



Universiteit
Leiden
The Netherlands

AI en de boardroom: aandachtsgebieden voor een verantwoordelijke en succesvolle ontwikkeling en gebruik van AI

Wuisman, I.S.; Lückerath-Rovers, M.; Ees, H. van; Kaptein, M.

Citation

Wuisman, I. S. (2023). AI en de boardroom: aandachtsgebieden voor een verantwoordelijke en succesvolle ontwikkeling en gebruik van AI. In M. Lückerath-Rovers, H. van Ees, & M. Kaptein (Eds.), *Jaarboek Corporate Governance 2023-2024* (pp. 137-154). Deventer: Wolters Kluwer. Retrieved from <https://hdl.handle.net/1887/3728377>

Version: Publisher's Version

License: [Leiden University Non-exclusive license](#)

Downloaded from: <https://hdl.handle.net/1887/3728377>

Note: To cite this publication please use the final published version (if applicable).

10 AI en de boardroom: Aandachtsgebieden voor een verantwoordelijke en succesvolle ontwikkeling en gebruik van AI

Iris Wuisman

1. Inleiding

Het zal de meeste bestuurders niet zijn ontgaan dat de ontwikkelingen op het gebied van kunstmatige intelligentie ('AI') het afgelopen jaar zijn geëxplodeerd. De multifunctionele toepassingsmogelijkheden en toegenomen beschikbaarheid van generatieve AI ('GenAI')¹ hebben als katalysator gewerkt. Het zorgt ervoor dat het gebruik van AI potentieel in alle gelederen van een onderneming van toegevoegde waarde kan zijn. Zo kan GenAI marketingcontent creëren, interacties met klanten ondersteunen, de researchafdeling van informatie voorzien en programmeurs helpen bij het schrijven van software. Daarnaast kan AI op allerlei manieren geïntegreerd worden in producten en diensten die een onderneming op de markt brengt. AI moet echter niet ingezet worden puur vanwege de wens of drang iets met de technologie te doen omdat men het gevoel heeft de boot te missen zonder goed na te denken over de implicaties hiervan. Tegelijkertijd is het belangrijk ook de kop niet in het zand te steken en te denken dat het wel goed komt binnen de organisatie.

In dit hoofdstuk bespreek ik vanuit het perspectief van een aantal kernaandachtsgebieden wat bestuurders in overweging kunnen nemen in de (strategische) gesprekken over deze ontwikkelingen, als zij richting en support geven aan de implementatie van kunstmatige intelligentie of daarop toezicht houden. Deze aandachtsgebieden zijn onderdeel van wat ik het P+AI-model noem en betreffen: (1) Purpose & Principles, (2) Problems & Products/Projects, (3) Perspectives, (4) People, (5) Planet, (6) Processes & Procedures, (7) Partnerships, (8) Performance & Progress en (9) Protection. Vanwege de gelimiteerde beschikbare ruimte beperk ik mij tot enkele centrale topics. Hierbij gaat het zowel om de integratie van AI in producten of diensten die door de onderneming op de markt worden gebracht als het gebruik van kunstmatige intelligentie in interne bedrijfsprocessen onder andere voor automatisering. Ik besteed daarbij tevens aandacht aan relevante wetgeving(sinitiatieven) op het gebied van digitalisering en duurzaamheid zonder

¹ Generatieve AI is een vorm van kunstmatige intelligentie die in reactie op een door de gebruiker ingegeven opdracht nieuwe content creëert op basis van bestaande data. Dit soort AI opent nieuwe mogelijkheden voor allerlei soortige contentcreatie en werkwijzen.

te veel in te gaan op de details hiervan.² Deze regels bieden (in de toekomst) voor een groot deel de kaders en randvoorwaarden voor de ontwikkeling en inzet van kunstmatige intelligentie.

2. Purpose & Principles

De purpose van een onderneming omvat de reden waarom een onderneming bestaat, de bestaansgrond. Er is al veel geschreven over purpose en misschien zijn we inmiddels zelfs al wel een beetje purpose-moe. Door de vergaande impact van AI op mens en milieu wordt de relevantie van purpose versterkt en dient deze op de agenda van bestuurders (en wetgevers) te blijven staan. Enerzijds betreft dit de bredere maatschappelijke discussie over wat corporate purpose zou moeten inhouden (enge of brede opvatting) en de rol van wet- en regelgeving daaromtrent. Anderzijds betreft dit het interne gesprek binnen de onderneming over de toekomstbestendigheid: heeft de onderneming, gelet op de gewijzigde (markt)omstandigheden vanwege kunstmatige intelligentie, in de huidige vorm nog wel bestaansrecht in de toekomst? Welke verantwoordelijkheid staat de onderneming voor? Zal de corporate purpose van de onderneming moeten wijzigen door de digitale transitie? En in het verlengde daarvan, welk effect heeft de technologie en de adoptie daarvan op het business model? Ik bespreek deze beide kanten. Hierbij besteed ik tevens aandacht aan de ‘Principles’, de kernwaarden en beginselen van de onderneming.

2.1 *Interne purpose-gesprekken tegen de achtergrond van AI-ontwikkelingen*

De digitale transformatie verandert de manier waarop we leven, waarop we werken, hoe we met elkaar interacteren en communiceren. De ontwikkeling van GenAI geeft daaraan een nieuwe dynamiek. De opkomst van de Metaverse³ als onderdeel van de digitale evolutie creëert een realiteit waarin de digitale en de fysieke wereld steeds verder met elkaar verstrengeld raken. In deze wereld verandert wat van waarde is en daarmee veranderen onze waardesystemen. Veel van de bestaande interacties zijn gebaseerd op de waarde van kennis en de waarde van tijd. Door de nieuwe mogelijkheden van GenAI is dit aan het verschuiven omdat kennis breder en sneller toegankelijk is. De AI-applicaties worden zo ontwikkeld dat het aanvoelt alsof je met een mens interacteert en daarin wordt succesvolle vooruitgang geboekt. Dit maakt het gebruik van GenAI laagdrempelig en in bepaalde contexten een goed alternatief voor interactie met natuurlijke personen. Ondertussen is de verificatie

2 Over het thema Technologie & ESG heb ik een drieluik geschreven in het *Tijdschrift Ondernemingsrecht*. Hierin heb ik een uitgebreide analyse opgenomen van de relevante Europese wet- en regelgeving (initiatieven): I.S. Wuisman (2023), ‘Technologie (AI & blockchain) en ESG; de Europese benadering van de digitaliseringstransitie – Deel I’, *Ondernemingsrecht* 2023/54; I.S. Wuisman (2023), ‘Technologie (AI & blockchain) en ESG; de Europese benadering van de digitaliseringstransitie – Deel II’, *Ondernemingsrecht* 2023/64; en I.S. Wuisman (2023), ‘Technologie (AI & blockchain) en ESG; de Europese benadering van de digitaliseringstransitie – Deel III’, *Ondernemingsrecht* 2023/65.

3 De Metaverse is een online wereld waarin mensen elkaar kunnen ontmoeten en kunnen interacteren via avatars. Aan deze virtuele realiteit kunnen ook bedrijven deelnemen.

van informatie en kennis door de komst van AI gecompliceerder geworden door de toename van misinformatie en deepfakes. Het karakter van het internet is daardoor aan wijziging onderhevig alsook het vertrouwen daarin. Dat heeft ook effect op de interacties die (via het internet) met bedrijven plaatsvinden. Als gevolg van deze ontwikkelingen veranderen ook de klantenvoorkeuren en -behoeften. Aan de ene kant bestaat de behoefte aan personalisatie van klantproposities gebaseerd op de persoonlijke gegevens, aan de andere kant is er ook de behoefte aan het behouden van privacy.

