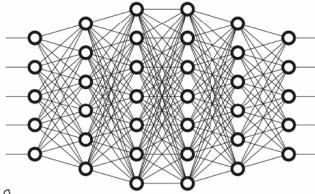
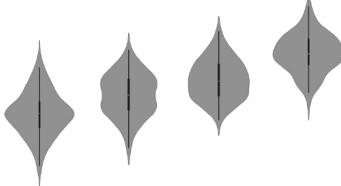
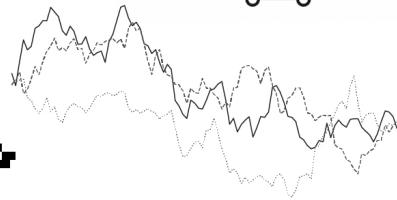
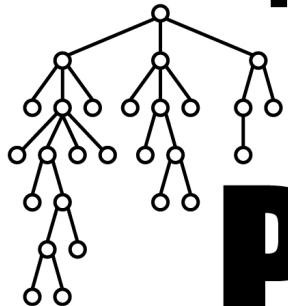


$((p \rightarrow q) \wedge p) \models q$



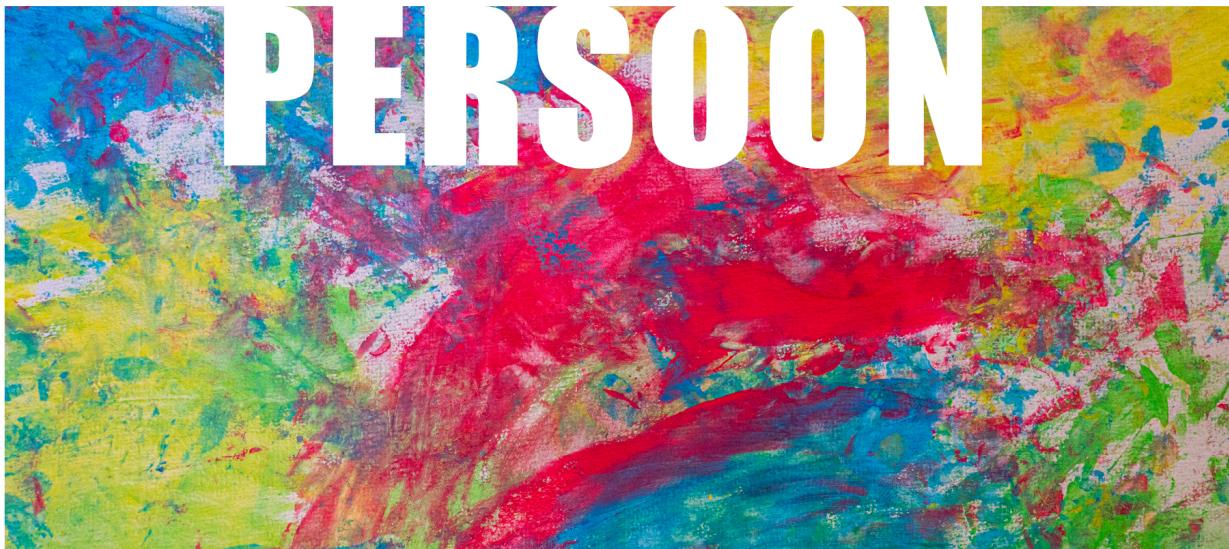
5 4 1 4 0 8 4 1 5 3 8 0 7 7 4 8 1 6 0 0  
8 7 2 6 9 7 0 9 7 4 5 7 6 8 9 5 7 6 0 8  
3 1 7 9 5 1 5 2 6 9 8 5 2 3 9 2 0 8 1 7  
1 0 1 6 6 2 8 2 9 6 6 7 5 6 8 6 4 6 7 9  
5 5 2 4 0 0 5 0 0 3 9 3 8 7 2 3 1 6 1 5  
9 8 2 4 8 1 0 5 8 5 9 4 3 0 1 1 3 5 4 4  
8 3 7 5 2 6 3 8 8 / 4 1 9 1 5 4 6 8 7 1  
3 8 4 0 6 1 3 1 1 0 5 2 0 0 9 5 1 1 0 9  
1 4 0 1 6 0 8 4 5 1 4 6 0 9 8 9 8 9 7 7  
9 8 0 2 3 9 3 4 8 3 4 7 2 1 4 9 3 9 0 5



# PATROON

*tussen* *en*

# PERSON



## Tussen patroon en persoon

*een rede[n] voor  
symbiotische (artificiële) intelligentie*

Jeroen Linssen

# Inhoudsopgave

<b>Voorwoord: Denken in patronen</b>	<b>4</b>
<b>1 Introductie: Over AI</b>	<b>7</b>
1.1 Definities van AI . . . . .	7
1.2 Doorbraken in AI . . . . .	13
1.3 Discussies over AI . . . . .	16
<b>2 Uitdagingen: Het belang van wederzijds begrip</b>	<b>19</b>
2.1 De huidige hype en schaduwzijdes van AI . . . . .	19
2.2 Van hype naar hoop: AI bij Ambient Intelligence . . . . .	22
2.2.1 Het proces centraal . . . . .	22
2.2.2 Het begrip centraal . . . . .	25
<b>3 Toekomst: Grenzen vervagen</b>	<b>29</b>
3.1 Oneindige technologische groei? . . . . .	29
3.2 De horizon van AI . . . . .	31
3.3 De horizon van mensen . . . . .	35
<b>4 Actie: Samen verantwoordelijk</b>	<b>38</b>
4.1 Terug naar het nu: what will (A)I do? . . . . .	38
4.1.1 Aanbevelingen over AI . . . . .	38
4.1.2 Het lectoraat in actie . . . . .	39
4.1.3 Samen in actie . . . . .	41
4.2 Huiswerk: leven is leren . . . . .	41
4.2.1 Geletterdheid verbeteren . . . . .	42
4.2.2 Methodologisch te werk . . . . .	42
4.2.3 Kritisch nadenken over AI . . . . .	43
4.2.4 De rode draad . . . . .	43
<b>Nwoord</b>	<b>45</b>
Dankwoord . . . . .	45
Reflectie . . . . .	46
Media-aanraders . . . . .	48
<b>Over deze rede</b>	<b>51</b>
Over Jeroen Linssen . . . . .	51

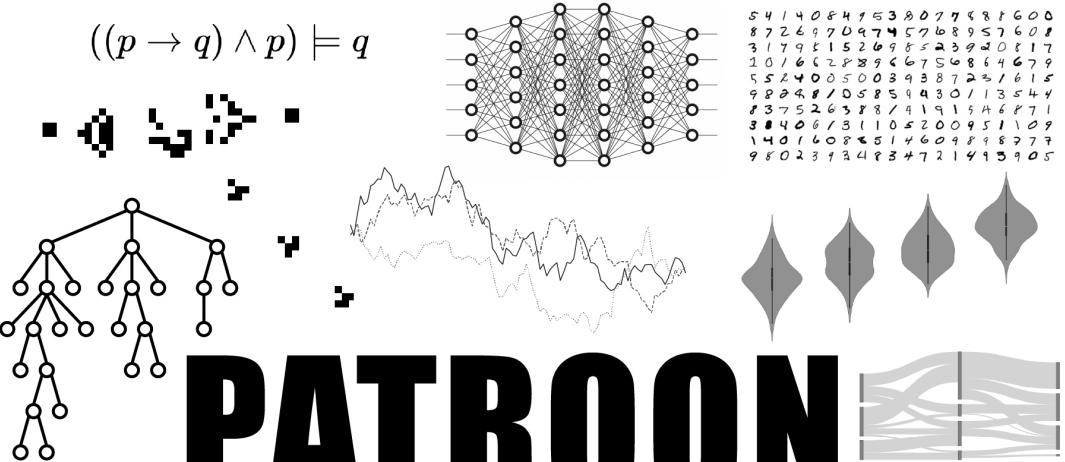
Open science . . . . .	52
<b>Referenties</b>	<b>53</b>
Copyright . . . . .	59

# **Voorwoord: Denken in patronen**

Dit is de lectorale rede van Jeroen Linssen, lector Symbiotic AI bij het lectoraat [Ambient Intelligence](#) van [Hogeschool Saxion](#). De volledige titel van deze rede luidt als volgt.

**Tussen patroon en persoon: een rede(n) voor symbiotische (artificiële) intelligentie**

Deze rede is bedoeld om als website te worden gelezen. Voor het gemak van de lezer zijn ook een [pdf-versie](#) en een [epub-versie](#) (voor e-readers) beschikbaar.



# PATROON

*tussen en*



‘It is possible I had some presentiment of my future.’ (Wolfe, 1998)

Patronen zijn overal. In de natuur, de vormen van planten en dieren, in onze leefomgeving, de manieren waarop wij onze wereld inrichten. In de tijd, de seizoenen, de etmalen. In alle talen die we spreken, lezen en schrijven. In abstractere zaken zoals gevoelens, metaforen en verhalen.<sup>1</sup> Patronen geven aan dat er een bepaalde regelmaat in iets zit, iets dat voorspelbaar

<sup>1</sup>Ook deze rede valt vaak terug op bepaalde patronen en metaforen, in de hoop dat bepaalde concepten op die manier helderder worden. In metaforen schuilt echter ook het risico dat patronen op een onbedoelde manier geïnterpreteerd worden.

is. Deze volgordes maken dat alles bestaat in samenhang en al dan niet stabiel evenwicht. Er zit een logica in de bouwstenen van onze wereld. Wij, als mensen, vormen ieder een wereldbeeld op basis van deze patronen. Dankzij onze intelligentie zijn wij in staat bepaalde taken uit te voeren, daarmee nieuwe patronen te maken, en hiervan te leren. Intelligentie is echter niet alleen aan mensen voorbehouden: er zijn tal van organismen die intelligent gedrag vertonen. Maar is het dan ook mogelijk om *systemen te maken* die patronen kunnen herkennen en genereren? Jazeker. Dit is precies waar het vakgebied *kunstmatige of artificiële intelligentie* (**AI**) om draait: het onderzoeken en bouwen van intelligente systemen.

De rode draad van deze rede is **symbiose**: het samen leven van twee levensvormen. Mensen en technologie zijn onlosmakelijk verbonden: dankzij technologie heeft de mensheid zich kunnen ontwikkelen. Maar technologie heeft omgekeerd ook invloed op de mensheid. Mijn boodschap is dat wij, als mensen, verantwoorder en duurzamer met AI om moeten gaan door beter te leren samenleven met deze technologie. Ik bedoel symbiose hier explicet als *mutualistisch*: zowel menselijke als artificiële intelligentie zou baat moeten hebben bij mijn voorgestelde samenwerking tussen beiden. We willen als mensen betere, slimmere systemen die begrijpen wat wij doen en nodig hebben, voor een verbetering van onze kwaliteit van leven. Dit vereist dat deze systemen de juiste patronen kunnen herkennen en nieuwe patronen kunnen maken die aansluiten op onze verwachtingen. Andersom hebben wij de verantwoordelijkheid om zelf deze systemen ook te snappen, zodat wij begrijpen wat de systemen doen en nodig hebben. We streven naar **wederzijds begrip** en *hybride intelligentie*. Alleen als we deze keerzijden van dezelfde medaille voldoende aandacht geven, zullen we komen tot een werkelijke symbiose tussen menselijke en artificiële intelligentie.

In deze rede leg ik uit hoe we tot zulke **symbiotische (artificiële) intelligentie** kunnen komen en wat we binnen ons lectoraat Ambient Intelligence hieraan onderzoeken. Ter introductie geeft het eerste hoofdstuk in deze rede een blik op de ontwikkeling van AI, waarna de rede over gaat naar huidige uitdagingen, de (mogelijke) toekomst van AI, en een oproep tot actie. Ieder hoofdstuk begint met een samenvatting van vijf zinnen in klare taal, zodat jij, de lezer, weet wat er komen gaat. Mijn streven was om de tekst zo helder mogelijk op te schrijven, zodat iedereen er zoveel mogelijk van kan leren.

Ik hoop dat deze rede bijdraagt aan een beter begrip van AI, bewustwording van de (on)mogelijkheden van deze technologie, en jou aan het denken zet. We moeten daarvoor samen nieuwe patronen maken en daarvoor vele rode draden spinnen. De rode draad van deze rede is slechts één lijn in dit geheel, en ook in deze tekst blijkt dat veel andere draden verweven zijn met de hoofdlijn.<sup>2</sup> Ik verwacht niet dat AI op zichzelf alle uitdagingen en transities die ons nog te wachten staan, zal oplossen. Dit wordt een taak die samenwerking eist tussen alle vakgebieden en sectoren. Zoals vaker wordt gezegd: de beste manier om de toekomst te voorspellen, is door deze te maken.<sup>3</sup> Laten we dat samen doen.

---

<sup>2</sup>Ik kan me er niet van weerhouden voetnoten met extra gedachten te tikken voor deze extra draden.

<sup>3</sup>De oorsprong van deze uitspraak is niet compleet te herleiden, maar waarschijnlijk was het [Dennis Gabor](#).

# 1 Introductie: Over AI

## Samenvatting

AI is geen magie: het is een vakgebied dat zich richt op het ontwikkelen en onderzoeken van intelligente systemen. Hiervan zijn veel verschillende vormen, zoals zoekmachines, beeldherkenning en het genereren van teksten. AI is gebouwd op basis van algoritmes: instructies die computers kunnen uitvoeren. In de loop der jaren zijn veel doorbraken geweest in dit veld en gaat de discussie ook over de effecten van deze technologie. In deze rede licht ik een aantal uitdagingen toe die vandaag de dag en in de nabije toekomst aangepakt moeten worden.

Er zijn tal van manieren om een rede als deze te beginnen. Minstens zoveel als het aantal redenen om met *Artificiële Intelligentie* (AI) bezig te zijn. AI is een technologie die toegepast kan worden om processen te verbeteren, van zorg tot industrie en van energietransitie tot sociaal werk. Het kan gebruikt worden om betere diagnoses te stellen voor ziektes (Kumar e.a., 2023), automatisch te regelen hoe stroomvoorzieningen optimaal draaien (Shi e.a., 2020), en inzicht geven in maatschappelijke problemen (Nishant e.a., 2020).<sup>1</sup> AI kan ook gebruikt worden om te onderzoeken wat intelligentie nu precies is, bijvoorbeeld om beter te begrijpen hoe ‘leren’ werkt (Modha e.a., 2011). Iedereen zal een eigen voorstelling kunnen maken van wat AI wel of niet is. Mijn doel bij deze rede is het verbeteren van het begrip van iedereen over AI en uitleggen hoe ons lectoraat Ambient Intelligence bijdraagt aan een verantwoorde, duurzame inzet en ontwikkeling van AI. Voordat ik in ga op wat we *nu* zouden moeten doen met AI, belicht ik in dit hoofdstuk eerst hoe we tot hier gekomen zijn en wat volgens mij belangrijke lessen zijn, door te praten over definities van, doorbraken in en discussies over AI.<sup>2</sup> Daarna gaat deze rede verder met uitdagingen van vandaag de dag (Hoofdstuk 2, reflectie), de mogelijke toekomst van AI (Hoofdstuk 3, speculatie) om tenslotte af te sluiten met een oproep tot actie (Hoofdstuk 4, activatie).

## 1.1 Definities van AI

Als vakgebied bestaat AI officieus sinds 1956, toen de **Dartmouth Summer Research Project on Artificial Intelligence** in de zomer van dat jaar gehouden werd op Dartmouth

<sup>1</sup>Hierom wordt AI tegenwoordig bestempeld als *key enabling technology* (KET). Andere **KETs** zijn bijvoorbeeld fotonica, nanotechnologie, robotica en quantum computing.

<sup>2</sup>Ook hiervoor geldt het patroon dat, wie niet leert van de geschiedenis, gedoemd is deze te herhalen.

College in de Verenigde Staten.<sup>3</sup> Natuurlijk waren daarvoor al allerlei pogingen gedaan om intelligentie te begrijpen en te automatiseren, of dat nu in fictieve werken was (Cave & Dihal, 2019) of in wetenschappelijke (Russell & Norvig, 2016). **Alan Turing** wordt bijvoorbeeld als één van de grondleggers genoemd, met zijn Turing-machine die alle mogelijke berekeningen kan uitvoeren.<sup>4</sup> Voordat ik Turings gedachten belicht, terug naar het moment waarop AI als dusdanig benoemd werd: het Summer Research Project in 1956. De insteek hiervan was als volgt:<sup>5</sup>

‘Wij stellen voor dat er in de zomer van 1956 een 2-maanden durende studie van 10 personen naar kunstmatige intelligentie wordt uitgevoerd aan Dartmouth College in Hanover, New Hampshire. De studie moet worden uitgevoerd op basis van de veronderstelling dat elk aspect van leren of een ander kenmerk van intelligentie in principe zo nauwkeurig kan worden beschreven dat een machine het kan simuleren. Er zal een poging worden gedaan om te ontdekken hoe machines taal kunnen gebruiken, abstracties en concepten kunnen vormen, problemen kunnen oplossen die nu alleen voor mensen zijn weggelegd, en zichzelf kunnen verbeteren. Wij denken dat er een significante vooruitgang kan worden geboekt in een of meer van deze problemen als een zorgvuldig geselecteerde groep wetenschappers er een zomer lang samen aan werkt.’ (McCarthy e.a., 2006)

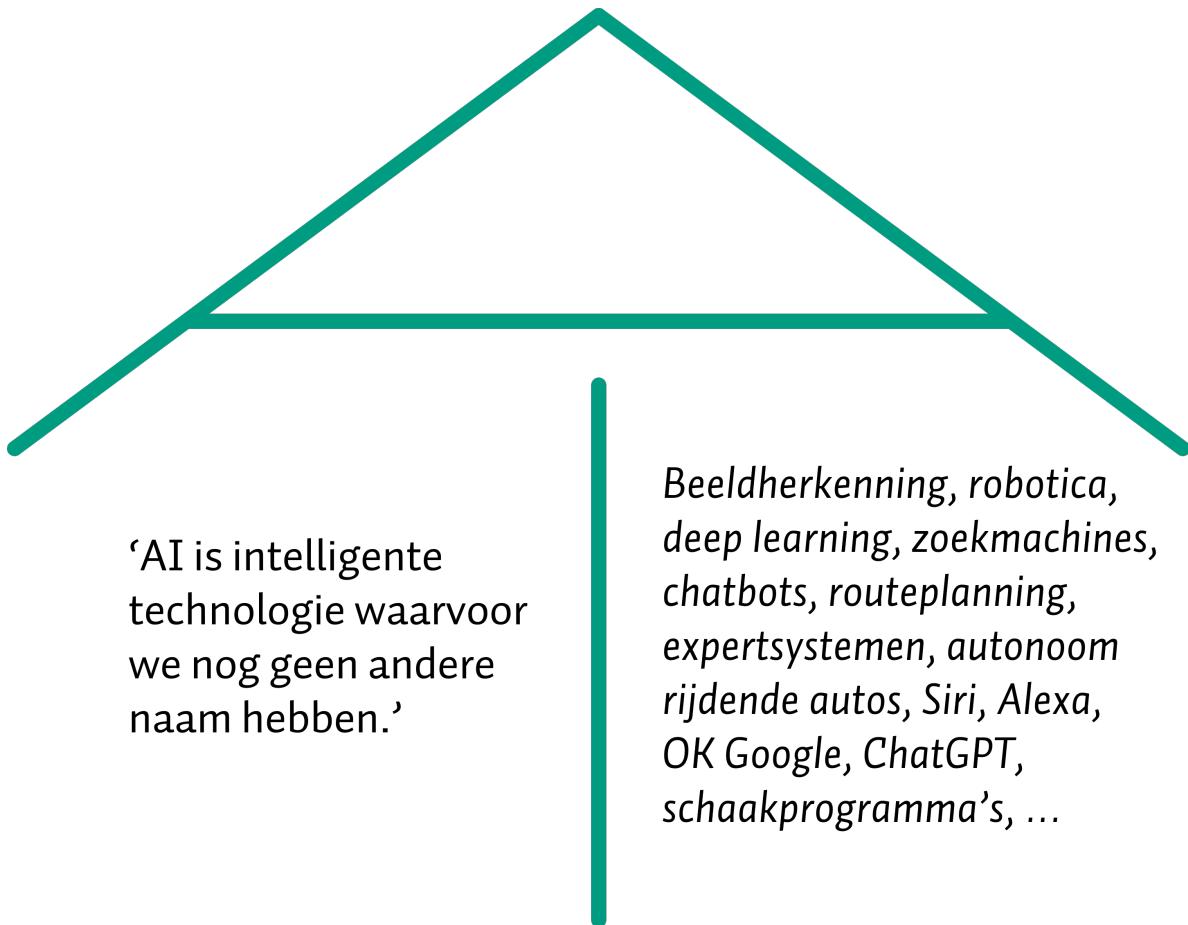
Aan ambitie geen gebrek, zo erkenden de betrokkenen ook in een terugblik hierop, vijftig jaar later (Moor, 2006). De kern van AI blijkt hieruit wel duidelijk, namelijk het nabootsen of simuleren van intelligentie. John McCarthy, één van de oprichters van de zomerworkshop, definieerde AI simpelweg als ‘de wetenschap en techniek van het maken van intelligente machines’. Dit is een brede, inclusieve definitie die wat mij betreft nog steeds het beste werkt om het hele vakgebied samen te vatten. Een andere blik op AI is dat het draait om *patroonherkenning* en daarmee ook het genereren van patronen. In de jaren na Dartmouth heeft AI namelijk vele ontwikkelingen doorgemaakt (Buchanan, 2005). Allerlei onderzoeksgebieden zijn sindsdien ontstaan, zoals robotica, expertsystemen, beeldherkenning, redeneersystemen en zelflerende systemen (Russell & Norvig, 2016). AI is daarom een parapluterm, wat ook af te zien valt aan de letters A en I.

---

<sup>3</sup>Met deze introductie doe ik veel onderzoek en innovaties binnen AI te kort (zoals genetische algoritmen (Mitchell, 1998), agent-based modellen (Macal, 2016) en Bayesiaanse netwerken (Heckerman, 1998)), maar anders wordt deze sectie te lang. Een uitspraak die mij is bijgebleven van mijn studiejaren is dat het lezen van het boek *Gödel, Escher, Bach* van Daniel Hofstadter in theorie een propedeusediploma in AI zou moeten opleveren (Hofstadter, 1999). Ook van mijn kant is dit een aanrader, nog steeds.

<sup>4</sup>Turing leidde een bijzonder maar ook tragisch leven. Hij legde het grondwerk voor het vakgebied van theoretische informatica en hielp de Britten de Enigma-code kraken tijdens de Tweede Wereldoorlog. Vanwege zijn homoseksuele geaardheid werd hij berecht en stierf op 41-jarige leeftijd door cyanide-vergiftiging, al dan niet zelf toegebracht (Copeland, 2024a).

