

Samenvatting

Data & ethiek in de publieke sector: van principes naar praktijk

Het probleem

De nieuwe technologieën op het gebied van data zijn een uitstekend en belangrijk middel voor overheidstaken. Echter, de transitie van het digitaliseren van overheidsfuncties naar een afhankelijkheid van digitale data voor het functioneren van een overheid, brengt compleet nieuwe uitdagingen met zich mee op het gebied van ethiek. De ethiek van de publieke sector spitst zich van origine meer toe op het gedrag van ambtenaren dan op het gedrag van publieke instellingen naar het publiek toe. Dit laatste is zeer relevant in het kader van de uitdagingen van big data. De **ontwikkelingen in de private sector illustreren** dat ondanks het bestaan van een systeem van 'checks en balances' het nog steeds gebruikelijk is dat data gebruikt worden op een manier die, ingrijpt op privacy, discrimineert tegenover minderheden en beschermde groepen, of schadelijk is op een andere manier. De publieke sector heeft met veel van deze risico factoren te maken en kampt daarnaast met een gebrek aan transparantie.

Het gebruik van digitale data door de Nederlandse overheid is uitgebreid, maar voor veel departementen bevindt het gebruik van geavanceerde data-analyse technieken zich nog in een relatief vroeg stadium. Er zijn echter signalen dat de overheid op het gebied van 'data science' in toenemende mate gebruik maakt van externe capaciteit, dat wil zeggen een samenwerking tussen de overheid, wetenschap, en partners uit de private sector. Dit is bijvoorbeeld het geval op het gebied van 'predictive policing', de aanpak van belastingfraude en het detecteren van fraude met publieke diensten en toeslagen. Op die gebieden worden standaard grootschalige gegevensverzamelingen gebruikt en werkt de overheid vaak samen met bedrijven gespecialiseerd in technologie die de nodige expertise en vaardigheden aanbieden. Bovendien is het waarschijnlijk dat veel van de big data die mogelijkheden bieden voor echte innovatie voor overheidstaken in de nabije toekomst afkomstig zijn van **bronnen buitenom de overheid**, zoals bijvoorbeeld sociale media of burger initiatieven. Deze samenwerkingsverbanden en diversiteit aan bronnen zullen het in kaart brengen en identificeren van de verschillende typen data analyse die worden gebruikt bemoeilijken. Ook wordt het lastig te detecteren wanneer er sprake is van een situatie die extra oplettendheid vereist in het kader van aspecten die problematisch zijn of kunnen zijn op het gebied van ethisch verantwoord gebruik van data.

We constateren dat er zich een **gat voordoet in het institutioneel toezicht** in de publieke sector op het gebied van ethiek en data: het redeneringsproces dat noodzakelijk is om juridische regels in verschillende situaties toe te passen, risico's te beoordelen en te evalueren hoe verschillende belangen en rechten te balanceren. De huidige visie op toezicht, gebaseerd op economische efficiëntie en naleving van gegevensbeschermingswetgeving, stimuleert een visie op verantwoordelijkheid die is gericht op de korte termijn en naleving van de wet en die ontoereikend is op de lange termijn voor de democratische en sociale implicaties van data analyse als een gereedschap van de overheid. Momenteel ontbreekt een ethisch kader dat een brug kan slaan tussen de twee soorten regels en dat een zinvolle en praktische handreiking kan bieden wanneer interne en externe partners samenwerken aan projecten terwijl zij uiteenlopende ideeën hebben over wat wel of niet toelaatbaar is op het gebied van experimenteren met data.

Samenvatting

Modellen, principes en praktijk

In dit rapport voeren wij een model aan voor toezicht op ethiek in de context van data gebruik door de overheid, waarin verantwoording centraal staat en dat het idee van verantwoord data gebruik als basis heeft om een systeem van checks en balances te creëren. Een dergelijk kader omvat ook maatregelen om transparantie te creëren en is tot op zekere hoogte gericht op het publiek. We identificeren verschillende modellen voor institutioneel toezicht op de data-gedreven overheid:

1. **Extern (ad hoc) toezicht**, door middel van klokkenluiders of ‘waakhonden’ van de burgermaatschappij;
2. **Intern toezicht op gerechtelijk niveau**, zoals bij inlichtingendiensten die geadviseerd kunnen worden door een gerechtelijk orgaan met betrekking tot interne regels;
3. **Onafhankelijk overheidstoezicht**, zoals bij het ‘National Statistician’s Data Ethics Advisory Committee’ uit het Verenigd Koninkrijk;
4. **Intern overzicht op het niveau van gemeenten/departementen**, zoals de Commissie Persoonsgegevens Amsterdam;
5. **Externe controle op het datagebruik van de overheid**, waarvoor een nieuw overheidsorgaan zal moeten worden gecreëerd dat controles uitvoert in alle overheidssectoren, mogelijk gebaseerd op het model van EUROSAI (the European Organisation of Supreme Audit Institutions);
6. **Projecten publiekelijk toegankelijk maken**, bijvoorbeeld door het plaatsen van data of processen om data in kaart te brengen op websites die toegankelijk zijn voor het publiek.

Vanuit internationaal perspectief gezien zijn er verschillende modellen: we richten ons op het Verenigd Koninkrijk en Frankrijk waar zich vergelijkbare problemen voordoen als in Nederland, maar waar voor verschillend manieren van aanpak wordt gekozen. Het Verenigd Koninkrijk heeft zes principes vastgesteld voor data gebruik op het overheidsniveau: 1) Begin vanuit een duidelijke behoefte van de gebruikers (van een systeem) en vanuit het publiek belang; 2) Gebruik de minimale hoeveelheid data om het publiek belang te bereiken; 3) Creëer robuuste ‘data science’ modellen; 4) Ben bedacht op de publieke opinie; 5) Wees zo open en verantwoordelijk als mogelijk; 6) En houdt data veilig. De kracht -en zwakte- van deze aanpak is dat het functioneert op een **hoog niveau van abstractie** en niet specifiek gericht is op een bepaalde praktijk of manier van gebruik van data.

In tegenstelling tot de benadering van het Verenigd Koninkrijk heeft de Franse overheid een wet (Loi Numerique) met drie belangrijke elementen: ten eerste stimuleert het gebruik van data door de overheid en burgers in het publiek belang, ook als middel voor economische groei; ten tweede creëert het bescherming van platformen; ten derde bevat het een agenda voor het creëren van universele toegang tot digitale technologie. De Franse aanpak biedt verscheidene mogelijkheden voor bescherming die ten doel heeft om gelijkheid in het digitale domein te bevorderen. Het voordeel van deze aanpak is dat het **regels geeft die gehandhaafd kunnen worden** en die datafificatie bevorderen en een systeem van checks en balances in het leven roepen. Echter, deze wet biedt geen richtlijnen hoe **maatschappelijke risico’s van data science te identificeren of aan te pakken** en zal daarom nader worden gevormd door rechterlijke uitspraken en zal niet zozeer handvaten bieden voor het modelleren van overheidstaken op basis van nieuwe data bronnen.

Deze verschillende voorbeelden roepen de vraag op wat voor type toezicht zowel verantwoordelijkheid –waar actoren *ex ante* nadenken over hun rol, de effecten en het toekomstig gebruik van de relevante dataset- als verantwoording –het *ex-post* structureren van reacties op problemen- creëert. Momenteel

Samenvatting

wordt ethiek omtrent data ad hoc opgepakt wanneer zich problemen voordoen. De vraag is wat voor type toezicht ethische kwesties omtrent data systematisch kan aanpakken, op een manier die relevant is voor de uitdagingen die big data oproept.

Voorstel: een model voor ethisch gebruik van big data binnen de overheid

Wij stellen een model voor dat bestaat uit drie elementen:

1. Basis principes

De eerste stap is de kernwaarden die we willen beschermen en bevorderen in data science door de overheid te definiëren. Hier moet een keuze worden gemaakt tussen het opstellen van concrete *regels* of algemene *principes* en of de te beschermen waarden *algemeen* of *domein specifiek* zouden moeten zijn.