Door deze ontwikkelingen verschuiven de grenzen van het speelveld van ondernemingen en zijn nieuwe spelers in zicht. Het vergt van ondernemingen continue aandacht voor de vraag of de bestaansgrond van de onderneming gelet op deze ontwikkelingen wel toekomstbestendig is. Waar raken deze wijzigingen de onderneming? Heeft het bestuur eigenlijk wel goed zicht op wat de onderneming op het gebied van AI doet? Dient de onderneming te veranderen omdat de producten en/of diensten irrelevant worden of omdat door de veranderende omstandigheden het ‘probleem’ waarvoor het bedrijf de oplossing biedt van karakter wijzigt (zie hiervoor paragraaf 3)? Op welke wijze wil de onderneming aan deze steeds digitaler wordende wereld deelnemen? Wat vindt de onderneming belangrijk daarin? Dit betreft dus niet alleen het gesprek over welke kansen en mogelijkheden de technologische innovaties bieden en wat de meest profijtelijke businesscase is die daarbij hoort. Het gaat ook over wat die businesscase over de brede linie betekent. Wat zijn de implicaties daarvan in brede zin en hoe verhoudt de onderneming zich daartoe? In aanvulling op de begrenzing van wet- en regelgeving, zal de onderneming ook zelf kaders hebben voor de bedrijfsuitoefening. Deze kaders zijn ingebed in de gekozen purpose. Wat betekenen de mogelijke strategische richtingen als gevolg van de technologische ontwikkelingen tegen de achtergrond van deze kaders? Het is van belang deze kaders in het licht van de veranderende wereld kritisch te blijven onderzoeken, toe te passen als ze nog steeds ondersteund worden of bij te stellen als dat beter bij de onderneming en de veranderende context past. Met andere woorden: het zijn dynamische kaders. Hierbij is ook van belang wat de maatschappelijke opvatting daarover is. Dit hangt nauw samen met de maatschappelijke discussie over purpose.

2.2 *Bredere maatschappelijke discussie over corporate purpose*

De purpose van een onderneming omvat zoals gezegd de reden waarom een onderneming bestaat. Sec genomen kan dit dus van alles betreffen als het maar niet strijdig is met de wet. De laatste jaren is corporate purpose steeds meer benaderd vanuit het perspectief dat ondernemen dient plaats te vinden vanuit een bepaald besef van verantwoordelijkheid. De kaders van de purpose van een onderneming hangen af van het regelgevend kader waarbinnen de onderneming opereert en de keuzes die de onderneming zelf maakt. Het toepasselijke (ondernemings)recht is bepalend voor de zorgvuldigheidsnorm en het richtsnoer dat bestuurders in acht moeten nemen bij besluitvorming en derhalve in hoeverre er rekening gehouden dient te worden met de belangen van alle stakeholders. Het ondernemingsrecht van sommige juris-

dicties, zoals het Nederlandse ondernemingsrecht, verlangt van ondernemingen dat zij bij het uitvoeren van hun activiteiten de belangen van de bij die onderneming betrokkenen niet onevenredig schaden. Of dit samenvalt met het nemen van verantwoordelijkheid valt te bediscussiëren. Dat er in Nederland stappen worden gemaakt, ten aanzien van de maatschappelijke verantwoordelijkheid van ondernemingen in aanvulling op het richtsnoer voor bestuurders voortvloeiend uit het Cancun-arrest, komt duidelijk naar voren in de Shell-zaak.⁴ Dit kan ondersteunend zijn aan een verantwoordelijke AI-ontwikkeling en -gebruik door ondernemingen. Naast dit richtsnoer moeten bestuurders bij de besluitvorming voldoen aan de statutaire doelomschrijving en een mogelijk door het bestuur opgestelde purpose-verklaring. Deze purpose-verklaring kan, afhankelijk van wat de onderneming zelf kiest, in het licht van de hierboven beschreven bredere benadering uitgebreider zijn dan alleen de bestaansgrond doordat deze ook informatie kan bevatten over de (maatschappelijke) verantwoordelijkheid die de onderneming neemt of wil nemen en de waarden en beginselen die daarbij leidend zijn.⁵ Ik kom op die waarden en beginselen straks terug bij de bespreking van de ‘Principles’.

2.3 Impact van AI

Het belang van de (discussie over) corporate purpose in het licht van kunstmatige intelligentie, is gelegen in de vergaande impact die deze technologie (mogelijk) heeft. Het potentiële verdienmodel van kunstmatige intelligentie is enorm en kan veel aandeelhouderswaarde alsook maatschappelijke waarde creëren. Tegelijkertijd kan deze technologie een betekenisvolle negatieve impact teweegbrengen voor bijvoorbeeld eigen werknemers of consumenten en eindgebruikers. Het is dus van belang dat een onderneming goed doordrongen is van deze kant van de medaille. De effecten doen zich op individueel niveau voor, maar ook op groepsniveau en zelfs op systemisch niveau. Dit betekent dat op de langere termijn een verandering van het socio-economische systeem kan plaatsvinden als gevolg van de implementatie van deze technologie. Negatieve effecten van AI betreffen onder andere de onwenselijke beperking van de menselijke autonomie. Kunstmatige intelligentie kan een negatief effect hebben op de mogelijkheid zelf invulling te geven aan wat belangrijk is en waarden te creëren die richting geven aan eigen beslissingen en acties. Deze effecten kunnen dwang, manipulatie en misleiding betreffen, zoals het verstrekken van fake news en desinformatie, het limiteren van keuzemogelijkheden en het uitvoeren van bepaalde nudgingactiviteiten waardoor mensen een bepaalde kant op worden geduwd. De persoonlijke vrijheid komt dan in het gedrang. Het kan ook gaan om ongewenste monitoring, toegangsrestricties, discriminatie en stigmatisering. Het gebruik van AI kan banen van mensen overbodig maken en de

4 Rechtbank Den Haag 26 mei 2021, ECLI:NL:RBDHA:2021:5337. Het is afwachten wat in hoger beroep wordt geoordeeld, de hoorzittingen staan gepland voor april 2024. Zie voor het Cancun-arrest: HR 4 april 2014, ECLI:NL:HR:2014:804.

5 Zie bijdrage L. Timmerman (2021), ‘Purpose in de managementleer en het vennootschapsrecht’, p. 41-47, in: J.M. de Jongh e.a. (red.), *Bestuursbesluiten*, Uitgave vanwege het Instituut voor Ondernemingsrecht, nr. 123, Deventer: Wolters Kluwer 2021 en L. Timmerman (2021), ‘Corporate Governance Codes, Bevoegdheden, Taken, Gedragsregels en Purpose’, p. 1-10, in: M. Lückers-Rovers e.a. (red.), *Jaarboek Corporate Governance 2020-2021*, Deventer: Wolters Kluwer 2021.

fysieke veiligheid in gevaar brengen. Zo bij elkaar genomen klinkt dit best heftig en dat is het ook. Dat het heel fout kan gaan, heeft de toeslagenaffaire aangetoond. De effecten kunnen relatief gezien ook minder ingrijpend of onzichtbaar(der) zijn, maar nog steeds ongewenst. De inzet van AI kan bijvoorbeeld het werkplezier bij medewerkers wegnemen omdat interactie met collega's voortaan wordt vervangen door interactie met computers. Er kan bij werknemers en ook bij klanten een onderhuids onbehagen groeien als gevolg van een '*big brother is watching you*-gevoel'. Dit kan in de werkomgeving gepaard gaan met angst voor represailles in het kader van promotiemogelijkheden en salariëring. Het kan ook zorgen voor een gevoel van onrechtvaardigheid bij medewerkers omdat waarde en betekenis binnen een bedrijf niet alleen afhankelijk zijn van zaken die gemakkelijk te meten zijn. Wantrouwen jegens de werkgever kan vervolgens ontstaan en het welzijn van werknemers kan er significant door afnemen. Bij klanten kan dit gevoel van onrechtvaardigheid zich bijvoorbeeld voordoen bij dynamische prijsstellingen omdat men ten opzichte van anderen meer betaalt omdat de prijs is afgestemd op het persoonlijke digitale profiel.

