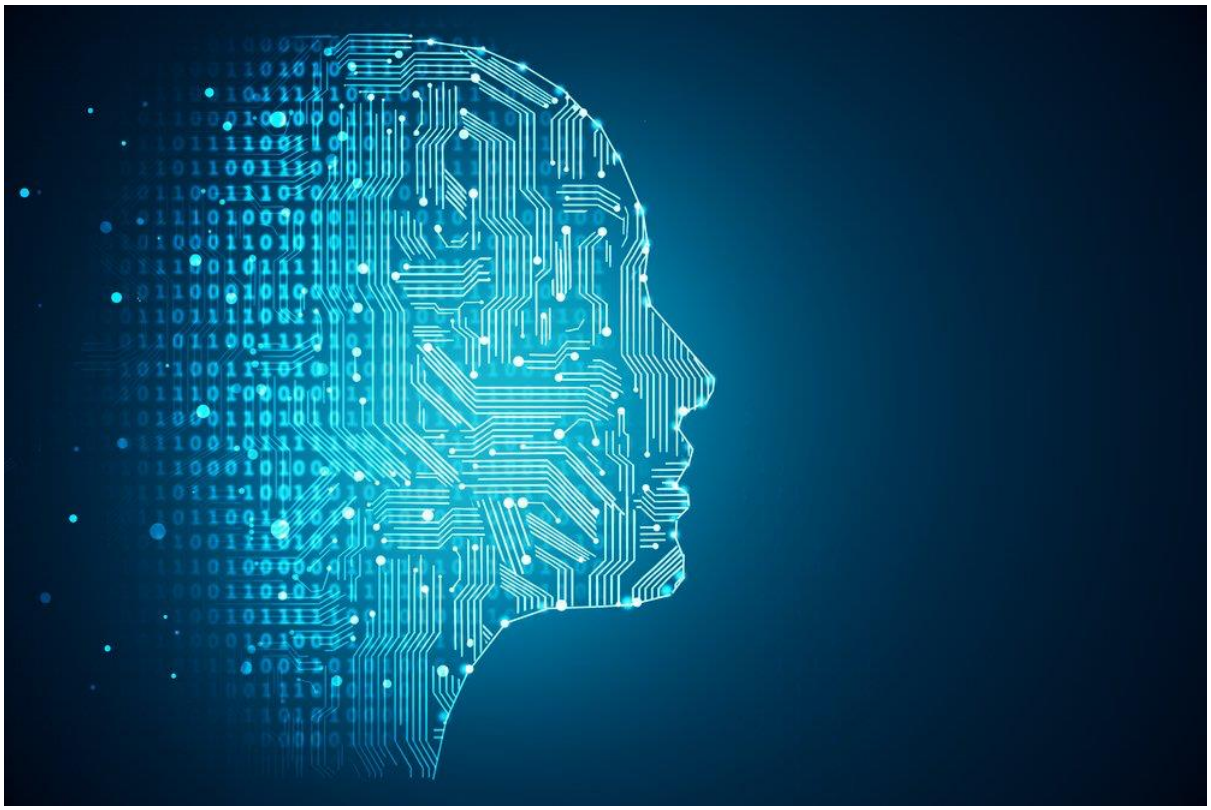
Kunstmatige Intelligentie: gevaarlijk of niet?



Lennert van Splunder
Geschreven in mei 2023

Voorwoord

Dit is een informatief rapport over Kunstmatige Intelligentie (KI), geschreven door een student aan de opleiding Informatica van de Hogeschool Rotterdam. Het rapport biedt waardevolle informatie voor personen die geïnteresseerd zijn in technologie en KI. De informatie in het rapport over KI en de daarmee samenhangende risico's zijn relevant aangezien KI al een aanzienlijk deel van ons dagelijks leven uitmaakt en naar verwachting in de toekomst een nog grotere rol zal spelen.

Samenvatting

De hoofdvraag van dit rapport is; wat zijn de risico's van kunstmatige intelligentie? Het doel is om een antwoord te geven op deze vraag met behulp van deelvragen. Met twee deelvragen wordt het duidelijk wat kunstmatige intelligentie is en wat de mogelijkheden ervan zijn om vervolgens voorbereid in te gaan op de risico's.

Kunstmatige intelligentie (KI) verwijst naar computersystemen die taken uitvoeren die menselijke intelligentie vereisen, zoals leren en probleemoplossing. Het maakt gebruik van technieken zoals machine learning, neurale netwerken en deep learning. KI kan worden onderverdeeld in zwakke KI, die specifieke taken uitvoert, en sterke KI, die menselijke cognitieve taken kan nabootsen. In de basis is KI een manier om computers slim te maken, want het is gebaseerd op de kennis die mensen aan computers toeschrijven.

KI wordt al veelvuldig toegepast in het dagelijks leven vanwege de vele mogelijkheden die het fenomeen biedt. Daar is nog niet iedereen zich van bewust en beschrijft dit rapport verschillende mogelijkheden en toepassingen van KI. Er worden niet alleen hedendaagse toepassingen beschouwd maar ook toekomstige. Zo wordt er bijvoorbeeld al KI in semi-autonome voertuigen gebruikt, zoals Tesla-auto's. Maar kan KI ook ingezet worden in volledig-autonome auto's en daar in de toekomst mogelijk het aantal verkeersongevallen mee verminderen. Verder worden andere toepassingen en mogelijkheden van KI in gebieden als wetenschap, de wapenwedloop, robotica en productie besproken op ongeveer dezelfde manier met voorbeelden.

