

Verslag van een openbaar verhoor

De parlementaire enquêtecommissie Fraudebeleid en Dienstverlening heeft op 2 oktober 2023 **de heer Koemans** als getuige gehoord, die werd bijgestaan door de heer Kramer.

Voorzitter: Belhaj

Griffier: Sjerp

Aanwezige leden van de commissie: Aartsen, Azarkan, Belhaj, Maatoug, Van Nispen, Van Raan, Sylvana Simons en Slootweg.

Aanvang 14.00 uur.

De **voorzitter**:

Ik verzoek de griffier om de volgende getuige van vandaag, de heer Koemans, naar binnen te leiden.

(De heer Koemans wordt binnengeleid door de griffier.)

De **voorzitter**:

Welkom, meneer Koemans. U wordt vandaag gehoord als getuige. Dit verhoor vindt plaats onder ede. U heeft ervoor gekozen de belofte af te leggen. Ik verzoek u met de belofte te bevestigen dat u de gehele waarheid en niets dan de waarheid zal zeggen.

(In handen van de voorzitter legt de heer Koemans de belofte af.)

De **voorzitter**:

Dan staat u nu onder ede en mag u plaatsnemen.

Goedemiddag, meneer Koemans. Ditmaal heet ik u van harte welkom namens de hele enquêtecommissie Fraudebeleid en Dienstverlening. De commissie onderzoekt hoe en waarom de overheid op een bepaalde manier fraude heeft bestreden en tekort is geschoten in de dienstverlening.

Ik wil u vragen om tijdens het verhoor uw microfoon aan te laten staan. Ik zie al dat hij een tikkeltje hoger mag; helemaal goed. Ik wil u erop wijzen dat de bijstandsverlener niet namens u het woord dient te nemen tijdens dit

verhoor. Het verhoor zal afgenomen worden door de leden Maatoug en Simons. Voordat we gaan beginnen, controleer ik even of het duidelijk is.

De heer **Koemans**:

Alles is duidelijk. Dank u wel.

De **voorzitter**:

Dan stel ik voor we gaan beginnen. Ik geef het woord aan mevrouw Maatoug.

Mevrouw **Maatoug**:

Goedendag, meneer Koemans.

De heer **Koemans**:

Goedemiddag.

Mevrouw **Maatoug**:

U werd van 1 juni 2013 tot 4 juni 2016 als manager Analytics van Deloitte ingehuurd door de Dienst Toeslagen van de Belastingdienst. Daar hield u zich bezig met de doorontwikkeling van het risicoclassificatiemodel van de Dienst Toeslagen en het overdragen van kennis over de werking van het model aan medewerkers van de Dienst Toeslagen. In dit verhoor wil de commissie het met u hebben over de opdrachtverlening aan Deloitte, de samenwerking tussen Deloitte en de Dienst Toeslagen en de werking van het risicoclassificatiemodel van de Dienst Toeslagen.

Voordat we daarnaartoe gaan, willen we het volgende aan u vragen. Welke trainingen heeft u gekregen toen u begon bij Deloitte, en hoe ging uw onboarding?

De heer **Koemans**:

Ruwweg klopt het wat u in de inleiding zei, met de kanttekening dat ik medio juni 2013 bij dit project betrokken ben geraakt. Ik ben op 1 juni in dienst gekomen bij Deloitte. Ik ben in dienst gekomen om dat te doen waarmee ik bij mijn vorige werkgever en in mijn studie ervaring had opgedaan, namelijk het analyseren van data. Medio juni kon ik, zoals ik zei, meewerken aan het eerste project, waaraan een aantal collega's al kort werkzaam was, namelijk ...

Mevrouw **Maatoug**:

Meneer Koemans, we gaan even terug naar die 1 juni, toen u in dienst kwam

bij Deloitte. Welke trainingen heeft u gekregen bij uw nieuwe werkgever Deloitte, toen u daar begon als nieuwe medewerker?

De heer **Koemans**:

Nou, wat mij in ieder geval nog bijstaat, zijn trainingen rondom onafhankelijkheid, omdat wij ook collega's in de audit, in de accountancydiensteverlening hebben. Voor de rest staan mij concreet geen onboardingstrainingen meer bij.

Mevrouw **Maatoug**:

Oké, dus u had geen projectmanagement of technische trainingen bij uw nieuwe functie aan het begin.

