



De (on)mogelijkheden van kunstmatige intelligentie in het onderwijs

In opdracht van:

Ministerie van Onderwijs, Cultuur & Wetenschap

Project:

2018.068

Publicatienummer:

2018.068.1828 v1.0.116

Datum:

Utrecht, 21 januari 2019

Auteurs:

ir. Tommy van der Vorst

ir. Nick Jelacic

mr. Marc de Vries

Julie Albers



Inhoudsopgave

Samenvatting	5
1 Introductie.....	9
1.1 Aanleiding.....	9
1.2 Onderzoeksvragen	9
1.3 Methode	10
1.4 Leeswijzer	10
2 Wat kan er?.....	13
2.1 Wat is AI?.....	13
2.2 Huidige en toekomstige toepassingen.....	17
2.3 Ontwikkeling in het buitenland	20
2.4 Eigenschappen van AI-toepassingen	22
2.5 Conclusie: toepassingsscenario's in het onderwijs	27
3 Wat mag er?.....	31
3.1 Toerekening van juridisch relevante gebeurtenissen aan personen.....	31
3.2 Verkenning van de juridisch relevante gebeurtenissen	32
3.3 Toepasselijke generieke regelingen	32
3.4 Toepasselijke sectorale regelingen	41
3.5 Conclusie: mogelijke juridische obstakels voor AI in het onderwijs.....	46
4 Wat willen we?.....	49
4.1 Mogelijke impact van AI in het onderwijs	49
4.2 Ethische afwegingen	50
4.3 Beleidsontwikkeling in het buitenland	53
5 Conclusie	57
5.1 Kansen van AI in het onderwijs	57
5.2 Risico's van AI in het onderwijs	58
5.3 Aanbevelingen voor beleid	59
6 Referenties	63
Bijlage 1. Overzicht interviewrespondenten	65
Bijlage 2. Overzicht van relevante regelingen.....	67
Generieke regelingen	67
Onderwijsregelingen	68
Versiegeschiedenis	69

Samenvatting

Het ministerie van Onderwijs, Cultuur en Wetenschap heeft Dialogic *innovatie & interactie* gevraagd een onderzoek uit te voeren naar de mogelijkheden van kunstmatige intelligentie (*artificial intelligence*, AI) in het onderwijs. Het doel van dit onderzoek is het verwerven van inzicht in hoe AI momenteel in het onderwijs wordt gebruikt, en hoe het in de komende 5-8 jaar gebruikt zal gaan worden (*wat kan er?*). Daarnaast onderzoeken we welke juridische aspecten daarbij spelen (*wat mag er?*). Tot slot identificeert het onderzoek de grootste risico's en kansen die verbonden zijn aan toepassing van AI in het onderwijs (*Wat willen we?*).

Wat kan er: wat zijn relevante toepassingen van AI in het onderwijs?

Het is niet eenvoudig om één definitie te geven van AI, specifiek als het gaat om toepassingen in het onderwijs. Uit ons onderzoek komt naar voren dat met name wanneer het gaat om toepassingen die (1) cognitieve taken automatiseren en (2) daarbij gebruik maken van grote hoeveelheden data en datagedreven methoden, er interessante, onopgeloste vraagstukken bestaan.

In het onderwijsproces maken leraren continu naar eigen inzicht beslissingen ten aanzien van (onder andere) de gehanteerde methode, de lesstof, de manier waarop een leerling wordt benaderd, et cetera. Uiteindelijk worden door leraren ook enkele formele beslissingen gemaakt: wat is het rapportcijfer, en mag een kind over naar de volgende klas? AI's kunnen docenten op verschillende manieren ondersteunen bij deze beslissingen. We onderscheiden een viertal scenario's als meest waarschijnlijk voor de komende 5-8 jaar: (1) een AI als onderwijsassistent, (2) AI voor learning analytics, (3) AI voor personalisering van het onderwijs, en (4) AI voor toetsing.

Kan AI een docent volledig vervangen? Niet in de nabije toekomst.

Wanneer AI zó slim is geworden dat het de docent kan vervangen, dan zou dat het onderwijs in theorie sterk kunnen verbeteren: iedere leerling kan dan immers 'privéles' worden geboden. Zo ver is het echter nog niet: de verwachting is dat dergelijke *artificial general intelligence* (waarbij een AI het niveau van intelligentie van een mens evenaart) nog minstens enkele decennia op zich laat wachten. Dit neemt echter niet weg dat minder intelligente AI, die vandaag al beschikbaar is, de leraar zodanig kan ondersteunen dat deze meer tijd per leerling kan besteden, of efficiënter te werk kan gaan.

Wat mag er: wat zijn juridische knelpunten bij toepassing van AI in het onderwijs?

Toepassing van AI in het onderwijs raakt aan diverse *generieke regelingen*. Sommige van die regelingen zijn van aanvullend recht, dat betekent dat daar gemakkelijk en veelal per contract van kan worden afgeweken, zoals bijvoorbeeld de regelingen van het Auteurswet en het Databankenrecht.

De flexibiliteit van genoemde privaatrechtelijke regelingen vindt men veel minder terug in de regelingen van *publiekrecht*, zoals de regels over openbaarheid en hergebruik van bestuurlijke respectievelijk overheidsinformatie. Toepasselijkheid van deze regelingen kunnen ontwikkelaars ervan weerhouden samenwerkingsverbanden met instellingen aan te gaan als dit zou betekenen dat hun kennis op straat komt te liggen. Aanvullende contracten over intellectuele rechten kunnen hier (gedeeltelijk) de helpende hand bieden. Overigens zal men er daarbij wel voor moeten waken dat samenwerking met marktpartijen uiteindelijk geen machtsposities doen ontstaan die in strijd kunnen zijn met regels over mededinging.

Besluiten die instellingen nemen worden uiteraard volledig beheerst door de algemene regels van het bestuursrecht, waaronder de algemene beginselen van behoorlijk bestuur. Toepassing van deze beginselen op AI-toepassingen kan lastig zijn, omdat deze inherent botsen met de *black box* die AI nu eenmaal schept. Desondanks lijkt dit geen onoplosbaar probleem te vormen.

De regels over de *bescherming van persoonsgegevens* (AVG) vormen mogelijk wél een knelpunt. De aard van de persoonsgegevens in het onderwijs en het karakter van AI gaan (juridisch gezien) heel slecht samen, en het lijkt heel moeilijk voorstelbaar dat grootschalige AI toegestaan is zonder wettelijke regeling. Uiteraard verandert dit geheel als de te gebruiken gegevens geen persoonsgegevens meer zijn.

Mocht er wat mis gaan dan lijken er geen noemenswaardige problemen bij de toepassing van de regels over *aansprakelijkheid*. Dit regime is zeer open en flexibel zodat zij die door AI-toepassingen benadeeld zijn niet op de tocht hoeven te staan.

De onderwijs-specifieke sectorale regels laten op zich veel vrijheid, en lijken niet echt harde verboden te bevatten. Wel zal steeds binnen de geest en letter van de beginselen van onderwijsrecht gehandeld moeten worden. Daar komt nog bij dat de belangen van het kind steeds een centraal en ethisch ijkpunt zullen moeten vormen en de vraag zal steeds zijn of dat met AI het geval is.

In meer praktische zin vormt de sectorale regelgeving wel een obstakel: voor *top down* sturing op inhoud van onderwijs bestaat geen sterk arrangement, nog afgezien van het feit dat dit heel gevoelig ligt. Ook de beweging van onderaf is een lastige: weliswaar kunnen er bij individuele instellingen initiatieven ontstaan maar het poldermodel en de diversiteit aan verschillende regelingen op het niveau van de instelling zal een grootschalige beweging waarschijnlijk in de weg staan. Dit nog afgezien van de mogelijke aversie van docenten (en instellingen), aan wiens belangen en positie toepassing van AI direct kan raken.

Wat willen we: wat zijn de belangrijkste kansen en risico's van toepassing van AI in het onderwijs?

Kansen van AI-toepassingen in het onderwijs

Toepassing van AI in het onderwijs biedt, zo volgt uit het onderzoek, mogelijkheden om de volgende positieve effecten te bewerkstelligen:

- Het verminderen van de werkdruk van docenten door toepassing AI ter ondersteuning van (administratieve) taken.
- Gepersonaliseerd leren: onderwijs beter laten aansluiten bij de leerling, met zowel betere uitkomsten als een beter leerproces.
- Het ondersteunen van de docent met holistische, onderbouwde inzichten (learning analytics).
- Het verbeteren van de manier van toetsen van kennis.
- Het vergroten van de effectiviteit van digitale leermiddelen, ook in synergie met andere technologieën, zoals VR en serious games.

Risico's

Uit het onderzoek volgt dat er bij de toepassing van AI in het onderwijs zou kunnen leiden tot de volgende mogelijk negatieve effecten, tenzij hiervoor mitigaties worden ingericht:

- Sommige onderwijsdoelen kunnen door AI in het geding raken, wanneer de focus op techniek en het aspect 'kennisoverdracht' te groot is.

- Bias in mensen vertaalt door naar data, die vervolgens wordt overgenomen door AI.
- Een verslechterd vooruitzicht bij 'docent' als beroep (voor huidige en toekomstige docenten).
- Afhankelijkheid van black-box modellen (unexplainable AI) versus verantwoordelijkheid van de docent.
- Toepassing van AI, terwijl andere basisvoorzieningen nog niet op orde zijn.
- Een machtsverschuiving bij de producenten van lesmateriaal.

Aanbevelingen

Om toepassingen van AI in het onderwijs te stimuleren is er een aantal aspecten waarop beleid kan worden gevormd, en/of waarop beleidsmakers (specifiek het ministerie van OCW) actie kunnen ondernemen:

1. Het stimuleren van acceptatie van AI in het onderwijs door leraar, leerling en ouder.
2. Het vergroten van de digitale vaardigheden van onderwijzend personeel.
3. Het inrichten van een datainfrastructuur.
4. Het faciliteren van experimenten van inzet van AI in het onderwijs.
5. Het stimuleren van een multidisciplinaire aanpak bij de ontwikkeling van AI.
6. Het ontwikkelen van een keurmerk voor verantwoorde toepassing van AI in het onderwijs.
7. Het initiëren van vervolgonderzoek naar AI in het onderwijs.

1 Introductie

1.1 Aanleiding

De inzet van technologie in het onderwijs kent een lange traditie. In de jaren tachtig van de vorige eeuw vonden computers hun weg naar het klaslokaal, gevolgd door het internet en (tegenwoordig) allerlei andere apparaten zoals tablets, digiborden en smartphones. Parallel aan deze ontwikkeling speelt, ook buiten de onderwijssector, een toename van de hoeveelheid verzamelde en verwerkte data. [1] Dit leidt steeds meer tot de inzet van intelligente systemen, die patronen kunnen herkennen in grote hoeveelheid data, en steeds beter in staat zijn om menselijk gedrag, en vooral menselijk redeneren, na te bootsen. Deze systemen kunnen daardoor taken zelfstandig uitvoeren, of mensen helpen bij het uitvoeren van taken. Het verzamelbegrip voor deze ontwikkeling is *kunstmatige intelligentie*, *artificial intelligence*, afgekort als AI.

Het ministerie van Onderwijs, Cultuur en Wetenschap heeft Dialogic *innovatie & interactie* gevraagd een onderzoek uit te voeren naar de mogelijkheden van AI in het onderwijs. Het doel van dit onderzoek is het verwerven van inzicht in hoe AI momenteel in het onderwijs wordt gebruikt, en hoe het in de komende 5-8 jaar gebruikt zal gaan worden (*wat kan er?*). Daarnaast onderzoeken we welke juridische aspecten daarbij spelen (*wat mag er?*). Tot slot identificeert het onderzoek de vijf grootste risico's en kansen die verbonden zijn aan toepassing van AI in het onderwijs (*Wat willen we?*).

1.2 Onderzoeksvragen

De doelstelling van het onderzoek is het verwerven van inzicht in hoe AI momenteel in het onderwijs wordt gebruikt en welke toepassingen in de toekomst worden verwacht. Daarbij wordt gevraagd welke juridische aspecten er (mogelijk gaan) spelen bij dit gebruik en welke ethische aspecten van belang zijn. De antwoorden op deze vragen zijn bedoeld om beleidsvorming op het vlak van het gebruik van AI in het onderwijs te ondersteunen. In dit onderzoek is deze onderzoeksvraag als volgt opgesplitst:

1. **Wat kan er?** Binnen deze vraag staat centraal welke toepassingen er nu beschikbaar zijn en hoe de ontwikkelingen AI in het onderwijs naartoe gaan (zowel nationaal als internationaal). Hierbij kijken we specifiek naar voor welke doeleinden AI wordt ingezet en wat dit betekent voor leerling, docent, instelling en het onderwijsstelsel.
2. **Wat mag er?** Om deze onderzoeksvraag te beantwoorden brengen we in kaart welke juridische handelingen er plaatsvinden bij het toepassen van AI in het onderwijs. Vervolgens worden deze handelingen afgezet tegen bestaande juridische kaders. Onder de juridische kaders vallen zowel generieke als onderwijs specifieke kaders in acht genomen. Naast juridische kaders brengen we in deze stap ook de ethische kaders in kaart.
3. **Wat willen we?** In deze onderzoeksvraag staan de kansen en risico's van AI centraal. Daarnaast brengen we in kaart welke beleidsopties er zijn om deze kansen optimaal te benutten en risico's te mitigeren.

1.3 Methode

Ter beantwoording van de onderzoeksvraag is een combinatie van onderzoeksmethoden gehanteerd. De verkenning van toepassingen en gebruik van AI in het onderwijs is grotendeels gebaseerd op literatuurstudie en interviews met diverse betrokkenen binnen en buiten het onderwijsveld. Om de finale set kansen en risico's te bepalen is tot slot een workshop georganiseerd.

Interviews

De set geïnterviewde personen bestond (onder andere) uit academici welke onderzoek uitvoeren op het gebied van AI, onderwijskunde en onderwijsinnovatie. Daarnaast zijn schoolleiders, docenten, en ontwikkelaars van AI-toepassingen in het onderwijs geïnterviewd om een zo compleet mogelijk beeld te scheppen van de mogelijkheden van AI in het onderwijs. Tot slot is er gesproken met ethici.

In dit onderzoek is niet gesproken met leerlingen en studenten. Het is aannemelijk dat acceptatie door leerlingen en studenten een cruciaal aspect is van het daadwerkelijk succesvol toepassen van AI in het onderwijs. Ons inziens is het echter weinig zinvol om leerlingen en studenten te bevragen zonder concrete toepassingen voor te leggen. Omdat het een verkennend onderzoek betreft is ervoor gekozen om personen te spreken met een overzicht over het veld. In dit onderzoek hebben we het leerlingperspectief meegenomen door met voorstaande academici te praten die onderzoeken hebben uitgevoerd met digitale leermiddelen waarin de leerling centraal staat.

De selectie van geïnterviewde personen is deels tot stand gekomen op basis van reputatie en veronderstelde expertise, en deels op basis van *snowball sampling*. Snowball sampling is een methode waarbij deelnemers aan het onderzoek andere deelnemers aandragen voor vervolginterviews. Het voordeel van deze methode is dat lastig te bereiken populaties via referenties toch deelnemen aan het onderzoek. In de bijlage is een lijst met geïnterviewde personen opgenomen.

Literatuurstudie

De analyses ten aanzien van de internationale context, technische (on)mogelijkheden en juridische aspecten zijn voornamelijk gebaseerd op (wetenschappelijke) literatuur. De gehanteerde methode is te beschrijven als *map-and-reduce*: allereerst is zoveel mogelijk materiaal verzameld rondom toepassingen van AI in het onderwijs, waarna (deels op basis van de kennis opgedaan uit interviews) de informatie is gestructureerd en gekaderd.

Voor wat betreft de analyse van de juridische aspecten is een systematische aanpak gehanteerd, waarbij allereerst een duiding van AI in het onderwijs is gevormd. Vervolgens wordt deze uitgewerkt naar juridisch relevante handelingen en actoren, waarna is bepaald welke regelingen daarop van toepassing zijn. Deze inventarisatie leidt vervolgens tot een bespreking (per handeling en regeling) van de juridische implicaties.

1.4 Leeswijzer

Het onderzoeksrapport is gestructureerd volgens de drie onderzoeksvragen. Hoofdstuk 2 gaat in op de vraag "wat kan er?" en beschrijft de gehanteerde definitie van AI en een systematische verkenning van toepassingen van AI in het onderwijsveld. In hoofdstuk 3 gaan we in op de vraag "wat mag er?" door te kijken naar juridische knelpunten die (kunnen gaan) spelen bij inzet van AI in het onderwijs. In hoofdstuk 4 gaan we in op de vraag "wat willen we?", waarbij aandacht uitgaat naar onderwijs- en beleidsdoelen en ethische aspecten van

de inzet van AI in het onderwijs. In hoofdstuk 5 geven we tot slot de belangrijkste conclusies uit dit onderzoek weer, geformuleerd als kansen en risico's van AI in het onderwijs.

In de tekst zijn op diverse plaatsen gearceerde vakken te vinden. Deze bevatten enerzijds achtergrondinformatie (bijvoorbeeld een nadere toelichting op een begrip), en anderzijds bevindingen. Deze bevindingen zijn niet direct antwoorden op de onderzoeksvraag, maar zijn wel relevant gezien het achterliggende vraagstuk.

2 Wat kan er?

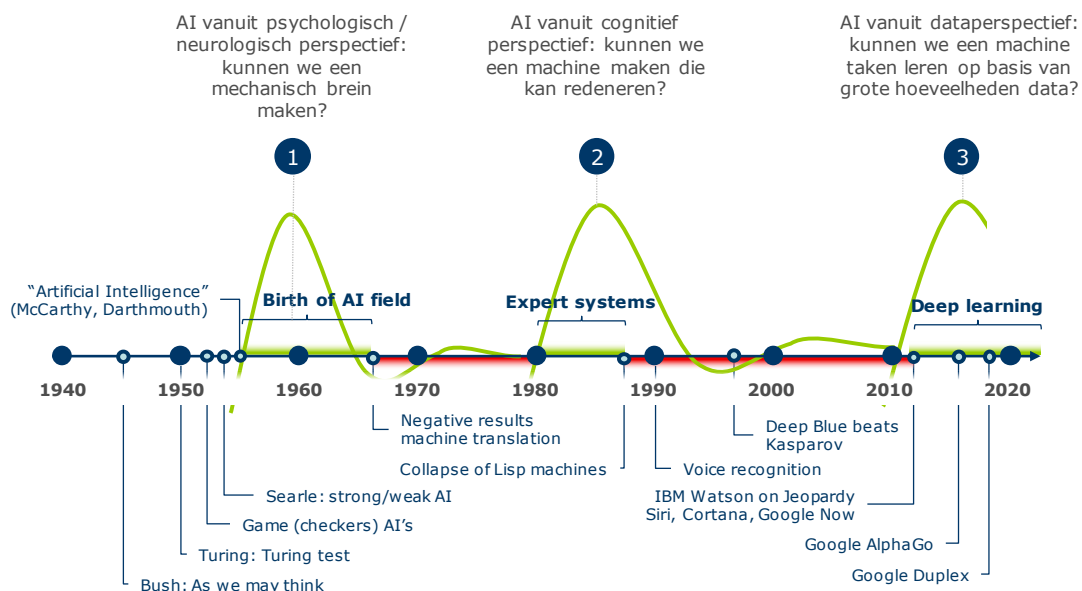
In dit hoofdstuk bespreken we de toepassingen van AI in het onderwijs. Hiervoor dienen we eerst AI goed te duiden. Hiervoor kijken we eerst naar de geschiedenis van AI om vervolgens een werkdefinitie van AI te formuleren. Hierna kijken we naar de ontwikkelaars van AI toepassingen in het onderwijs en specifieke toepassingen die binnen de tijdhorizon worden ontwikkeld. Vervolgens extraheren we verschillende elementen waar deze AI toepassingen aan voldoen. Tot slot geven we een aantal realistische en bruikbare scenario's waarin AI zou kunnen worden toegepast.

2.1 Wat is AI?

Om een goed en compleet begrip van AI te krijgen dient de ontwikkeling van AI in een historische context te worden geplaatst. Op basis van de historische ontwikkeling van geven we een werkdefinitie van AI in het onderwijs. Deze werkdefinitie moet niet te ruim zijn om te voorkomen dat technologieën onterecht als AI worden aangemerkt en niet te smal om ervoor te zorgen dat sommige AI-technologieën buiten de definitie vallen. Deze werkdefinitie is niet normatief maar wordt slechts gebruikt om het onderzoek te kaderen.

2.1.1 Geschiedenis

De geschiedenis van AI gaat verder terug dan menigeen zou vermoeden: de droom van het automatiseren van menselijk gedrag en redeneervermogen is terug te herleiden tot de oudheid. In Hellenistische mythes zien we bronzen automaten zoals Talos, die Kreta beschermde tegen piraten. Ook in de Middeleeuwen zien we ideeën van AI terug in de verhalen rondom de Golem van Praag. Moderne ideeën rondom AI gaan terug naar de opkomst van de computer met sleutelfiguren zoals Turing, Walters en Minsky.

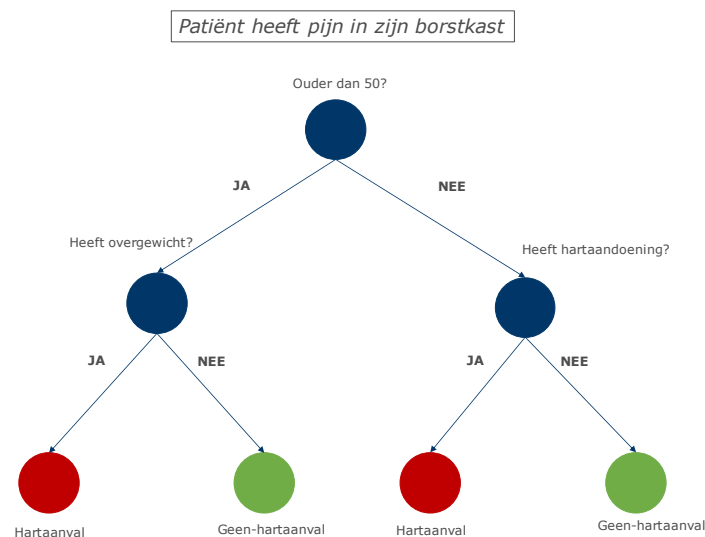


Figuur 1 Overzicht van de geschiedenis van AI en de drie 'waves' die daarin te onderscheiden zijn

De term "kunstmatige intelligentie" werd in 1956 geïntroduceerd door wetenschapper John McCarthy. Het wetenschappelijke veld heeft sindsdien een aantal cycli gekend; hoogtepunten waarin AI sterk werd gehypet, gevolgd door teleurstelling en kritiek. Het AI-veld heeft tot nu

toe drie van dergelijke grote ervaringen gehad. De eerste ervaring speelde rond de jaren 50 en 60, rond de periode dat de term in leven werd geroepen, gedreven door pioniers bij MIT en Stanford. In de jaren 70 werd er echter flink in de onderzoeksbudgetten gesneden, omdat AI in de praktijk niet in staat bleek om (bijvoorbeeld) teksten vanuit het Russisch naar het Engels te vertalen – destijds een toepassing waar vraag naar was en waarvan werd verwacht dat AI deze zou kunnen vervullen. Het klassieke voorbeeld hierbij is de vertaling van “*De geest is gewillig, maar het vlees is zwak*” in het Russisch, naar “*de wodka is goed maar het vlees is bedorven*” in het Engels, door een AI.