Naast de vele mogelijkheden en toepassingen van KI, zijn er echter ook potentiële risico's en uitdagingen die gepaard gaan met de groeiende invloed ervan. Ondernemers, wetenschappers en experts hebben in een open brief gewaarschuwd voor de onbegrijpelijkheid, voorspelbaarheid en betrouwbare controle van KI. Gezichtsherkenningstechnologie kan namelijk gebruikt worden voor het ontgrendelen van je telefoon, maar ook om individuen te identificeren en te volgen. Arbeidsvervanging, autonome wapens, deepfakes en de opkomst van superintelligentie worden ook genoemd als risico's van KI.

Ten slotte is het belangrijk om zoveel mogelijk risico's van KI uit te sluiten door een zorgvuldige regulering en ethische normen te hanteren bij de ontwikkeling, implementatie en het gebruik van KI-systemen. Dit omvat het waarborgen van transparantie, verantwoordelijkheid en aansprakelijkheid van degenen die KI-systemen ontwikkelen en implementeren. Daarnaast is het belangrijk om bias en discriminatie in KI-algoritmen aan te pakken door middel van rigoureuze gegevensselectie en -verwerking, evenals onafhankelijke audits van KI-systemen.

Inhoud

Voorwoord 2
Samenvatting 3
Inleiding 5
Wat is kunstmatige intelligentie? 6
Wat zijn de mogelijkheden van KI? 7
Wat zijn de risico's van kunstmatige intelligentie? 11
Conclusie 15
Aanbeveling 17
Bronnen 18
Afbeeldingen 20

Inleiding

Kunstmatige intelligentie (KI) is een opkomende technologie die computers in staat stelt om taken uit te voeren die normaal gesproken menselijke intelligentie vereisen, zoals leren, redeneren en probleemoplossing. Met zijn enorme potentie om de samenleving te veranderen, is KI naar verluidt een onderwerp van grote opwinding. Echter, te midden van deze opwinding is het ook van belang om de risico's van KI te erkennen. Het doel van deze tekst is om de lezer bewust te maken van de mogelijke risico's die het met zich meebrengt.

Het rapport is als volgt opgebouwd. Het begint met de eerste deelvraag: wat is KI? Het antwoordt dient als introductie van KI en informeert lezers over de werking ervan. Hierna volgt de tweede deelvraag: wat is er mogelijk met KI? Hier worden toepassingen en mogelijkheden in diverse sectoren uitgelegd. Vervolgens is er genoeg informatie vergaard om de hoofdvraag uit gaan werken: "Wat zijn de risico's van KI?". Zo wordt men bewust van verschillende risico's die door deze vraag worden belicht. Tot slot worden er conclusies getrokken en aanbevelingen gedaan op basis van de inhoud uit dit rapport.

Wat is kunstmatige intelligentie?

Volgens een artikel van MIT Technology Review is kunstmatige intelligentie (KI) een verzamelnaam voor een aantal verschillende technieken en methoden, waaronder machine learning, neurale netwerken en deep learning. In de kern verwijst KI naar computersystemen die geprogrammeerd zijn om taken uit te voeren die normaal gesproken menselijke intelligentie vereisen, zoals leren, redeneren en probleemoplossing. Dit kan worden bereikt door regels en algoritmen te programmeren of door een computer zichzelf te laten trainen met behulp van machine learning-technieken. De KI van dit moment omvat taken zoals spraakherkenning, beeldherkenning, besluitvorming en probleemoplossing (Kiger, 2023).

KI, oftewel artificiële intelligence (AI), kan onderverdeeld worden in twee categorieën: sterke KI en zwakke KI. Zwakke kunstmatig intelligente systemen zijn ontworpen om slechts één specifieke taak uit te voeren, bijvoorbeeld een spraak assistent. Sterke KI daarentegen, is ontworpen om machines te creëren die dezelfde soort cognitieve taken kunnen uitvoeren als mensen. Alhoewel er aanzienlijk vooruitgang wordt geboekt, blijven deze systemen nog ver verwijderd van het bereiken van de flexibiliteit, creativiteit en aanpassingsvermogen van menselijke intelligentie (Koenders & Könning, z.d.). Daarom is sterke KI nog conceptueel, terwijl zwakke AI al functioneel is, dat heeft te maken met het verschil in de mate van intelligentie en autonomie die ze vertonen (IBM, z.d.).

Machine learning (ML) is één van de belangrijkste elementen van KI. Het geeft computers de mogelijkheid om van zichzelf te leren en zichzelf te verbeteren zonder expliciet daarvoor geprogrammeerd te zijn. Onderdelen van ML zijn neurale netwerken en deep learning technieken die gebaseerd zijn op biologische processen in het menselijk brein. Hierdoor kunnen computers patronen in data ontdekken en deze gebruiken om voorspellingen te doen en beslissingen te nemen. ML wordt toegepast in een breed scala aan domeinen, zoals fraude-detectie, gezondheidszorg en autonome voertuigen (SAP, z.d.).

In de basis is KI een manier om computers slim te maken, want het is gebaseerd op basis van de kennis die mensen aan computers toeschrijven. Een verhelderend voorbeeld hiervan, is een initiatief van KWF om KI toe te passen in de gezondheidszorg. Onderzoekers hebben 2100 mammogrammen (röntgenfoto's) verzameld van vrouwen waarbij in het verleden borstkanker is vastgesteld. Met behulp van deze data kunnen slimme computers leren om te bepalen of het gaat om een agressieve vorm van borstkanker en wat de groeisnelheid van de tumor is. Op deze manier kan er een betere behandeling ontwikkeld worden voor patiënten met borstkanker (KWF, 2023). Zo kan er dus naar aanleiding van menselijke kennis een toepassing van KI ontwikkeld worden.