De heer **Koemans**:

Nee, dat staat mij niet bij.

Mevrouw **Maatoug**:

En trainingen over de code of conduct, de cultuur van het bedrijf of kennis over hoe dingen gaan bij Deloitte?

De heer **Koemans**:

Ja, die wel.

Mevrouw **Maatoug**:

Kunt u daar wat over vertellen?

De heer **Koemans**:

Dat zijn trainingen die online worden aangeboden, die meestal worden afgesloten met een examen van een aantal vragen. Dat is hoe die trainingen ruwweg in zijn werk gaan bij ons.

Mevrouw **Maatoug**:

Meneer Koemans, kunt u iets vertellen over de inhoud van die trainingen, over bijvoorbeeld de code of conduct?

De heer **Koemans**:

Nee. Daar ben ik ook niet op voorbereid. Ik kan daar nu zo uit mijn hoofd zonder voorbereiding niks over zeggen.

Mevrouw **Maatoug**:

U kunt het zich niet herinneren en weet niet wat de code of conduct is van het bedrijf waar u werkzaam bent?

De heer **Koemans**:

Volgens mij heb ik u aangegeven dat ik u nu niet iets onder ede kan verklaren over de inhoud van de trainingen die meer dan tien jaar geleden hebben plaatsgevonden.

Mevrouw **Maatoug**:

Oké, goed dat u dat aangeeft. Ik ga u ook niet vragen dingen te vertellen die u niet weet. Maar kunt u dan in algemene zin iets zeggen over die trainingen en de code of conduct van Deloitte?

De heer **Koemans**:

Nou, ik vind dat een heel ingewikkelde vraag, in die zin dat wij inderdaad een code of conduct hebben voor hoe wij ons gedragen op projecten en voor de manier waarop wij samenwerken met onze klanten en met onze collega's. Daar zullen die trainingen over gegaan zijn. Maar het staat mij niet meer bij wat die trainingen van tien jaar geleden wel of niet precies in zich hadden.

Mevrouw **Maatoug**:

Oké. U gaf net aan dat u medio juni 2013 begint met het project bij Toeslagen. Op 26 april 2013, dus daarvoor al, wordt Deloitte door de Dienst Toeslagen ingehuurd om te helpen bij de doorontwikkeling van het risicoclassificatiemodel. Vanaf uw aanwezigheid sinds medio juni 2013 tot aan 4 januari 2016 heeft u bij dat team gezeten. Kunt u vertellen wat uw functie van manager Analytics inhield?

De heer **Koemans**:

Ja. Wij hanteren voor elke opdracht die wij bij Deloitte aannemen, een structuur met een partner of director die uiteindelijk de eindverantwoordelijkheid draagt. Dat was in deze opdracht iemand anders. Mijn rol als manager bestaat eruit om het Deloitteteam operationeel aan te sturen, gecombineerd met het zijn van aanspreekpunt voor de opdrachtgever, in dit geval Belastingdienst/Toeslagen, bij gesprekken over de werkzaamheden van het Deloitteteam en de voortgang van de verstrekte opdracht.

Mevrouw **Maatoug**:

Kunt u vertellen wat "operationeel management" betekent?

De heer **Koemans**:

Dat bedoelde ik op die manier dat de twee Deloittecollega's die met mij aan het project werkten, in principe onder mijn aansturing werkten. Ik gaf aan wat ik verwachtte dat zij zouden doen, welke analyses zij moesten uitvoeren en hoe zij, in dit voorbeeld, de vraag van de klant in de SAS-software die voor de rekenmodule gebruikt werd, konden implementeren.

Mevrouw **Maatoug**:

De collega's van Deloitte waren sinds 26 april al erbij betrokken. Welke overdracht heeft u gekregen toen u zelf begon?

De heer **Koemans**:

Er waren inderdaad collega's al langer bij betrokken. Er is een overlap geweest met degene die voor mij die managerrol had.

Mevrouw **Maatoug**:

Welke overdracht heeft u van die collega gekregen?

De heer **Koemans**:

Een overdracht in de zin dat wij een periode gelijktijdig aan dat project hebben gewerkt en dat toen samen hebben gedaan. Daarin is mij uitgelegd wat er belangrijk was. Er is mij natuurlijk ook uitgelegd wat onze opdracht was.

Mevrouw **Maatoug**:

Wat was de opdracht?

