



Artificial Intelligence en Ethiek In gesprek met: Dasha Simons (IBM)



Interview door: *Mirjam Veen-Hendriks (VCO), Raimond Stappers (VCO)*

Introductie

De RegTech Tafel van het VCO heeft Artificial Intelligence & Ethiek gekozen als thema voor het jaar 2023. De ontwikkelingen rond AI gaan snel. Waarom is dit een belangrijk onderwerp voor Compliance Officers en wat wordt er hieromtrent van hen verwacht? Om antwoorden te vinden op deze vragen gaan wij dit jaar in gesprek met een aantal professionals op het gebied van AI en Ethiek.



Dasha Simons, Managing consultant Trustworthy AI bij IBM Nederland.

Wie is Dasha Simons?

Ik werk bij IBM als managing consultant Trustworthy AI. Ik adviseer klanten hoe zij AI op een zo betrouwbaar mogelijke manier zo ver mogelijk kunnen ontwikkelen. Of het uitlegbaar is, eerlijker en transparanter, dat zijn de gebieden waar vooral mijn expertise ligt. Dat kan binnen verschillende sectoren zoals de financiële wereld, de overheid, consumentengoederen. Dat is de ene kant van mijn rol. Het andere deel van mijn werkzaamheden is meer op educatie gefocust. Ik ben verantwoordelijk voor het trainen van onze consultants in EMEA & UKI (dat is UK en Ierland) bij IBM op het gebied van betrouwbare AI.

Wie ben je privé?

Mensen zullen mij omschrijven als een nieuwsgierige, vrij avontuurlijke en filosofisch denker die op zoek is naar de betekenisvolle discussies in het leven. Ik ben in Nederland geboren maar ben half Russisch en half Nederlands, zoals mijn naam wellicht verraad.

Hoe ben je op deze plek terechtgekomen?

Ik heb Industrieel Ontwerpen gestudeerd, eerst aan de TU Eindhoven en daarna in Delft. Ik vond kunstmatige intelligentie altijd al interessant en wilde graag op dit onderwerp afstuderen. Wanneer krijg je de tijd om 6/7 maanden volledig op een onderwerp te focussen? Ik kwam erachter dat er heel veel ethische vraagstukken spelen rond kunstmatige intelligentie.

Bij IBM heb ik een afstudeeronderzoek gedaan naar designmethodes en technieken die je kunt gebruiken voor het ontwikkelen voor eerlijkere AI-systemen. Hierbij zijn meerdere principes van belang zoals: je mag niet discrimineren, het moet uitlegbaar zijn, het moet transparant zijn. Dit zijn mooie begrippen maar ze zijn ook abstract. Als je iemand een lening wil toekennen, wat betekent dan niet discrimineren binnen het AI-systeem? Uiteindelijk maak je altijd onderscheidt tussen aanvragen.

Wat je je dan afvraagt is wat vinden we een goed onderscheid en wat vinden we een niet passend onderscheid? Je hebt natuurlijk de wet maar je hebt wellicht ook dingen die je als bedrijf of overheidsinstantie alsnog niet wilt. Hoe kun je bij deze vraagstukken data scientists ondersteunen? Daar ging mijn onderzoek over. Zo ben ik bij IBM beland bij de consultancy tak. Daarbij denken mensen wellicht, waarom IBM? Dat is omdat IBM niet enkel gericht is op consultancy. Er worden ook echt dingen gebouwd. IBM combineert consultancy met 'hard tech' skills. Daarnaast is het businessmodel niet gericht op de verkoop van data en dat is in lijn met de AI-ethische waarden die ik zelf belangrijk vind.

Wat zijn je drijfveren?

Het is denk ik een typisch Millennial gedachte, maar ik wil graag bijdragen aan een betere wereld van morgen. Wat ik heel tof vind is dat je enerzijds heel technisch bezig bent maar toch ook sociaal maatschappelijk bijdraagt aan deze wereld. Je houdt je bezig met de laatste technologische ontwikkelingen wat intellectueel heel prikkelend is. Tegelijkertijd kan je het vertalen naar vragen als: Waar willen we naartoe als maatschappij? Hoe kunnen we een maatschappij maken die eerlijk goederen verdeelt? Je komt op een bepaald moment uit op hele abstracte sociale vraagstukken. Bijvoorbeeld wat is eerlijkheid op het vlak van het toekennen van een lening? Maar dit wil je weer vertalen naar wat betekent dit technisch? En dan weer terug, als het AI-systeem is implementeert, dan is het vervolgens interessant om te kijken naar het sociaal maatschappelijke geheel. Juist deze vraagstukken zijn enorm boeiend.