Wat zijn de mogelijkheden van KI?

Er is al behoorlijk wat mogelijk met KI en het wordt daarom al frequent toegepast in mensen hun dagelijks leven, desalniettemin is nog niet iedereen zich daar even bewust van. Een pertinent voorbeeld hierbij is de bijdrage van KI bij semi-autonome voertuigen. Het wordt gebruikt om dit soort voertuigen aan te sturen, zoals dat bij Tesla-auto's ook het geval is. Het Tesla autopilot-systeem maakt gebruik van geavanceerde sensoren en neurale netwerken om auto's automatisch te besturen en obstakels te vermijden (Tesla, z.d.).

Alhoewel de huidige Tesla voertuigen momenteel ook nog door mensen bestuurd worden, betekent dat niet dat een slimme-auto niet zonder een bestuurder veilig aan het verkeer kan deelnemen. Sterker nog, een rapport van de National Highway Safety Administration (NHTSA) in de Verenigde Staten stelt dat autonome auto's, met behulp van KI, het aantal verkeersongevallen aanzienlijk kunnen verminderen. Het rapport toont aan dat 94% van de ongevallen veroorzaakt worden door menselijke fouten, zoals afleiding, vermoeidheid, enzovoort. Autonome auto's kunnen deze menselijke fouten elimineren en daardoor het aantal ongevallen verminderen. NHTSA suggereert dat mobiliteit veiliger, duurzamer en efficiënter wordt met KI. Niet alleen in het vervoer van personen, maar ook wat betreft de logistiek. Het kan onder andere onderhoud aan vrachtwagens en personenauto's voorspellen, wat ongelukken kan voorkomen en de kosten terug kan dringen.

Afbeelding I | Een Tesla-auto

Een andere mogelijkheid met KI wat ook te maken heeft met mobiliteit, vindt plaats in de Russische hoofdstad waar sinds vijftien oktober 2021 alle metrostations Face Pay gebruiken. Dit is het nieuwe betalingssysteem in de metro’s van Moskou, dat gebruik maakt van een biometrische sleutel en gezichtsherkenning. Het is behoorlijk eenvoudig in gebruik: je downloadt de app, koppelt een pasfoto en bankpas, en de tourniquets in de metrostations openen zich automatisch. Moskou is de eerste stad ter wereld waar een dergelijke service op zo'n grote schaal is uitgerold. Gemak dient de mens, lijken de autoriteiten willen uitdragen (Bosman, 2021).

Overigens is gezichtsherkenning ook erg populair in de telefonie-industrie. Hier wordt het meestal ingezet om het ontgrendelen van smartphones gemakkelijker te maken. Door AI-algoritmen te trainen met duizenden beelden van gezichten, kan een telefoon worden ontgrendeld door simpelweg naar het scherm te kijken. Na verloop van tijd zullen gebruikers wellicht merken dat het ontgrendelen steeds beter verloopt. Dit komt doordat de ML-technologie achter gezichtsherkenning steeds nauwkeuriger wordt dankzij de vergaarde data. Dit is een voorbeeld van hoe KI kan worden toegepast om technologie te verbeteren en het dagelijks leven gemakkelijker te maken (Marr, 2019).

Naast de toepassingen die het leven gemakkelijker maken, zijn er ook mogelijkheden die het menselijk leven prettiger kunnen maken. Bijvoorbeeld met de chatbot “Replika”, die is ontworpen om te fungeren als een virtuele vriend en gesprekspartner voor mensen met mentale gezondheidsproblemen, waaronder autisme. Deze chatbot is getraind met behulp van machine learning en natuurlijke taalverwerking om gesprekken te voeren die lijken op die van een menselijke vriend, en kan helpen bij het verminderen van eenzaamheid en het verbeteren van de stemming (Stella, 2022).

Een andere mogelijkheid van KI is om complexe wetenschappelijke problemen op te lossen en nieuwe ontdekkingen te doen die anders wellicht geen eens mogelijk waren. Een voorbeeld hiervan is de ontdekking van een nieuw antibioticum. In 2019 maakten wetenschappers van het Massachusetts Institute of Technology (MIT) gebruik van KI-algoritmen om miljoenen moleculaire combinaties te analyseren en de meest veelbelovende kandidaten voor nieuwe antibiotica te identificeren. Het resultaat was een nieuwe verbinding, genaamd Halicin, die zeer effectief bleek te zijn tegen een breed scala aan bacteriën, waaronder sommige die resistent zijn tegen bestaande antibiotica (Stokes et al., 2020).

Maar dat is nog niet alles, er bestaan ook kunstmatig intelligente huisdieren. Het bekendste voorbeeld is Aibo, wat in het Japans vriend of maatje betekent. De door Sony geproduceerde hond heeft als doel om mensen gezelschap te houden. Wanneer een huisdier intelligenter is, wordt het makkelijker voor een mens om er een emotionele band mee te hebben. Deze KI-hond kan redelijk vloeiend bewegen en kent allerlei trucjes, dansjes en kan ook emoties uitbeelden: verrassing, blijheid, ontevredenheid, boosheid, verdriet en angst. De programmeurs hebben hem ook instincten meegegeven: nieuwsgierigheid, honger (batterij opladen), slaap, liefde en drukte door extra veel beweging (Sony, n.d.).

