Kennissessie AI: hoe snel gaat het en wat moet de overheid?

Een sessie propvol informatie over AI en hoe de overheid daarmee om kan gaan. Hoe snel gaan de ontwikkelingen? Welke vormen van AI zijn er? Wat is de verantwoordelijkheid van de overheid? Welke instrumenten en kaders zijn er? Mogen rijksambtenaren gebruik van ChatGTP? AI-adviseur Joost van der Burgt van het Rijks ICT Gilde (RIG) had antwoorden op alle vragen.

Vijf jaar geleden leek kunstmatige intelligentie (AI) een speeltje van technici. Deze maand debatteerde het kabinet over generatieve AI. Met deze woorden opende Richard Vielvoije, directeur Rijksorganisatie ODI, de kennissessie over AI. Vervolgens introduceerde hij Joost van der Burgt, adviseur AI bij het Rijks ICT Gilde (RIG).

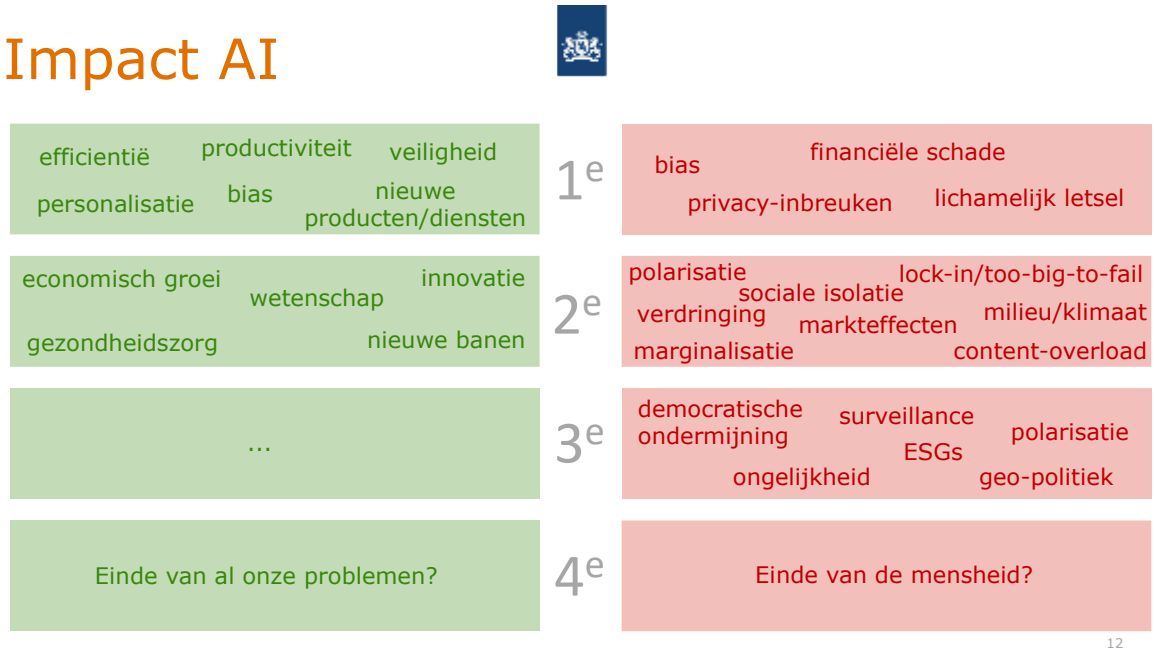
Kunstmatige intelligentie (AI) is complexe techniek waar vooral datacenters voor nodig zijn (en geen robots). In vrijwel ieder overheidsrapport over AI staat de frase “vanwege toegenomen rekenkracht en data”. Het belang van die zin wordt onderschat, want die rekenkracht neemt exponentieel toe. Elke twee jaar worden computers ongeveer twee keer zo sterk. Deze voorspelling uit de jaren zestig (de Wet van Moore) geldt nog steeds. Op dit moment is de krachtigste computer in staat tot 1,2 triljoen (18 nullen) berekeningen per seconde. Dat is het equivalent van 12.000 sterke pc’s. Per seconde.