In de jaren tachtig zien we Japan sterk inzetten op AI om haar industrie vooruit te helpen. De Verenigde Staten en het Verenigd Koninkrijk volgen snel. In deze periode ligt de nadruk op expertsystemen. Expertsystemen *emuleren* bepaalde specifieke taken/handelingen van mensen. Bij deze systemen was de intelligentie nog compleet met de hand geprogrammeerd en kon de AI nog geen nieuwe taken leren zonder dat een mens de nieuwe regels programmeerde. De systemen zijn het best te beschrijven als een voorgeprogrammeerde ‘beslisboom’ die systematisch in software wordt doorlopen. Figuur 2 toont een voorbeeld van een dergelijke beslisboom – een expertstelsysteem zou uiteraard bestaan uit een veel grotere ‘boom’ en daarmee beslissingen faciliteren.



Figuur 2 De toepassing van een beslisboom op de eerste hulp: heeft iemand een hartaanval bij pijn in de borst?

Doordat AI de betekenis van woorden niet intrinsiek begrijpt, maar slechts werkt op basis van regels, blijken de expertsystemen uiteindelijk niet zo goed te werken als gehoopt. In de jaren 90 neemt de aandacht voor AI daarom wederom sterk af.

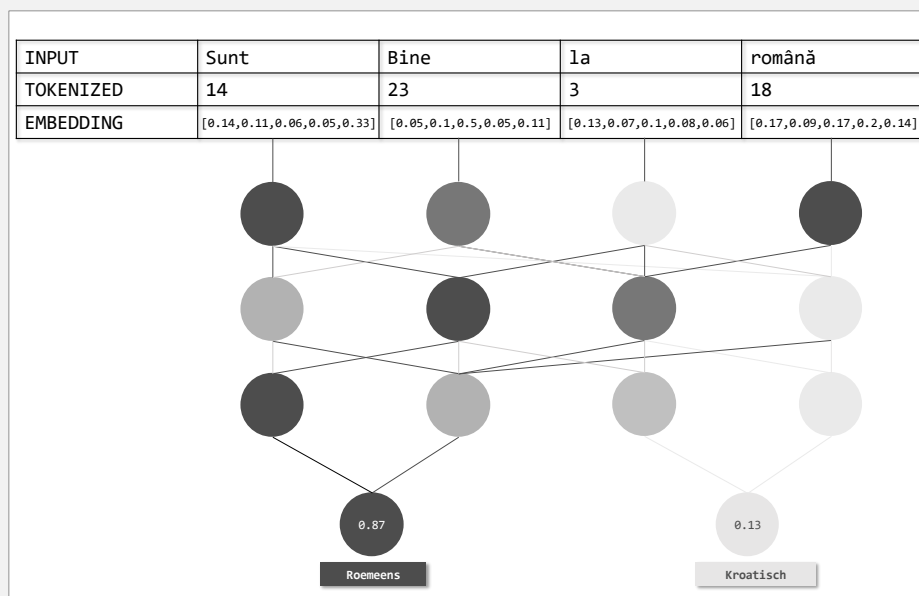
Recentelijk (anno 2011) heeft AI een nieuwe ervaring gehad door inbreng van onderzoekers als Andrew Ng, Geoffrey Hinton en Yann LeCun. Zij ontwikkelden *deep learning* – technieken waarmee een sprong kon worden gemaakt in de intelligentie van algoritmes. AI-toepassingen die tot dan toe onmogelijk werden geacht, bleken ineens haalbaar. Neem bijvoorbeeld het door Google ontwikkelde *AlphaGo*, dat in 2016 de wereldkampioen Lee Sedol in het spel Go versloeg. Tot dan toe werd gedacht dat alleen met menselijke intelligentie (en intuïtie) het spel Go op hoog niveau kon worden gespeeld. Het spel Go kent namelijk 10^{172} mogelijke zetten – een veelvoud van het aantal zetten dat mogelijk is bij schaken. Dit zijn te veel zetten om met traditionele methodes (bijv. met een *minimax tree search*) de meest optimale strategie te bepalen.

Achtergrond 1. Machine learning: de 'magie' achter hedendaagse AI

AI kan op verschillende manieren worden gerealiseerd. In de tijd van de expert systemen was de intelligentie van het systeem met de hand geprogrammeerd. De ontwerper van het expertsysteem moest alle mogelijke handelingen van de AI zelf specificeren. Bijgevolg was de intelligentie van de expertsystemen vrij beperkt: alle situaties waarvoor de ontwerper geen regel had voorzien konden niet door het systeem worden behandeld. Tegenwoordig zijn AI systemen niet *regelgedreven* maar *datagedreven*: de AI-systemen leren zelf de regels uit de data. De onderliggende algoritmes van deze zelflerende systemen worden doorgaans *machine learningalgoritmes* genoemd.

Een speciale klasse van machine learning is *deep learning*. Bij deep learning wordt er iteratief een mapping tussen de input en de output van een model gezocht in de vorm van een serie aan wiskundige transformaties. Deze transformaties zijn geïnspireerd op de manier waarop ons menselijke brein werkt: een zogenaamd 'neuraal netwerk'. In onze hersenen prikkelen onze zintuigen hersencellen – de 'neuronen'. Deze neuronen sturen, afhankelijk van de prikkel, wel of geen signaal naar andere neuronen. Honderden miljarden neuronen leiden samen tot intelligent gedrag.

Figuur 3 toont een eenvoudig voorbeeld van een neuraal netwerk zoals 'gesimuleerd' in een deep learningmodel. De invoer bestaat in dit voorbeeld uit woorden (gecodeerd naar getallen en vervolgens naar 'stimuli' voor het neurale netwerk). Drie lagen 'neuronen' combineren deze stimuli tot een voorspelling, in dit geval of de tekst Roemeens of Kroatisch is.



Figuur 3 Een vereenvoudigde weergave van een neuraal netwerk dat kan bepalen of een zin in het Roemeens of Kroatisch is geschreven.

Een vuistregel die geldt bij machine learning is dat de prestatie van het model verbetert naar mate er meer data beschikbaar is voor het trainen van het model. Als gevolg van steeds verder gaande digitalisering van processen wordt veel data gegenereerd die hiervoor kan worden ingezet. Dit in combinatie met alsmaar toenemende computerkracht en academische doorbraken zorgt ervoor dat de ontwikkelingen op het gebied van machine learning exponentieel groeien.

Deep learning heeft buiten het onderwijs tal van toepassingen die menselijk presteren benaderen of zelf verbeteren. In de medische wereld wordt deep learning gebruikt voor

het diagnosticeren van huidkanker en het classificeren van CT scans. In zelfrijdende auto's wordt deep learning gebruikt om op basis van camera's de richting en snelheid van de auto te bepalen. In de strijd tegen neopnieuws zet onder andere Facebook deep learning in om de authenticiteit van teksten te beoordelen.

Wat is het summum van AI? In *science fiction* zien we toekomstbeelden waarin machines zich gedragen en voordoen als mensen, en in veel gevallen zelfs over superieure intelligentie beschikken. Een niveau van AI met een intelligentie gelijk aan die van mensen wordt *artificial general intelligence* of AGI genoemd; is de AI slimmer dan de mens, dan spreken we van *artificial superintelligence* of ASI. Futurologen beweren dat AGI en met name ASI wel eens de laatste uitvinding van de mens kunnen zijn. Een AGI of ASI zou zelfs een bedreiging voor de mensheid kunnen zijn, daar we als mens de superieure intelligentie van een dergelijke AI niet meer kunnen bijhouden. [2]

2.1.2 Werkdefinitie

Zoals hierboven beschreven kent AI een rijke historie en heeft het veld zich meermaals opnieuw "uitgevonden". De definitie van AI in de praktijk lijkt vooral gedreven door wat gezien de stand van de techniek mogelijk is. Tegelijkertijd is het doel van dit onderzoek ook om vooruit te kijken: wat zullen de mogelijkheden van AI zijn binnen de komende vijf tot acht jaar? Een goede werkdefinitie van AI is daarom in het kader van dit onderzoek cruciaal. Er zijn verschillende 'aanvliegroutes' om een dergelijke definitie te formuleren.

Toepassingsmogelijkheden van AI: het automatiseren van cognitieve taken

In deze definitie wordt AI gedefinieerd door *wat* het kan en niet *hoe*. Onder deze definitie is technologie agnostisch: het maakt niet uit *hoe* het systeem de taak automatiseert. Of dit gebeurt met de hand gecodeerde regels of aan de hand van zelflerende algoritmes is in deze definitie niet van belang. Als kritiek op deze definitie wordt vaak gesteld dat deze te breed is: een simpele rekenmachine kan ook worden gezien als het automatiseren van cognitieve taken (namelijk rekenen) terwijl intuïtief gezien een rekenmachine niet als 'intelligent' wordt bestempeld.

Onderliggende technologie van AI: het exploiteren van grote hoeveelheden data

Vanuit dit perspectief wordt AI gedefinieerd in termen van de achterliggende methode. In tegenstelling tot het 'automatiseren van cognitieve taken' als definitie, is deze definitie niet agnostisch ten opzichte van de gebruikte technologie. Niet alleen voorspellende statistiek maar ook beschrijvende statistiek, gevisualiseerd in bijvoorbeeld dashboard, wordt in deze definitie als AI bestempeld. Expertsystemen worden in deze definitie echter *niet* gezien als kunstmatig intelligent.

Ten opzichte van de mens: alles wat een computer nog niet kan

Deze definitie bekijkt AI vanuit de huidige tijdsgeest. AI betreft alle vormen van (menselijke) intelligentie die op dit moment niet kunnen worden gerealiseerd door een computer te programmeren. Voorstanders van deze definitie stellen dat wanneer een cognitieve taak wordt geautomatiseerd, dit vaak resulteert in golf van pessimisme. Zo werd voorheen gedacht dat een computer nooit een mens zou kunnen verslaan in schaken. Nadat schaakcomputer Deep Blue Kasparov versloeg werd er direct gesteld dat de methode van Deep Blue niet *echte* intelligentie betrof, maar louter *brute force*. Het gehanteerde algoritme loopt immers simpelweg alle mogelijke zetten langs, en kiest daaruit de vervolgzet die, gezien alle gesimuleerde spelverlopen, de hoogste winkans oplevert. Vervolgens wordt er een nieuwe

benchmark gesteld: er kan nu pas worden gesproken van een 'echte' AI wanneer een computer deze nieuwe taak kan uitvoeren.

Eenzelfde trend zien we bij de hype rondom deep learning. Hoewel met deep learning tal van cognitieve taken kunnen geautomatiseerd en deze methoden (lichtelijk) geïnspireerd zijn door de werking van het brein, kunnen deze systemen niet 'denken zoals mensen'. De verschillende toepassingen van deep learning zijn daarmee wederom gereduceerd tot "niet AI" terwijl men voorafgaand aan de introductie van deep learning zou hebben gezegd dat zaken als beeldherkenning en het vertalen van teksten wel degelijk zaken zijn waarvoor AI nodig zou zijn.

Een algemene opmerking bij bovenstaande is dat moet worden opgepast met antropomorfisme (het toekennen van menselijke eigenschappen aan niet-menselijke objecten). Een AI 'denkt' in veel gevallen niet als een mens, maar voor een gebruiker kan dit wel zo lijken, en het ligt dan ook voor de hand om AI's langs dezelfde prestatie-meetlatten te leggen als een mens die dezelfde taak zou uitvoeren.

Gehanteerde definitie

In deze rapportage wordt AI gedefinieerd als: "*het automatiseren van cognitieve taken in het onderwijs aan de hand van datagedreven methodes*". Deze definitie leent elementen uit elk van de hierboven besproken definities. Uit de 1^e definitie nemen we de *wat* (cognitieve taken) mee; uit de 2^e definitie de *hoe* (statistische algoritmes op grote hoeveelheden data). We kijken daarnaast naar toepassingen die binnen de komende 5 tot 8 jaar mogelijk zijn. Hierdoor vallen zaken als *artificial general intelligence* buiten de scope van het onderzoek en gaat het in feite om AI zoals nu mogelijk gemaakt door recente ontwikkelingen op het vlak van deep learning.

Bevinding 1. Er is niet één definitie van AI te geven, maar de vraagstukken in het onderwijs bevinden zich op het specifieke snijvlak van algoritmes en data.

Het is niet eenvoudig om één definitie te geven van kunstmatige intelligentie, specifiek als het gaat om toepassingen in het onderwijs. Uit ons onderzoek komt naar voren dat met name wanneer het gaat om toepassingen die (1) cognitieve taken automatiseren en (2) daarbij gebruik maken van grote hoeveelheden data en datagedreven methoden, er interessante, onopgeloste vraagstukken bestaan.

2.2 Huidige en toekomstige toepassingen

Er zijn diverse soorten organisaties actief die toepassingen voor het onderwijs ontwikkelen op basis van AI. Enerzijds zien we dat sommige uitgevers een AI-laag toevoegen aan digitale leermiddelen. Daarnaast experimenteren diverse startups met toepassingen van AI in het onderwijs. De grote techgiganten, zoals Google en Facebook, zijn dominante spelers in ontwikkeling van AI,¹ maar zien we nog beperkt op de Nederlandse onderwijsmarkt acteren. Dit komt waarschijnlijk omdat AI in het onderwijs specifieke contextuele data vereist (in dit geval van Nederlandse leerlingen) die niet eenvoudig beschikbaar is.

Academisch onderzoek op het gebied van AI is niet noodzakelijkerwijs fundamenteel van aard. Uit experimenten wordt ook praktische kennis opgedaan door te experimenteren met verschillende toepassingen van AI in het onderwijs. Ook zien we dat scholen en docenten zelf zijn gaan experimenteren met AI. IT-leerkrachten die ervaring met AI hebben passen

¹ Veel bekende systemen voor machine learning zijn door Google en Facebook ontwikkeld. Ook worden tal van bekende AI-onderzoekers door deze bedrijven ingehuurd om AI-onderzoek uit te voeren.

dit toe om het werk te vereenvoudigen. Zij-instromers die ervaring hebben opgedaan met AI in het bedrijfsleven kunnen deze skills het onderwijs in brengen.

Op hoofdlijnen kan een AI in het onderwijs worden ingezet bij het overbrengen van kennis, het toetsen van kennis, bij het informeren van docenten van prestaties leerlingen. Hieronder geven we een overzicht van toepassingen van AI in het onderwijs zoals geïdentificeerd (en geclusterd) tijdens het onderzoek.

Adaptieve leersystemen

Een adaptief leersysteem is een digitaal leermiddel dat zich zo goed als mogelijk aanpast aan de leerling, zodat het leerproces beter verloopt, en/of de uitkomsten hiervan beter worden. Om dit te realiseren bestaat een adaptief leersysteem uit de volgende componenten:

1. **Expertmodel.** Dit is een model dat alle informatie bezit welke uiteindelijk aan de leerling zal moeten worden overgedragen. Dit model is ook op de hoogte van de moeilijkheidsgraad van de informatie, en de hiërarchische structuur tussen de concepten.
2. **Studentmodel.** Dit model houdt bij wat de leerling al aan informatie heeft geleerd en is op welk niveau de leerling is. In feite vindt hier een modellering plaats van de wijze waarop een leerling leert.
3. **Instructiemodel.** Dit model is gespecialiseerd in onderwijsmethoden. Het model weet wat de beste manier van feedback geven is aan een specifieke leerling, en wanneer deze bepaalde didactische strategieën zoals *productive failure* moet toepassen (eerst de student zelf laten proberen voordat het goede antwoord wordt gegeven). Het model combineert daarbij informatie uit het expertmodel en het studentmodel.

Bovengenoemde modellen kunnen op basis van zowel relatief eenvoudige statistische methoden (zoals Bayesiaanse statistiek) als op basis van geavanceerde deep learningtechnieken worden geïmplementeerd.

De drie componenten worden uiteindelijk samengebracht in een *user interface* waar de student mee interacteert. Het ontwerpen van een adaptief leersysteem is niet eenvoudig. Bij een adaptief leersysteem zijn er meerdere doelen die tegelijkertijd moeten worden geoptimaliseerd. Soms zijn deze doelen conflicterend, en soms is er een trade-off tussen de doelen nodig. Een adaptief leersysteem kan bijvoorbeeld als doel hebben om de student zo snel mogelijk zo complex mogelijke kennis aan te leren. Maar hoe ervaart de leerling dit? Als de vragen te makkelijk zijn, leert de leerling niet snel genoeg, maar als de vragen te moeilijk zijn raakt de student gedemotiveerd, wat op lange termijn weer kan leiden tot een vermindering in prestaties.

Daarnaast kan een adaptief leersysteem 'op hol slaan'. Het kan bijvoorbeeld voorkomen dat een leerling over het algemeen goed is in economie, maar heeft moeite met het specifieke onderwerp inflatie. Een (slecht ontworpen) adaptief leersysteem zou deze leerling na verloop van tijd bijna alleen maar inflatie-oefeningen voor kunnen schotelen. Op de toets worden echter alle onderwerpen bevraagd, en heeft de student niet goed genoeg geoefend voor de andere (non-inflatie) onderdelen.

Knowledge tracing

Knowledge tracing is de methode die binnen adaptieve leersystemen wordt gebruikt om bij te houden hoe goed de leerling de stof beheerst. Het systeem probeert het kennisniveau van een leerling te modelleren, en werkt dit model continu bij, naarmate de leerling meer oefent

in de leeromgeving. Een knowledge tracing-model wordt meestal gerealiseerd aan de hand van Bayesiaanse statistiek (*Bayesian Knowledge Tracing*). In het Bayesiaanse model wordt de kennis bepaald op basis van de volgende 4 parameters:

1. De kans dat de leerling de kennis bezit voordat deze begint met oefenen
2. De kans dat de leerling de kennis demonstreert nadat deze heeft kunnen oefenen
3. De kans dat de leerling een fout maakt terwijl deze de stof wel kent
4. De kans dat de leerling het goede antwoord geeft terwijl deze de stof niet kent.

Vervolgens kan aan de hand van de bovenstaande 4 parameters worden berekend of de leerling de kennis heeft overgenomen en goed kan toepassen. Ook deep learningmodellen kunnen worden ingezet voor knowledge tracing (*Deep Knowledge Tracing*). In een paper gepubliceerd door onderzoekers aan de universiteit van Stanford is gebleken dat een Deep Knowledge Tracing-model beter is in het modelleren van de kennis van de student, maar (veel) minder inzichtelijk is ten opzichte van het Bayesiaanse model. [3] Iets dergelijks geldt in algemene zin voor deep learning-gebaseerde modellen ten opzichte van meer conventionele statistische methoden: een deep learningmodel 'werkt', maar het is niet evident waaróm het werkt. Bij een eenvoudiger statistisch model is de weg van inputparameter tot output te traceren en te begrijpen, bij deep learningmodellen niet meer.

Intelligent tutoring systems

Een *intelligent tutoring system* (ITS) is een systeem dat gepersonaliseerde feedback geeft gedurende het leren. Een ITS is vaak onderdeel van een adaptief leersysteem (het instructiemodel). Een ITS neemt informatie uit het expertmodel en het studentmodel om zo bij te houden welke instructie het meest succesvol is geweest bij het overbrengen van de kennis. Wanneer een student geen voortgang meer lijkt te boeken kies het ITS het de meest geschikte uitleg voor de student.

Automated essay scoring

Automated essay scoring (AES) is een methode om geschreven teksten zoals werkstukken en essays te beoordelen zonder tussenkomst van een mens. AES maakt gebruik van *natural language processing* om een geschreven tekst te beoordelen op inhoud en schrijfstijl. Bij automated essay scoring wordt eerst een model opgesteld op basis van eerder, met de hand, beoordeelde teksten. Vervolgens worden machine learning methodes bepaald in hoeverre nieuwe teksten overeenkomen met, of elementen bevatten van, de teksten die zijn gebruikt bij het opstellen van het model.

Analytical dashboards

Naarmate de adoptie van digitale leermiddelen in het onderwijs toeneemt zien we ook dat er steeds meer data wordt gegenereerd door deze systemen. Deze data kan worden gebruikt om de docent te informeren over de prestaties van leerlingen. Vaak gebeurt dit in zogenaamde analytische dashboards. Deze dashboards visualiseren de voortgang en prestaties van leerlingen. Denk hierbij aan taartgrafieken die aangeven hoeveel procent van de leerlingen een bepaalde vaardigheid al beheerst, of een lijngrafiek die de prestaties van een leerling over tijd weergeeft.

Hoewel deze beschrijvende statistiek op zichzelf niet noodzakelijk erg 'intelligent' is, zien we dat steeds vaker gebruik wordt gemaakt van *predictive modelling* – het voorspellen van de indicatoren. Denk hierbij aan het koppelen van informatie uit een leerlingvolgsysteem met persoonlijkheidskenmerken om zo een voorspellend model te maken van het presteren van een leerling binnen en buiten de klas.

2.3 Ontwikkeling in het buitenland

In deze paragraaf bespreken we relevante AI ontwikkelingen in het buitenland. Deze ontwikkelingen zijn op twee manieren geïdentificeerd. Enerzijds is bottom-up, vanuit de technologie, gekeken wat er gebeurt op het gebied van AI en onderwijs. Denk hierbij aan toepassingen die aan de orde komen in bijlagen aan vooraanstaande conferenties zoals Artificial Intelligence in Education Conference (AIED), of de International Conference on Intelligent Tutoring Systems (ITS).

Anderzijds is er 'top-down', vanuit een selectie van Europese landen, gekeken welk ontwikkelingen er plaatsvinden op het gebied van AI in het onderwijs. Voor deze landen geldt dat de beleidscontext waarin deze technologieën worden ontwikkeld nog tot zekere hoogte te vergelijken met Nederlandse context. Ook niet-Europese landen met een uitgesproken AI-visie zoals China, de VS en Canada zijn meegenomen in de internationale beschouwing.

Op het gebied van specifieke applicaties zien we enerzijds eenzelfde soort trend als in Nederland. Zo worden er in Zweden en het VK adaptieve leersystemen ontwikkeld door bedrijven zoals Sana Labs, Knewton en Kwiziq. In Frankrijk zien we een substantiële markt in EdTech. Het Franse Observatoire EdTech identificeert 384 Franse bedrijven die zich bezig houden met technologische innovaties in het onderwijs [4]. Hiervan zijn niet alle start ups bezig met AI maar we zien dat ruim 29% procent van de Franse EdTech bedrijven zich bezighoudt met adaptief leren (Tabel 1).

Tabel 1 Overzicht Franse EdTech markt. Eén toepassing kan meerdere technologieën incorporeren.

Tehnologieën Franse EdTech	Aantal	Percentage
Innovatieve inhoud	202	53%
Serieuze game / leer-apps	139	36%
Adaptief leren	112	29%
Digitaal leeraanbod (MOOCS)	109	28%
Learning Management System / CMS	99	26%
Social media / samenwerkingstools	89	23%
Aggregatieplatform en distributie van diensten	73	19%
Technologies de back / middle office, verbindingen	69	18%
Blended learning	70	18%
Ondersteunt / apparatuur	34	9%
Virtuele / toegevoegde realiteit	32	8%
IoT / digitale schoolplaatsen	22	6%
anders	23	6%
Totaal	384	100%

Classroom monitoring

In China lijken de toepassingen van AI onbegrensd. We zien hier dat AI niet alleen wordt ingezet bij het leren en het toetsen, maar ook dat leerlingen worden gemonitord door een AI die aan de hand van camera's bijhoudt welke leerling aan het opletten is [5] [6]. Op basis van deze informatie kan een leerling worden aangesproken, of kan de docent andere methode inzetten om het aandacht niveau te verhogen. Ook in Europese landen zien we aandacht voor en onderzoek naar dergelijke technologie. Zo zien we dat in Zwitserland en

Frankrijk er wordt geëxperimenteerd met technologieën die studenten analyseert met camera's om real-time feedback te geven aan de docent.