Bovendien kunnen AI-applicaties er ook voor zorgen dat mensen steeds meer vertrouwen op AI-aanbevelingen bij het nemen van beslissingen. Dat kan vervolgens tot gevolg hebben dat het vertrouwen in de eigen capaciteiten, het kritisch reflecteren op zaken en het vermogen om te leren steeds meer afnemen. Hierdoor kunnen binnen ondernemingen belangrijke zaken over het hoofd worden gezien met alle gevolgen van dien. Het zorgt ook voor systemische risico's. Bijvoorbeeld in de situatie dat in een bepaalde sector over de hele linie bedrijfsbeslissingen genomen worden op basis van een handjevol dominante generatieve Amerikaanse AI-modellen (Large Language Models; 'LLM's') die een bias bevatten of misinformatie verschaffen. Naast de impact op sociale aspecten, heeft kunstmatige intelligentie ook impact op het klimaat & milieu ('E-impact'). Dit is voor velen nogal onzichtbaar. Van zoekopdrachten en vragen aan een generatieve tekst- of image AI-tool zoals ChatGPT, Bard of Stable Diffusion, tot aan automatisering van activiteiten die vroeger manueel werden uitgevoerd. Het kost allemaal energie en er wordt een groot beroep gedaan op (koel)water. Veel energie. Veel water. En dat levert problemen op tegen de achtergrond van het veranderende klimaat en de schaarste van water in sommige gebieden. En ook hier raakt AI de verantwoordelijkheid van bedrijven. Daarbij gaat het niet alleen om grote tech-bedrijven. Hoewel de ontwikkeling van de LLM's, vanwege de benodigde grote investeringen niet voor veel andere partijen dan de BigTech is weggelegd, kunnen de AI-tools die op de basis van de LLM's worden gebouwd ook een niet te onderschatten sociale- en milieu-impact hebben.

De mogelijke implicaties van de ontwikkeling en de inzet van AI zijn derhalve divers en meer of minder zichtbaar. Sommige effecten openbaren zich op de korte termijn, andere worden pas duidelijk na een langere tijd. Het ene effect zal vaker op de loer liggen dan het andere effect en dit alles is afhankelijk van de context waarin AI wordt ontwikkeld en gebruikt. Daarbij is het van belang te beseffen dat ondernemingen die vanuit de kern goed willen doen, door de wijze van uitvoering van AI-projecten toch de mist in kunnen gaan. Een voorbeeld hiervan zijn de negatieve

effecten als gevolg van bias. Bias in de gebruikte data is een universeel risico en kan funest zijn. Hierdoor kan het ook bij AI-ontwikkeling of -inzet door ondernemingen met de beste intenties, goed fout gaan. Een onderneming kan daarnaast ook het goede trachten te doen door een AI-systeem te ontwikkelen met een positieve E-impact. Maar tegelijkertijd kan dit systeem een negatieve sociale impact hebben. Omgekeerd kan ook. Wat dan te doen? Het antwoord is niet altijd zo voor de hand liggend en kan een lastig ethisch dilemma veroorzaken.

Het bovenstaande betekent een aantal dingen. GenAI en andere vormen van AI kunnen in een netwerkmaatschappij machtige tools zijn. Corporate purpose kan ervoor zorgen dat deze macht gelimiteerd wordt en dat er verantwoordelijk met deze macht wordt omgegaan. Dit is enerzijds een taak voor de wetgever en anderzijds nog steeds zeer afhankelijk van de insteek van de onderneming zelf. Tevens geeft de corporate purpose richting als wordt gezocht naar een oplossing voor ethische kwesties. De digitaliseringstransitie kan wat mij betreft niet los gezien worden van de duurzaamheidstransitie. De digitale transformatie binnen een onderneming zal verweven moeten zijn met de wijze waarop de onderneming omgaat met ESG-zaken. Duurzaamheid is de laatste tijd steeds meer verschoven van de flanken van ondernemingen naar de as. Dat zou ook moeten gelden voor digitalisering en dan het liefst gezamenlijk aangevlogen.

2.4 *Responsible AI*

Dit komt onder andere tot uiting in *Responsible AI* ('Verantwoorde AI'), een belangrijk thema dat in Europa volop in de aandacht staat. De voorgestelde Europese wetgeving, de AI-Act, bevat allerlei regels vanuit een risico-gebaseerde benadering op AI-systeemniveau om waarborgen te creëren tegen ongewenste effecten. De functie die een AI-systeem vervult alsook het specifieke doel en de modaliteiten waarvoor het systeem wordt gebruikt, bepalen welke regels van toepassing zijn. In de com-promistekst van het Europees Parlement zijn enkele beginselen opgenomen voor ethische en betrouwbare AI.⁶ Deze gelden voor alle AI-systemen en foundationmodellen maar worden per risicocategorie verder uitgewerkt. Deze principes omvatten beginselen die betrekking hebben op de volgende gebieden: (1) *human agency and oversight*, (2) *technical robustness and safety*, (3) *privacy and data governance*, (4) *transparency*, (5) *diversity, non-discrimination and fairness* en (6) *social and environmental well-being*. Het is een 'best-effort'-verplichting van aanbieders en gebruikers van AI-systemen om AI-systemen te ontwikkelen en te gebruiken in lijn met deze principes (ik kom terug op deze beginselen in paragraaf 2.6).

Hoewel de AI-verordening nog niet is aangenomen (verwachting is eind 2023) en dus nog niet helemaal duidelijk is wat de precieze regels zullen zijn, zullen hoog-risicosystemen moeten zijn omgeven met kwaliteits- en risicomanagementsystemen die voldoen aan bepaalde eisen (ik bespreek deze nader in paragraaf 7). De voor-ge-

6 https://www.europarl.europa.eu/doceo/document/TA-9-2023-0236_EN.html.

stelde AI-verordening bevat geen regels die zich specifiek richten op bestuurders. Veel AI-systemen worden niet gereguleerd door de AI-Act. De verordening laat het dan aan de ondernemingen zelf over invulling te geven aan de wijze waarop zij een verantwoordelijke ontwikkeling en gebruik van AI willen realiseren.⁷ Een deel hiervan wordt opgevangen door (voorgestelde) duurzaamheidswetgeving, zoals de Corporate Sustainability Reporting Directive ('CSRD') en de Corporate Sustainability Due Diligence Directive ('CSDDD'). De CSRD verlangt transparantie over de elementen van de strategie die gerelateerd zijn aan of impact hebben op duurzaamheidsaspecten alsook over de inbedding van die aspecten in de governance van de onderneming. Daarbovenop komt een verplichte uitleg over hoe het risicomanagementsysteem van de onderneming omgaat met duurzaamheid. Dit zorgt ervoor dat grotere ondernemingen met de billen bloot moeten. Dit betreft ook de relatie tussen de strategie en de impact, risico's en kansen van de onderneming die verband houden met kunstmatige intelligentie tegen de achtergrond van klimaat & milieu en de mens. Het is mogelijk dat de CSRD-wetgeving en de daarbij behorende rapportage-standaarden ondernemingen aanzetten om ten aanzien van de gebieden waar zij onderpresteren, een tandje bij te zetten. Hoewel de CSRD dus geen gedragsregels bevat, kan zij wel degelijk een dergelijke uitwerking hebben in de praktijk. De voorgestelde CSDDD bevat daar bovenop regels die gaan over de verantwoordelijkheid van bedrijven in de keten en wellicht, afhankelijk van de uitkomst van de triloog, ook een zorgplicht voor bestuurders. Dat kan in een keten van kunstmatige intelligentie nog best ingewikkeld zijn. De toepasselijkheid van deze duurzaamheidswetgeving is echter afhankelijk van de grootte van de ondernemingen en een eventuele beursnotering (de toepassingsreikwijdte is dus anders dan bij de AI-Act) waardoor veel kleinere bedrijven zowel bij de AI-Act als bij de duurzaamheidswetgeving buiten schot blijven. En dan komen we weer uit bij de corporate purpose als anker voor de onderneming. Een heldere purpose biedt de onderneming, ook de kleinere of misschien juist de kleinere ondernemingen, houvast bij de uitdagingen die kunstmatige intelligentie teweegbrengt en kan zorgen voor helderheid over de verantwoordelijkheid die de onderneming voorstaat. Voor de integratie van de purpose in de onderneming zijn de kernwaarden en beginselen van de onderneming van belang. In sommige opvattingen zijn deze onderdeel van de purpose. Wat daar ook van zij, het staat vast dat zij in ieder geval nauw aan elkaar gelieerd zijn.

