VERSLAG VAN EEN DESKUNDIGENBIJEENKOMST

Vastgesteld 8 november 2021

De vaste commissie voor Justitie en Veilligheid (J&V) heeft op 5 oktober 2021 gesprekken gevoerd over:

- Grip op algoritmische besluitvorming bij de overheid. De rol van de Eerste Kamer (Rathenau Instituut);

- Een blik op de toekomst van verantwoorde AI: mens over AI over AI (presentatie Catholijn Jonker);

- Eisen aan AI (presentatie dr. Melanie Rieback).

Van deze gesprekken brengt de commissie bijgaand geredigeerd woordelijk verslag uit.

De voorzitter van de vaste commissie voor Justitie en Veiligheid,
De Boer

De griffier van de vaste commissie voor Justitie en Veiligheid,
Van Dooren

Voorzitter: Recourt
Griffier: Van Dooren

Aanwezig zijn acht leden der Kamer, te weten: Crone, Dittrich, Faber-van de Klashorst, Jorritsma-Lebbink, Kluit, Nanninga, Nicolaï en Stienen,

alsmede mevrouw Jonker en mevrouw Rieback.

Aanvang 09.00 uur.

Technische ontwikkelingen AI: wat is AI, wat kan met AI, kansen en risico's / uitlegbare AI, eisen aan data

De **voorzitter**: Goedemorgen. Welkom bij deze eerste bijeenkomst van de Eerste Kamer over kunstmatige, of artificiële, intelligentie. Het is een technische briefing, de eerste van drie. Vandaag is de inleidende briefing. Ik heet iedereen van harte welkom, mede namens Gala Veldhoen, die hier vandaag helaas niet kan zijn. Ik ben vereerd op deze plek te mogen zitten, waar normaal de Voorzitter zit. Die zit nu in de zaal en ik mag zijn plek innemen. Dat is een voorrecht.
Het gaat vandaag over mens en machine. De grootste samenvatting, met name voor de Eerste Kamer, is vooral: welke regelgeving moet er nu komen om mens en machine op een goede manier samen te laten werken, op een manier die niet discriminerend is en werkt voor burgers? We willen dus inzicht geven in de aard en uitdagingen van kunstmatige intelligentie. Er wordt bijzondere aandacht gegeven aan het streven om uitlegbare, veilige, privacyvriendelijke en onbevooroordeelde AI-systemen te bouwen en te gebruiken. Deze briefing dient ter voorbereiding van de twee verdiepende sessies, een volgende week en een in de week na het reces, over de rol van de Eerste Kamer bij het "controleren en medevormgeven van algoritmische overheidsbesluitvorming", zoals hier staat genoteerd.
De eerste spreker vandaag is mevrouw Catholijn Jonker. Zij is hoogleraar interactieve engineering aan de Technische Universiteit Delft. Zij zal uitleggen wat artificiële intelligentie precies is, welke soorten systemen er bestaan en wat we van verantwoorde AI mogen verwachten. Ook licht ze toe waarom het complex is om sommige AI-systemen uit te leggen en hoe ze daar in haar onderzoek oplossingen voor ontwikkelt. Ik weet niet precies hoelang dit gaat duren, maar we hebben er in ieder geval ongeveer het eerste uur voor gereserveerd. Mogelijk zal het iets korter duren.
Deze bijeenkomst duurt formeel tot 11.00 uur, maar ik denk dat we iets eerder moeten stoppen, vanwege de verplichtingen die we allemaal elders hebben. De tweede spreker is

mevrouw Melanie Rieback. Zij is CEO van Radically Open Security, zeg ik met mijn steenkolenengels, een non-profit computer security consultancy. Zij zal spreken over de eisen die de samenleving aan kunstmatige intelligentie mag stellen, bijvoorbeeld ten aanzien van vooroordelen, veiligheid en privacy, en over de politieke, maatschappelijke en technische uitdaging om aan die eisen te voldoen.
Dan ben ik door mijn introductie heen, ware het niet dat ik ook nog onze externe gasten van de Rekenkamer, de Raad van State en de Tweede Kamer moet welkom heten, die hier fysiek dan wel via de livestream aanwezig zijn. Tot slot zeg ik nog dat wij zeer veel dank verschuldigd zijn aan het Rathenau Instituut, dat ons heel goed heeft geholpen bij de inhoudelijke en praktische voorbereiding van deze bijeenkomst. Mevrouw Jonker, mag ik u het woord geven?

Presentatie 1: Een blik op de toekomst van verantwoorde AI: mens over AI over AI

Gesprek met:
- mevrouw Catholijn Jonker (hoogleraar aan de TU Delft en Universiteit Leiden)

Mevrouw **Jonker**: Dank u wel, meneer de voorzitter. Het is natuurlijk een voorrecht om in deze illustere zaal zulk illuster publiek toe te mogen spreken. Ik ben inderdaad hoogleraar kunstmatige intelligentie, zowel in Delft als in Leiden. Ik houd mij bezig met interactieve intelligentie en met explainable AI, zoals we dat noemen. Ik laat u maar gelijk wennen aan de termen waarmee ik tussendoor waarschijnlijk steeds zal gooien. Ik ben iemand die kunstmatige intelligentie en artificial intelligence rustig als AI afkort, maar ik zie dat in de maatschappij tegenwoordig vaker gebeuren, dus ik denk dat u dat wel redt.
Ik wil u een blik op de toekomst geven, op het onderzoek dat natuurlijk niet alleen ik, maar vele anderen met mij proberen te doen om verantwoorde kunstmatige intelligentie te creëren. De visie die ik vandaag aan u presenteer, is: mens over AI, en AI over AI, dus mens over AI over AI. Dus ik maak een stapeling. Als u de slides van tevoren al bekeken heeft, zult u dat misschien zien.
Ik wil hierbij graag ook even noemen dat we in Delft een Design for Values Institute hebben opgericht, precies met dit doel, namelijk om technologie die we ontwikkelen tot ondersteuning te maken van de normen en waarden van onze maatschappij, ten dienste van sociale waarden.
Daarnaast wil ik nog even wijzen op het embleem rechtsboven op de eerste slide, dat staat voor: Hybrid Intelligence. Hybrid Intelligence is een Zwaartekrachtprogramma. Dat is een groot onderzoeksprogramma dat tien jaar duurt, waarin we ons precies richten op dit soort

vraagstellingen. Hoe kunnen we intelligentie maken die de mens ondersteunt bij zijn taken en daarin zelfs als een soort partner fungeert?
Ik ga door naar de tweede slide, waarin dat toekomstbeeld opnieuw is neergezet. U kunt dat natuurlijk niet zien, maar dit was ooit een geanimeerde presentatie. Aan de rechterkant stonden er allemaal van die akelige dingen, waarvan u zich zeker ook bewust bent, die ook met AI worden bereikt. Denk aan het produceren van fake news, het kunnen manipuleren van opinies op sociale media en het bekende trollingeffect, waarbij mensen via het internet, via sociale media, worden gepest en getreiterd. Daar zitten aspecten onder die ook met kunstmatige intelligentie te maken hebben.
Daarnaast zijn er natuurlijk voorbeelden te over van kunstmatige intelligentie die wordt ingezet met alle beste bedoelingen, maar waarbij zich ineens side-effects voordoen die helaas toch een heel akelige uitwerking hebben. Dat is de discriminerende werking van AI. Ik zal het voorbeeld noemen van een experiment waarin een chatbot werd gelanceerd door een van de grotere techreuzen. Ik geloof dat die binnen twee dagen in staat was om fascistische taal uit te slaan en iedereen uit te schelden die in de buurt kwam. Dat is natuurlijk niet de bedoeling. Zij zijn zich ook rot geschrokken en hebben het natuurlijk teruggetrokken. Het geeft wel aan dat er nog heel veel te doen is op het gebied van kunstmatige intelligentie.
Op deze slide staan de uitgangsprincipes van The Hybrid Intelligence Centre, dat we hebben opgericht om de autonomie van de mens te versterken, om de mens rijkere ervaringen te geven, om de mens in staat te stellen nieuwe activiteiten te ontwikkelen die hij daarvoor niet kon ontwikkelen en om de democratie te versterken en niet te ondermijnen. Kortom, wij streven naar een AI die is afgestemd op onze menselijke waarden. Met dat toekomstbeeld wil ik u graag iets meer gaan vertellen over kunstmatige intelligentie als partneridee. Op de volgende slide ziet u daar een aantal voorbeelden van, eigenlijk om het begrip van AI als partner langzaam op te bouwen. Deze voorbeelden zijn heel bewust gekozen, omdat we de intelligentie van AI willen benadrukken, maar ook de mogelijkheid om autonome systemen te ontwikkelen met eigen intelligentie. Hoe kan dat dan goed samenwerken met de mens?
Het voorbeeld linksboven is een ruiter te paard. Het paard heeft duidelijk een eigen intelligentie en een eigen wil. Het is in staat om hard te lopen, hoog te springen en grote lasten te dragen en te trekken. Het heeft zijn waarde in de duizenden jaren in het verleden voor ons bewezen. Maar het interessante van dit beeld, van die ruiter op dat paard, is wel heel bijzonder. U ziet een paar kleine teugels lopen. Als u die op dit plaatje wilt zien, moet u wel goed kijken. Met die teugels en met de knieën kan de ruiter het paard sturen. Als het paard echt niet wil, heeft de ruiter een enorme klus om dat paard de goede kant op te krijgen. Het gaat eigenlijk niet met geweld. Op het moment dat we een paard met geweld

trainen, zal het minder goed functioneren dan wanneer we een echt partnerschap met dat paard ontwikkelen. Dan gebeurt er iets moois. Dan krijg je een situatie waarin het paard met ruiter hoger springt dan een paard zonder ruiter, omdat de intelligentie van mens en dier samengevoegd worden. De mens heeft een beter vermogen om diepte te zien en kan daarmee de afstand tot het object en de hoogte van het object beter inschatten dan het paard. Dat is natuurlijk het soort voorbeelden waarvan we denken: daarmee leg je uit wat een partnerschap tussen twee echt essentieel verschillende stukken intelligentie zou kunnen behelzen.
Het middelste voorbeeld is aan de ene kant schattig, maar aan de andere kant geeft dit een kind dat bijvoorbeeld een epileptische aandoening heeft een stuk vrijheid in leven. Het verrijkt de mogelijkheden, omdat het veilig is voor het kind om zijn leven te voeren en de hond er is om aan te duiden: nu moet je gaan zitten, want er komt een aanval. Dat is voor het kind zelf niet mogelijk en voor de hond wel. De hond als partner hoef ik niet verder uit te leggen.
Met dat beeld gaan we naar het beeld rechtsonder. Rechtsonder ziet u een robotje, een NAO, in interactie met een kind. Er worden op het moment veel onderzoeken gedaan om kinderen bijvoorbeeld bij te staan die in het ziekenhuis verblijven om te leren omgaan met diabetes, een aandoening die ze op jonge leeftijd kunnen krijgen. Ze moeten leren hoe ze met gezonde voeding om moeten gaan, hoe ze met hun activiteiten om moeten gaan. Maar ze zitten in een ziekenhuis. Niet iedereen heeft altijd tijd voor ze. Het zou toch wel heel fijn zijn als er een buddy voor ze is. Dit robotje wordt dan ingezet als een buddy in de zin dat het spelletjes met ze kan spelen, ook spelletjes gerelateerd aan diabetes, en daarbij extra kennis kan overdragen aan het kind. Echter, de robot zelf is een zeer wankel stukje techniek. Echt goed lopen is er echt nog niet bij voor deze NAO. Goed kijken waar hij is, vindt hij ook erg lastig. Dat kan het kind veel beter. Ook daar hebben we voorbeelden gezien dat het kind zich ontfermt over de robot en zorgt dat de robot op de juiste plekken komt te staan om spelletjes met ze te spelen en niet omvalt. Zo ontstaat er een band tussen twee verschillende vormen van intelligentie die allebei wat aan elkaar hebben. Dat is het beeld dat we voor ogen hebben.
Maar wat is kunstmatige intelligentie dan eigenlijk? Het bestaat al heel lang. Op de volgende slide ziet u een voorbeeld waarbij een brein is afgebeeld. Vanuit het menselijke brein zijn er twee paden ontwikkeld, al vanaf de jaren zestig en zelfs wat eerder. Het ene is gebaseerd op de boekenkennis, dus de kennis die wijzelf in een vorm hebben gegoten waarin die voor ons leesbaar en toegankelijk is. Dat is de zogenaamde kennistechnologische benadering. De andere benadering is geweest dat we kijken naar wat ons brein nou zo bijzonder maakt, hoe dat in elkaar zit, of we dat kunnen namaken, kunnen simuleren en of we dan ook intelligentie krijgen, maar dan in een kunstmatige vorm. Dat blijkt inderdaad zo te zijn. Als u