De heer **Koemans**:

Onze opdracht was tweeledig. Die bestond eigenlijk uit twee stromen. De ene stroom was: meewerken aan het zelflerend maken van het al bestaande model voor risicoclassificatie voor de huur- en kinderopvangtoeslag. Dat deden we door zogenaamde trainingscases te gebruiken om uiteindelijk een statistisch model te kunnen schatten, dus om die op dat moment het model te laten zijn. De stroom de andere kant op is om vervolgens het model op basis van die trainingscases toe te passen op alle toeslagaanvragen, zodat het risico bij die toeslagaanvragen kon worden ingeschat.

Mevrouw **Maatoug**:

Kunt u de commissie vertellen wat de samenwerkingsovereenkomst was tussen Deloitte en de Dienst Toeslagen?

De heer **Koemans**:

Het contract dat er lag, moet u formeel zien als een detacheringsoopdracht. Wij noemen dat loan staff. Wij waren dus eigenlijk extra inhuur, extra handjes om de data-analysen die werkzaam waren bij Toeslagen, het grootste gedeelte van de tijd vier mensen, te helpen.

Mevrouw **Maatoug**:

U gaf net aan dat een collega van u de eerste fase had gedaan. U heeft even samen tegelijkertijd gewerkt. Daarna nam u het operationele management over.

De heer **Koemans**:

Ja.

Mevrouw **Maatoug**:

Waarbij was die collega betrokken geweest?

De heer **Koemans**:

Die was vanaf het begin van de inhuur van Deloitte erbij betrokken geweest, heeft hij mij verteld.

Mevrouw **Maatoug**:

Dus ook bij de totstandkoming van het risicoclassificatiemodel.

De heer **Koemans**:

Ik heb begrepen dat er daarvoor ook al een risicoclassificatiemodel was, maar in ieder geval is hij erbij betrokken geweest, heeft hij mij gezegd, vanaf 26 april 2013.

Mevrouw **Maatoug**:

Oké. Hoe waren de taken verdeeld tussen de Dienst Toeslagen en Deloitte?

De heer **Koemans**:

Een belangrijke en goede vraag. Het is inderdaad goed om aan te geven dat er een governance, een taakverdeling was.

Mevrouw **Maatoug**:

Welk woord gebruikt u? Governance?

De heer **Koemans**:

Governance, ja. "Taakverdeling" is volgens mij de best mogelijke

Nederlandse vertaling. Mijn rol, en ook van mijn Deloittecollega's, was vooral gericht op die twee stromen die ik schetste, dus het zelflerend maken de ene kant op en het kunnen toepassen de andere kant op. Voor beide stromen hoorde daar natuurlijk bij dat de data die bij Toeslagen aanwezig waren op één plek, in die rekenmodule verzameld werden. Dat moet allemaal in rij en gelid worden gezet, zou ik dat willen noemen, dus allemaal in het juiste formaat gemaakt worden, zodat het model het kan gebruiken en zodat het model ook toegepast kan worden op die aanvragen. Telkens werd dat model natuurlijk opnieuw geschat, met nieuwe trainingscases ...

Mevrouw **Maatoug**:

Meneer Koemans, ik onderbreek u even. We gaan straks wat uitgebreider op het model in. U heeft net heel duidelijk aangegeven wat uw taak was in de governance, als ik het juiste woord gebruik, in die taakverdeling. Wat lag er dan bij de Dienst Toeslagen?

De heer **Koemans**:

Daar lag dan in principe al het andere, dus zaken als de interne afstemming over het model, privacy en cybersecurity. Dat zijn zaken die ook van belang zijn en die dus bij Toeslagen lagen.

Mevrouw **Maatoug**:

En die governance was ook afgesproken en opgeschreven?

De heer **Koemans**:

Die was in ieder geval afgesproken. Daar is ook naar geleefd. Ik kan me zo geen document herinneren waar dat in was opgeschreven, behalve dan de formele opdrachtbevestiging aan Deloitte waarin we die detachingsafspraken hebben gemaakt.

Mevrouw **Maatoug**:

En in die detachingsafpraak stond dan niet de governance of de taakverdeling?

De heer **Koemans**:

In die detachingsafpraak staat wat er van de Deloittemensen verwacht wordt en wat de voorwaarden zijn voor inhuur.