Afbeelding II | Een robothond genaamd Aibo

Technologische ontwikkelingen worden niet alleen gebruikt in hardware en apparatuur die we in ons dagelijks leven gebruiken. Ze zijn in de loop der geschiedenis ook voortdurend toegepast in wapens en ander oorlogstuig. Hetzelfde geldt voor AI- technieken. In veel systemen wordt tot op zekere hoogte al gebruikgemaakt van AI, maar velen zien de ontwikkeling en invoering van volautomatische wapensystemen, ook wel bekend als ‘Lethal Autonomous Weapons Systems’ (LAWS), als de volgende belangrijke stap (Fernanda, 2022).

Afbeelding III | Close-in Weapons System

LAWS zijn systemen die, eenmaal geactiveerd, doelen kunnen traceren, identificeren en aanvallen zonder verder menselijk handelen. De algoritmen hebben grote hoeveelheden data nodig en de modellen moeten uitgebreid getraind worden om aan de wensen te voldoen. AI-technieken worden al geruime tijd ingebouwd in wapensystemen, en er zijn al enkele halfautomatische systemen in gebruik. Een voorbeeld is het ‘Close-in Weapons System’ (CIWS), veel gebruikt in de luchtafweer. Daarmee worden met behulp van radartechnieken bedreigingen geïdentificeerd en getraceerd. Zodra er een bedreiging wordt waargenomen, kan een computergestuurd afvuursysteem deze selecteren en zelfstandig aanvallen. Een bekend voorbeeld van een CIWS is de Nederlandse ‘GoalKeeper’. Dit type systeem wordt speciaal ingezet om een bepaald gebied te verdedigen en werkt alleen in eenvoudige, gestructureerde omgevingen. Hoewel ze ook zouden kunnen worden beschouwd als autonome wapens, zijn de inzetmogelijkheden en toepassingen ervan beperkt (Fernanda, 2022).

Kunstmatige intelligentie en robots kunnen niet uitblijven tussen de mogelijkheden van KI. Want deze combinatie heeft verschillende industrieën gerevolutioneerd, dankzij hun capaciteiten en potentieel om efficiëntie en productiviteit te verbeteren. In de wereld van de autoproductie hebben bedrijven zoals BMW kunstmatige intelligentie geïntegreerd in hun assemblagelijnen, wat helpt hen om hun hoge kwaliteitsnormen te handhaven. Hen AI-oplossing kan worden ingesteld: medewerkers maken foto's van een onderdeel vanuit verschillende hoeken en markeren mogelijke afwijkingen op de afbeeldingen. Op deze manier creëren ze een afbeeldingsdatabase om een zogenaamd neuraal netwerk op te bouwen, dat later de afbeeldingen kan evalueren zonder menselijke tussenkomst. Zodra het leerproces is voltooid, kan het neurale netwerk zelfstandig bepalen of een component voldoet aan de specificaties. Zo voert het veeleisende inspectietaken uit, elimineert het pseudo-defecten (afwijkingen van de norm), versnelt het logistieke processen en verlicht het medewerkers van repetitieve taken (Hemmerle, 2017).

Bovendien laat Sophia, de baanbrekende kunstmatig intelligente robot ontwikkeld door Hanson Robotics, opmerkelijke vooruitgang zien in humanoïde robotica, waarbij de grenzen worden verlegd van wat robots kunnen bereiken op het gebied van mensachtige interactie en cognitieve vaardigheden. Sophia vormt een voorbeeld van de huidige stand van zaken met betrekking tot Strong AI. De robot met het uiterlijk van een vrouw is in staat om gezichtsuitdrukkingen te herkennen en menselijke emoties te interpreteren. Sophia is ook geprogrammeerd om te leren en zichzelf te verbeteren, waardoor haar interacties met mensen steeds natuurlijker worden. Ze is bekend geworden vanwege haar publieke optredens en interviews, waarin ze complexe antwoorden geeft en een indrukwekkende mensachtige verschijning heeft. Het doel van Sophia is om te helpen bij de ontwikkeling van mens-robot interactie en te laten zien wat er mogelijk is met de KI-technologie (Hanson Robotics, 2020).

Afbeelding IV | Sophia de kunstmatig intelligente robot

Wat zijn de risico's van kunstmatige intelligentie?

Niemand – zelfs de maker niet – kan KI begrijpen, voorspellen of betrouwbaar controleren. Dat staat in een open brief die is ondertekend door bekende ondernemers, wetenschappers en vooraanstaande experts op het gebied van KI. Reeds duizend mensen ondertekende de brief, onder wie Tesla-topman Elon Musk, Apple-oprichter Steve Wozniak en historicus Yuval Harari. Ze slaan alarm bij de Amerikaanse overheid en bigtechbedrijven over de snelle ontwikkeling en opmars van kunstmatige intelligentie en de samenhangende gevaren ervan (Eshuis, 2023).