het pad daarboven bekijkt en inzoomt, dan ziet u een enorm netwerk van neuronen, knopen in onze hersens, die een ontzettend hoge graad van verbinding hebben met andere knopen in ons brein. Dat kun je namaken. Aan de rechterkant ziet u zo'n gestileerd voorbeeld. Het blijkt dat die kunstmatige neurale netwerken inderdaad het vermogen hebben om associaties te leren. Dat is iets waar ons brein fantastisch goed in is, terwijl het formuleren van kennis voor ons brein best lastig is, tenminste in een vorm dat we het in boeken kunnen opschrijven.
Als we het zo in die boeken opschrijven en daar regels uit trekken, dan komen we aan de rechterkant bij de kennisgebaseerde systemen. In de jaren zestig waren de computers niet krachtig en waren er weinig data beschikbaar. Onze hele maatschappij was nog niet gedigitaliseerd. Op dat moment was het gewoon onmogelijk om te komen tot kunstmatige neurale netwerken die sterk genoeg waren om complexere vraagstukken te leren begrijpen, terwijl we heel snel aan de slag konden gaan met die kennisgebaseerde systemen. In die tijd ging de ontwikkeling van de kennisgebaseerde systemen veel harder dan die van de kunstmatige neurale netwerken. Het heeft jaren geduurd voordat we zover waren dat we een doorbraak konden bereiken in de kunstmatige intelligentie. Nu ik toch in jullie midden ben, wil gelijk een pleidooi houden voor fundamenteel onderzoek; je weet vaak niet waar kennis op den duur toe leidt.
We gaan naar de volgende slides. Ik weeg die twee nog even tegen elkaar af. Aan de ene kant is het voordeel van de kunstmatige neurale netwerken dat we ze met gewone computers kunnen simuleren. We konden ze al maken voordat we speciale hardware daarvoor gingen maken. Ze zijn adaptief. Dat is natuurlijk het karakteristiek waaraan we het meest hebben. Ze kunnen nieuwe associaties maken. Daarmee hebben ze het vermogen om continu bij te leren. Aan de andere kant waren de computers in de jaren zestig niet krachtig genoeg om er echt voortgang mee te bereiken. Het laatste stuk is nog steeds een issue: het is toch wel heel erg lastig om ze te ontwerpen. Dat vraagt op zich een eigen discipline, een eigen kunde.
Dan kijken we naar de kennisgebaseerde kunstmatige intelligentie. Het voordeel was dat we ter plekke wisten hoe we verder moesten. We hadden de computers en we konden hiermee aan de slag. Ze waren effectief, ze zijn efficiënt en een groot voordeel is: je weet wat ze weten, want we hebben er precies dat zelf in gestopt. Ze zijn controleerbaar en stuurbaar. Het vervelende is alleen -- dan komen we bij de tegenargumenten -- dat het uitvragen van experts duur is en langzaam. Ik zal er een voorbeeldje van geven dat zich al in de jaren negentig, eind jaren tachtig, voordeed.
Twee teams hadden de taak gekregen om te leren een vliegtuigje te besturen. Het ene ging aan de slag met een aantal piloten en experts om deze mensen uit te vragen en de kennis op te schrijven, dus via de onderste weg. Het andere koppelde de besturing van het

vliegtuigje aan een computer en zette daar de piloot gewoon achter: gaat u maar vliegen. Drie weken waren 30 mensen bezig met het eerste systeem, het kennisgebaseerde systeem, en het ding vloog nog niet. Het andere vloog binnen een dag. Dat zijn natuurlijk de voordelen waar we vandaag de dag onze vruchten van plukken, maar ook onze lasten van hebben. Dus je weet wat ze weten, maar het uitvragen van experts is duur, het kost heel veel tijd, het onderhoud is dan ook arbeidsintensief en het vervelende is dat ze niet per se adaptief zijn. De doorbraak van machinelearning hebt u allemaal meegemaakt in uw leven. U kent de verschillende spelen waarin dit ineens naar voren kwam met als laatste het winnen van het spel Go, wat geleid heeft tot de ontwikkeling van deep learning, de diepe neurale netwerken die we nu zien. Dat kunt u rustig doorlezen. Ik ga door naar de volgende slide.

Gaan we met behulp van die diepe neurale netwerken aan de slag, dan hebben we de kracht van het snel kunnen leren als we voldoende correcte data hebben. Die "als" is wel een hele grote en belangrijke "als". Dan kan je snel fantastische resultaten bereiken. Maar de tegenvoorbeelden, de voorbeelden waarin het fout gaat, ziet u op de volgende slide. Dit is een plaatje van een panda. Als je er een kleine ruisfactor overheen legt, is het voor ons nog steeds een panda, maar wordt het door het algoritme ineens geclassificeerd als een gibbon. Hetzelfde gebeurt met een spraaksignaal daarboven dat er toch behoorlijk hetzelfde uitziet links en rechts, maar ineens van "how are you" "open the door" wordt. Het voorbeeld rechtsboven is van het stopbord: een paar plakkertjes erop en helaas, de zelfrijdende auto's zijn in de war en negeren het stopbord. Nou ja, ze negeren het niet; ze zien het niet. Daaronder ziet u nog het bekende voorbeeld van Amazon. Die probeerden hun selectieproces van nieuwe medewerkers te automatiseren en eigenlijk ook te verbeteren. En wat bleek: een onderselectie van minderheden bij sollicitaties, in feite op basis van historische onderselectie. Rechtsonder is het voorbeeld dat wij in computer science noemen "bullshit in, bullshit out", excusez le mot: onzin erin, dan komt er ook onzin uit. Dat is precies wat daar aan de hand was. Maar als je er goede data in stopt, je blijft bevestigen wat het algoritme doet en je de data die je op die manier gekozen hebt terug in het algoritme voert, blijft het algoritme leren. Het kan zich daarmee op den duur toch op andere paden gaan begeven dan je eerst bedoeld hebt. Om er een zeilterm bij te gebruiken: het algoritme kan verlijeren.

Ik ga naar de volgende slide om het laatste punt nog even duidelijk te maken. Ik heb met het Rathenau wat discussie gehad over de vraag of we deze slide moesten laten staan. Ik zei dat ik de slide belangrijk vond, en wel om het volgende: ik heb u een paar minuten geleden verteld dat het zo lastig is om uit ons brein op te halen welke kennis erin zit. Er is geen knop waarop u kunt drukken, waarna onze kennis eruitrolt. Onze kennis extraheren gaat alleen maar door heel intensief met de mens om te gaan, door voorbeelden voor te leggen, uit te

vragen, de mens er zelf over na te laten denken, nieuwe associaties te laten maken en zo de kennis van die persoon te eliciteren.
Dat is ook het probleem van de neurale netwerken. Bij neurale netwerken is het heel erg ingewikkeld om te achterhalen waar kennis is opgeslagen en welke kennis dat dan eigenlijk is. Dat is hier het voorbeeld van. De kennis over zo'n netwerk van honderdduizenden knopen zit gedistribueerd opgeslagen. Je kunt niet een knoop aanwijzen die zegt: dit is het stukje waarin de kennis zit die hier leidt tot deze classificatie. Sterker nog, het is ook een voordeel. Dat geldt voor ons brein ook. Als je een kleine hersenbeschadiging oploopt, dan kunnen je hersens dat eigenlijk nog heel goed opvangen. Zo geldt het in zo'n neuraal netwerk ook. Het gedistribueerd opslaan van kennis heeft ook z'n voordelen, maar het heeft ook dat nadeel van "hoe weet je wat erin staat?". Hoe weet je op grond waarvan dat netwerk beslist? Dat weten we dus niet.
We gaan naar de volgende slide. Daarin is nog even voor u opgesomd welke AI-systemen er in grote lijnen zijn. Je ziet aan de linkerkant de kennisgebaseerde systemen en aan de rechterkant de zelflerende systemen. Binnen die zelflerende systemen kun je weer klassiek leren en diep leren hebben. Daar kom ik straks nog wel op terug. Dat diep leren hebben we in feite net behoorlijk besproken.
We gaan naar de volgende slide. Als we het hebben over zelflerende systemen, dan zijn dat systemen die verbanden in grote hoeveelheden data ontdekken en op basis daarvan een kans berekenen dat er een bepaald patroon in zit of dat het systeem, of het bord, waar het naar kijkt, binnen een bepaalde klasse valt. Voorbeelden van vragen die wij aan die systemen voorleggen zijn de volgende. Wat is de kans dat iemand geschikt is voor een baan vanuit het Amazonvoorbeeld? Wat is de kans dat iemand borstkanker heeft? Wat is de kans dat iemand gefraudeerd heeft? Et cetera. Die vragen zijn in grote lijnen gebaseerd op kansberekening. Ouderwetse statistiek is enorm belangrijk voor het ontwikkelen van kunstmatige intelligentie.
Als je kijkt naar het klassieke voorbeeld van machinelearning, dan bepalen wij als mensen van tevoren welke aspecten, welke features zoals dat in ons vakgebied genoemd wordt, belangrijk zijn. Als een systeem bijvoorbeeld zelf moet leren om geschreven taal te herkennen, dan zijn er elementen waarvan mensen hebben gezegd dat je zou kunnen kijken naar hoeveel kromming er in dat symbool zit waar je naar kijkt, hoeveel rechte lijnen daarin zitten, en wat dan de kans is dat het een 0 of dat het een 1 is. Op die manier leer je die systemen, met behulp van je eigen idee over hoe je naar dat soort symbolen kan kijken, wat de aspecten zijn, wat de symbolen zijn, die ze moeten leren. Daar is natuurlijk ook heel veel gebruik van gemaakt. Het voordeel daarvan is dat die enigszins stuurbaar en controleerbaarder zijn, want er staat toch meer op kennisniveau geschreven kennis in.