Mevrouw **Maatoug**:

Hoeveel medewerkers van Deloitte waren van april 2013 tot januari 2016 werkzaam voor de Dienst Toeslagen?

De heer **Koemans**:

De kern van onze inzet zit vooral in de periode vanaf de start, dus 26 april 2013 – ik zoek ondertussen even naar het precieze cijfer dat u wilt weten – tot de zomer van 2014. Daarna is er nog wat incidentele maar veel kleinere dienstverlening geweest. Ik zoek even op hoeveel mensen dat waren in dat eerste jaar, dus tijdens de kern van onze inzet. Is dat vooral het antwoord op uw vraag?

Mevrouw **Maatoug**:

Nee, dat heeft u nu aangegeven, maar ik vraag naar de hele periode april 2013 tot januari 2016. Neemt u de tijd, hoor, als u het even moet opzoeken.

De heer **Koemans**:

Ik kom op tien mensen in totaal, inclusief mezelf, als de "buitenomtrek" van iedereen die er op enig moment bij betrokken is geweest vanaf 26 april. Ik heb van collega's begrepen dat er ook nog twee collega's in de voorfase betrokken zijn geweest, dus vóór 26 april 2013, maar niet meer vanaf daarna.

Mevrouw **Maatoug**:

Dus tien mensen in totaal, geeft u heel duidelijk aan, meneer Koemans, en het meest intensief in die eerste periode tot de zomer van 2014.

De heer **Koemans**:

Ja, met nadruk: zeker niet alle tien tegelijk. Het was in mijn ogen een klein team waarmee we werkten in de kernperiode van juni 2013 tot de zomer van 2014.

Mevrouw **Maatoug**:

Daarin deed u het operationele management en stuurde u twee collega's aan.

De heer **Koemans**:

Ja.

Mevrouw **Maatoug**:

We hadden het net over de taakverdeling, de governance, die wel een beetje terugkwam in de detacheringsovereenkomst maar niet in een stuk was vastgelegd. Zijn er stukken, zoals verslagen van projectoverleggen of voortgangsverslagen over de samenwerking tussen Deloitte en de Dienst Toeslagen?

De heer **Koemans**:

Heel weinig. Daar zijn er wel enkele van, vooral uit de beginfase waarin het model ontwikkeld is.

Mevrouw **Maatoug**:

U helpt ons door specifiek te zijn. Heel weinig, maar dus wel een aantal uit de beginfase. Wat staat er dan in die verslagen van projectoverleggen of in voortgangsverslagen? Welke dingen zijn wel vastgelegd?

De heer **Koemans**:

Ik weet niet precies wat daarin staat. Dat heeft er vooral mee te maken dat onderdeel van die detacheringsafspraken was dat het hele project op de systemen van Toeslagen, van de Belastingdienst werd uitgevoerd, met een eigen inlognaam en e-mailadres voor onze medewerkers. Een deel van de vastlegging zult u zeker daar kunnen aantreffen, maar daar heb ik nu niet de beschikking over.

Mevrouw **Maatoug**:

Oké. U gaf aan: heel weinig. Dus heel veel voortgang, in de vorm van verslagen van projectoverleggen en voortgangsverslagen, is dus niet vastgelegd. Kan ik dat concluderen uit uw woorden?

De heer **Koemans**:

Dat mag u concluderen, ja.

Mevrouw **Maatoug**:

Oké. Kunt u de commissie vertellen wat de basis was voor de betaling van Deloitte? Ik beloof u: we gaan later heel erg de inhoud in, maar we willen goed begrijpen wat ...

De heer **Koemans**:

Dat was onze urenregistratie, die wij bijhielden. Op basis daarvan werd per e-mail met de opdrachtgever afgestemd wat de gemaakte uren waren van een afgelopen maand. Daar volgde dan een factuur over.

Mevrouw **Maatoug**:

En de medewerkers hielden de gemaakte uren bij?

De heer **Koemans**:

Ja.

Mevrouw **Maatoug**:

Als er een nieuwe applicatie wordt gebouwd of wordt doorontwikkeld, wordt er vaak eerst een functioneel ontwerp gemaakt. Hierin staat beschreven wat de applicatie moet kunnen en hoe de applicatie moet werken. Is er een functioneel ontwerp van het risicoclassificatiemodel gemaakt? Weet u dat?

