



aan DB - Directie Beraad CBS, DT DRI
van Beleidsstaf DRI

onderwerp Evaluatie beleidskader AI - DB versie 2 - Definitieve versie
datum Februari 2024

Het gebruik van algoritmes en het beleidskader AI

De overheid gebruikt steeds vaker algoritmes. Bijvoorbeeld om burgers snel te helpen bij de aanvraag van voorzieningen, of bij de aanpak van maatschappelijke opgaven als de energietransitie. Geregeld ontvangen wij, als CBS, hulpvragen van overheidsorganisaties over het gebruik van algoritmes. Bijvoorbeeld: Zit er een bias in de dataset (die als input dient voor een AI-model)? Naar aanleiding hiervan heeft het CBS het beleidskader Kunstmatige Intelligentie (hierna: AI) opgesteld. Het beleidskader AI dient als hulpmiddel om te bepalen welke dienstverlening het CBS wel en niet levert.

Deze notitie is een evaluatie van het beleid 1,5 jaar na invoering. Hoe goed voorziet het beleid in het beoogde doel? Welk AI gerelateerd beleid is er nog meer bij het CBS en welk beleid ontbreekt nog? Wat gebeurt er om ons heen qua AI-beleid?

We starten met een samenvatting van het huidige AI beleidskader, vervolgens bespreken we de opzet van de evaluatie. Deze bestaat uit twee delen: (1) Bevindingen en aanbevelingen afwegingskader verzoeken inzet algoritmes en (2) Bevindingen en aanbevelingen AI beleid. In de bijlage is een contextsectie over AI toepassingen in het statistische proces opgenomen.

Samenvatting beleidskader AI

In het eerste deel van het kader wordt AI gedefinieerd en is er een overzicht van de verschillende lijnen waarlangs het CBS met AI te maken heeft. Vervolgens stelt het kader dat de voorwaarden voor het CBS om statistiek te produceren ook gelden voor het gebruik van AI. Zo dienen algoritmes, net als bijvoorbeeld regressiemodellen, controleerbaar en reproduceerbaar te zijn. Indirect wordt hiermee aangegeven dat aanvullend beleid (nog) niet nodig is. Wat wel nodig is, is een additioneel kader dat focust op de hulpvragen rondom algoritmes en duidelijk maakt welke AI werkzaamheden het CBS kan en mag uitvoeren. Vervolgens geeft het beleidsdocument een dergelijk afwegingskader. We lichten deze hieronder toe.

Afwegingskader verzoeken inzet algoritmes

Het kader onderscheidt vier type hulpvragen. Binnen elke typering wordt vervolgens categorisch aangegeven welke dienstverlening we wel en niet leveren. Elke hulpvraag die we ontvangen komt, in theorie, overeen met een van de geschetste situaties. Zo kan er een passende afweging worden gemaakt of we van dienst kunnen en mogen zijn. In enkele van deze situaties is dit niet direct duidelijk en is de beslissing contextafhankelijk.

Ter verduidelijking: Advies (*type hulpvraag*) over een model voor statistisch gebruik (*categorie*) geven we wel. Maar zodra het een voorspellend model (en daarmee een andere *categorie*) betreft, is de beslissing contextafhankelijk. Er worden vervolgens concrete vervolgstappen genoemd om alsnog tot een goede afweging te komen. De belangrijkste stap daarvan is het voorleggen van de casus aan team juridisch en de ethische commissie.



Opzet van de evaluatie

Aanpak

In deze evaluatie kijken we in eerste instantie in hoeverre het afwegingskader bruikbaar is geweest bij de afhandeling van de hulpvragen over het gebruik van algoritmes. We gaan na in hoeverre de organisatie bekend is met het beleidskader en in welke mate het zijn doel bereikt.