2.5 *Principles en AI*

De grondbeginselen of kernwaarden reflecteren de purpose en zorgen er tegelijkertijd voor dat de purpose kan worden geoperationaliseerd (*values*). Dit zijn de idealen of concepten die de onderneming nastreeft. Het vertegenwoordigt aldus de gewenste ethische benadering van de onderneming in haar handelen en geeft ook vorm aan de impact die de onderneming wil realiseren. De kernwaarden komen (in meer of mindere mate) tot uiting in de cultuur en het gedrag. Dit zie

⁷ Zie voor een uitgebreide bespreking hiervan Deel II van mijn drieluik over Technologie en ESG: *supra* noot 2.

ik als de wijze waarop de personen die onderdeel uitmaken van de onderneming handelen en de verwachtingen daarover in de groep (*norms*). Het omvat derhalve (collectieve) gedragspatronen binnen de onderneming. Het gedrag in de top van de onderneming heeft, zoals inmiddels algemeen bekend, een belangrijke uitwerking op het gedrag in de rest van de onderneming. De onderneming heeft daarnaast formele organisatiestructuren veelal vastgelegd in beleidsdocumenten (*governance*) waarbinnen (delen van) die gedragspatronen plaatsvinden. De effectiviteit of het succes van de onderneming wordt onder andere bepaald door de mate waarin de kernwaarden, cultuur & gedrag en governance op een lijn liggen en in balans zijn. Dat is voor veel ondernemingen nog geen vanzelfsprekendheid.

2.6 *AI-beginselen*

De toegenomen aandacht voor kernwaarden, cultuur & gedrag en governance is in het kader van kunstmatige intelligentie van grote waarde. Dit komt door de vele ethische vraagstukken die AI met zich kan brengen en het feit dat er nogal veel partijen en personen betrokken kunnen zijn bij de ontwikkeling en inzet van AI. De kernwaarden en cultuur beïnvloeden de wijze waarop de organisatie omgaat met deze ethische vraagstukken en de beslissingen die als gevolg daarvan worden genomen. Verschillende bedrijven hebben inmiddels AI-beginselen of -richtsnoeren opgesteld. Voor deze beginselen en/of richtsnoeren kunnen ondernemingen inspiratie op doen bij een veelheid aan AI-codes opgesteld door organisaties of instituties zoals de OESO (*OECD AI Principles*⁸), de Europese Commissie (*AI-Act* (zie hierboven), *Europese digitale rechten en beginselen*⁹ en *Ethische Richtsnoeren voor Betrouwbare KI*¹⁰ *AI*), ISO/IEC (*verschillende standaarden*¹¹) en IEEE (*verschillende standaarden*¹²) en er zijn er nog veel meer. Deze bevatten over het algemeen vergelijkbare beginselen. Wanneer een onderneming de eigen benadering van AI in ogenschouw neemt, zal het goed zijn om de voor de onderneming geldende AI-beginselen vast te stellen en duidelijk te maken op welke wijze deze zijn verankerd in de kernwaarden. De beginselen zijn over het algemeen high-level en de kunst is om deze op een goede wijze in de onderneming ingebed te krijgen. Dat is de uitdaging waar veel ondernemingen momenteel voor staan. Het begint in ieder geval met een heldere uitleg van de beginselen en van de strategie hoe deze beginselen geoperationaliseerd worden in de onderneming. Dit heeft betrekking op de invulling van de verschillende P's (aandachtsgebieden) en de samenhang daartussen. Deze bespreek ik in de volgende paragrafen. Daarbij is het van belang dat de kernwaarden, cultuur & gedrag en governance in belangrijke mate bepalen welke financiële beslissingen en risico's worden genomen en waar de aandacht naar uit gaat binnen de onderneming. Het geeft inzicht wat binnen de onderneming van waarde wordt geacht. De heersende cultuur heeft ook grote invloed op hoe er wordt aangekeken tegen zaken op het moment dat er iets fout gaat. Het bepaalt in belangrijke mate op welke wijze er

8 <https://oecd.ai/en/ai-principles>.

9 <https://digital-strategy.ec.europa.eu/en/policies/digital-principles>.

10 <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>.

11 <https://www.iec.ch/ai>.

12 <https://standards.ieee.org/initiatives/autonomous-intelligence-systems/>.

naar de P's wordt gekeken en de wijze waarop die functioneren. Stel een AI-project heeft uiteindelijk na de lancering (grote) negatieve effecten in de samenleving. De organisatiecultuur zal bepalen waar de focus ligt bij het analyseren van wat er is misgegaan en waar er moet worden verbeterd. Zijn de purpose en de kernwaarden onvoldoende ingebed (*Purpose & Principles*)? Is de mislukking gelegen aan een gebrek aan goede processen en/of procedures (*Processes & Procedures*)? Of ligt het aan de mensen (*People*)? Of is de onderneming met de foute partners in zee gegaan (*Partnerships*)? Of zijn er te weinig perspectieven van interne en externe stakeholders meegenomen in het AI-project (*Perspectives*)? Of kijkt de onderneming met name naar wat er op het gebied van prestatie-indicatoren is gedaan (*Performance & Progress*)? Het is goed om bewust te zijn hoe de bedrijfscultuur zich verhoudt tot het systeem van de verschillende P's en waar de aandacht naar uitgaat. Door hier kritisch op te zijn en al deze aandachtsgebieden in het oog te houden, kan de onderneming blinde vlekken voorkomen wanneer AI-projecten worden geëvalueerd.

2.7 *Invloed van AI op de bedrijfscultuur*

Naast het feit dat de bedrijfscultuur invloed heeft op de wijze waarop AI wordt ontwikkeld en gebruikt, heeft AI ook invloed op de bedrijfscultuur. Zo kan er, zoals voornoemd, angst ontstaan bij werknemers op het moment dat AI wordt gebruikt in de werkomgeving. Dit kan te maken hebben met angst dat medewerkers worden vervangen door AI-applicaties, angst door een algemene onzekerheid over hoe de digitale toekomst eruit komt te zien en onzekerheid met betrekking tot beslissingen die zij moeten nemen bij de ontwikkeling en de inzet van AI. Naast angst kan de integratie van AI in werkprocessen ook invloed hebben op de menselijke interacties binnen het bedrijf. Vertrouwensstructuren kunnen veranderen, doordat medewerkers bijvoorbeeld sneller de uitkomst van een AI-supportsysteem vertrouwen dan menselijke output. Omdat ik benieuwd was naar de reactie van een GenAI-tool op de vraag welke invloed AI kan hebben op de bedrijfscultuur, heb ik deze vraag aan ChatGPT gesteld. Opvallend was dat in het antwoord de nadruk lag op de positieve effecten van AI. Het antwoord stelde dat AI-systemen beslissingen nemen die ethische implicaties hebben en dat die in beschouwing moeten worden genomen. De conclusie die ChatGPT hieraan verbond, is dat dit een cultuur teweeg kan brengen waarin ethische overwegingen en een verantwoord gebruik van AI worden geprioriteerd. En zo werden eigenlijk allerlei effecten van AI omgezet in een 'positieve' uitkomst voor de cultuur: het in ogenschouw nemen van bias zou leiden tot een cultuur waarin diversiteit en inclusiviteit wordt gepromoot, implementatie van AI-tools vergt samenwerking met een computer waardoor er een samenwerkingscultuur zou ontstaan, AI-toepassingen kunnen risico's veroorzaken, waardoor dit zou leiden tot een cultuur van risicobewustzijn. En zo waren er wel twaalf aspecten van positieve invloed. Het zou mooi zijn als dit inderdaad de uitkomst zou zijn, maar dat gaat natuurlijk niet vanzelf. Het is goed om als onderneming, door het doen van onderzoek, inzicht te krijgen wat de invloed van AI is op de eigen bedrijfscultuur. De negatieve implicaties kunnen dan worden voorkomen of verminderd en de positieve kunnen worden ondersteund.