In Europa zien we dat *classroom monitoring*-technieken vooral worden toegepast in het kader van academisch onderzoek. In China worden de technologieën al op grote schaal geïmplementeerd. Merk op dat toepassing van de systemen in klaslokalen in China de aandacht in het klaslokaal niet lijken te verbeteren; gevonden verbeteringen zijn voornamelijk toe te schrijven aan effecten zoals omschreven in Jeremy Bentham's *Panopticon*. Een panopticon is een gevangenis waarbij sociale controle wordt uitgeoefend door de gevangenen het gevoel te geven constant in de gaten te worden gehouden. Een reactie van een Chinese student spreekt boekdelen: "*Ik durf niet meer afgeleid te zijn nadat de camera's in het klaslokaal zijn geïnstalleerd. Het is alsof een paar mysterieuze ogen me constant in de gaten houden*" [7]

Augmented reality, virtual reality & simulaties

In Duitsland zien we onderzoek waarin AI wordt geïntegreerd met *augmented reality* door middel van *smart glasses* zodat de leerling direct instructies voor zijn ogen te zien krijgt. Duitse onderzoekers voeren onder andere onderzoek uit naar het integreren van augmented reality met intelligent tutoring systems. Ook in Canada wordt er ook onderzoek gedaan naar hoe simulaties en virtual reality gebruikt kunnen worden in combinatie met AI om de leerervaring te verrijken. In deze simulatie omgevingen krijgen leerlingen praktische instructie van een AI waarna ze meteen aan de slag kunnen met een probleem.



Figuur 4 Met Augmented reality en AI instructies krijgen over hoe schakelaars werken (Deutsches Forschungszentrum für Künstliche Intelligenz)

Teachable robots

Een andere toepassing van AI is om met robots nieuwe didactische methoden te realiseren. Zo worden er nu AI-robots ontwikkeld die door leerlingen nieuwe informatie of vaardigheid moeten worden aangeleerd. Dankzij het *protegé effect* leren de leerlingen door de lesstof uit te leggen deze zelf ook beter te begrijpen. [8] Het protegé effect veronderstelt dat studenten sneller kennis eigen maken wanneer de student het geleerde kan uitleggen aan een andere student. Bewijs voor het protegé effect zien we ook buiten het klaslokaal: Leerlingen met jongere broers of zussen hebben vaak een hoger IQ dan leerlingen die dat niet hebben [9]. Een teachable robot kan 'oneindig' uitleg krijgen, terwijl een mens wellicht geneigd is om net te doen of deze het uitgelegde begrijpt om geen reputatieschade te leiden.



Figuur 5 Een teachable robot ontwikkeld door onderzoekers EPFL [10]

2.4 Eigenschappen van AI-toepassingen

De ene toepassing van AI in het onderwijs is de andere niet – dat volgt ten minste uit de hierboven benoemde voorbeelden. Om beleid te kunnen vormen ten aanzien van AI-toepassingen in het onderwijs is het dus belangrijk om te analyseren welke aspecten van AI-toepassingen relevant en onderscheidend zijn: hoe zijn de toepassingen systematisch in te delen? Hieronder geven we de belangrijkste onderscheidende aspecten.

Didactisch versus pedagogisch

Veel AI-toepassingen in het onderwijs zijn gericht op het automatiseren van didactische activiteiten. Denk hierbij aan het selecteren van lesstof (typisch in adaptieve leersystemen), het nakijken van toetsen (zoals automated essay scoring), en het geven van constructieve feedback aan leerlingen (intelligent tutoring systems). Daarnaast wordt van docenten en leerkrachten niet alleen verwacht dat zij puur en alleen kennis overdragen. Ook verwachten we dat leerlingen leren samenwerken, zelfstandigheid, normen en waarden en andere niet didactische competenties. Een AI kan ook voor deze doeleinden worden ingezet. Voor een groepsopdracht kan een AI op basis van de sterktes en zwaktes van individuele leerlingen bijvoorbeeld de meest diverse selectie van studenten maken. [11]

Specifiek of generiek

Een AI kan worden ingezet om kennis van bepaald vak over te dragen of een specifieke competentie aan te leren. Deze specifieke AI toepassingen worden al ingezet om bijvoorbeeld een leerling te laten oefenen met rekenen. Een vakoverstijgende AI toepassing is niet alleen bedoeld om één specifieke competentie aan te leren maar neemt het gehele presteren van de leerling mee over alle vakken. Zo kan deze AI niet alleen de inhoud van één bepaald vak voor een leerling personaliseren maar ook het gehele curriculum. Waar op het PO de leerkracht al in zekere mate op een vakoverstijgende manier bepaald welke leerling waar aan werkt is dit op het VO en HO minder eenvoudig. Op het VO en HO zien we dat het voor docenten minder eenvoudig is om een goed beeld te hebben van wat er buiten de eigen lessen om gebeurt. Om dit op te vangen worden docenten op het VO omgedoopt tot mentor om deze vakoverstijgende taken uit te voeren. Een AI kan deze mentor rol ondersteunen door het presteren van een leerling over de vakken heen volgen en waar nodig te signaleren.

Vaste leerjaren en niveaus versus een continuüm

In traditioneel onderwijs worden de onderwerpen en de moeilijkheidsgraad van de lesstof bepaald door het leerjaar en niveau waar de leerling zich bevindt. De leerling wordt pas toegelaten tot een bepaald leerjaar of niveau als deze over een selectie van vakken een

minimale prestatie heeft gehaald. Zo wordt op het VWO van een leerling verwacht dat deze zowel wiskunde als een vreemde taal op (minstens) op hetzelfde niveau beheerst. Wanneer een AI wordt ingezet in de vorm van een adaptief leersysteem kan dit de praktische noodzaak voor het onderscheiden van leerjaren en niveaus overbodig maken. Een AI-toepassing kan het mogelijk maken dat een leerling het ene vak op een geavanceerder niveau beheerst dan een ander vak. Een AI-toepassing maakt het zo mogelijk om te certificeren per vak, in plaats van certificering over een selectie van vakken.

Autonoom versus ondersteunend

AI-toepassingen in het onderwijs verschillen sterk in de mate van autonomie waarover zij beschikken. In het onderwijs zien we grofweg drie niveaus van autonomie van AI-toepassingen:

1. De AI maakt prestaties van leerling inzichtelijk aan de docent
2. De AI interpreteert en geeft informatie over prestaties van leerlingen aan de docent
3. De AI geeft rechtstreeks aanbevelingen aan leerling (en wellicht ouder) zonder tussenkomst van de docent.

Wanneer een AI ingezet wordt om te signaleren en informeren heeft deze een relatief lage mate van autonomie. De gebruiker (de docent) kan in deze situatie zelf bepalen of deze de informatie gebruik of naast zich neerlegt. Ook kan de gebruiker in deze situatie zelf bepalen hoe er moet worden gehandeld op de informatie. Denk bij deze informerende AI bijvoorbeeld aan een dashboard met geavanceerde patroonherkenning, die het presteren van leerlingen inzichtelijk maakt. In een adaptief leersysteem is de AI in zekere zin al autonoom. Een adaptief leersysteem kan worden ingezet om te laten bepalen: welke leerling, op welk moment, welk lesmateriaal, op welke manier krijg. In de toekomst zou men nog autonomere systemen kunnen tegenkomen die zonder tussenkomst van een mens kunnen handelen.

Daarnaast speelt er ook nog de vorm waarop de AI interacteert met mensen. Bij het inzichtelijk maken van prestaties worden er keuzen gemaakt. Hoe wordt de informatie gepresenteerd, hoe wordt het presteren van de leerling inzichtelijk gemaakt (over tijd of tussen leerlingen?). Daarnaast veronderstelt dit ook dat een docent de kennis en kunde heeft om deze gegevens correct te interpreteren. Denk hierbij aan hoe kansen moeten worden geïnterpreteerd, wat een betrouwbaarheidsinterval is, wat er wordt bedoeld met statistische significantie, wat een type 1 en 2 fout zijn, enzovoorts. Ook op het gebied van informeren zijn er verschillende afwegingen die kunnen worden gemaakt. Er moet worden nagedacht over hoe informatie wordt *geframed*.

Bevinding 2. Framing van informatie en uitkomsten maakt of breekt AI in het onderwijs

De manier waarop mensen informatie interpreteren is afhankelijk van hoe de informatie wordt gepresenteerd. Dit effect wordt framing genoemd. Ook informatie die uit een AI komt veroorzaakt framingeffecten. Neem bijvoorbeeld een AI die een docent informeert over de voortgang van leerlingen binnen een bepaald vak. Stel er is een leerling die moeite heeft met het onderwerp, maar wel flinke stappen vooruit maakt. De AI kan in zo een situatie op verschillende manieren de docent informeren over het presteren van deze leerling:

1. Leerling X presteert 10% onder het gemiddelde van de rest de klas.
2. Leerling X is sinds het begin van oefenen 50% beter geworden.

Beide uitspraken zijn in dit geval waar maar de eerste vorm van informeren heeft een negatievere ondertoon. Bij het ontwerp van een AI moet op voorhand worden nagedacht

over welke informatie het meest nuttig is voor de docent en waarbij een leerling het meeste tot zijn recht komt.

Nieuwe versus traditionele onderwijsvormen

Een AI-toepassing kan zowel worden ingezet om onderwijs in de huidige (veelal klassikale) werkvorm te ondersteunen, maar ook om nieuwe werkvormen te faciliteren.

Uiteraard kunnen de digitale leersystemen die vandaag de dag worden gebruikt vaak al zowel binnen als buiten het klaslokaal worden gebruikt. Ook andere toepassingen die aansluiten bij de huidige manier van werken zijn denkbaar. AI kan bijvoorbeeld worden gebruikt om binnen administratieve systemen verzuim te identificeren. Ook kan AI worden ingezet om roosters te optimaliseren.

Een stap verder richting nieuwe werkvormen gaat *blended learning*. Bij blended learning wordt traditioneel onderwijs gecombineerd met digitale leermiddelen. AI wordt hierbij niet alleen ingezet in de digitale leermiddelen maar om het klaslokaal te monitoren met allerlei sensoren virtual reality en gamificatie kunnen worden ingezet om de aandacht en motivatie van de student te verhogen. Ook VR kan buiten het klaslokaal worden ingezet. Met VR kunnen ook meer praktische vaardigheden worden aangeleerd dan met bijvoorbeeld een adaptief leersysteem.

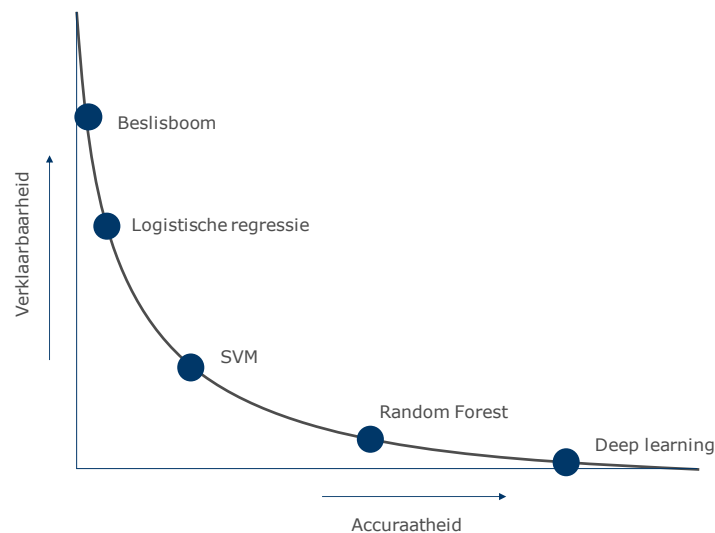
Tot slot zou AI leren op afstand kunnen verbeteren. Met de opkomst van Massive Online Open Courses (MOOC's) worden, door de digitale aard van het materiaal, enorme hoeveelheden data verzameld. Deze data leent zich uitstekend als input voor een AI. Binnen MOOC's worden verschillende soorten gegevens verzameld over studenten:

1. **Gegevens over het leerproces:** terwijl studenten bezig zijn met studeren worden er allerlei gegevens verzameld zoals: welke oefening de student hiervoor maakte, hoe lang de student aan een oefening besteedt, en welke uitleg de student hierbij kreeg. Deze gegevens kunnen als input dienen van een adaptief leersysteem.
2. **Gegevens over de student:** Denk hierbij aan persoonskenmerken zoals geslacht, land van herkomst, en leeftijd maar ook welke andere MOOC's de student op het platform gevolgd heeft.
3. **Koppeling met externe bronnen:** Voor internet gebaseerde diensten zijn koppelingen met andere systemen vaak eenvoudiger te realiseren. Veel MOOC platformen bieden de mogelijkheid om in te loggen met een facebook account (zie Figuur 6). Wanneer dit heeft plaatsgevonden kan het MOOC platform zijn data bereiken met social-media gegevens. Zo kan binnen een MOOC data incorporeren die buiten het platform om wordt gegenereerd.

Figuur 6 Mogelijkheid om in te loggen met een Facebookaccount bij drie grote MOOC platformen (van links naar rechts: Udemy, Coursera, Kahn Academy)

Accuraatheid & verklaarbaarheid

Bij AI-toepassingen is in feite altijd sprake van een voorspellend model. Daarbij is een voorspelling vrijwel nooit 100% accuraat (waarmee uiteraard niet is gezegd dat een mens het beter zou doen!). Daarbij speelt veelal een *trade-off* tussen de accuraatheid van een model en de inzichtelijkheid van een model. Complexere AI-algoritmes (met name de recent beschikbaar gekomen *deep learning* algoritmes) zijn lastiger te begrijpen, maar veel effectiever. Figuur 2 toont de verhouding tussen verklaarbaarheid en accuraatheid van een aantal populaire algoritmes voor machine learning.



Figuur 7 Verhouding tussen verklaarbaarheid en accuraatheid en de positie van enkele populaire machine learning algoritmes op deze curve.

Naarmate machine learning in toenemende mate wordt ingezet om menselijke taken te automatiseren wordt in toenemende mate verklaarbaarheid van deze modellen verwacht. Vragen zoals: *waar is een model goed in?*, *waar is een model slecht in?* en *hoe komt een*

model tot een voorspelling? moeten kunnen worden beantwoord als een model in de praktijk wordt ingezet.

Dit heeft ervoor gezorgd dat in recente jaren methoden zijn ontwikkeld om complexe machine learning modellen inzichtelijk te maken. Zo kan met Shapley Additive Explanation (SHAP) worden berekend hoe een verandering in de input de output van een model beïnvloed. [12] Op deze manier kan zelfs voor de meest geavanceerde modellen aangetoond worden welke eigenschappen het model meeneemt in een besluit en op welke wijze. Daarnaast is de SHAP-methode toepasbaar op vrijwel elk algoritme, en is hiermee een geünificeerde methode om de voorspellingen van een model te begrijpen.

Op het gebied van convolutionele neurale netwerken zijn flinke vooruitgangen geboekt in het inzichtelijk maken van deze modellen. Zo kunnen met *feature visualization* de individuele lagen van een neurale netwerk inzichtelijk worden gemaakt. Over het algemeen blijkt hieruit dat in de eerdere lagen van het netwerk simpele geometrische patronen worden herkend zoals randen en hoeken. Naarmate we verder in het netwerk kijken zien we complexere combinaties van deze geometrische patronen ontstaan. Elke laag van het netwerk combineert de patronen uit de eerdere lagen om zo complexere eigenschappen (features) te creëren. Figuur 8 toont deze opbouw van complexiteit van het neurale netwerk.



Figuur 8 Feature visualisation binnen de lagen van een convolutioneel neurale netwerk. Links: de 'eerste' lagen van het netwerk. Rechts: de 'laatste' lagen van het netwerk [13]

Modulariteit en portabiliteit

Kenmerkend aan veel AI-toepassingen is dat data wordt geïntegreerd uit verschillende bronnen. Dat maakt dat aanbieders geneigd zijn om totaaloplossingen aan te bieden, waarbij AI-toepassing en dataset geïntegreerd zijn. Het koppelen van verschillende systemen die in het onderwijsveld worden gebruikt, en het uitwisselen van data daartussen, is (ook buiten AI-toepassingen) over het algemeen lastig.

Op landelijk niveau zijn initiatieven opgestart tot standaardisatie van verschillende aspecten van (voornamelijk) digitale leersystemen. Dergelijke standaarden faciliteren bijvoorbeeld het bijhouden van één "leerlingidentiteit" tussen verschillende systemen. Op vlak van AI zien we nog nauwelijks dergelijke standaardisatie en veel fragmentatie.

De discussie omtrent modulariteit en fragmentatie vindt ook plaats in de bredere context omtrent (digitale) leermiddelen. Recent heeft een grote scholengroep, teneinde minder afhankelijk te worden van de markt, nieuw inkoopmodel ontwikkeld. Dit model stelt de scholen in staat om digitale leermiddelen en, ter ondersteuning, regulier lesmateriaal direct bij de uitgeverijen in te kopen. Dit wijkt af van de gangbare procedure waarbij lesmateriaal, digitaal en regulier, in een pakket bij een distributeur worden aangeschaft. Dit leidde tot weerstand van de distributeurs, welke een rechtszaak aanspanden tegen de scholengroep, en verloren. [14] De relevantie voor AI in het onderwijs is dat deze als gevolg van de uitspraak eenvoudiger modulaair (door verschillende partijen) kan worden aangeboden.

2.5 Conclusie: toepassingsscenario's in het onderwijs

In het onderwijsproces maken leraren continu naar eigen inzicht beslissingen ten aanzien van (onder andere) de gehanteerde methode, de lesstof, de manier waarop een leerling wordt benaderd, et cetera. Uiteindelijk worden door leraren ook enkele formele beslissingen gemaakt: wat is het rapportcijfer, en mag een kind over naar de volgende klas? AI's kunnen docenten op verschillende manieren ondersteunen bij deze beslissingen. We onderscheiden een viertal scenario's als meest waarschijnlijk voor de komende 5-8 jaar: (1) een AI als onderwijsassistent, (2) AI voor learning analytics, (3) AI voor personalisering van het onderwijs, en (4) AI voor toetsing.

Een AI als onderwijsassistent

Op basis van AI en recente ontwikkelingen op het gebied van *computer vision* en *natural language processing* zouden bepaalde eenvoudige, administratieve taken van docenten kunnen worden uitgevoerd door een AI. Op hoofdlijnen zien we daarbij de volgende drie toepassingen:

1. *Metadatering van content.* Met slimme algoritmes wordt (digitaal) leermateriaal gecategoriseerd, en wordt de docent ondersteund bij het selecteren van lesmateriaal. De docent is daarnaast beter in staat lesmateriaal te personaliseren.
2. *Specifieke micro optimalisatie in digitale leermiddelen.* Digitale leermiddelen schotelen leerlingen dagelijks grote hoeveelheden materiaal en oefeningen voor. Met AI kunnen deze systemen betere beslissingen maken over het te presenteren materiaal. Herkent het systeem bijvoorbeeld dat een leerling goed in staat is om kwadratische vergelijkingen op te lossen, dan kan een ander onderwerp worden behandeld. Een AI kan deze beslissingen op een veel fijnmaziger niveau maken dan een docent (die zou daarvoor immers van alle leerlingen zeer veel informatie moeten onthouden).
3. *Patroonherkenning/datamining in administratieve systemen.* Een AI kan bepaalde structurele problemen herkennen in datasets. Denk bijvoorbeeld aan het diagnosticeren van leerafwijkingen (zoals dyslexie of dyscalculie) op basis van werkstukken die leerlingen hebben geproduceerd. Een AI zou ook inzet kunnen worden om in administratieve systemen leerlingen te identificeren met een grotere kans op schooluitval.
4. *Monitoren klaslokaal.* Een AI kan wanneer deze is uitgerust met sensoren ook het klas lokaal monitoren. Zo kan een AI onenigheden tussen leerlingen in het klaslokaal herkennen of de aandacht van leerlingen kunnen monitoren. Op deze manier kan een AI de docent feedback geven over activiteiten die plaatsvinden buiten de digitale leermiddelen.

In dit scenario, levert de AI (net als de klassenassistent) af en toe een oordeel zonder tussenkomst van de docent. De docent laat sommige taken over aan de assistent, maar besteedt niet alles uit. Vertrouwen is gebaseerd op hoe goed de AI in het verleden heeft gepresteerd. Vertrouwen van de leraar in een AI kan worden onderbouwd op basis van inzicht in welke fouten in het verleden door het algoritme zijn gemaakt. De klassenassistent heeft daarnaast ook een keurmerk (opleiding).

AI voor learning analytics

Allereerst zien we een groot aantal toepassingen waarbij met slimme algoritmes inzichten worden verkregen in het leerproces (*learning analytics*). Deze kunnen de docent in staat stellen om vroegtijdig problemen te signaleren of om beter onderbouwde beslissingen te

maken. Ook het informeren van docenten over de ontwikkeling van leerlingen en het identificeren van leerlingen die een grotere kans hebben uit te vallen worden gezien als toepassing van AI.

Personalisering van onderwijs

Het is denkbaar dat een AI bepaalde *cognitieve taken* van een docent volledig kan overnemen, en daarbij (vrijwel) autonoom tot een beslissing komt. Hierbij kan worden gedacht aan het automatiseren van het nakijken en beoordelen van toetsen, feedback geven aan leerlingen, en het uitkiezen van leermateriaal. Dergelijke toepassing van AI zou vergaande personalisering van onderwijs mogelijk maken. Tot op de dag van vandaag is (regulier) onderwijs in Nederland voornamelijk klassikaal ingericht: één docent onderwijst meerdere leerlingen tegelijkertijd, en kiest daarbij zelf hoe de aandacht te verdelen. Men lijkt het er, zowel in de onderwijssector als in wetenschappelijke literatuur, echter over eens dat één-op-één onderwijs in theorie tot betere leerresultaten zou leiden. [15] Dergelijk onderwijs is echter zeer kostbaar, alleen al gezien het aantal leraren dat nodig zou zijn.

AI voor toetsing

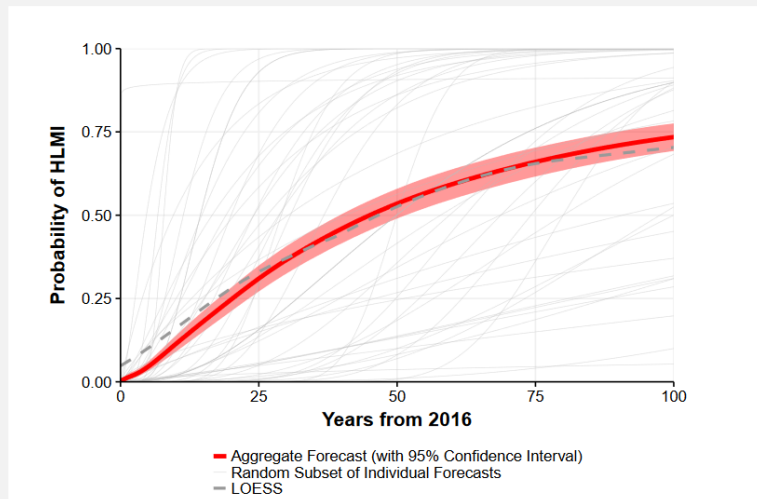
Met AI is het, op basis van de huidige stand van techniek, mogelijk om bepaalde nakijktaken van een docent te automatiseren. Het is denkbaar dat de technologie zich zodanig verder ontwikkelt, dat deze op basis van gegevens over het leerproces kan inschatten of een leerling bepaalde stof beheerst. Uiteindelijk zou met inzet van AI kunnen worden 'getoetst zonder toetsen'.

Bevinding 3. Kan AI een docent volledig vervangen? Niet in de nabije toekomst.

Wanneer AI zó slim is dat het de docent kan vervangen, dan zou dat het onderwijs in theorie dus sterk kunnen verbeteren. Zo ver is de technologie echter nog niet: de verwachting is dat dergelijke *artificial general intelligence* (waarbij een AI het niveau van intelligentie van een mens evenaart) nog minstens enkele decennia op zich laat wachten. Dit neemt echter niet weg dat minder intelligente AI, die vandaag al beschikbaar is, de leraar zodanig kan ondersteunen dat deze meer tijd per leerling kan besteden, of efficiënter te werk kan gaan.