We gaan naar de volgende slide. Als we dat alles afzetten tegen deep learning, dan gaat het er dus om dat de machine bij deep learning zelf bepaalt welke aspecten worden meegenomen, welke features van belang zijn. Zo zie je hier een voorbeeld staan van gezichtsherkenning voor risicoclassificaties. Welke features worden er dan gebruikt? Is dat de huidskleur, de bril, haardracht, een hoofddoek, of is het de achtergrond? Als je iemand bijvoorbeeld altijd voor een blauw scherm zet als ze iets gedaan hebben, dan krijg je er ook een andere uitkomst uit.
Het vervelende van die deep-learningmachines is dat ze zowel moeilijk stuurbaar als niet uitlegbaar zijn. Ik hoop u dat net aan de hand van het voorbeeld van dat ingewikkelde schema uitgelegd te hebben. We hebben geen idee waarom die systemen zeggen wat ze zeggen, maar we weten wel dat we er preciezere of betere successcores mee halen dan met die kennisgebaseerde, zelflerende systemen. We hebben dus een beetje het probleem dat voor ons het formuleren van kennis erg lastig is. Als je dat dan weer op dat voorbeeld terugzet, zie je dat als je naar die verschillende AI-systemen kijkt, niet elke vorm van AI even zorgelijk is qua uitlegbaarheid en stuurbaarheid, maar dat je daar dan wel weer op inlevert qua prestaties. Dat is ook een belangrijk aspect. De kennisgebaseerde systemen zijn dus stuurbaar, controleerbaar en uitlegbaar, en de deep-learningmachines zijn moeilijk stuurbaar, moeilijk controleerbaar en grof gezegd ook niet uitlegbaar.
Ik hou het tempo erin. Als ik te snel ga, dan mag u mij interrumperen. Dat ben ik bij colleges ook gewend, dus aarzelt u niet.
Ik ga verder met datavraagstuk, met het voorbeeld dat ik net aanhaalde over die classificatie. Rechtsboven ziet u een foto van een hond die wel wat wolfachtige kenmerken heeft. Het vraagstuk ging over een systeem dat getraind werd in het herkennen van wolven. Deze husky werd daarbij geclassificeerd als wolf. Dat is niet zo heel gek, want er zijn zat mensen die dat ook zouden doen. Je zou daar dus bij kunnen zeggen: nou, dat snap ik eigenlijk wel; het lijkt er best op, dus ik laat het erbij zitten. Maar het loont toch echt om door te graven waarop het nou misging. In dit geval bleek het helemaal niet te gaan over het plaatje dat erop stond, maar over de hoeveelheid sneeuw die achter het dier zichtbaar was. Die hoeveelheid sneeuw maakte dat de hond geclassificeerd werd als wolf. Dat zijn de pijnlijke voorbeelden op grond waarvan je je bewust wordt van het punt dat als je geen idee hebt wat het algoritme doet, je mogelijk te maken krijgt met beslissingen van zo'n systeem waar je echt niet blij van wordt.
Dus als we kijken naar het datavraagstuk, dan moet je gewoon heel simpel vaststellen dat het machinelearning algoritme de systematische fouten die in de trainingsdata zitten reproduceert. Is het nou een fout om bij zo'n plaatje sneeuw op de achtergrond te hebben? Dat zien wij toch niet? Daar denk je als mens toch niet aan? Dus hoe weet je nou wat goede

data zijn? De enige manier waarop je daar uiteindelijk achter kan komen, is door dit soort experimenten tot in den treure uit te voeren, uit te werken en te blijven analyseren.
Dan is er nog het aspect dat niet-relevante maar wel gecorreleerde patronen onterecht betekenis krijgen. Dat is zeker het geval als we het hebben over discriminatie. En zeker bij overheidssystemen is dat natuurlijk van groot belang.
Nou weet ik vanuit mijn eigen leven als wetenschapper dat je als vrouw toch vaak een nadeel hebt ten opzichte van je mannelijke collega's. Het is herhaaldelijk vastgesteld dat papers geschreven door bijvoorbeeld Marie Jonker -- om maar even mijn eigen naam een beetje te verbasteren -- toch een aanzienlijk lagere kans hebben om geaccepteerd te worden dan de artikelen van mijn collega Pieter Jonker. Pieter Jonker bestaat trouwens echt. Is dat eerlijk? Dan kunnen we zeggen: dat geeft niet, dan mag je alleen maar je voorletters gebruiken; dan doen we dat. Dat werkt niet. Het blijkt dat de mannelijke collega's rustig de naam Pieter Jonker blijven gebruiken en ik dan fijn met mijn C.M. Jonker aan kom zetten. Nou, dat heeft dat algoritme helaas zo door. En dat algoritme is ons algoritme in ons brein. Daar kunnen we niets aan doen. We weten het niet eens. Mijn collega's zijn echt niet van plan om te discrimineren, maar dat zit gewoon in het onderbewuste. Dus: hoe wil je dat aanpakken? Lastig. Dan kun je ook nog zeggen: nou, dan halen we hele namen weg. Maar ja, als ik één keer over een bepaald onderwerp heb geschreven, is de kans groot dat ik weer over dat onderwerp schrijf. De kans dat dat paper van mij is, is dan relatief groot.
Zo zit het ook met de data over persoonsgegevens. Het is niet voldoende om een voornaam of een postcode weg te laten. Als je wilt voorkomen dat iemand direct identificeerbaar is als behorend tot een bepaalde categorie, dan moet je naar zo veel factoren kijken. Het is echt mind-boggling.
Dus waar staan we nu? Laten we maar naar de volgende slide gaan, want u heeft nog meer te doen vandaag: de AI anno nu. We hebben kunstmatige neurale netwerken. Die zijn mogelijk geworden door de toegenomen rekenkracht van computers. Maar we hebben het fenomeen dat bekendstond onder de term "black box AI": we weten niet wat erin zit. En we hebben de kennistechnologische AI, die haar waarde in de loop van de decennia bewezen heeft, maar ook haar beperkingen. Langzamerhand zijn we bewust van de kracht van die black box AI en beginnen we ons ook bewust te worden van de beperkingen.
Dus hoe kunnen we nu verder? We gaan naar de volgende slide. Ik en mijn collega's zijn op weg naar uitlegbare machinelearning, dus naar kunstmatige neurale netwerken op een uitlegbare manier. Een van de manieren waarop wij dat nu proberen te doen, en ook anderen, is om te kijken of we een leesbaar model uit die black box kunnen trekken. Want laten we wel wezen: dat black-boxalgoritme is zeer krachtig; het doet het echt beter, ook op die classificatietaken, dan wanneer we puur met kennisgeoriënteerde systemen aan de slag gaan. Maar ik wil wel weten waar die classificatie op gebaseerd is, dus dan kan ik

machinelearning gebruiken, en dan de kennisgebaseerde machinelearning, om er een kennismodel uit te trekken, iets wat ik wel kan lezen en waardoor ik weet op welke regels en op welk niveau die machine beslissingen neemt.
We gaan naar de volgende slide. Daarvoor hebben we een experimentje opgezet. Spelletjes zijn voor kunstmatige intelligentie zeer belangrijk, zoals u misschien aan dat schaakvoorbeeld en het go-voorbeeld al gezien had, omdat je het eindeloos vaak kan herhalen. Tja, dit is ook weer zo'n voorbeeldje: "geode strategie". Dit heb ik niet getikt, hoor. Ik heb echt "goede" getikt, maar de kunstmatige intelligentie maakt daarvan dan, omdat ik meestal Engels schrijf, "geode" van. Goed, interessant. We nemen dus zo'n spelletje, we laten er een black-boxmachinelearningalgoritme op los, en dat leert dan hoe het dat spelletje het beste kan spelen. In dit geval moet het rode muntjes verzamelen en de blauwe en groene zien te vermijden. Daar heb je natuurlijk data voor nodig. Nou, dat is het leuke van spelletjes: die kun je eindeloos genereren om te gaan spelen.
Kunnen we dan -- zie de volgende slide -- daaruit leren wat dat algoritme weet? Wat we doen, is daar een algoritme op zetten dat in drie fasen leert. In de eerste fase leert het de hoofdlijnen van de strategie. Daar staan een paar van die regeltjes onder. Grof gezegd gaat het over in welke gevallen je moet springen, de jump, en dat je dan naar rechts moet gaan; dat zijn zo'n beetje de aspecten waar het hier over gaat. En daar staan herkenbare dingen in: er staan bricks in de weg, en ik zie een rood muntje enzovoort. Dat kan ik lezen, terwijl het netwerk dat daarboven afgebeeld staat, dat nu natuurlijk maar een symbolische representatie is, voor ons absoluut niet te lezen is. Ik heb dan in hoofdlijnen een idee van wat het algoritme doet, maar is dat dan ook wat het algoritme doet? Hoe goed is die samenvatting dan? Zo komen we tot een vorm van hybride intelligentie: we koppelen menselijke intelligentie aan kunstmatige intelligentie, om op een interactieve manier die kennis te verfijnen. We gaan samen met de machine op zoek naar uitzonderingen op de samenvatting: in welke omstandigheden moet het dan toch net anders? En daar kun je interactief heel lang op doorvragen, natuurlijk. Dat kun je door de mens laten doen, maar daar kun je natuurlijk ook wel weer intelligentie op zetten. Ik kan wel degelijk intelligentie trainen om naar de uitzonderingen te gaan zoeken. Zo gebruik ik AI om vat te krijgen op AI.
We komen bij de voorlaatste slide. Zo komen we dus tot een situatie waarin het mogelijk wordt dat mensen samen met AI naar AI kijken, en over AI kunnen beslissen en het sturen. Dus de kennisgebaseerde AI bouw ik boven op de black box AI, en op grond van de kennisgebaseerde AI kan ik een interactie met de mens aangaan. Dus die geeft de classificatie door van het black-boxalgoritme, en als de mens zegt "waarom denk je dat ik dit zou moeten doen?", dan kan er een uitleg komen op basis van die kennisgebaseerde AI. Het kan dan zijn dat die mens zegt: nee, niet in dit geval; dat is niet correct. Nou, dan kan het algoritme gelijk terug gaan leren en vragen: waarom dan? Als de mens daar dan een

uitleg over geeft, dan kan de machine dat nog een keer samenvatten en op grond daarvan nieuwe data genereren, waarmee het black-boxalgoritme weer wordt aangestuurd en gaat bijleren. Er wordt nu gevraagd: is het dan nog een black-boxalgoritme? Ik zou zeggen: ik heb een raam gemaakt in de black box, en daar komt licht in.
Goed, de takeaway message. Door die manier van hybride intelligentie kun je combinaties maken van de kunstmatige neurale netwerken, die machinelearningalgoritmes, de kennisrepresentatie en de menselijke intelligentie. En daarmee komen we tot een punt dat ook centraal ligt onder het Hybrid Intelligence Centre, namelijk eentje waarin kunstmatige intelligentie een intelligentie is die samenwerkt met mensen, en niet gericht is op het vervangen van mensen. Door AI over AI te zetten, kunnen we de black box openmaken, er een raam in zetten, en door de mens daarmee samen te laten werken, kunnen we het bestuurbaar en controleerbaar gaan maken. Maar daar zijn we nog niet.
Voor uw verdere interesse kunt u terecht op de websites die hieronder staan, of natuurlijk bij een van onze onderzoekers.
Dank u wel.

De **voorzitter**: Dank ú wel, professor Jonker. Er is nu ruimte voor vragen over de presentatie, voordat we mevrouw Rieback het woord geven. En dat zijn dan ook weer vragen aan u, maar het is misschien ook nog mogelijk dat u onderling opmerkingen of aanmerkingen heeft. Ik kijk nu naar de aanwezigen voor het stellen van vragen. Dat moet bij de interruptiemicrofoon, anders kunnen de mensen thuis niet meeluisteren.

De heer **Nicolaï** (PvdD): Dank u wel. Ik probeer het te begrijpen. Ik zag de eerste sheets en dacht "daar begrijp ik niks van", maar na het verhaal begreep ik het een stuk beter. Er is een ding dat ik nog steeds niet begrijp, en dat is de sheet waarin u het heeft over de machine die zelf bepaalt welke aspecten belangrijk zijn. Want daar zit volgens mij de kern van het hele vraagstuk. Ik kan me nog voorstellen dat je dingen ergens in het systeem stopt. Dan ben je ook verantwoordelijk voor welke kenmerken je erin gestopt hebt, en kan je vervolgens, als er dingen uit komen waar je vragen over hebt, denken: heb ik er wel de goede kenmerken in gestopt? Het gevaar ontstaat voor mijn gevoel op het moment dat zo'n machine zelf gaat bepalen welke kenmerken bepalend zijn. Hoe moet ik het zien dat de machine dat zelf kan bepalen?