Vervolgens verbreden we onze blik en kijken we naar de huidige tendens in AI toepassingen, zowel voor (innovatief) onderzoek als beleid. Het gebruik van kunstmatige intelligentie als hulpmiddel, ook in het statistisch proces, is volop in ontwikkeling. Waar houden we ons op dit moment mee bezig? Welk AI gerelateerd beleid is er nog meer bij het CBS en welk beleid ontbreekt nog? Wat is op hoofdlijnen het rijksbrede of Europese AI-beleid? Sluiten we hier goed op aan of kunnen we dat verbeteren?

Informatie verzamelen

We hebben informatie opgehaald bij ongeveer 20 medewerkers uit verschillende vakgebieden binnen het CBS. Informatie over het internationale en rijksbrede AI beleid is voornamelijk verkregen via deskresearch. Er is gesproken met:

1. Statistisch onderzoekers, zij ontvangen de hulpvragen en leiden de AI-gerelateerde projecten, zoals Risico op Armoede, Vaardig met Vaardigheden of Eerlijke Algoritme;
2. medewerkers die zich bezighouden met innovatie, data science en methodologie; en
3. beleidsadviseurs en managers.

1. Bevindingen en aanbevelingen afwegingskader verzoeken inzet algoritmes

Medewerkers die te maken hebben met hulpvragen over het gebruik van algoritmes zijn goed bekend met het beleidskader AI. Het afwegingskader wordt ingezet en indien nodig is aan de ethische commissie gevraagd om zich over de vraag te buigen. Wel zijn er meerdere opmerkingen over het gebruikersgemak van het afwegingskader. De structuur van het beleidskader maakt het lastig om snel te vinden wat je zoekt, ook de hoeveelheid tabellen lijken de afweging juist complexer te maken dan nodig. De uitgewerkte voorbeelden aan het eind worden heel nuttig bevonden. De gang naar de ethische commissie is niet voor iedereen direct logisch, maar wordt wel ervaren als een goed werkende oplossing om tot een beslissing te komen.

Meerdere projectleiders geven aan de categorie dat het CBS zelf een algoritme inzet, te missen. Andere geïnterviewden geven aan dat we met elkaar moeten bespreken wat we wel en niet als een algoritme beschouwen. Een eigenschap van kunstmatige intelligentie is o.a. het zelflerende aspect. Het gevaar bestaat om een stukje programmeertaal te snel en daarmee onterecht als een algoritme te zien. Algoritmes die een handmatige werkinstructie automatiseren zijn bijvoorbeeld uitgezonderd van het algoritme register.

Aan de hand van bovenstaande bevindingen komen we tot de volgende aanbevelingen:

- Maak een zelfstandig beleid voor het gebruik van algoritmes met daarin als onderdeel het afwegingskader voor de hulpvragen.
- Sluit hierbij zoveel mogelijk aan bij het rijksbrede beleid of leg uit waarom het CBS afwijkt.



- Start vooraf met de vraag wat we nu precies verstaan onder een algoritme. Wat zien we als algoritme en wat als geautomatiseerde rekenregels. Een gezamenlijk herijkt startpunt is cruciaal voor het toekomstige beleid in dit dossier.
- Stel, in acht neming van de vorige aanbeveling, aanvullend richtlijnen op over de eigen inzet van algoritme in het statistische en werkproces. Maak hierbij geen onderscheid tussen regulier of aanvullende onderzoek (ASD).
- Neem in de update een link op naar het algoritme register (zie onder kopje 'AI toepassing en beleid CBS'). Hiermee heb je een volledig beleidskader voor een definitie en het gebruik van algoritmes.
- Hernoem het nieuwe beleidsdocument zodat de scope direct duidelijk is.
- Ga na of het afwegingskader handzamer gepresenteerd kan worden, door bijvoorbeeld te bekijken of de verschillende tabellen niet ineen geschoven kunnen worden.
- Evaluer het beleid hoogfrequent om actueel te blijven. Evaluer dan ook of de ethische commissie nog steeds het meest voor de hand liggende adviesorgaan is om te raadplegen.