Wat wil je uiteindelijk bereiken? Wat is je ambitie?

Op het vlak van werk lijkt het me mooi om bij te dragen aan de ontwikkeling van een sterk front in Europa op het gebied van kunstmatige intelligentie dat voldoet aan onze waarden. Ik sta erg voor een Europese aanpak van dit onderwerp. We kunnen meer op Europees niveau bereiken de komende jaren naast de nationale Nederlandse aanpak. De combinatie van academisch onderzoek en de industrie boeit mij sterk. Daarbinnen lijkt het me fantastisch om een onderzoeksafdeling te leiden zodat je een visie kan neerzetten om daarmee het veld verder te brengen zonder het oog op de praktijk en de toepasbaarheid te verliezen.

Als we nu even terugkijken, waar ben je trots op?

Ik denk toch wel de Trustworthy AI training. Dat is wel iets waar je echt impact mee maakt. Tijdens een van de trainingen was er een data scientist, die zich op dat moment bewust werd van het feit dat AI-systemen ook kunnen discrimineren. Dat dat iets is waar hij op moet letten en dat hij het in het vervolg anders aan gaat pakken. De training laat je een andere manier van denken zien, nieuwe inzichten. Het is tof als je dat bewustzijn kan brengen.



Wat zie jij op dit moment als de belangrijkste ontwikkelingen op gebied van AI?

Er is elke dag zoveel gaande. Mijn research collega's zijn alleen al anderhalf uur per dag bezig om te kijken welke tools er op de markt zijn gekomen die specifieke dag. Dat geeft wel een beetje een beeld van hoe snel het gaat.

Generatieve AI gaat een grotere rol spelen. Dat doet het al maar dat wordt nog groter. Verschillende bedrijven zijn hier al mee aan het experimenteren. Generatieve AI is bijvoorbeeld ChatGPT. Generatieve AI genereert iets nieuws in plaats van dat je bijvoorbeeld een deterministische voorspelling maakt. Dat kan een afbeelding zijn of tekst. Heel veel typewerk wat we doen kan worden geautomatiseerd. Denk aan contracten opstellen, reviewen van documenten, samenvattingen maken van bijvoorbeeld nieuwe regelgeving.

Maar hier zijn ook ethische risico's aan verbonden. Wat je ziet is dat we steeds meer zijn gaan vertrouwen op foto en video als bewijsmateriaal om te zien of iets daadwerkelijk is gebeurd. Je kunt nu al een foto maken via AI waarbij Poetin en Merkel samen dansen. Momenteel kan je de namaakfoto nog onderscheiden van de echte foto maar dat verschil wordt steeds kleiner. Hoe zorgen we ervoor dat je je eigen keuzes kan maken zonder dat het beïnvloed wordt door misinformatie? Dit wordt steeds lastiger.

Hoe kijk jij aan tegen de kennis die men heeft op gebied van AI?

Ik ben blij met de kennisontwikkeling van de afgelopen paar jaar. Mijn focus is vooral op de ethische aspecten en daarbinnen zie ik dat meer en meer mensen zich bezighouden met dit onderwerp. Zelfs mijn fysiotherapeut vraagt er nu naar. Het leeft veel meer, het is mainstream aan het worden. Dat maakt het makkelijker om het er over te hebben, ook voor gesprekken over de afwegingen. Er is meer begrip voor de ethische vraagstukken. AI is er genoeg werk te doen voor de implementatie ervan!

Wat kun je zeggen over de impact buiten de financiële sector?

De zware regulering maakt dat de financiële sector een voorloper is op dit vlak. Dit maakt echter ook dat er soms weerstand is, omdat data scientists vaak al continu bezig zijn met model documentatie terwijl ze juist bezig willen zijn met het bouwen en ontwikkelen. Verder is de overheid natuurlijk veel met onderwerp bezig. Je merkt het ook aan de telecommunicatiesector die gereguleerd en de meeste grotere bedrijven zijn hier inmiddels proactief mee aan de slag gegaan.

Welke valkuilen zie jij voor partijen die bezig zijn met AI?