Mevrouw **Jonker**: Ja, dat is een uitstekende vraag. Het punt is natuurlijk, zoals ik in het begin al zei: het ontwerpen van die diepe neurale netwerken is een lastige zaak; om dat goed in elkaar te zetten, heb je echt kennis van zaken nodig. Maar kennis van zaken is niet genoeg; het is ook een kwestie van intelligent proberen, zoeken, uitproberen en kijken hoe

het algoritme zich ontwikkelt, dan weer op stop kunnen drukken en een ander feature aangeven. Mensen zijn natuurlijk op de manier begonnen, door zelf die features, die aspecten, aan te geven, en op die manier het algoritme te trainen en daar bepaalde successen mee te halen. Maar ja, als je er een paar maanden mee bezig bent geweest om dat algoritme beter te krijgen en je hebt dus honderd keer een andere feature bedacht en je bent weer aan het trainen geslagen, dan ga je je als mens natuurlijk op een gegeven moment afvragen: kan dat niet slimmer? En aan een kunstmatige-intelligentieonderzoeker is dat een heel logische vraag om te stellen. Dit zijn dan ook typisch de mensen die dan denken: oké, ik schrijf daar een ander algoritme op, dat zorgt dat het zelf alle mogelijke features gaat exploreren. Dat betekent dat na enig zoeken en denken die architectuur dusdanig werkt, dat de machines inderdaad zelf gingen bepalen wat ze deden.
Nou, wat doe je dan in feite? Je traint ze door aan te geven: dit was een correcte classificatie of geen correcte classificatie. Dat is, zeg maar, de eerste tussenstap; het gaat nog verder. En dan kun je een reinforcement-backpropagationalgoritme gebruiken om tegen het algoritme te zeggen: daar zat je fout, dus je moet de manieren waarop je jezelf veranderd hebt om deze conclusie te bereiken, achteruitdraaien. En als je een goede conclusie hebt getrokken, dan komt daar een zelfversterkend element in, dus dan worden die verbindingen in die neurale netwerken verstevigd. Maar ja, die manier van leren gaat best aardig, maar we willen natuurlijk eigenlijk ook weten wat er in grote hoeveelheden data zit waarvan we niet eens weten of er patronen in zitten. Van daaruit is de volgende stap gemaakt door te zeggen: kijk maar of je een patroon kunt vinden; elk patroon dat jij kunt vinden, vinden wij interessant. Dat is de meest vrije vorm van machinelearning. En ja, dan weet je niet wat ie doet, maar aan de andere kant weten wij als mensen dat eigenlijk ook niet. Ook de menselijke systemen -- die heeft u zien staan in de voorbeelden -- waarbij de mensen de features bepaald hebben, maken fouten. En waar ligt dat nu aan? Ligt dat aan de data die we ingevoerd hebben, ligt dat aan de features die wij daar als ontwerpers hebben neergezet, of ligt dat aan de manier waarop het algoritme ingezet wordt? Dat zijn natuurlijk allemaal aspecten waar we met elkaar naar moeten kijken, om die te verbeteren.

Mevrouw **Kluit** (GroenLinks): Ik borduur hier een beetje op voort. Ik werk zelf in de mobiliteitswereld. Daar wordt al langer gewerkt met algoritmes en artificial intelligence, met geluk en minder geluk. De data veranderen, de omvang verandert ook, de manier waarop we omgaan met de data in artificial intelligence verandert. Dus dat zijn allemaal moving targets. En als je weet hoeveel alleen al één auto aan informatie de datawolk in stuurt, dan vraag ik me weleens af: als we dat gele blokje nou willen laten meedenken met die black box, wat heb je dan nodig aan uitvoeringskracht om dat goed te kunnen doen, dus om die patronen te herkennen, maar ook om snel door te hebben dat er iets in die dataontwikkeling

anders gaat? En voor degenen die het niet weten: een auto stuurt elke seconde tienduizenden informatiepunten naar leveranciers toe; dat is echt enorm.

Mevrouw **Jonker**: Dat is een hele goede vraag. Er zitten twee stukken aan uw vraag. Het eerste stuk betreft de ontwikkeling van de goed functionerende yellow box: het kennissysteem. Ook die zal uit moeten vragen aan het deepmachinelearningssysteem om die kennis te achterhalen. Dat uitvragen kost tijd en dus rekenkracht. Een ruwe schatting is dat het er nog een keer bovenop komt. We zullen ofwel de benodigde tijd om zo'n algoritme te ontwikkelen moeten verdubbelen ofwel de machine krachtiger moeten maken om dat in dezelfde tijd uit te rekenen. Het energiegebruik gaat dus ook nog steeds omhoog. De bijbehorende opwarmingsfactor van de aarde speelt dus wel degelijk mee. Het andere stuk is het moment waarop het systeem staat. Op dat moment is de efficiëntie van de algoritmes prima, zowel van het gele boxje als van het black-boxalgoritme. Dus daar zit de pijn niet meer, maar wel in de ontwikkeling.

Mevrouw **Kluit** (GroenLinks): Misschien moet ik de vraag iets anders stellen. Mijn vraag is meer gericht op de menskant of op de beheersbare, controleerbare kant: hoe zorg je ervoor dat er aan de publieke kant voldoende kennis en instanties zijn die kunnen meedenken of kunnen toetsen in de black boxen, zodat we ongelukken voorkomen? Het is waar dat mensen ook fouten maken, maar mensen zijn voor mensen wel makkelijker te volgen dan een black box waarin er elke seconde zoveel data bij komt.

Mevrouw **Jonker**: Ja, dat is een hele goede vraag. De vraag had ik niet goed begrepen. Het Design for Values Institute richt zich precies op die vraag. Als je weet dat je zo'n soort systeem zou willen maken, wil je dat vanaf dag één ontwerpen op een manier dat het dat makkelijker maakt. Het ontwerp voor waarden en het ontwerp voor betrouwbaarheid en voor transparantie moet je vanaf het eerste moment meenemen. Dat is één. Ten tweede moet je zorgen voor voldoende experts om mee te blijven kijken in de teams die zo'n algoritme toepassen.

Mevrouw **Kluit** (GroenLinks): Stel dat we het grijze gebied tussen Facebook -- dat is een actueel thema -- en de overheid een beetje willen managen en monitoren of dat we daar zelfs op kunnen ingrijpen. Hoeveel mensen in Nederland, om het daartoe te beperken, moeten wij dan hierop zetten? In de kabinetsonderhandelingen wordt gesproken over miljarden. Ik kan me voorstellen dat investeringen hierin broodnodig en heel handig zouden zijn, maar waar hebben we het dan over?

Mevrouw **Jonker**: Miljarden, niet alleen van ons, maar ook van onze buurlanden. We moeten samen hierin optrekken. Wat dat betreft is de EU toch een heel mooi samenwerkingsverband waarin sociale normen en waarden en de rechten van het publiek en het individu in een grotere balans staan dan in sommige andere werelddelen. Dit betekent dat het onderzoek naar het verantwoord ontwerpen van kunstmatige-intelligentiesystemen echt prioriteit moet krijgen. Dat gaat heel veel tijd en energie kosten.

De **voorzitter**: Dank u. Het is nu 9.45 uur. Ik wil om 9.50 uur verder met de volgende inleiding. Meneer Crone.

De heer **Crone** (PvdA): Ik hoorde toevallig gisteren op een congres dat heel veel wetenschappers geen geld meer krijgen voor onderzoek, tenzij ze het onder het labeltje AI plakken. "AI" wordt een trefwoord om geld uit de potten te krijgen.

Mevrouw **Jonker**: Ja, dat klopt.

De heer **Crone** (PvdA): In dit huis geldt gelukkig de traditie dat wij domme vragen mogen stellen, maar dat ministers naar huis moeten als ze domme antwoorden geven. Zover bent u gelukkig nog niet.

Mevrouw **Jonker**: Misschien moet ik dan ook wel naar huis!

De heer **Crone** (PvdA): Ik wilde vragen naar een heel ander aspect van deep learning. Ik hoorde dat er nu al zo veel rekenkracht mogelijk is dat als je een verslag wil hebben van een voetbalwedstrijd, bijvoorbeeld Ajax tegen Cambuur, je tijdens de wedstrijd al makkelijk via alle netwerken kunt bijhouden hoe laat er een doelpunt viel, wie het heeft gemaakt enzovoorts. Eén minuut na de wedstrijd kan er al een verslag zijn, dat zelfs in de stijl van Mart Smeets geschreven kan zijn.

Mevrouw **Jonker**: Ja, dat klopt.

De heer **Crone** (PvdA): De machine kan zowel de stijl van Mart Smeets leren als hoe de wedstrijd is afgelopen. Maar weet ik dan thuis nog of het artikel van Mart Smeets of van een ander is? Dat geldt natuurlijk ook voor overheidsdocumenten. Kan iemand een overheidsdocument ontvangen dat niet door de overheid is geschreven, maar door een ander? Gaat iemand dus foute dingen doen? Kunnen we daar een keurmerk voor

verzinnen? Het gaat mij nu niet om een oplossing, maar dat we iets doen. Kan dit inderdaad al? Ik hoor dat dit vrij simpel is.

Mevrouw **Jonker**: Ja, dit kan en behoort tot de huidige stand van zaken. De eerste die serieus een automatische samenvatting maakte, was een jongen van een jaar of 12. Hij was binnen de kortste keren miljonair.

De heer **Crone** (PvdA): Wordt er aan keurmerken gedacht? Nu krijg je al mailtjes van banken, maar in de toekomst denk je: ik open er nooit meer een. Dat gaan we natuurlijk niet doen. Wordt er gedacht aan keurmerken? Dit komt misschien later nog terug, want we hebben nog veel hoorzittingen.

Mevrouw **Jonker**: In onze technologie speelt natuurlijk navenant het hele punt van blockchaining een rol. Hopelijk komt er op den duur ook quantum computing, waarmee we veel meer grip gaan krijgen op dit soort fenomenen. Maar certificering heeft zeker onze aandacht. Als je het niet structureel aanpakt, blijft het een wapenwedloop tussen de boeven en goeieriken. Zeker.

De **voorzitter**: Dank u wel. Nog een hele korte vraag van mij. Wanneer zou het venster in die black box waar u naar op zoek bent, operationeel worden? Want de black box is er al en gaat niet meer weg. Wanneer zouden we daar een venster in kunnen hebben dat ook toepasbaar is?

Mevrouw **Jonker**: Dat is altijd moeilijk te zeggen. Het is eigenlijk zoals we altijd met kunstmatige intelligentie hebben: als het voor één specifiek systeem is, zullen we het relatief snel kunnen, maar het wordt lastig als we het in algemene zin moeten doen. De veralgemeniseerbaarheid van een oplossing naar meerdere problemen is moeilijk. Voor een specifiek probleem zal het vrij snel zijn. Ik denk dat we er ook rustig op moeten blijven inzetten om zowel onderzoek te doen naar wat het in directe zin voor een bepaalde situatie mogelijk maakt als fundamenteler onderzoek te doen naar wat het in algemenere zin mogelijk maakt.

De **voorzitter**: Dank u.

Presentatie 2: Eisen aan AI

Gesprek met:

- mevrouw Melanie Rieback (CEO/medeoprichter van Radically Open Security)

De **voorzitter**: Dan kom ik bij u, mevrouw Rieback. Ik wil u vragen om uw presentatie te geven.