2. Bevindingen en aanbevelingen AI beleid

Eventuele toekomstige aanvullingen op het AI beleid van het CBS dienen in lijn te zijn met het Rijksbrede AI beleid. Voordat we tot de aanbevelingen overgaan schetsen we hieronder een globaal beeld van het AI beleid en de AI strategie van opeenvolgend Europa, de Nederlandse overheid en het CBS. Voor meer informatie zijn er verwijzingen opgenomen.

Europese AI strategie en AI verordening

De Europese Unie (EU) heeft als doel om een leidende rol te spelen in de ontwikkeling en toepassing van AI. Ze streeft ernaar om investeringen in AI-onderzoek en – innovatie te vergoten. Tegelijkertijd wil ze veiligheid en acceptatie waarborgen. Er worden ethische AI-richtlijnen en normen ontwikkeld. In het Europese Parlement zijn de standpunten voor de onderhandelingen van een AI-verordening¹ vastgesteld. Hiermee zijn we een stap dichterbij de eerste uitgebreide AI-wet ter wereld.

Digitaliseringsbeleid Nederlandse overheid en AI

Ook in Nederland is het uitgangspunt dat de Nederlandse samenleving optimaal gebruik gaat maken van AI: Het ministerie van Binnenlandse Zaken en Koninkrijksrelaties (BZK) bevordert een verantwoorde inzet van AI. In de Werkagenda Waardengedreven Digitaliseren² is uitgewerkt hoe dit kabinet de digitale transitie in goede banen wil leiden. Het gebruik van AI neemt hierin een prominente plek³. Twee onderwerpen die hierin op dit moment voor het CBS van belang zijn, zijn de standpunten over gebruik van algoritmes en generatieve AI. Onverantwoord gebruik van algoritmes heeft in bijvoorbeeld de toeslagenaffaire laten zien welke verstreckende gevolgen dit kan hebben. Een van de voorgestelde maatregelen is het opstellen van een algoritmeregister. Deze verplichting voor ministeries (nog niet voor ZBO's)

¹ Zie nieuwsitem Europees Parlement: [AI-verordening: eerste regels voor artificiële intelligentie](#)

² Zie Digitaliseringsbeleid: [Werkagenda Waardengedreven Digitaliseren](#)

³ Zie Digitaliseringsbeleid: [AI \(Artificiële Intelligentie\)](#)



is een voorloper van de AI verordening. Zie voor meer informatie de volledige kamerbrief⁴ 'algoritme reguleren' van juli 2023.

Voor het gebruik van generatieve AI heeft de staatssecretaris een voorlopig standpunt ingenomen. Summier samengevat staat hierin dat zowel de kansen als de risico's van generatieve AI-toepassingen worden gezien. Vervolgens wordt het gebruik van deze toepassingen puntsgewijs ingekaderd. Waarvan mogelijk het belangrijkste standpunt is dat niet gecontracteerde toepassingen zoals ChatGPT, BRAD en Midjourney over het algemeen niet aantoonbaar aan de geldende privacy- en auteursrechtelijke wetgeving voldoen. Zodoende is het gebruik hiervan door Rijksorganisaties (of in opdracht daarvan) in beginsel niet toegestaan. Maar lees vooral de gehele kamerbrief⁵ voor een het volledige standpunt.

Ten slotte is er door meerdere departementen gezamenlijk, gewerkt aan een overheidsbrede visie op de impact van generatieve AI. Deze schetst een integraal beeld van de impact die deze technologie heeft op onze maatschappij en de uitdagingen en mogelijkheden die hiermee zijn gemoeid en wordt medio januari gepresenteerd.

AI toepassing en beleid CBS

Naast het beleidskader AI hebben we een beleidsstuk voor de werkwijze bij synthetische data⁶. Beide vastgesteld in het najaar van 2022. Recent is er in het DB gesproken over het algoritme register. Het CBS steunt het bijhouden ervan, waarbij gekozen is om te starten met een eigen register⁷. Het beleidskader AI stelt dat AI toepassingen dienen te voldoen aan dezelfde wet en regelgeving als de statistiekproductie. Er zijn geen aanvullende kaders of handreikingen over het gebruik van generatieve AI toepassingen. In de afgelopen periode is er echter steeds meer aandacht voor onderzoeken naar de mogelijkheden van AI technieken in onze werkprocessen.

