

Mening: Bias die AI beïnvloedt.

Uit het Cambridge Handbook of Artificial Intelligence [1]: “ *It is widely agreed that current AI systems have no moral status. We may change, copy, terminate, delete, or use computer programs as we please; at least as far as the programs themselves are concerned. The moral constraints to which we are subject in our dealings with contemporary AI systems are all grounded in our responsibilities to other beings, such as our fellow humans, not in any duties to the systems themselves.*”

Oftewel, de systemen hebben zelf geen invloed op hoe zij keuzes maken en het is de mens die het systeem aanleert.

Een bias is een afwijking in de resultaten (output) van bepaalde machine learning algoritmes (Kantarci, A.,2020). Dit kan door menselijke fouten of persoonlijke meningen komen. Een racist met een afkeur voor mensen met een aziatische afkomst zal deze mensen sneller hierdoor laten beïnvloeden.

Bias kan ook komen door te weinig data, bijvoorbeeld een bevolkingsgroep waar weinig medische data over is zal niet gerepresenteerd worden door de data van de andere bevolkingsgroepen (Harvard Business Review, 2020).

Een mogelijke manier om in de toekomst een bias te voorkomen, is door blinde tests uit te voeren en zo te testen of er een bias aanwezig is (Harvard Business Review, 2020). Zo kunnen deze uit het systeem verwijderd worden. Alleen voor menselijke bias kan dit natuurlijk niet zo uit de ontwikkelaar ‘verwijderd’ worden, dat is wel waar het echte probleem ligt.

Valt je eigen mening samen met je professionele taak? Moet je je (professionele) medewerking verlenen aan deze techniek?

Elk bedrijf die een AI-toepassing maakt heeft een code of conduct, bovendien zijn de ‘human rights, well-being and ’ uit de algemene principes van de IEEE [2], er om er voor te zorgen dat een professional zijn of haar mening niet zomaar kan toepassen in de AI.

Dus bijvoorbeeld een ontwikkelaar die vindt dat vrouwen geen baan moeten hebben in de IT, kan er voor zorgen dat een recruitment-AI een bias heeft op vrouwen. Terwijl dit niet de professionele taak is. Dus, nee je eigen mening valt niet samen met je professionele taak. Je kan als professional altijd weigeren als het echt niet overeenkomt.

We hebben allemaal onze eigen normen en waarden, maar in dit geval vind ik dat je als professional moet zorgen dat er alles aan gedaan moet worden om een bias te voorkomen.

Bronnen

- [1] The Ethics of Artificial Intelligence, Cambridge Handbook of Artificial Intelligence, <https://dlo.mijnhva.nl/d2l/le/content/203189/viewContent/425792/View>
- [2] The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems, <https://dlo.mijnhva.nl/d2l/le/content/203189/viewContent/563704/View>
- [3] Radboud University Nijmegen. (2018). Discrimination, artificial intelligence, and algorithmic decision-making.

<https://rm.coe.int/discrimination-artificial-intelligence-and-algorithmic-decision-making/1680925d73>

- [4] A Simple Tactic That Could Help Reduce Bias in AI. (2020, 4 november). Harvard Business Review. <https://hbr.org/2020/11/a-simple-tactic-that-could-help-reduce-bias-in-ai>
- [5] Kantarci, A. (2020, 5 november). Bias in AI: What it is, Types & Examples, How & Tools to fix it. appliedAI. <https://research.aimultiple.com/ai-bias/#what-is-ai-bias>