De heer **Koemans**:

Niet dat ik weet. Wij werkten in kortcyclische periodes, zou je kunnen zeggen, in een soort OTAP-straat. Het model werd ontwikkeld, het werd getest, het werd geaccepteerd of niet en vervolgens werd het in productie genomen.

Mevrouw **Maatoug**:

Hoe heeft u, toen u begon medio juni, het model zo snel mogelijk tot u genomen? Op welke manier heeft u dat kunnen doen?

De heer **Koemans**:

Door in de software te kijken. Ik had bij mijn vorige werkgever zeven jaar ervaring met het zo goed als elke dag programmeren in de programmeertaal die ook bij de Dienst Toeslagen gebruikt werd.

Mevrouw **Maatoug**:

Welke programmeertaal was dat?

De heer **Koemans**:

Dat is SAS, s-a-s in hoofdletters.

Mevrouw **Maatoug**:

Op de website van Deloitte staat Risk & Regulatory Analytics als een van de diensten waarvoor Deloitte kan worden ingehuurd. Er staat hier onder andere dat de experts van Deloitte de meest geavanceerde statistische methoden gebruiken om afwijkingen zoals fraude te detecteren. Viel uw opdracht bij de Dienst Toeslagen onder de dienst Risk & Regulatory Analytics of een voorloper daarvan?

De heer **Koemans**:

Dan overvraagt u mij. Dat is typisch een vraag die u aan de eindverantwoordelijke partner of director zou kunnen stellen. Ik kan u het antwoord niet geven.

Mevrouw **Maatoug**:

U werd ingehuurd om het risicoclassificatiemodel van de Dienst Toeslagen door te ontwikkelen. Was dit een bijzondere opdracht of is dit voor Deloitte redelijk gangbare dienstverlening?

De heer **Koemans**:

Ehm ... Ik vind dit een ingewikkelde vraag. U vraagt mij als het ware gevoelsmatig antwoord te geven. Dat vind ik ingewikkeld onder ede.

Mevrouw **Maatoug**:

Nou, niet gevoelsmatig, maar naar uw inschatting, vanuit uw ervaring. Was dit een heel bijzondere klus? Was dit een gangbare klus? Leek die op de andere klussen, gezien vanuit uw portfolio in de jaren dat u werkte als ...

De heer **Koemans**:

Dit leek zeker op andere klussen. Collega's hadden hier ook ervaring mee. In die zin was het een gangbare opdracht. Ik denk ... Ja.

Mevrouw **Maatoug**:

Was het ook gangbaar om zo weinig vast te leggen, zoals we eerder hebben vastgesteld?

De heer **Koemans**:

Terugkijkend denk ik dat dat zeker anders had moeten. Dat gaat tegenwoordig zeker ook echt heel anders. In mijn ogen was het op dat moment gangbaar om het op deze manier te doen. Ik keek er op dat moment niet van op, maar op dit moment gaan dat soort projecten en ook de vastlegging daarvan zeker anders.

Mevrouw **Maatoug**:

Dus de werkwijze van Deloitte was, als er klussen gedaan werden rondom data analytics, ook op dat moment niet om logs bij te houden, beslissingen rond het model vast te leggen en aan verslaglegging te doen. Dat was dus de reguliere werkwijze van Deloitte bij dit soort projecten.

De heer **Koemans**:

Ik kan u in ieder geval aangeven dat het op mij niet vreemd overkwam op dat moment. Hoe dat op dat moment in andere projecten ging, daar kan ik u niks over zeggen.

Mevrouw **Maatoug**:

Week de dienstverlening aan de Dienst Toeslagen op onderdelen af van wat gebruikelijk is bij uitvoeringsinstanties of was ook dat een gemiddelde gang van zaken, vanuit uw kennis en ervaring?

De heer **Koemans**:

Kunt u iets specifieker zijn? Op welk punt bedoelt u dat?

Mevrouw **Maatoug**:

Dus de detacheringsovereenkomst en de dienstverlening aan de Dienst Toeslagen in het kader van dit project.

De heer **Koemans**:

Dat is gebruikelijk. Er zijn een aantal vormen waarin we dit type dienstverlening aanbieden. Dit is zeker ook een gebruikelijke vorm, ja.

Mevrouw **Maatoug**:

Ik geef het woord aan mijn collega, mevrouw Simons.