<sup>5</sup>Vertaald uit het Engels.



Figuur 1.1: AI als parapluterm

Een cynische definitie van AI zou ook kunnen zijn: intelligente technologie waar we nu nog geen andere naam voor hebben. Het is namelijk zo dat sommige technieken doorontwikkeld worden en een eigen naam krijgen. Soms wordt dan gezegd dat deze techniek AI is, waarbij het gevaar bestaat dat er geen andere vormen van AI zouden bestaan. Een recent voorbeeld is *ChatGPT*, dat gebruik maakt van AI-technologie, maar slechts één vorm hiervan is.<sup>6</sup> Andere voorbeelden zijn navigatiesystemen (routeplanning m.b.v. Google Maps, TomTom), spraakherkenning (spraak naar tekst in schrijfprogramma’s, Apple Siri), zoekmachines (Bing, DuckDuckGo), zelfrijdende auto’s, en virtuele karakters in computergames. Tig vormen van AI, en veel sub-vakgebieden waar de meeste mensen zich niet bewust van zijn.

Ik grijp de kans niet aan om een complete uitleg te geven van alle mogelijke vormen van AI,

---

<sup>6</sup>Taalmodellen op basis van *transformers* die weer voortkomen uit *deep learning*, meer hierover in het volgende hoofdstuk.

maar geef wel een uitleg van de principes.<sup>7</sup> Een belangrijke opmerking vooraf: mensen maken AI-technologie en deze is in de basis te begrijpen. **AI is geen magie.** AI is logisch opgebouwd en maakt gebruik van *algoritmes*: instructies die een computer kan uitvoeren. Deze algoritmes kunnen zo simpel zijn als ‘ALS dit, DAN dat’: ‘ALS de temperatuur in de woonkamer onder de 18,5 graden Celsius komt, DAN moet de verwarming actief worden.’ Ook kunnen algoritmes geprogrammeerd worden als complete grammatica’s met allerhande regels die getest kunnen worden. Op zichzelf kunnen zulke regels met moeite intelligent genoemd worden, maar een verzameling hiervan kan complex en intelligent(er) gedrag vertonen. Zo zijn er algoritmes om op basis van een kaart met afstanden tussen steden efficiënt de kortste route te vinden van A naar B. Ook zijn er algoritmes die woorden in zinnen herkennen en op basis daarvan een indicatie geven of de tekst positief of negatief geladen is.

Opvallend hierbij is dat bij deze aanpak de intelligentie vastligt in het algoritme. Er wordt geen nieuwe kennis toegevoegd, alle gevallen moeten beschreven zijn voordat het algoritme gestart wordt. Het is regelgebaseerd, volgens een bepaalde logica. Er vindt dus geen ‘leren’ plaats. Dit is een belangrijk onderscheid in AI, wat door de opmars van beschikbare data vaak vergeten wordt: het onderscheid tussen lerende systemen en statische systemen.<sup>8</sup> Dit verdeelt AI grofweg in twee kampen: **symbolische AI** en **connectionisme**. Beiden gaan uit van hoe menselijke intelligentie werkt. Symbolische AI gaat uit van een basis van begrijpelijke regels en vastgelegde logica.<sup>9</sup> Het bekijkt intelligentie dus op een hoog niveau van abstractie dat door mensen leesbaar is. Dit kan effectief zijn, maar heeft als nadeel bijvoorbeeld dat het schrijven van de regels veel werk kan kosten in complexe situaties.<sup>10</sup>

---

<sup>7</sup>Een voorzetje voor het nawoord, voor diegenen die meer te weten willen komen over de details van AI: de [Nationale AI-cursus](#) met een uitstekende uitleg, en de diepgaandere [AI-trainingssessies van Saxion](#) die we vanuit ons lectoraat Ambient Intelligence ontwikkeld hebben.

<sup>8</sup>Daarbij is ook nog het verschil te maken tussen systemen die eenmalig leren en als ze draaien niet verder leren, en de systemen die steeds nieuwe data verzamelen en op basis daarvan nieuwe dingen leren. Dit laatste wordt vaak verwacht van AI, maar is juist vaak *niet* het geval. De volgende hoofdstukken lichten dat in meer detail toe.

<sup>9</sup>Dit wordt liefkozend **GOFAI** genoemd: Good Old-Fashioned AI.

<sup>10</sup>Wat moet wel en niet beschreven worden (Shanahan, 2016)? Kan intelligentie los van een lichaam gezien worden (Chrisley, 2003)?

## Symbolische AI

```

moeder_van_kind(alice, suzan).
vader_van_kind(tom, suzan).
vader_van_kind(tom, inge).
vader_van_kind(mike, tom).

broer_of_zus(X, Y) :- ouder_van_kind(Z, X), ouder_van_kind(Z, Y).

ouder_van_kind(X, Y) :- vader_van_kind(X, Y).
ouder_van_kind(X, Y) :- moeder_van_kind(X, Y).

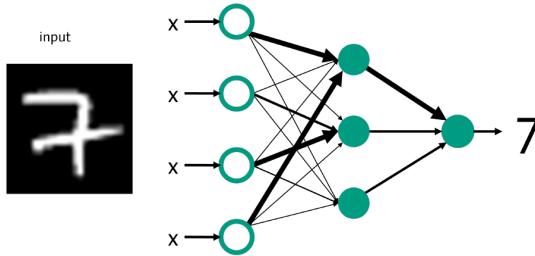
voorouder(X, Y) :- ouder_van_kind(X, Y).
voorouder(X, Y) :- ouder_van_kind(X, Z), voorouder(Z, Y).

?- broer_of_zus(suzan, inge).
Yes

?- voorouder(inge, alice).
No

```

## Connectionisme

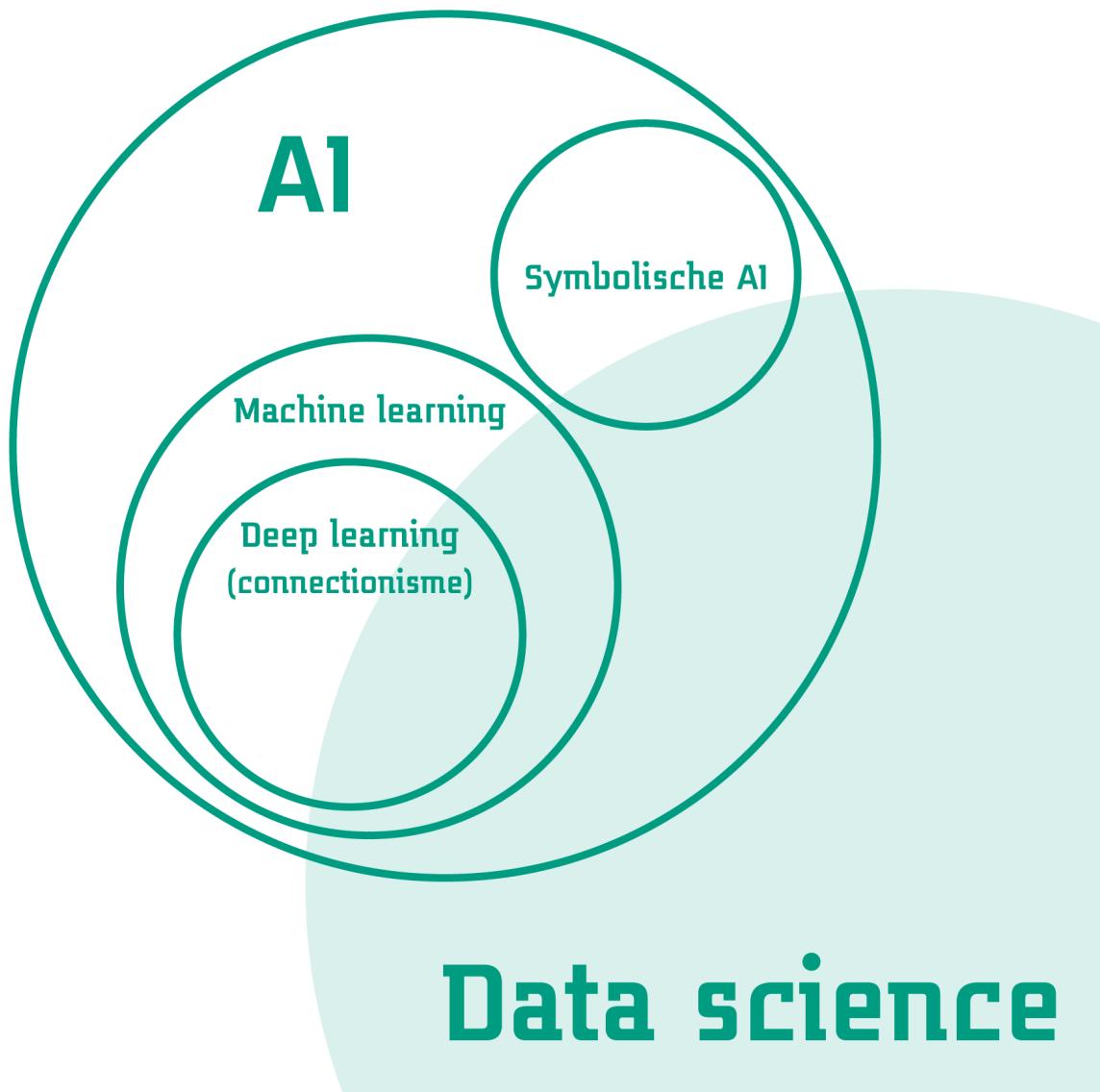


Figuur 1.2: Voorbeelden van symbolische AI en connectionisme. Links is uitgeschreven hoe met de programmeertaal Prolog familieverbanden achterhaald kunnen worden. Rechts staat een schematische weergave van een neuraal netwerk dat een handgeschreven cijfer herkent.

Connectionisme gaat uit van een lager niveau van abstractie van intelligentie, door netwerken te beschrijven die lijken op de structuren in hersenen. Hierbij is het idee dat er (kunstmatige) neuronen met elkaar in verbinding staan die elkaar kunnen activeren op basis van signalen die ze doorgeven. Door verschillende signalen met elkaar te combineren, kunnen vervolgens complexe functies worden beschreven. Dit werkt door de neuronen in een netwerk met elkaar te verbinden en te trainen met data. Dit is **deep learning**, een vorm van **machine learning**: het een computer aanleren van een bepaalde taak zonder deze van tevoren geprogrammeerd te hebben. Om het schoolvoorbeeld te geven van de MNIST-dataset: stel, we willen een netwerk trainen om handgeschreven cijfers te herkennen.<sup>11</sup> We geven het netwerk een plaatje van een cijfer als input en vragen het om als output te geven wat de waarschijnlijkheden zijn dat het een bepaald getal is. De verbindingen in het netwerk bepalen door de activiteiten van de signalen in het netwerk wat deze waarschijnlijkheden zijn. Klopt de output van het netwerk niet? Dan 'leert' het netwerk door de activiteiten aan te passen dankzij achterliggende wiskundige formules die geprogrammeerd zijn. Op den duur wordt dit netwerk steeds beter naarmate het meer voorbeelden van de cijfers te zien krijgt.<sup>12</sup> Bij deep learning wordt dus in het begin enkel de vorm van het netwerk beschreven en traint het zichzelf op basis van data. De intelligentie komt voort uit het leerproces dat plaatsvindt.

<sup>11</sup>Zie [de originele dataset](#) (LeCun, 1998).

<sup>12</sup>Dit is *supervised learning*, waarbij er ook *unsupervised learning* bestaat, met andere doelen, bijvoorbeeld om onbekende data te analyseren (Russell & Norvig, 2016). Gebruikelijk is dat er duizenden voorbeelden getoond moeten worden aan netwerk voordat het goed werkt.



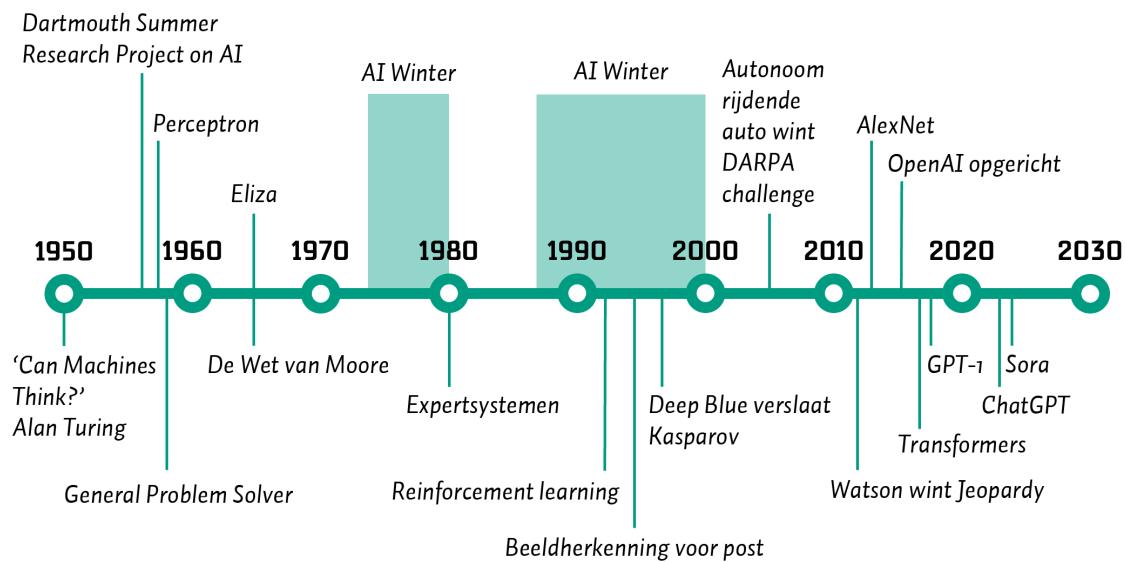
Figuur 1.3: De verhoudingen tussen AI, Symbolische AI, Connectionisme, machine learning en deep learning, en data science

Kort gezegd kan het vakgebied van AI met deze stromingen als volgt bekijken worden. Er zijn twee grote stromingen binnen AI, waarbij de symbolische AI uitgaat van het van tevoren uit-denken van patronen en het connectionisme deze patronen laat ontdekken. Het connectionisme heeft in de vorm van deep learning de laatste tien jaar veel aandacht heeft gekregen, dankzij computers die al het rekenwerk aan kunnen. Data science maakt gebruik van technieken uit

deze wereld om op basis van data nieuwe inzichten te verwerven. Dat kan door modellen te trainen met connectionisme en via regelgebaseerde logica, gecombineerd met technieken uit onder andere statistiek en informatica. Maar daarvoor waren er ook al tal van AI-technieken die zorgden voor nieuwe inzichten en toepassingen. Deze belicht ik in de volgende sectie.

## 1.2 Doorbraken in AI

In de afgelopen 80 jaar is er veel gebeurd in het vakgebied AI. Men is van ver gekomen, en tegenwoordig draaien vele systemen op AI-gebaseerde technologie, of we ons daarvan nu bewust zijn of niet. In vogelvlucht noem ik hier een aantal hoogte- en dieptepunten voor een historisch perspectief. Voor een volledige(re) geschiedenis van AI zijn er tal van werken beschikbaar.<sup>13</sup>



Figuur 1.4: Tijdlijn van AI in de afgelopen 80 jaar

Alan Turing verdient de eer om benoemd te worden met zijn doorbraak in denkwerk in 1950:<sup>14</sup>

Ik stel voor om de vraag te overwegen: 'Kunnen machines denken?' (Turing, 1950)

<sup>13</sup>De Encyclopedia Britannica legt het bondig uit, met tal van verwijzingen (Copeland, 2024b), maar ook de Wikipedia is hiervoor een goed startpunt (Wikipedia contributors, 2024a, 2024d). De website [AI Navigator](#) laat visueel een hele reeks doorbraken zien met meer detail.

<sup>14</sup>Vertaald uit het Engels.

Dit was de aanzet voor het gedachte-experiment de **Turing-test**.<sup>15</sup> Hierin moet een ondervrager achterhalen in welke van twee kamers zich een mens of een machine bevindt. De ondervrager kan alleen vragen stellen en krijgt daarop antwoord vanuit beide kamers. De machine zal zich proberen voor te doen als een mens, terwijl de mens de ondervrager zal proberen te overtuigen dat die een mens is. Uit deze simpele voorzet ontstond een heel plan met ideeën om machines menselijke intelligentie te laten imiteren en kwamen er vele tegen-argumenten naar voren, al helemaal met de taalmodellen van vandaag de dag (Pinar Saygin e.a., 2000; Sejnowski, 2023).

Overigens was **Ada Lovelace** ook een persoon die eer verdient.<sup>16</sup> Zij zag al in 1843 in dat machines niet alleen voor rekenwerk gebruikt hoefden te worden, maar ook voor andere, menselijke problemen (Fuegi & Francis, 2003). Maar, zoals gezegd, pas in de jaren '50 van de 20e eeuw zou er echt werk gemaakt worden van AI als vakgebied, dankzij het Dartmouth Summer Research Project (Moor, 2006). Dit project was tamelijk ambitieus, maar legde enkel het grondwerk voor tal van onderzoek dat daarna zou moeten gebeuren. In 1958 werd de basis gelegd voor deep learning Rosenblatt (1958) met een klein kunstmatig neuraal netwerk genaamd **perceptron** om de werking van zulke netwerken aan te tonen. In 1959 werd de **General Problem Solver** ontwikkeld die met behulp van intelligente zoekalgoritmes bepaalde problemen kon oplossen (Newell e.a., 1959). In 1965 werd **Eliza** ontwikkeld, de eerste chatbot, die een psychotherapeut simuleerde door ingetypte zinnen te lezen met geprogrammeerde patroonherkennung (Weizenbaum, 1966). In de latere jaren '60 werden verbeteringen voor deep learning ontwikkeld (Amari, 1967) en tegelijkertijd werden tekortkomingen van zulke netwerken benoemd. Begin jaren '70 heerste er nog enthousiasme, zoals Minsky gezegd zou hebben:<sup>17</sup>

‘Binnen drie tot acht jaar hebben we een machine met generieke intelligentie van een gemiddelde mens. Ik bedoel een machine die Shakespeare kan lezen, een auto kan smeren, kantoorpolitiek kan spelen, een mop kan vertellen, een gevecht kan aangaan. Op dat moment zal de machine zichzelf gaan ontwikkelen met fantastische snelheid. Binnen een paar maanden zal het een genie zijn en een paar maanden daarna zullen diens krachten onvoorstelbaar zijn.’ (Darrach, 1970)

Maar niet heel veel later deed een nieuw begrip diens intrede: de eerste **AI Winter** begon. Alle successen tot dan toe kwamen niet in de buurt bij het beeld dat geschatst werd. De intelligentie van de ontwikkelde systemen was beperkt en er waren duidelijke limitaties. Met name dankzij verloren vertrouwen in deze technologie door wat nu de Defense Advanced Research Projects Agency (DARPA, van de Verenigde Staten) heet, hield de financiering voor het onderzoek grotendeels op. Deze AI Winter hield aan tot 1980, waarin **expertsystemen** commercieel inzetbaar werden (Buchanan & Smith, 1988). Hieraan werd al gewerkt in de jaren '60: het zijn systemen waarin kennis in regelvorm werd opgeslagen, waarna een efficiënt algoritme kon achterhalen welke kennis van belang was om een bepaald probleem op te lossen en daar de stappen voor te schetsen. In de jaren '80 werden door grote bedrijven miljarden euro's gespendeerd om zulke systemen te ontwikkelen en hun proces te optimaliseren. Veel van

<sup>15</sup> Destijds de *Imitation Game* genoemd.

<sup>16</sup> Niet alleen omdat zij benoemd wordt de eerste computerprogrammeur geweest te zijn (Fuegi & Francis, 2003).

<sup>17</sup> Vertaald uit het Engels.

deze systemen hadden echter hun mankementen, want ze waren bovenal duur om te onderhouden vanwege alle regels (en uitzonderingen daarop) die handgeschreven moesten worden. Dit zorgde voor een tweede AI Winter, die zou aanhouden van 1990 tot het begin van het nieuwe millennium. Intussen zouden er wel nieuwe doorbraken zijn, zoals **reinforcement learning** in het begin van de jaren '90. Deze techniek zorgt ervoor dat een zogenaamde 'agent' diens omgeving kan waarnemen en kan plannen om acties uit te voeren die dichter bij diens doel brengen (Kaelbling e.a., 1996). Dit gebeurt door middel van het geven aan positieve of negatieve beloningen voor acties in die omgeving. In 1996 en 1997 speelde de computer **Deep Blue** tegen Gary Kasparov, wereldkampioen schaak, en versloeg deze in hun rematch (Campbell e.a., 2002). Dit was een mijlpaal doordat deze gebeurtenis liet zien dat een computer beter was dan een mens in een activiteit waarvoor men heel intelligent zou moeten zijn. Deep Blue's algoritmes baseerden zich op 700.000 schaakwedstrijden van grootmeesters in schaak en een verzameling van specifieke openingszetten, waarmee de machine 200 miljoen posities per seconde kon evalueren.

De volgende ontwikkeling die zorgde voor de huidige AI Lente waren neurale netwerken voor deep learning. De ideeën hiervoor bestonden al in de jaren '60, maar uitwerkingen hiervan waren gelimiteerd door rekenkracht. Vanaf het begin van de 21e eeuw werd rekenkracht beschikbaar die ervoor zorgde dat de netwerken konden groeien van tientallen neuronen in een paar lagen tot miljoenen neuronen in tientallen lagen.<sup>18</sup> Vanaf ongeveer 2010 werden toepassingen in beeldherkenning hiermee enorm verbeterd. De mijlpaal hier was **AlexNet**, een deeplearning-netwerk dat in de beroemde ImageNet Challenge voor beeldherkenning als enige neurale netwerk meedeed, en won met 10 procentpunt verschil met de tweede plaats (Krizhevsky e.a., 2012). Vanaf daar was het hek van de dam en begonnen meer en meer teams en onderzoekers zich met deep learning te bezighouden, net als bedrijven waarvoor commerciële inzet van zulke technieken nu haalbaar ging worden.<sup>19</sup> Deep learning leent zich niet alleen voor beelden goed, maar ook voor tekst, geluid, tijdsignalen, en vele andere soorten data (LeCun e.a., 2015). In het volgende hoofdstuk wordt duidelijk wat hier voor haken en ogen aan zitten, maar de mogelijkheden van deze technieken zijn duidelijk legio.