Mevrouw **Rieback**: Goedemorgen iedereen. Ik ben Melanie Rieback. Ik ben CEO en medeoprichter van Radically Open Security, een sociale onderneming op het gebied van computerbeveiliging. Trouwens, jullie horen het vast: ik ben Amerikaanse van afkomst, maar ik woon wel al twintig jaar in Nederland. Ik ben begonnen als academicus en ben voormalig universitair docent informatica bij de Vrije Universiteit. Ik heb een jaar of zeven gewerkt aan security en privacy, met name van "radio frequency identification"-technologie. Tussen ongeveer 2006 en 2010 was ik bezig met onderwerpen als privacy van de ov-chipkaart. Voor jullie is dat waarschijnlijk ook a blast from the past. Ik werk ook voor de faculteiten van Singularity University, zowel in Silicon Valley als in Nederland. Mijn bedrijf doet eigenlijk iets heel bijzonders. Wij doneren 90% van onze winst aan het goede doel. De laatste 10% is onze cashflowbuffer. Zo houd ik een goedlopend bedrijf.
We zijn een consultancy van ongeveer 40 mensen en hebben meer dan 150 klanten gehad, inclusief de overheid. Bijvoorbeeld het ministerie van VWS heeft ons ingehuurd om een pentest uit te voeren op de CoronaMelder. Jullie zijn daar vast mee bekend. Volgens mij hebben jullie mijn bedrijf in die ene sessie 22 keer genoemd. Ook hebben we werk gedaan voor de Europese Commissie. We hebben pentesten uitgevoerd op andere contact tracing apps in het kader van COVID-19, zoals Immuni in Italië en ProteGO in Polen. Dat hebben we ook gedaan voor de European Framework Gateway Services, dus het EU Interoperability Framework. Daarnaast hebben dat gedaan voor de Google/Apple Exposure Notification API. Verder hebben we laatst een security audit losgelaten op de digitale vaccinatiepaspoorten van de EU. We zijn dus bezig met nogal wat hoogprofielopdrachten. Wij werken met Google, met Mozilla, met het Open Tech Fund in de Verenigde Staten en met Wikimedia Foundation. In Nederland werken we voor energiebedrijven als TenneT, Eneco en Stedin. We doen daar pentesten, van power transformers tot websites. Voor de rest doen we ook veel ander werk, van internet exchanges in de core internet infrastructure tot en met non-profitorganisaties waarvoor wij werken op een non-profitbasis, tegen kostprijs. Dat doen we voor non-profit ngo's en civil society.
In de eerste zes jaar van mijn bedrijf hebben we meer dan een half miljoen gedoneerd aan de stichting NLnet, die open source en digitaleburgerrechtenorganisaties ondersteunt, zoals de Bits of Freedoms en IFF's van deze wereld, en technologie zoals Tor, WireGuard, Jitsi en DNSSEC, eigenlijk alles voor een beter open internet. Ik heb mijn best gedaan om een bedrijf te maken dat geoptimaliseerd is voor sociale impact. Dat komt natuurlijk ook vanuit

mijn achtergrond als academicus. Ik probeer erachter te komen hoe wij business kunnen gebruiken als een vorm van positief activisme. Ik heb veel prijzen gewonnen voor mijn werk. CIO Magazine heeft mij de meest innovatieve IT-leider van Nederland genoemd. Ik was ook een van de finalisten voor de EU Prize for Women Innovators. De KvK heeft Radically Open Security het 50ste meest innovatieve mkb-bedrijf van Nederland genoemd. Dat is allemaal validatie voor onze manier van bedrijfsvoering, die toch een beetje vreemd is. Ik denk dat ik daardoor ook een ander perspectief heb en een iets ander verhaal kan vertellen.
We gaan nu naar de volgende slide. Ik hoef jullie niet te vertellen dat onze levens steeds digitaler worden. Dat komt ook nog in een stroomversnelling terecht door covid. We deden al veel online, van winkelen en boodschappen doen tot een date vinden, en natuurlijk e-mailen en sociale media. Maar nu komt met het work from home natuurlijk ook nog bijna alle bedrijfsvoering online. Dat verhoogt natuurlijk wel de stakes, met de beveiliging en de privacy van onze data. Governance blijft ook wel een puntje. Internet of things is natuurlijk ook wel langer aan de gang. Maar ook met AI -- en daar gaat het gesprek van vandaag over -- komt het een beetje in de knel.
De overheid wordt ook digitaal. Dat gebeurt natuurlijk ook al langer. De overheid gebruikt ook AI, voor alles van verkeersboetes tot fraudeopsporing. Het is natuurlijk geen toeval dat ik hier op de slide een foto van een toeslagenenvelop van de Belastingdienst heb. Ik hoef jullie daar ook niks over te vertellen. AI wordt ook gebruikt bij risicobeoordeling van ex-gevangenen of huidige gevangenen, voor de vraag of zij losgelaten mogen worden of met parole mogen. Denk ook aan de controle van immigratie: wie laten we toe en wie niet? Bij militaire wapens, inclusief drones, worden beslissingen die gemaakt zijn door AI echt een kwestie van leven en dood. AI wordt ook gebruikt voor voorspellend politiewerk. Sommige van deze punten zijn best controversieel, maar goed.
De volgende slide gaat over de eerste dreiging: gebrek aan transparantie. Daar heeft Catholijn het ook over gehad. AI is vaak, niet altijd, maar wel soms een black box, een volledige black box. Dat is natuurlijk ook wel een probleem. Catholijn heeft het natuurlijk ook wel gehad over dat voorbeeld van Amazon met de hiring criteria, die uiteindelijk alleen maar bestaande vooroordelen hebben ondersteund. Inderdaad, garbage in, garbage out. De makers van de systemen hebben bij Amazon geprobeerd om dat probleem uit het hiringsysteem te halen, maar uiteindelijk hebben ze dat systeem volledig weggegooid omdat de makers van het systeem het probleem niet eens op konden lossen.
We gaan naar de volgende slide. De volgende dreiging is gegevensverzameling, ofwel datahonger. Natuurlijk heb je met deeplearningsystemen een goede trainingset van data nodig. Het probleem is dat zowel de overheid als bedrijven erg happig zijn om data te verzamelen en te houden. We zijn een beetje hoarders, want we denken altijd: misschien heb ik het ooit nodig. Bovendien wordt wel vaker gezegd dat voor de kern van onze

businessmodellen data de nieuwe olie vormen. Hoe verwachten wij dan dat we gaan doen aan dataminimalisatie als data juist in de kern van ons businessmodel zitten?
Dat geeft natuurlijk wel een aantal problemen, want met zo veel data krijg je natuurlijk ook kans op een datalek. Dat komt tegenwoordig bijna elke dag voor. Als je de krant leest, dan zie je dat. Het beveiligen van alle data is bijna onmogelijk. Statistisch gezien zijn er 16 bugs per 1.000 lines of code. Wij zwemmen natuurlijk in een zee van complexiteit. Er is zo veel software om ons heen dat er statistisch gezien altijd kwetsbaarheden in zullen zitten. Uiteindelijk is de enige manier om een datalek te voorkomen de data niet hebben. Ik zeg dit natuurlijk ook als iemand die nu al zevenenhalf jaar de leiding heeft over een pentestbedrijf. Het is ons nog niet één keer niet gelukt om in een systeem in te breken. Het enige waar wij niks mee kunnen, is iets wat trivial, klein, is of iets wat bijna totaal geen functionaliteit heeft. Voor de rest zijn systemen altijd stuk te maken.
Natuurlijk is er ook een spanningsveld tussen de datahonger en de GDPR, ofwel de AVG. Er worden natuurlijk eisen gesteld aan bedrijven, en ook aan overheidsinstanties, en er is een aantal verplichtingen. Wij moeten kijken hoelang we de gegevens bewaren. We moeten ook toegang geven tot die gegevens aan mensen, aan burgers, zodat zij de data kunnen bekijken en ook veranderingen aan kunnen brengen. Maar goed, hoe vaak denk je dat dat in de praktijk daadwerkelijk gebeurt? Als ik wil weten wat voor data een data broker over mij heeft verzameld, dan zijn er wel blogposts of media te vinden die over dat onderwerp gaan. Maar uiteindelijk is het bijna onmogelijk om daar toegang toe te krijgen, laat staan om er ook nog echt verbeteringen in aan te brengen. Daarnaast is het met GDPR bijna onmogelijk om compliant te zijn.
Daar zijn een paar redenen voor. Ten eerste omdat vooral kleine bedrijven bijna geen resources hebben. Als je af en toe een pentest uit kan voeren, dan is dat natuurlijk wel fijn. Bovendien worden de regels bijna niet nageleefd. Wie gaat van mkb naar mkb om te vragen: hoeveel data heb je nog?
Er is ook heel weinig jurisprudentie, omdat de wetgeving er relatief nog niet zo lang is. Zonder die jurisprudentie weten bedrijven niet precies waar de grens is van wat ze wel en niet kunnen. Er zijn niet genoeg voorbeelden waardoor ze echt 100% weten hoe ze compliant kunnen worden. Uiteindelijk krijgen wij wel consultancybedrijven die met een checklist komen. Ze denken een beetje: als je dit, dit en dit doet, dan ben ik compliant, check! Maar uiteindelijk is dat ook een beetje bullshit. Het is wel een goed businessmodel, maar … Wat GDPR wel goed heeft gedaan, is de discussie aanwakkeren onder het bedrijfsleven en binnen de overheid. Ik denk dat mensen zich nu meer bewust zijn van dataverzameling en dataminimalisatie. Het is mooi gespreksvoer, maar uiteindelijk moeten we iets aan de enforcement doen.

De **voorzitter**:
Mevrouw Rieback, ik heb een korte interruptie. Deze bijeenkomst is voor iedereen om op niveau te komen, maar u steekt op sommige punten hoog in. Een pentest, zo begrijp ik, is een penetratietest. Dan kijkt u of u in de systemen kunt komen. Vat ik dat goed samen?

Mevrouw **Rieback**:
Ja, penetratietest is vaktaal voor ethisch hacken. Ik heb een bedrijf met hackers. Overheden en bedrijven betalen ons om dingen stuk te maken, en hun daarna uit te leggen wat wij hebben gedaan en hoe zij die fouten kunnen herstellen.

De **voorzitter**:
AVG kent, denk ik, iedereen, maar GDPR is, denk ik, ook niet voor iedereen een bekende afkorting. Of ga ik nou te ver?

Mevrouw **Rieback**:
Dat is de Europese privacywetgeving.

De **voorzitter**:
Dank u. Gaat u verder.

Mevrouw **Rieback**: Ik werk al twintig jaar in computerbeveiliging. Ik kan zeggen dat er op het gebied van privacy al best wat privacy enhancing technologieën ontwikkeld zijn. Gedurende mijn eigen carrière heb ik ook VPN's ontwikkeld. Toen ik op de VU werkte, was iedereen bezig met technieken om data te pseudoanonimiseren en dat soort zaken. Het zit allebei op het gebied van pseudoanonimisering, zoals dataminimalisatie. De technieken zijn er. Die zijn er eigenlijk al tien, vijftien, twintig jaar, maar het probleem is dat die technieken ook toegepast moeten worden. Daar zie ik het gebrek. Volgens mij hebben we de tools, maar de motivatie en de wil om ze te gebruiken missen af en toe.
Volgende slide. Vooroordelen, dus bias, vormen een andere bedreiging. Ik weet niet of jullie de geweldige documentaire Coded Bias op Netflix hebben gezien. Als jullie die niet gezien hebben: het is echt een aanrader. Het gaat over een AI graduate student op MIT in de Verenigde Staten. Zij is African American en probeert met AI een gezichtsherkenningsalgoritme te trainen, maar uiteindelijk werkt dat algoritme niet goed, omdat zij zwart is. Uiteindelijk doet zij een blank masker op en opeens doet dat algoritme het veel beter. Het geeft natuurlijk al aan dat de mensen die die systemen ontwikkelen, een bepaald wereldbeeld hebben. Dat wereldbeeld wordt verwerkt in die software. Ze doen het niet eens expres. Soms proberen ze zelfs bewust om dat wereldbeeld eruit te halen, maar

