Dasha Simons: “AI moet betrouwbaar zijn en dat begint bij de mens”

03-11-2025 10:00



Wanneer we praten over kunstmatige intelligentie, gaat het vaak over wat de technologie kan. Maar Dasha Simons van IBM richt zich juist op de vraag hoe AI zich hoort te gedragen en hoe we het wenselijk willen inzetten. Op inct.spiratie 2025 zal ze hier in haar closing keynote op ingaan.

Als consultant op het gebied van AI & Ethiek en AI & Governance bij IBM helpt ze organisaties om AI op een betrouwbare, transparantere en verantwoordelijke manier in te zetten. “AI kan enorm veel, maar zonder goede waarborgen versterkt het vaak juist bestaande vooroordelen en ongelijkheden.”
Toen IBM jaren geleden begon met het ontwikkelen van richtlijnen voor betrouwbare AI, stond het onderwerp nog in de kinderschoenen. Simons: “We hebben met klanten gekeken hoe we AI-systemen op een betrouwbare manier kunnen inrichten, van governance-structuren en ethische adviesraden tot beleid en uitlegbaarheidslagen in de technologie zelf.”

Verborgen vooroordelen

Die noodzaak is volgens haar alleen maar groter geworden. “Je ziet dat AI vaak bestaande vooringenomenheid versterkt. Neem het voorbeeld van Amazon, waar een wervingssysteem werd getraind op historische data van werknemers. Omdat het merendeel daarvan man was, leerde het systeem dat ‘goede werknemers’ vooral

mannen waren. En dus werden vrouwen automatisch uitgesloten.” En nu al helemaal met de nieuwe modellen voor video-generatie zoals Sora 2 en het brede gebruik van ChatGPT voor vriendschappelijke en romantische relaties.

Daarom werkt Simons met bedrijven aan vraagstukken als eerlijkheid, uitlegbaarheid en privacy. “Bijvoorbeeld bij uitlegbaarheid gaat het erom dat je begrijpt waarom een AI-systeem een bepaalde beslissing neemt. Is iemand afgewezen omdat hij niet op een schaatsvereniging zat, of omdat hij vijf jaar relevante werkervaring miste? Dat maakt nogal verschil.”

De dunne lijn tussen innovatie en risico

Een belangrijk thema is privacy en copyright. Zeker nu AI-systemen gevoed worden met enorme hoeveelheden data, is de vraag: waar blijft die informatie? Van wie is deze informatie? “Veel modellen worden nog steeds getraind op gebruikersdata. Je kunt instellingen uitzetten of modellen lokaal draaien, maar helemaal zeker weet je nooit of die data niet tóch ergens belanden. Het blijft een kwestie van vertrouwen.”

Voor uitgevers en mediabedrijven leidt dat tot een paradox, zegt Simons. “AI is tegelijk een kans en een bedreiging. Je kunt het gebruiken om teksten te genereren of te herschrijven, maar diezelfde tools halen hun informatie ook weer uit jouw content. Het businessmodel komt daardoor onder druk te staan.”
Toch is ze optimistisch: “Ik denk dat we AI alleen bewust kunnen inzetten, nadenkend op de manier waarop, en welke impact dat heeft op de samenleving, maar dan moeten we dit wel doen.”

Naar een nieuwe ethiek van creatie

Naast haar werk bij IBM doet Simons promotieonderzoek aan de Universiteit van Amsterdam, gericht op de ethische aspecten van video-AI. “Denk aan video-analyse, video-generatie en de vraag wat dat betekent voor creativiteit, weerspiegeling van de realiteit en authenticiteit.”
Ze ziet enorme mogelijkheden, bijvoorbeeld gepersonaliseerde kinderboeken die automatisch worden voorgelezen, maar heeft ook onbeantwoorde vragen. “Als AI alles kan genereren, wordt het steeds moeilijker om echte vakmanschap te herkennen. Alles kan op elkaar gaan lijken. De uitdaging wordt straks om nog het echte onderscheid te vinden.”

“Betrouwbare AI begint bij de mens,” besluit Simons. “We moeten systemen bouwen die eerlijk, uitlegbaar en veilig zijn. Technologie is nooit neutraal. Het is zo ethisch als de mensen die het ontwerpen.”

Dasha Simons zal inct.spiratie 2025 afsluiten : AI kan veel. Verantwoordelijkheid dragen hoort daar niet bij. Dit is een prikkelende keynote die scherp en speels de ethische rafelranden blootlegt.

Er zijn nog een paar plaatsen beschikbaar

MEER INFO & AANMELDEN

David Huijzer