Wat we de afgelopen jaren hebben meegemaakt is de opkomst van **generative AI** zoals grote taalmodellen of **Large Language Models** (LLMs) (Zhao e.a., 2023), met als bekendste voorbeeld **ChatGPT**, en **text-to-image**-modellen die tekst kunnen omzetten in beelden, zoals **DALL-E** en **Midjourney** (Zhang e.a., 2023). LLMs zijn gebaseerd op de **transformer**-architectuur die ervoor zorgt dat tekst efficiënt gerepresenteerd kan worden (Vaswani e.a., 2017). Op basis van grote hoeveelheden tekst kunnen dan modellen worden gemaakt die verbanden in de tekst vastleggen en op die manier een 'begrip' hebben van die tekst. De text-to-image-modellen werken dan weer op basis van modellen die getraind zijn om ruis uit plaatjes te verwijderen, waarbij er een tekst bij het plaatje hoort die beschrijft wat er op het plaatje staat. Dit idee is in het extreme doorgedreven, door op ruizige plaatjes een model te trainen en

<sup>18</sup>Dit kon dankzij zogenaamde **Graphical Processing Units**, die videokaarten in computers die parallelle berekeningen doen voor beelden en uitstekend geschikt zijn voor parallelle berekeningen in neurale netwerken.

<sup>19</sup>Dat wil zeggen, voor complexere taken. Beeldherkenning voor bijvoorbeeld adresherkenning op brieven bestaat al sinds de jaren '90 (LeCun e.a., 1995).

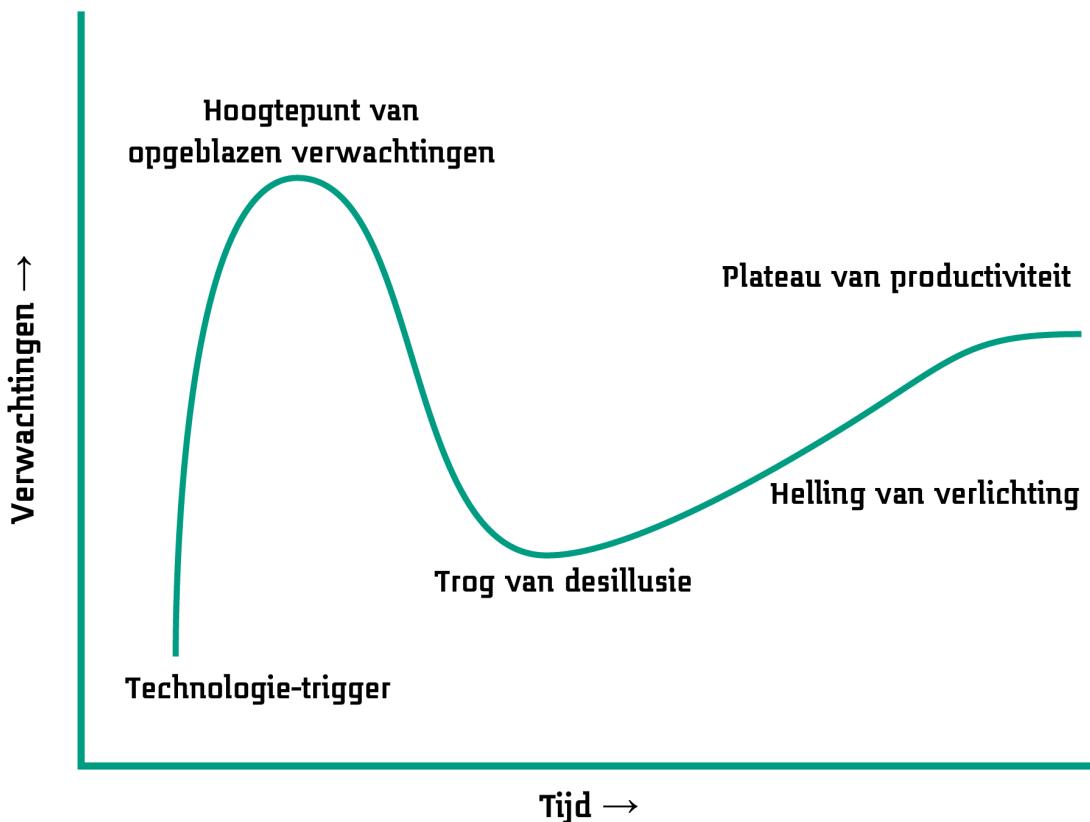
vervolgens een plaatje met alleen ruis en een beschrijving te geven aan zo'n model (Ling Yang e.a., 2023). De impact van deze twee technologieën is gebleken op veel vlakken en behandel ik in het volgende hoofdstuk. Duidelijk mag zijn dat AI nu al een lange historie heeft en er nog veel te gebeuren staat.

### 1.3 Discussies over AI

Wat de doorbraken van AI ons in ieder geval laten zien, is dat er vaak een duidelijke verwachting is van wat AI wel en niet kan. Dat kan heel fundamenteel en theoretisch zijn, zoals de vraag of AI bewustzijn kan hebben. Het kan ook heel praktisch zijn, namelijk of AI een bepaalde taak kan uitvoeren. Wat in beide gevallen vaststaat, is dat verwachtingen vaak hooggespannen zijn. Dat blijkt uit de AI winters die eerder plaatsvonden: men verwachtte veel van AI, en ook op korte termijn. Dit is een typisch gegeven voor nieuwe technologie, wat beschreven kan worden met de *hype cycle*, bekend dankzij onderzoeks- en adviesbureau Gartner.<sup>20</sup> Het principe hiervan is het beste uit te leggen met deze figuur:

---

<sup>20</sup>Zie [de uitleg](#) van deze methode.



Figuur 1.5: De hype-cycle van nieuwe technologie

Voor nieuwe technologieën kunnen vijf fasen beschreven worden:<sup>21</sup>

1. Technologie-trigger: de eerste verhalen over een technologie doen de ronde.
2. Hoogtepunt van opgeblazen verwachtingen: men raakt meer dan enthousiast, de hype is op diens hoogtepunt.
3. Trog van desillusie: de technologie lost de verwachtingen niet in, wat blijkt uit falende implementaties en experimenten.
4. Helling van verlichting: bepaalde implementaties werken toch en worden doorontwikkeld.
5. Plateau van productiviteit: adoptie van de technologie door het grote publiek.

Gartner probeert met invullingen van deze hype cycle inzicht te bieden in welke technologieën op welke plek in de curve zitten. Zo ook voor AI, en alle afzonderlijke technieken in dit veld. Volgens recente voorspellingen van Gartner zit bijvoorbeeld *Generative AI* (de taalmodellen

---

<sup>21</sup>Vertaald uit het Engels.

en plaatjesgeneratoren) al voorbij het hoogtepunte van fase twee (Gartner, Inc., 21 August 2024). Een inzicht dat belangrijk is bij dergelijke voorspellingen komt van de hand van Roy Amara, onderzoeker bij verschillende Amerikaanse universiteiten en futuroloog.<sup>22</sup>

**De Wet van Amara:** We hebben de neiging om het effect van een technologie op de korte termijn te overschatten en het effect op de lange termijn te onderschatten.

Het plateau van productiviteit kan uiteindelijk hoger liggen dan we verwachten als we gedessillusioneerd zijn. En wellicht is de impact van een technologie anders dan verwacht, wat maar al te vaak gebeurt.

Als we kijken naar wat we uiteindelijk kunnen verwachten bij AI, is de hamvraag hoe intelligent deze technologie werkelijk kan worden. Is het theoretisch en ook praktisch mogelijk dat AI slimmer wordt dan mensen? Deze fundamentele vraag splitst het AI-veld ook op in twee grote kampen: zij die de **strong AI** aanhangen versus de **weak AI** (B. Liu, 2021). De laatste betekent zoveel als AI die toegespitst is op één toepassing, zoals het herkennen van plaatjes. Dit wordt ook wel *narrow AI* genoemd. De lading hiervan is dat AI goed kan zijn in afzonderlijke taken en hiermee mensen kan ondersteunen, maar geen generieke taken kan oppakken. Sommigen zijn ervan overtuigd dat dit voor AI nooit mogelijk zal zijn.<sup>23</sup> Aan de andere kant van het spectrum zit de strong AI, waaronder AI valt die wel degelijk generieke taken kan oppakken, dan wel bewustzijn kan hebben of superintelligent kan zijn (Bostrom, 2015). Eén van de centrale aannames hierbij is dat, als we de hersenen van een mens perfect kunnen maken in termen van input, output, en het complete neurale netwerk daartussen, we in staat zijn menselijke intelligentie te maken (Cole, 2004). Kort door de bocht gezegd is het idee dat hersenen een computer zijn. Functioneel gezien zou er dan namelijk geen verschil moeten zijn tussen een ‘normale’ mens en een nagemaakte mens die draait op andere *hardware*.

Dit zijn grote discussies en zeker geen uitdagingen die ik in deze rede of in ons lectoraat als geheel wil aanpakken.<sup>24</sup> Maar ze zetten wel de toon voor de verwachtingen over AI als technologie en einddoelen die we wel of niet willen najagen. In Hoofdstuk 2 richt ik me op praktische uitdagingen voor de komende jaren die we tegemoet gaan, en waar we ons volgens mij op zouden moeten richten.

---

<sup>22</sup>Er zijn weinig goed toegankelijke bronnen te vinden over Amara; de beste plek om meer te weten te komen is via Wikipedia (Wikipedia contributors, 2024c).

<sup>23</sup>Ik laat de discussies hierover voor nu buiten beschouwing, maar in essentie gaat hier over de mogelijkheden van AI begrip en bewustzijn kan hebben. Zie bijvoorbeeld het gedachte-experiment van John Searle over de *Chinees Kammer* (Cole, 2004), iets waar bijvoorbeeld Daniel Dennett het faliekant mee oneens is (Dennett, 2017).

<sup>24</sup>Ik zal er wel op reflecteren in Hoofdstuk 3.

## 2 Uitdagingen: Het belang van wederzijds begrip

### Samenvatting

Snappen wij wat AI kan en hoe dit gebeurt? En snapt AI andersom ons als mensen ook? Met de huidige hype van AI worden mogelijkheden van deze technologie duidelijk, maar er zijn ook schaduwzijdes. Ons lectoraat Ambient Intelligence zet zich in voor wederzijds begrip. We maken het proces achter AI transparant en zorgen dat AI aansluit bij de wensen van mensen.

In de introductie van deze rede beschrijf ik wat AI als vakgebied definieert en voortstuwt: het maken en onderzoeken van intelligente machines. Dit is nog steeds de overkoepelende uitdaging. Met de recente AI Lente van **Generative AI** zijn sommige onderdelen van deze uitdaging duidelijker geworden en anderen weer ondergesneeuwd. In dit hoofdstuk benoem ik deze uitdagingen door eerst een aantal casussen te benoemen, waarna ik de specifieke uitdagingen concreet maak met voorbeelden van projecten van ons lectoraat Ambient Intelligence.

### 2.1 De huidige hype en schaduwzijdes van AI

Een rechter die ChatGPT gebruikt om de uitspraak te onderbouwen.<sup>1</sup> Chatbots die gebruikt worden om desinformatie te verspreiden in de aanloop van de Amerikaanse verkiezingen in November 2024.<sup>2</sup> AI-modellen die op basis van auteursrechtelijk beschermd materiaal worden getraind.<sup>3</sup> *Big Tech* investeren meer dan 100 miljard dollar in de eerste helft van 2024 voor de infrastructuur die nodig is voor AI.<sup>4</sup> De eerste AI-reclamespot wordt gepubliceerd, gemaakt voor een speelgoedwinkelketen.<sup>5</sup> Dit is een aantal voorbeelden van nieuwsberichten die deze zomer langskwamen. Eerder zagen we al voorbeelden van een AI-model dat een kunstwedstrijd won<sup>6</sup> en een Google-onderzoeker die beweerde dat een chatbot bewustzijn had.<sup>7</sup> Geen dag gaat voorbij, of er is weer niewtje over AI dat al dan niet opgeblazen wordt.

---

<sup>1</sup>Zie [Tweakers](#).

<sup>2</sup>Zie [AI Forensics](#) en [EuroNews](#).

<sup>3</sup>Zie [NOS](#).

<sup>4</sup>Waaronder Microsoft, Amazon, Apple, Meta, zie de [Financial Times](#).

<sup>5</sup>Zie het [Financieel Dagblad](#) en [ToysRUs](#).

<sup>6</sup>Zie de [New York Times](#).

<sup>7</sup>Zie de [Washington Post](#).

Dankzij relatief makkelijk toegankelijke websites en applicaties kunnen plots veel mensen gebruikmaken van AI om teksten en beelden te genereren. Er zijn tal van afgeleide diensten ontstaan die op basis van AI-modellen, van chatbots die gesprekken tussen historische figuren simuleren tot applicaties die papieren schetsen omzetten naar werkende websites.<sup>8</sup> Terugkijkend naar de *hype cycle* zou je wellicht verwachten dat we nu massa-adoptie hebben van deze vorm van AI. Dat is tot op zekere hoogte ook zo, maar gaandeweg worden de gebreken en schaduwzijdes helderder. Voor bepaalde toepassingen werkt AI enorm goed, maar juist met deze globale aandacht wordt er (gelukkig!) nauwlettender gekeken naar de ontwikkeling van deze technologie.

Waar deze AI-modellen namelijk behoeft aan hebben is **data**, veel data. In de introductie benoem ik dat machine learning baat heeft bij het zien van vele voorbeelden, bijvoorbeeld duizend plaatjes om een object goed te kunnen herkennen. Voor LLM's als GPT geldt hetzelfde: deze worden getraind op veel voorbeelden van tekst. Hieruit worden verbanden gehaald met woorden die elkaar met enige waarschijnlijkheid opvolgen, en dan niet met een paar woorden achter elkaar, maar met reeksen van duizenden woorden. Vervolgens kan een dergelijk model voorspellen welk woord waarschijnlijk een bepaalde reeks moet opvolgen. Hierin zoekt het het patroon dat past bij de context: wat zijn de stijl van de tekst en de woorden die vaak gebruikt worden? De hoeveelheid data die bijvoorbeeld gebruikt is om het GPT-4-model te trainen, wordt geschat op 10 biljoen woorden of 300 terabyte (Schreiner, 11 July 2023).<sup>9</sup> Dit gebeurt op basis van grote datasets die openbaar te verkrijgen zijn.<sup>10</sup>

Hier dient zich een eerste schaduwzijde aan: mag een dergelijk model op basis van deze data getraind worden? Hoe zit het met auteursrecht? Media als de New York Times spanden rechtszaken hierom aan, omdat miljoenen van hun artikelen onrechtmatig gebruikt zouden zijn voor het trainen van chatbots.<sup>11</sup> Net zoals voor tekst geldt dat ook voor beeld en alle data die gebruikt wordt om zulke modellen te trainen, zoals voor Midjourney en Stable Diffusion.<sup>12</sup>

Wat ten tweede vaak vergeten wordt, is dat met zulke hoeveelheden data ook veel mensenwerk gepaard gaat. Een model met zulke complexiteit wordt niet ‘even’ getraind. Ten eerste moet bekijken worden of de data die gebruikt worden van goede kwaliteit zijn. Dus veel van deze data wordt bekijken door mensen. Dit is uitdagend en vaak ook repetitief werk, wat ervoor zorgt dat zulke klussen uitbesteed worden. Een onderzoek van het tijdschrift Time liet bijvoorbeeld al begin 2023 zien dat dit werk via bepaalde bureaus naar landen ging waar het loon laag ligt, zoals Kenia.<sup>13</sup> Hierin ging het om teksten die bekijken en gecensureerd moesten worden, om-

<sup>8</sup>Future Tools heeft een lijst van bijna 3.000 diensten die gebaseerd zijn op AI.

<sup>9</sup>OpenAI werd opgericht als non-profit organisatie om AI ten bate van mensen te maken, maar inmiddels is er ook een commerciële tak die diensten als ChatGPT uitbaat. Niet alles wat ontwikkeld wordt is openbaar en reproduceerbaar, dankzij ‘zowel het competitieve landschap en de veiligheidsimplicaties van grote modellen als GPT-4’ (vertaald uit het Engels) (OpenAI, 2023). Met wat cynisme valt OpenAI in mijn ogen dan inmiddels ook *ClosedAI* te noemen.

<sup>10</sup>Zoals CommonCrawl met 2,3 miljard websites in augustus 2024 en RefinedWeb met ongeveer 400 miljard woorden (Penedo e.a., 2023).

<sup>11</sup>Zie de [New York Times](#).

<sup>12</sup>Zie [AP News](#).

<sup>13</sup>Zie [Time](#).

dat deze gingen over strafbare zaken zoals kindermishandeling, bestialiteit en moord. Ook voor onschuldigere taken worden mensen ingezet, om bijvoorbeeld te beoordelen of zoekresultaten in de Amazon-webwinkel goed genoeg zijn, of bepalen wat in politieke filmpjes wordt gezegd. Ook hiervoor worden AI-modellen gebruikt die suggesties doen, maar wel gevalideerd moeten worden door mensen.<sup>14</sup> De mensen die dit uitvoeren worden **klikwerkers** en in het Engels **ghostworkers** genoemd. Die laatste benaming is erg toepasselijk, want het is een vrij onzichtbaar gedeelte van het werk dat vrijwel geheel via internet wordt uitgevoerd. Een schatting uit 2023 van de Wereldbank over het aantal klikwerkers benoemt namelijk dat tussen de 150 en 430 miljoen mensen die dit (in deeltijd) uitvoeren.<sup>15</sup> Los van de humanitaire bedenkingen die hierbij mogelijk zijn, is ook af te vragen hoe efficiënt en goed dit werk gebeurt. Wellicht kost het meer moeite om de modellen te trainen dan zij uiteindelijk opleveren.]

Er zijn nog meer nadelen te benoemen van AI, zoals het energieverbruik dat van het trainen en uitvoeren van de modellen. Schattingen hiervan zijn lastig, maar bijvoorbeeld het genereren van een enkel plaatje met een AI-model kost net zoveel energie als het opladen van een smartphone, ongeveer 3 kilowattuur (Luccioni e.a., 2024). Dit werpt verdere uitdagingen op voor klimaatdoelstellingen.[Helemaal gecombineerd met datacenters, die ongeveer 1-1,5% van het globale energieverbruik verbruiken.<sup>16</sup> Daarnaast zijn de sociaal-maatschappelijke doelen waarvoor AI kan worden ingezet natuurlijk al lang en breed uitgemeten in fictie, maar ook in realiteit. Typisch voorbeeld is het verspreiden van desinformatie, propaganda, en andersoortig materiaal door kwaadwillenden (Bird e.a., 2023; Wach e.a., 2023).

Zo is de impact van AI plots groter dan wellicht gedacht. Gelukkig zijn er ook acties die hoop bieden. Net zoals dat AI kan worden ingezet voor doelen met minder goede intenties, kan het ook hiertegen worden ingezet. Er zijn onderzoeken naar modellen die desinformatie kunnen detecteren in plaats van creëren (Aïmeur e.a., 2023). Ook wordt er gewerkt aan efficiëntere AI-modellen, die een factor tien of honderd kleiner zijn en net zo goed presteren (Choudhary e.a., 2020). Hoewel klikwerkers op de manier van hun repetitieve taken een vorm van inkomen hebben, kunnen voor hen nog betere werkomstandigheden gecreëerd worden, of modellen ontwikkeld worden die deze taken ondersteunen (Barbudo e.a., 2023). En een initiatief dat ik verder graag benoem is **GPT-NL**, een project om een Nederlands open taalmodel te ontwikkelen dat zo open mogelijk is, ook om een eerlijker alternatief voor onder andere ChatGPT te zijn.<sup>17</sup>

Regulering van AI en alles wat met data gepaard gaat zie ik als noodzakelijk. Vanuit de Europese Unie zijn onlangs twee initiatieven hieromheen gelanceerd: de Data Act en de AI Act.<sup>18</sup> De Data Act stelt regels op voor het gebruik van en toegang tot data die in de EU wordt gemaakt. De AI Act gaat in op de ontwikkeling van AI-systeem en de verantwoordelijkheid

---

<sup>14</sup>Zie het [NRC](#).

<sup>15</sup>Zie [Worldbank](#).

<sup>16</sup>Zie [IEA](#).

<sup>17</sup>Zie [SURF](#).

<sup>18</sup>Zie de [AI Act](#) en de [Data Act](#).

hiervoor. Hierin wordt explicet gemaakt wat voor AI-gebaseerde technologie verboden is, bijvoorbeeld systemen die grote aantallen mensen kunnen manipuleren of emoties herkennen op werkplekken en onderwijsinstellingen. Daarnaast worden systemen geklassificeerd op basis van hun risico, waarbij onder andere goede documentatie en risicobeheersing verplicht zijn, net als voor *generieke* AI-systemen.

## 2.2 Van hype naar hoop: AI bij Ambient Intelligence

AI is een breed vakgebied en kan in alle sectoren en domeinen een toegevoegde waarde hebben. Als lectoraat zijn wij bij Ambient Intelligence *domein-agnostisch*: wij kijken vanuit de techniek, met de mens centraal, naar hoe we AI-technologie kunnen inzetten. Dat wil zeggen dat onze uitdagingen abstracter zijn dan specifieke beroepsvragen zoals het perfect identificeren van een ziekte of het voorspellen wanneer een machine kan uitvallen.

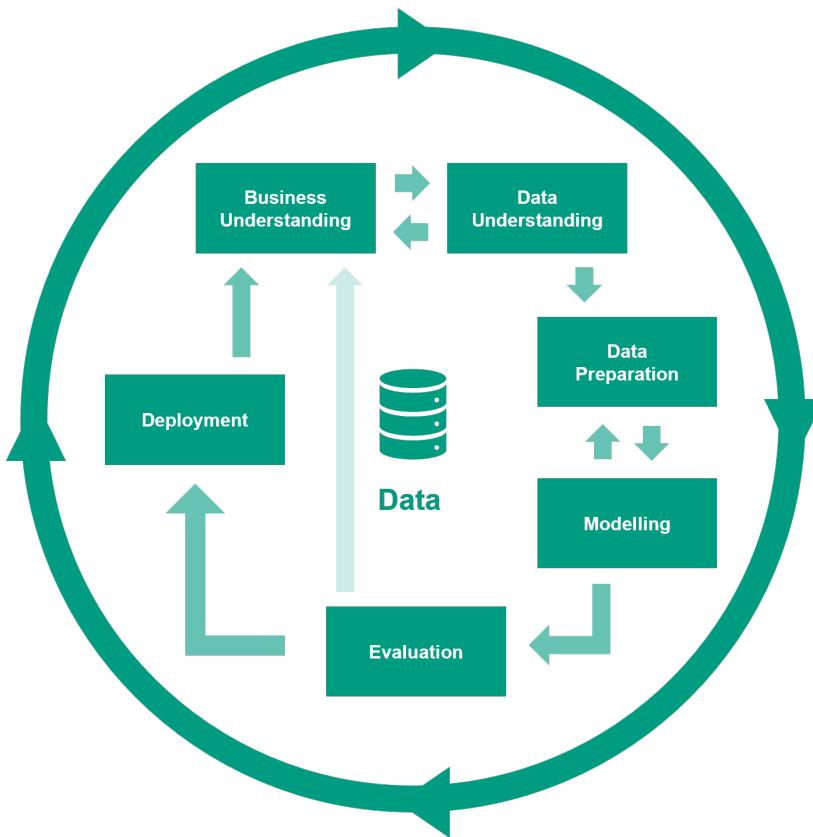
Centraal hierbij staat kennis van zaken als het gaat om AI. Er is meer begrip nodig van AI, wat erachter zit, en we hoe ermee kunnen omgaan. *Transparantie* is hierin een sleutelwoord: hoe werken zulke systemen, waarvan zijn ze afhankelijk, en wat of wie beïnvloeden ze? We moeten daarom toewerken naar een hoger niveau van **AI-geletterdheid**, in het verlengde van digitale geletterdheid.<sup>19</sup> De perceptie van wat AI is en zou kunnen of *moeten* zijn, dat is een belangrijk maatschappelijk vraagstuk. Ik zeg vaker: *AI is een tool, een stuk gereedschap*. Als je enkel een hamer hebt, lijkt alles op een spijker. Is AI wel het goede gereedschap voor hetgeen je wilt oplossen? En hoe gebruik je AI dan wel (of niet)?