Het is onduidelijk wanneer *artificial general intelligence* (AGI) wordt gerealiseerd. Optimisten stellen dat gegeven de doorbraken op het gebied van *deep reinforcement learning* zo snel gaan dat het slechts een kwestie van rekenkracht voordat AGI gerealiseerd kan worden. Pessimisten stellen er eerst fundamentele doorbraken moeten worden gedaan zodat een AI kan 'denken als een mens'. Hierbij horen ook zaken zoals menselijk redeneer en abstractievermogen.

Uit een survey van Nick Bostrom in 2013 onder AI-experts blijkt dat een optimistische voorspelling (10% waarschijnlijkheid) voor een AGI is 2022, een realistische voorspelling (50% waarschijnlijkheid) is 2040, en een pessimistische voorspelling (90% waarschijnlijkheid) is 2075. [16] In 2017 is een soortgelijke studie uitgevoerd door Grace et al [17]. Uit dit onderzoek het waarschijnlijk is dat dat AI binnen aanzienlijke tijd beter presteert dan mensen op de volgende taken: vertalen van teksten (2024), het schrijven van middelbare-school essays (2026), het besturen van een vrachtwagen (2027), het schrijven van een bestseller (2049) en het werken als chirurg (2053). In deze studie komt naar voren dat over 50 jaar de kans meer dan 50% dat er een *human level machine intelligence* (HLMI) bestaat. Figuur 9 toont de per jaar de waarschijnlijkheid op HLMI. De grijze lijnen zijn de voorspellingen van de individuele respondenten. De rode lijn is de geaggregeerde voorspelling. De LOESS lijn is een non-parametrische regressie door alle observaties.



Figuur 9 De kans op Human Level Machine Intelligence volgens deelnemers NIPS en ICML conferenties in 2015. [17]

3 Wat mag er? ²

Binnen de doelstelling van het onderzoek kijken we in dit hoofdstuk naar de tweede vraag: welke juridische aspecten er spelen bij het gebruik van AI-toepassingen in het onderwijs (wat mag er?). We vragen ons daarbij af welke grenzen het in Nederland toepasselijke recht opwerpt bij het toepassen van AI in het onderwijs en inventariseren mogelijke belemmeringen.

In paragraaf 3.1 adresseren allereerst kort hoe we tegen een AI-toepassing moeten aankijken vanuit een perspectief van *rechtspersonen*. Vervolgens duiden we welke gebeurtenissen bij het toepassen van AI juridisch relevant zijn. Daarnaast lopen we de toepasselijke rechtsgebieden af die van toepassing zijn op de juridische relevante gebeurtenissen, waarbij we een onderscheid maken tussen algemene regelingen (paragraaf 3.3) en het sectorspecifieke onderwijsrecht (paragraaf 3.4). Paragraaf 3.5 bevat tot slot de belangrijkste conclusies van deze analyse.

3.1 Toerekening van juridisch relevante gebeurtenissen aan personen

Vragen over 'wat moet en wat mag?' worden beheerst door het recht. Naar Nederlands recht kunnen alleen personen drager zijn van rechten en/of plichten. Deze rechtssubjecten zijn onder te verdelen in natuurlijke personen (een mens van vlees en bloed) en rechtspersonen (een juridische fictie, zoals een BV, een stichting (zoals sommige scholen) of de Staat (zoals het ministerie van OCW)). Kunnen we AI onder een van deze twee scharen?

Men zou kunnen betogen dat een AI kenmerken van een persoon vertoont: het is in staat zelfstandig te leren en zelfstandig besluiten te nemen. Aan de andere kant, AI is zeker niet van vlees en bloed, en het Nederlandse stelsel van rechtspersonen is gesloten: iets kan alleen een rechtspersoon zijn als de wet dat bepaalt, en AI wordt daarin niet genoemd. De conclusie is daarom dat AI-infrastructuur geen natuurlijk of rechtspersoon is, en daarmee geen zelfstandig drager van rechten plichten kan zijn. Dit vormt dus het onbetwiste uitgangspunt in deze rapportage.

Dit impliceert ook dat achter 'het handelen van een AI' altijd een persoon valt aan te wijzen aan wie deze handelingen toegewezen kunnen worden, en die de daarmee samenhangende rechten (bijvoorbeeld rechten van intellectuele eigendom) en/of plichten (bijvoorbeeld aansprakelijkheid voor AI-gebaseerde beslissingen waar schade uit voortvloeit) draagt. Hierbij kan aangesloten worden bij de wettelijke regelingen die gelden voor hulpzaken en hulppersonen, en wettelijke vertegenwoordigers (zoals bijvoorbeeld voor handelingsonbekwamen).

Bevinding 4 Een persoon is verantwoordelijk voor een AI, maar kan die verantwoordelijkheid niet zomaar nemen.

Zoals hierboven toegelicht is het op dit moment juridisch zo dat een AI geen rechtspersoon is, en daarmee op zichzelf geen drager van rechten noch plichten. De consequentie is dat een (rechts- of natuurlijk) persoon aansprakelijk rechten en plichten draagt namens een AI-toepassing, zoals aansprakelijkheid. Op zichzelf is dat niets nieuws: een persoon draagt immers ook de 'rechten en plichten' van ieder ander stuk gereedschap dat deze gebruikt. Gebruikt een autobestuurder bijvoorbeeld *cruise control* en leidt dit tot een ongeluk, dan is (als de cruise control geen defect bevat) de bestuurder daarvoor verantwoordelijk te

² Dit hoofdstuk is primair geschreven door Marc de Vries, partner bij The Green Land, waarmee Dialogic ter uitvoering van dit onderzoek heeft samengewerkt.

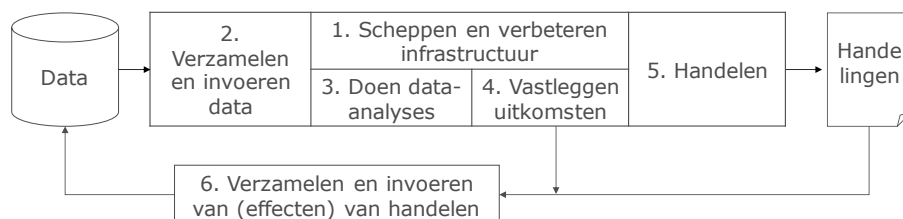
houden. De bestuurder vertrouwt de autofabrikant dat deze een degelijke cruise control heeft ontworpen. Daarnaast is de functionaliteit van de cruise control voor een bestuurder relatief overzichtelijk, maakt deze geen autonome beslissingen (anders dan het besturen van het gaspedaal op basis van snelheid), en is de cruise control in principe altijd uit te schakelen.

Bij AI-toepassingen is het echter veel lastiger om verantwoordelijkheid te nemen: een AI-toepassing is vaak intransparant en maakt soms wél autonome beslissingen. In de analogie met auto's zou het gaan om een AI die de auto bestuurt. Hoewel de bestuurder de verantwoordelijkheid heeft om (met de handen aan het stuur) de AI te 'controleren', is deze daar nooit voor opgeleid, en heeft de bestuurder, op basis van de huidige opleiding, geen begrip van de condities die ervoor kunnen zorgen dat een AI de verkeerde beslissing maakt.

3.2 Verkenning van de juridisch relevante gebeurtenissen

Nu we vastgesteld hebben dat aan het toepassen van AI altijd een persoon (in de juridische zin) gekoppeld kan worden, kunnen we nagaan welke juridische regimes van toepassing zijn bij het toepassen van AI in het onderwijs. Daarvoor moeten we dan eerst vaststellen wat de juridisch relevante gebeurtenissen zijn (paragraaf 3.2).

Bij het ontwikkelen, toepassen en verbeteren van AI zijn – binnen de kaders van dit onderzoek – de volgende juridisch relevante gebeurtenissen te onderscheiden (Figuur 10). We zijn daarbij uitgegaan van een situatie waarbij de instelling data aanlevert en een leverancier de AI ontwikkelt en toepast.



Figuur 10 Schematisch overzicht van juridisch relevante gebeurtenissen bij toepassing van AI in het onderwijs

Hieronder lopen we deze handelingen af, en gaan we na welke juridische regelingen op deze handelingen van toepassing zijn. Hierbij zij wel opgemerkt dat de scheidslijnen tussen de diverse gebeurtenissen niet altijd even hard zijn, en daarmee tevens de (al dan niet) toepasselijkheid van bepaalde regelingen daarop. Bijlage 2 bevat een schematisch overzicht van alle toepasselijke regimes per handeling (inclusief het onderwijsrecht).

3.3 Toepasselijke generieke regelingen

In deze sectie beschrijven we eerst de *generieke regelgeving*, dat wil zeggen regels die van toepassing zijn op de geïdentificeerde relevante gebeurtenissen *ook al zouden ze buiten het onderwijsdomein plaatsvinden*. Verderop, in paragraaf 3.4, komen de sectorspecifieke regelingen van onderwijsrecht aan bod.

Hierbij merken we nog op dat, gegeven de uitvraag van het Ministerie en de beschikbare dagen, het hierbij uiteraard gaat om een eerste inventarisatie en analyse op hoofdlijnen. Zeker waar het de onderwijsregels aangaat zal men steeds in concrete gevallen (bij toepassing van AI) moeten nagaan welke specifieke regels van toepassing zijn en moeten worden nageleefd en of er (dus) juridische obstakels zijn voor de uitvoering van de plannen.

3.3.1 Scheppen en verbeteren van AI-infrastructuur

Contractenrecht

De bouw van de AI-infrastructuur zal in een bepaalde rechtsverhouding tot stand komen. Het kan zijn dat dit op grond van opdracht (mogelijk via een aanbesteding) aan een ontwikkelaar plaatsvindt. Het kan ook zijn dat een koop- en/of licentieovereenkomst de grondslag vormt. In ieder geval hebben we hier te maken met regels uit het civiele recht, het Burgerlijk Wetboek (BW) in het bijzonder. Het BW kent als uitgangspunt contractsvrijheid, hetgeen veel ruimte laat voor instelling en leverancier zaken naar goeddunken te regelen.

Auteursrecht

Bij het maken van de AI-infrastructuur zullen er op grond van de Auteurswet (Aw) auteursrechten kunnen ontstaan. Deze ontstaan van rechtswege, de schepping van het werk is voldoende. Dat werk moet dan wel een eigen, intellectuele schepping zijn en het persoonlijk stempel van de maker dragen.^{3,4}

Kijken we naar het ontwikkelen van AI-toepassingen dan kunnen deze auteursrechten dus rusten op de software maar ook op bijvoorbeeld grafische interface, et cetera.⁵ Bijzondere aandacht verdient de vraag hoe het zit met creaties van de AI-toepassing *ze/ve*. We hebben hierboven vastgesteld dat deze rechten in ieder geval aan een persoon toebehoren. Hierbij valt te denken aan de instelling die de AI-toepassing gebruikt, dan wel de leverancier van de AI-toepassing.

De Aw bevat '*by default*' voorzieningen hoe deze auteursrechten worden toebedeeld aan makers en andere rechtsverkrijgers. Hiervan kan per contract worden afgeweken. Bij het ontwikkelen van AI-toepassingen zal het van groot belang zijn dit goed te regelen, met name in situaties waarin de instelling samen met de leverancier een ontwikkelrelatie aangaat waarbij de een de intelligentie levert en de ander de data. (zie hierna ook databankrechten). Voor auteursrechten van de overheid bestaan bovendien bijzondere regels: eerste openbaarmaking zonder voorbehoud leidt ertoe dat auteursrechten niet meer kunnen worden ingeroepen.

Mededingingsrecht

In zeer specifieke gevallen zou de Mededingingswet (Mw) van toepassing kunnen zijn, als de afspraken tussen de instelling(en) (en/of het ministerie) en de leverancier tot een machtspositie leiden doordat de facto de beschikkingsmacht over (mogelijk unieke) data gemonopoliseerd wordt dan de mededinging anderszins beperkt wordt (of kan worden). Deze situatie kan voorkomen worden door goede contractuele afspraken te maken over (beschikkingsmacht) over data en auteursrechten.⁶

³ HR 04-01-1991, ECLI:NL:HR:1991:ZC0104, m.nt. D.W.F. Verkade (Romme/Van Dale), r.o. 3.4.

⁴ HR 30-05-2008, ECLI:NL:HR:2008:BC2153 (Endstra/Nieuw Amsterdam c.s.), r.o. 4.5.1..

⁵ Artikel 10 Aw, deze lijst is echter niet limitatief: ook niet genoemde categorieën werken kunnen auteursrechtelijk beschermd zijn.

⁶ Overigens is het marktgedrag van de AI-leverancier het aangrijpingspunt, dit zal de instelling niet snel in de gevarenczone brengen van artikel 101 of 102 van het Verdrag betreffende de werking van de Europese Unie (of de nationale pendanten in de Mededingingswet).

3.3.2 Verzamelen en invoeren van data

Databeschermingsrecht

AI-toepassingen draaien op grote hoeveelheden data, die ze vervolgens verwerken om te kunnen analyseren. Zodra deze gegevens persoonsgegevens bevatten zullen de regels van de AVG in acht genomen moeten worden.⁷ Deze regeling zal daarom eveneens kunnen spelen bij de gebeurtenis 2 tot en met 6 (en mogelijk ook bij gebeurtenis 1).

De begrippen 'persoonsgegeven' en 'verwerken' zijn ruim gedefinieerd, en dat betekent dan ook dat de AVG op praktisch iedere handeling met de persoonsgegevens van toepassing is. Daar komt nog eens bij dat privacy enorm belangrijk is binnen het onderwijs: het gaat om gegevens van kinderen waarvoor een nog stringenter regime geldt.⁸ Voorts is het bepalen voor wie de verplichtingen van de AVG nu precies gelden op zich al een puzzel, vooral gezien de vele verschillende personen die bij de ontwikkeling en toepassing van AI betrokken kunnen zijn.

De AVG kent een aantal beginselen waarmee rekening moet worden gehouden bij het ontwikkelen en toepassen van AI en die waarschijnlijk al problematisch zijn:

1. **Verenigbaarheid:** persoonsgegevens niet worden verwerkt voor een doel dat onverenigbaar is met dat waarvoor de gegevens in de eerste zijn verkregen. Direct bij de ontwikkeling van AI zal dit al een probleem kunnen vormen, data zullen immers veelal niet specifiek zijn ingewonnen om een AI te helpen ontwikkelen.
2. **Juiste gegevens:** De verwerkingsverantwoordelijke mag alleen correcte persoonsgegevens verwerken. Dit zal bij AI gemakkelijk onzichtbaar een rol kunnen spelen omdat de ingebouwde intelligentie ontwerpfouten kan bevatten die een impact hebben op de te genereren data en intelligentie.
3. **Dataminimalisatie:** er mogen niet meer persoonsgegevens verwerkt worden dan noodzakelijk is voor het doel. Maar de kern van AI is nu juist dat het doel vaak (nog) niet duidelijk is zodat ook geen maat kan worden gezet op het gebruik van de data.

Uiteraard is voor iedere verwerking van persoonsgegevens een grondslag nodig. Dit kan zijn de vervulling van de publieke taak of behartiging van een gerechtvaardigd belang dat zwaarder weegt dan het privacybelang van de betrokkene(n). Er geldt evenwel een uitzondering voor wetenschappelijke doeleinden (maar dat moeten dan natuurlijk ook echt wetenschappelijke doeleinden zijn).

De AVG bevat voorts speciale regels als het gaat om '*geautomatiseerde individuele besluitvorming*' dat wil zeggen als persoonsgegevens worden gebruikt om tot een bepaalde beslissing te komen en deze beslissing *uitsluitend* is gebaseerd op geautomatiseerde verwerking van persoonsgegevens (dus zonder menselijke tussenkomst). Betrokkenen hebben het recht om niet aan GIB te worden onderworpen als i) dat rechtsgevolgen voor hen heeft of ii) het hen anderszins in aanzienlijke mate treft. Daarvoor gelden dan wel weer wat

⁷ Deze is van toepassing wanneer de verwerkingsverantwoordelijke of verwerker in de EU is gevestigd, of, als dat niet zo is, persoonsgegevens over betrokkenen in de EU worden verwerkt (als goederen of diensten aan betrokkene worden aangeboden). (art. 3 AVG). Dit zal steeds het geval zijn als AI in het Nederlandse onderwijs wordt toegepast.

⁸ Niet voor niets hebben uitgevers, digitale onderwijsdienstverleners en sectororganisaties in 2106 een landelijk privacyconvenant over de omgang met persoonsgegevens bij gebruik van (digitale) leermiddelen en administratiesystemen gesloten.

uitzonderingen maar het is zeer de vraag of toepassing in het onderwijsdomein daarop ziet. Hierbij speelt ook de vraag wanneer sprake is van een besluit. Immers bij toepassing van AI worden in feite continu (kleine) beslissingen genomen – het vaststellen van correlaties – die uiteindelijk tot de logica van een besluit leiden. De vraag is dus waar de knip te leggen en welke menselijke interventies in te bouwen. De antwoorden hierop zullen nog gegeven moeten worden (door de rechter dan wel op termijn door de wetgever).

Tenslotte bevat de AVG tal van transparantie-eisen. Zo moeten betrokkenen proactief worden geïnformeerd over de verwerking van hun persoonsgegevens (art. 13 en 14 AVG). Het is zeer de vraag of dit in de praktijk te realiseren zal zijn. Immers, (door)ontwikkeling van AI vereist continue verwerking die op basis van resultaten steeds weer aangepast moet worden.

Dit alles overziend lijkt het zeer de vraag of AI-gebruik kan maken van persoonsgegevens in het onderwijsdomein. Het lijkt realistischer de data van alle persoonskenmerken te ontdoen zodat de regels van de AVG niet langer van toepassing zijn. Weliswaar verliest men dan uiteraard veel waarde maar op een hoger aggregatieniveau zullen patronen niettemin ook zichtbaar worden.

Databankenrecht

Bij het verzamelen van data ter voeding van de AI-toepassing kunnen zogenaamde databankrechten ontstaan op grond van de Databankenwet (Dw). Deze regeling kan dus spelen bij de gebeurtenissen 2, 4 en 6.

Net als het auteursrecht ontstaat het databankenrecht van rechtswege en richt zich op de inhoud van een verzameling gegevens waarin substantieel geïnvesteerd is. Op grond van dit recht is het niet toegestaan, zonder toestemming van de producent van de databank, de databank of een (in kwalitatief of kwantitatief opzicht) substantieel gedeelte daarvan op te vragen en/of te hergebruiken. Dit recht beoogt de gedane investeringen in de databank te beschermen voor een periode van vijftien jaren, te rekenen vanaf het moment van productie dan wel de ter beschikkingstelling aan het publiek. Belangrijke wijzigingen van de inhoud van de databank (in kwalitatieve of kwantitatieve zin) doen een nieuwe beschermingstermijn aanvangen, zodra er sprake is van een nieuwe substantiële investering. Net als voor auteursrechten bestaan voor databankrechten van de overheid bijzondere regels: eerste openbaarmaking zonder voorbehoud leidt ertoe dat deze rechten niet meer kunnen worden ingeroepen.

Past men deze regels toe op AI-toepassingen dan is uiteraard de eerste vraag wie nu de investering doet. Het kan immers goed zijn dat de AI *zelf* nieuwe data produceert en de vraag is dan of dit gezien kan worden als een verzameling waarin geïnvesteerd is en, zo ja, bij wie de rechten dan moeten berusten: bij de leverancier van de data of de maker van de algoritmes? Daar komt nog eens bij dat de AI dikwijls gebruik zal maken van data uit verschillende bronnen (en van verschillende databankgerechtigden). Gegeven deze onzekerheden is raadzaam dit contractueel goed te regelen.

Openbaarheids- en hergebruikrecht

Ook zullen de regelingen die zien op de openbaarheid van bestuurlijke informatie (Wet openbaarheid bestuur (Wob)) en het hergebruik van overheidsinformatie (Wet hergebruik overheidsinformatie (Who)) mogelijk van toepassing zijn bij het toepassen van AI. Deze toepasselijkheid zal steeds kunnen ontstaan als informatie zich 'materialiseert' en deze komt te berusten bij de instelling en/of het ministerie. Dit zal dus kunnen plaatsvinden bij de handelingen 2, 4, 5 en 6.

De Wob ziet op bestuurlijke informatie. Onderwijsdata zullen daar veelal onder vallen. Artikel 10 van de Wob bevat evenwel uitzonderingen, de zogenaamde absolute en relatieve weigeringsgronden. Op die grond zullen grote verzamelingen persoonsgegevens (van bijvoorbeeld leerlingen of leraren) niet openbaar zijn. De toepassing van AI verandert hier niets aan. Vraag is wel hoe het zit met de intelligentie die de AI levert en de resultaten daarvan. Als deze openbaar zou zijn, zou dit een obstakel kunnen vormen voor het aangaan van overeenkomsten van leveranciers, zeker als deze intelligentie het resultaat is van eigen investeringen en ook in andere domeinen gebruikt zou kunnen worden. Aan de andere kant zou men ook kunnen betogen dat de onderliggende algoritmes juist wel openbaar zouden moeten zijn omdat inzichtelijk moet zijn hoe besluiten gebaseerd op AI tot stand komen.

Met andere woorden, hier zit een spanningsveld. Mogelijk zou artikel 10 lid 2 sub g hier uitkomst kunnen bieden. Dit artikel beschermt de belangen van degene die onevenwichtig benadeeld zou kunnen worden door de openbaarmaking.

Het bereik van de hergebruikregeling wordt voor een groot deel bepaald door de Wob: alleen openbare informatie kan herbruikbaar zijn. Bovendien mogen er op die informatie geen rechten van derden rusten. Dat schept dus een nieuwe vraag: hoe om te gaan met gemengde data? Immers de toepassing van AI zal ertoe leiden dat echte overheidsinformatie en data van derden door elkaar gehusseld gaan worden. Daar komt nog bij dat het resultaat van dat husselen tot nieuwe data leidt: zijn deze herbruikbaar of niet (omdat er rechten van derden op rusten). Trekt men deze lijn door dan zou overheidsinformatie wel eens een bedreigde diersoort kunnen worden.

Ook hier ligt de oplossing weer in het maken van heldere afspraken over eigendomsrechten – als deze uitsluitend bij de overheid liggen dan ontstaan er geen extra complicaties – dan wel het verkrijgen van licenties op grond waarvan de overheid data (waarop derden rechten hebben) openbaar mag maken.

3.3.3 Doen van data-analyses

Bij het uitvoeren van de AI-processen – het analyseren van de data door de AI-infrastructuur op zoek naar patronen – zal weer de AVG-toepassing zijn wanneer het om de verwerking van persoonsgegevens gaat (zie hierboven).

3.3.4 Verzamelen van uitkomsten

Ditzelfde zal gelden voor het verzamelen van de uitkomsten. Daarnaast is het denkbaar dat ook nu weer databankrechten ontstaan (zie hierboven). Ook zullen de Wob en de Who van toepassing kunnen zijn nu immers de toepassing van AI neerslaat in informatie neergelegd in documenten (zie hierboven).

3.3.5 Handelen op basis van uitkomsten

Zodra de instelling gaat handelen op basis van de AI – en dan in het bijzonder net nemen van besluiten – zal de Algemene wet bestuursrecht (Awb) van toepassing zijn, inclusief de algemene beginselen van behoorlijk bestuur en ook het BW in het bijzonder de regels over aansprakelijkheid. Als deze handelingen neerslaan in documenten dan zijn ook weer de Wob en de Who van toepassing (zie hierboven). Hieronder richten we ons dus uitsluitend op de toepasselijkheid van de Awb en het BW en de rechtsbescherming.

Algemene wet bestuursrecht

Zodra een instelling een besluit neemt in de zin van artikel 1:3 Awb, moet deze aan tal van procedurele en inhoudelijke eisen voldoen. We behandelen hier de relevante algemene – dat

wil zeggen de niet sectorspecifieke - die in de Awb zijn neergelegd of gebaseerd zijn op ongeschreven recht en die kunnen spelen bij het nemen van besluiten met behulp van een AI-toepassing. We gaan er daarbij vanuit dat een bestuursorgaan gebruik maakt van een AI-toepassing bij het nemen van een besluit. Uiteraard neemt het bestuursorgaan zelf het besluit, niet de AI-toepassing (zie hierboven).