3. Problems & Products/Projects

Vanuit de purpose wordt de bedrijfsstrategie opgesteld waaraan doelstellingen en over het algemeen KPI's worden gekoppeld. Bij AI-initiatieven, of dit nu producten/diensten of interne processen betreft, gaat het in essentie om de wijze waarop AI wordt toegepast om waarde te creëren. En waarde is een breed begrip dus het gaat om inzicht in welke waarde wordt gecreëerd en welke waarde er in de onderneming wordt nagestreefd. AI-projecten vormen geen uitzondering op de algemene project- of productcycli. Zij zijn wel gecompliceerd vanwege de invloed van zowel technische als niet-technische factoren en de betrokkenheid van veel verschillende mensen met uiteenlopende belangen. Daarnaast is het tempo van technische ontwikkeling hoog. Dit brengt een bepaalde mate van urgentie met zich.

Het uitgangspunt is dat er een probleem bestaat dat opgelost moet worden. Vanuit problemen (of anders gezegd; behoeften) die bestaan, zal de onderneming op zoek gaan naar producten/diensten of AI-toepassingen die binnen de purpose en de kernwaarden passen. Daarbij dient de onderneming zich telkens de vraag te stellen: is AI de beste oplossing voor dit probleem en zo ja, welke opties zijn er binnen het AI-spectrum en hoe verhouden deze zich tot de purpose, strategie en kernwaarden? Hier zullen de vragen naar de technische geschiktheid, compliance en de sociale-ethische aanvaardbaarheid worden gesteld. De vervolgvraag zal zijn wat het potentieel van dit product of deze dienst is, en of het schaalbaar is? Om antwoord te kunnen geven op deze vragen zal onderzoek gedaan moeten worden naar de impact (zie hiervoor de paragrafen 2, 4, 5 en 6) en de verhouding tot het risicoprofiel van de onderneming. De verschillende opties zullen bekeken moeten worden tegen de achtergrond van de capaciteiten van de onderneming. Wie en wat zijn er nodig om het initiatief te realiseren, hoe moet het productteam eruitzien en welke functies dienen allemaal betrokken te worden? Is er een platform(team) in de organisatie die de back-end technologische systemen bij elkaar brengt voor een specifieke business-toepassing of interne dienst? Het kan zijn dat de onderneming geen of onvoldoende capaciteiten heeft om het initiatief uit te voeren. Dan zal duidelijk moeten worden welke investeringen benodigd zijn om de vereiste capaciteit te creëren of dient er naar samenwerkingspartners gezocht te worden. Een raamwerk of richtlijnen waarin de stappen zijn opgenomen van dit proces waarin AI een rol speelt, zal project- en productmanagers/owners houvast bieden. De uitkomst van dit proces zal vervolgens de basis kunnen vormen voor de prioritering van AI-initiatieven.

4. Perspectives

Betekenisvolle stakeholderengagement ('SE') is een belangrijk aandachtspunt bij AI-initiatieven. Vanuit de duurzaamheidswet- en regelgeving wordt vol ingezet op het betrekken van relevante stakeholders (of hun vertegenwoordigers) bij het in kaart brengen van de impact, de risico's en kansen van de onderneming. Hierbij dient gerapporteerd te worden over de gekozen vormgeving van de SE (het vormt

een onderdeel van de verplichte assurance), wie de operationele eindverantwoordelijkheid hiervoor heeft en op welke wijze de effectiviteit wordt geanalyseerd. In de nieuwe AI-Act is de aandacht voor de betrokkenheid van stakeholders bij de identificatie van risico's summierder. Op basis van de nu bekende voorgestelde regels zal de nieuwe AI-Act voor bepaalde AI-systemen en foundationmodellen regels bevatten waarin SE een rol speelt maar deze zijn niet zo uitgebreid en (nog) niet consistent.¹³

Voor AI-projecten is het van belang dat stakeholders al worden betrokken voordat er besloten wordt om bepaalde diensten of producten op de (of een andere geografische) markt te brengen of om AI in de bedrijfsprocessen te integreren. Het is van belang SE op te nemen in de verschillende stadia van de project/productcyclus. Om de SE binnen de onderneming te ondersteunen kan er gebruik worden gemaakt van speciale SE-beleidsdocumenten of handboeken gericht op AI-projecten, tools die de SE kunnen ondersteunen, training en evaluatie. Hierbij zal de onderneming allereerst kennis moeten hebben van de verschillende beschikbare methodologieën van SE in de productcyclus. In de meer traditionele benadering wordt gebruikgemaakt van een *User-centered designbenadering* (identificeren en verzamelen van data over gebruikersbehoeften bij het ontwerp). Dit is een beperkte vorm van SE. In het kader van betekenisvolle SE schuift dit steeds meer op naar *participatory design, development, deployment and monitoring methods* en *participatory data stewardship practices*¹⁴ (directe input en beslissingsinvloed van stakeholders bij de verschillende stadia van de productcyclus en dataverwerking). Keuzes hierin zijn afhankelijk van de doelstelling die ten grondslag ligt aan de SE. Voor betekenisvolle AI-stakeholderbetrokkenheid zal de nadruk meer komen te liggen op een vorm die recht doet aan een gedeelde doelstelling. Dus een doelstelling die zowel tegemoetkomt aan de belangen van de onderneming als die van de stakeholder. Kunstmatige intelligentie kan vervolgens op verschillende wijze behulpzaam zijn bij SE, denk aan chatbots en analysetools.

Ter ondersteuning van ondernemingen zijn verschillende raamwerken voorhanden zoals de OESO Due Diligence Handreiking voor Maatschappelijk Verantwoord Ondernemen.¹⁵ Naast de algemene richtlijnen, is ook een aantal richtlijnen specifiek

13 Dit zou in de trilog nog kunnen wijzigen. Daarnaast kan stakeholderbetrokkenheid in de CEN-CENELEC standaarden, de standaarden die momenteel worden ontwikkeld in het kader van de AI-Act (zie paragraaf 7), een grotere rol toebedeeld krijgen bij de vormgeving van het risicomanagementsysteem: [https://ec.europa.eu/transparency/documents-register/detail?ref=C\(2023\)3215&lang=en](https://ec.europa.eu/transparency/documents-register/detail?ref=C(2023)3215&lang=en).

14 Ada Lovelace Institute, 'Participatory data stewardship, data for the public good: A framework for involving people in the use of data', 2021: https://www.adalovelaceinstitute.org/wp-content/uploads/2021/11/ADA_Participatory-Data-Stewardship.pdf.