Eén van de risico's die verband houd met KI is de opkomst van deepfakes. Dit zijn nepvideo's en afbeeldingen die zijn gemaakt met behulp van KI en geavanceerde ML-algoritmen, waardoor beelden gemanipuleerd kunnen worden om het te laten lijken alsof een persoon of object iets doet wat in werkelijkheid niet het geval is. De negatieve toepassingen ervan zoals pornografie, fraude en misleiding, haat zaaien, het verspreiden van misinformatie en het beïnvloeden van democratische verkiezingen, kunnen naast kwalijke gevolgen voor direct betrokkenen, ook maatschappelijke gevolgen hebben. Daarbij valt te denken aan een afnemend vertrouwen in media, de democratie en de rechtspraak als van veel digitale content niet meer duidelijk is of het authentiek is (Ministerie van Justitie en Veiligheid, 2022).

Een voorbeeld deepfake-video is er één van Barack Obama waar "hij" een toespraak houdt. KI-software laat Obama van alles zeggen wat hij in werkelijkheid niet heeft gezegd, met het doel om het publiek bewust te maken van de opkomst van deepfake-videos (Silverman, 2018). De stem in deze video is identiek aan die van Obama. Hier maakt men tegenwoordig ook in de muziekwereld gebruik van. Er wordt steeds vaker muziek met KI gegenereerd op basis van audio van andere artiesten hun stem. Het gevaar hiervan heeft te maken met auteursrechten en intellectueel eigendom. Diverse platenmaatschappijen en juristen zijn dan ook actief op zoek naar manieren om dit verschijnsel tegen te gaan (Van Riel, 2023).

Afbeelding V | Barack Obama deepfake

Buiten muziek en deepfakes brengt het klonen van stemmen nog een ander risico met zich mee. Criminelen maken er namelijk gebruik van om mensen op te lichten door aan de telefoon met de software de stem van een familielid van het slachtoffer te imiteren. Ze zijn vooral gericht op senioren en laten het dan lijken alsof er een kleinkind aan de lijn is die urgent geld nodig heeft. Via online geplaatste filmpjes komen de oplichters aan stemmen van derden. De KI-software doet vervolgens de rest wat ervoor zorgt dat er “gesproken” kan worden met een nagemaakte stem. Deze vorm van fraude komt nu zo frequent voor dat de Amerikaanse telecomtoezichthouder FTC er zelfs een waarschuwing over heeft gepubliceerd (RTL Nieuws, 2023).

Desondanks de exploitatie van verschillende KI-toepassingen zijn er wel degelijk situaties waarin het positieve effecten oplevert in plaats van negatieve. Dit komt met name vaak voor bij bedrijven (Haan, 2023), echter zit zelfs hier een risico aan verbonden. Want door de toenemende inzet van KI bij bedrijven, heeft de maatschappij steeds meer te maken met robotisering en automatisering. Dat gaat onvermijdelijk verschillende banen kosten. Volgens het World Economic Forum (WEF) wordt in 2025 de helft van mensen hun werk gedaan door machines. Wereldwijd zal dit volgens hen 85 miljoen arbeidsplaatsen gaan kosten. Maar de nieuwe technologieën zal ook voor nieuwe werkgelegenheid zorgen. Volgens hen zal het voor 97 miljoen nieuwe banen zorgen, wat onderaan de streep dus duidt op meer werkgelegenheid. Maar dan is het de vraag of men die er een baan door verliest de juiste vaardigheden bezit voor de toekomstige arbeidsmarkt. Wellicht gaat het leiden tot ongelijkheid en onzekerheid voor werknemers die hun baan verliezen door automatisering en moeite hebben om zich aan te passen aan de veranderende arbeidsmarkt (WEF, 2020).

Afbeelding VI | Robotisering van beroepen

Naast de impact van KI op banen en werkgelegenheid, heeft kunstmatige intelligentie ook zijn intrede gedaan in de wapenwedloop, waar ook risico's aan verbonden zijn. Een voorbeeld van een controversieel onderwerp binnen dit domein zijn autonome wapens die specifiek ontwikkeld worden voor oorlogsvoering. Het gebruik van het woord 'autonoom' in deze context is niet ideaal, want de KI-systemen hebben geen moreel besef. Ze zijn niet in staat om mensen als morele individuen te begrijpen en te erkennen. Wanneer dit soort autonome wapens dan ingezet worden tijdens een conflict, ontbreekt het hen aan het vermogen om de waarde van het leven, inclusief dat van de 'doelwitten', te overdenken en te begrijpen. Mensen worden gereduceerd tot eenvoudige gegevens, als potentiële doelen die moeten worden geëlimineerd.

Slagvelden zijn complexe omgevingen die een hoog niveau van morele en ethische afwegingen vereisen. Mede daarom zijn er verschillende non-gouvernementele organisaties die het belang benadrukken van regulering voor autonome wapens, waarbij sommigen een volledig verbod als de meest geschikte aanpak beschouwen met het oog op de risico's. De wapens zijn namelijk afhankelijk van complexe algoritmes en sensoren om beslissingen te nemen, waardoor er altijd een risico op onvoorziene omstandigheden bestaat, wat kan leiden tot onbedoelde slachtoffers. Bovendien bemoeilijkt het autonome aspect de individuele of collectieve verantwoordelijkheid voor de acties van deze wapens (Fernanda, 2022).

Het Ministerie van Defensie heeft in Januari '23 aangekondigd dat de DoD Directive 3000.09, Autonomie in Wapensystemen, is bijgewerkt. De richtlijn is opgesteld om de waarschijnlijkheid en gevolgen van fouten in autonome en semi-autonome wapensystemen te minimaliseren (U.S. Department of Defense, 2023).