Het zorgvuldigheidsbeginsel

Allereerst moet een bestuursorgaan voordat het een besluit neemt de nodige kennis omtrent de relevante feiten en de af te wegen belangen vergaren. Dit vergaren van feiten en belangen moet dus in de AI ingebouwd zitten. Die feiten en belangen zijn communicerende vaten en zullen steeds individueel moeten worden afgewogen, hetgeen normaal in het hoofd van de ambtenaar gebeurt. Die schakel gaat er (gedeeltelijk) uit en wordt in feite gebaseerd op abstracte en anterieure ideeën van de ontwikkelaar van de AI-toepassing.

Het transparantiebeginsel

Vervolgens moet het voor burgers helder zijn welke informatie de overheid heeft gebruikt bij het nemen van een besluit te nemen. Dit betekent dat de burger niet alleen moet kunnen weten *welke* data aan de AI-toepassing gevoerd zijn maar ook *hoe* deze geredeneerd heeft om tot het besluit te komen (dus goed beschouwd of de keuzes kloppen en eerlijk zijn). Dit geheel zal dus inzichtelijk moeten worden gemaakt of in ieder geval uitgelegd moeten worden aan de burger (het aanleveren van een stukje programmeercode zal dus niet volstaan).

Het gelijkheidsbeginsel

Verder moeten gelijke gevallen gelijk worden behandeld en ongelijke gevallen moeten ongelijk worden behandeld naar gelang de mate waarin zij verschillen. Wederom klemmt het hier dat de weging van de gelijkheid vooraf bepaald wordt, terwijl er altijd weer omstandigheden kunnen zijn die *dit specifieke geval* anders maakt.

Het vertrouwensbeginsel

De burger moet erop kunnen vertrouwen dat gewekte en gerechtvaardigde verwachtingen niet worden beschaamd. Dit impliceert dat de overheid grip moet houden op AI-toepassingen en evidente fouten signaleert en herstelt. Dit geldt temeer nu een AI-toepassing in staat is zelf te leren en dus het maken van ongecorrigeerde fouten, zonder ingrijpen, steeds vaker zal maken. Met andere woorden er zal een governancestructuur, en zullen governanceprocessen, op AI moeten worden ingericht om dit beginsel na te leven.

Het motiveringsbeginsel

Tenslotte moet op grond van de artikelen 3:46, 3:47 Awb ieder besluit een motivatie bevatten waaruit blijkt hoe bovenstaande beginselen in besluitvorming zijn nageleefd. De uitdagingen die hierboven zijn beschreven klinken dus ook door in dit motiveringsbeginsel.

In twee recente uitspraken heeft de Afdeling bestuursrechtspraak van de Raad van State de toepassing van deze beginselen nader uitgewerkt.

In een zaak uit 2017 werd de Afdeling verzocht zich uit te spreken over de vraag hoe reële rechtsbescherming te realiseren wanneer besluiten door AI worden genomen. Het ging hier om besluiten van de provincie Noord-Brabant onder de Natuurbeschermingswet 1998 tot het exploiteren en/of uitbreiden en wijzigen van zes verschillende agrarische bedrijven waartegen de belangengroep 'Behoud de Peel' opkwam. Bij de beantwoording overwoog de Afdeling dat als de burger niet kan controleren op basis waarvan tot een besluit is gekomen, AI een soort *black box* wordt. En omdat een gebrek aan inzicht in de gemaakte keuzes en gebruikte gegevens en aannames ertoe kan leiden dat er dan een ongelijkwaardige procespositie van partijen ontstaat, zal de overheid die keuzes en gegevens op een passende wijze openbaar

moeten maken, zodat de burger die eventueel gemotiveerd kan betwisten en toegang heeft tot rechtsbescherming.⁹

Dit jaar - het ging hier om een het tracébesluit 'Blankenburgverbinding' waartegen beroep werd ingesteld - heeft de Afdeling dit nog wat aangescherpt door te stellen dat de overheid aan de burger die opkomt tegen een besluit inzage moet geven in de *input* van de AI en moet aantonen op welke maatwerkgegevens een besluit is gebaseerd, op basis van artikel 8:42 Awb. De Afdeling maakt daarbij een onderscheid tussen twee soorten inputdata: (a) de *maatwerk* inputdata - deze moeten waarneembaar beschikbaar moeten worden gesteld door de overheid - en (b) *standaard* inputdata die alleen te hoeven worden verstrekt als de burger daarom vraagt.¹⁰

Zo zal de burger inzicht hebben in de inputkant, oftewel de voorkant van de besluitvorming als de outputkant, namelijk wat het zelflerende algoritme heeft gegenereerd - een individueel besluit. Wat de bestuursrechter dus niet heeft gezegd, is dat het bestuursorgaan ook de broncode van de AI-toepassing openbaar moet maken, dus de wijze waarop in de AI inputdata zijn verwerkt en welke keuzes er in dat kader al gemaakt zijn.^{11 12}

Aansprakelijkheidsrecht

Hoe zit het met aansprakelijkheid? Als op basis van een AI-toepassing een bestuursorgaan een besluit heeft genomen dat achteraf onjuist blijkt te zijn en als dit onjuiste besluit terug te voeren is op het gebruik van de AI-toepassing (en dit besluit heeft tot schade geleid bij een burger) dan zijn de regels over aansprakelijkheid uit het BW van toepassing.

Uiteraard zijn er ook andere situaties denkbaar waarbij de schade niet voortkomt uit een besluit (in de zin van de Awb) maar uit een andere gedraging. We komen hier nog op terug in de paragraaf over de sectorale regels. Daarnaast speelt uiteraard ook wat de instelling zou kunnen doen tegenover de leverancier van de AI-toepassing. Dit kan zijn omdat de instelling wordt aangesproken en verhaal wil halen bij de leverancier of de instelling kan simpelweg zelf menen dat er niet of onvoldoende gepresteerd is en daarom de leverancier aanspreken. We laten deze casus verder voor wat zij zijn omdat ze niet specifiek zijn voor toepassing in AI in het onderwijs.

Terug dus naar de casus waarin een instelling, met behulp van een AI-toepassing een besluit heeft genomen en de betrokken leerling daarvan schade heeft ondervonden. De eerste vraag is dan wat de rechtsgrond is waarop aansprakelijkheid gebaseerd kan worden. Dit hangt af van de vraag of er tussen de instelling en de leerling sprake is van een overeenkomst - in dat geval is de grondslag de toerekenbare tekortkoming - of niet - in dat geval is de grondslag de onrechtmatige daad. Meestal wordt aangenomen dat er geen sprake is van een overeenkomst maar recent is er een rechter geweest die het bestaan van een overeenkomst aannam tussen een bijzondere vmbo-school en een leerling.¹³ Omdat de grondslag uiteindelijk niet veel uitmaakt voor onze casus, kiezen we voor de grondslag onrechtmatige daad,

⁹ ABRvS, 17 mei 2017, ECLI:NL:RVS:2017:1259, r.o. 14.3 en 14.4.

¹⁰ ABRvS, 18 juli 2018, ECLI:NL:RVS:2018:2454, r.o. 23.1, 23.4 en 23.5.

¹¹ ABRvS, 17 mei 2017, ECLI:NL:RVS:2017:1259, pp.13-14.

¹² Dit is in lijn met de in Frankrijk sinds 2016 geldende *Loi pour une République numérique* (Loi no.2016-1321). Deze bepaalt dat een besluit dat is genomen op basis van algoritmes de belanghebbende hierover moet informeren en dat op verzoek de belangrijkste regels en de kenmerken van de uitvoering aan de belanghebbende worden medegedeeld (B.M.A. Van Eck, Geautomatiseerde ketenbesluiten & rechtsbescherming (dissertatie Tilburg 2018)).

¹³ Hof Arnhem-Leeuwarden 26 januari 2016, ECLI:NL:GHARL:2016:502.

omdat deze wat meer vraagt van de vordering uit hoofde van de toerekenbare tekortkoming (waarbij alleen maar de wanprestatie en de schuld moet worden aangetoond).

Artikel 6:162 BW geeft de vijf cumulatieve vereisten voor de vestiging van de aansprakelijkheid uit hoofde van een onrechtmatige daad. Er moet dan sprake zijn van: een onrechtmatige daad, schuld, schade, causaal verband tussen de handeling en de schade; en relativiteit. Omdat voor de onderdelen schade en relativiteit geen specifieke AI lijken te spelen, richten we ons uitsluitend op de drie resterende onderdelen: onrechtmatigheid, schuld en causaliteit.

Onrechtmatigheid

Artikel 6:162 kent drie soorten gedragingen die onrechtmatig kunnen zijn:

1. een inbreuk op een recht
2. een doen of nalaten in strijd met een wettelijke plicht
3. doen of nalaten in strijd met hetgeen volgens ongeschreven recht in het maatschappelijk verkeer betaamt.

Met name deze laatste - het zogenaamde betamelijkheidsbeginsel - zal een rol kunnen spelen bij de beoordeling of toepassen van AI onrechtmatig was jegens de leerling, namelijk of de instelling tekort geschoten is bij het voldoen aan haar zorgplicht.¹⁴ Deze norm wordt sterk ingekleurd door de hierboven beschreven regels van de geschreven en ongeschreven bestuursrecht, en dan met name het zorgvuldigheidsbeginsel dat bestuursorganen in acht moeten nemen bij het nemen van besluiten, op de voet van artikel 1:3 Awb. Zijn deze dus niet nageleefd, dan is aan het vereiste van onrechtmatigheid voldaan. Denk hierbij dus aan situaties waarin de AI:

1. Relevante feiten buiten beschouwing heeft gelaten;
2. Relevante belangen buiten beschouwing heeft gelaten;
3. Onjuiste input heeft gebruikt;
4. Een onjuiste redenering toepast;
5. Een foute beslissing neemt;
6. Uit eerdere besluiten het gerechtvaardigd vertrouwen is gewekt dat het besluit anders uit zou vallen;
7. Het besluit niet of onvoldoende gemotiveerd is.

Schuld

Aansprakelijkheid vergt voorts dat de daad toe te rekenen is aan de 'dader'. Artikel 6:162 lid 3 BW bepaalt dat zo'n toerekening kan plaatsvinden indien de daad te wijten is aan de schuld van de dader of aan een oorzaak die op grond van de wet of de in het verkeer geldende opvattingen voor zijn rekening komt.

Bij AI kunnen tal van actoren hebben bijgedragen aan de AI-toepassing: producenten van de AI-systemen, de leveranciers van de data, de programmeur van de software, de gebruiker van de toepassing, de eigenaar van de toepassing etc. Daarnaast zou de AI ook geïntegreerd kunnen zijn in een object (een zaak). In de relatie tussen leerling die schade heeft geleden door het besluit (gebaseerd op de AI-toepassing) en de instelling is dit evenwel niet van

¹⁴ De andere twee onrechtmatigheidsgronden zijn over het algemeen veel simpeler toe te passen: als de instelling iets stuk maakt dat van de leerling is of geen onderwijs aanbiedt dan is die gedraging onrechtmatig.

belang. De overheid is namelijk aansprakelijk voor fouten van hulppersonen en hulpzaken als ware het haar eigen fouten en is als bezitter aansprakelijk voor gebrekkige zaken.¹⁵

Dit laat overigens onverlet dat de gedraging van de leverancier van de AI-toepassing op zichzelf ook onrechtmatig kan zijn tegenover de leerling. Anders gezegd, de leerling zou *naast* de instelling, ook deze leverancier(s) kunnen aanspreken.

Causaliteit

Uiteraard moet er een verband bestaan tussen de daad en de schade. Dit wordt getest door de vraag te stellen: was de schade ook ontstaan als de gedraging niet had plaatsgevonden? Als het antwoord daarop nee is, dan ontbreekt dit zogenaamde *conditio sine qua non* verband en zal de schade niet voor vergoeding in aanmerking komen.

Dit kan een lastige zijn bij de toepassing van AI. In de eerste plaats is een kenmerk van AI dat het zelf leert op basis van *feed back loops* en het op een gegeven moment niet meer duidelijk is wat nu de input data, de criteria en het denkproces zijn op basis waarvan de AI tot een besluit komt. Daar komt nog bij dat de AI alleen kan leren op basis van heel veel data, dikwijls uit diverse bronnen. Dit leidt ertoe dat de relatie tussen de schade en het schadeveroorzakende feit steeds lastig te leggen zal zijn. Nu is het uitgangspunt dat de benadeelde moet stellen en aannemelijk maken dat het *conditio sine qua non* verband bestaat, maar de rechter kan de benadeelde te hulp schieten door de bewijslast om te keren. Zeker nu het ontstaan van deze *black box* te wijten is aan de leverancier (en daarmee de instelling), zal de rechter daartoe geneigd zijn.

Overigens zou het ook nog denkbaar zijn de aansprakelijkheid te baseren op zogenaamde gevaarzetting. Gevaarzetting is, op zijn Hollands gezegd het in het leven roepen van een gevaarlijke situatie voor personen of zaken of het achterwege laten van voorzorgsmaatregelen met betrekking tot dat gevaar. Als gevaarzetting aanwezig is – dus als er een gevaarlijke situatie bestaat – dan is men op grond van de betamelijkheidsnorm verplicht anderen niet bloot te stellen aan een groter risico dan onder de gegeven omstandigheden redelijkerwijs verantwoord is. Met andere woorden: niet iedere gevaarzetting is onrechtmatig.

Wanneer wordt de gevaarzetting dan wel onrechtmatig? Dat hangt af van de omstandigheden van het geval, waarbij met name gekeken moet worden naar een viertal factoren, aldus de Hoge Raad in het beroemde Kelderluik arrest.¹⁶ De hoofdregel uit deze (en opvolgende) jurisprudentie is dat men ernaar moet streven goede en afdoende maatregelen te nemen die gericht zijn op het voorkomen van gevaarzetting. Deze zorgplicht varieert naar gelang de omstandigheden, maar er is een zestal vuistregels:

1. naarmate het waarschijnlijker is dat de potentiële slachtoffers minder oplettend zijn, dient een hogere zorgplicht in acht genomen te worden;
2. naarmate de ernst (met name letselschade) en omvang van de mogelijke schade groter is, dient een hogere zorgplicht in acht genomen te worden;
3. naarmate de kans op schade groter is, dient een hogere zorgplicht in acht genomen te worden;
4. naarmate de gedraging gevaarlijker is, dient een hogere zorgplicht in acht genomen te worden;

¹⁵ Zie onder meer de artikelen in het BW: 6:170 (ondergeschikten), 6:171 (opdrachtnemers), 6:172 (vertegenwoordigers), 6:173 (gebrekkige zaken) en 6:76 (hulppersonen).

¹⁶ HR 5 november 1965, NJ 1966, 136 (*Kelderluik*)

5. naarmate het nemen van bepaalde voorzorgsmaatregelen, zowel op zichzelf beschouwd, als in relatie tot mogelijke schade, minder bezwaarlijk is qua kosten, tijd en moeite, bestaat een sterkere gehoudenheid tot het treffen van preventieve maatregelen;
6. daarbij is niet van belang of men het gevaar dat zich heeft verwezenlijkt, concreet heeft voorzien. Voldoende is dat met het moeten nemen van bepaalde veiligheidsmaatregelen in het algemeen – naar objectieve maatstaven – bepaalde (gezondheids)gevaaren getracht wordt te voorkomen. Waar dat niet mogelijk of te bezwaarlijk is, moet voor dat mogelijke gevaar gewaarschuwd worden. Daarbij is van belang of verwacht mag worden dat de waarschuwing ertoe zal leiden dat het gevaar vermeden wordt.

Als er niettemin een rechter aan te pas moet komen zal deze moeten afwegen of de genomen veiligheidsmaatregelen voldoende geweest zijn, door na te gaan welke veiligheidsmaatregelen genomen zijn en dat af te zetten tegen welke genomen hadden moeten worden. Te beoordelen naar de stand van zaken ten tijde van de gedraging.

Relevant hierbij is dat AI van nature fouten maakt om daar vervolgens van te leren. Dus de gevaarstelling zit als het ware ingebouwd in de toepassing, zeker aan het begin. Dit vereist uiteraard extra voorzorgsmaatregelen van degene die zich van de AI-toepassing bedient. Daarnaast kan men van iemand op wie AI wordt toegepast, vanwege de complexiteit niet al te veel oplettenheid verwachten.

Rechtsbescherming

Rechtsbescherming in het onderwijs is versnipperd. Men kan terecht bij de gewone bestuursrechter en de burgerlijke rechter, maar daarnaast zijn er ook vaak tal van klachtencommissies. Voorts zijn er vaak tal van (min of meer verplichte) interne voortrajecten, bijvoorbeeld bij een beroepsinstantie van een school. Niet iedere handeling van een openbare school is een besluit in de zin van de Awb (bijvoorbeeld het besluit om een leerling na te laten blijven of een extra les te volgen). De bestuursrechter toetst aan de Awb, de relevante onderwijsregels, de beginselen van behoorlijk bestuur en eventuele specifieke beleidsregels van de instelling zelf. Kernvraag daarbij is steeds of de school de beslissing op zorgvuldige wijze heeft genomen, dat wil zeggen: na deugdelijk feitenonderzoek en adequate belangenafweging.

3.3.6 Verzamelen en invoeren van (effecten) van handelen

De laatste stap in het proces betreft het verzamelen en invoeren van de uitkomsten van de AI-toepassing en het (her)invoeren daarvan in het proces. Er vindt immers een continue stroom van data plaats vanuit het proces naar de AI-infrastructuur die leidt tot het leggen van verbanden tussen handelingen in het proces en effecten daarvan. Als hierbij persoonsgegevens worden gebruikt is de AVG van toepassing en kunnen wederom databankrechten ontstaan op grond van de Dw. Tenslotte, daar waar informatie materialiseert, zullen de ook de regelingen omtrent openbaarheid (Wob) en hergebruik (Who) van toepassing kunnen zijn.

3.4 Toepasselijke sectorale regelingen

Bij de toepassing van AI in het onderwijs zijn, naast de generieke regelingen, uiteraard ook onderwijs-specifieke regelingen van toepassing en omdat alle relevante gebeurtenissen plaatsvinden binnen het kader van het onderwijs zijn op al deze gebeurtenissen de sectorale regels van toepassing.

Uiteraard, zoals hierboven reeds vermeld, moeten we daarbij steeds in gedachten houden dat veruit de meeste handelingen die in het kader van het openbaar onderwijs worden verricht, beheerst worden door de algemene regels van het Bestuursrecht, in het bijzonder de Awb (zie hiervoor paragraaf 3.3.5). De focus in deze paragraaf ligt dus volledig op *die* regels van het onderwijsrecht die *daarbovenop* een verschil maken.

We starten met een korte beschrijving van de inhoud en structuur van het onderwijsrecht in paragraaf 3.4.1. Vervolgens gaan we in paragraaf 3.4.2 na welke beginselen en regels het gebruik van AI inkaderen en mogelijk beperken.

3.4.1 Structuur en inhoud onderwijsrecht

Onderwijsrecht is al het recht dat betrekking heeft op onderwijsinstellingen, en de rechtsrelaties regelt tussen betrokkenen in het onderwijs: de overheid, schoolbesturen, ouders, leerlingen en personeel. Het is een zogeheten functioneel rechtsgebied: in plaats van één centrale wet, bestaat het onderwijs recht uit verschillende takken van het recht. Een groot deel is gebaseerd op publiekrecht, omdat het over grondwettelijke vraagstukken gaat – zoals de vrijheid van onderwijs, vastgelegd in artikel 23 van de Grondwet, en daarmee ook besluiten die de overheid neemt ten aanzien van schoolbesturen, of die schoolbesturen zelf nemen. Die besluiten worden, zoals hiervoor gezegd, geregeld in de Awb. [18, pp. 14-15] Ook het privaatrecht is deels van toepassing op het onderwijs.

Behalve in het bestuursrecht, is het onderwijsrecht verder geregeld in een aantal verschillende ‘lagen’ van wetten. Bovenaan staan de sectorwetten: de Wet op het primair onderwijs (WPO) en de Wet op het voortgezet onderwijs (WVO). De WPO is nader uitgewerkt in een aantal besluiten; de WVO is een kaderwet wat betekent dat belangrijke onderdelen zijn opgenomen in Algemene Maatregelen van Bestuur (AMvB). Daarnaast zijn er sectoroverstijgende themawetten, zoals de Inspectiewet, de Leerplichtwet, de Wet op het onderwijstoezicht (WOT) of de Wet medezeggenschap op scholen (WMS).

Daaronder in de rangorde bestaat weer een scala van regelingen en regels die de onderwijsinstelling of betrokken partijen juridisch binden. Zo zijn er bijvoorbeeld horizontale regels die binnen een instelling gelden (in tegenstelling tot verticale regels, tussen overheid en burger of instelling), zoals bijvoorbeeld een professioneel of medezeggenschapsstatuut. Er zijn codes voor goed bestuur, gedragscodes voor medewerkers, of convenanten – zoals een landelijk privacy convenant.

3.4.2 Relevante sectorale kaders voor de toepassing van AI

AI-toepassingen kunnen in feite alle bestaande rechtsrelaties in het onderwijs raken. Denk daarbij aan de relatie leerling - instelling, docent - leerling, instelling – ouder, instelling – docent, ouder – docent, maar bijvoorbeeld ook aan de relatie ministerie – instelling, inspectie – instelling. Binnen de kaders van dit onderzoek zou het te ver voeren al deze relaties te ontleden, en conform de opdracht hebben we onze focus gelegd op de kern van het onderwijsproces, dus datgene dat in de klas plaatsvindt: het aanbieden van lesstof en het afnemen daarvan. We vragen ons daarbij af wat nu de relevante kaders zijn om vervolgens de vraag te beantwoorden in hoeverre deze kaders het gebruik van AI in het onderwijs beperken.

Ter structurering daarvan hebben we een onderscheid gemaakt tussen mogelijke beperkingen op (1) het niveau van de beginselen die het onderwijssysteem beheersen, (2) het niveau van de inhoud van het onderwijs (3) het niveau van individuele regels en (4) het niveau van de sturingsmogelijkheden.

Op het niveau van beginselen

Het onderwijsrecht wordt gedreven door een aantal grondbeginselen. Deze zijn vastgelegd in 'hogere' regelingen zoals verdragen en de Grondwet (Gw). Zo zijn daar:

1. Niemand mag *het recht op onderwijs* en bestaande onderwijsvoorzieningen worden ontzegd. Dat betekent ook dat het onderwijs in al zijn fases en geledingen open moet staan voor personen die daarvoor de bekwaamheid bezitten. Toegelaten leerlingen moeten vervolgens ook daadwerkelijk dat onderwijs kunnen voltooien.¹⁷
2. *De garantiefunctie*: de overheid moet ervoor zorgen dat niemand feitelijk gedwongen is zijn kinderen naar een bijzondere school te sturen.¹⁸
3. *De vrijheid van stichting*: het geven van onderwijs is vrij, behoudens het toezicht van de overheid.¹⁹
4. *De vrijheid van richting*: in het bijzonder onderwijs heeft men de vrijheid om een bepaalde godsdienst of levensovertuiging in het onderwijs tot uitdrukking te brengen en op die grond leerlingen en personeel te selecteren.²⁰ Dat beleid moet dan wel plaatsvinden op basis van de grondslag van de instelling en in de praktijk daadwerkelijk gestalte krijgen, het beleid moet noodzakelijk zijn voor de verwezenlijking van de grondslag en consequent worden uitgevoerd en gehandhaafd.²¹
5. *De vrijheid van inrichting*: in het bijzonder onderwijs heeft men de vrijheid om eigen onderwijsmethodes te kiezen, een eigen bestuursmodel, kwaliteitsdoelstellingen etc.).
6. Ook in het openbaar onderwijs gelden de principes van richting en inrichting, maar deze vrijheid is wel meer begrensd. Deze zogenaamde pedagogische autonomie is daarbij overigens niet gebaseerd op de Grondwet maar op ongeschreven beginselen.
7. En, last but not least, staat naast het onderwijzen ook *de bescherming van het kind* centraal.