Als we kijken naar generatieve AI dan vinden velen het een tof onderwerp om mee te werken. Wat we soms zien is dat een innovatie team met AI gaat experimenteren zonder dat ze kijken naar de processen waarbinnen je het moet passen of zonder rekening te houden met de skills sets die de

medewerkers nodig hebben die ermee moeten werken. Juist de dingen er omheen, die moeten bijdragen aan de waarde-generatie, worden niet altijd juist beoordeeld.

Een tweede voorbeeld is: Zeker in grotere organisaties zijn data scientists werkzaam bij diverse afdelingen zoals marketing, HR of productontwikkeling. Wat je ziet is dat deze allemaal aan de slag gaan met AI zonder dat daar een bedrijfsbrede filosofie/ beleid achter zit. Er ontbreekt een centrale plek waar het overzicht is, die kan borgen dat wordt voldaan aan de wet en die bewaakt welke rol de organisatie speelt in AI-land. Zijn we proactief en willen we meer aan corporate social responsibility doen of voldoen we enkel aan de wet, de minimumnorm. Deze vraagstukken spelen op dit moment. Ik denk niet dat je op een centrale plek bezig moet zijn met AI. De data scientists van de afdelingen kunnen prima aan de hand van de visie en strategie werken. Ze hoeven daarvoor niet naar een centrale plek binnen het bedrijf. Wat wel goed is, is dat op een centrale plek de modellen gemonitord worden om te zien of aan het interne beleid wordt voldaan.

Wat je verder ziet is dat er wel AI en ethiek principes worden opgesteld maar hoe je dat daadwerkelijk meet binnen je modellen is nog in ontwikkeling binnen bedrijven.

Wat is de rol van de compliance officer volgens jou?

Wat we eigenlijk adviseren is dat er een AI-Ethics Board is. De compliance officer heeft uiteraard zitting in deze board. Daarnaast zie je dat sommige bedrijven AI & Ethiek beleggen bij compliance en je ziet bedrijven die het beleggen in de 'Tech-hoek'. Beide hebben voor- en nadelen. Binnen compliance kan het risico zijn dat het als een 'moetje' gezien wordt. Het risico bij Tech is dat zij het niet zo serieus neemt als compliance. Er is een sterke behoefte aan mensen die de wet- en regelgeving goed kunnen doorgronden in combinatie met goede technische skills.

Welke afdelingen zouden in een Ethics Board vertegenwoordigd moeten zijn?

Ik ben een voorstander dat AI & Ethiek op Boardniveau wordt belegd. Dergelijke ethische en strategische besluiten kunnen niet enkel door data scientists gemaakt worden. Dat moet op een hoger niveau bepaald worden. Dus managing board (strategie), compliance, CIO, legal, privacy, innovatie. Innovatie juist ook omdat zij anders naar de problematiek kijken. Daarnaast is het goed als iemand vanuit de commercie betrokken is bij de ontwikkeling. Wat je ook vaak ziet is dat iemand vanuit communicatie of HR betrokken is.

Afhankelijk van de inrichting binnen de organisatie speelt compliance een rol bij het ontwikkelen van een monitoringsraamwerk voor de aankomende wet- en regelgeving. Denk aan de European AI Verordening.

Heb je nog een laatste tip voor compliance officers?

Compliance officers kunnen de organisatie begeleiden bij het bepalen van de visie en strategie ten aanzien van betrouwbare AI-ontwikkeling. Er is een sterke behoefte aan expertise die aan de ene kant de wet- en regelgeving begrijpt



en aan de andere kant de techniek achter AI begrijpt en de vertaalslag kan maken. Dus mocht je nog weinig/geen AI-kennis hebben, raad ik aan om een training hierover te volgen!

Compliance is natuurlijk belangrijk bij de bewaking van het voldoen aan de wettelijke norm. Er is ook een rol voor compliance voor het meehelpen te bepalen van de interne norm en het bewaken van het ethisch kompas van de organisatie. Niet alleen van wat er niet kan maar ook wat betekenen je bedrijfswaardes in de AI-wereld? De compliance officer kan de organisatie hierbij ondersteunen.



Dit interview is een publicatie van de Vereniging van Compliance Professionals; sinds 2001 dé beroepsvereniging voor professionals op het gebied van Ethiek, Compliance en Integriteit.