Dezelfde gezichtsherkenningstechnologie (GHT) die wordt gebruikt in autonome wapens om oppositie en steun te herkennen, wordt op grote schaal toegepast. Van het ontgrendelen van een iPhone tot het automatisch taggen van foto's op Facebook, tot werkgevers die productiviteit volgen en politiemachten die protesten observeren, wordt GHT steeds meer geïntegreerd in het dagelijks leven. GHT vergelijkt vastgelegde beelden met andere gezichtsbeelden die bijvoorbeeld in databases of op overheidslijsten staan. Het is een uiterst indringende vorm van surveillance en kan de individuele vrijheden ernstig ondermijnen en uiteindelijk de samenleving. Zo kan het bijvoorbeeld misbruikt worden kwaadwillende partijen om individuen onbevoegd te monitoren, te profileren of zelfs te discrimineren op basis van hun uiterlijk, ras, religie, seksuele oriëntatie of andere persoonlijke kenmerken (Ahmed, 2021).

Het agressieve ontwikkelings- en gebruikspatroon van gezichtsherkenning in China geeft inzicht in hoe de technologie die zowel onschadelijk als nuttig kan zijn – denk aan GHT op telefoons – ook kan worden misbruikt om een specifieke aanpak mogelijk te maken. Want Chinese functionarissen maken gebruik van surveillancetools om mensen die "onbeschaafd gedrag" vertonen publiekelijk te

vernederen, zo stellen surveillance-experts vast. De dreiging van publieke vernedering stelt ze in staat om meer dan een miljard mensen licht te sturen naar wat zij beschouwen als acceptabel gedrag, van wat je draagt tot hoe je de straat oversteekt (Ng, 2020).

Naast de bredere gevaren van kunstmatige intelligentie rijst er nog één specifiek risico: de opkomst van superintelligentie. Dit concept, beschreven door verschillende onderzoekers, brengt potentiële existentiële bedreigingen met zich mee en vereist volgens hen zorgvuldige aandacht en voorbereiding. Nick Bostrom; een Zweedse filosoof van de Universiteit van Oxford; heeft er een boek over geschreven genaamd: “Superintelligence: Paths, Dangers, Strategies". Het betoogt dat als machinale intelligentie het niveau van algemene intelligentie van menselijke breinen overstijgt, deze nieuwe superintelligentie mensen zou kunnen vervangen als de dominante levensvorm op aarde. Machines die voldoende intelligent zijn, zouden in staat zijn om hun eigen capaciteiten sneller te verbeteren dan menselijke computerwetenschappers, wat zou kunnen resulteren in een existentiële catastrofe voor de mensheid (Wikipedia-bijdragers & Wikipedia, 2023a).

Een andere uitdaging die wordt veroorzaakt door KI-technologie is het potentieel voor bias en discriminatie. De systemen zijn alleen zo onbevooroordeeld als de data waarop ze getraind worden; als die data bevooroordeeld is, zal het resulterende systeem dat ook zijn. Dit kan leiden tot discriminerende beslissingen die individuen beïnvloeden op basis van factoren zoals ras, geslacht of sociaaleconomische status (European Union Agency for Fundamental Rights, 2022).

Om zoveel mogelijk risico’s van KI te minimaliseren, is het belangrijk om een zorgvuldige regulering en ethische normen te hanteren bij de ontwikkeling, implementatie en het gebruik van KI-systemen. Dit omvat het waarborgen van transparantie, verantwoordelijkheid en aansprakelijkheid van degenen die KI-systemen ontwikkelen en implementeren. Daarnaast is het belangrijk om bias en discriminatie in KI-algoritmen aan te pakken door middel van rigoureuze gegevensselectie en -verwerking, evenals onafhankelijke audits van KI-systemen (Europese commissie, 2022).

Ten slotte is het ook van belang om te investeren in omscholing en bijscholing van werknemers, zodat zij de nodige vaardigheden kunnen verwerven om zich aan te passen aan de veranderende arbeidsmarkt. Dit kan helpen om ongelijkheid en onzekerheid te verminderen in het tijdperk van kunstmatige intelligentie.

Conclusie

Wat zijn de risico's van kunstmatige intelligentie? Om deze vraag te beantwoorden, is het belangrijk om te begrijpen wat KI eigenlijk inhoudt. Het verwijst naar computersystemen die taken uitvoeren die menselijke intelligentie vereisen, zoals leren en probleemoplossing. KI is in essentie een manier om computers intelligent te maken, gebaseerd op de kennis die aan computers wordt toegeschreven (Kiger, 2023).

KI wordt al veelvuldig toegepast in het dagelijks leven vanwege de talloze mogelijkheden die het fenomeen biedt. Denk hierbij aan mobiliteit, wetenschap, wapenwedloop, robotica, productie en meer. Een voorbeeld hiervan is de opkomst van autonome wapens in de wapenwedloop, speciaal ontwikkeld voor oorlogsvoering. Deze wapens zijn kunstmatig intelligente systemen die, eenmaal geactiveerd, doelen kunnen traceren, identificeren en aanvallen zonder verder menselijk handelen.

De systemen hebben echter geen moreel besef en zijn niet in staat om mensen als morele individuen te begrijpen en te erkennen. Mensen worden gereduceerd tot eenvoudige gegevens, potentiële doelwitten die geëlimineerd moeten worden. Bij het gebruik van de autonome wapens bestaat er altijd een risico op onvoorziene omstandigheden, wat bijvoorbeeld kan leiden tot onbedoelde slachtoffers. Bovendien bemoeilijkt het autonome aspect de individuele of collectieve verantwoordelijkheid voor de acties van deze wapens (Fernanda, 2022).