Uit het bovenstaande blijkt dat het Nederlands onderwijsstelsel levensbeschouwelijk en pluriform is, algemeen toegankelijk en algemeen beschikbaar en het openbaar onderwijs een basisvoorziening is die voor eenieder openstaat. Het spreekt voor zich dat een AI-toepassing deze beginselen geen geweld mag aandoen. AI zal dus nooit iemand het recht op onderwijs mogen ontzeggen. Evenmin mag AI ertoe leiden dat de vrijheid van inrichting en richting illusoir worden.

Verder zal het steeds zo zijn dat de belangen van het kind centraal moeten staan en zodra een AI-toepassing daar afbreuk aan dreigt te doen dit juridisch, maar vooral ook ethisch onacceptabel zal zijn. Dat klinkt uiteraard ook door in de regels over het verwerken van de

¹⁷ Artikel 2 Eerste Protocol van het Europees Verdrag voor de Rechten van de Mens en EHRM 23 juli 1968, Publ. ECHR, Series A, vol. 6 (Belgische Taalzaak). Dit beginsel is uitgewerkt in lagere regelingen bijvoorbeeld artikel 46 lid 2 WPO.

¹⁸ Artikel 23 lid 4 van de Grondwet.

¹⁹ Artikel 23 lid 2 van de Grondwet.

²⁰ Artikel 23 lid 5 en 6 van de Grondwet.

²¹ Artikel 1 van de Grondwet, de Algemene Wet Gelijke Behandeling en HR 22 januari 1988, AB 1988, 96, NJ 1988/981 (Maimonides)

persoonsgegevens van kinderen dat aan stringente eisen is onderworpen waarbij een ander (grootschalig) gebruik dan het gebruik waarvoor de gegevens verzameld zijn – vaak zal AI dat willen doen - nauwelijks mogelijk is.

Op het niveau van de inhoud van het onderwijs

Het bevoegd gezag is verantwoordelijk voor de zorg voor de kwaliteit van het onderwijs: de deugdelijkheidseisen.²² Op bepaalde onderdelen vult de sectorale wetgeving deze kwaliteitsnormen verder in. Zo moet het onderwijs zodanig zijn ingericht dat dat leerlingen een ononderbroken ontwikkelingsproces kunnen doorlopen en het moet afgestemd zijn op de voortgang en ontwikkeling van de leerlingen. Er is een specifieke zorgplicht voor leerlingen die extra zorg behoeven, dus het onderwijs moet gericht zijn op individuele begeleiding die is afgestemd op de behoeften van de leerling (artikel 8 lid 4 WPO).

De sectorale wetten bepalen voorts dat de minimale inhoud van het lesprogramma's moet worden uitgewerkt in het Besluit Kerndoelen (artikel 9 lid 9 WPO en 11b WVO). Deze kern-doelen zijn vervolgens geclusterd onder thema's. Daarnaast bepaalt de Wet referentieniveaus Nederlandse taal wat leerlingen moeten kunnen en kennen op bepaalde momenten in hun schoolcarrière.

De sectorale wetten bevatten voorts procedurele voorschriften ter zake van de te leveren kwaliteit, waaronder de wijze waarop het bevoegd gezag bewaakt dat die kwaliteit wordt gerealiseerd, de vaststelling welke maatregelen ter verbetering van de kwaliteit nodig zijn en maatregelen en instrumenten om te waarborgen dat het personeel zijn bekwaamheid onderhoudt. Hieronder valt ook het beleid ten aanzien van aanvaarding van materiele of geldelijke bijdragen indien het bevoegd gezag daarbij verplichtingen op zich neemt waarmee leerlingen zullen worden geconfronteerd tijdens schooltijd.

Voorts zijn er tal van regels over het onderwijsproces en de inrichting, bijvoorbeeld over de duur en omvang van het onderwijs en de verdeling over dagen (artikel 8 lid 9 WPO en artikel 6g en 11c WVO). Ook de eisen aan de toetsing zijn nader uitgewerkt in sectorale wetten. Voor het VO werkt het Eindexamenbesluit de verplichte inhoud nader uit (inhoud, afname, beoordeling, uitslag, herkansing en diplomering).

Als men mogelijke AI-toepassingen tegen deze regels aanlegt dan kunnen we constateren dat de vrijheid van instellingen aan de ene kant groot is, maar aan de andere kant aan specifieke kwaliteitseisen moet voldoen, zowel qua inrichting van bestuur, de inhoud van onderwijs, de inrichting daarvan, het voldoen aan de onderliggende processen en de uiteindelijke toetsing. Dit perkt uiteraard de vrijheid van het toepassen van AI maar tegelijkertijd kan geconstateerd worden dat deze regels dusdanig open geformuleerd zijn dat er geen materiële blokkades zijn voor het toepassen van AI, zolang men maar binnen de getrokken lijntjes blijft.

Op het niveau van individuele regels

Dat neemt uiteraard niet weg dat het toepassen van AI in het onderwijs niet tal van gedetailleerde juridische vragen zal oproepen. Het valt buiten het kader van dit onderzoek om deze allemaal – voor zover dat al mogelijk zou zijn – af te lopen. Als illustratie enkele voorbeelden:

²² Deze normen over de inhoud van het onderwijs zijn vastgelegd in art. 8 WPO jo. Art. 9 WPO en art. 11b WVO.

1. De Leerplichtwet bepaalt dat een leerling na inschrijving geregeld de school bezoekt, dat wil zeggen dat deze feitelijk de lessen bijwoont (artikel 4 lid 2 en artikel 4c lid 2 van de Leerplichtwet). Dit heeft uiteraard te maken met de aanwezigheid van de leraar op de school: om onderwijs te kunnen aanbieden moet de leerling aanwezig zijn als de leraar lesgeeft. Maar wat nu als een AI toepassing de rol van de leraar (gedeeltelijk) overneemt? Is er dan nog een noodzaak dat de leerling op school aanwezig is? Sensoren en andere meetinstrumenten kunnen de benodigde waarnemingen doen en real time respons geven daarop. Daarbij moet men zich wel realiseren dat een leerling door aanwezig te zijn op school (en dus in een groep) ook tal van andere (sociale) vaardigheden aanleert.
2. In het onderwijsrecht is er een aparte rechtspositie voor de ouder(s) van de leerling. De relatie tussen ouder(s) het kind en de instelling waarop deze onderwijs geniet is in feite een driehoeksrelatie. Deze relatie is gedeeltelijk wettelijk geregeld en wordt gedeeltelijk overgelaten aan partijen om verder in te vullen. Zo heeft de instelling een actieve informatieplicht over de vorderingen van de leerling (artikel 6 lid 6 en artikel 11 WPO en artikel 23b WVO). Ook op grond van de AVG hebben ouders tal van rechten waar het de verwerking van persoonsgegevens aangaat door de instelling. Daarnaast hebben ouders ook inzagerecht in het leerlingendossier. Maar hoe dit in te richten bij de toepassing van AI? Wat moet gemeld worden? Wat moet inzichtelijk gemaakt worden? Valt hier de werking van de AI ook onder? Kunnen/moeten ouders toestemming geven voor het gebruik van AI dat verder gaat dan de AVG toestaat?
3. Het kader van de onderwijsinspectie valt uiteen in het waarderingskader (het object van het toezicht) en het onderzoekskader (het proces van toezicht). Het waarderingskader omvat onder meer de kwaliteitsgebieden onderwijsproces, onderwijsresultaten, en onderwijszorg. Deze gebieden bevatten standaarden waar een instelling aan moet voldoen. De inspectie kan op basis hiervan een oordeel vormen over het functioneren van de instelling. Op dit alles zal een AI dus impact hebben, als die bijvoorbeeld gebruikt wordt om leerlingen te toetsen. Wat daaruit komt, moet door een Inspectie beoordeeld kunnen worden, en de kwaliteit van het onderwijs moet door het bevoegd gezag gewaarborgd worden. Hiervoor is het dus nodig dat het duidelijk is wat een AI doet en wat er op welke manier uitkomt. Moet hiervoor de regels worden aangepast? Heeft de inspectie extra bevoegdheden nodig gezien de black box die ontstaat?

Op het niveau van de sturing

Het Nederlands stelsel wordt gekenmerkt door een hoge mate van vrijheid 'op de werkvloer' maar dan wel gekoppeld aan een hoog poldergehalte op dat niveau.

Top down sturing vanuit het ministerie in de richting van de instellingen is nauwelijks geregeld. Als tegenwicht voor de vrijheid die instellingen hebben is er weliswaar een toezichtarrangement gecreëerd in de vorm van extern toezicht door een inspectie vastgelegd in de Wet op het onderwijstoezicht (WOT). Maar dit is in hoge mate reactief en een instelling moet (langdurig) uit de bocht vliegen voordat de inspectie inhoudelijk zal ingrijpen. Ook de

rechterlijke macht is terughoudend bij het toetsen van besluiten van een instelling, bijvoorbeeld over overgaan of blijven zitten.²³

Dit geldt ook in zeker zin ook voor de leraar: deze heeft professionele ruimte om de lesstof aan te bieden, uiteraard binnen de wettelijke grenzen en niettegenstaande de verantwoordelijkheid van het schoolbestuur voor de onderwijskwaliteit en het functioneren van de school. In de artikelen 31a WPO en 32e WVO wordt onder meer bepaald dat de leraar *zelf* verantwoordelijk is voor de beoordeling van onderwijsprestaties van leerlingen. Het professioneel statuut werkt deze zeggenschap verder uit.

Op het niveau van de individuele instelling zijn er evenwel zal er tal van horizontale *checks and balances*. Zo moet aan een instelling moet een medezeggenschapsraad (MR) zijn verbonden zodat leerlingen, ouders en personeel medezeggenschap hebben. De Wet medezeggenschap op scholen (WMS) regelt dit. Op grond van artikel 8 WMS heeft de MR tal van informatierechten zodat deze haar taken kan verrichten en bevoegdheden kan uitoefenen. Artikel 10 WMO geeft aan waarover het bevoegd gezag van de instelling instemming dan wel advies moet verkrijgen van de MR. Instemming behoeven onder meer: (voorgenomen) besluiten over verandering van onderwijskundige doelstellingen, het school/leerplan, examenregelingen, (niet wettelijk geregelde) materiele of geldelijke bijdragen. Advies behoeven onder meer: vaststelling van het lesrooster, deelname aan een onderwijskundig project, duurzame samenwerking. Op die voet zal het aangaan van samenwerkingsverbanden met marktpartijen – denk hierbij aan de ruil van (gebruik van) data tegen de beschikbaarstelling van AI-toepassingen (of een korting op de koopprijs of een licentie) waarschijnlijk instemming van de MR behoeven. Wordt het gebruik van AI structureel, dan mag verwacht worden dat dit (in ieder geval) aan het school/leerplan raakt.

Uit deze constellatie volgt dat het grootschalig centraal gestuurd invoeren van AI-toepassingen een illusie is. Dit zal uiteraard erg gevoelig liggen bij de instelling maar zeker ook bij de docenten. Bij afwezigheid van een dergelijk coördinatiemechanisme zal de invoering dus *bottom up* moeten plaatsvinden. Maar op het niveau van de instellingen is het evenmin één partij die de dienst uitmaakt. Naast het bevoegd gezag zijn er de individuele leraren die een grote macht hebben. En daarnaast hebben we de nog de ouders, de leerlingen en de medezeggenschapsraad. Daar komt nog bij dat de posities van deze partijen worden beschreven (en beschermd) door tal van lagere en vaak *verschillende* regelingen (besluiten, statuten etc.) met tal van eigen processen (met beroepsmogelijkheden) die afgestemde besluitvorming – een aantal scholen wil bijvoorbeeld gaan samenwerken – over een gevoelig onderwerp als dit niet gemakkelijk zullen maken.

Kortom, de Nederlandse governancestructuur in het onderwijs zowel op systeem als op instellingsniveau zal naar verwachting een *praktisch* obstakel vormen voor het grootschalig uitrollen en toepassen van AI.

3.5 Conclusie: mogelijke juridische obstakels voor AI in het onderwijs

Onderstaand schema (Tabel 2) geeft een overzicht van de verkaveling van de regelingen over de diverse gebeurtenissen.

²³ Zie onder meer artikel 8:4 Awb lid 3 onder b dat bepaalt dat geen beroep openstaat tegen beoordelingsbeslissingen genomen in het openbaar onderwijs ten aanzien van het kennen en kunnen van leerlingen.

Tabel 2 Verkaveling van gebeurtenissen met betrekking tot AI in het onderwijs over regelingen

	Aw	Dw	AVG	Awb	Wob en Who	BW	Mw	Onderwijsrecht
1. Scheppen en verbeteren van AI infrastructuur	✓					✓	✓	✓
2. Verzamelen en invoeren van data		✓	✓		✓			✓
3. Doen van data-analyses			✓					✓
4. Verzamelen van uitkomsten		✓	✓		✓			✓
5. Handelen op basis van uitkomsten			✓	✓	✓	✓		✓
6. Verzamelen en invoeren van (effecten) van handelen		✓	✓		✓			✓

Het geheel overziend kan geconcludeerd worden dat de vraag 'wat mag er?' steeds zal afhangen van de concrete toepassing van AI in het onderwijs. Niettemin kan hier in algemene zin het volgende over gezegd worden:

1. Toepassing van AI in het onderwijs raakt aan diverse generieke regelingen van diverse pluimage. Sommigen van die regelingen zijn van aanvullend recht, dat betekent dat daar gemakkelijk en veelal per contract van kan worden afgeweken, zoals bijvoorbeeld de regelingen van het Auteurswet en het Databankenrecht.
2. De flexibiliteit van genoemde privaatrechtelijke regelingen vindt men veel minder terug in de regelingen van publiekrecht, zoals de regels over openbaarheid en hergebruik van bestuurlijke respectievelijk overheidsinformatie. Toepasselijkheid van deze regelingen kunnen ontwikkelaars ervan weerhouden samenwerkingsverbanden met instellingen aan te gaan als dit zou betekenen dat hun kennis op straat komt te liggen. Aanvullende contracten over intellectuele rechten kunnen hier (gedeeltelijk) de helpende hand bieden. Overigens zal men er daarbij wel voor moeten waken dat samenwerking met marktpartijen uiteindelijk geen machtsposities doen ontstaan die in strijd kunnen zijn met regels over mededinging.
3. Besluiten die instellingen nemen worden uiteraard volledig beheerst door de algemene regels van het bestuursrecht, waaronder de algemene beginselen van behoorlijk bestuur. Toepassing van deze beginselen op AI-toepassingen kan lastig zijn omdat deze inherent botsen met de black box die AI nu eenmaal schept. Maar het lijkt geen onoplosbaar probleem.
4. De regels over de bescherming van persoonsgegevens (AVG) zijn dit mogelijk wel. De aard van de persoonsgegevens in het onderwijs en het karakter van AI gaan heel slecht samen en het lijkt heel moeilijk voorstelbaar dat grootschalige AI toegestaan is zonder wettelijke regeling. Uiteraard verandert dit geheel als de te gebruiken gegevens geen persoonsgegevens meer zijn.
5. Mocht er wat misgaan dan lijken er geen noemenswaardige problemen bij de toepassing van de regels over aansprakelijkheid. Dit regime is zeer open en flexibel zodat zij die door AI-toepassingen benadeeld zijn niet op de tocht hoeven te staan.

6. De onderwijs-specifieke sectorale regels laten op zich veel vrijheid en lijken niet echt harde verboden te bevatten. Wel zal steeds binnen de geest en letter van de beginselen van onderwijsrecht gehandeld moeten worden. Daar komt nog bij dat de belangen van het kind steeds een centraal en ethisch ijkpunt zullen moeten vormen en de vraag zal steeds zijn of dat met AI het geval is.
7. In meer praktische zin vormt de sectorale regelgeving wel een obstakel: voor top down sturing op inhoud van onderwijs bestaat geen sterk arrangement, nog afgezien van het feit dat dit heel gevoelig ligt. Ook de beweging van onderaf is een lastige: weliswaar kunnen er bij individuele instellingen initiatieven ontstaan maar het poldermodel en de diversiteit aan verschillende regelingen op het niveau van de instelling zal een grootschalige beweging waarschijnlijk in de weg staan. Dit nog afgezien van de mogelijke aversie van docenten (en instellingen), aan wiens belangen en positie toepassing van AI direct kan raken.

Bevinding 5 Er zijn concrete knelpunten op juridisch vlak die aandacht verdienen

Hierboven wordt feitelijk geconcludeerd dat het juridisch kader, in strikte zin, ruimte laat om AI toe te passen. Die toepassing raakt echter wel aan tal van fundamentele, welhaast ethische vragen, omdat het een impact heeft op het krachtenveld waar het juridisch regime in feite een uitvloeisel van is. Denk hierbij bijvoorbeeld aan:

- de relatie tussen de leerling en de leraar: de leraar krijgt concurrentie, de leerling krijgt een alternatief;
- de relatie tussen de instelling en de leverancier van de AI: toegang tot en gebruik van met AI gegenereerde kennis wordt cruciaal;
- de relatie tussen de leraar en de instelling: de instelling krijgt mogelijk meer zeggenschap over inhoud en proces, omdat het de gesprekspartner is van de leverancier van de AI;
- de relatie tussen het Ministerie (inclusief Inspectie) en de instellingen (als geheel): toezicht- en controlemogelijkheden veranderen en bescherming van sommige belangen zal op systeemniveau moeten plaatsvinden.

Omdat het krachtenveld verandert, is het goed denkbaar dat de belangen die de rechtsregels beschermen op de tocht komen te staan en/of dat er nieuwe belangen bijkomen waarin de huidige rechtsregels niet voorzien. Het zou daarom raadzaam zijn een 'impactanalyse' uit te voeren waarbij *de waarden* (die we willen blijven dienen) centraal stellen en vervolgens nagaan of deze gevaar lopen en daarom versterkt of anders ingevuld moeten worden.

Dit zou praktisch ingevuld kunnen worden door de diverse stakeholders uit dit krachtenveld samen te brengen en hen te confronteren met een aantal scenario's (toepassingen van AI) en vast te stellen 'waar ze op aanslaan': dat immers vertegenwoordigt de waarde waar ze voor staan. Aansluitend kan dan onderzocht worden wat er volgens hen zou moeten gebeuren waarbij het hele instrumentarium van interventies – regelgeving, contracten, aanbesteding, praktische maatregelen - ingezet zou mogen worden.

4 Wat willen we?

In dit hoofdstuk gaan we nader in op mogelijke beleidsopties omtrent de ontwikkeling en adoptie van AI in het onderwijs. Eerst brengen we de impact voor onderwijsinstellingen, leerlingen, docenten en het onderwijsstelsel in kaart. Vervolgens gaan we nader in op ethische aspecten omtrent AI in het onderwijs. Hieruit volgt of er beleid nodig is, bijvoorbeeld om bepaalde gewenste effecten te stimuleren en ongewenste effecten te beperken. Tot slot kijken we naar beleidsontwikkeling in het buitenland om te zien waar en op welke manier maatregelen worden genomen om AI in het onderwijs te sturen.

4.1 Mogelijke impact van AI in het onderwijs

4.1.1 AI en onderwijsinstellingen

AI kan worden geschaard onder de digitale leermiddelen. Om AI succesvol toe te kunnen passen moeten er aan soortgelijke randvoorwaarden worden voldaan. Op het laagste niveau dient er een goede digitale infrastructuur zijn. De digitale infrastructuur bestaat ten minste uit een breedbandige internetverbinding en apparaten voor leerlingen en docenten. Ook dient er data beschikbaar te zijn om AI-modellen te trainen. Dit is data uit administratieve systemen, (meta)data over leermiddelen, en data die wordt gegenereerd door het gebruik van (digitale) leermiddelen.

De randvoorwaarden voor succesvolle AI-toepassingen zijn niet alleen technisch van aard. Ook dienen docenten, schoolleiders en bestuurders de meerwaarde van AI-toepassingen kunnen inzien en *verantwoord* kunnen inzetten. Hiervoor is het noodzakelijk dat personeel in het onderwijs voldoende digitale vaardigheden bezitten. Vervolgens kunnen er vakspecifieke AI-toepassingen worden ingezet in het leerproces. Naarmate men voldoende ervaring heeft met het verantwoord toepassen en een AI kan een instelling gaan naar een holistische integratie van verschillende AI-systemen.

4.1.2 AI en leerlingen

De impact van AI zal verschillen tussen de verschillende onderwijssectoren. In het PO kan AI ertoe leiden dat leerlingen minder vaak een klassikale uitleg zullen krijgen van lesstof, maar vaker een uitleg die gepersonaliseerd is naar moeilijkheidsgraad en leerstijl. Dit betekent, afhankelijk van de ontwikkelingen op het gebied van digitale leermiddelen, dat leerlingen in het klaslokaal meer tijd achter een computer (of tenminste een scherm) zullen doorbrengen. Desondanks verwachten we dat er nog steeds veel persoonlijk contact (nodig) zal zijn tussen leerling en docent.

In het VO zal de impact van AI voor de leerling minder groot zijn ten opzichte van het PO. AI-toepassingen zullen daar met name terug te vinden zijn in de werkvormen waar nu al zelfstandig wordt gewerkt. AI-systemen zijn op dit moment (en in de nabije toekomst) vooral geschikt om feitelijke kennis aan te leren en te toetsen. Vanwege de discrete aard van bètavakken zal de impact van AI voor deze vakken groter zijn dan voor alfavakken. Een AI is beter in het beoordelen of een wiskundesom correct is opgelost dan of de argumenten in een betoog hout snijden.

In het HO kan de impact van MOOC's mogelijk zeer groot zijn. Wanneer MOOC's AI gaan integreren is het niet onaannemelijk dat via een MOOC dezelfde kennis en vaardigheden: sneller, tegen lagere kosten, en van een meer prestigieuze onderwijsinstelling kunnen

worden vervaardigd dan in het traditionele hoger onderwijs. Wanneer dit het geval is zal fysieke aanwezigheid niet meer nodig zijn en kunnen studenten daardoor eenvoudiger het curriculum van internationale onderwijsinstellingen volgen.

4.1.3 AI en docenten

In het Nederlandse onderwijssysteem neem de docent een centrale rol in. De docent mag zelf bepalen hoe hij of zij de les inricht. Het ontwerp van een AI moet daarom vanuit een design thinking perspectief worden benaderd. Bij Design thinking staat de wens van de gebruiker centraal. Om AI succesvol in het onderwijs te implementeren moet de AI flexibel zijn in het gebruik en moet aansluiten bij verschillende leerstijlen van een docent.

De grootste impact van AI zal, naast uiteraard de gebruikers van de toepassingen (leerlingen en studenten) worden ervaren door docenten. Een deel van de taken van docenten kunnen worden overgenomen door AI, een deel van de taken zullen belangrijker worden ten opzichte van de huidige situatie, en er zullen ook een aantal nieuwe taken bijkomen. Onder de taken die kunnen worden overgenomen zien we voornamelijk het uitkiezen van lesmateriaal en het nakijken van toetsen. Ook het geven van feedback op basis van prestaties kan worden overgenomen door een AI. Doordat een deel van de taken kunnen worden overgenomen door een AI maakt dit ruimte vrij voor de docent om meer aandacht te besteden aan taken die lastig/niet over te nemen zijn door een AI. Denk hierbij aan meer sociale taken als begeleiding en coaching. Ook zullen er door de toepassing van AI nieuwe taken ontstaan. Om een AI verantwoord toe te kunnen passen dient een docent output van een AI kunnen interpreteren en vertalen naar concrete handelingen.

