren. De eerste systemen waren instructiegebaseerd, een programmeur programmeerde ze dus regel per regel. Dit soort systemen staan vol met mogelijke situaties en de bijbehorende antwoorden; als situatie A optreedt, dan moet het systeem actie B nemen.

Die systemen waren makkelijk begrijpbaar en transparant voor mensen, maar kenden ook enkele nadelen. Ze vereisten enerzijds enorm veel tijd om bepaalde oplossingen te programmeren, en anderzijds zijn de oplossingen voor heel wat problemen moeilijk vooraf neer te schrijven in een reeks instructies. Bedenk maar eens dat je instructies moet neerschrijven over hoe je foto's van honden of katten van elkaar moet onderscheiden, wat vandaag een standaard toepassing is van AI. Verder ageren regelgebaseerde systemen super stabiel in een gekende context, maar ondernemen ze geen actie als de programmeur geen regel voor een bepaalde situatie vooraf definieerde. Dat is iets wat problematisch is, gezien onze wereld vaak complex en moeilijk te voorspellen is.

Figuur 0.2 **RELATIE TUSSEN VERSCHILLENDE BUZZWORDS**



Machine learning (ook wel afgekort als ML) bood op deze nadelen een antwoord. Bij machine learning leren computers op basis van data en passen ze zich aan door ervaring, zonder dat een programmeur de regels definieert. Zo slaagt deze techniek erin om op voordien ongekende situaties te reageren, wat een enorm voordeel is. Tegelijk echter maakt het de reden achter een beslissing weinig transparant, een probleem dat we later in het boek in detail bespreken. Het grootste deel van wat we vandaag onder AI-systemen rekenen valt onder machine learning.

Binnen machine learning vind je dan weer neurale netwerken. Dat is een heel populaire techniek om machines te doen leren, een techniek die de complexe structuur van het menselijke brein nabootst. Ons brein bestaat namelijk uit neurale netwerken van lagen neuronen, die informatie doorgeven aan andere lagen neuronen. Een systeem dat we kopiëren naar computers via neurale netwerken, en dat we verder in het boek in detail bespreken. Deep learning is een naam voor het leren via neurale netwerken. Het woord 'deep' voegden we de laatste jaren daaraan toe, want bij grote hoeveelheden complexe data bestaan er heel veel lagen neuronen in het netwerk. Het systeem wordt dus steeds dieper.

Deep learning en neurale netwerken zijn dus een subdomein van machine learning, en machine learning is op zijn beurt een subdomein van AI. In de praktijk verwijst AI echter vooral naar machine learning, wat op zijn beurt vooral verwijst naar neurale netwerken en deep learning omwille van het succes van deze technieken.

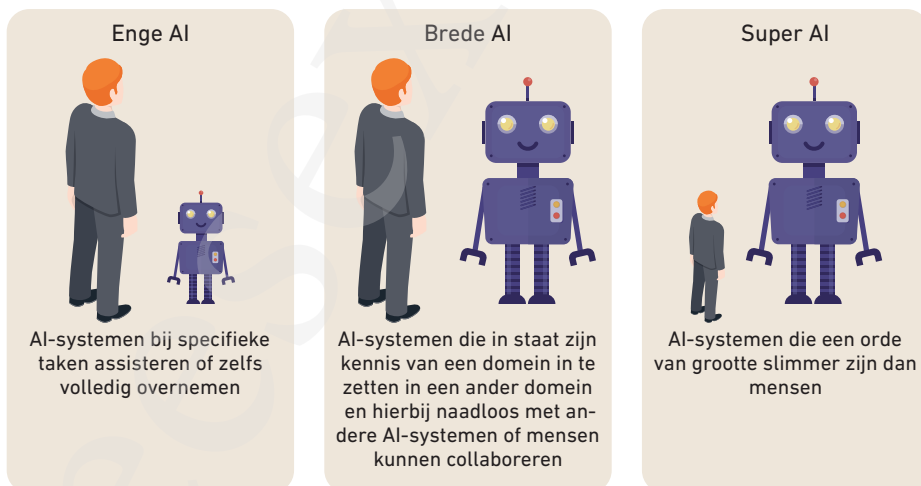
VAN ENGE NAAR SUPER AI

Er zijn verschillende manieren om AI-systemen in te delen. De meest gebruikte onderverdeling is die tussen enge of nauwe AI (in het Engels 'narrow'), brede AI (in het Engels 'general') AI en super AI (zie figuur 0.3).⁴

ENGE AI

Enge of nauwe AI-systemen bevatten oplossingen die één welomschreven taak in een specifieke context extreem goed uitvoeren (zoals een spamfiltertool of het Japanse systeem dat komkommers in categorieën indeelt). Nauwe AI-systemen zijn hyperspecialisten in hun eigen kleine domein. In die ene welomschreven taak in een heel specifieke context zijn ze vaak beter dan mensen. Buiten die context is het systeem daarentegen volstrekt wereldvreemd. Het Japanse komkommersysteem kan in sommige gevallen

Figuur 0.3 DRIE NIVEAUS



dit type groenten beter indelen dan mensen, maar zolang je het niet opnieuw traint, zal het nooit spam uit je mail kunnen filteren. Het Japanse systeem zal er zelfs niet in slagen om Europese komkommers te klasseren omdat die er anders uitzien dan hun Japanse tegenhangers.

Daar staan we nu. Alle AI-systemen vandaag zijn nauwe AI. Ze zijn goed in het verwerken van informatie en het zoeken van patronen, maar ze bezitten geen redeneervermogen en handelen niet buiten hun specifieke context. AI biedt op die manier veel waarde, maar het is nog ver verwijderd van wat we in de volksmond intelligent zouden noemen.

Neem bijvoorbeeld ons menselijk redeneervermogen. Stel dat we op safari gaan in Zuid-Afrika, en er plots een leeuw voor onze neus staat. Onze zintuigen geven signalen door aan onze hersenen, zoals het beeld van de leeuw of het gebrul ervan. Onze hersenen bewerken en interpreteren die signalen, zodat we weten dat er een leeuw voor onze neus staat. We begrijpen wat we zien, en we analyseren op basis van die informatie of we tot actie overgaan. We besluiten in de meeste gevallen dan om weg te lopen. Als we daarentegen een leeuw in de zoo zien, achter tralies, dan zal onze reactie helemaal anders zijn. We lopen dus niet weg.

De meeste commerciële AI-systemen bezitten dit soort intelligentie niet automatisch. Ze zijn goed in het verwerken van signalen, en het herkennen van patronen. Maar het redeneren ontbreekt. Een AI-systeem leert misschien erg efficiënt foto's van een leeuw herkennen. Maar het zal maar erg moeilijk het algemene redeneervermogen en begrip van de context ontwikkelen die mensen doet beseffen dat een leeuw op de savanne gevaarlijk is.

BREDE AI

Mensen bezitten dus een graad van algemene intelligentie waaraan AI-systemen niet kunnen tippen. In sommige taken, zoals het spelen van bordspelletjes, verslaat een AI-systeem ons misschien keer op keer. Maar buiten die context is dat systeem minder intelligent dan een kleuter. Niettemin werken onderzoekers ook aan brede of general AI. Simpel gezegd zou een breed AI-systeem kunnen schakelen tussen verschillende taken en verschillende functies automatisch moeten kunnen aansturen. Bijvoorbeeld door een slimme combinatie te maken van informatie uit verschillende sensoren. Net zoals een mens dat doet. Wat makkelijker gezegd is dan gedaan, want een computer zou daarbij ook in staat moeten zijn om flexibel en rationeel