### 2.2.1 Het proces centraal

Dit is dan ook wat we voor ogen hebben bij ons lectoraat Ambient Intelligence: we onderzoeken hoe ondernemers, het onderwijs, organisaties en overheden met AI uit de voeten kunnen. Eén van de pilaren voor ons datagedreven onderzoek draait dan ook om het proces. Onderliggend hieraan gaat het ook om omgang met data, dus streven we naar verbeterde **datageletterdheid**.<sup>20</sup> Hierbij gaat het onder andere om het begrijpen welke data nodig kan zijn, deze beheren, en de toegevoegde waarde van data kunnen inschatten. Data zijn uitdagend: meten is zeker niet altijd weten. Voorop staat altijd dat goede vragen gesteld moeten worden. Wat wil je met de data? Is de data beschikbaar, of moet deze nog opgehaald worden? Moet er iets gemeten worden en zo ja, hoe dan? Hoe betrouwbaar zijn die metingen? Weet je zeker dat alles gemeten is? Welke dingen meet je wellicht niet, en ben je je niet bewust van dat deze van invloed zijn op je vraag? Dit zijn vragen die vragen om een goede aanpak. Bij ons lectoraat hanteren we voor de analyse van data een proces dat **CRISP-DM** genoemd wordt: het CRoss-Industry Standard Process for Data Mining (Schröer e.a., 2021).

<sup>19</sup>Net zoals dat we onze algemene taalvaardigheid in Nederland zouden moeten verbeteren.

<sup>20</sup>Datageletterdheid is onderdeel van het raamwerk voor Digitale Competentie DigComp van de Europese Unie (Carretero e.a., 2017), zie [het framework](#).



Figuur 2.1: CRISP-DM

CRISP-DM is context-onafhankelijk: het kan in ieder domein gebruikt worden. Het is een methode om iteratief te ontdekken wat de waarde van data is en deze vervolgens te realiseren. Juist voor het toepassen van AI is een methode als CRISP-DM essentieel.<sup>21</sup> Dit proces bestaat uit zes fasen: Business Understanding, Data Understanding, Data preparation, Modelling, Evaluation en Deployment. In ieder hiervan worden verschillende taken opgepakt, bijvoorbeeld om te achterhalen wat het doel is in de ‘bedrijfscontext’. AI toepassen is waarschijnlijk nooit het doel, maar kan een middel zijn, bijvoorbeeld om te voorspellen wanneer fouten zich kunnen voordoen in een productieproces en deze vervolgens te voorkomen. Of om te analyseren in welke groep een revaliderende patiënt zich bevindt, om hierop een behandelplan aan te passen. In de fase Modelling kan AI toegepast worden als de context zich daarvoor leent: er is genoeg data van voldoende kwaliteit voor een bepaald algoritme. Voordat men op dat punt aankomt, moeten de data, als die er al zijn, verkend en opgeschoond worden (Chu e.a., 2016). En daar komen de vele vragen weer. Zijn alle variabelen die geanalyseerd moeten worden wel gerepresenteerd in de data? Wie of wat heeft de data verzameld? Wellicht wijzen de antwoorden op *bias*: een (onbewuste) voorkeur voor of beïnvloeding van de data (Ntoutsi e.a., 2020). Dit alles

<sup>21</sup>Er zijn natuurlijk nog andere, vergelijkbare methodes, zoals SEMMA en KDD (Shafique & Qaiser, 2014).

kan er al direct voor zorgen dat een AI-techniek mogelijk verkeerde uitkomsten gaat genereren. Zoals dat in jargon heet: **garbage in, garbage out**.

AI toepassen is geen kwestie van data aan een algoritme geven. Juist om dit idee duidelijk te krijgen hebben we als lectoraat gewerkt aan **AI-trainingssessies** als resultaat van het project **Data in Smart Industry**.<sup>22</sup> Deze zijn beschikbaar als open-access materiaal, met ondersteunende bronnen en oefeningen via  [deze website](#).<sup>23</sup> Deze uitdagingen van dataverwerking waren bijvoorbeeld het geval in projecten over het automatisch detecteren van objecten in spoorwegomgevingen, waarbij honderden gigabytes aan data beschikbaar waren uit laserscans.<sup>24</sup> Hoe nauwkeurig was deze data, en is dat wel nodig? Welke algoritmes zijn van toepassing? Hoe trainen we een efficiënt model? Wat zijn goede manieren om de data te annoteren voordat we kunnen trainen?<sup>25</sup>

Wat duidelijk mag zijn, is dat er veel stappen gezet moeten worden om tot een goede data-analyse en inzet van AI te komen. Tot op zekere hoogte kan dit geademtiseerd worden, want de eerste fasen vragen wel degelijk om mensenwerk. Deze automatisering is een belangrijk onderwerp in het onderzoek van het lectoraat Ambient Intelligence. Een typische manier hiervoor, die verweven is met CRISP-DM, is **DataOps** en specifiek **Machine Learning Operations** (MLOps) (Kreuzberger e.a., 2023).<sup>26</sup> MLOps is een set van praktijken die gericht zijn op het betrouwbaar en efficiënt in productie brengen en onderhouden van machine learning-modellen. Het combineert aspecten van machine learning, DevOps en data-engineering om het hele levenscyclusbeheer van machine learning te stroomlijnen. MLOps omvat onder andere de automatisering van alle stappen in het verwerken van data, versiebeheer hiervan, het geademtiseerd testen en implementeren van modellen, en inzicht geven in modelprestaties.

Naar de *best practices* op dit gebied voeren wij toegepast onderzoek uit in een project als **DataFlow**. Hierin draait het om het vaststellen van de huidige methodes, tools en technieken om efficiënter met data om te gaan, specifiek als het gaat om data-analyse. Er is namelijk een veelvoud aan verschillende raamwerken, softwarepakketten en *libraries*, zoals de figuur hieronder laat zien.<sup>27</sup> Feitelijk teveel om in één keer een afgewogen keuze te maken. Daarom trekken we in dit project op met bedrijven om tot een selectiemethode te komen die hen, en anderen, zal helpen bij het maken van die keuze.

---

<sup>22</sup>RRAAK-mkb [Data in Smart Industry](#), 2017 – 2020, met [Top-Up-project AIoTValley](#).

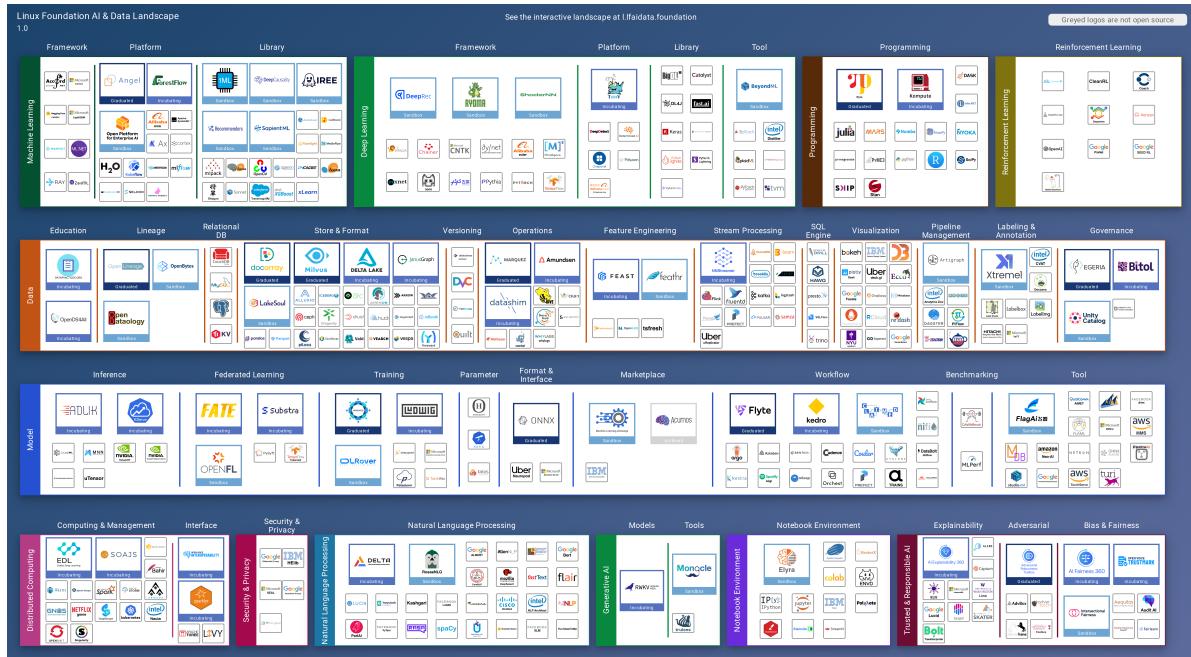
<sup>23</sup>Ook is CRISP-DM-documentatie beschikbaar via verschillende plekken: [Originele CRISP-DM documentatie](#) en [IBM-documentatie over CRISP-DM](#).

<sup>24</sup>TechForFuture [Digital Twinning voor Spoorontwerp](#), 2022 – 2024, en TechForFuture [Digitalisatie Bovenledingen en Draagconstructies](#), 2020 – 2021.

<sup>25</sup>En kunnen we het werk van klikwerkers verbeteren of vervangen?

<sup>26</sup>Ook samengevat onder de noemer **AutoML** (He e.a., 2021).

<sup>27</sup>Zie [LF AI & Data Foundation](#).



Figuur 2.2: Landschap van AI, met dank aan de [LF AI & Data Foundation](#)

En als je eenmaal met AI aan de slag kunt gaan, met grote hoeveelheden data, dan is er keuze te over voor alle verschillende technieken, zoals ik in het vorige hoofdstuk al benoemde. Afhankelijk van het type data (tekst, beeld, sensormetingen, etc.) en het type probleem (classificatie, regressie, tekstgeneratie, etc.), zijn bepaalde algoritmes wel of minder geschikt. Ook wat dit betreft zijn er mogelijkheden om deze selectie te automatiseren, of in ieder geval het zoeken naar de goede instellingen van een specifiek algoritme (Kerschke e.a., 2019; Li Yang & Shami, 2020).

Maar als we alles zo goed als mogelijk willen automatiseren, hoe houden we als mensen dan grip op het proces?

## 2.2.2 Het begrip centraal

Hierin zit de balans: wat automatiseren we in de machine, en wat laten we expliciet over aan mensen?

Als lectoraat pakken we niet alle uitdagingen van AI zelf aan. Hiervoor werken we allicht samen met andere lectoraten, universiteiten, bedrijven en organisaties, want zij hebben iets dat wij niet voor hen kunnen invullen: domeinexpertise. De domeinexpertise van Ambient Intelligence zelf richt zich op IT, data en AI. Deze technologische inzichten passen we toe in verschillende domeinen, maar we hebben niet al die kennis in pacht. Domeinexpertise is

cruciaal, bijvoorbeeld om te begrijpen hoe data geïnterpreteerd moeten worden. Is een 7 goed of niet? Als een patiënt revalideert in drie maanden, wat betekent dat? We zien opeens geen data meer in een bepaalde periode, wat is er aan de hand?<sup>28</sup> Begrip van data en de uitkomsten van AI is niet iets dat de machine direct voor ons kan doen. Het kan ons van alles voorschotelen, maar het is aan ons om het te interpreteren.

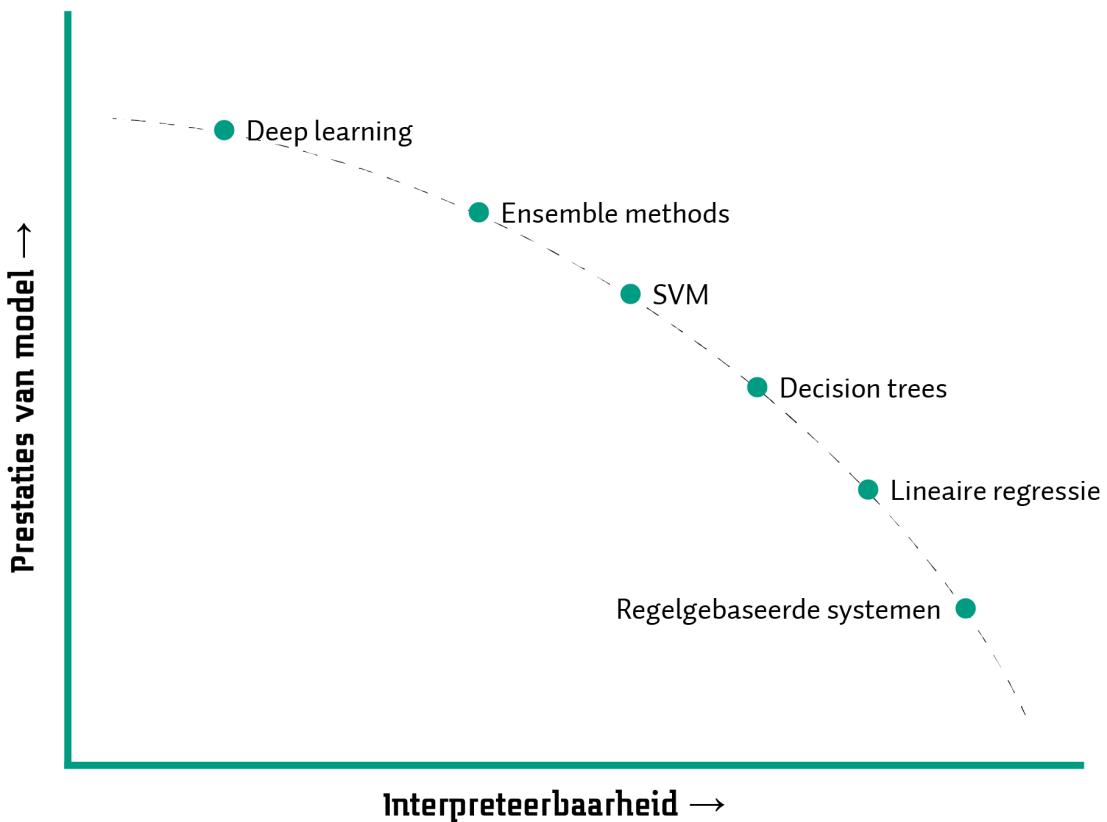
**Human-centered AI** is gericht op het zoeken naar die balans tussen mens en machine (Shneiderman, 2020). De aanpak hierbij is dat onderzocht moet worden welke niveaus van controle mensen moeten hebben. Soms moet de balans meer naar de één of ander uitwijken, om te voorkomen dat mensen overbelast raken met details (specifieke berekeningen van een deeplearningalgoritme) of dat het AI-model een ongewenste afweging maakt (beslissen om wel of niet te opereren). Dit is natuurlijk verbonden aan de ethische kanten van AI, waar gelukkig de laatste jaren meer en meer aandacht voor begint te komen. Dit is niet per se onderzoek wat het lectoraat Ambient Intelligence uitvoert. Wij zijn technischer van aard, maar houden wel deeglijk rekening met deze implicaties en principes van verantwoorde of **responsible AI** (Dignum, 2019). Het gaat hierbij om de ELSA: de ethische, legal (juridische) en sociomaatschappelijke aspecten van AI. Zulke aspecten dienen kritisch worden afgewogen door ontwikkelaars en gebruikers van AI. De AI Act van de EU is een goede stap naar regulering van AI, zodat deze op een verantwoorde manier ontwikkeld wordt.

Als wij als mensen willen begrijpen wat AI inhoudt, is geletterdheid over dit onderwerp gewenst, zoals ik eerder benoemde. Maar we kunnen verder gaan. We kunnen systemen ontwikkelen die de werking van AI uitlegbaar maken. Dit idee is gevangen in de term **Explainable AI** (XAI) (Arrieta e.a., 2020; Dwivedi e.a., 2023): AI die begrijpelijk of interpreerbaar is, waardoor mensen de uitkomsten van een AI-model beter op waarde kunnen schatten. AI kan gebruikt worden om data om te zetten in informatie op verschillende manieren: de algoritmes die bepaalde taken kunnen uitvoeren. Sommige algoritmes produceren modellen die simpel zijn of goed te volgen, zoals een formule die een rechte lijn aangeeft, bijvoorbeeld om aan te geven dat gemiddeld, hoe meer oppervlakte een huis heeft, hoe hoger de vraagprijs voor dat huis is. Zo'n verband wordt al gauw moeilijker uit te drukken als er meerdere variabelen een rol gaan spelen.<sup>29</sup> Bijvoorbeeld als er tientallen, misschien honderden variabelen zijn die van invloed zijn op de prijs van een huis. Maar ook beslisbomen kunnen gemaakt worden met AI, waarin er keuzes worden gemaakt door in een soort flowchart te kiezen op basis van de waarde van bepaalde variabelen. Met deep learning wordt het al een flink stuk lastiger: hiermee worden modellen gemaakt die bestaan uit tientallen lagen met duizenden neuronen. Een extreem voorbeeld is GPT-4, waarvan geschat wordt dat het model 1,8 biljoen parameters bevat (Schreiner, 11 July 2023).

---

<sup>28</sup>Dit voorbeeld is uit het leven gegrepen: het bleek dat er geen data werd verzameld tijdens de zomerstop van een bedrijf.

<sup>29</sup>En allerhande aparte verbanden kunnen gevonden worden, als je maar zoekt. Correlatie is geen causatie. Zie de voorbeelden van [Spurious Correlations](#).



Figuur 2.3: Uitlegbaarheid van bepaalde modellen

Met zulke complexe modellen is het voor mensen in eerste instantie niet mogelijk om te begrijpen wat er in het model gebeurt. Hierom worden het dan ook **black box**-modellen genoemd (Guidotti e.a., 2018). Dat het niet traceerbaar is hoe een model tot een uitkomst komt kan leiden tot verschillende dingen. Wellicht neemt iemand wat het model produceert simpelweg voor waar aan, misschien ten onrechte, wat zorgt voor risicovolle gevolgen. Of iemand die twijfels heeft bij de uitkomst, vertrouwt het toch niet, en laat het model voor wat het is. Hierbij ontstaat er een gemiste kans als het model toch betrouwbare resultaten geeft. Er is dus noodzaak voor XAI die op een manier mensen laat interpreteren wat in het model gebeurt (Molnar, 2022).<sup>30</sup> Dit kan bijvoorbeeld door grafisch weer te geven welke variabelen het belangrijkst zijn in de beslissing van een model, zodat een mens kan interpreteren of hier een logische verband tussen zit. Of er kan voor beeldherkenning getoond worden welke gedeeltes in een plaatje belangrijk zijn voor het identificeren van een ziekte (Borys e.a., 2023). Er wordt

<sup>30</sup>Het boek [Interpretable Machine Learning](#) is beschikbaar onder een openaccess-licentie.

nu ook onderzocht hoe, met dank aan taalmodellen, een uitleg gegeven kan worden over de uitkomsten van AI-modellen (X. Wu e.a., 2024).

In ons lectoraat is dit onderwerp een centraal aandachtspunt om de mens centraal en in controle te houden. Twee van onze onderzoekers promoveren hier dan ook op in samenwerking met de Universiteit Twente. Het PhD-onderzoek van Iris Heerlien gaat over ontwerprichtlijnen voor XAI in het gezondheidsdomein. Juist in deze sector is het essentieel dat werkwijzes transparant zijn om te onderbouwen waarom bepaalde keuzes gemaakt worden. AI-gebaseerde modellen kunnen sommige ziektebeelden nauwkeurig analyseren, maar niet altijd met voldoende menselijk inzicht in deze modellen. Daarom richt Iris zich op het opstellen van richtlijnen waarbij de gebruiker, de zorgprofessional, centraal staat. Vanuit diens perspectief gedacht zijn bepaalde vormen van uitleg over een AI-model beter toepasbaar dan anderen, afhankelijk van de vaardigheden van die persoon. Hiernaast voert Annemarie Jutte PhD-onderzoek uit naar het uitlegbaar maken van het gehele proces rondom de automatisering van machine learning (MLOps). Voordat de AI-modellen eenmaal gebruikt kunnen worden, ondergaan de data namelijk een lang proces (zie vorige sectie). Van verkenning van de data of deze wel geschikt zijn voor analyse, naar het aanpassen van data in geschikte formaten, tot uiteindelijk het in gebruik nemen en resultaten tonen van een model. Dit onderzoek is ook **user-centered**, maar dan op de ontwikkelaars van AI-systemen, die in deze automatisering grip moeten blijven hebben op de verwerking data door AI.

Aanverwant onderzoek doen we ook naar de samenwerking van intelligente systemen die kennis opslaan en weer kunnen delen. Zeker met de vergrijzing die een rol speelt in vele sectoren, is het belangrijk om bestaande kennis te borgen. Als we AI in staat willen stellen kennis van mensen te gebruiken, ligt het intuïtief voor de hand die kennis op de één of andere manier te verzamelen. Hiernaar doen we onderzoek in de context van kennisrepresentatie in de maakindustrie.<sup>31</sup> Mensen zijn zich vaak niet bewust van alle kennis die ze hebben, ze hebben impliciete kennis. Deze kan gevangen worden door mensen bewust te laten uitspreken of opschrijven, waarbij we AI-gebaseerde technologie deze uitspraken, woorden of visuele acties laten analyseren. Als we in staat zijn structuur en patroon te herkennen in de werkzaamheden van mensen, kunnen we dit vangen in modellen. Met deze modellen kunnen we vervolgens nieuwe arbeidskrachten sneller opleiden en mensen ondersteunen in hun bezigheden, door er bijvoorbeeld voor te zorgen dat er geen stappen in procedures worden overgeslagen.

Dit zijn slechts enkele voorbeelden van ons AI-onderzoek bij Ambient Intelligence.<sup>32</sup> Verantwoordelijke inzet van AI is een hoeksteen bij al onze projecten. De uitdaging die we met Ambient Intelligence tegemoet gaan, is zorgen voor wederzijds begrip. Mensen die snappen hoe de machines werken, en machines die snappen hoe mensen werken. De verwachtingen van deze twee werelden moeten samenkommen en de verklaringen voor hun gedrag daarop aansluiten. Dan komen we tot **symbiotische AI**: samen levende, lerende intelligenties.