4.1.4 AI en het onderwijsstelsel

De impact van AI op het onderwijsstelsel kan potentieel groot zijn. Binnen het huidige onderwijsstelsel worden leerlingen uit praktische noodzaak gegroepeerd op competenties. Dit maakt het eenvoudiger voor docenten om op niveau les te geven. Wanneer een AI wordt ingezet voor gepersonaliseerd leren zal de nadruk van generieke 'onderwijsniveaus' (vmbo, havo, vwo) naar een nadruk op vakken.

AI kan ook de noodzaak van gestandaardiseerde toetsen wegnemen. Eén kritiek op gestandaardiseerde toetsen is dat dit een momentopname betreft, en niet in alle gevallen een goede representatie is van het presteren van een leerling. Met AI kunnen de prestaties van een leerling gedurende het gehele leerproces worden gemonitord. Hierdoor is het niet meer nodig om te toetsen op specifieke momenten en ontstaat een beter beeld van de capaciteiten van een leerling.

4.2 Ethische afwegingen

Los van de vraag of bepaalde AI-toepassingen *mogelijk* of juridisch *toegestaan* zijn, is het belangrijk om te bepalen of bepaalde toepassingen wel *wenselijk* en *ethisch verantwoord* zijn. In de onderwijssector en daarbuiten leeft het beeld dat deze keuze uiteindelijk bij docenten en scholen ligt. We bespreken hieronder enkele ethische aspecten die specifiek relevant zijn bij de toepassing van AI in het onderwijs.

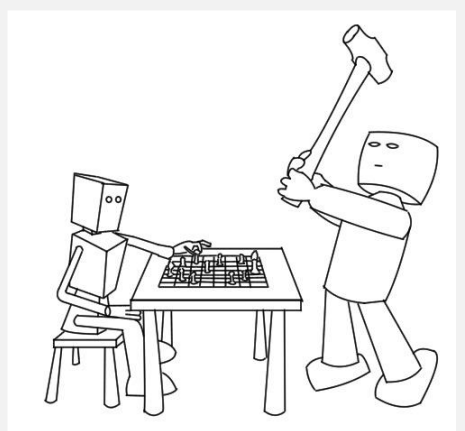
4.2.1 Relevantie

Normen en waarden weerspiegelen de wensen van een samenleving. Niet alle normen zijn vastgelegd in de wet. Normen en waarden zijn niet alleen van toepassing op vormen van gebruik maar kunnen ook (bewust of onbewust) verwerkt zijn in een technologie. Zo kan een ingenieur bij het ontwerp bepaald gebruik voor ogen hebben (in lijn met de ingenieur

zijn normen en waarden) die niet noodzakelijk hetzelfde zijn als die gebruiker van de gebruiker.

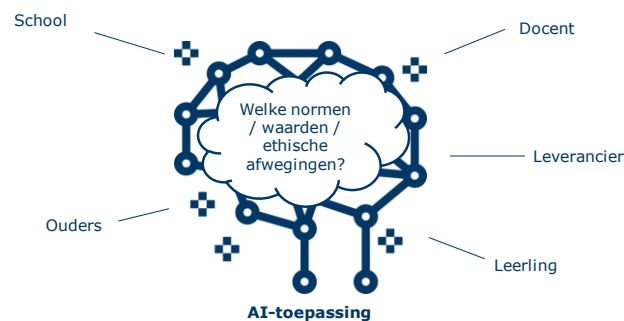
Achtergrond 2. Ethiek en AGI

Elizier Yudkowsky (2001) beschrijft een aantal mogelijke ongewenste uitkomsten door verkeerd gebruik van AI. [42] Wanneer sociale normen en waarden niet expliciet in een AI verwerkt zijn kan een simpele instructie aan een AI zeer schadelijke gevolgen hebben. Yudkowsky illustreert dit aan de hand van een AI die de vrij onschuldige opdracht heeft: "Speel zo goed mogelijk schaak". Wanneer de AI geen sociale normen en waarden kent kan de AI zomaar al onze infrastructuur vernietigen en in zichzelf verwerken om meer computerkracht te verkrijgen (en dus beter te kunnen schaken). Wanneer deze AI tegen een ander schaakcomputer speelt is het zomaar mogelijk dat de AI berekent dat de grootste kans om te winnen wordt gerealiseerd door zijn tegenstander te vernietigen. In deze twee situaties heeft een vrij onschuldige taak ("speel zo goed mogelijk schaak") vergaande ongewenste gevolgen.



Zoals eerder aangegeven is een dergelijk scenario, dat AGI veronderstelt, in de nabije toekomst niet aan de orde. Desondanks stimuleert het denken over dergelijke extreme scenario's oordeelvorming rondom inzet van minder geavanceerde vormen van AI.

In de onderwijscontext zijn verschillende stakeholders betrokken (primair de docent/instelling, leerling en ouders) welke ieder hun eigen normen en waarden hebben. In een AI-toepassing zitten eveneens (impliciet) normen en waarden besloten die niet noodzakelijkerwijs overeenstemmen met de normen en waarden van de verschillende betrokkenen (Figuur 11).



Figuur 11 Verschillende stakeholders in het onderwijs en de normen en waarden van een AI-toepassing

4.2.2 Afwegingen

In dit onderzoek zijn verschillende ethische afwegingen geïdentificeerd die relevant zijn bij het gebruik van AI in het onderwijs. Aan de hand van deze set kunnen betrokkenen een oordeel vellen over AI-toepassingen door de afwegingen te beoordelen binnen het eigen normen en waardenkader. Merk op dat veel van de genoemde afwegingen niet nieuw zijn, maar ook in algemene zin spelen (hetzij voor AI-toepassingen, hetzij voor onderwijs in het algemeen).

Bevinding 6. Als het om AI gaat, meten we met twee maten.

De introductie van AI-toepassingen in diverse sectoren wakkert de maatschappelijke discussie over ethische aspecten aan. Opvallend is daarbij de wijze waarop de ethische discussie wordt geframed, en langs welke meetlat AI-toepassingen worden gelegd. We zien in de praktijk dat aan AI-toepassingen in sommige gevallen *hogere* eisen worden gesteld dan aan mensen in soortgelijke scenario's. Twee voorbeelden zijn hierbij illustratief:

- Amazon kwam in het nieuws doordat een AI-toepassing die zij hadden ontwikkeld ter ondersteuning van het aannameproces bleek te discrimineren. De AI was echter getraind op historische gegevens, en repliceerde slechts een bestaande bias in het menselijke selectieproces. Hoewel de discussie over de AI werd gevoerd, had deze net zo goed over het menselijke proces kunnen worden gevoerd.
- Onlangs was een geval in het nieuws waarbij een man dronken achter het stuur van een zelfrijdende auto zat – de auto bleef autonoom rijden en kon door de politie slechts met moeite worden stilgezet. Er ontstond discussie over het feit dat de auto bleef doorrijden, en over de vraag of zelfrijdende technologie wel klaar is voor gebruik in de echte wereld. In de discussie werd voorbij gegaan aan het feit dat, was de bestuurder in een *normale* auto gestapt, dit waarschijnlijk niet goed was afgelopen.

Transparantie, verklaarbaarheid en voorspelbaarheid

Wanneer een AI wordt ingezet ter ondersteuning van, of voor het volledig autonoom maken van beslissingen, dan delegeert de gebruiker in feite een deel van de ethische afwegingen aan een AI. Wanneer een bestuurder van een zelfrijdende auto de controle overdraagt aan de auto is deze ethisch gezien niet 'ontslagen' van de gevolgen. Het is daarom belangrijk dat de gebruiker hiervan op de hoogte is, de mogelijke uitkomsten overziet, en kan beoordelen of deze passen binnen zijn ethisch kader. Van een AI-toepassing moeten de volgende zaken bekend zijn:

- Op welke data is de AI-toepassing getraind, en is het mogelijk dat zich hierin bepaalde biases voordoen (die kunnen leiden tot een bias in beslissingen of conclusies van de AI)?
- In hoeverre blijft de AI leren, en hoe is gewaarborgd dat er geen bias ontstaat?
- Hoe is het achterliggende algoritme van de AI opgebouwd, en hoe gaat de AI om met voorbeelden die het niet eerder heeft gezien? Onder welke omstandigheden is het algoritme wel of niet valide?

Onomkoopbaarheid

Een AI moet (zover dat mogelijk is) een objectieve keuze maken en mag niet van buitenaf beïnvloedbaar zijn. Wanneer een AI te manipuleren is kan hier misbruik van worden gemaakt. Wanneer een AI om de tuin kan worden geleid door leerlingen worden de uitkomsten van AI gebruik niet meer in lijn met het oorspronkelijke gebruik. Een van de meest voorkomende kritieken op AES-systemen is dat deze systemen studenten belonen voor bepaald woordgebruik en niet zozeer op inhoud. Zo kan een AES-systeem om de tuin worden geleid door taalgebruik in te zetten dat wordt geassocieerd met complexe redeneringen en niet te focussen op de inhoud. Denk hierbij aan overmatig gebruik van woorden zoals "niettemin" en "desondanks" in plaats van "maar" of "toch".

Privacy en zelfbeschikking

Het gebruik van AI-toepassingen kan strijdig zijn met privacy en zelfbeschikking van betrokkenen – het (deels) modelleren van een leerling kan worden gezien als een inbreuk hierop, zeker wanneer geen beperkingen worden gesteld aan het mogelijke gebruik van een dergelijk model.

4.2.3 Ethische denkkaders

Bovenstaande aspecten kunnen vanuit verschillende ethische denkkaders worden benaderd en ingevuld. Vanuit een Kantiaans perspectief worden vooraf regels bepaald, welke een AI-toepassing in geen enkel geval zal mogen overtreden: het doel heiligt nooit middelen. In een meer utilitaristisch perspectief telt primair de *uitkomst*: als een onvoorspelbare AI wel aantoonbaar substantieel beter presteert dan alternatieven, zou het wellicht juist *onethisch* zijn om deze *niet* in te zetten. Deugdethiek redeneert tot slot vanuit de intrinsieke eigenschappen van de toepassing, wat in deze context betekent dat juist de ontwerpprincipes van de AI langs de ethische meetlat moeten worden gelegd. In de deugdethiek staan daarnaast doelen (teleologie) centraal. Vanuit dit perspectief kan men spreken van ethisch gebruik van AI wanneer deze voor *deugdelijke* doeleinden wordt ingezet. Hieronder gaan we dieper in op hoe deze ethische denkkaders kunnen worden toegepast op een intransparant adaptief leersysteem.

Wanneer een AI geïntegreerd is met een adaptief leersysteem kan, vanuit het Kantiaanse perspectief, niet van ethisch gebruik worden gesproken wanneer de achterliggende algoritmes niet bekend zijn. Een AI-toepassing zal, hoe goed deze ook is, namelijk altijd aan de ethische principes moeten voldoen. Zelfs al leert de leerling aanzienlijk sneller in de adaptieve omgeving is het gebruik niet ethisch verantwoord omdat het niet transparant is hoe de leerstof wordt gepersonaliseerd.

In een utilitaristisch perspectief wordt er een afweging gemaakt tussen de voor- en nadelen. In het geval van een intransparant adaptief leersysteem zal de afweging moeten worden gemaakt of de verbetering in het leren opweegt tegen het intransparante. Wanneer er consensus is tussen alle stakeholders dat het gebruik van de intransparante AI dusdanig veel opbrengsten met zich meebrengt ten opzichte van een situatie zonder AI het ethisch gerechtvaardigd is om de AI toe te passen.

In de deugdenethiek staan de doelen centraal. In het geval van de intransparante adaptief leersysteem kan men zich afvragen of het voor een deugdelijk doel wordt ingezet. Zet een school het systeem in om de prestaties van de leerling te verbeteren of om te bezuinigen op de leerkracht populatie? In het eerste geval zou men kunnen spreken van een deugdelijk doel, in het tweede geval kan men zich afvragen of het toepassen van intransparante systemen voor kostenreductie een deugdelijk doeleinde is.

4.3 Beleidsontwikkeling in het buitenland

AI is een mondiale ontwikkeling: hoewel het onderwijs in Nederland op een specifieke manier is ingericht hoeven niet alle technologische innovaties van eigen bodem te komen. De drie landen met de sterkste AI-regio's in Europa zijn het Verenigd Koninkrijk, Duitsland en Frankrijk. [19] Daarnaast komt Amsterdam ook voor in de top 10 AI startup hubs van Europa, samen met London, Berlin, Paris, Madrid, Stockholm, Copenhagen, Barcelona en Dublin. We zien internationaal dat er op het gebied van AI en onderwijs drie trends spelen:

1. Ontwikkeling specifieke applicaties van AI (zie paragraaf 2.3 (Ontwikkeling in het buitenland))
2. Beleid om experimenten met, en adoptie van AI in het onderwijs te stimuleren

3. Beleid omtrent AI-literacy

Beleid om experimenten met, en adoptie van AI in het onderwijs te stimuleren

In recente jaren zien we steeds meer beleid ter bevordering van de ontwikkeling en adoptie van AI in het onderwijs. Concreet zien we dat de drie Europese leiders op het gebied van AI beleid voeren om AI in het onderwijs te stimuleren. In het VK zien we in de Industrial Strategy dat er 20 miljoen pond beschikbaar wordt gesteld voor initiatieven die AI integreren in de publieke diensten. Deze publieke diensten zijn niet alleen gelimiteerd tot onderwijs maar ook in de gezondheidszorg en op het gebied van transport en milieu.

In Duitsland is er een speciaal Educational Technology Lab is opgericht door het Duitse onderzoekscentrum voor AI (DFKI). In dit EdTech lab wordt er geëxperimenteerd met toepassingen van AI in het onderwijs. Frankrijk is wederom een speciaal geval. De Franse president Emmanuel Macron heeft de ambitie uitgesproken om van Frankrijk een AI wereldspeler te maken die kan concurreren met de VS en China. In lijn van deze ambitie zien we dat de Franse overheid heeft aangekondigd om in de komende 4 jaar ruim 1.8 miljard dollar te investeren in het Franse AI-ecosysteem. Daarnaast zien dat het Franse ministerie voor onderwijs 1 miljard euro in 3 jaar zal investeren in de ontwikkeling van digitale leermiddelen²⁴. Dit geld zal geïnvesteerd worden in:

1. Training van leraren
2. Digitale infrastructuur in de klas zoals digitale connectiviteit en devices voor leerlingen.
3. Innovatie en incubatie van nieuwe applicaties

In Finland zien we de Finse overheid anticipeert op de transformerende impact die AI gaat hebben op het onderwijs (en onderwijsstelsel). De Finse overheid merkt op dat er reeds hoogwaardige AI-systemen bestaand die voornamelijk worden toegepast in het onderwijzen van bèta vakken zoals wiskunde. De Finse overheid is daarom bezig met het ontwikkelen van nieuwe manieren van onderwijzen. In het rapport *Finland's age of Artificial Intelligence: Turning Finland into a leading country in the application of artificial intelligence* [20] drie onderwijsinnovaties genoemd die door toedoen van AI kunnen worden gerealiseerd:

1. **De ontwikkeling van nieuwe online vakken.** Deze nieuwe online vakken zijn niet alleen nieuw in de manier waarop ze worden gegeven (online i.p.v. fysiek) maar ook door de innovatieve inhoud van het vak. Denk hierbij aan vakken zoals programmeren.
2. **Nieuwe vormen van kwalificatie.** De noodzaak voor toetsen en diploma's neemt af. Er zal een grotere nadruk komen op kwalificatie voor specifieke vaardigheden.
3. **Virtuele instellingen.** Fysieke scholen zijn geen vereiste meer om onderwijs te volgen. Onder virtuele instellingen verstaan we bijvoorbeeld MOOC-platformen.

De Finse overheid is hierin vrij uniek. Hoewel het merendeel van de landen een strategie ontwikkeld op het gebied van AI zien we dat er maar weinig aandacht is voor de rol van AI in het onderwijs. Wanneer er aandacht is voor AI dan is het vaak het *subject* en niet het *object*. Men is van mening dat AI moet worden onderwezen, maar spreekt zich niet uit over het toepassen van AI in het onderwijs.

²⁴ Deze 1 miljard euro zal niet alleen voor AI toepassingen maar voor digitale leermiddelen breed worden ingezet.

Beleid omtrent AI-literacy

Vrijwel alle EU-lidstaten hebben zich uitgesproken over de noodzaak van onderwijs over AI. Hoewel dit niet noodzakelijkerwijs betekend dat dit onderwijs ook met een AI zal plaatsvinden is dit wel een stap naar algemene AI-literacy en bewustzijn. Ook de docenten van de toekomst zullen onderwijs over AI krijgen. Er heerst nu een hype omtrent AI en goede scholing over de (on)mogelijkheden is nodig voor verantwoord AI gebruik. Onder AI-literacy verstaan we onder andere: kennis van algoritmes, kennis van data, de kansen van AI, de risico's van AI, mogelijke bijeffecten die kunnen ontstaan door toepassen van AI.

5 Conclusie

AI is een technologie die op verschillende manieren voor veel verschillende doeleinden kan worden ingezet in het onderwijsdomein. Hieronder gaan we op de 5 grootste kansen en risico's voor de toepassing van AI in het onderwijs die in dit onderzoek zijn geïdentificeerd.

5.1 Kansen van AI in het onderwijs

Het verminderen van de werkdruk van docenten door toepassing AI ter ondersteuning van (administratieve) taken

De grootste kans voor AI in het onderwijs is de rol die AI kan spelen bij het oplossen van problemen die leerkrachten ervaren. In recente jaren heeft de docent vaak ontevredenheid getoond met betrekking tot de hoge werkdruk die wordt ervaren in het onderwijs. Denk hierbij aan initiatieven zoals 'PO in actie'. Deze verhoogde werkdruk is mede ontstaan door de additionele administratieve taken die leerkrachten bij het bestaande takenpakket hebben gekregen als door de verminderde instroom van PABO studenten. AI kan de docent ondersteunen door (administratieve) taken te automatiseren en zo de werkdruk te verlagen.

Taken waarvan we verwachten dat AI deze in de nabije toekomst sterk kan automatiseren en/of vergemakkelijken zijn met name gerelateerd aan nakijkwerk (bijvoorbeeld het markeren van sterke en zwakke punten in een essay, waarna de docent primair deze punten beoordeelt) en samenstelling van het lesmateriaal (met behulp van automatische classificatie van inhoud).

Gepersonaliseerd leren: onderwijs beter laten aansluiten bij de leerling, met zowel betere uitkomsten als een beter leerproces

Ook zien we dat de rol die AI kan spelen in gepersonaliseerd leren als grote kans wordt aangemerkt. Docenten hebben maar beperkte tijd en aandacht en kunnen om deze reden niet elke leerling individueel onderwijzen. AI heeft deze begrenzing niet. Hierdoor kan een AI het onderwijs beter laten aansluiten bij de wensen van de leerlingen. Dit heeft als resultaat dat de docent zijn/haar aandacht beter kan richten op 'probleemleerlingen', en de leerling op eigen tempo en niveau door de lesstof gaat. Het automatiseren van taken zal gestaag verlopen. We verwachten dat het zal beginnen bij het uitvoeren van kleine taken zoals het uitkiezen en oefenen van lesmateriaal. Denk hierbij aan verbeteringen binnen adaptieve leersystemen door de implementatie van deep learning algoritmes.

Het ondersteunen van de docent met holistische, onderbouwde inzichten (learning analytics)

Ook kan AI de docent ondersteunen door data te combineren en inzichtelijk te maken. Met deze *learning analytics* kan de docent holistische en onderbouwde inzichten in leerlingen krijgen. AI kan vooroordelen in gedrag blootleggen en zo ook het onderwijs eerlijker maken ten opzichte van etniciteit of geslacht. Denk hierbij aan het doen van een schooladvies: een AI die niet op de hoogte is van etniciteit of geslacht kan deze variabelen ook niet meenemen in het advies. Het is reeds bekend dat leerlingen met een niet-westerse migratieachtergrond over het algemeen een lager schooladvies krijgen dan autochtone leerlingen. Een AI kan mogelijke (onbewuste) vooroordelen van een docent corrigeren.

Het verbeteren van de manier van toetsen van kennis

Op het gebied van toetsing kan AI van een periodieke naar een continue toetsing. Als kritiek op gestandaardiseerde toetsen wordt vaak genoemd dat deze geen goede representatie is van de kennis die een leerling bezit maar slechts een momentopname is. Met AI kan continue het kennisniveau van een leerling worden bijgehouden zonder dat hier een formeel toetsmoment aan te pas komt.

Het vergroten van de effectiviteit van digitale leermiddelen, ook in synergie met andere technologieën, zoals VR en serious games

Tot slot kan AI de effectiviteit van bestaande digitale leermiddelen worden vergroot door binnen deze leermiddelen door een met de hand geprogrammeerde laag te vervangen met een AI laag. In combinatie met nieuwe technologieën, zoals VR, AR en serious games, kan er synergie plaatsvinden door de toepassing van AI. Deze technologieën creëren een virtuele ruimte die volledig te controleren is, waardoor de leeruitkomsten sterk kunnen worden geoptimaliseerd. Een mogelijk risico hierbij is dat kennis die wordt opgedaan in een virtuele omgeving in beperkte mate generaliseert naar 'de echte wereld', en dat andere, minder geformaliseerde kennis, die een leerling opdoet op school (bijvoorbeeld op sociaal vlak) onderbelicht wordt.

5.2 Risico's van AI in het onderwijs

Sommige onderwijsdoelen kunnen door AI in het geding raken wanneer de focus op techniek te groot is.

Wanneer AI te veel wordt benaderd vanuit de data/techniek bestaat er een risico dat AI toepassingen niet meer aansluiten bij onderwijsdoelen. Zo kan een AI die geoptimaliseerd is op kennisoverdracht andere elementen zoals motivatie niet hebben geïncorporeerd. Ook bestaat er het risico dat buitenlandse technologiebedrijven de Nederlandse onderwijsmarkt betreden. Het is niet vanzelfsprekend dat deze bedrijven dezelfde onderwijsdoelen voor ogen hebben. Hierdoor kunnen we de controle verliezen op het onderwijs.

Wanneer het onderwijs compleet geautomatiseerd is en leerlingen zonder tussenkomst van mensen het onderwijstraject doorloopt kan men zich afvragen of het onderwijs zo niet onmenselijk wordt. Dit risico wordt groter wanneer er simulaties worden ingezet om lesmateriaal over te brengen. Hierbij bestaat de kans dat kennis en vaardigheden geleerd in een simulatie niet in te zijn in de realiteit.

Bias in mensen vertaalt naar data, die vervolgens wordt overgenomen door AI.

Een groot risico bij het inzetten van AI heeft betrekking op de data die beschikbaar is voor het trainen van modellen. Naast de beschikbaarheid van data is er ook het risico dat de data die beschikbaar is vooroordelen bevat. Historische data bevat vaak vooroordelen ten opzichte van etniciteit en geslacht. Het risico bestaat hier dat AI deze vooroordelen overneemt en zelfs versterkt. Hierdoor kan er een feedback-loop ontstaan die structureel bepaalde minderheden benadeeld.

Een verslechterd vooruitzicht bij 'docent' als beroep (voor huidige en toekomstige docenten).

Het risico bestaat dat de docent mogelijk nieuwe vaardigheden nodig heeft om AI goed in te kunnen zetten. Op dit moment is data gedreven onderwijs geen onderdeel van lerarenopleidingen terwijl een AI wel een basiskennis op dit gebied van de gebruiker verwacht. Indien docenten niet meekunnen in deze ontwikkeling bestaat er de kans dat de docent van de toekomst een data scientist is met een minder sterke onderwijskundige achtergrond. Het

risico dat 'de docent' wordt weg-geautomatiseerd is echter zeer klein. Wil een AI alle taken kunnen overnemen van een docent, dan is een vorm van AGI nodig, waarvan we in dit onderzoek vaststellen dat dit in de nabije toekomst niet realistisch is.

Afhankelijkheid van black-box modellen (unexplainable AI) versus verantwoordelijkheid van de docent.