Zo zijn er meerdere innovatieve werkgroepen op het gebied van data science en AI. Zoals de werkgroep synthetische data, de werkgroep Machine Learning, de werkgroep Large Language Models en de werkgroep webscraping. Zij doen allen innovatief onderzoek naar, en toepassing van, AI als hulpmiddel in het statistische proces. Ze staan in verbinding met elkaar via de data science community. Zie onderstaand contextkader voor meer uitleg over deze AI toepassingen.

We maken bijvoorbeeld gebruik van algoritmes in het ASD-onderzoek 'proactieve dienstverlening armoede'. In (de voorlopers van) dit project is met behulp van een AI-model onderzocht welke kenmerken samenhangen met een verhoogde kans om arm te blijven of worden. Met als doel dat gemeenten dit kunnen inzetten voor effectief armoedebelid.

Daarnaast lopen er bij CCN verschillende initiatieven rondom de inzet van AI, met als doel om de CBS content beter vindbaar te maken. Denk hierbij aan de mogelijkheden van de zoekmachine van de CBS website of het generen van goede 'keywords' (metadata) aan onze artikelen.

⁴ Zie Digitaliseringsbeleid: [Kamerbrief verzamelbrief algoritmes reguleren.pdf](#)

⁵ Zie Digitaliseringsbeleid: [Kamerbrief-over-voorlopig-standpunt-voor-rijksorganisaties-bij-het-gebruik-van-generatieve-ai](#)

⁶ Zie DB bespreking [Beleid Synthetische Data DB.pdf](#)

⁷ Zie DB besluit 29 augustus 2023: [DB-besluiten 2023-0829.pdf](#)



Bevindingen AI initiatieven en AI beleid

Samengevat wordt er zowel vanuit Europa als door het kabinet het gebruik van AI bevorderd. Landelijke of Europese richtlijnen en verordeningen moeten ervoor (gaan) zorgen dat dit transparant en verantwoord gebeurt. In de werkgroepen die vallen onder de data science community wordt gekeken naar ideeën voor en de mogelijkheden van concrete toepassing van (generatieve) AI. De focus van deze werkgroepen ligt vooral op (on)mogelijkheden van de methoden en technieken. Het is belangrijk om deze werkgroepen de ruimte en de middelen te geven om te innoveren, tegelijkertijd is het goed om bewust te zijn van de risico's van deze toepassingen. Hierbij zouden we waarschijnlijk meer gebruik kunnen maken van de hulpmiddelen die door de rijksoverheid al op dit gebied ontwikkeld zijn. Om hierin stappen te kunnen zetten, worden de volgende aanbevelingen gedaan:

- Centraliseer beschikbare informatie, maak het makkelijk om goed geïnformeerd te zijn. Concreet: Breng beleid, afwegingskader en besluiten rondom AI samen op een plek, bijvoorbeeld op een wikipagina op intranet. Werkgroepen kunnen deze onderwerpen zo ook efficiënter agenderen.
- Ontwikkel bovenstaand informatiepunt door naar een soort van 'toolbox' of 'framework' voor het gebruik van algoritmes en generatieve AI toepassingen. Richt het flexibel, en daarmee toekomstbestendig, in. Documenten van verschillende volwassenheid en doeleinden kunnen elkaar aanvullen. Zoals een overzicht van rijksbrede besluiten en hulpmiddelen op het onderwerp, of een checklist om bewustwording rondom risico's bij het gebruik van AI toepassingen te vergroten. Het inzichtelijk maken van risico's van (generatieve) AI en hoe hiermee om te gaan is een belangrijk onderdeel van het 'framework'.
- Zorg voor meer coördinatie. Zorg dat er zicht is op alle nieuwe en lopende verschillende initiatieven van AI. De data science community zou deze rol kunnen vervullen. Hiermee creëer je ook een centraal aanspreekpunt om te toetsen of innovatieve ideeën (nog) in lijn zijn met het rijksbrede of CBS beleid.
- Zorg bij de opvolging van bovenstaande aanbeveling voor een beheersbare en toetsbare inrichting (governance). Maak hierbij een directe link tussen de data science community en CBS bestuurders: Benoem een dossierhouder datascience of AI in het DB. Het voorstel is om dit de hoofddirecteur DRI te laten zijn.