---

<sup>31</sup>In het RAAK-mkb-project [Borging van vakkennis in de maakindustrie met Augmented Reality](#).

<sup>32</sup>Al ons onderzoek is te vinden via [de website van Ambient Intelligence](#).

# 3 Toekomst: Grenzen vervagen

## Samenvatting

Hoe ver gaan we met AI? Met wat fantasie kunnen we allerlei innovaties bedenken en zelfs verwachten dat AI slimmer wordt dan mensen. Voordat we zover komen, zijn er nog vele onderwerpen die onderzocht moeten worden. Het meest belangrijke is dat we onszelf de vraag stellen wat we willen bereiken met AI, en wat ons mens maakt. AI kunnen we zien als een goede samenwerkingspartner als we hier zelf slim mee omgaan.

Maar hoe dan verder? Wat valt er te verwachten van AI de komende jaren en zelfs decennia? Komen we in een winter terecht, of beginnen hierna de zomerdagen van AI? Koffiedik kijken kent risico's, en ik zal me niet wagen aan concrete voorspellingen voor de jaren hierna. Wel belicht ik hieronder verwachtingen die anderen hebben, en durf ik het aan om te speculeren over de relatie tussen mens en machine.

## 3.1 Oneindige technologische groei?

Ik benoemde eerder al dat mensen onlosmakelijk verbonden zijn met technologie. Vanaf het moment dat we bepaalde gereedschappen (*tools*) zijn gebruiken, hebben we ons verder kunnen ontwikkelen en heeft onze evolutie sindsdien veel meer mentaal dan fysiek plaatsgevonden (Dennett, 2017). Van een paar miljoen jaar geleden tot nu zijn er veel uitvindingen gedaan, van stenen bijlen tot quantumcomputers. Als we deze historie bekijken zien we echter ook een patroon. Technologie versnelt. Dat wil zeggen: uitvindingen van nieuwe gereedschappen en vorderingen in technologie lijken elkaar steeds sneller op te volgen. Er zit bijvoorbeeld minder tijd tussen het kunnen gebruiken van elektriciteit (1825, de electromagneet) en de eerste digitale computer (1941, de Z3 (Rojas, 1997)), dan tussen de start van de Bronzen en IJzeren Tijdperken (3.300 en 1.200 voor Christus). De technologische groei zien we bijvoorbeeld ook in de industriële revoluties: rond het jaar 1800 maakte de industrie automatisering door. Allerlei processen in de maakindustrie konden plotseling veel sneller plaatsvinden. Daarna kwamen een revolutie rond het jaar 1900 dankzij communicatienetwerken en het gebruik van elektriciteit. De derde revolutie in deze lijn kwam dankzij digitale computers in de tweede helft van de 20e eeuw. En nu, in de vierde industriële revolutie, draait het om de verbinding van de analoge met de digitale wereld dankzij vakgebieden als het **Internet of Things** (IoT) en AI (Lasi e.a., 2014).

Kijkend naar de afgelopen vijftig jaar zien we ook dat het gebruik van computers duidelijk is toegenomen. Over de hele wereld gebruikte in 2023 ongeveer 69% van de bevolking smartphones.<sup>1</sup> Als we denken aan de bevolkingsgroei van onze planeet, zien we ook dat dat sneller en sneller gaat. Voor computerchips geldt eenzelfde soort groei volgens de **Wet van Moore**: het aantal transistoren op chips verdubbelt iedere twee jaar (Wikipedia contributors, 2024b). Dit is gegroeid van 2.300 transistoren in de Intel 4004 (1971) naar 7,5 miljoen transistoren in de Pentium II (1997) tot meer dan 100 miljard transistoren in de Ryzen EPYC (2024).

Dit soort groei wordt exponentieel genoemd: hoe groter iets wordt, hoe sneller het groeit. Dit is een lastig te overzien patroon. Mensen kunnen lineaire groei prima begrijpen, maar exponentiële groei is lastiger om te bevatten. Als het gras in een tuin iedere week vijf centimeter groeit, snappen we snel wanneer we moeten maaien om er nog overheen en niet doorheen te lopen. Als het gras exponentieel zou groeien, en iedere week zou verdubbelen in lengte, dan zou gras van 5 cm na een maand 80 cm lang zijn en na een jaar iets meer dan 225 miljard kilometer.<sup>2</sup> Andersom gedacht, als een vijver na een maand voor 10% is dichtgegroeid met planten, en deze planten verdubbelen iedere dag, hoe lang duurt het dan nog voor de vijver overvol zit? Slechts vier dagen.<sup>3</sup> Hoeveel tijd er daarvoor is gespendeerd aan die verdubbeling in groei is niet belangrijk en daardoor kan iets wat schijnbaar klein is onverwacht snel groot worden.

Deze versnellende snelheid van technologie wordt door sommigen gezien als iets dat gaat leiden tot een **technologische singulariteit**. Als onze computers steeds sneller worden en technologieën zoals AI zich ook sneller ontwikkelen, wat gebeurt er dan op een gegeven moment? Een explosie van intelligentie, dankzij machines die zelf leren en zichzelf verbeteren. Intelligentie is een ruim begrip, en zien we dat computers op bepaalde vlakken al intelligenter zijn dan mensen. Allicht gaat de discussie dan ook meer over **Artificial General Intelligence** (AGI): generieke artificiële intelligentie. AGI betekent dat deze kunstmatige intelligentie niet louter één taak beter kan dan mensen, maar verschillende taken, en op een hoger niveau van begrip, waardoor deze kan leren van diens eigen ervaringen.

De explosie van intelligentie is een hypothese van verschillende wetenschappers (Chalmers, 2016; Good, 1966; Vinge, 1993). Deze singulariteit is een punt in de tijd waarna ontwikkelingen niet meer voor te stellen zijn, omdat deze gedreven zullen zijn door bovenmenseleijke intelligentie.<sup>4</sup> Een ander, vaak aangehaald werk hierover, komt van Ray Kurzweil, die inschatte dat deze gebeurtenis in 2045 zal plaatsvinden (Kurzweil, 2005). Al deze verwachtingen zijn gebaseerd op een interpretatie van exponentiële groei, en met name dat deze niet stopt. Hoe en of zulke super-intelligentie gaat ontstaan wordt onderbouwd met argumenten die uitgaan van betere hardware en software. In andere woorden: snellere computers of efficiëntere algoritmes.

---

<sup>1</sup>Met een totaal van bijna 7 miljard smartphones, zie [Statistica](#).

<sup>2</sup> $5 * 2^4 = 80$  en  $5 * 2^{52} = 22.517.998.136.852.480$  cm. Een aloud verhaal over rijst op een schaakbord laat dit verband ook mooi blijken (Wikipedia contributors, 2024e).

<sup>3</sup> $10 * 2^4 = 160$ .

<sup>4</sup>Vinge benoemt ook de term **Intelligence Amplification** (IA) om inderdaad een soort mens/machine-symbiose te duiden die hierin zou kunnen helpen (Vinge, 1993).

Kritiek op de singulariteit hierop is legio, zelfs van Moore, wiens wet vaak wordt aangehaald (IEEE Spectrum, 2008). Vraagtekens worden geplaatst bij het patroon van exponentiële groei, want het is niet zeker of deze zo doorgaat of afvlakt.

Mijn mening over de singulariteit is dat we niet van tevoren dienen uit te sluiten dat het mogelijk is. Dat is net zo kwalijk als met volkomen zekerheid zeggen dat het wel gebeurt. Wat volgens mij belangrijker is, is onze houding ten opzichte van dit idee en hoe we omgaan met ontwikkelingen van AI. Los van de praktische tegenargumenten op basis van technologie waar we nu beeld van hebben, kunnen we ook theoretisch kijken naar de mogelijkheden en uitdagingen van het ontstaan van **super-intelligentie** (Bostrom, 2015). Wat iedere intelligente ‘levensvorm’ aanjaagt zijn doelen als zelfbehoud en het nastreven van een bepaald ‘nut’. Dat laatste is wederom een groot onderwerp, maar voor veel levensvormen draait het om het behouden van de soort en dus reproductie. Wat een hypothetische AI als nut inzet, is wellicht wat die AI al dan niet bewust initieel wordt meegegeven. Een infameus gedachte-experiment hierover verhandelt over een *paperclip maximizer* wiens doel het is om zoveel mogelijk paperclips te maken (Bostrom, 2020). Met een beetje AI in dat systeem besluit het dat mensen wel eens dit doel in gevaar kunnen brengen en dat mensen gemaakt zijn van atomen waarmee paperclips gemaakt kunnen worden. De gevraagde conclusie van dat verhaal is duidelijk en niet bedoeld om schrik aan te jagen, maar om bewustzijn te creëren over ethische risico’s van het inzetten van technologie waarvan de gevolgen niet bekend zijn.

Voordat we zo ver vooruitblikken, lijkt het mij verstandig om eerst te kijken welke onderzoeksinnovaties nu gaande zijn. Het is goed om verder vooruit te kijken, maar ook om de blik op de grond voor ons te hebben.

## 3.2 De horizon van AI

In de vorige sectie haalde ik speculaties aan over de richting waarin AI uiteindelijk zou kunnen groeien. Voordat we die horizon bereiken, is het goed om naar de kortere termijn te kijken. AI als vakgebied staat duidelijk niet stil. De generatieve AI die de afgelopen jaren aandacht krijgt kwam niet uit de lucht vallen; hieraan werd al jaren gewerkt. De taalmodellen van onder andere ChatGPT zijn gebaseerd op werk uit 2017 over de specifieke deeplearning-architectuur **transformers** (Vaswani e.a., 2017). De plaatjesgeneratoren komen voort uit onderzoek naar **generative adversarial networks** uit 2014. Hierin worden twee AI-modellen tegenover elkaar gezet, de een om nieuwe plaatjes te genereren op basis van een dataset, de ander om te beoordelen of deze plaatjes behoren tot de originele dataset. Gaandeweg worden beide modellen steeds beter in hun taak, met als resultaat een goede plaatjesgenerator (Goodfellow e.a., 2014; Reed e.a., 2016).

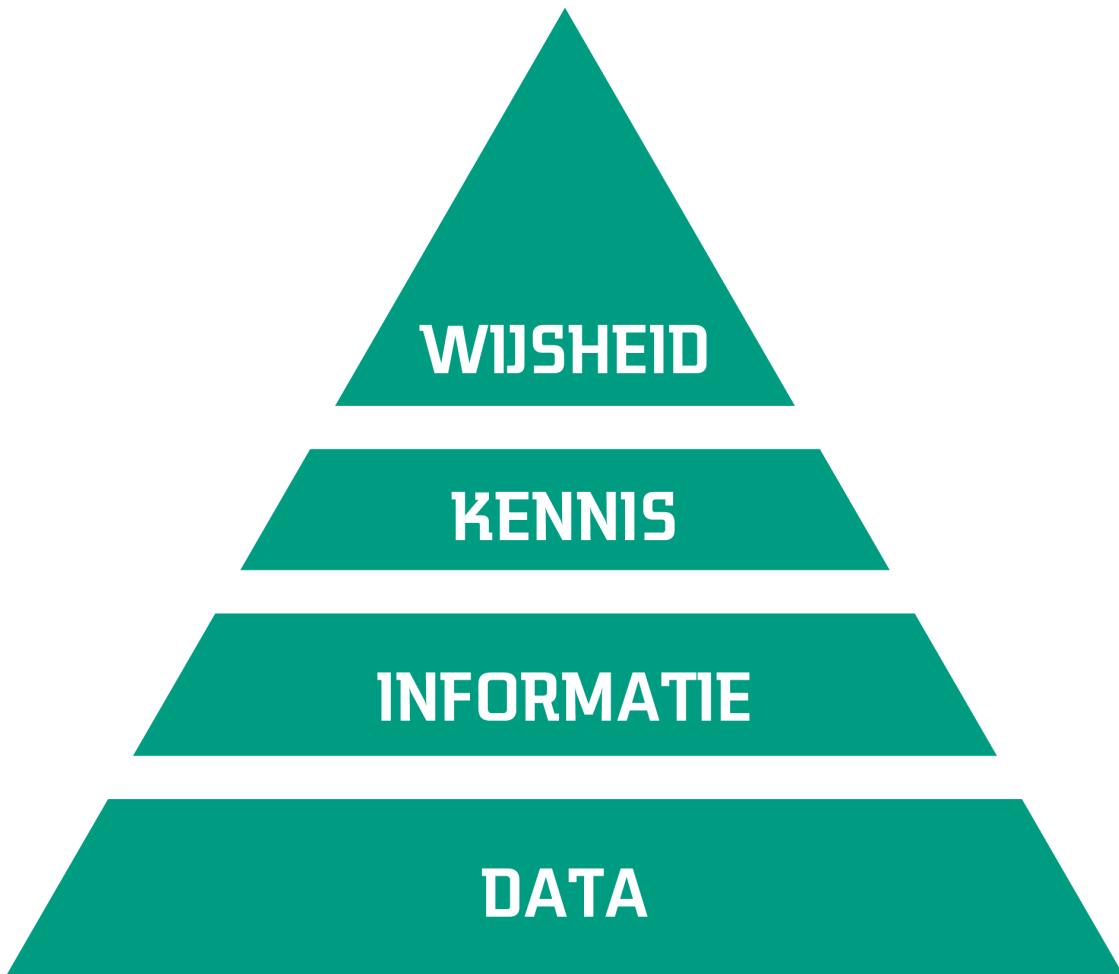
De vraag is natuurlijk of we nu ook kunnen bepalen welke van zulke innovaties de komende vijf á tien jaar een evengrote rol gaan spelen. Het veld van AI is breed, dus ik zal zeker niet alles kunnen benoemen. Wat ik belangrijk vind, zijn zaken die raken aan een veranderend perspectief op AI. Ten eerste twee nieuwe manieren om naar data te kijken.

Data is een voedingsbron voor AI, en AI-modellen hebben als doel die data om te zetten in informatie. Het *garbage-in-garbage-out*-principe vat het belang van goede data samen. **Data-centric AI** is een eerste nieuwe blik op de zaken: hierin wordt gekeken naar AI-technieken om data te verbeteren voordat de data gebruikt wordt een AI-model te trainen (Zha e.a., 2023). Het idee om data te bekijken en aan te passen voordat het aan een AI-algoritme gegeven wordt is niet nieuw. Maar het inzetten van AI om die data te verbeteren biedt wel degelijk vele mogelijkheden. Een groot probleem is namelijk dat veel data slecht of niet ‘gelabeld’ is: het is niet duidelijk wat de data betekent, bijvoorbeeld wat voor iets er op een plaatje staat, of of een bepaalde waarde in een grafiek slecht of goed is. Hiervoor zijn technieken in opkomst om toch met weinig of laagkwalitatieve data goede resultaten te behalen, zoals **semi-supervised learning** (Van Engelen & Hoos, 2020).<sup>5</sup>

Een tweede nieuwe blik op data gaat over de verhouding van data tot informatie, kennis en wijsheid. Dit is minder technisch ingestoken en gaat over het doel dat je met data wilt bereiken. Dit valt uit te beelden in een piramide vorm, waarop iedere volgende laag bouwt op die daaronder (Frické, 2019). Deze opbouw van niveaus is simplistisch, maar het laat ons wel op een andere manier kijken naar hoe we AI kunnen gebruiken.

---

<sup>5</sup>Dit is een soort middenvorm tussen *supervised* en *unsupervised learning*, waarmee bijvoorbeeld wordt gekeken naar de verdeling van de data en wat dit kan zeggen over individuele datapunten.



Figuur 3.1: DIKW-pyramide

Een nauwe blik op AI en data-analyse is namelijk dat we data omzetten in informatie. Deze informatie laat zien wat er aan de hand is, bijvoorbeeld dat er een kat op een foto gedetecteerd wordt, of een machine een bepaalde grenswaarde gaat overstijgen en daarmee kapot zou kunnen gaan. Een niveau hoger zit echter kennis: een kat is gedetecteerd op een foto, waarom is dat zo? Wijsheid is vervolgens het inzicht waarmee een vervolgactie gepland kan worden: de machine moet stopgezet en onderhouden worden. Deze **DIKW**-pyramide laat zien dat we op basis

van data verschillende soorten interpretaties kunnen doen. Dit helpt ons om beter te duiden waarvoor we een AI-systeem wel of niet kunnen gebruiken. Geeft zo'n systeem geen uitleg over een uitkomst? Dan is het niet goed genoeg om bepaalde beslissingen op te baseren. Zegt het niets over het effect van een bepaalde vervolgactie? Dan is het waarschijnlijk niet goed genoeg om een aanbeveling over te nemen.

Als we het dan over kennis hebben, is het ook belangrijk om kennis van mensen te gebruiken. Maar hoe digitaliseren we die kennis? Het vakgebied van **knowledge representation** (kennisrepresentatie) gaat hierover: hoe slaan we kennis op en kunnen we deze vervolgens gebruiken met computers? Hier komt de klassieke, symbolische AI weer om de hoek kijken in de vorm van **knowledge graphs** (Ji e.a., 2021). In plaats van datagedreven allerlei data te verzamelen, gaat dit over verbanden tussen concepten. Bijvoorbeeld dat deze zin bestaat uit woorden en leestekens, en dat die woorden bestaan uit letters. Dit kan worden uitgedrukt in een graaf die die verbanden laat zien, en waarin vervolgens een algoritme patronen kan opsporen en relaties kan doorgronden. Om kennis op een eenvoudige manier uit mensen te krijgen, kunnen bijvoorbeeld taalmodellen ingezet worden om deze verbanden vast te leggen, op basis van interviews en geschreven documenten (Pan e.a., 2024). Dit is een mooie samensmelting van de symbolische AI (uit de jaren '70) en de nieuwe aanpak van grote taalmodellen op basis van *transformers* (uit de jaren '10). Daarin zie ik grote kansen om ook weer beter verklaarbaar gedrag uit AI-systemen te krijgen, omdat inzichten uit zulke methodes weer op een uitlegbare manier naar mensen vertaald kunnen worden.

Ik vermoed ook dat er meer aandacht gaat gekomen voor meer verantwoord gebruik en ontwikkeling van AI.<sup>6</sup> Wetgeving en regulering zijn gaande om beter in handen te houden wat met wiens data gebeurt, zoals in de Data Act en AI Act van de EU. **Responsible AI** begint al meer en meer een begrepen concept te worden (Dignum, 2019). Ook voor **generative AI** zoals taalmodellen en de plaatjesgeneratoren wordt onderzocht wat wel en niet wenselijk is (Gu, 2024).

Daarnaast is er ook wat te zeggen voor duurzamere AI (Van Wynsberghe, 2021). Wat ik eerder aanhaalde is dat grotere AI-modellen meer rekenkracht en dus meer energie nodig hebben. De rekenkracht wordt ieder jaar efficiënter, maar de vraag naar die rekenkracht stijgt ook, en harder, dus er is almaar meer energie nodig (C.-J. Wu e.a., 2022). Het trainen van die modellen is maar één aspect van de levenscyclus van data voor AI. Andere fases zijn de verzameling tot experimentatie, het trainen van modellen, optimalisatie daarvan, en uiteindelijk het gebruik van een model. In ieder van die fases is winst te behalen, bijvoorbeeld via betere algoritmes. Maar een andere aanpak van de infrastructuur kan ook verbetering brengen. Op dit moment zijn veel AI-gebaseerde diensten namelijk afhankelijk van cloud-omgevingen om in te draaien. Daarvoor moet data over en weer gestuurd worden via het internet. Met efficiëntere algoritmes en AI-modellen kan AI echter ook lokaal toegepast worden (Deng e.a., 2020). Dit concept wordt **edge intelligence** genoemd, waarbij de intelligentie dus niet elders, in de cloud, zit, maar op een eigen systeem. Dit kent voordelen zoals onafhankelijkheid van internetverbinding en

---

<sup>6</sup>Misschien is dat wel *wishful thinking*, maar de ontwikkelingen zijn er gelukkig.

veiligheid vanwege geen communicatie met andere systemen. Vanuit duurzaamheidsoogpunt vergt dit echter ook weer een investering in kleinere, en meer computers.

Wat de komende tijd op de voorgrond gaat staan is de bredere inzet van de modellen die aan de huidige generatieve AI ten grondslag liggen: **foundation models** (Zhou e.a., 2023). Dit zijn modellen die getraind zijn met niet één doel voor ogen: ze kunnen generiek worden ingezet en zijn daarmee een stap naar **Artificial General Intelligence** (AGI). Dat wil zeggen dat er verschillende taken aan een model gegeven kunnen worden, zoals bij een taalmodel het schrijven van een tekst, het samenvatten daarvan of het geven van een antwoord op een vraag. Deze modellen hebben echter nog limitaties, maar die zijn op dit moment steeds minder van toepassing zijn door nieuwe ontwikkelingen. Huidige voortgang wordt bijvoorbeeld geboekt voor multimodale foundation modellen die getraind worden op zowel tekst als beeld en audio (Li e.a., 2024). Hoe snel deze modellen compleet generiek worden is een open vraag, ook omdat de definitie van AGI nu niet toereikend is om hier iets definitiefs over te zeggen.

De vraag is dan ook hoe slim AI echt is. Bij AI wordt al gauw gedacht dat zulke systemen zelf leren. Dat is echter nog niet zo. Er gaat veel werk zitten in het ophalen van de juiste data, deze verwerken, annoteren, mee experimenteren, dan het analyseren, het trainen van modellen met algoritmes, het evalueren, en uiteindelijk het inzetten. Met veel van deze stappen gaat mensenwerk gepaard. Hoe leren werkt is een vraag waarop het vakgebied **meta-learning** zich stort: het automatisch laten ontdekken wat een goede aanpak is om een model te trainen (Hospedales e.a., 2021).<sup>7</sup>

Waar echter voor de komende tijd de intelligentie in zit, is volgens mij toch echt de samenwerking tussen mensen en machines.<sup>8</sup> Hoe begrijpen we elkaar beter en kunnen wij als mensen AI optimaal en verantwoord inzetten? Dit wordt het elders **hybride intelligentie** genoemd (Akata e.a., 2020), waar ik kies voor symbiotische intelligentie. Dit licht ik toe in de volgende sectie over de volgende stappen voor mensen op weg naar onze horizon.

### 3.3 De horizon van mensen

Wat maakt ons mens?

Een vraag die ik heb uitgesteld om te beantwoorden is deze, en in het bijzonder zaken over emoties en bewustzijn. Sluitende antwoorden ga ik niet geven, en wellicht zijn die er ook niet.<sup>9</sup>

<sup>7</sup>Zie ook het concept van **Gödel machines** die hun eigen code kunnen herschrijven om zichzelf te verbeteren (Schmidhuber, 2007).

<sup>8</sup>Technieken als semi-supervised learning lenen zich hier goed voor.