het is moeilijk tot onmogelijk om dat te doen. Uiteindelijk worden de vooroordelen verwerkt in de broncode, maar ook in de algoritmes en de training data sets. Die vooroordelen, die biases, uiten zich op verschillende manieren, bijvoorbeeld op basis van ras, economische toestanden, geslacht of geografie. Je ziet het ook met predictive policing. We proberen de bias op ras er wel uit te halen, maar als ik ga kijken naar de Bijlmer of Geuzenveld in Amsterdam, kom ik toch een beetje op dezelfde plek uit.
Volgende slide. Nog een dreiging. Als wij over ethiek in kunstmatige intelligentie praten, dan hebben mensen het al snel over het bekende trolleyprobleem. Voor degenen die dat niet kennen, geef ik een voorbeeldje. Je hebt een trolleybus. Op het ene spoor staan drie oude mensen en op het andere spoor staat één jonge persoon. Je moet dan kiezen welke persoon je doodrijdt. Maar goed, dit is wel een behoorlijk contrived voorbeeld. Je maakt je er een beetje makkelijk van af als je naar die theoretische voorbeelden gaat kijken en zegt: o ja, dit is AI-ethiek. Nee, dat is het eigenlijk niet. Naar mijn mening gaat het echte probleem over bedrijfsvoering. De hele wereld heeft last van de problematiek rondom bedrijven. Burgers hebben er last van. Overheden hebben er last van. Bovendien zijn overheden afhankelijk van bedrijven voor hun eigen bedrijfsvoering. Uiteindelijk komen die businessproblemen terug bij de overheid en daar hebben wij ook heel veel last van. Het probleem is dat bedrijven niet transparant zijn en niet democratisch zijn, maar wel een rol spelen in de democratie. Ik weet niet of jullie de krant hebben gelezen. Een whistleblower van Facebook heeft net een heleboel data en documenten aan de Wall Street Journal geleverd. Haar conclusie is dat niemand buiten Facebook weet wat er binnen Facebook aan de hand is. Die vrouw, die klokkenluider, is een expert in algoritmes, net als academici zoals Catholijn. Zij heeft kennis van zaken, maar toch is zij weggegaan bij Facebook om klokkenluider te worden.
De kennis is wel in huis, maar toch beslist de leiding om het anders aan te pakken. Dat komt natuurlijk door de perverse prikkels van de commercie. Zelfs aandeelhouders kunnen dit soort bedrijven, zoals Facebook, niet besturen. Zij oefenen bijna geen controle uit. Specifiek bij Facebook zie je dat Mark Zuckerberg een soort supervotingconstructie met A- en B-aandelen heeft. Mark Zuckerberg heeft tien stemmen voor elke stem die de andere aandeelhouders hebben. Dat betekent dat Mark Zuckerberg bij elkaar controle heeft over 60% van de stemmen binnen Facebook, al heeft hij waarschijnlijk maar zo'n 30% van de aandelen. De aandeelhouders, inclusief hedge funds, hebben geprobeerd een aandeelhoudersresolutie in te dienen om dat votingstelsel te veranderen naar "één aandeel, één vote". Mark Zuckerberg heeft dat zelfstandig weggestemd. Zelfs de aandeelhouders kunnen het gedrag van Facebook niet sturen. Dat betekent dat niemand het kan. Zelfs de mensen die daar werken, zien dat de enige manier om daar verbeteringen in aan te brengen, is om naar de Wall Street Journal te stappen met een lading documenten. Dit is

naar mijn mening het grote probleem met kunstmatige intelligentie. Dus nee, dat is inderdaad geen technologieprobleem. Maar goed, ik ga even verder.
Volgende slide. Ik wil ook even praten over mijn eigen ervaring met het spanningsveld tussen de grote techbedrijven en de Europese Commissie. Ik heb eerder verteld dat ik door de Europese Commissie gevraagd ben om werk te doen voor COVID-19 contact tracing apps. Ik noemde even de Google/Apple Exposure Notification API. Als pentestbedrijf hebben wij toestemming nodig om iets te testen. Dat heet een vrijwaringsverklaring. Wij noemen het in vaktaal een "get out of jail free card". Als die niet getekend is, zijn wij voor de computercriminaliteitswetten natuurlijk strafbaar. Wij blijven een klein computerbeveiligingsbedrijf, dus we zijn met de Europese Commissie naar Google en Apple gestapt om te zeggen: wij willen de Google/Apple Exposure Notification API testen en wij willen graag dat jullie deze pentestwaiver gaan tekenen. Van Apple hebben wij helemaal niks teruggehoord. Van Google hebben we wel iets teruggehoord, maar ze wilden even intern overleggen. Een maand later hadden we nog niets gehoord, dus pingden we weer even met de vraag hoe het zat. "O ja, wij moeten het escaleren naar het management. We komen er later op terug." Toen waren we weer een paar maanden verder. Uiteindelijk hadden wij er een gesprek met de Europese Commissie over, want deze Google/Apple Exposure Notification API komt op de telefoons van iedereen in de wereld. Dit is echt een kwestie van groot civiel belang, want er wordt continu data verzameld. We hebben totaal geen inzicht erin. Bovendien zit het echt in de kern van het besturingssysteem, dus het is ook heel gevoelig voor fouten en beveiligingsproblemen. Het is belangrijk voor de maatschappij om deze toepassing, deze software, te testen. We kunnen het wel reverse engineeren, maar dan komen wij mogelijk legaal, wettelijk, in de problemen. We hebben de Europese Commissie dus gevraagd: als Google een rechtszaak tegen ons gaat aanspannen, komen jullie dan op voor de legale kosten? Zij hadden zoiets van: het is een politieke en ingewikkelde kwestie; we willen natuurlijk wel onze goede relatie met Google behouden.
Er zijn natuurlijk uitzonderingen op reverse-engineering-wetten, bijvoorbeeld interoperability. Ik heb dus ook even gesnuffeld bij het ministerie van VWS en dezelfde vraag gesteld: jullie kunnen wel reversen bij de coronamelder, maar kunnen wij misschien bij jullie aanhaken om dit te doen? Politiek! Ingewikkeld! We zijn op die manier dus ook niet verder gekomen. Uiteindelijk hebben we overlegd hoe wij nu verder moeten, omdat wij op zo'n belangrijk stuk technologie niet legaal een security audit kunnen uitvoeren. Niemand wil ons hierbij ondersteunen. Het enige wat ik kan bedenken, is dat we het toch gewoon moeten doen en dat we, als er uiteindelijk een rechtszaak komt, maar moeten hopen dat alle advocaten van de EFF onze kant op springen en dat wij misschien jurisprudentie kunnen maken. We hebben echt letterlijk dat gesprek gehad binnen mijn bedrijf. Maar goed, ik wilde de relatie

tussen de EU en big tech natuurlijk niet beschadigen, dus ik ben uiteindelijk toch naar achteren gestapt. Maar dit zijn voor mij pijnpunten. Dit zouden voor jullie ook pijnpunten moeten zijn. Dit laat heel goed zien wat de machtsverhoudingen zijn tussen big tech en overheden, zelfs op Europees niveau.
Goed. De volgende slide: oplossingen. Ik denk ten eerste aan open source. De Nederlandse overheid is daar eigenlijk al jaren mee bezig. Het komt en het gaat. Het is natuurlijk niet volledig opensourcevriendelijk, maar Linus Torvalds zegt: "Many eyes make bugs shallow." Dus hoe meer mensen kunnen kijken naar software, hoe groter de kans dat daar uiteindelijk problemen uit gaan komen. Maar dit gaat niet altijd goed in Nederland.
Ik ga nog een leuk voorbeeld geven vanuit mijn eigen ervaring. Ik ben een keer gevraagd voor een RFP (request for proposal) om mee te doen om een offerte in te dienen voor het pentesten van een stemcomputer, die uiteindelijk stemmen van Nederlanders in het buitenland moest registreren. Het bedrijf dat dat systeem heeft gemaakt, heeft ook de scope van die penetratietest bepaald en is naar ons toegekomen met de vraag of wij een offerte kunnen indienen voor een pentest van één week. Ze wilden ook een hackathon maken. Ik wilde een belafspraak maken en toen ik de persoon van dat consultancybedrijf aan de lijn had, vroeg ik: we hebben het nu over stemmachines, maar vind je niet dat er méér nodig is dan een pentest van één week en een amateuristische hackathon?
Dit zou promotieonderzoek moeten zijn. Dit heeft echt jarenlang onderzoek nodig. Maar het probleem is dat ze aan mij vragen om iets in te dienen. Stel dat ik een offerte indien en daarop een penetratietest uitvoer, dan is de scope zo klein dat ik in een week bijna niks ga vinden. Dat is niet genoeg tijd. Ze kunnen dan wel zeggen dat Radically Open Security naar hun stemmachine heeft gekeken en niks heeft gevonden, maar ik doe niet mee met dat soort dingen. Ik heb een e-mail gestuurd naar die bedrijven, waarin ik zei dat ik weiger om een offerte in te dienen, want ik vind het uiteindelijk niet respectvol voor de democratie. Ik was bijna geneigd om naar een reporter te stappen met de hint om een Wob-verzoek op dit systeem in te dienen, maar dat heb ik uiteindelijk gelaten. Maar het punt is dat dit vaak is hoe het gaat met de automatisering en IT-systemen van de overheid.
Het komt steeds weer terug op die bedrijfsproblemen, dus het komt dan ook vaak terug bij bedrijven. Wij kunnen ook kijken naar die supervotingconstructies. Bedrijfsproblemen hebben bedrijfsoplossingen nodig, dus we moeten weg van dat soort aandelenconstructies. Als wij meer sociale ondernemingen kunnen stimuleren, dan denk ik dat wij ook beter gedrag gaan krijgen van bedrijven. Want als zij gecommitteerd zijn om minder winst te maken, dan denk ik dat impact uiteindelijk de key motivator gaat zijn. Het klinkt een beetje contra-intuïtief, maar ik denk dat wij, in plaats van ons te hard te focussen op de technische oplossingen, ook kunnen kijken naar die bedrijfsoplossingen.

Het vierde punt op deze slide, technologie, kan ook een oplossing zijn. Ik zeg, als iemand die al twintig jaar in de computerbeveiliging werkt, ook als voormalig assistant professor informatica: ik denk dat bijna alle technologieën die wij nodig hebben waarschijnlijk al bestaan. Het probleem zit hem veel meer in hoe zij toegepast worden.
Als conclusie bij deze laatste slide wil ik het volgende zeggen. Ik heb de laatste twintig jaar van mijn carrière privacy-enhancing technologies (PETs) ontwikkeld. Uiteindelijk ben ik gefrustreerd geworden, omdat ik het echt helemaal zat was om technologische pleisters te ontwikkelen en te plakken op businessproblemen. Ik ben daar een beetje van weg gestapt, want ik denk dat businessproblemen businessoplossingen nodig hebben. Daarom heb ik Radically Open Security gestart als non-profit-business. Daarom heb ik drie jaar geleden Nonprofit Ventures opgericht om als incubator andere non-profit-bedrijven te starten. Ik ben ook helemaal gedoken in economische filosofieën zoals post-growth en degrowth, omdat ik denk dat daar de oplossing in zit. Ik zou heel graag willen dat zowel de Eerste als de Tweede Kamer en alle politieke partijen meer aandacht zouden richten op die nieuwe economische gedachtestromen, want ik denk dat ook voor AI daar de oplossing in zit.
Ik heb mijn punt gemaakt. Laten wij in die laatste twee sessies niet tevreden zijn met technische oplossingen. Technische oplossingen zijn populair omdat zij relatief gezien makkelijk te maken zijn en omdat het "do something syndrome"-meespeelt, zoals wij dat in de computerbeveiliging noemen. Dat is het gevoel dat we iets hebben gedaan. Een technologie lijkt heel sexy, maar het is veel moeilijker om die bedrijven aan te pakken. Het is ook politiek moeilijker om de bedrijven aan te pakken. Maar dat is volgens mij de oplossing. Dank je wel.