15 <http://mneguidelines.oecd.org/Oeso-Due-Diligence-Handreiking-voor-maatschappelijk-verantwoord-ondernemen.pdf>. Deze handreiking bevat de randvoorwaarden waaraan betekenisvolle stakeholderengagement dient te voldoen. Dit houdt bijvoorbeeld in dat het dient te gaan om continue interactieve processen gebaseerd op goede wil en een open communicatiewijze en op wederzijdse betrokkenheid bij de invulling en uitvoering daarvan. Hierbij wordt rekening gehouden met mogelijke barrières voor kwetsbare en gemarginaliseerde stakeholdergroepen. De onderneming spant zich in om haar besluitvorming te baseren op de opvattingen van de getroffen stakeholders die in de dialoog naar voren komen. Van belang zijn de volgende elementen: (1) Vaststelling welke stakeholders worden betrokken met aandacht voor mogelijke bias (identificatie, analyse, mapping, prioritering van stakeholders), (2) een onderbouwde keuze voor een bepaalde vorm van dialoog met helderheid over het proces, dit proces voldoet aan bepaalde beginselen (zoals inclusiviteit, integriteit en openheid), (3) informatievervalsing gedurende het gehele proces en (4) mechanismen via welke stakeholders hun zorgen of klachten kenbaar kunnen maken.

voor AI ontwikkeld. Soms worden deze apart gepresenteerd, soms opgenomen in een raamwerk voor impactassessments waar SE een onderdeel van vormt.¹⁶ De operationalisatie van de SE binnen de onderneming hangt sterk af van de beschikbare tijd, resources en support van het hoger management. Bij het ontwikkelen van technologieën zal het niet altijd even makkelijk zijn om de verschillende stakeholders gedurende het gehele traject te betrekken. Door de enorm snel veranderende omstandigheden is tijdsgebrek kenmerkend voor AI-projecten. Hierbij zal er veelal worden geprioriteerd op basis van de interne risico-inschatting. Ook voor de vormgeving van de SE kan een risico-gebaseerde benadering worden toegepast. Daarbij dient men beducht te zijn op de eigen bias in het kader van de vaststelling van dat risico. Veel van de SE-activiteiten vinden op de werkvloer plaats. Meer ingrijpende use-cases en beslissingen zullen, eventueel via een ethische commissie, bij de bestuurders terechtkomen. Uit onderzoek blijkt dat de volgende uitdagingen gedeeld worden door ondernemingen als het gaat om SE bij AI-projecten:¹⁷ top-down organisatorische kaders bepalen voor een belangrijk deel wie van de stakeholders betrokken worden, de betrokkenheid betreft vaak de 'hoe'-vraag maar niet de 'of'-vraag, dus of een AI-toepassing wel op de markt gebracht dient te worden en de SE-verantwoordelijken binnen de onderneming zijn veelal de regisseur van alle aspecten van de SE waardoor er weinig input van stakeholders is hieromtrent. Er zullen naar verwachting steeds meer ervaringen worden gedeeld ten aanzien van best practices. Het is ook goed om het delen van best practices binnen de onderneming zelf te organiseren en stimuleren.

5. People

People heeft betrekking op het sociale aspect van AI. Hieronder kunnen vallen de eigen medewerkers, de werknemers in de keten, gemeenschappen en ten slotte consumenten en eindgebruikers. Algemene effecten heb ik in paragraaf 2.3 besproken. Ik beperk me hier tot de eigen medewerkers. De AI-ontwikkelingen hebben op verschillende wijze invloed op medewerkers. De impact zal in de eerste plaats afhangen van de functie die de medewerkers bekleden en in hoeverre en op welke wijze zij te maken krijgen met de toepassing van AI. Daarnaast zal die impact ook nauw samenhangen met de karakters van de individuen, hun eigen opvattingen over (de implicaties van) AI en de groepsdynamiek. Hoe ervaren zij integratie van AI in bedrijfsprocessen of hoe gaan zij om met moeilijkheden en onduidelijkheden die zij tegenkomen in het kader van de ontwikkeling van producten met een AI-component? Sommige medewerkers zullen het misschien best een beetje spannend

16 Zie bijvoorbeeld: <https://ecnl.org/sites/default/files/2023-03/Final%20Version%20FME%20with%20Copyright%20%282%29.pdf> en Alan Turing Institute, 'Human Rights, Democracy, and the Rule of Law Assurance Framework for AI Systems: A proposal prepared for the Council of Europe's Ad hoc Committee on Artificial Intelligence', 2021: <https://www.turing.ac.uk/news/publications/human-rights-democracy-and-rule-law-assurance-framework-ai-systems>.

17 F. Delgado e.a., 'Stakeholder Participation in AI: Beyond "Add Diverse Stakeholders and Stir"', Workshop on Human-Centered AI at Conference on Neural Information Processing Systems (NeurIPS 2021): https://www.researchgate.net/publication/355841587_Stakeholder_Participation_in_AI_Beyond_Add_Diverse_Stakeholders_and_Stir.

vinden en kunnen onzeker zijn over de eigen capaciteiten ten aanzien van de omgang met AI of het begrip ervan. Ook vergt implementatie van AI nauwe samenwerking met collega's met een andere disciplinaire achtergrond. Het is dan soms zoeken naar het spreken van dezelfde taal en ook dat kan tot onzekerheid leiden. Deze groep met bepaalde angsten of onzekerheden kan tegelijkertijd heel enthousiast zijn over AI. Dat betekent vruchtbare grond voor het *upskillen* of *reskillen* van deze medewerkers door bijvoorbeeld trainingen en het identificeren waar zij hun talenten kunnen en willen inzetten. Het vormen van *communities of practice*, dit zijn groepen van mensen die informeel met elkaar verbonden zijn door een gedeelde passie, expertise of wens, kan dan goed werken. In deze groepen worden kennis en ervaring gedeeld om te leren en om te gaan met praktijkuitdagingen of -problemen.

Ten opzichte van de eerste groep kan er ook een groep zijn die misschien al wat meer van AI afweet maar minder te spreken is over de mogelijke effecten. Deze groep heeft twijfels over de ontwikkeling en inzet ervan. Hoewel een continu proces van leren in de context van technologie voor iedereen van belang is, zal de voornoemde skills-training voor deze groep niet de hoogste prioriteit hebben. Het gesprek aangaan met deze groep is wat belangrijk is. Wat zijn de zorgen precies en wat ligt aan deze zorgen ten grondslag? Hoe kan de onderneming daarmee omgaan, wat kan de onderneming ervan leren? En zo zijn er binnen de onderneming nog wel meer 'soorten groepen' te vinden als het gaat om de verhouding tot AI. Het is derhalve belangrijk om inzicht te krijgen wat de specifieke situatie binnen de onderneming is als het gaat om de houding en capaciteiten in het kader van AI en hoe de verdeling tussen de groepen is. Dan kunnen de activiteiten die nodig zijn voor een succesvolle transitie, hierop worden aangepast. Ook wordt dan duidelijker wat en hoeveel tijd er nodig is voor deze transitie. Over de hele linie geldt dat medewerkers baat hebben bij activiteiten die bewustzijn creëren met betrekking tot de gevolgen van AI en hoe de onderneming zich daartoe verhoudt. Dat betreft zeker niet alleen de negatieve impact. Een onderneming kan er ook baat bij hebben bewustzijn te creëren ten aanzien van de positieve effecten. Dit betekent dat duidelijk moet worden hoe de purpose, de kernwaarden en AI-richtlijnen zich verhouden tot deze effecten. Het houden van workshops met medewerkers vanuit de verschillende functies waarin casestudies worden behandeld, kan dan goed helpen. Daardoor komen de purpose en de kernwaarden tot leven en blijven het geen lege mooie woorden. Ook wordt er gewerkt aan het onderling begrip en bewustzijn van elkaars rollen.