Hoe snel gaat het?
De ontwikkelingen gaan veel sneller dan zelfs de deskundigen konden vermoeden. In 1996 versloeg IBM’s supercomputer Deep Blue de schaakgrootmeester Gary Kasparov. In 2016 versloeg AlphaGo van Google’s DeepMind meervoudig wereldkampioen Go, Lee Sedol. Go is een veel complexer spel dan schaken. In 2018 creëerde DeepMind een nieuwe versie van AlphaGo, die de oude versie 100-0 kon verslaan. De AI AlphaStar (ook van DeepMind) behaalde in 2019 het hoogste niveau in het supercomplexe online spel StarCraft II, op basis van dezelfde informatie die de menselijke spelers kregen. Spellen als go, poker en StarCraft vragen om meer dan het kennen van de regels, meer dan statistiek en kansberekening. Strategie is nodig, bluf en inzicht in het gedrag van andere spelers. Met de huidige rekenkracht is dat al mogelijk. Als we deze ontwikkeling doortrekken, zitten we over een paar jaar naar schatting bijna op dezelfde rekenkracht als het menselijk brein.

Computers gaan net zo snel
Diezelfde exponentiele groei zien we in ook in gewone computers. Zij bereiken na ongeveer 15 jaar het niveau van de snelste supercomputer van vandaag de dag, mobiele telefoons nog eens 9 jaar later. De superkracht van nu hebben we dus over 24 jaar in de broekzak.
Maar wacht, er is meer!
Sinds 2020 wordt AI ingezet voor praktische toepassingen, zoals het voorspellen van de structuur van eiwitten. Dit heeft enorme impact in de medische wereld. Kort daarna volgde

generatieve AI als Midjourney, Dall-E (beeld) en ChatGTP (tekst). AI gaat nu verder dan machines die maar één ding kunnen, zoals schaken of go spelen.
“Als data het nieuw goud is, dan is AI de nieuwe elektriciteit,” zei Andrew Ng in 2016. De oprichter van Deeplearning. Hij bedoelde daarmee dat AI de basistechnologie is voor de vierde grote industriële revolutie. Overdreven? In 2008 deden de top 5 grootste bedrijven ter wereld in olie en goud (tot 2008), maar tegenwoordig staan in de top 10 alleen maar technologiebedrijven.

Impact van AI
De impact van AI op ons leven en werken is afhankelijk van waar ze wordt toegepast. Dat geldt voor de kansen en voor de risico’s. Het risico op het einde van de mensheid zoals we die kennen (zie illustratie) wordt vaak afgedaan als doemdenken, maar negeer het niet, als overheid.



AI en de overheid als wetgever
De AI Verordening van het Europese parlement komt eraan. De AI Act is de eerste wetgeving op het gebied van AI. Ook de AVG zegt het nodige over algoritmische systemen (verbod op onacceptabele bias) maar de AI Act gaat veel verder. Inmiddels praten ook de G7 en de G20 over de noodzaak om wereldwijd afspraken te maken over AI. In Nederland kennen we daarnaast de werkagenda waardegedreven digitaliseren die onze publieke waarden wil borgen en algoritmes reguleren.

Kaders voor de overheid als AI-gebruiker
De overheid gebruikt ook AI en niet altijd goed. Denk aan SyRI, de fraudedetectiemethode van de gemeente Rotterdam en natuurlijk het toeslagenschandaal. Aan de andere kant

ontwikkelde de overheid ook het algoritmeregister, om iedereen inzage te geven in de rekenregels die worden gebruikt en welke data worden verzameld. In het register staan onder andere de berekening huurtoeslag en de kadastrale kaart. Op dit moment zijn 15 van de 118 systemen zelflerende algoritmen, maar wordt vast meer wanneer het register verplicht wordt. Ook wordt gewerkt aan een implementatiekader, waarvan een voorstel dit najaar hopelijk naar de Tweede Kamer gaat.