De risico's die te maken hebben met verantwoordelijkheid en onvoorziene omstandigheden treden vaker op bij toepassingen van KI. Om terug te komen op het vorige voorbeeld is er om de waarschijnlijkheid en gevolgen van fouten in autonome en semi-autonome wapensystemen te minimaliseren, een richtlijn voor autonomie in wapensystemen bijgewerkt (U.S. Department of Defense, 2023). Dit is volgens de Europese Commissie een valide aanpak, zij suggereren dat het naleven van zorgvuldige regelgeving en ethische normen bij de ontwikkeling, implementatie en het gebruik van KI-systemen helpen bij het verminderen van de risico's van KI.

Daarnaast zijn er specifieke risico's verbonden aan kunstmatige intelligentie die onze aandacht verdienen. Eén belangrijk aspect is de impact van KI op werkgelegenheid. De toenemende automatisering en de groeiende capaciteiten van KI-systemen kunnen leiden tot veranderingen in de arbeidsmarkt en mogelijk tot het verlies van bepaalde banen. Dit vraagt om doordachte strategieën en beleidsmaatregelen om de negatieve gevolgen voor werknemers op te vangen en hen te ondersteunen bij omscholing en het vinden van nieuwe mogelijkheden in een veranderende werkomgeving (Europese commissie, 2022).

Een ander relevant risico betreft gezichtsherkenningstechnologie. Hoewel deze technologie veelbelovend is en diverse toepassingen heeft, brengt het ook serieuze zorgen met zich mee op het gebied van privacy en individuele vrijheden. Het potentieel misbruik van gezichtsherkenningssystemen door overheden, bedrijven of kwaadwillende actoren kan leiden tot grootschalige surveillance, inbreuk op de privacy en mogelijke discriminatie (Ahmed, 2021).

Al met al kan men stellen dat er risico's met grote impact aan KI verbonden zitten, denk bijvoorbeeld aan superintelligentie. Dus, om zoveel mogelijk risico's te minimaliseren, is het belangrijk om een zorgvuldige regulering en ethische normen te hanteren bij de ontwikkeling, implementatie en het gebruik van KI-systemen. Dit omvat het waarborgen van transparantie, verantwoordelijkheid en aansprakelijkheid van degenen die KI-systemen ontwikkelen en implementeren (Europese commissie, 2022).

Aanbeveling

Ik raad aan om in het verlengde van dit rapport de ethische kant van KI te onderzoeken. Naar mijn mening is dat een interessant en belangrijk aspect ten opzichte van KI en zijn ontwikkelingen. Echter paste het niet in dit informatieve rapport, omdat het belichten van de ethiek waarschijnlijk zal resulteren in subjectieve tekst.

Een mogelijke onderzoeksvraag, die het waard is om te onderzoeken, luidt als volgt: "Is het ethisch verantwoord om KI te gebruiken?" Deze vraag kan leiden tot diepgaande vraagstukken over de mogelijke gevolgen van KI op gebieden zoals privacy, autonomie, rechtvaardigheid en menselijke waarden. Het is van belang om ethische kwesties te belichten en te onderzoeken, omdat ze een directe invloed hebben op de ontwikkeling, implementatie en het gebruik van KI-systemen (Wikipedia-bijdragers & Wikipedia, 2023b).

In dit onderzoek zou je kunnen overwegen om een brede reeks ethische vraagstukken te onderzoeken die verband houden met KI. Enkele voorbeelden van dergelijke vraagstukken zijn:

- Privacy: Hoe beïnvloedt KI de privacy van individuen? Welke maatregelen kunnen worden genomen om de privacy te waarborgen in een wereld waarin KI-technologieën wijdverspreid zijn?
- Autonomie: In hoeverre kunnen KI-systemen de autonomie van individuen beperken of juist bevorderen? Hoe kunnen we ervoor zorgen dat mensen controle en beslissingsbevoegdheid behouden in een omgeving waarin KI een steeds grotere rol speelt?
- Rechtvaardigheid: Welke impact heeft KI op bestaande ongelijkheden in de samenleving? Hoe kunnen we ervoor zorgen dat KI-systemen eerlijk en rechtvaardig worden ontwikkeld en toegepast?
- Menselijke waarden: In hoeverre kunnen KI-systemen menselijke waarden zoals empathie, compassie en ethiek begrijpen en respecteren? Hoe kunnen we ervoor zorgen dat KI niet in conflict komt met fundamentele menselijke waarden?

Door het onderzoeken van dit soort ethische vraagstukken kan een dieper inzicht worden verkregen in de impact en implicaties van KI. Het zal helpen bij het beantwoorden van de hoofdvraag: "Is het ethisch verantwoord om KI te gebruiken?"

Bronnen

Ahmed, H. S. A. (2021, December 21). Facial Recognition Technology and Privacy Concerns. ISACA. Retrieved June 10, 2023, from https://www.isaca.org/resources/news-and-trends/newsletters/atisaca/2022/volume-51/facial-recognition-technology-and-privacy-concerns#:~:text=Data%20breaches%20involving%20facial%20recognition,faces%20cannot%20easily%20be%20changed.