Dit betekent dus kort gezegd dat er aandacht moet zijn voor de verschillen tussen medewerkers als het gaat om capaciteiten maar ook de impact die AI op hen heeft gelet op de verschillende karakters. Onzekerheden kunnen gepaard gaan met schaamte. Dit kan tot gevolg hebben dat er veel in de onderstroom gebeurt. Overigens hebben deze verschillen niet alleen betrekking op medewerkers in de lagere niveaus van de organisatie. Dit geldt zeker ook voor het leiderschapsteam en de bestuurders. Een open communicatie is van groot belang om de transitie die voor de deur staat goed te laten verlopen waarbij er niemand vergeten wordt. Heb aandacht voor (groepen van) medewerkers waarvoor het wellicht allemaal wat snel

gaat. Neem ze mee en geef ze een bepaalde verantwoordelijkheid zodat ze onderdeel worden van de transitie.

Heldere communicatie over de wijze waarop AI in de processen van de onderneming wordt gebruikt alsook in de producten en diensten die het bedrijf levert, is belangrijk voor vertrouwen. Ook hierbij dient rekening te worden gehouden met de verschillen in kennis en expertise op het gebied van technologie en de kennis en informatiebehoeften van de medewerkers. En bij dit proces is het van belang dat er voldoende aandacht wordt besteed aan de verandermanagementcapaciteiten binnen de organisatie. Dit is uiteindelijk van cruciaal belang voor het succes van de transitie. In dit kader wordt in toenemende mate een Chief Transformation Officer aangesteld in het bestuur, die verantwoordelijk is voor het aansturen en ondersteunen van het transitieproces binnen de onderneming. Dat is dus niet de Chief Information/Technology Officer (CI/TO) of de Chief Digital Officer, maar iemand die sterk is in programmamanagement en synergie creëert, bruggen kan bouwen en ervoor kan zorgen dat er binnen de onderneming verandercapaciteiten worden geborgd. Veelal heeft de CTO ook een belangrijk strategisch element in zijn takenpakket waardoor hij/zij een belangrijke factor is in het vormgeven van de digitale strategie als integraal onderdeel van de bedrijfsstrategie en de commercialisatie van AI. De CTO kan worden ondersteund door *champions* of *scrummasters* die ervoor zorgen dat de verandering binnen de onderneming landt. Een ander aandachtsgebied is het aantrekken van nieuw talent. De arbeidsmarkt voor ‘tech-talent’ is krap en wordt alleen maar krappere. Dit kan leiden tot de situatie dat er nieuwe medewerkers worden aangenomen die eigenlijk niet zo goed passen bij de organisatie omdat er geen goed ander alternatief is. Ook hier geldt dat de organisatie goed moet nadenken hoe hier mee om te gaan omdat dit tech-talent voor een groot deel verantwoordelijk is voor de operationalisatie door de *values-by-design*-benadering.

6. Planet

Aan het begin van de periode dat er in toenemende mate aandacht was voor de implicaties van AI ging deze vooral uit naar de (directe) impact op de mens, de sociale implicaties. Het werd dan ook veelal met ‘*Trustworthy AI*’ aangeduid. Steeds meer verschuift die aandacht nu tevens naar de impact van AI op de planeet en dus naar ‘*Sustainable AI*’ in de brede zin. Duidelijk wordt dat de inzet van deze technologie niet zonder gevolgen is voor het klimaat & milieu. Hoewel er steeds meer initiatieven zijn om bijvoorbeeld het energie- en waterverbruik van datacenters te verminderen (zie bijvoorbeeld de recente vermindering in het energieverbruik met 85% bij een datacenter van AMS-IX, het Amsterdamse internetknooppunt¹⁸), zal de technologie altijd een (eco)footprint hebben. De nieuwe voorgestelde wetgeving voor AI-systemen (met name de compromistekst van het Europees Parlement) vereist dat er inzicht wordt gegeven in deze footprint en verlangt van sommige AI-ontwikkelaars

¹⁸ <https://www.ams-ix.net/ams/news/upgrade-at-ams-ix-result-in-85-power-reduction>.

zelfs dat zij de AI-systemen op de meest (*state-of-the-art*) energie-efficiënte wijze ontwikkelen en dat zij ervoor zorgen dat het energieverbruik tijdens de inzet daarvan wordt gemeten en gedocumenteerd. Het uitvoeren van een *Environmental Impact Assessment* zal dus waarschijnlijk voor bepaalde AI-systemen verplicht worden gesteld. Toch zou het gelet op de context ook voor andere systemen belangrijk zijn om inzicht te krijgen in de E-impact. Het is niet uitgesloten dat dit in de toekomst alsnog in wetgeving opgenomen wordt omdat hierdoor ook de ESG-rapportage van ondernemingen wordt vergemakkelijkt, bijvoorbeeld in het kader van de Scope 1 en 3 emissie-rapportages. Daarnaast geldt dat ook gebruikers bewust om moeten gaan met de impact van het gebruik. Ondernemingen zouden moeten nadenken op welke wijze zij in de besluitvormingsprocessen over AI en daarnaast in het kader van individueel gebruik van technologieën zoals GenAI, aandacht kunnen geven aan deze E-impact.

7. Processes & Procedures

De nieuwe AI-wetgeving zal voor hoog-risico AI-systemen en daarnaast foundationmodellen bepalen waaraan de processen en de procedures die het ontwerp, de ontwikkeling, de inzet, de inkoop, de documentatie en de monitoring van AI omgeven. Daarnaast bevat deze wetgeving regels ten aanzien van een conformiteitsbeoordeling. Deze regels zullen verder uitgewerkt worden in standaarden. De Europese Commissie heeft in mei 2023 een verzoek uitgevaardigd aan de Europese Standaarden Organisaties, CEN (*European Committee for Standardisation*) en CENELEC (*European Committee for Electrotechnical Standardisation*), om standaarden te ontwikkelen op het terrein van tien aandachtsgebieden.¹⁹ Het idee is dat deze nauw aansluiten bij de ISO/IEC-standaarden. Voor bepaalde AI-systemen zijn de regels in specifieke product- of sectorwet- en regelgeving opgenomen. De AI-Act bevat regels ten aanzien van verschillende aspecten van het kwaliteitsmanagementsysteem waarin een AI-systeem is ingebed. De onderneming dient tevens de verantwoordelijkheden van het management en medewerkers in het kader van de AI-kwaliteitsmanagementactiviteiten vast te stellen. De regels voor het risicomanagementsysteem zijn gedetailleerder en geven specifiek invulling aan de vereiste risicomitigatiemaatregelen die gericht zijn op data & datagovernance, transparantie & informatievervalsing, menselijk toezicht, juistheid, robuustheid & cybersecurity en technische documentatie. Het betreft aldus zowel organisatorische processen als technische maatregelen.

Aanbieders en gebruikers van AI-systemen die niet binnen de reikwijdte van deze regels vallen, worden aangemoedigd om toch vergelijkbare maatregelen te nemen. Van belang is dat de onderneming inzicht krijgt waar de verantwoordelijkheid ligt als het gaat om AI-initiatieven en wat de escalatieprocedures zijn. Steeds vaker

¹⁹ Dit zijn: datagovernance and data quality, record keeping, provision of information and transparency, robustness, accuracy, cybersecurity, human oversight, riskmanagement and testing, conformity assessment and quality management.

stellen ondernemingen een *Ethics board*, *AI Ethics Committee* of *Data & AI Ethics Council* in. Deze gremia spelen een belangrijke rol in het operationaliseren van de kernwaarden en ethische richtsnoeren door uitleg te geven over de interpretatie van deze waarden en bijvoorbeeld het beoordelen van en adviseren over materiële use-cases. Zij kunnen bijvoorbeeld ondersteuning bieden als er sprake is van botsende waarden bij de ontwikkeling en inzet van AI. Het is in dit soort gevallen dat de high-level waarden concreet moeten worden gemaakt. Zij kunnen zodoende de riskfunctie ondersteunen als het gaat om specifieke AI-risico's en daarnaast bij besluitvorming over aanvaardbaar risico. Zij kunnen bovendien het bestuur van strategisch advies voorzien en kunnen ondersteuning bieden in het kader van de commerciële strategie. Afhankelijk van hun takenpakket fungeren zij in sommige gevallen ook als het contactpunt voor klokkenluiders.