Hoog, laag, geen risico

De AI Act onderscheidt vier risiconiveaus: onacceptabel, hoog, beperkt en minimaal. Massasurveillance en manipulatie zijn voorbeelden van onacceptabel gebruik van AI. Voor de overheid is met name de lijst met hoog risico belangrijk. Dit zijn toepassingen die een hoog risico vormen voor de gezondheid en veiligheid of de fundamentele rechten van personen. Daar is een lijst aan risicobeheersende regels voor opgesteld, die vergelijkbaar is wat betreft scope en diepgang met de eisen waaraan de financiële sector moet voldoen. Eisen gaan steeds vaker over besluitvorming en governance. Uiteindelijk moet een hele organisatie doordrenkt zijn van het besef dat verantwoord gebruik van AI van iedereen is, niet alleen van de technische afdeling. Inzet van AI heeft een sterk ethisch component: houd altijd het doel voor ogen en maakt duidelijk wat wel en niet acceptabel is.

IAMA-vragenlijsten

Impact Assessment Mensenrechten en Algoritmes (IAMA) is een vragenlijst die helpt bepalen of een algoritme mogelijk schade toebrengt aan mensen. Dat is iets van de hele organisatie. Uit de hele rits rollen en functies die nodig zijn om deze belangrijke vragen te beantwoorden zitten maar twee technische functies. De anderen zijn bijvoorbeeld juristen, bestuur, HR en burgerpanels. Ook hier weer een parallel met de financiële sector: veel mensen worden geacht kennis te hebben van de risico's, dat zou ook zo moeten zijn voor algoritmes.

De IAMA-vragenlijst is geen dwangbuis maar een behulpzaam instrument dat waar nodig is aan te passen voor de eigen organisatie en context. Verantwoordelijkheid voor algoritmes stopt niet bij de aanschaf van systemen. Ook als de afnemer is de overheid verantwoordelijk, samen met de leverancier. Die kan zich niet verstoppen achter excuses van intellectueel eigendom om informatie niet te delen.

Wel of geen ChatGTP?
Op dit moment zijn er nog geen handreikingen voor het gebruik van generatieve AI (zoals ChatGTP) door rijksambtenaren. Voorzichtigheid is troef, zoals voor alle technologieën. Een vuistregel is om ze te gebruiken als inspiratie. Gebruik geen zaakspecifieke of privacygevoelige informatie. Zie ChatGTP als een welbespraakte maar onervaren stagiair. Het klinkt goed, maar klopt lang niet altijd. Overigens wordt niet alle data die ChatGTP gebruikt openbaar gemaakt, maar eigenaar OpenAI behoudt zich het recht voor alle data te mogen gebruiken om het model te trainen. Pas dus op hiermee.

Summer Course AI
Meer weten? Doe mee met de Summer Course AI, een training van twee dagen op 14 en 28 september 2023. Stuur een mail naar i-partnerschap@rijksoverheid.nl

Meer vragen? Mail Joost.
Het Rijks ICT Gilde ontwikkelt zich tot een informeel competence centre op het gebied van AI. Neem dus gerust contact op met vragen, bijvoorbeeld over bias-toetsing in de praktijk en hoe om te gaan met generatieve AI. Het RIG wil overheden actief ondersteunen en bedienen met handige tools om zelf aan de slag te gaan.
Joost.burgt@rijksoverheid.nl

Leestips
Boeken:

- Weapons of Math Destruction (Cathy O'Neil)
- Van Aristoteles tot algoritme (Guido van der Knaap)
- Human Compatible (Stuart Russell)
- The Alignment Problem (Brian Christian)

Blogs:

- Can you make AI fairer than a judge? Play our courtroom algorithm game (Karen Hao Jonathan Stray)
- Everything you've ever wanted to know about machine learning (Cassie Kozyrkov)
- What Is ChatGPT Doing … and Why Does It Work? (Stephen Wolfram)
- The AI Revolution: The Road to Superintelligence (Tim Urban)

Rapporten:

- The Fairness Handbook (Gemeente Amsterdam)
- Discriminatie door risicoprofielen: Een mensenrechtelijk toetsingskader (College voor de Rechten van de Mens)
- Aandacht voor Algoritmes (Algemene Rekenkamer)
- Handreiking Non-discriminatie by design (Ministerie van Binnenlandse Zaken en Koninkrijksrelaties)

Vragen uit de chat

Q. Heb je een eenvoudige uitleg van exponentiële groei?

A. Exponentiële groei betekent dat iets niet met constante hoeveelheid, maar steeds sneller toeneemt naarmate de tijd vordert. Een voorbeeld hiervan is dat de rekenkracht van computers ongeveer iedere anderhalf jaar verdubbelt.

Q. Wat gaat quantum computing betekenen voor AI?

A. In potentie veel. Een van de mogelijke toepassingen van quantum computing is een aanzienlijke versnelling van het oplossen van optimalisatievraagstukken. Machine learning is een optimalisatievraagstuk bij uitstek. In de toekomst is het daarom wellicht mogelijk om grotere modellen te trainen dan met traditionele systemen mogelijk is of om bestaande modellen veel sneller te trainen. Echter, het duurt nog wel een (groot) aantal jaren voordat quantum computing zover is ontwikkeld dat het voor dit soort toepassingen gebruikt kan worden.

Q. AI is een heel breed begrip. Ga je ook nog in op de verschillende soorten AI (narrow, generic, super, etc.)?

A. Artificial Narrow Intelligence (ANI), ook wel zwakke AI genoemd, is een soort AI die is ontworpen om een specifieke taak uit te voeren, zoals spraakherkenning of het spelen van een schaakspel. Artificial General Intelligence (AGI), ook wel sterke AI genoemd, is een soort AI die dezelfde cognitieve vaardigheden heeft als een mens. Het kan dus elke intellectuele taak uitvoeren die een mens kan doen. Artificial Super Intelligence (ASI) verwijst naar een soort AI die niet alleen alles kan wat een mens kan, maar ook nog eens beter is dan mensen in alle opzichten: het kan sneller denken, heeft meer kennis, kan beter problemen oplossen, en ga zo maar door.

Q. Wat is het doel van het Algoritmeregister?

A. Door middel van het Algoritmeregister kunnen burgers, maatschappelijke organisaties en media de overheid kritisch volgen en bevragen, en controleren of zij zich aan de regels houdt. Het Algoritmeregister kan op die manier een belangrijke bijdrage leveren aan het beter uitlegbaar maken van de toepassing en uitkomst van algoritmes.

Q. Hoe ver is het implementatiekader?
A. Er is nog geen officiële update, maar naar verwachting wordt er na de zomer een eerste (voortgangs)rapportage aan de Tweede Kamer opgeleverd.

Q. Zijn het IAMA en andere verplichtingen nog praktisch uitvoerbaar?
A. Het is begrijpelijk dat al deze verplichtingen onuitvoerbaar lijken, omdat ze nu plots bovenop al bestaande verplichtingen worden gestapeld. Anderzijds is het ook een kwestie van deze zaken eigen maken en inbedden in de normale bedrijfsvoering. De financiële sector kent een mooie tegeltjeswijsheid die hier (met name met alle schandalen uit het recente verleden) ook mooi opgaat: *"If you worry about the cost of compliance, consider the cost of non-compliance"*. (Als je je zorgen maakt over de kosten van het voldoen aan de regels, denk dan eens aan de kosten van het niet voldoen).