De **voorzitter**: Dank u wel voor dit perspectief, het zeker aanvullende perspectief. Ik kijk naar de aanwezigen of er vragen zijn aan mevrouw Rieback. Meneer Dittrich van D66.

De heer **Dittrich** (D66): Hartelijk dank voor uw inbreng en uw visie. Ik begrijp dat u zegt dat er een politieke oplossing, of eigenlijk politieke sturing, nodig is voor de tekortkomingen die u ziet. In de Eerste Kamer houden wij ons natuurlijk met name met wetten bezig. Mijn vraag aan u is: kunt u aangeven op welke manier een wet de tekortkomingen die u ziet, zouden kunnen sturen of beperken in de problematiek? Begrijpt u mijn vraag? Ik vind hem zelf erg ingewikkeld. Nog één keer: hoe kan de wet u helpen om de problemen die u ziet op te lossen?

Mevrouw **Rieback**: Ik ga nu praten vanuit mijn eigen uitzichtpunt. Ik ben natuurlijk geen politicus, dus ik denk niet zo vaak na over wetgeving. Het probleem met wetten is dat, als je ze maakt, ze ook nageleefd moeten worden. Ik denk wel dat de wetten prikkels kunnen

geven, maar je ziet het nu ook met die CoronaCheck-app. Ik ben best vaak uit eten geweest en ik ben niet één keer in een restaurant gecontroleerd. Ik denk het meer een kwestie is van de informele politieke invloed, die jullie uit kunnen oefenen. Ik denk dat er meer vragen en discussies moeten komen over bijvoorbeeld het start-up-ecosysteem. Dat klinkt als heel iets anders dan kunstmatige intelligentie. De start-ups van vandaag zijn de grote bedrijven van morgen. Het probleem is dat wij vanaf het begin de verkeerde normen en waarden in onze bedrijven proppen.

Het probleem is dat we mensen hebben helemaal aan de top in Nederland, zoals bij Techleap -- dat is een beetje de organisatie tussen het start-up-ecosysteem en de overheid -- mensen zoals Prins Constantijn, die continu roepen dat Nederland een land moet worden van unicorns. Jullie moeten beseffen hoe ongelooflijk problematisch dit is, maar daar is echt bijna geen discussie over. Er is een econoom, Mariana Mazzucato, die praat over wat zij noemt "the entrepreneurial state". Uiteindelijk is die overheid dus dé grootste durfkapitalist die er bestaat. Als je bijvoorbeeld kijkt naar alle subsidies, dan gaan die naar mooie onderzoeken zoals die van Catholijn, maar ook naar veel andere dingen. Het probleem is dat wij de kosten externaliseren naar de overheden en uiteindelijk naar de belastingbetalers, maar dat de winsten allemaal worden geprivatiseerd. Dat zit in de kern van het Silicon Valley-businessmodel van capital-scale-exit. Dat wordt echt overal geleerd.

Het eerste wat ik zou willen -- dat heeft niets met wetten te maken -- is dat iemand even met Techleap gaat praten over de problemen van dit schadelijke businessmodel. Het eerste probleem is dat aan alle oprichters wordt verteld dat kapitaal verplicht is, dat je gewoon helemaal geen bedrijf kunt starten zonder investeerder. Ik ben zevenenhalf jaar geleden natuurlijk ook door een incubator heen gegaan met mijn bedrijf. Uiteindelijk is onze definitie van succes het verkopen van het bedrijf. Denk even na, ook qua data governance, over wat voor impact dit heeft. De Nederlandse overheid heeft dit bijvoorbeeld gezien bij Fox-IT, een bekende grote speler in de computerbeveiliging in Nederland. Wat is er gebeurd? Er zijn monitoring black boxes met AI geweest in alle hoeken van de overheid, in Defensie en in alle grote Nederlandse bedrijven. Wat is er uiteindelijk gebeurd? Fox-IT is overgenomen door de Britse NCC Group. Op een gegeven moment hebben de hele Nederlandse overheid en het bedrijfsleven zoiets van: o shit, al onze data zijn nu eigendom van de Britten.

Dus wat gebeurt er? We kijken naar de eerste de beste Nederlandse start-up die toch nog Nederlands is. RedSocks is een spin-off geweest van Fox-IT en heeft ook investeerders gehad. Uiteindelijk zijn alle partijen overgestapt naar RedSocks, want ze dachten: dat is een mooie, prachtige Nederlandse start-up. Dat bedrijf is exponentieel gegroeid. Twee jaar later lees ik in de krant: gefeliciteerd, RedSocks is overgenomen door de Roemenen. Wij zijn ons niet bewust van tot hoever wij Nederland aan het verkopen zijn aan buitenlandse investeerders. Ik woon in Amsterdam. Dat is ook voor mij natuurlijk een pijnpunt. Als ik een

telewinkel in de stad zie en nadenk over hoe onbetaalbaar woningen hier zijn, dan denk ik: hier moet discussie over komen.
Laatst was er dat Nationaal Groeifonds. Ze zochten groeiplannen voor Nederland. De bekende Britse econoom Tim Jackson, die post-growth heeft bedacht, heeft samen met mij en Fair Capital Partners een proposal ingediend voor het Groeifonds, dat we het "Post Growth Plan" hebben genoemd. In ons proposal zeggen wij: "Kunnen jullie even nadenken over de vraag of groei eigenlijk wenselijk is voor Nederland? Zo niet, dan zijn er hier een paar alternatieven." Ik had gehoopt dat we in ieder geval tot de tweede ronde zouden komen, zodat ik een gesprek zou kunnen voeren met de ministeries van Economische Zaken en Financiën. Verrassend genoeg zijn we niet eens door de eerste ronde gekomen, maar ze hebben wel de commitment gemaakt om alle proposals op de website te publiceren. Dus dat is wel gebeurd. Ik doe nu de open uitnodiging aan alle ministers en Kamerleden en iedereen van het ministerie: ik heb hier een heleboel ideeën over en ik wil heel graag met jullie praten.

De **voorzitter**: Dank, we hebben het gehoord.

Mevrouw **Stienen** (D66): Dank aan jullie allebei voor jullie inleiding. Ik moet zeggen dat ik een beetje het gevoel heb dat ik nu in een computergame zit met allerlei informatie en draadjes. Ik wil met de commissie en de collega's iets delen vanuit mijn rol als vertegenwoordiger in de delegatie naar de Parlementaire Assemblee van de Raad van Europa. We hebben daar vorig jaar een hele dag gehad over artificial intelligence. Daar is de conclusie getrokken dat "the Assemblee strongly beliefs that there is a need to create a crosscutting regulatory framework for AI with specific principles based on the protection of human rights, democracy and the rule of law". Dat is een aanvulling op de vraag van mijn collega Boris Dittrich, namelijk dat we een internationaal framework nodig hebben, dus niet alleen maar in de Nederlandse wetgeving. Ik ben benieuwd of u daarmee bekend bent. Ik zal u straks even linken via LinkedIn en u in verbinding brengen met de mensen die daar in Straatsburg mee bezig zijn. Maar de "regulatory frameworks" lijken ook een beetje op internationaal maatschappelijk verantwoord ondernemen: wil je het aanmoedigen of heb je echt regelgeving nodig op internationaal vlak?

Mevrouw **Rieback**: Ik denk dat we het van alle kanten moeten proberen. Ik vind het zeker een goede zaak dat de Europese Commissie nu bezig is met AI en ethiek in wetgeving. Ter voorbereiding had ik even gekletst met Catelijne Muller, die ook bezig was met het maken van ... Ik denk dus dat er experts zijn in Nederland die jullie kunnen helpen met het voorbereiden van dat soort wetgeving. Natuurlijk ben ik geen advocaat en ook geen

politicus, dus dat is niet mijn area of expertise, maar ik vind wel dat daar ook wetgeving voor moet zijn. Dat geeft natuurlijk ook prikkels aan de rest van de maatschappij om dingen na te leven. Maar voor de rest kijk ik meer vanuit het perspectief van een voormalig academicus en van iemand uit het bedrijfsleven.

Mevrouw **Jonker**: Ik wil graag iets aanvullen. Ik denk dat jullie beider perspectief daarbij van groot belang is. Er komt Europese regelgeving op dit gebied, die AI Act. Inderdaad hebben wij aan Catelijne Muller een hele goede om daar samen naar te kijken. In het kader van de aanstaande wetgeving vraag ik mij af in hoeverre die wetgeving niet uiteindelijk met name een killer is voor onze eigen kleinere industrie, terwijl de grote giganten er gewoon overheen kunnen lopen omdat ze in staat zijn om de boetes van zich af te houden en de rechtszaken dermate lang uit te spinnen dat ze daar gewoon mee wegkomen. Ik ben bang dat we met name de kleine industrie, de opkomende eigen industrie, daarmee belemmeren, en dat de grote AI-giganten in staat zijn om daar gewoon mee weg te komen. Zij kunnen die rechtskosten immers wel betalen.

De **voorzitter**: Dank voor deze aanvulling en zorg. Ik zie in eerste instantie mevrouw Faber van de PVV.

Mevrouw **Faber-van de Klashorst** (PVV): Ten eerste bedankt voor uw interessante inbreng, maar ik werd ook enigszins getriggerd toen u begon over de coronanotificatiemelder, over die API. Die is hier namelijk uitvoerig besproken. U geeft aan dat u die bedrijfsstructuur enzovoort zat bent, maar wij zijn de overheid soms ook een beetje zat. Daar bedoel ik mee dat die API gratis aangeboden is door Google en Apple. Als iets gratis is, ben jij het product, zeg ik altijd. Dan moet je altijd een beetje opletten. Er zijn zelfs overeenkomsten gesloten. Die waren zodanig dat Google en Apple die eenzijdig konden wijzigen, zonder dat de overheid daar iets over te zeggen had. Moet de overheid niet eens op haar hoofd gaan krabben en zich afvragen of er wel dergelijke overeenkomsten gesloten moeten worden met Google en Apple? Dit zijn namelijk wel hele grote jongens. Moet je wel jouw lot in hun handen leggen? Is het ook niet een mentaliteitsprobleem, niet alleen van de bedrijven maar ook van de overheden? Overheden faciliteren het namelijk. Deze Kamer is erop geattendeerd dat deze API gratis was en dat dat een risico is. Toch wordt daarmee ingestemd. Moeten we daarbij dan ook niet naar onszelf kijken? Hoe ziet u dat, vraag ik u via de voorzitter.

Mevrouw **Rieback**: Ik ben het op heel veel punten met u eens. Ik denk dat overheden het een beetje moeilijk hebben met deze onderwerpen. Ik denk namelijk dat de overheid relatief

gezien weinig verstand heeft van technologie. Soms ziet de overheid weinig alternatieven. Er wordt gezegd dat het gratis wordt aangeboden, maar ik vind dat het meer gedwongen op onze telefoons is gekomen. Het is niet opt-in geweest. Het is gepushed door een update, of wij het nou wilden of niet. De CoronaMelder en dat soort applicaties zijn wel opt-in, maar uiteindelijk is de onderliggende basis dat niet.
Ik kom even terug op die machtsverhouding tussen de grote techbedrijven en de overheid. Ik denk dat de overheid heel vaak meegaat met die techbedrijven, omdat het heel vaak een situatie van onmacht is. Jullie willen het misschien anders, maar zien geen andere manier. Ik denk dat het daarop neerkomt.

De **voorzitter**: Mevrouw Jonker, ook op deze vraag?

Mevrouw **Jonker**: Ja, gerelateerd aan de hele automatisering op zich en vooral ook aan robotisering, dat een verwant vraagstuk is. Het klinkt zo mooi, maar uiteindelijk is die software van bedrijven. Ik denk dat dat aansluit bij jouw opmerking, Melanie. Als je mensen vervangt door kunstmatige intelligentie, robotisering of automatisering, dan geef je daarmee de macht direct weg aan die bedrijven. Die krijgen we niet meer terug.