2. Kader voor verantwoording: de 'control loop'

Voor de kern van het kader voor verantwoording verwijzen wij naar het model van het A4Cloud project, dat verder gaat dan alleen naleving van de wet om ethische reflectie en respect voor de informatie die wordt gebruikt in analyses te bewerkstelligen. In dit model moet een organisatie die verantwoording aflegt 1) uiteenzetten wat deze doet; 2) uitvoeren wat uiteengezet is; 3) monitoren hoe deze handelt; 4) verschillen tussen wat deze uiteen heeft gezet te doen (normen) en wat daadwerkelijk gedaan wordt (gedrag) remediëren, en 6) uitleg geven en de acties rechtvaardigen.

3. Handhaving en toezicht

Nadat de relevante normen (bijvoorbeeld, gegevensbescherming en domein specifieke wetgeving) en principes (zoals gedefinieerd in laag 1) uiteen zijn gezet, tezamen met andere belangrijke kwesties zoals het voorkomen van oneigenlijk gebruik van bevoegdheden door het terug zien op activiteiten in een later stadium, belanden we bij de structuren die de kern vormen van de verantwoordingslaag. In deze laag moeten keuzes worden gemaakt betreffende **handhaving en toezicht**. Moet de omvang van het verslaggeven *uitgebreid* zijn (bijvoorbeeld een impact assessment) of *marginaal* (beschrijving van de doelen en methoden)? Moet de verslaggeving *intern* of *extern* gericht zijn? Moet een handhavingsorgaan *intern* of *extern*; *ad hoc* (bijvoorbeeld voor een specifiek project) of *structureel* (bijvoorbeeld een Publieke Data Science Autoriteit die data science projecten in de publieke sector overziet) zijn? De afstand tot de project organisatie is ook een belangrijke keuze: moet handhaving en toezicht plaatsvinden op het niveau van de sector, een ministerie, een laag van de overheid, of op een ander niveau? Deze keuze voor een niveau en positie heeft verdere implicaties voor de vaardigheden en kennis die van een dergelijk orgaan vereist zijn.

Samenvatting

Aanbeveling: aanpak gericht op sectorniveau

Voor Nederland is een model noodzakelijk dat overzicht biedt voor zowel ministeries en gedecentraliseerde overheden (zoals de 388 gemeenten) en dat kan meegroeien met het data gebruik van de overheid. Op basis van deze criteria lijkt de meest geschikte oplossing te zijn om op **centraal niveau principes** vast te stellen die vervolgens toegepast worden door middel van **kaders op sectorniveau**. Door middel van toezicht op sectorniveau kan de control loop, zoals hierboven beschreven, functioneren op verschillende niveaus en mogelijk in connectie tot bestaande structuren, zoals het CBS dat nu al een trefpunt vormt voor operationele informatie met betrekking tot activiteiten met data tussen gemeenten.

Een uitdaging is om een te grote werklast bij één orgaan, in termen van werkdruk door het uitoefenen van toezicht, te voorkomen. Een dergelijke werkdruk zou het onmogelijk maken om het volledige scala aan ethische kwesties omtrent datagebruik te adresseren. In plaats hiervan maakt het gebruik van **meerdere structurele opties op sectorniveau** het mogelijk om zowel horizontale verantwoording (bewustzijn van principes, ervoor zorgen dat afdelingen intern rapporteren en aan het direct daarboven liggende management) als toezicht van bovenaf (via wetgeving, toezichtsorganen op een hoger niveau, externe organen en/of verantwoording aan de burgermaatschappij) te bewerkstelligen. Er is een voordeel verbonden aan het implementeren van een **publieksgericht element** van toezicht, of dit nu via openbare rapportages gebeurt of door middel van de burgermaatschappij, namelijk dat dit binnen de maatschappij legitimiteit voor de data projecten creëert en kan bijdragen aan van het debat over bewustzijn van het gebruik van data door de overheid. De ideale uitkomst van deze structurele benadering is om ethiek omtrent data, vergelijkbaar met Privacy by Design, te veranderen in een 'mindset' en doorlopend proces in plaats van een taak die voldaan of nageleefd moet worden. Hoewel het tijd kost om een verandering in mindset te bewerkstelligen is het proces van onderhandelen, discussiëren en definiëren van ethische principes, essentieel om ervoor te zorgen dat verantwoording legitiem is en te handhaven is.