Q. Wordt er al nagedacht over hoe managers , politici, etc. bij de overheid worden getraind over AI en de vereisten?
A. Daar wordt ongetwijfeld hier en daar al over nagedacht, maar hierover is nog geen handreiking/richtlijn beschikbaar voor zover ik weet. Dat gezegd hebbende is dit wel iets waar organisaties zelf over na zouden moeten denken en waar wellicht ook vanuit een centrale rol ondersteuning voor kan worden geboden.

Q. Transparantie voor de burger? Gaat toch primair om uitlegbaarheid van besluiten cf beginselen van behoorlijk bestuur?
A. De verplichtingen en beperkingen van transparantie zijn een open discussie. Het kunnen uitleggen van besluiten is een minimale vereiste, maar waarschijnlijk wil je ook enige transparantie kunnen bieden over wat voor een type systemen (algoritmes) er worden gebruikt, hoe deze (op hoofdlijnen) werken, welke data hiervoor wordt gebruikt, en verantwoording over de prestaties en mogelijke bias ervan.

Q. besteedt IAMA ook aandacht aan de zuiverheid, kwaliteit van de data? Was een goede ppt tijdens een vorige IIR dag
A. Jazeker! In stap 2A "Data(input)" komt dit uitgebreid aan bod.

Q. Hoe ga je om met off the shelf producten waarin vormen van AI wordt toegepast. Dan wordt het lastig om veel van deze vragen te beantwoorden.
A. Als gebruiker van AI ben je volgens de aankomende AI Verordening primair verantwoordelijk voor het correct functioneren ervan. Hoewel het gebruik van toekomstige keurmerken en standaarden hierbij kan helpen zul je te allen tijde inzicht moeten hebben in en (direct of indirect) controle moeten kunnen uitoefenen op het juist functioneren van deze producten. Dit betekent ook dat aanbieders hiervan voldoende informatie moeten kunnen verstrekken om deze verantwoordelijkheid te kunnen nemen.

Q. Wat moet een rijksambtenaar in het hier en nu als stelregel hanteren? Bv. gebruik wel/niet chatGPT?
A. Zie ChatGPT als een zeer welbespraakte maar grasgroene stagiair. En bedenk dat deze stagiair soms ook per ongeluk jouw vragen op social media zet. Dus denk goed na wat je aan ChatGPT vraagt, en controleer altijd(!) de output goed (jij bent daar zelf uiteindelijk verantwoordelijk voor).

Q. Onacceptabele bias? Is er dan ook acceptabele bias?
A. Het compleet uitbannen van bias is in de praktijk een illusie. Bovendien zijn er veel verschillende uitingsvormen van bias, en zal het reduceren van een bepaalde vorm van bias vaak de toename van een andere vorm van bias hebben. Het gaat er dus om te bepalen wat de meest relevante/belangrijkste vorm van bias is die je wilt beperken (bijv. het verschil tussen bepaalde demografische groepen in het risico dat een onschuldig iemand als hoog-risico wordt aangeduid). Wanneer een bepaalde groep een tweemaal zo hoge foutmarge kent als bepaalde andere groepen is er hoogstwaarschijnlijk sprake van onacceptabele bias. Wanneer dit verschil tussen groepen echter bijvoorbeeld maar 5% is dan zou deze mate van bias (afhankelijk van de context) mogelijk prima te verantwoorden zijn.

Q. Zijn bedrijven verplicht om een IAMA (oid) te gaan gebruiken? Wat zou een motivatie voor bedrijven zijn om dat WEL te gaan doen?
A. Dat zal moeten blijken uit de definitieve versie van de AI Verordening, maar het Europees Parlement heeft hiervoor wel een amendement ingediend. Even afwachten dus. En zelfs wanneer dit niet verplicht zou worden zou het voor bepaalde bedrijven nog steeds een goed idee kunnen zijn, met name wanneer deze bedrijven ook overheden bedienen met hun diensten en producten.