<sup>9</sup>Een grapje dat vaker wordt gemaakt is dat wat mensen en computers onderscheidt, het beantwoorden van **CAPTCHA**'s is. CAPTCHA staat voor Completely Automated Public Turing test to tell Computers and Humans Apart. Bekende voorbeelden zijn de reCAPTCHA-testen van Google die gebruikt worden op websites, bijvoorbeeld door een aantal plaatjes aan te klikken waarop auto's staan. Dit wordt dan gebruikt om te bevestigen dat een mens probeert in te loggen in plaats van een softwarereprogramma. En daarnaast wordt dit ook gebruikt om informatie in te winnen om machinelearningmodellen te trainen: zo goed als

Ik haalde eerder de **Turing-test** aan, één van de eerste gedachte-experimenten om te bepalen of je praat met een mens of een computer. Dit gaat over het verschil tussen menselijke en kunstmatige intelligentie, wat een nauwere vraag is dan wat ‘mens zijn’ betekent. Kan een computer net zo menselijk zijn als mensen? De Turing-test zelf gaat alleen over taal, en slechts geschreven taal in de originele vorm van het experiment (Turing, 1950). Met de huidige **large language models** (LLM’s) lijkt het erop dat de Turing-test in sommige gevallen verslagen kan worden, waarbij een computer niet altijd onderscheiden kan worden van een mens (Elkins & Chun, 2020; Jones & Bergen, 2024). Huidige **text-to-speech**-software heeft inmiddels zulke realistische uitspraak dat deze niet meer van menselijke spraak te onderscheiden is (Tan e.a., 2024). Gekoppeld aan LLM’s zou dit er ook voor kunnen zorgen dat de Turing-test ook in gesproken vorm door een computer verslagen zou kunnen worden. Een ander inzicht is dat de Turing-test mensen vooral een spiegel voorhoudt. Wat je bespreekt met een computer hangt vooral af van wie je zelf bent, dus het zal iets zeggen over je eigen blik op menselijkheid (Sejnowski, 2023).

Dus een vorm van AI kan menselijk gedrag vertonen. De manier waarop dit gebeurt komt steeds dichter bij wat we ervaren als mensen: in tekst, spraak, en wellicht al snel ook in beeld.<sup>10</sup> Wat mij betreft is de interessantere vraag echter: waarom? Waarom willen we AI? Hierop zijn vele antwoorden te geven, die te maken hebben met de brede doelstellingen van het vakgebied. Ten eerste om systemen te maken die beter werken dan wat er nu is, efficiënter zijn, en nieuwe taken kunnen uitvoeren. Ten tweede om menselijke intelligentie beter te begrijpen. Ten derde: omdat het kan. Dit laatste is vaak een drijfveer van technologie. Het kan, dus we proberen het. Maar waarom willen we AI die menselijk is? Moet een AI-gebaseerd systeem per se menselijke intelligentie hebben om diens werk goed te doen? Moeten emoties onderdeel zijn van dit systeem?<sup>11</sup> Kan of moet een AI-systeem een slechte dag hebben, net als mensen dat zo hebben? Moeten AI-systeem ook menselijke normen en waarden hebben? Dit zijn gesloten vragen, maar het enige antwoord is ‘het ligt eraan.’ Context is uiterst belangrijk, dus hoe een AI-systeem wordt ingezet bepaalt wat nodig is.

Ik verwacht dat we anders naar intelligentie gaan kijken. Hopelijk met meer begrip van wat intelligentie inhoudt, maar ook welke normen en waarden we daaraan hangen. Respect voor elkaars (menselijke) intelligentie, maar voor andersoortige intelligenties. Op den duur verschuift, verwacht ik, het algehele beeld van wat mensen mensen maakt. We zijn verbonden met technologie, soms omdat het nodig is, vaak omdat we het gewend zijn. Mijn verwachting is dan ook dat wij steeds minder onderscheid gaan maken tussen mens en machine. Onze technologie maakt wie wij zijn, en andersom. Deze symbiose zou een **co-evolutie** genoemd kunnen worden (Lee, 2020). Mens en machine groeien naar elkaar toe, iets wat het **transhumanisme** aanhangt: het verbeteren van de mens dankzij technologie (Huxley, 2015). Wij veranderen ons

---

gratis annotatie van data door honderden miljoenen mensen op het internet (Von Ahn e.a., 2008).

<sup>10</sup>Zoals met **Sora van OpenAI** (Y. Liu e.a., 2024). Tastbare, fysieke AI in een realistische, menselijke vorm (humanoïde robots) lijkt echter nog ver weg (Mara e.a., 2021; Seaborn e.a., 2023).

<sup>11</sup>Affectieve Intelligentie? Met huidige chatbots op basis van LLM’s kunnen dialogen gevoerd worden die lijken op interpersoonlijke relaties en als zodanig gezien worden door mensen. Dit kan zover gaan als een liefdesverhouding met zulke chatbots (Brandtzaeg e.a., 2022.)

leven met behulp van AI, en wij veranderen AI ook door nieuwe ontwikkelingen. Hoe we ons verhouden tot AI is een kernvraag (Siemens e.a., 2022). Of intelligentie menselijk is of niet kan op den duur minder belangrijk worden. Misschien wegen de meningen van AI op een dag net zo zwaar als die van mensen. Daarvoor zijn dan wel goede kaders nodig om te bespreken wat standpunten zijn van menselijke en kunstmatige intelligentie. Onderbouwing van argumenten is vanuit een rationeel standpunt noodzakelijk. Fundamentele discussies over het toekennen van bepaalde eigenschappen aan AI, zoals bewustzijn (McDermott, 2007), zullen blijven doorgaan. Dit zijn uiteindelijk geen technische problemen, maar menselijke uitdagingen waarover we het gesprek moeten blijven voeren.

Gaat AI ons leven verbeteren? Dat ligt aan onszelf. Onderliggend aan technologie en de ontwikkeling daarvan liggen normen en waarden. Deze duiden wat we belangrijk vinden en zijn ook subjectief. Het publieke debat hierover zou ruimte moeten laten voor ieders mening, en vooral ook constructief moeten zijn (Brauner e.a., 2023; Fast & Horvitz, 2017). Er is noodzaak voor transparantie over wat AI wel en niet kan, en wat daarmee gepaard gaat. Om deze boodschap nogmaals te herhalen: er is hogere AI-geletterdheid nodig. Geïnformeerde discussie hierover helpt. Zowel voor de AI-wetenschapper die niet alle meningen en gevoelens van mensen hierover kan weten, als voor de mensen die niet weten wat er allemaal met AI gepaard gaat.

Als we verder in de toekomst proberen te kijken, is er veel mogelijk en vooral onzeker. Als de (rode) lijn tussen mens en machine vervaagt, volgt wellicht een ultieme symbiose. Dit mogelijke punt ligt in de toekomst, misschien ver weg en misschien dichter bij dan we denken. Daarom is het belangrijk dat we verstandiger en verantwoorder omgaan met AI. We sturen zelf hoe en of we daar willen komen. Het beste moment om te beginnen met deze discussie was 70 jaar geleden; het op één na beste moment is nu. In het volgende hoofdstuk belicht ik hoe we dat kunnen gaan doen.

# 4 Actie: Samen verantwoordelijk

## Samenvatting

Deze rede richt zich op bewustwording, maar is ook een aanzet tot actie. Mijn drie aanbevelingen over AI zijn: verbeter wederzijds begrip, onderzoek de impact van AI, en pas een goede methode toe. Ons lectoraat Ambient Intelligence zet hiervoor stappen samen met het onderwijs, andere hogescholen, en het bedrijfsleven. Het is voor iedereen belangrijk diens AI-geletterdheid te verbeteren. Een aantal startpunten daarvoor is beschikbaar als huiswerk na het lezen van deze rede.

### 4.1 Terug naar het nu: what will (A)I do?

Het is tijd voor actie. Het lijkt wellicht flauw om dit bij ieder individu te leggen, maar ik wil onderstrepen dat je jezelf niet moet onderschatten. Stel jezelf de vraag: *What will (A)I do?*

Onderzoek en wetenschap dienen de maatschappij. Als lectoraat komen we tot nieuwe inzichten op specifieke onderwerpen, zoals AI. We signaleren wat er gebeurt in fundamenteel onderzoek, testen dit in echte casussen, en bekijken hoe we bedrijven verder kunnen helpen hiermee. Dit is toegepast onderzoek, dat leidt tot aanbevelingen voor verantwoord en onderbouwd gebruik van deze technologie. Hieronder doe ik een aantal aanbevelingen over AI, geef ik aan wat ons lectoraat Ambient Intelligence gaat doen, en wat we gezamenlijk zouden moeten doen.

#### 4.1.1 Aanbevelingen over AI

Mijn aanbevelingen over AI vallen samen te vatten in drie punten.

##### 1. Verbeter wederzijds begrip.

In de symbiose tussen mens en machine moeten zij elkaar's patronen herkennen en interpreteren. Er is dus noodzaak voor verantwoordelijkheid van begrip van beide kanten. Mensen moeten leren hoe algoritmes toegepast worden, waarop zij gebaseerd zijn, en wat de mogelijkheden en tekortkomingen ervan zijn. Dit vergt een hogere **AI-geletterdheid** en kan geholpen worden door onderzoek naar **Explainable AI**. Hiermee zijn mensen beter in staat AI op waarde te schatten. Machines, de AI-systemen, moeten snappen wat mensen bedoelen, wat hun intenties

en verwachtingen zijn. De inzet van algoritmes moet worden aangepast aan de belangen van mensen, op een verantwoorde en dus duurzame manier. Dit betekent dat het proces deels participatief moet plaatsvinden, zodat belanghebbenden goed betrokken worden. Op technisch vlak betekent dit dat bepaalde stappen in het proces zo goed als mogelijk geautomatiseerd kunnen worden, zoals via **DataOps** en **MLOps**. Voor de machines betekent dit dat zij meer **mensgeletterd** moeten worden.

## 2. Onderzoek de impact van AI.

De onderzoeksagenda rondom AI moet verscherpt worden. Er is niet alleen behoefte aan korte-termijndenken voor de huidige hype van AI met alle generatieve vormen. We moeten een pad schetsen voor de impact van AI op langere termijn. Als we dat niet doen, onderschatten we de mogelijke effecten van AI, zoals de **Wet van Amara** ons waarschuwt. Initiatieven rondom het verantwoord omgaan met onze data, zoals de **Data Act** en de **AI Act** van de EU, ondersteunen dit. Als we nu nadenken over welke richting we opsturen, helpt dat ons doem-scenario's zo goed als mogelijk in te schatten. Op technisch vlak dienen we onderzoek te doen naar concepten als **responsible by design** en de integratie van normen en waarden in het ontwerp van AI-systemen (Díaz-Rodríguez e.a., 2023; Jobin e.a., 2019).

## 3. Pas een goede methode toe.

AI is geen magie, maar het is ook geen  $1 + 1 = 2$ . Er zijn vele *unknown unknowns*, zaken waarvan we niet weten dat we ze niet weten. Wat we kunnen doen is dit gestructureerd aanpakken. Wetenschappelijke methodes helpen daarbij en gaan per definitie gepaard met discussies en onzekerheden. Dit betekent niet dat we zulke processen moeten loslaten, maar dat we iteratief moeten bekijken wat wel en niet werkt. Enige terughoudendheid en reflectie op de ontwikkeling van AI is in mijn ogen belangrijk. De insteek is niet om te treuzelen hierop, maar juist van tevoren de (on)mogelijkheden en risico's van AI in kaart te brengen. Methodes zoals **CRISP-DM** zijn hiervoor een basale aanzet. De integratie van AI in complexe vraagstukken vraagt echter veel meer, namelijk de betrokkenheid van alle belanghebbenden. Enkel een bestuurlijk-technische aanpak hiervoor is zeker niet voldoende. Dit maakt een multidisciplinair, overkoepelend proces noodzakelijk en dit verdient nog veel meer (toegepast!) onderzoek.

### 4.1.2 Het lectoraat in actie

Met ons lectoraat Ambient Intelligence zetten we stappen die deze aanbevelingen volgen. Samenwerking is van belang en focus ook. We richten ons op de technische aspecten van de symbiose tussen mens en machine, waarbij we uitgaan van verantwoorde, human-centered principes voor de ontwikkeling en toepassing van AI. Zoals eerder en te meer benoemd zijn **Explainable AI** (XAI) en **MLOps** de twee pilaren van ons AI-gerichte onderzoek. Er zijn verschillende onderwerpen die hieraan grenzen en die we onderzoeken. **Edge intelligence** omvat de studie

naar efficiënte inzet voor intelligente systemen die op zichzelf staan, op kleine systemen die weinig energie verbruiken en deels onafhankelijk zijn van de cloud (Deng e.a., 2020). Ook staat ***immersive data-visualisatie voor data-exploratie*** op onze roadmap, waarin we onderzoek doen naar alternatieve manieren om data te verkennen, zoals door 3D-visualisaties in ***virtual reality*** (Kraus e.a., 2022; Qin e.a., 2020).

In drie verschillende SPRONG-programma's doen we op dit moment gezamenlijk met andere lectoraten en andere organisaties onderzoek. Een **SPRONG-programma** dient ertoe deze partners samen te brengen tot één grote onderzoeks groep, in een traject van acht jaar. In de SPRONG **Data Engineering en Management in Dataketens (DEMAND)** richten we ons op de noodzaak om veilig, verantwoord en efficient met data om te gaan. Datagedreven werken met AI kan immers alleen als het data-fundament op orde is. Samen met Saxion, en de HAN en Fontys hogescholen onderzoeken we onderwerpen als data management, data engineering en AI engineering om deze kennis te ontwikkelen en delen. In de SPRONG **Digital Driven Manufacturing (DDM)** onderzoeken we hoe we de maakindustrie met AI en robotica verder kunnen digitaliseren, om deze efficiënter in te richten en het professionals in deze sector te ondersteunen. Dit is een samenwerking tussen Saxion en Windesheim hogescholen, gericht op onderwerpen als autonome navigatie in industriële omgevingen, predictief onderhoud en procesautomatisering. In onze derde SPRONG **GROUNDED** bekijken we hoe we een gezonde, veilige en duurzame leefomgeving kunnen ontwikkelen met behulp van sensoren, data-analyses en mens-machine-interactie. Dit pakken Saxion en de Haagse Hogeschool samen op door onderzoek te doen naar nanotechnologie voor nieuwe sensoren om data te verzamelen, data beschikbaar te maken in dataplatforms, en analyses van deze met AI uit te voeren en interpreteerbaar te maken. Via deze SPRONGen en met hun partners ontwikkelen en toetsen we onze kennis op data- en AI-gebied, en delen deze kennis met belanghebbenden.

Als lectoraat verbreden en verdiepen we onze verbinding met onderwijs. In de nieuwe bacheloropleiding **Applied Data Science & AI** en de master **ICT: Software Engineering**, en de bachelor **HBO-ICT** werken de docent-onderzoekers van ons lectoraat mee om relevant en toekomstgericht onderwijs over AI op te zetten en uit te voeren. Dit reikt van introducties tot machine learning en CRISP-DM als methodologie tot diepgaande uitleg over deep learning en MLOps. Ook streven we ons onderwijsmateriaal goed deelbaar te maken binnen Saxion, want AI is voor ieder vak van belang. Daarnaast onderzoeken we hoe we AI voor onderwijs in kunnen zetten, zoals wat de mogelijkheden zijn voor chatbots die docenten en studenten ondersteunen. We werken mee aan onderwijs over AI en AI voor onderwijs.

In Deventer openen we tijdens mijn installatie als lector op 17 oktober ons **Data+AI Lab**, als samenwerking tussen de lectoraten Ambient Intelligence en **Smart Cities**, vanuit de SPRONG DEMAND. De doelstelling van het Data+AI Lab is kennis over data en AI ontwikkelen, bundelen en beschikbaar maken. De hoofdvraag is dan ook hoe we de symbiotische samenwerking tussen mens en machine kunnen ondersteunen. Dit doen we vanuit principes voor datagedreven werken: verantwoord, duurzaam en human-centred. Van belang zijn het duiden van de juiste processen en *methodes*, de beschikbare *tools* zoals softwareframeworks en -pakketten, en de toepasbare *technieken*, namelijk algoritmes voor machine learning, deep learning, en andersoortige

AI. Vanuit het onderzoek dat we in dit lab uitvoeren met onze docent-onderzoekers, bedrijven, overheden en studenten, stellen we hiervoor **best practices** op. Ook experimenteren we met cloud computing voor rekencapaciteit. Kernvraag hierbij is hoe we op een gestructureerde, efficiënte, en reproduceerbare manier met data en AI omgaan. Wat we aan kennis ontwikkelen, brengen we waar mogelijk uit als *open* resultaten. Dat wil zeggen: toegankelijk voor geïnteresseerden in het kader van **open science**. Ook zien we kansen om open data te genereren en toegankelijk te maken voor verder onderzoek. Al onze opgedane kennis bundelen we in een *Body of Knowledge, Skills & Attitude (BoKSA)* voor data en AI. De visie voor het Data+AI Lab is dat het een platform wordt voor onderzoek en onderwijs binnen Saxion, maar ook voor belanghebbenden daar buiten. Als je met data en AI aan de slag wilt, is het lab een uitstekend vertrekpunt waar je kunt kennismaken met deze technologie en ook terecht kunt voor verdere verdieping. Hiermee streven we ook naar een hechtere (learning) community voor AI binnen Saxion, want we moeten samen aan de slag.

#### 4.1.3 Samen in actie

AI inzetten lukt ons niet alleen. Niet als onderzoekers, technici, beleidsmakers, docenten, studenten, bedrijven, overheden, burgers afzonderlijk. AI vergt actie van ons allen. Iedereen heeft belangen die verschillen van die van anderen of daar deels mee overlappen. We even allemaal onze eigen rode draad. Iedere discipline van onderzoek en iedere sector kent een ander perspectief op de zaken. Om AI goed te kunnen ontwikkelen en toe te passen, is een multidisciplinaire aanpak essentieel. Dit is niet een kwestie van samenkomen en even brainstormen, maar een proces van de lange adem. Zoals ik hierboven aanhaalde: men moet meer snappen van AI en het gebruik van data en er moet een langetermijnvisie komen die gedragen wordt door de samenleving. Om verder te komen met AI waarbij iedereen ervan profiteert, moet ieders belang vertegenwoordigd kunnen zijn. Laat deze rede een oproep zijn tot zulke Activistische Intelligentie.

### 4.2 Huiswerk: leven is leren

Deze rede is niet af. Dat wil zeggen: het vakgebied AI heeft nog genoeg te ontdekken. Er is nog veel te doen, en dat lukt ons als lectoraat, als onderzoekers, als Saxion, niet in ons eentje. Daarvoor hebben we ook jou, de lezer, nodig. Niet om nieuwe algoritmes te ontwikkelen of data te analyseren, maar om bewust(er) te worden van AI. Daarvoor geef ik hieronder huiswerk mee, als oefening voor de lezer.<sup>1</sup> Mijn persoonlijke motto is: leven is leren. Ik hoop dat jij er ook voor open staat om meer over AI te weten te komen, om het verantwoorder te gebruiken.

---

<sup>1</sup>Hiervan komt niets op het tentamen, het is allemaal facultatief.

#### **4.2.1 Geletterdheid verbeteren**

Het verbeteren van geletterdheid is essentieel. In de basis is taalgeletterdheid het fundament: woorden, zinnen, teksten leren begrijpen en interpreteren.<sup>2</sup> In het verlengde hiervan is **datageletterdheid** vandaag de dag meer en meer belangrijk. Zo ook **AI-geletterdheid**. De impact van digitale technologie is verreikend. Het wordt ook wel de **digitale transitie** genoemd, omdat vele zaken erdoor veranderen, zowel in de privé- als de professionele sfeer. Pas als je beter snapt hoe deze technologieën werken en waarin ze verwerkt zijn, kun je inzien wat wel en niet beïnvloed kan worden door data en AI. Dat betekent niet dat je heel specifiek hoeft te weten hoe bepaalde algoritmes werken, maar wel wat de voorwaarden en limitaties zijn hiervan. Moet er veel data beschikbaar zijn? Is een bepaald algoritme betrouwbaar? Welke effecten kan een AI-model hebben op werkprocessen? Met inzicht hierin kun je beter een beeld vormen over AI en beslissen of de inzet hiervan aansluit bij je eigen normen en waarden. Zo kun je geïnformeerd beslissen wat verantwoord, duurzaam en wenselijk is.

Het huiswerk hiervoor is als volgt: poets je kennis van data en AI op. Hiervoor zijn vele bronnen beschikbaar, dus hieronder een kleine selectie:

- [De Nationale AI-cursus](#) geeft op een begrijpelijke manier uitleg over de basisconcepten van AI, en heeft ook cursussen voor specifieke [toepassingsbieden](#) als de zorg, de creatieve industrie en onderwijs.
- [De Saxion AI-trainingssessies](#) zijn een aantal jaren geleden verzorgd door het lectoraat Ambient Intelligence, met uitleg over methodologie voor data-analyse, en diepgaandere sessies voor machine learning en deep learning.
- [De Data Literacy Project](#) heeft een aantal cursus die basisbegrippen binnen datageletterdheid uitleggen en hoe je hierin vervolgstappen kunt zetten.
- De online cursussen met verschillende diepgang op [Coursera](#), via [Stanford University](#), [MIT Open Learning](#), en op vele andere plekken.

#### **4.2.2 Methodologisch te werk**

Aan de slag gaan met data en AI is geen simpele kwestie. Het is niet zo makkelijk als data ergens uploaden en een AI-model terugkrijgen, of een korte instructie schrijven voor een chatbot. Het proces om data en AI in te zetten is complexer en vereist veel (menselijk) denkwerk. In het geheel is de toepassing van AI om modellen te genereren maar een klein gedeelte van dit werk. Veel meer tijd wordt gespendeerd aan het bepalen van het doel, verzamelen van data, bepalen van datakwaliteit, selectie van toepasbare algoritmes, valideren dat de modellen naar verwachting zijn, implementeren van deze modellen, integreren in bestaande diensten en processen, en het onderhoud hiervan op langere termijn. En in dit proces worden sommige stappen met opgedane inzichten opnieuw gedaan tot er een wenselijk resultaat uit komt. Kort gezegd: er komt nogal wat bij kijken. Hierin is een goede methode dus van belang, om ervoor te

---

<sup>2</sup>Hierin schuilt ook al een uitdaging, met naar schatting [2,5 miljoen laaggeletterden in Nederland in 2024](#).

zorgen dat alle nodige stappen worden uitgevoerd en gecontroleerd. Hiervoor zijn verschillende methodologien beschikbaar, zoals CRISP-DM, SEMMA en KDD (Azevedo & Santos, 2008), met nog vele afgeleiden en alternatieven (Martínez-Plumed e.a., 2019).

