

Deskundigenbijeenkomst data & analytics

Midden 2015 las ik de *Investeringsagenda* van de Belastingdienst. De *Investeringsagenda* bood een heldere analyse en visie voor de aanpak van de modernisering van de belastingdienst. De probleemanalyse maakte helder hoe de belastingdienst haar eigen te grote hoeveelheid werk creëert. Hoe ze vast zit in een vicieuze cirkel die doorbroken moet worden. 'Met 700.000 brieven per week is de communicatie met belastingplichtigen old school, sluit niet meer aan bij de digitale tijd, is 'aanbodgericht' en sluit niet aan bij de beleving van belastingplichtigen (het deelproces van de belastingdienst is leidend). Ook organiseert de belastingdienst zelf onhandige pieken in het eigen werk en juridificeert ze te veel.' De belastingdienst wil dit omdraaien en met data-analyse en ICT haar professionals ondersteunen om veel slimmer te werken.

Wat mij aansprak in deze visie uit 2015 is dat data science en organisatieverandering aan elkaar verbonden worden. Data science draait niet om meer dan 'een leuk inzicht'. Het transformeert de werkwijze van organisaties. Uiteindelijk doel is om met data science meer publieke waarde te creëren.

Inmiddels zijn we ruim twee jaar verder. Het vakgebied data science ontwikkelt zich net als de maatschappelijk discussie over de toepassing daarvan. Ter voorbereiding op de deskundigenbijeenkomst data & analytics is ons gevraagd enkele punten naar voren te brengen. Punten die voor de Commissie Financiën van belang zouden kunnen zijn bij de toepassing van data science bij de belastingdienst.

1. Publieke sector en belastingdienst moeten transparanter zijn over de wijze waarop ze data science toepassen

Vertrouwen tussen de overheid en haar burgers is cruciaal voor een goed functionerende rechtstaat. Ook private partijen zoals Microsoft stellen het behoud van *vertrouwen* centraal in hun doelstelling. De belastingdienst heeft in de media een naam als overheidsorganisatie die gretig is om data te verkrijgen en te analyseren. De artikelen in de correspondent van [Maurits Martijn en Dmitry Tokmetzis](#) illustreren dit.

De belastingdienst zou meer vooraf en pro-actief kenbaar moeten maken welke data zij verzamelt en analyseert. Op de huidige site van de belastingdienst staan wel algemene uitspraken over de beweging die de belastingdienst maakt maar niet de data die wordt gebruikt en analyses die de belastingdienst uitvoert. Hoe zorgt de belastingdienst ervoor dat zij enerzijds maximaal oog heeft voor de privacy van de burger en anderzijds de compliance optimaliseert? Kortom: hoe de belastingdienst de principes van Privacy by design toepast.

2. De impact van algoritmes vraagt ook om onafhankelijk assessment van de kwaliteit

In de Verenigde Staten wordt al langer gewerkt aan de onafhankelijke oordeelsvorming over de toepassing van data science. De Association for Computing Machinery, US Public Policy Council (USACM) heeft hier heeft hier een visie en principes voor ontwikkeld. Kern is dat de toepassing van data science – in het bijzonder wanneer dit maatschappelijke en persoonlijke impact heeft - onafhankelijk moet kunnen worden beoordeeld. Ik quote hier een aantal onderdelen uit de visie.

‘Computer algorithms are widely employed throughout our economy and society to make decisions that have far-reaching impacts, including their applications for education, access to credit, healthcare, and employment. The ubiquity of algorithms in our everyday lives is an important reason to focus on addressing challenges associated with the design and technical aspects of algorithms and preventing bias from the onset’.

‘There is also growing evidence that some algorithms and analytics can be opaque, making it impossible to determine when their outputs may be biased or erroneous.’

‘This set of principles, consistent with the ACM Code of Ethics, is intended to support the benefits of algorithmic decision-making while addressing these concerns. These principles should be addressed during every phase of system development and deployment to the extent necessary to minimize potential harms while realizing the benefits of algorithmic decision-making’.