De **voorzitter**:
Dank u wel. Mevrouw Kluit, GroenLinks.

Mevrouw **Kluit** (GroenLinks) (GroenLinks):
De boodschap over degrowth en postgrowth triggerde mij wel een beetje. Als ik het goed begrijp, accepteren wij via ons scale-upmodel, start-upmodel en verkoop-de-boelmodel eigenlijk dat bedrijven kunnen groeien en dat strategische posities worden weggegeven zonder dat wij dat doorhebben. Daarnaast is het belangrijk om te vermelden dat wij bedrijven zo ver laten groeien dat wij als overheid onze grip daarop kwijt zijn. U zegt ook dat wij eigenlijk onvoldoende kennis en kunde hebben als politici of als politiek om dat goed te kunnen volgen. Kunt u daar nog wat meer over zeggen? Postgrowth en degrowth worden al snel aan links gelabeld, maar dat hoeft niet per definitie. Kunt u misschien iets meer zeggen over hoe wij dit anders zouden moeten doen? En zijn er ook bepaalde groottes bij bedrijven die wij nog kunnen accepteren of momenten waarop we weten dat we de grip op die bedrijven als overheid kwijtraken? Dat vraag ik mijzelf af als ik dit hoor.

De **voorzitter**: Dank. Ik wil u vragen te antwoorden in relatie tot AI.

Mevrouw **Rieback**: Wat zei u, sorry?

De **voorzitter**: In relatie tot artificial intelligence, om te voorkomen dat we een algemeen economisch debat gaan voeren.

Mevrouw **Rieback**: Helaas vind ik toch wel dat een grotere context voor artificial intelligence heel veel uitmaakt. Dit betreft alweer de populaire discussie over hoe belangrijk die unicorns zijn. Een unicorn is natuurlijk een bedrijf met een valuatie van 1 miljard dollar. Je moet die unicorns als plofkippen zien. Als start-up ga je natuurlijk heel veel investeringsgeld in je bedrijf pompen, zodat het er kunstmatig, snel, juicy, lekker en aantrekkelijk uitziet, totdat de "liquidatie" van het bedrijf plaatsvindt -- zo noemen ze dat echt. Dat is het woord voor de exit, als je het aan investeerders vraagt. Op dat moment wordt die plofkip doodgemaakt. De waarde wordt uit het bedrijf gehaald en uiteindelijk blijft er een soort van karkas over. Het probleem is dat 90% van die unicorns verliesmakend is. 90%! Heel veel politici beseffen dat niet. Op een gegeven moment hebben ze zo'n economische schaal dat ze wel winstmakend worden, maar dat gebeurt maar in 10% van de gevallen.
Het probleem is dat er bij verliesmakende bedrijven zonder een echt businessmodel eigenlijk sprake is van het subsidiëren van een monopolie. Dit gebeurt trouwens ook met overheidsgeld en door onze pensioenfondsen, die wegens lage rentes nu heel wanhopig op zoek zijn naar betere investeringen. Het is dus allemaal een beetje complex. Er is dus een soort van pump-and-dumpscheme op de openbare markten. Die plofkippen worden gehypet, zo van "o, dit is echt geweldig", maar uiteindelijk wordt dat verliesmakende ding gedumpt aan mensen die het verkopen. Dat zijn niet alleen retailinvesteerders, maar ook pensioenfondsen en investeringsfondsen. Als we in een groeiende economie zitten komt het allemaal goed, maar stel dat het maart 2020 is en de economie in een dip komt, dan is het altijd de laatste investeerder die de schade oploopt. We moeten kijken naar de manier waarop dat hele systeem in elkaar zit. Zo'n pump-and-dumpscheme zou eigenlijk illegaal moeten zijn. In andere landen zijn daar ook proposals voor. Als politicus zou ik mij dus ook focussen op dat soort punten, want ik denk dat daar concrete dingen uit te halen zijn.

De **voorzitter**: Mevrouw Jonker op dit punt.

Mevrouw **Jonker**: Ik zal het vanuit een iets andere hoek proberen te bekijken. Technologie is nooit waardevrij. Bij artifical intelligence, zeker daar waar het wordt ingezet voor besluitvorming, gaat het er natuurlijk om in hoeverre die besluitvorming aansluit bij onze normen-en-waardensystemen. Als je je realiseert dat AI vanaf de oorsprong al beïnvloed is door de waarden van de makers, zonder dat ze daar negatieve bedoelingen bij hebben, dan hebben we hier te maken met een wicked problem. Technisch kun je het misschien voor een

heel eind voor elkaar krijgen, maar mijn collega heeft absoluut gelijk dat het probleem niet alleen daar zit. Als je nieuwe technologie wilt introduceren, zeker op dit soort besluitvormingsprocessen, dan moet je je realiseren dat je je hele proces opnieuw moet ontwerpen, dat je je hele proces van stap 1 af moet ontwerpen op een manier die aansluit bij de impact die die technologie gaat hebben op je besluitvormingsproces en op de mensen die ermee omgaan.

De **voorzitter**: Dank voor deze stevige conclusie, waarover we kunnen doordenken.

Mevrouw **Jorritsma-Lebbink** (VVD): Daar ben ik zeer voor. Laat daar geen misverstand over bestaan. Ik ben tegen monopolies. Door de technologie zijn hele grote monopolies ontstaan. Het zal wereldwijd moeten gebeuren, want als Europa alleen lukt dat niet, laat staan als Nederland alleen. Er zal nieuwe competitiewetgeving moeten komen. Eigenlijk is het een beetje hetzelfde als we hebben gezien aan het begin van de vorige eeuw. Toen zijn er allerlei bedrijven opgesplitst, omdat men toen ook monopolies kreeg door nieuwe technologische ontwikkelingen. Ik denk dat dit moet gebeuren, maar we moeten niet de illusie hebben dat wij dat in Nederland in ons eentje kunnen doen. Maar we kunnen wel helpen bij het op gang krijgen van de discussie.
Ik heb niet zo veel last van die covidchecker, ik heb veel meer last van de desinformatie die ertoe leidt dat heel veel mensen zich niet laten vaccineren. Wat doen we daaraan? Wat kunnen we daar nou nog meer aan doen dan we nu al doen?

Mevrouw **Jonker**: Dat is recht in mijn hart geschoten; ik kan het niet anders zeggen. Het onderzoek waar ik mij mee bezig hou, gaat met name over het ondersteunen van deliberatieve processen. Wij bekijken of we mensen met behulp van eerdergenoemde samenvattingstechnologie snel toegang kunnen geven tot waar de kern van een debat zit. Bovendien willen we samenvattingen waardegebaseerd maken: vanuit welke waarden, vanuit welke achterliggende belangen zegt iemand iets en wil iemand een oplossing bereiken? We zien heel vaak dat mensen het op lange termijn best wel met elkaar eens zijn, maar er is onenigheid over het pad ernaartoe. Hoe komt dat? Wat zijn daar de onderliggende waardes van? Kunnen we met behulp van onze kennistechnologische AI en onze machinelearningtechnologie de mensen de middelen geven om snel toegang te krijgen tot waar de kern van dit debat zit?

Mevrouw **Rieback**: Ik wil ook een antwoord op deze vraag geven. De klokkenluider heeft hierover meer informatie gegeven. Facebook had intern een nieuw algoritme geïntroduceerd om kwalitatief betere sociale interacties te bevorderen. Uiteindelijk betekende dat dat er

meer reacties komen op bepaalde posts. Hoe controversiëler, hoe meer ruziemakend de content is, hoe hoger die content wordt gepromoot. Facebook heeft gekozen om dat algoritme te maken en te implementeren, omdat het meer interactie betekent en meer reclameverkoop, dus uiteindelijk meer geld. We zien heel duidelijk dat winstmaximalisatie hierbij conflicteert. Dit is het probleem van inflammatory content. De enige manier die wij hebben om Facebook te laten luisteren, is het introduceren van maatregelen waardoor het bedrijf financieel pijn lijdt. Dat kan anticoncurrentiewetgeving zijn, dat kunnen hele hoge boetes zijn. Hoe meer boetes jullie dat soort bedrijven opleggen, hoe meer jullie het bedrijfsbestuur kunnen verbeteren. Dat is uiteindelijk de oplossing.

De **voorzitter**: Dank u wel. Mevrouw Nanninga, u bent de laatste in de rij.

Mevrouw **Nanninga** (Fractie-Nanninga): Dank, voorzitter. Ik dank de aanwezigen voor het delen van hun tijd en kennis met ons. Die technologie gaat natuurlijk altijd veel sneller dan wij hier besluiten kunnen nemen. Voordat iets door de Eerste Kamer is, laat staan voordat iets Europees geregeld is … Wetgeving maken is een ontzettend stroperig en traag verhaal. Tegen de tijd dat die wetgeving er is, is de hele technologie misschien alweer drie stappen verder. Is het dan niet zinvoller om meer in te zetten op economische wetgeving voor bepaalde bedrijfsmodellen? Wilt u daar een reflectie op geven?

Mevrouw **Jonker**: Het idee is meestal snel daar. Een proof of concept, een prototype, kunnen we meestal heel snel maken. Maar sommige zaken vragen echt meer tijd, op z'n minst een promotieonderzoek. Andere zaken kosten nog veel meer tijd om echt goed door te akkeren, om echt uit te vinden hoe je die technologie tot in de details moet beheersen en hoe je de effecten ervan moet begrijpen. Kunstmatige intelligentie heeft het vermogen om zelf testen te genereren en om zelf data te creëren waarop het zichzelf kan trainen. Daarmee wordt de versnellingsfactor er bovenop gezet. Aan de andere kant is het dermate complex dat er erg veel tijd nodig is om echt uit te zoeken wat die complexiteit teweegbrengt. Nogmaals, het proof of concept gaat razendsnel, maar zorgen dat we echt weten waar we het over hebben, kost heel veel tijd en onderzoek.

Mevrouw **Rieback**: Technologie gaat soms snel, maar soms ook niet. Vaak gaat het om bouwstenen die er al sinds de jaren 70 zijn. Of technologie nou zo snel beweegt, weet ik dus niet. De toepassingen wel. Het is hetzelfde met wetgeving. Jullie hebben hele mooie bouwstenen. Het gaat meer om de toepassing van die wettelijke bouwstenen. Dat is wat telt. Het gaat uiteindelijk om de compliance rondom die wetten. De interpretatie beïnvloedt heel vaak het gedrag. Ja, jullie kunnen natuurlijk aan wetgeving werken, maar ik denk dat jullie

veel meer kunnen op cultureel niveau. Denk aan compliance officers. Die zijn vaak conservatief. Ze vinden zichzelf een beetje saai en willen echt de regels naleven. Maar denk aan het doel van wetten. Vaak zijn wetten er om mensen en de maatschappij te beschermen. Als jij bezig bent met compliance, heb je uiteindelijk de regels en de geest van de regels. Denk aan het concept van social compliance. Stel je neemt een aantal kunststudenten die je dropt in een compliance-afdeling. Wie gaat nou echt kijken waar een wet over gaat? Het is hetzelfde met security.

De **voorzitter**: Mag ik u vragen af te ronden?

Mevrouw **Rieback**: Heel kort. Culturele problemen hebben culturele oplossingen nodig. Jullie zitten in een prachtige positie in het publieke oog om een culturele verandering te bewerkstelligen.

De **voorzitter**: Dank u wel voor de mooie laatste woorden van u beiden. Dank dat u ons op vliegniveau heeft gekregen, zodat wij volgende week verder kunnen. Volgende week om 9.00 uur gaan we verder. Over de manier waarop, krijgt u nog bericht. Waarschijnlijk wordt het een combinatie van fysieke en digitale aanwezigheid.
Nogmaals heel veel dank. Ik wens u een goede dag.

Sluiting: 10.52 uur.