Bijlage: Contextsectie

Huidige AI toepassingen en technieken in het statistische proces

Er zijn verschillende definities van kunstmatige intelligentie (AI). In de kern gaat het eigenlijk altijd over de simulatie van menselijke intelligentieprocessen door machines, met name computers. In het beleidskader wordt de definitie zoals omschreven door de Nederlandse AI Coalitie (NLAI) gehanteerd: "Het vermogen van een systeem om externe gegevens correct te interpreteren, om te leren van deze gegevens, en om deze lessen te gebruiken om specifieke doelen en taken te verwezenlijken via flexibele aanpassing".

Hieronder volgt een overzicht van de onderwerpen binnen data science waar op dit moment bij het CBS aandacht voor is. Om de evaluatie van het beleidskader in de juiste context te plaatsen is het handig om de benadering van deze onderwerpen te kennen zoals de auteur deze ziet.



Algoritme en AI

Een algoritme is een stapsgewijze reeks instructies of regels die wordt gebruikt om een bepaald probleem op te lossen of een taak uit te voeren. Algoritmes vormen de logische en wiskundige basis van AI, waardoor computers taken kunnen leren en uitvoeren door patronen te identificeren en aanpassingen te maken op basis van de data waarmee ze worden gevoed. In de context van AI spelen algoritmes een cruciale rol: Zij vormen de bouwstenen waarmee AI-systemen worden gebouwd en verbeterd.

Generatieve AI

Generatieve AI verwijst naar kunstmatige intelligentie-systemen die in staat zijn om nieuwe, originele inhoud te produceren die lijkt op iets wat door mensen is gemaakt. Deze systemen hebben geleerd patronen en structuren te herkennen in gegevens, en op basis daarvan kunnen ze nieuwe gegevens genereren die vergelijkbaar zijn met de inputgegevens. Recent bekend voorbeelden zijn muziekcompositie(-algoritme) en beeldgeneratoren. Maar ook tekstgeneratoren zoals GPT-3. GPT-3 is het model waarop **ChatGPT** is gebaseerd.

Generatieve AI kan binnen de statistiek worden toegepast in het genereren van nieuwe datasets die vergelijkbaar zijn met een bestaande dataset. Een bekend voorbeeld hiervan is **synthetische data**. Deze synthetische data kan weer worden gebruikt om algoritmes te trainen.

Een ander voorbeeld zijn **Large Language Models** (LLM's). Dit zijn geavanceerde kunstmatige intelligentiesystemen die zijn getraind op enorme hoeveelheden tekstgegevens. Deze modellen zijn in staat om natuurlijke taal te begrijpen en te genereren. Ze worden bijvoorbeeld gebruikt voor taken zoals samenvatten of chatbots.

Machine learning

Machine learning (**ML**) is een belangrijke tak van AI en is gebaseerd op algoritmes. Bij ML worden algoritmes gebruikt om modellen te trainen op basis van data, waarbij het algoritme zelf aanpassingen maakt om betere voorspellingen te doen.

Een voorbeeld hiervan waar het CBS op dit moment onderzoek naar doet is het schatten van de belangrijkste vorm van toegevoegde waarde van een bedrijf dat niet in de steekproef is opgenomen met behulp van een trainingsset die is samengesteld o.b.v. bedrijven die wel in de steekproef zitten.

Webscraping

Webscraping is een techniek binnen AI welke kan worden gebruikt om data te verzamelen. Deze techniek wordt bijvoorbeeld toegepast in het project Vaardig met Vaardigheden (competent NL).