Het huiswerk hiervoor is als volgt: gebruik een passend procesmodel dat dit proces ondersteunt voor je eigen situatie. De standaard binnen ons lectoraat is CRISP-DM, zoals eerder beschreven in Hoofdstuk 2. Dit is de **CRoss-Industry Standard Process for Data Mining**, en bestaat reeds 20 jaar (Schröer e.a., 2021). Het is generiek en goed toepasbaar voor ieder domein. Centraal hierin staan de data die geanalyseerd moeten worden, gedreven door een doelstelling die aan het begin concreet gemaakt moet worden. Een aantal bronnen die helpen bij het toepassen van CRISP-DM zijn:

- [Het IBM-overzicht van CRISP-DM](#) dat uitlegt geeft over de verschillende fasen en taken.<sup>3</sup>
- [De Saxion AI-trainingssessies](#) waarin de eerste sessie specifiek uitleg geeft over het gebruik van uitdagingen van CRISP-DM.

#### 4.2.3 Kritisch nadenken over AI

Waarom AI?<sup>4</sup> Een fundamentele, essentiële vraag: waarom, of eigenlijk *waarvoor* zou je AI willen gebruiken? Willen we blijven afsturen op oneindige groei, of zijn daar limieten? Maakt het ons leven makkelijker? Gaat onze kwaliteit van leven omhoog? Is AI een doel of een *tool*? En gebruiken we de hamer van AI om te bouwen of te vernietigen? Hiervoor valt niet zomaar één document of cursus aan te wijzen die zulke vragen beantwoordt.

Het huiswerk hiervoor is als volgt: bedenk welke impact AI op je eigen leven zou kunnen hebben. Waarvoor zou je het willen inzetten? Wat zou het verbeteren in je leven? En als je een specifieke toepassing in gedachten hebt, dan kunnen twee tools van Nederlandse makelij helpen hierop te reflecteren:

- [De Product Impact Tool](#) die onderzoek ondersteunt naar de impact van technische producten op mens, maatschappij en milieu (Dorrestijn, 2012, 2024).
- [De Technology Impact Cycle Tool](#) die op een vergelijkbare manier de impact van technologie helpt in te schatten (Vorst & Kamp, 2021).

#### 4.2.4 De rode draad

Tot slot wil ik meegeven dat we ieder onze eigen rode draad spinnen. Wat logisch lijkt voor jou, is dat misschien niet voor iemand anders. Zo ook voor AI: wat is logisch, wat is wenselijk, en wat is menselijk? Neem een stukje uit deze rede mee dat voor jou belangrijk is en houd rekening met anderen. De laatste huiswerkopdracht is dan ook het simpelst en het moeilijkst:

---

<sup>3</sup>Ook beschikbaar als [document](#).

<sup>4</sup>*Why AI?*

gebruik deze rede om samen iets moois te weven van onze individuele draden en te zorgen voor een toekomst met verantwoorde, duurzame AI.

# **Nawoord**

Hieronder benoem ik een aantal belangrijke zaken die niet aan bod kwamen in de lopende tekst: mijn dankwoord, persoonlijke reflectie, en aanraders voor verschillende media over AI.

## **Dankwoord**

Dat ik dit punt heb bereikt, is niet te danken aan alleen mezelf. Hiervoor heb ik veel steun ontvangen van vele mensen; teveel om op te noemen. Heel veel verschillende dingen hebben hiertoe geleid, dus laat ik in ieder geval de voornaamste benoemen.

Deze rede kwam tot stand met dank aan:

- Alle mensen van ons lectoraat Ambient Intelligence. Een creatieve, diverse, behulpzame groep. Samen verkennen we wat wel en niet mogelijk is via ons toegepast onderzoek. Er is nog genoeg te ontdekken, en ik heb er alle vertrouwen in dat we die reis met jullie kunnen maken. Speciale dank aan Bram en Danny voor hun uitgebreide input.
- Mijn mede-lector Wouter Teeuw, omdat ik dankzij hem veel heb kunnen leren over het wel en wee van onderzoek bij Saxion. We houden elkaar scherp en in balans.
- Alle collega's van onze Academie voor Creatieve Technologie, van alle lectoraten, opleidingen, en ondersteuning. In het bijzonder het lectoraat Smart Cities, waarmee we steeds meer samenwerken. Vanuit ons lectoraat Ambient Intelligence zien we vrijwel altijd overal kansen. In ons Data+AI Lab zie ik mooie dingen ontstaan.
- De directeur van onze Academie voor Creatieve Technologie, omdat hij mij deze rol kon aanbieden. In de regio Stedendriehoek (Apeldoorn, Deventer, Zutphen) ga ik AI op de kaart zetten.
- Het College van Bestuur, met name Anka Mulder, omdat zij heil zag in deze kans voor mij en voor Saxion.
- Alle mensen van Saxion die in meer of mindere mate hebben bijgedragen. De vraagstukken over IT en digitalisering vinden altijd plaats in een context. Zonder inzichten uit andere domeinen en andere perspectieven riskeren we dat we ons blindstaren op de techniek. Speciale dank aan Ronald voor het verwijzen naar Quarto, waarmee ik deze rede vormgaf.
- Mijn (schoon)familie voor hun steun, kritiek, en gevarieerde kijk op de zaken. Hartstikke mooi dat ik jullie heb en ik hoop dat jullie het ook andersom zo zien. Ieder doet diens eigen ding en dat is prachtig om te zien.

- Een speciale noot gaat uit naar mijn oma's en opa's. 'Ze zouden trots op je geweest zijn.' Ik zie het liever zo dat ik trots op hun ben. Men zegt dat je twee keer sterft. Eén keer als je lichaam ermee ophoudt, en één keer als je uit de herinnering van levenden treedt. Ik vergeet jullie niet.
- Leonie, mijn partner. Zij zou zeggen: AI is net als paardrijden, het gaat om samenwerking, een soort symbiose. We snappen elkaar steeds beter.
- Kleine Ivar, die ik zie leren en daarvan zelf ook leer. Zoals een sciencefictionschrijver twintig jaar geleden zei: de beste manier om een intelligentie te creëren, is om een nieuw mens op de wereld te zetten en te helpen leren.<sup>5</sup> Het is ongelooflijk mee te maken hoe ons kind zich ontwikkelt. Daar kan geen AI tegen op.
- En ten slotte bedankt aan causaliteit. Waar zouden we zijn zonder oorzaak-gevolg? Er zijn zoveel dingen waarvan ik wellicht niet eens weet dat ze hebben bijgedragen, zoveel rode draden die onzichtbaar zijn. Dankzij dat complexe samenspel van omstandigheden zijn we hier aangekomen en het zal blijken hoe alles zich verder ontwikkelt. Nogmaals, de toekomst voorspellen lukt het beste door hiervoor actie te ondernemen.

## Reflectie

Na iedere grote taak is reflectie een wijs iets. Ik had al snel een idee van waar deze rede over zou moeten gaan. Dat was eigenlijk meer een vraag dan een concreet voorstel, namelijk: wat is intelligentie en wat hebben wij er aan? En wat hebben we dan aan artificiële intelligentie? En hoe gaan we daar verstandig mee om? Dat samenspel, die balans, die spagaat wellicht tussen techniek en menswezen, dat is nogal wat.

Ik heb zo goed als mogelijk uit alle verschillende rode draden die ontsponnen uit deze gedachten een verhaal opgetekend. Ik ben dan ook niet van mening dat dit hét verhaal is, maar slechts één van de mogelijke vormen. Mijn aanpak hierbij was veel zaken proberen te benoemen en dat langzaamaan structuur te geven. Dat ging zeker niet lineair. Bij iedere zin die ik schrijf, kwamen er minstens twee nieuwe gedachten op. Als een draak waarbij je na het afhakken van ieder hoofd, er twee nieuwe aangroeiden. Schrijven is schrappen, en het is ook voldoening nemen met wat er staat. Wat ik meermaals zei tijdens dit proces was dan ook: 'het is nooit af.' Maar wat er nu staat is wel een verhaal dat gelezen kan worden. Het is niet een begin, ook geen eind, maar een stuk van rode draden dat ik geweven heb (met input van vele anderen, al dan niet direct). Het vervolg zou moeten zijn dat het anderen inspireert tot nadenken en tot actie.

Ik ben een *late adopter*, maar een *early investigator*. Ik wil eerst goed uitzoeken hoe iets zit en begrijpen wat het doet, wie er achter zit en waarvoor het al dan niet nodig en nuttig is, voordat ik het gebruik. Tijdens het werken aan deze rede ben ik vaker gevraagd naar mijn motivatie voor alles, en dan kwam vaker de opmerking naar voren dat ik me veel bezig houd met de ethische en morele kaders van AI. Dit staat voor mij buiten kijf: hoe kunnen we bezig

---

<sup>5</sup>Ik meen dat het Neal Stephenson was, maar hier blijf ik een bronvermelding schuldig.

zijn met technologie zonder me te baseren op bepaalde principes, normen en waarden? Als we dat niet zouden doen lopen we het risico onverantwoord bezig te zijn. Ik zal die boodschap er continu in blijven hameren: niet alles wat mogelijk is, is wenselijk. Eerst onderzoeken wat iets inhoudt, wat daarvan de mogelijke impact zou kunnen zijn, en daarna pas. Dat wil niet zeggen dat we jaren bezig moeten zijn voordat we iets implementeren, maar wel dat voorzichtigheid verstandig is. Liever iets minder dan exponentiële groei, als je het mij vraagt.

Het maken van deze rede was ook een experiment. Zoals wel vaker had ik meer ideeën dan tijd om deze uit te werken. Ik wilde iets schrijven dat begrijpelijk zou zijn en kan dienen als een introductie tot AI, maar ook een oproep tot bewustwording. Ik wilde bepaalde principes aanhangen. Vandaar dat ik géén gebruik heb gemaakt van *large language models* voor het schrijven van teksten in deze rede. Veel van deze modellen die in commerciële producten geïntegreerd zijn, zijn immers getraind op auteursrechtelijk beschermd materiaal. Daarnaast is het proces dat hiermee gemoeid is ook bijzonder ontransparant. Ook als het gaat om datakwaliteit van de teksten of plaatjes die in deze modellen gebruikt wordt, zijn er vele vragen te stellen. Ik was graag symbiotisch te werk gegaan om deze rede te schrijven, maar de LLM's die op dit moment beschikbaar zijn, voldoen niet aan mijn normen en waarden. Mijn hoop is dat de wetenschappelijke en commerciële sector hiermee beter zal omgaan in de toekomst. Wij dragen met het onderzoek bij ons lectoraat in ieder geval zo goed als mogelijk ons steentje bij.

Het schrijven van deze teksten was ook onderdeel van dit experiment: ik heb hiervoor [Quarto](#) gebruikt. Dit is een softwarepakket om op gestandaardiseerde wijze teksten te publiceren in verschillende formaten, zoals pdf, website, en epub, gebaseerd op een Markdown-variant. Ik had het niet eerder gebruikt, maar kan het nu aanraden voor diegenen die nastreven om open access te publiceren, en ook voor technische verslaglegging.

Daarnaast had ik nog wat verder meta-onderzoek willen doen naar deze rede zelf, bijvoorbeeld om de milieu-impact in kaart te brengen (wat is duurzamer: geprint of digitaal?) en om de teksten te analyseren (op taalniveau, sentiment, etc.). Dit laat nog even op zich wachten, als ik er al aan toe gaan komen. Voor eenieder die zich geroepen voelt om dit onderzoek aan te gaan: de rede zelf is beschikbaar onder een open-access-licentie, net als de broncode zelf.

Ik had in deze rede verder nog veel meer willen beschrijven, onder andere: fractals; chaos-theorie en entropie (in relatie tot informatietheorie); Lindenmayer-systemen en andere vormen van procedurele generatie als generatieve AI; filterbubbles; invloed van AI op wereldpolitiek dankzij bijvoorbeeld de gevolgen van deepfakes op de korte en lange termijn ('cheapfakes' vormen nu nog de meest impactvolle beïnvloeding); het feit dat [meer dan 50% van het internetverkeer in 2024 wordt gegenereerd door bots](#) (met eenderde van het totaal door kwaadwillende bots); hoe data in de vorm van browser cookies hebben geleid tot een vermarkting van het internet met een omzet van tientallen miljarden euro's; emergent gedrag van complexe (intelligente) systemen); cookies, EULA's, en andere zaken waarmee gebruikers van het internet en andere diensten ondoorzichtig worden behandeld; de driehoeksverhouding tussen mens, natuur en techniek; rationaliteit versus emotie versus logica - volgens mij kunnen emoties heel logisch zijn, maar zijn ze niet per se rationeel; de paradox van Jevons; artificiële kunst; Artificial Life; data management, data governance, data security, data warehouses, data lakes, FAIR

data, metadata; digital twins en alle definities daarvan; en nog veel meer dingen. Misschien de volgende keer.

## Media-aanraders

Mensen gebruiken verhalen om de wereld te begrijpen. Ik ben niet van mening dat wetenschappelijke literatuur de beste manier is om ideeën uit te leggen. Veel fictie boort nieuwe concepten aan die op hun beurt weer de wetenschap en de maatschappij beïnvloeden (en andersom). Ik gebruik dan ook liever de term *speculative fiction* dan *science fiction* om dit soort literatuur te benoemen. Ik ben er daarnaast van overtuigd dat geletterdheid gebaat is bij meer lezen. Zo ook voor AI-geletterdheid, dus ik draag graag een aantal werken voor die de geïnteresseerde lezer tot zich kan nemen. En niet alleen in de literatuur, maar ook in films en games komen op hun eigen manier relevante ideeën naar voren die nieuwe ideeën naar voren brengen, zowel ter inspiratie en als waarschuwing. De redelijke bekende boeken en films heb ik achterwege gelaten, dit zijn mijn persoonlijke aanraders die mij inspireren.

Dus, wil je meer te weten komen over...

- ...(Good Old-Fashioned) AI? Dan is [Gödel Escher Bach](#) van Douglas Hofstadter mijn definitieve aanrader. Tijdens mijn studie zeiden we al: als je dit boek hebt uitgelezen, ben je een propedeuse AI waardig.
- ...verschillende soorten van intelligentie?
  - [Blindsight](#) van Peter Watts is waarschijnlijk mijn favoriete boek, waarin ongelooflijk veel ideeën worden uitgediept rondom intelligentie, bewustzijn, transhumanisme, en dat op enigszins wetenschappelijke basis, met honderden referenties aan het eind van het boek. En het is beschikbaar als [open-access](#) publicatie.
  - [Vacuum Diagrams](#) van Stephen Baxter belicht allerlei vormen van intelligentie, samen met de evolutie van mensen over miljoenen jaren heen.
  - [The Sirens of Titan](#) van Kurt Vonnegut Jr. schetst hoe de mensheid dankzij een buitenaards bericht hun technologie kan verbeteren; de vraag is achter wat het bericht eigenlijk betekent.
  - [The Book of the New Sun](#) van Gene Wolfe speelt zich zo ver in de toekomst af dat onduidelijk is wat magie is en wat technologie. Het blijkt maar heel langzaam wat er aan de hand is. Dit boek is wat mij betreft hét voorbeeld dat speculatieve fictie literair kan zijn.
- ...transhumanisme?
  - [Diamond Dogs](#) van Alastair Reynolds verkent hoe ver een team onderzoekers gaat om hun eigen hersenen aan te passen om een buitenaards artefact te verkennen.
  - [Diaspora](#) van Greg Egan beschrijft hoe in de verre toekomst mensen in fysieke en digitale vorm bestaan, inclusief beschrijvingen in zesdimensionale ruimte; het is

maar goed dat de schrijver een complete website met wiskundige onderbouwingen bijlevert.

- [The Last Question](#) van Isaac Asimov is een kort verhaal over de laatste vraag die de mens ooit zal stellen.
  - [SOMA](#) van Frictional Games is een game die de vraag stelt wat bewustzijn inhoudt en doet dit met de hersenscan van de hoofdrolspeler die verongelukt is als vertrekpunt.
  - [Deus Ex](#) van Ion Storm pakt heel veel onderwerpen tegelijk op in een game met veel keuzevrijheid (voor die tijd): nanotechnologie, AI, heel veel overdreven complottheorieën, een wereldorde verstorende pandemie, en de keuze aan de speler om dit op een bepaalde manier op te lossen.
  - [2001: A Space Odyssey](#) van Arthur C. Clarke is een boek met een weergaloze verfilming door Stanley Kubrick dat hypotheseert dat de evolutie van mensen te danken is aan buitenaardse intelligentie, waar in de toekomst ook nog een botsing met AI door volgt.
- ...de invloed van (kunstmatige) taal op (kunstmatige) intelligentie?
    - [Story of Your Life](#) van Ted Chiang is een novelle over het ontcijferen van een buitenaardse taal die perceptie verandert. De verfilming [Arrival](#) beeldt dit goed uit, maar het boek is beter.
    - [Babel-17](#) van Samuel R. Delaney stelt dat taal een wapen kan zijn, gebaseerd op theorieën van [linguïstische relativiteit](#).
    - [The Diamond Age](#) van Neal Stephenson verhaalt over een boek dat een meisje helpt opgroeien door zich aan te passen aan haar omstandigheden.
  - ...AI die menselijk wordt?
    - [Blade Runner](#) van Ridley Scott is een verfilming van het boek [Do Androids Dream of Electric Sheep?](#) van Philip K. Dick, waarin synthetische mensen de vuile klusjes moeten opknappen; dat kan natuurlijk niet goed gaan. De sfeer die deze film neerzet is ongekend.
    - [Ghost in the Shell](#) verkent het thema identiteit in relatie tot belichaming daarvan.
  - Affectieve Intelligentie?
    - [HER](#) van Spike Jonze is een film waarin het lijkt te draaien om de opbloeide liefde van een man voor zijn computerbesturingssysteem, maar waar op de achtergrond iets groter speelt.
    - [I Dated a Robot](#) is een aflevering van de animatieserie Futurama, waarin de risico's van mens-machine-relaties en het kopiëren van menselijk bewustzijn worden belicht.
  - ...de gevaren en mogelijkheden van data en AI, in non-fictie-vorm?
    - [Superintelligence](#) van Nick Bostrom verkent op theoretische basis wat er nodig is voor superintelligentie en wat deze intelligentie zou kunnen drijven.

- [From Bacteria to Bach and Back](#) van Daniel Dennett is een historie van de evolutie van menselijk bewustzijn, en legt hierin ook de link naar mogelijk kunstmatig bewustzijn.
  - [The Model Thinker](#) van Scott E. Page legt uit hoe je met verschillende, datagedreven manieren naar problemen kunt kijken.
  - [Echt Nep](#) van Menno van Doorn, Sander Duivestein en Thijs Pepping belicht vele voorbeelden van des- en misinformatie.
  - [Ghost Work](#) van Mary L. Gray en Siddharth Suri geeft uitleg over klikwerkers.
- ...en als je hiervoor passende muziek zoekt, dan nog een aantal persoonlijke aanraders die deze thematiek aansnijden.
  - [Obsolete](#) van Fear Factory is een album dat het onheilspellende beeld schetst van machines die mensen overheersen, en dat mensen achterhaald zijn.
  - [Relentless Mutation](#) van Archspire pakt dit thema nog wat extremer op, door te verwoorden wat er kan gebeuren als transhumanisme een stap te ver gaat.
  - [01011001](#) van Ayreon is een opera van Nederlandse bodem met vele gastzangers, over uplifting van de mensheid door middel van technologie.
  - [Data Renaissance](#) van The Algorithm doet wat de naam zegt: een muzikaal gezicht geven aan de digitale transitie.

Laat deze media je inspireren dan wel waarschuwen. Voor verdere aanraders houd ik me aanbevolen!

# Over deze rede

Dit is de lectorale rede van Jeroen Linssen, lector Symbiotic Artificial Intelligence bij het lectoraat Ambient Intelligence van Hogeschool Saxion. Hij sprak deze rede uit op 17 oktober 2024, in Deventer.

## Over Jeroen Linssen



Dr. Jeroen Linssen (1987) is Lector Symbiotic AI bij het lectoraat [Ambient Intelligence](#) van de [Saxion University of Applied Sciences](#). Hij is gepromoveerd bij de vakgroep [Human Media Interaction](#) van de [University of Twente](#) (2017) na zijn studie [Cognitive Artificial Intelligence](#) aan de [Utrecht University](#) (2011). Zijn onderzoek richt zich op het ontwerpen en inzetten van systemen die met behulp van Kunstmatige Intelligentie (AI) beslissingen ondersteunen. Hij leidt het lectoraat Ambient Intelligence samen met Wouter Teeuw. Jeroens doel is om de intelligentie van AI-gebaseerde systemen te verbeteren zodat ze beter met mensen kunnen samenwerken – en vice versa. Zijn persoonlijke motto is: leven is leren.

**ORCID** <https://orcid.org/0000-0002-2626-1837>

Email [j.m.linssen@saxion.nl](mailto:j.m.linssen@saxion.nl)

Website Ambient Intelligence

## Open science

Met deze rede wil ik de principes achter open science en FAIR data ondersteunen. Het algemene idee hierachter is dat wetenschap toegankelijk moet zijn voor iedereen.<sup>6</sup> Wetenschap is niet een mening gebaseerd op gevoel, maar een wijze van methodisch zaken analyseren, interpreteren, en (h)erkennen welke onzekerheden er zijn. Daarom stel ik het onderzoek dat wij uitvoeren zo goed als mogelijk openbaar ter beschikking, dus ook deze rede. Vandaar dat de tekst en de bronbestanden beschikbaar zijn om te lezen en delen volgens onderstaande licentie.

De lectorale rede Tussen patroon en persoon door Jeroen Linssen is gelicenseerd onder CC BY-NC 4.0.

De bronbestanden zijn tevens beschikbaar onder dezelfde licentie via de repository [TussenPatroonEnPersoon](https://saxonami.github.io/TussenPatroonEnPersoon/). De website zelf is gehost via Github Pages, op <https://saxonami.github.io/TussenPatroonEnPersoon/>.

Voor het citeren van deze rede kan de volgende BibTeX-referentie gebruikt worden:

```
@misc{linssen2024tussenpatroonenpersoon,
  author = {Linssen, J.M.},
  title = {Tussen patroon en persoon: een rede(n) voor symbiotische (artificiële) intelligentie},
  year = {2024},
  howpublished = {\url{https://saxonami.github.io/TussenPatroonEnPersoon/}},
  doi = {10.5281/zenodo.13939704},
  publisher = {Ambient Intelligence, Saxion University of Applied Sciences, Deventer}
}
```

Of, in APA-formaat:

Linssen, J.M. (2024). *Tussen patroon en persoon: een rede(n) voor symbiotische (artificiële) intelligentie*. <https://saxonami.github.io/TussenPatroonEnPersoon/>; Ambient Intelligence, Saxion University of Applied Sciences, Deventer. DOI:[10.5281/zenodo.13939704](https://doi.org/10.5281/zenodo.13939704)

---

<sup>6</sup>Zie <https://www.openscience.nl/> voor meer toelichting over open science en GO FAIR over Findable, Accessible, Interoperable en Reusable data.