De belastingdienst zou de publieke sector een dienst bewijzen door zelf ook een dergelijk statement te laten maken zodat bij de toepassing van data science kan worden getoetst aan principes. Hier kan zij de samenwerking met universiteiten zoeken om de onafhankelijkheid te borgen.

3. Soft Sister mag best wat meer omarmd worden bij de belastingdienst

In zijn in 1948 geschreven roman 1984 schetst Orwell een onmenselijke dictatoriale eenpartijstaat die in alle opzichten volledig beheerst wordt door de *Partij*. De alom aanwezige leider van de Partij wordt *Big Brother* genoemd. Big Brother en Big Data worden door critici snel aan elkaar verbonden vooral als de overheid op grote schaal data gaat verzamelen om fraude en andere misdragingen tegen te gaan. Veel cases uit de Verenigde Staten in publieke organisaties gaan ook over het opsporen en voorkomen van fraude en het besparingspotentieel daarvan.

‘Soft Sister’ is de helpende overheid. Een voorbeeld: we weten dat in Nederland een groot deel van de mensen die in armoede leeft geen gebruik maakt van voorzieningen. Voorzieningen waar ze recht op hebben om de armoede te verlichten. Dat komt doordat deze gezinnen weinig ‘bureaucratische vaardigheden’ hebben en soms slecht kunnen

lezen en schrijven. De belastingdienst zou kunnen investeren om de soft sister meer te laten zien.

4. Voorkom een rommelige discussie over ethiek, privacy en dataveiligheid

Het is onze overtuiging dat organisaties veel slimmer kunnen werken wanneer ze gebruik maken van data science-technieken. Of het nu gaat om het slimmer inkopen van zorg, het personificeren van websites of het voorkomen van branden en incidenten: data science kan hier een groot verschil maken. Er zal geen gebied ontkomen aan de toepassing van data science. Het is goed dat de belastingdienst de kansen onderkent en haar kennis en competenties de afgelopen jaren heeft versterkt. Het is belangrijk dat de belastingdienst aangemoedigd wordt om deze technieken te blijven toepassen.

Veel publieke organisaties worstelen met het begrip privacy rondom datagedreven werken. In sommige gevallen wordt er te weinig aandacht besteed aan de juridische aspecten rondom het gebruik van data en informatie. Hierdoor kan het gebeuren dat een project stil wordt gelegd op het moment van implementatie. In andere gevallen wordt het argument: 'mag niet vanuit privacy' aangedragen om niet te innoveren met data. Hierdoor blijven inefficiënte processen onnodig voortbestaan. Naast de juridische vraag spelen ook ethische en organisatorische vragen een rol:

- Juridisch: Wat mogen we?
- Ethisch: Wat willen we? Willen we vergaand differentiëren in onze aanpak of liever standaardiseren?
- Dataveiligheid (hoe verwerken we de data op een zo veilig mogelijke manier?).
- Organisatorisch: Wat kunnen we? Kunnen we onze mensen en processen veranderen zodat deze aansluiten op een data gedreven aanpak?

De laatste organisatorische vraag is naar ons idee breder dan techniek, want ook processen, mensen en vaardigheden zijn cruciaal voor een succesvolle data science toepassing.

dr Jaring Hiemstra
03 07 2017

Over Ynformed

Ynformed is een middelgroot data science bureau uit Utrecht. Wij zijn gestart in 2013. Onze missie is om voor organisaties groei en duurzame oplossingen te creëren met data science. Wij werken samen met opdrachtgevers in teams waar onze data scientists, projectleiders en adviseurs het verschil maken. Wij bieden leertrajecten en masterclasses aan voor managers, professionals en data scientists, doen data science projecten en hebben ook een aantal in huis ontwikkelde (standaard) oplossingen voor organisaties.