8. Partnerships

Om het potentieel uit de mogelijkheden van digitalisering te halen zal het niet altijd verstandig zijn om als onderneming telkens zelf het wiel uit te vinden. Soms is dit gewoonweg ook niet mogelijk vanwege de gigantische investering die het vergt om bijvoorbeeld een LLM te ontwikkelen. Het kan ook een wijze bedrijfsbeslissing zijn om de ontwikkeling van bepaalde technologie aan anderen over te laten. De AI-ontwikkelingsketen bestaat veelal uit meerdere partijen waarbij gebruikt wordt gemaakt van data die uit veel diverse bronnen afkomstig is. Daarnaast zal er waarschijnlijk gebruik worden gemaakt van de cloud en ook voor de infrastructuur zal er worden samengewerkt met externe partijen. De mogelijke impact van de onderneming hangt daardoor nauw samen met de partijen waarmee wordt samengewerkt. De focus zal derhalve niet alleen maar moeten liggen op wat er binnenshuis gebeurt. De uitdaging hierbij is dus: met welke partner ga je in zee? Hoe maak je die keuzes? Voor procurement is van belang dat nauw wordt samengewerkt met verschillende functies binnen de onderneming. Zo zullen de business, legal/compliance, IT alsook data-teams allemaal bepaalde requirements opstellen waaraan de AI-technologie die wordt ingekocht, moet voldoen. Ook hiervoor geldt dat de onderneming een proces zal instellen om te zorgen dat met de juiste partners wordt samengewerkt. Het World Economic Forum heeft in juni 2023 een richtlijn gepubliceerd specifiek voor AI-procurement in de private sector (*Adopting AI responsibly: Guidelines for Procurement of AI Solution by the Private Sector*) die daarbij behulpzaam kan zijn.²⁰ Het kan de integratie van AI-specifieke aspecten in het inkoopproces ondersteunen. De aandacht voor IT-partnerships en het belang daarvan komt bijvoorbeeld ook terug in de Verordening betreffende de digitale operationele weerbaarheid van de financiële sector (DORA).²¹ Deze schrijft onder andere voor dat financiële instellingen beleidsmaatregelen dienen te hebben in het kader van het ICT-risico van derde-aanbieders.

20 <https://www.weforum.org/reports/adopting-ai-responsibly-guidelines-for-procurement-of-ai-solutions-by-the-private-sector/>.

21 <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32022R2554>.

9. Performance & Progress

Om ervoor te zorgen dat de onderneming de gestelde doelstellingen behaalt en de voorgestane verantwoordelijkheid op het gebied van AI realiseert, is het verstandig om bepaalde KPI's en metrics vast te stellen. Met de juiste KPI's zal de onderneming de eigen prestaties kunnen evalueren en voortgang kunnen monitoren. In het kader van AI helpt dit om vast te stellen of AI-initiatieven in lijn liggen met de strategische doelstellingen, kan er besluitvorming plaatsvinden over investeringen en maakt het de verwachtingen helder voor de betrokkenen bij het AI-initiatief. Dit zullen enerzijds metrics betreffen die waardecreatie meten, anderzijds gaat het om delivery-of-work en bijvoorbeeld model metrics. Bepaalde ondernemingen zullen dit al hebben gedaan, andere worden door onder andere de CSRD en de maatschappelijke aandacht hiervoor aangezet tot het nadenken over ESG-KPI's en metrics.²² AI kan ook gebruikt worden om de performance van KPI's zelf vast te stellen. Gebruikt de onderneming eigenlijk wel de juiste indicatoren, dienen deze niet gemoderniseerd te worden? Doordat AI bepaalde correlaties kan opsporen, die niet door personen kunnen worden vastgesteld, kunnen de KPI's worden geoptimaliseerd, nieuwe KPI's worden vastgesteld en de verbanden tussen KPI's worden blootgelegd.²³

10. Protection

Het aandachtsgebied Protection is gericht op de bescherming van de onderneming tegen risico's die voortvloeien uit het gebruik van AI-toepassingen. Dit betreft onder andere beveiliging van AI-toepassingen die bijvoorbeeld het doelwit kunnen zijn van cyberaanvallen. Het gevaar bestaat dat AI-systemen worden overgenomen door partijen met kwade wil. Ook is het mogelijk dat de AI-toepassingen worden voorzien van schadelijke input zoals data die desinformatie bevat. De impact van deze gevaren verschilt per sector of branche. Maatregelen om tegen deze ongewenste praktijken gewapend te zijn, betreffen de wijze waarop de software wordt ontwikkeld maar ook de toegang tot systemen. De voorgestelde AI-Act bevat regels ten aanzien van de beveiliging en er zal een specifieke standaard voor AI worden opgesteld door CEN-CENELEC in het kader hiervan. Aangezien het hoofdstuk van Sandra Konings in dit *Jaarboek Corporate Governance* uitgebreid ingaat op grip op cybersecurity, verwijs ik voor een verdere uitleg naar dat hoofdstuk.

22 Zie bijvoorbeeld de ESRS: ANNEX to the Commission Delegated Regulation (EU) ... supplementing Directive 2013/34/EU of the European Parliament and of the Council as regards sustainability reporting standards, ESRS1, application requirement 7. De delegated act is op 31 juli aangenomen door de Europese Commissie, maar nog niet gepubliceerd in de Official Journal: https://finance.ec.europa.eu/regulation-and-supervision/financial-services-legislation/implementing-and-delegated-acts/corporate-sustainability-reporting-directive_en en World Economic Forum (2020), 'Measuring Stakeholder Capitalism Towards Common Metrics and Consistent Reporting of Sustainable Value Creation', Whitepaper, september 2020: https://www3.weforum.org/docs/WEF_IBC_Measuring_Stakeholder_Capitalism_Report_2020.pdf.

23 <https://sloanreview.mit.edu/article/improve-key-performance-indicators-with-ai/>.

Conclusie

Ondernemingen kunnen niet meer om AI heen. Uitdagingen waar ondernemingen voor staan zijn onder andere het helder krijgen hoe de purpose en kernwaarden van de onderneming zich verhouden tot het AI-domein, zicht krijgen op de dynamiek tussen cultuur en AI en hoe de medewerkers zich verhouden tot AI. Ondernemingen zullen hun aandacht richten op het identificeren van de juiste business cases en de daarbij behorende investeringen, het realiseren van de juiste skills en capaciteiten bij de medewerkers, het organiseren van betekenisvolle stakeholderbetrokkenheid, het zorgen voor solide data en datamanagement, het in goede banen leiden van infrastructurele uitdagingen, het realiseren van een goede vertaling van complexe compliance aspecten als gevolg van overlappende en gestapelde wet- en regelgeving, het samenwerken met de juiste partners en dat alles op een veilige manier en in een geoliede governance-structuur. Het is nogal wat. Een succesvolle benadering van AI vergt derhalve een goed doordachte strategie, open communicatie en samenwerking. Aandacht voor de verschillen in achtergrond, expertise en percepties van medewerkers in het kader van AI is hierbij van belang en adequaat verandermanagement cruciaal. Het zal regelmatig schipperen zijn tussen tijdsdruk en beschikbare middelen enerzijds en het op verantwoorde wijze uitvoering geven aan AI-initiatieven anderzijds. Daarbij zullen hopelijk steeds meer ondernemingen toewerken naar een 'CSquadupleD': een Corporate Sustainable Digital Design, Development & Deployment. De wijze waarop door ondernemingen invulling wordt gegeven aan de verschillende aandachtsgebieden die onderdeel zijn van het P+AI-model, zal bepalen of zij daarin succesvol zullen zijn.