# Referenties

- Aïmeur, E., Amri, S., & Brassard, G. (2023). Fake news, disinformation and misinformation in social media: a review. *Social Network Analysis and Mining*, 13(1), 30.
- Akata, Z., Balliet, D., De Rijke, M., Dignum, F., Dignum, V., Eiben, G., Fokkens, A., Grossi, D., Hindriks, K., Hoos, H., e.a. (2020). A research agenda for hybrid intelligence: augmenting human intellect with collaborative, adaptive, responsible, and explainable artificial intelligence. *Computer*, 53(8), 18–28.
- Amari, S. (1967). A theory of adaptive pattern classifiers. *IEEE Transactions on Electronic Computers*, 3, 299–307.
- Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., García, S., Gil-López, S., Molina, D., Benjamins, R., e.a. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information fusion*, 58, 82–115.
- Azevedo, A., & Santos, M. F. (2008). KDD, SEMMA and CRISP-DM: a parallel overview. *IADS-DM*.
- Barbudo, R., Ventura, S., & Romero, J. R. (2023). Eight years of AutoML: categorisation, review and trends. *Knowledge and Information Systems*, 65(12), 5097–5149.
- Bird, C., Ungless, E., & Kasirzadeh, A. (2023). Typology of risks of generative text-to-image models. *Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society*, 396–410.
- Borys, K., Schmitt, Y. A., Nauta, M., Seifert, C., Krämer, N., Friedrich, C. M., & Nensa, F. (2023). Explainable ai in medical imaging: An overview for clinical practitioners–saliency-based xai approaches. *European journal of radiology*, 162, 110787.
- Bostrom, N. (2015). *Superintelligence: Paths, Dangers, Strategies*. Oxford University Press.
- Bostrom, N. (2020). Ethical issues in advanced artificial intelligence. *Machine Ethics and Robot Ethics*, 69–75.
- Brandtzaeg, P. B., Skjuve, M., & Følstad, A. (2022). My AI friend: How users of a social chatbot understand their human–AI friendship. *Human Communication Research*, 48(3), 404–429.
- Brauner, P., Hick, A., Philipsen, R., & Ziefle, M. (2023). What does the public think about artificial intelligence?—A criticality map to understand bias in the public perception of AI. *Frontiers in Computer Science*, 5, 1113903.
- Buchanan, B. G. (2005). A (very) brief history of artificial intelligence. *AI Magazine*, 26(4), 53–60.
- Buchanan, B. G., & Smith, R. G. (1988). Fundamentals of expert systems. *Annual review of computer science*, 3(1), 23–58.

- Campbell, M., Hoane Jr, A. J., & Hsu, F. (2002). Deep blue. *Artificial intelligence*, 134(1-2), 57–83.
- Carretero, S., Vuorikari, R., & Punie, Y. (2017). DigComp 2.1. *The Digital Competence Framework for Citizens. With eight proficiency levels and examples of use. Publications Office of the European Union*.
- Cave, S., & Dihal, K. (2019). Hopes and fears for intelligent machines in fiction and reality. *Nature machine intelligence*, 1(2), 74–78.
- Chalmers, D. J. (2016). The singularity: A philosophical analysis. *Science fiction and philosophy: From time travel to superintelligence*, 171–224.
- Choudhary, T., Mishra, V., Goswami, A., & Sarangapani, J. (2020). A comprehensive survey on model compression and acceleration. *Artificial Intelligence Review*, 53, 5113–5155.
- Chrisley, R. (2003). Embodied artificial intelligence. *Artificial intelligence*, 149(1), 131–150.
- Chu, X., Ilyas, I. F., Krishnan, S., & Wang, J. (2016). Data cleaning: Overview and emerging challenges. *Proceedings of the 2016 international conference on management of data*, 2201–2206.
- Cole, D. (2004). The Chinese Room Argument. In *The Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/entries/chinese-room/>; Metaphysics Research Lab, Stanford University.
- Copeland, B. (2024a). *Alan Turing*. <https://www.britannica.com/biography/Alan-Turing>
- Copeland, B. (2024b). *History of artificial intelligence (AI)*. <https://www.britannica.com/science/history-of-artificial-intelligence>
- Darrach, B. (1970). Meet Shaky, the first electronic person. *Life, November*, 58–68.
- Deng, S., Zhao, H., Fang, W., Yin, J., Dustdar, S., & Zomaya, A. Y. (2020). Edge intelligence: The confluence of edge computing and artificial intelligence. *IEEE Internet of Things Journal*, 7(8), 7457–7469.
- Dennett, D. (2017). *From Bacteria to Bach and Back: The Evolution of Minds*. W. W. Norton & Company.
- Dignum, V. (2019). *Responsible artificial intelligence: how to develop and use AI in a responsible way* (Vol. 2156). Springer.
- Díaz-Rodríguez, N., Del Ser, J., Coeckelbergh, M., Prado, M. L. de, Herrera-Viedma, E., & Herrera, F. (2023). Connecting the dots in trustworthy Artificial Intelligence: From AI principles, ethics, and key requirements to responsible AI systems and regulation. *Information Fusion*, 99, 101896.
- Dorrestijn, S. (2012). The Product Impact Tool. *Design for Usability Methods & Tools*, 111–119.
- Dorrestijn, S. (2024). *Ethisiek & Technologie, maar dan praktisch. Met bijdragen van de leden van het Saxion lectoraat Ethisiek & Technologie*. Ethisiek & Technologie, Saxion University of Applied Sciences, Deventer. <https://doi.org/10.5281/zenodo.12683806>
- Dwivedi, R., Dave, D., Naik, H., Singhal, S., Omer, R., Patel, P., Qian, B., Wen, Z., Shah, T., Morgan, G., e.a. (2023). Explainable AI (XAI): Core ideas, techniques, and solutions. *ACM Computing Surveys*, 55(9), 1–33.
- Elkins, K., & Chun, J. (2020). Can GPT-3 pass a writer's Turing test? *Journal of Cultural Analytics*, 5(2).

- Fast, E., & Horvitz, E. (2017). Long-term trends in the public perception of artificial intelligence. *Proceedings of the AAAI conference on artificial intelligence*, 31.
- Frické, M. (2019). The knowledge pyramid: the DIKW hierarchy. *Ko Knowledge organization*, 46(1), 33–46.
- Fuegi, J., & Francis, J. (2003). Lovelace & Babbage and the creation of the 1843 ‘notes’. *IEEE Annals of the History of Computing*, 25(4), 16–26.
- Gartner, Inc. (21 August 2024). *Gartner 2024 Hype Cycle for Emerging Technologies Highlights Developer Productivity, Total Experience, AI and Security*. <https://www.gartner.com/en/newsroom/press-releases/2024-08-21-gartner-2024-hype-cycle-for-emerging-technologies-highlights-developer-productivity-total-experience-ai-and-security>.
- Good, I. J. (1966). Speculations concerning the first ultraintelligent machine. In *Advances in computers* (Vol. 6, pp. 31–88). Elsevier.
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative adversarial nets. *Advances in neural information processing systems*, 27.
- Gu, J. (2024). Responsible generative ai: What to generate and what not. *arXiv preprint arXiv:2404.05783*.
- Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Giannotti, F., & Pedreschi, D. (2018). A survey of methods for explaining black box models. *ACM computing surveys (CSUR)*, 51(5), 1–42.
- He, X., Zhao, K., & Chu, X. (2021). AutoML: A survey of the state-of-the-art. *Knowledge-based systems*, 212, 106622.
- Heckerman, D. (1998). A tutorial on learning with Bayesian networks. *Learning in graphical models*, 301–354.
- Hofstadter, D. R. (1999). *Gödel, Escher, Bach: an eternal golden braid*. Basic books.
- Hospedales, T., Antoniou, A., Micaelli, P., & Storkey, A. (2021). Meta-learning in neural networks: A survey. *IEEE transactions on pattern analysis and machine intelligence*, 44(9), 5149–5169.
- Huxley, J. (2015). Transhumanism. *Ethics in Progress*, 6(1), 12–16.
- IEEE Spectrum. (2008). *Tech Luminaries Address Singularity*. <https://spectrum.ieee.org.saxon.idm.oclc.org/tech-luminaries-address-singularity>.
- Ji, S., Pan, S., Cambria, E., Marttinen, P., & Philip, S. Y. (2021). A survey on knowledge graphs: Representation, acquisition, and applications. *IEEE transactions on neural networks and learning systems*, 33(2), 494–514.
- Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature machine intelligence*, 1(9), 389–399.
- Jones, C. R., & Bergen, B. K. (2024). People cannot distinguish GPT-4 from a human in a Turing test. *arXiv preprint arXiv:2405.08007*.
- Kaelbling, L. P., Littman, M. L., & Moore, A. W. (1996). Reinforcement learning: A survey. *Journal of artificial intelligence research*, 4, 237–285.
- Kerschke, P., Hoos, H. H., Neumann, F., & Trautmann, H. (2019). Automated algorithm selection: Survey and perspectives. *Evolutionary computation*, 27(1), 3–45.
- Kraus, M., Fuchs, J., Sommer, B., Klein, K., Engelke, U., Keim, D., & Schreiber, F. (2022).

- Immersive analytics with abstract 3D visualizations: A survey. *Computer Graphics Forum*, 41, 201–229.
- Kreuzberger, D., Kühl, N., & Hirschl, S. (2023). Machine learning operations (mlops): Overview, definition, and architecture. *IEEE access*, 11, 31866–31879.
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25.
- Kumar, Y., Koul, A., Singla, R., & Ijaz, M. F. (2023). Artificial intelligence in disease diagnosis: a systematic literature review, synthesizing framework and future research agenda. *Journal of ambient intelligence and humanized computing*, 14(7), 8459–8486.
- Kurzweil, R. (2005). The singularity is near. In *Ethics and emerging technologies* (pp. 393–406). Springer.
- Lasi, H., Fettke, P., Kemper, H.-G., Feld, T., & Hoffmann, M. (2014). Industry 4.0. *Business & information systems engineering*, 6, 239–242.
- LeCun, Y. (1998). *The MNIST database of handwritten digits*. <http://yann.lecun.com/exdb/mnist/>
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444.
- LeCun, Y., Jackel, L., Bottou, L., Brunot, A., Cortes, C., Denker, J., Drucker, H., Guyon, I., Muller, U., Sackinger, E., e.a. (1995). Comparison of learning algorithms for handwritten digit recognition. *International conference on artificial neural networks*, 60, 53–60.
- Lee, E. A. (2020). *The coevolution: The entwined futures of humans and machines*. Mit Press.
- Li, C., Gan, Z., Yang, Z., Yang, J., Li, L., Wang, L., Gao, J., e.a. (2024). Multimodal foundation models: From specialists to general-purpose assistants. *Foundations and Trends® in Computer Graphics and Vision*, 16(1-2), 1–214.
- Liu, B. (2021). "Weak AI" is Likely to Never Become "Strong AI", So What is its Greatest Value for us? *arXiv preprint arXiv:2103.15294*.
- Liu, Y., Zhang, K., Li, Y., Yan, Z., Gao, C., Chen, R., Yuan, Z., Huang, Y., Sun, H., Gao, J., e.a. (2024). Sora: A review on background, technology, limitations, and opportunities of large vision models. *arXiv preprint arXiv:2402.17177*.
- Luccioni, S., Jernite, Y., & Strubell, E. (2024). Power hungry processing: Watts driving the cost of AI deployment? *The 2024 ACM Conference on Fairness, Accountability, and Transparency*, 85–99.
- Macal, C. M. (2016). Everything you need to know about agent-based modelling and simulation. *Journal of Simulation*, 10(2), 144–156.
- Mara, M., Stein, J.-P., Latoschik, M. E., Lugrin, B., Schreiner, C., Hostettler, R., & Appel, M. (2021). User responses to a humanoid robot observed in real life, virtual reality, 3D and 2D. *Frontiers in psychology*, 12, 633178.
- Martínez-Plumed, F., Contreras-Ochando, L., Ferri, C., Hernández-Orallo, J., Kull, M., La-chiche, N., Ramírez-Quintana, M. J., & Flach, P. (2019). CRISP-DM twenty years later: From data mining processes to data science trajectories. *IEEE transactions on knowledge and data engineering*, 33(8), 3048–3061.
- McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (2006). A proposal for the Dartmouth Summer Research Project on Artificial Intelligence, August 31, 1955. *AI magazine*, 27(4), 12–14.

- McDermott, D. (2007). Artificial intelligence and consciousness. *The Cambridge handbook of consciousness*, 117–150.
- Mitchell, M. (1998). *An introduction to genetic algorithms*. MIT press.
- Modha, D. S., Ananthanarayanan, R., Esser, S. K., Ndirango, A., Sherbondy, A. J., & Singh, R. (2011). Cognitive computing. *Communications of the ACM*, 54(8), 62–71.
- Molnar, C. (2022). *Interpretable machine learning*. Independently published.
- Moor, J. (2006). The Dartmouth College artificial intelligence conference: The next fifty years. *AI Magazine*, 27(4), 87–91.
- Newell, A., Shaw, J. C., & Simon, H. A. (1959). Report on a general problem solving program. *IFIP congress*, 256, 64.
- Nishant, R., Kennedy, M., & Corbett, J. (2020). Artificial intelligence for sustainability: Challenges, opportunities, and a research agenda. *International Journal of Information Management*, 53, 102104.
- Ntoutsi, E., Fafalios, P., Gadiraju, U., Iosifidis, V., Nejdl, W., Vidal, M.-E., Ruggieri, S., Turini, F., Papadopoulos, S., Krasanakis, E., e.a. (2020). Bias in data-driven artificial intelligence systems—An introductory survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 10(3), e1356.
- OpenAI. (2023). GPT-4 technical report. *arXiv preprint arXiv:2303.08774*.
- Pan, S., Luo, L., Wang, Y., Chen, C., Wang, J., & Wu, X. (2024). Unifying large language models and knowledge graphs: A roadmap. *IEEE Transactions on Knowledge and Data Engineering*.
- Penedo, G., Malartic, Q., Hesslow, D., Cojocaru, R., Cappelli, A., Alobeidli, H., Pannier, B., Almazrouei, E., & Launay, J. (2023). The RefinedWeb dataset for Falcon LLM: outperforming curated corpora with web data, and web data only. *arXiv preprint arXiv:2306.01116*.
- Pinar Saygin, A., Cicekli, I., & Akman, V. (2000). Turing test: 50 years later. *Minds and machines*, 10(4), 463–518.
- Qin, X., Luo, Y., Tang, N., & Li, G. (2020). Making data visualization more efficient and effective: a survey. *The VLDB Journal*, 29(1), 93–117.
- Reed, S., Akata, Z., Yan, X., Logeswaran, L., Schiele, B., & Lee, H. (2016). Generative adversarial text to image synthesis. *International conference on machine learning*, 1060–1069.
- Rojas, R. (1997). Konrad Zuse's legacy: the architecture of the Z1 and Z3. *IEEE Annals of the History of Computing*, 19(2), 5–16.
- Rosenblatt, F. (1958). The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review*, 65(6), 386.
- Russell, S. J., & Norvig, P. (2016). *Artificial intelligence: a modern approach* (4th dr.). Pearson.
- Schmidhuber, J. (2007). Gödel machines: Fully self-referential optimal universal self-improvers. In *Artificial general intelligence* (pp. 199–226). Springer.
- Schreiner, M. (11 July 2023). *GPT-4 architecture, datasets, costs and more leaked*. <https://the-decoder.com/gpt-4-architecture-datasets-costs-and-more-leaked/>.
- Schröer, C., Kruse, F., & Gómez, J. M. (2021). A systematic literature review on applying CRISP-DM process model. *Procedia Computer Science*, 181, 526–534.
- Seaborn, K., Barbareschi, G., & Chandra, S. (2023). Not only WEIRD but ‘uncanny’? A systematic review of diversity in Human–Robot Interaction research. *International Journal*

- of Social Robotics*, 15(11), 1841–1870.
- Sejnowski, T. J. (2023). Large language models and the reverse turing test. *Neural computation*, 35(3), 309–342.
- Shafique, U., & Qaiser, H. (2014). A comparative study of data mining process models (KDD, CRISP-DM and SEMMA). *International Journal of Innovation and Scientific Research*, 12(1), 217–222.
- Shanahan, M. (2016). The Frame Problem. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2016). <https://plato.stanford.edu/archives/spr2016/entries/frame-problem/>; Metaphysics Research Lab, Stanford University.
- Shi, Z., Yao, W., Li, Z., Zeng, L., Zhao, Y., Zhang, R., Tang, Y., & Wen, J. (2020). Artificial intelligence techniques for stability analysis and control in smart grids: Methodologies, applications, challenges and future directions. *Applied Energy*, 278, 115733.
- Shneiderman, B. (2020). Human-centered artificial intelligence: Reliable, safe & trustworthy. *International Journal of Human–Computer Interaction*, 36(6), 495–504.
- Siemens, G., Marmolejo-Ramos, F., Gabriel, F., Medeiros, K., Marrone, R., Joksimovic, S., & Laat, M. de. (2022). Human and artificial cognition. *Computers and Education: Artificial Intelligence*, 3, 100107.
- Tan, X., Chen, J., Liu, H., Cong, J., Zhang, C., Liu, Y., Wang, X., Leng, Y., Yi, Y., He, L., e.a. (2024). Naturalspeech: End-to-end text-to-speech synthesis with human-level quality. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.
- Turing, A. M. (1950). I.—Computing Machinery and Intelligence. *Mind*, LIX(236), 433–460. <https://doi.org/10.1093/mind/LIX.236.433>
- Van Engelen, J. E., & Hoos, H. H. (2020). A survey on semi-supervised learning. *Machine learning*, 109(2), 373–440.
- Van Wynsberghe, A. (2021). Sustainable AI: AI for sustainability and the sustainability of AI. *AI and Ethics*, 1(3), 213–218.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*.
- Vinge, V. (1993). Technological singularity. *VISION-21 Symposium sponsored by NASA Lewis Research Center and the Ohio Aerospace Institute*, 30–31.
- Von Ahn, L., Maurer, B., McMillen, C., Abraham, D., & Blum, M. (2008). re captcha: Human-based character recognition via web security measures. *Science*, 321(5895), 1465–1468.
- Vorst, R. van der, & Kamp, J.-A. (2021). The importance of a free, open, online Technology Impact Cycle Tool. *EUNIS'21*.
- Wach, K., Duong, C. D., Ejdys, J., Kazlauskaité, R., Korzynski, P., Mazurek, G., Paliszewicz, J., & Ziembka, E. (2023). The dark side of generative artificial intelligence: A critical analysis of controversies and risks of ChatGPT. *Entrepreneurial Business and Economics Review*, 11(2), 7–30.
- Weizenbaum, J. (1966). ELIZA—a computer program for the study of natural language communication between man and machine. *Communications of the ACM*, 9(1), 36–45.
- Wikipedia contributors. (2024a). *History of artificial intelligence — Wikipedia, The Free Encyclopedia*. [https://en.wikipedia.org/w/index.php?title=History\\_of\\_artificial\\_intelligence&oldid=113700000](https://en.wikipedia.org/w/index.php?title=History_of_artificial_intelligence&oldid=113700000)

- intelligence&oldid=1238546461.
- Wikipedia contributors. (2024b). *Moore's law* — Wikipedia, The Free Encyclopedia. [https://en.wikipedia.org/w/index.php?title=Moore%27s\\_law&oldid=1250511271](https://en.wikipedia.org/w/index.php?title=Moore%27s_law&oldid=1250511271).
- Wikipedia contributors. (2024c). *Roy Amara* — Wikipedia, The Free Encyclopedia. [https://en.wikipedia.org/w/index.php?title=Roy\\_Amara&oldid=1237425964](https://en.wikipedia.org/w/index.php?title=Roy_Amara&oldid=1237425964).
- Wikipedia contributors. (2024d). *Timeline of artificial intelligence* — Wikipedia, The Free Encyclopedia. [https://en.wikipedia.org/w/index.php?title=Timeline\\_of\\_artificial\\_intelligence&oldid=1240226624](https://en.wikipedia.org/w/index.php?title=Timeline_of_artificial_intelligence&oldid=1240226624).
- Wikipedia contributors. (2024e). *Wheat and chessboard problem* — Wikipedia, The Free Encyclopedia. [https://en.wikipedia.org/w/index.php?title=Wheat\\_and\\_chessboard\\_problem&oldid=1238847142](https://en.wikipedia.org/w/index.php?title=Wheat_and_chessboard_problem&oldid=1238847142).
- Wolfe, G. (1998). *The Book of the New Sun*. SFBC.
- Wu, C.-J., Raghavendra, R., Gupta, U., Acun, B., Ardalani, N., Maeng, K., Chang, G., Aga, F., Huang, J., Bai, C., e.a. (2022). Sustainable ai: Environmental implications, challenges and opportunities. *Proceedings of Machine Learning and Systems*, 4, 795–813.
- Wu, X., Zhao, H., Zhu, Y., Shi, Y., Yang, F., Liu, T., Zhai, X., Yao, W., Li, J., Du, M., e.a. (2024). Usable XAI: 10 strategies towards exploiting explainability in the LLM era. *arXiv preprint arXiv:2403.08946*.
- Yang, Li, & Shami, A. (2020). On hyperparameter optimization of machine learning algorithms: Theory and practice. *Neurocomputing*, 415, 295–316.
- Yang, Ling, Zhang, Z., Song, Y., Hong, S., Xu, R., Zhao, Y., Zhang, W., Cui, B., & Yang, M.-H. (2023). Diffusion models: A comprehensive survey of methods and applications. *ACM Computing Surveys*, 56(4), 1–39.
- Zha, D., Bhat, Z. P., Lai, K.-H., Yang, F., Jiang, Z., Zhong, S., & Hu, X. (2023). Data-centric artificial intelligence: A survey. *arXiv preprint arXiv:2303.10158*.
- Zhang, C., Zhang, C., Zhang, M., & Kweon, I. S. (2023). Text-to-image diffusion models in generative ai: A survey. *arXiv preprint arXiv:2303.07909*.
- Zhao, W. X., Zhou, K., Li, J., Tang, T., Wang, X., Hou, Y., Min, Y., Zhang, B., Zhang, J., Dong, Z., e.a. (2023). A survey of large language models. *arXiv preprint arXiv:2303.18223*.
- Zhou, C., Li, Q., Li, C., Yu, J., Liu, Y., Wang, G., Zhang, K., Ji, C., Yan, Q., He, L., e.a. (2023). A comprehensive survey on pretrained foundation models: A history from bert to chatgpt. *arXiv preprint arXiv:2302.09419*.

## Copyright

De figuur van het landschap van AI van LFAI in Hoofdstuk 2 is beschikbaar onder de [Apache 2.0-licentie](#).