Er bestaat het risico dat we afhankelijk worden van AI-systemen die we niet compleet begrijpen. Dit begrip komt van twee kanten. Zo moet zowel de verklaarbaarheid van machine learning modellen als de kennis van docenten (met betrekking tot AI) moeten worden vergroot om dit risico te overkomen. Een docent kan geen verantwoordelijkheid nemen voor systemen die niet te begrijpen zijn.

Toepassing van AI terwijl andere basisvoorzieningen nog niet op orde zijn.

AI vereist een aantal voorwaarden waaraan moet worden voldaan eer het succesvol kan worden toegepast. Zo is er een technische infrastructuur en data nodig voordat er pas aan AI kan worden gedacht. Wanneer er te snel AI wordt implementeert bestaat er de kans dat de AI niet aan de verwachtingen kan voldoen. Hierdoor kan AI, net als andere innovatieve leermiddelen, voor een grote teleurstelling zorgen. Deze teleurstelling kan op zijn duur weer zorgen voor een aversie van AI in het onderwijs en mogelijk ook andere innovatie belemmeren. Daarnaast zijn er op juridisch vlak een aantal belemmeringen (veelal met goede reden) die het onmogelijk maken om AI van de één op andere dag sterk door te voeren.

Een machtsverschuiving bij de producenten van lesmateriaal.

Er bestaan zorgen dat de introductie van AI de macht consolideert bij één partij. Deze partij kan een bestaande uitgever van lesmateriaal of een techgigant zijn. Recente ontwikkelingen lijken er echter op te wijzen dat dit risico te overzien is. In een recente uitspraak is bevestigd dat het eigendom van data die wordt gegenereerd in digitale leermiddelen eigendom is van de school, en niet noodzakelijk van de ontwikkelaar van het lesmateriaal. Het blijft echter een aandachtspunt voor beleid.

5.3 Aanbevelingen voor beleid

Het lijkt evident dat toepassing van AI in het onderwijs op diverse vlakken positieve impact kan hebben. Er lijkt vanuit de onderwijssector weinig weerstand, en zelfs enthousiasme, te zijn voor het inzetten van AI om bepaalde (met name administratieve en zeer specifieke) taken te automatiseren. Denk hierbij aan het nakijken van toetsen en het personaliseren van een curriculum. Tegelijkertijd zien we dat verdergaande inzet van AI vereist dat eerst bepaalde fundamentele juridische en ethische vragen moeten worden beantwoord.

Daarnaast staat de rol van de docent ten aanzien van de inzet van AI ter discussie. Aan de ene kant kunnen docenten zeer goed worden ondersteund met AI, en in staat worden gesteld om beter onderwijs te geven. Aan de andere kant wordt van docenten gevraagd verantwoordelijkheid te nemen voor de beslissingen van een (ondoorzichtige) AI, om te gaan met data en het op waarde schatten van conclusies uit dergelijke systemen, en de inzet van AI te kunnen uitleggen aan betrokkenen.

Om toepassingen van AI in het onderwijs te stimuleren is er een aantal aspecten waarop beleid kan worden gevormd, en/of waarop beleidsmakers (specifiek het ministerie van OCW) actie kunnen ondernemen:

1. Het stimuleren van acceptatie van AI in het onderwijs door leraar, leerling en ouder.

We stellen hierbij voor om scholen te vragen een visie op te stellen ten aanzien van inzet van AI-toepassingen binnen het eigen onderwijs. Deze visie maakt, zowel naar het eigen personeel als naar ouders (en indirect leerlingen) duidelijk hoe ver de school gaat, en wat het ambitieniveau is. Uiteraard moeten de scholen worden ondersteund bij het opstellen van een dergelijke visie. Een dergelijke AI-visie zal minstens de volgende elementen moeten omslaan:

- Waarom zet de school AI in?
- Voor welke taken wordt AI *wel* ingezet (en waarvoor *niet*)?
- Hoe wordt de verklaarbaarheid van de AI gewaarborgd?
- Hoe wordt er gewaarborgd dat de AI niet bevooroordeeld is?
- Hoe wordt er gewaarborgd dat onderwijsdoelen niet in het geding raken door AI?
- Hoe wordt de privacy van de leerling gewaarborgd?

2. Het vergroten van de digitale vaardigheden van onderwijzend personeel.

Digitale vaardigheden zijn een vereiste om AI verantwoord in te kunnen zetten. Dit is nodig voor alle onderwijssectoren (PO/VO/HO). Onder deze digitale vaardigheden verstaan we niet alleen het omgaan met een computer maar ook specifiek data vaardigheden die nodig zijn om de output van een AI te kunnen interpreteren. Denk hierbij aan sterke statistische basis en algemene kennis van enkele krachtige machine learning algoritmes. Concreet betekend dit dat er drie acties kunnen worden ondernomen:

- Er moet meer aandacht komen voor digitale vaardigheden op docenten opleidingen. Maak digitale vaardigheden een verplichting zoals als didactische en pedagogische vaardigheden.
- De huidige docent populatie moet worden bijgeschoold om verantwoord AI in te kunnen zetten.
- Het stimuleren van zij-instromers met een achtergrond in data of AI. Dit kan binnen een school een kettingreactie op gang zetten die de adoptie van AI kan bevorderen.

3. Het inrichten van een datainfrastructuur.

In de regel geldt dat AI-toepassingen veel data nodig hebben om optimaal te kunnen presteren. De beschikbaarheid van data is daarmee een cruciale factor voor de ontwikkeling van AI-toepassingen binnen het Nederlandse onderwijs. Mogelijk is een rol voor OCW weggelegd als het gaat om het stimuleren van het beschikbaar maken van data, en het stellen van kaders hierbij, ten behoeve van onderzoek naar en ontwikkeling van AI-toepassingen in het onderwijs.

Daarnaast speelt er nog de angst voor bevooroordeeldheid binnen data. OCW kan helpen met het verzamelen en openbaar stellen van deze databronnen die gebruikt kunnen worden om onbevooroordeelde AI modellen te trainen.

4. Het faciliteren van experimenten van inzet van AI in het onderwijs.

Hierbij moet het beleid voorzien in kaders (hoe ver mogen experimenten gaan). Beschikbaarheid van data voor analyse is een van de belangrijkste knelpunten welke vanuit beleid zou kunnen worden aangepakt. Experimenteren betekent ook dat er niet te strenge eisen moeten worden gesteld aan de *uitkomsten* van een pilot. Daarnaast moet er ook geld beschikbaar zijn om deze experimenten uit te kunnen voeren. Binnen bestaande onderzoeksbeurzen zou er extra ruimte moeten komen voor AI-experimenten.

5. Het stimuleren van een multidisciplinaire aanpak bij de ontwikkeling van AI.

Een multidisciplinaire aanpak is cruciaal om te waarborgen dat onderwijsdoelen niet in het geding raken, de wensen van de gebruiker in acht worden genomen, en aan de hoogste technologische standaarden wordt voldaan. Om deze multidisciplinaire aanpak te stimuleren kunnen er nieuwe onderzoeksbeurzen worden ontwikkeld die niet slechts gefocust zijn op één domein.

6. Het ontwikkelen van een keurmerk voor verantwoorde toepassing van AI in het onderwijs.

Om de complexiteit achter AI te vereenvoudigen voor stakeholders zoals scholen (als inkoper van de toepassingen), ouders en leerlingen, zou er een keurmerk kunnen worden ontwikkeld, waarbij een AI-toepassing aan bepaalde voorwaarden moet voldoen. Zo kan een gebruiker die geen AI-expert is wel een verantwoorde keuze maken bij de selectie van een AI-toepassing. Een dergelijk keurmerk zou aan tenminste de volgende aspecten aandacht moeten besteden:

- Het stimuleren van standaardisatie, dataportabiliteit, interoperabiliteit, teneinde de markt voor AI-oplossingen voor het onderwijs goed te laten functioneren. Merk hierbij op dat vergaande 'top down' standaardisatie innovatie juist zou kunnen afremmen.
- Het uitwerken van juridische kaders (scheppen van duidelijkheid) en het oplossen van juridische knelpunten bij de inzet van AI in het onderwijs.
- Kaders op het gebied van transparantie en verklaarbaarheid van een AI. In welke mate moet een AI transparant zijn, welke data is erin gegaan, met welk algoritme wordt het geanalyseerd, en aan welke eisen moet een verklaring van een AI-model voldoen. Praktisch kan OCW bijdragen aan het opstellen van benchmarks waaraan een AI minimaal aan moet voldoen. Denk hierbij aan benchmarks voor het nakijken van essays.

7. Het initiëren van vervolgonderzoek naar AI in het onderwijs.

In het voorliggende onderzoek is een verkenning uitgevoerd van de mogelijkheden van AI in het onderwijs. Hierbij zijn naast antwoorden ook een groot aantal nieuwe vragen gevonden, waarop op dit moment geen (goed) antwoord bestaat. We zien de volgende (niet uitputtende lijst van) richtingen voor vervolgonderzoek:

- Vanwege de omvang van het onderzoek konden niet alle stakeholders betrokken worden. Vervolgonderzoek kan meer inzoomen op de acceptatie van AI door leerlingen en ouders.
- Op juridisch vlak kan er in vervolgonderzoek dieper ingegaan worden op de impact van AI op specifieke onderwijs regelingen.

- Het buitenlandse beleid omtrent AI in het onderwijs is nog in zijn kinderschoenen. In de toekomstig onderzoek kan er worden gekeken of er dan wel beleid wordt gevoerd op dit onderwerp en welke *best-practices* kunnen worden geïdentificeerd.
- AI die vandaag al beschikbaar is kan de leraar zodanig kan ondersteunen dat deze meer tijd per leerling kan besteden, of efficiënter te werk kan gaan. Hoe een AI de werkdruk kan verlagen is een richting voor vervolgonderzoek.

6 Referenties

- [1] Dialogic (2014). *Big data in het onderwijs en de wetenschap* [www.dialogic.nl] Utrecht: Dialogic.
- [2] Barrat, J. (2015). *Our Final Invention: Artificial Intelligence and the End of the Human Era* St. Martin's Griffin.
- [3] Piech, C., Bassen, J., Huang, J., Ganguli, S., Sahami, M., Guibas, L., en Sohl-Dickstein, J. (2015). *Deep Knowledge Tracing* Advances in Neural Information Processing Systems .
- [4] Cap Digital (2018). *EdTech Observatoire* [www.observatoire-edtech.com] Paris,
- [5] Chen, S. (2018). *China's schools are quietly using AI to mark students' essays ... but do the robots make the grade?* South China Morning Post.
- [6] Liang Jun, B. (2018). *Facial recognition used to analyze students' classroom behaviors* China: People's Daily Online.
- [7] Connor, N. (2018). *Chinese school uses facial recognition to monitor student attention in class* [www.telegraph.co.uk] Beijing : Telegraph.
- [8] Chase, C.C., Chin, D.B., Oppezzo, M.A., en Schwartz, D.L. (2009). *Teachable Agents and the Protégé Effect: Increasing the Effort Towards Learning* [aaalab.stanford.edu] vol. 18, pp. 334-352.
- [9] T. Bjerkedal, P.K. G. A. S. J. I. (2007). *Intelligence test scores and birth order among young Norwegian men (conscripts) analyzed within and between families*
- [10] Hood, D., Lemaignan, S., en Dillenbourg, P. (2015). *When Children Teach a Robot to Write: An Autonomous Teachable Humanoid Which Uses Simulated Handwriting* Portland, Oregon, USA: Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction .
- [11] Collaboration.ai (2018). *Education: shape the minds of the future by adding intelligence to the way students work and collaborate together* [www.collaboration.ai]
- [12] Lundberg, S.M., en Lee, S.I. (2017). *A unified approach to interpreting model predictions* Advances in Neural Information Processing Systems.
- [13] Google (2017). *Google AI Blog: Feature Visualization* [ai.googleblog.com]
- [14] Rechtbank Den Haag (2018). *ECLI:NL:RBDHA:2018:13949*
- [15] Bloom, B.S. (1984). *The two sigma problem: the search for methods of group instruction as effective as one-to-one learning* [web.mit.edu] vol. 13, pp. 4-16.
- [16] Müller, V.C., en Bostrom, N. (2014). *Future Progress in Artificial Intelligence: A Survey of Expert Opinion* Berlin: Springer.
- [17] Grace, K., Salvatier, J., Dafoe, A., Baobao, Z., en Evans, O. (2018). *When Will AI Exceed Human Performance? Evidence from AI Experts* Journal of Artificial Intelligence Research.
- [18] P. Huisman, Red., (2017). *Basisboek onderwijsrecht. Een inleiding op de onderwijswet- en regelgeving in primair en voortgezet onderwijs* 2 red., Den Haag: Sdu Uitgevers.
- [19] Europese Commissie (2018). *The European Artificial Intelligence landscape: workshop report* [ec.europa.eu] Brussel: Europese Commissie.
- [20] Steering Group of the Artificial Intelligence Programme (2017). *Finland's Age of Artificial Intelligence* [urn.fi] Helsinki: Ministry of Economic Affairs and Employment, Finland.
- [21] Holstein, K., McLaren, B., en Aleven, V. (2018). *Student Learning Benefits of a Mixed-Reality Teacher Awareness Tool in AI-Enhanced Classrooms* [link.springer.com] vol. 10947, Springer, Cham.
- [22] Lelei, D., en McCalla, G. (2018). *How to Use Simulation in the Design and Evaluation of Learning Environments with Self-directed Longer-Term Learners* [link.springer.com] vol. 10947, Cham: Springer.

- [23]Månsson, K., en Haake, M. (2018). *Teaching Without Learning: Is It OK With Weak AI?* [link.springer.com] vol. 10948, Cham: Springer.
- [24]HM Government (2017). *Industrial Strategy: Building a Britain fit for the future* [assets.publishing.service.gov.uk] London: Crown.
- [25]Government Office for Science (2015). *Artificial intelligence: opportunities and implications for the future of decision making* [assets.publishing.service.gov.uk] london: Crown.
- [26]Piech, C.B. J. H. J. G. S. S. M. G. L. J. & S. J. (2015). *Deep knowledge tracing*In Advances in Neural Information Processing Systems.
- [27]Brouwer, B. (2017). *De robot als componist* [www.ie-forum.nl]
- [28]Čerka, P., Grigienė, J., en Sirbikyte, G. (2015). *Liability for damages caused by artificial intelligence*Elsevier.
- [29]. *Copyrightable Authorship: What Can Be Registered*3 red.
- [30]Davies, C.R. (2011). *An evolutionary step in intellectual property rights – Artificial intelligence and intellectual property*Elsevier.
- [31]Eck, M.v. (2018). *Geautomatiseerde ketenbesluiten en rechtsbescherming: Een onderzoek naar de praktijk van geautomatiseerde ketenbesluiten over een financieel belang in relatie tot rechtsbescherming*Tilburg: Tilburg University.
- [32]Commission, E. (2008). *Artificial intelligence: Commission outlines a European approach to boost investment and set ethical guidelines* [europa.eu]
- [33]European Parliament, Committee on Legal Affairs (2017). *Report with recommendations to the Commission on Civil Law Rules on Robotics* [www.europarl.europa.eu]
- [34]Hartlief, T. (2018). *Van knappe koppen en hun uitvindingen* [www.njb.nl]
- [35]Kortmann, B. (2018). *Intelligente systemen en rechtssubjectiviteit*
- [36]Naves, J. (2018). *Data in de rechtspraktijk*
- [37]Pels Rijcken & Drooglever Fortuijn adovcaten en notarissen (2018). *Whitepaper: Juridische aspecten van AI & machine learning*
- [38]Prins, C. (2018). *Bestuursrecht en digitalisering*
- [39]Van Bergen, M., en Mevissen, C. (2018). *Kunst(matige intelligentie): auteursrecht voor robots?* [ictrecht.nl]
- [40]van den Hoven van Genderen, R. (2016). *Robot als rechtspersoon, juridische noodzaak of legal science fiction?*
- [41]van den Hoven van Genderen, R. (2016). *De slimme robot is geen apparaat, maar een rechtspersoon* [www.nrc.nl]
- [42]Van Schaick, A. (2018). *Robot als rechtssubject* vol. 2018,
- [43](1992). *Algemene wet bestuursrecht* [wetten.overheid.nl]
- [44](1912). *Auteurswet* [wetten.overheid.nl]
- [45](1981). *Wet op het primair onderwijs* [wetten.overheid.nl]
- [46](1963). *Wet op het voortgezet onderwijs* [wetten.overheid.nl]
- [47](2007). *United Nations Convention on the Use of Electronic Communications in International Contracts*New York,
- [48](2018). *Edtech Observatory: Discover French Edtech ecosystem* [www.observatoire-edtech.com]
- [49]Yudkhowsky, E. (2001). *Creating Friendly AI 1.0: The Analysis and Design of Benevolent Goal.*

Bijlage 1. Overzicht interviewrespondenten

Naam	Rol	Organisatie
Aimee van Wynsberghe	Assistant professor of ethics and technology, lid High-Level expert group on AI, mede-oprichter Dutch Alliance for Artificial Intelligence	TU Delft
Albert Jagt	Chief Technology Officer	The learning network
Bert Bredeweg	Associate professor Informatics Institute	Universiteit van Amsterdam
Carla Haelermans	Associate professor onderwijs economie	TIER, Universiteit Maastricht
Christien Bok	Teamhoofd Onderwijsdiensten	SURF
Dennis Reidsma	Assistant professor Human Media Interaction group	Universiteit Twente
Derk Reneman	Wethouder	Gemeente Haarlemmermeer
Frank Léoné	Associate professor Cognitive psychology, associate professor Donders center for cognition, associate professor Donders institute for Brain, Cognition and Behavior	Radboud Universiteit
Ger Vennegoor	Schoolleider	Carmel College
Ilja Cornelisz	Assistant professor, tevens oprichter Amsterdam Center for Learning Analytics (ACLA)	Vrije Universiteit Amsterdam
Jaap van den Berg	CIO Office	DUO
Jaeques Koeman	Directeur	Edia
Jelmer de Ronde	Deelnemer Jong Talent-programma	SURF
Marlon Duarte	Docent	Dendron College
Martijn Janssen	Docent	Dendron College
Matthijs Moed	Projectmanager	SURFsara
Michael van Wetering	Strategisch adviseur innovatie	Kennisnet
Paul Vermeulen	CEO	Q-Sense
Rudy Jonkers	Docent	Carmel College
Wietse van Bruggen	Adviseur	Kennisnet

Bijlage 2. Overzicht van relevante regelingen

Hieronder een korte beschrijving per regeling (met een onderverdeling tussen generieke en sectorale regelingen), waarbij steeds de juridisch relevantie wordt geduid, de personen (burgers, bedrijven, overheden) die erbij betrokken zijn en de aard van de regeling (dwingend recht of aanvullend recht (bij het laatste kan men zaken contractueel (aanvullen en afwijkend) regelen)).

Generieke regelingen

Regelingen	Juridische relevantie	Betrokkenen	Karakter regeling
Algemene wet bestuursrecht (Awb)	Ieder besluit van een instelling is aan de regels van de Awb gebonden.	Bestuursorganen aan de ene kant en personen aan de andere kant	Grotendeels dwingend recht
Wet openbaarheid van bestuur (Wob)	Of informatie openbaar is, wordt in grote mate bepaald door de Wob. Zodra data onder instellingen komen te berusten vallen deze in principe onder het regime van de Wob	Bestuurs- en overheidsorganen aan de ene kant en personen aan de andere kant	Grotendeels dwingend recht
Wet hergebruik overheidsinformatie	De Who geeft regels die overheid in acht moet nemen als het hergebruik van informatie toestaat. Informatie moet dan wel openbaar en vrij van rechten van derden zijn.	Overheidsorganisaties aan de ene kant en personen aan de andere kant	Grotendeels dwingend recht
Algemene Verordening Gegevensbescherming (AVG)	Als de data waarop AI wordt toegepast persoonsgegevens bevat dan wel oplevert, is de handeling een verwerking in de zin van de AVG, met alle rechtsgevolgen van dien (doelbinding, informatieplichten etc.). Mogelijk zullen sommige data 'bijzondere persoonsgegevens' zijn, waarop een zwaarder verwerkingsregime rust	Natuurlijke personen en verantwoordelijken (= zij die persoonsgegevens verwerken) en bewerkers (= zij die in opdracht van verantwoordelijken verwerken)	Grotendeels dwingend recht
Algemeen deel Burgerlijk wetboek (BW)	Regelt onder meer het aangaan van contracten op basis waarvan AI-infrastructuur wordt gemaakt, met alle rechtsgevolgen van dien (nakomingsverplichtingen, etc.). Regelt ook (gevolgen van) schending van verplichtingen, waaronder onrechtmatige toepassing van AI	Alle personen	Verbintenissenrecht, waaronder contractenrecht, biedt veel ruimte en aanknopingspunten voor aanvullende afspraken
Auteurswet (Aw) en Databankenwet (Dw)	Bij het scheppen van AI-infrastructuur en het verzamelen van data kunnen deze intellectuele eigendomsrechten ontstaan. Dit leidt ertoe dat openbaarmaking en veelvoudig in de regel toestemming van	Rechthebbenden hebben een 'absoluut recht' (in te roepen tegen een ieder).	Veel ruimte om aanvullende of afwijkende afspraken te maken

Regelinen	Juridische relevantie	Betrokkenen	Karakter regeling
	de rechthebbende(n) behoeft. Via licenties kunnen aanvullende afspraken gemaakt worden		
Mededingingsrecht, waaronder wet Markt & Overheid (in Mw) Europees Verdrag Hergebruikrichtlijnen	Activiteiten van de portaalhouder kunnen een marktactiviteit zijn, met alle gevolgen van dien (discriminatieverbod, verbod misbruik te maken van een machtspositie etc.). Hergebruikregels kleuren ruimte verder in (verbod exclusiviteit, discriminatie etc.)	Marktpartijen	Grotendeels dwingend recht
Aanbestedingsrecht	Het opzetten van AI-infrastructuur (door de overheid) kan onder de aanbestedingsregels vallen	Aanbestedende overheidsorganisatie	Grotendeels dwingend recht

Onderwijsregelingen

Regelinen	Juridische relevantie	Betrokkenen	Karakter regeling
Grondwet (Gw)	Geeft grondbeginselen aan van (onder meer) het onderwijsrecht	Iedere persoon en de overheid	Dwingend recht
Wet Primair Onderwijs (WPO)	Algemene sectorregeling die onder meer doelstellingen, inrichting, kwaliteit, zorgplan, positie personeel, ouders en leerlingen beschrijft	Diverse partijen betrokken bij primair onderwijs	Grotendeels dwingend recht
Wet Voortgezet Onderwijs (WVO)	Regelt verschillende onderdelen van het voortgezet onderwijs en diverse onderwerpen die ook in de WPO geregeld worden. Belangrijk verschil: WVO is een kaderwet, waardoor sommige onderwerpen in AMvBs geregeld zijn	Diverse partijen betrokken bij voortgezet onderwijs	Grotendeels dwingend recht
Wet Medezeggenschap (WMS)	Voorbeeld van een themawet, regelt medezeggenschapsverhoudingen	Diverse geledingen binnen instellingen	Grotendeels dwingend recht
Inspectiewet (WOT)	Regelt bevoegdheden van de onderwijsinspectie	Instellingen, inspectie	Grotendeels dwingend recht
Leerplichtwet (Lpw)	Regelt verplichtingen van ouders om hun kinderen naar school te sturen	Ouders, instelling, kinderen	Grotendeels dwingend recht
Diverse lagere regelingen	Besluit bekostiging WPO, Besluit vernieuwde kerndoelen WPO, en AMvBs zoals het Inrichtingsbesluit WVO en het Eindexamenbesluit	diversen	Afhankelijk van de regeling

Versiegeschiedenis

Versie	Wijzigingen
v1.0.115	(Definitieve versie rapport)
v1.0.116	Ontbrekende naam toegevoegd in overzicht interviewrespondenten

Contact:

Dialogic innovatie & interactie
Hooghiemstraplein 33-36
3514 AX Utrecht
Tel. +31 (0)30 215 05 80
www.dialogic.nl