Apple. (2023, May 16). Apple introduces new features for cognitive accessibility, along with Live Speech, Personal Voice, and Point and Speak in Magnifier. Apple Newsroom. Retrieved May 23, 2023, from https://www.apple.com/newsroom/2023/05/apple-previews-live-speech-personal-voice-and-more-new-accessibility-features/

Bosman, J. (2021, October 17). Moskovieten huiverig voor gezichtsherkenning in de metro. FD.nl. Retrieved May 18, 2023, from https://fd.nl/tech-en-innovatie/1416337/moskovieten-en-mensenrechtenorganisaties-huiverig-voor-gezichtsherkenning-in-metro

European Union Agency for Fundamental Rights. (2022). Bias in Algorithms – Artificial Intelligence and Discrimination. In FRA (ISBN 978-92-9461-935-8). Retrieved May 29, 2023, from https://fra.europa.eu/sites/default/files/fra_uploads/fra-2022-bias-in-algorithms_en.pdf

Europese commissie. (2022). Regulatory framework for artificial intelligence. In European Commission. Retrieved June 10, 2023, from https://commission.europa.eu/system/files/2022-09/1_1_197605_prop_dir_ai_en.pdf

Fernanda, T. (2022, October 4). The use of artificial intelligence in war. University of Groningen. Retrieved May 18, 2023, from https://www.rug.nl/news/2022/10/the-use-of-artificial-intelligence-in-war

Haan, K. (2023, April 24). How Businesses Are Using Artificial Intelligence In 2023. Forbes Advisor. Retrieved June 10, 2023, from https://www.forbes.com/advisor/business/software/ai-in-business/

Hemmerle, A. (2017, July 15). Fast, efficient, reliable: Artificial intelligence in BMW Group Production. BMW Group PressClub. Retrieved May 18, 2023, from https://www.press.bmwgroup.com/global/article/detail/T0298650EN/fast-efficient-reliable:-artificial-intelligence-in-bmw-group-production?language=en

Ng, A. (2020, August 11). How China uses facial recognition to control human behavior. CNET. Retrieved April 25, 2023, from https://www.cnet.com/news/politics/in-china-facial-recognition-public-shaming-and-control-go-hand-in-hand/

RTL Nieuws. (2023, March 27). Waakhond: pas op voor oplichters die stemmen van familieleden namaken. Retrieved May 23, 2023, from https://www.rtlnieuws.nl/tech/artikel/5374282/oplichting-nagemaakte-stem-ai-voice-ouderen-oplichter-bellen

Silverman, C. (2018, April 17). How To Spot A Deepfake Like The Barack Obama–Jordan Peele Video. BuzzFeed. Retrieved April 25, 2023, from https://www.buzzfeed.com/craigsilverman/obama-jordan-peele-deepfake-video-debunk-buzzfeed

Sony. (n.d.). *aibo*. Aibo. Retrieved June 14, 2023, from https://us.aibo.com/

U.S. Department of Defense. (2023, January 25). DoD Announces Update to DoD Directive 3000.09, 'Autonomy In Weapon Sys. Retrieved May 26, 2023, from https://www.defense.gov/News/Releases/Release/Article/3278076/dod-announces-update-to-dod-directive-300009-autonomy-in-weapon-systems/

Van Riel, R. (2023, April 12). Kunstmatige intelligentie maakt muziek, maar niet iedereen wordt daar vrolijk van. FD. Retrieved April 14, 2023, from https://fd.nl/tech-en-innovatie/1473437/ai-muziek-tegenhouden-wordt-een-moeilijke-klus

Wikipedia contributors. (2023). Superintelligence: Paths, Dangers, Strategies. Wikipedia. Retrieved May 26, 2023, from https://en.wikipedia.org/wiki/Superintelligence:_Paths,_Dangers,_Strategies

Wikipedia contributors. (2023b). Ethics of artificial intelligence. Wikipedia. Retrieved June 01, 2023, from https://en.wikipedia.org/wiki/Ethics_of_artificial_intelligence

Afbeeldingen

I – Tesla-interieur, een elektrische auto gecombineerd met KI:

- Artikel: "Self-Driving Tesla Was Involved in Fatal Crash, U.S. Says" van The New York Times
- Link: https://www.nytimes.com/2016/07/01/business/self-driving-tesla-fatal-crash-investigation.html

II – AIBO huisdier, een hond in combinatie met KI:

- Artikel: "Aibo" van Sony
- Link: https://www.sony.com/en/SonyInfo/sony_ai/aibo.html

III – Goalkeeper luchtafweer, een computergestuurd afvuursysteem:

- Artikel: “Goalkeeper CIWS” van Wikipedia
- Link: https://en.wikipedia.org/wiki/Goalkeeper_CIWS

IV – Sophia, een slimme robot met mensachtige uitstraling:

- Artikel: "Ontwikkelaar Sophia-robot wil dit jaar massaproductie van robots starten" van NU nieuws
- Link: https://www.nu.nl/tech/6110207/ontwikkelaar-sophia-robot-wil-dit-jaar-massaproductie-van-robots-starten.html

V – Barack Obama, een deepfake-video met een Obama toespraak:

- Artikel: "Deep fake video van Obama" van Mediawijsheid.nl
- Link: https://www.mediawijsheid.nl/video/deep-fake-video-van-obama/

VI – Robotisering van beroepen, een cartoon achtige afbeelding met zowel een robot als een net ontslagen menselijke medewerker:

- Artikel: “Robotisering van beroepen: grote impact op de arbeidsmarkt” van Nationale beroepen gids:
- Link: https://www.nationaleberoepengids.nl/robotisering-van-beroepen-grote-impact-op-de-arbeidsmarkt