Mevrouw **Simons**:

Dank u wel, collega, en goedemiddag, meneer Koemans. Ik wil het nu met u hebben over de eisen op het gebied van de bescherming van persoonsgegevens zoals die in de toen geldende Wet bescherming persoonsgegevens waren opgenomen. Mijn eerste vraag aan u is: waren in de opdrachtverlening voor de doorontwikkeling van het risicoclassificatiemodel eisen op het gebied van de bescherming van persoonsgegevens opgenomen?

De heer **Koemans**:

Nee. Ik schetste net aan uw collega dat wij werden ingehuurd als een team van experts die het wiskundige model, het statistische model zelflerend kwamen maken, plus nog die andere stroom die ik u schetste.

Mevrouw **Simons**:

En de opdrachtverlener, in dit geval de Dienst Toeslagen, heeft er op dat moment niet bij Deloitte, bij u en uw collega's, op aangedrongen om u ervan te vergewissen dat ook de bescherming van persoonsgegevens een belangrijke factor moest zijn?

De heer **Koemans**:

Wel in de zin dat het hele model en alle persoonsgegevens uiteraard nooit de

IT-omgeving van de Belastingdienst hebben verlaten. Uiteraard hebben wij die niet geanalyseerd op de systemen van Deloitte. Dat doen we ook nooit. Dus in die zin wel.

Mevrouw **Simons**:

Als ik u goed begrijp, zegt u dat het werken binnen de omgeving van Toeslagen dan wel de Belastingdienst eigenlijk al een vorm van het borgen van de bescherming van die persoonsgegevens was.

De heer **Koemans**:

Dat is in ieder geval een manier om dat te doen, ja. Ik denk dat u het met mij eens kunt zijn dat die gegevens bij een andere organisatie waren beland als dat niet zou gebeuren. Ik denk dat niemand dat zou willen.

Mevrouw **Simons**:

Heeft Deloitte zich ervan vergewist of het model dat Deloitte hielp bouwen, überhaupt voldeed aan de toen geldende Wet bescherming persoonsgegevens?

De heer **Koemans**:

Ik schetste net aan uw collega de governance, de taakverdeling, waarin deze aspecten duidelijk belegd waren bij de Dienst Toeslagen. Ik heb er gedurende de looptijd van die opdracht ook nooit aan getwijfeld dat die taakverdeling nageleefd werd of dat er over dit soort aspecten werd nagedacht. In die zin hebben we ons er zeker van vergewist. Jazeker. Daar komt bovenop dat wij natuurlijk ook ...

Mevrouw **Simons**:

Excuus. Ik zal de vraag nog eens herhalen: heeft Deloitte zich ervan vergewist dat het model dat Deloitte hielp bouwen, voldeed aan de toen geldende Wet bescherming persoonsgegevens? In uw antwoord zei u zojuist: op basis van de governance-afspraken gingen wij ervan uit dat die verantwoordelijkheid lag bij de Dienst Toeslagen. Mijn vraag is of u zich er desondanks van vergewist heeft dat er ook gehandeld werd volgens de toen geldende Wet bescherming persoonsgegevens.

De heer **Koemans**:

Ik zeker. Daar werd natuurlijk ook wel over gesproken met de opdrachtgever. Ik weet dat er in die tijd ook een privacy-impactassessment is uitgevoerd en dat er overleg is geweest met het College bescherming persoonsgegevens.

Mevrouw **Simons**:

Heeft u de uitkomsten van die privacy ... Wat zei u ook alweer?

De heer **Koemans**:

Privacy impact assessment.

Mevrouw **Simons**:

Dat is een PIA, toch?

De heer **Koemans**:

Ja.

Mevrouw **Simons**:

Heeft u die gezien?

De heer **Koemans**:

Nee. Ik ben wel geïnterviewd door een externe partij die dat privacy impact assessment heeft opgesteld, als ik het zo mag noemen.

Mevrouw **Simons**:

Maar het resultaat, het daadwerkelijke product, heeft u niet gezien?

De heer **Koemans**:

Dat klopt.

Mevrouw **Simons**:

Heeft Deloitte de opdracht van de Dienst Toeslagen getoetst aan een normenkader of een code voor goed bestuur?

De heer **Koemans**:

Er wordt bij ons altijd getoetst of er aan de voorwaarden voor opdrachtacceptatie voldaan wordt. Maar ik schetste u al dat ik in de loop van de opdracht hierbij betrokken ben geraakt, dus ik ben niet betrokken geweest bij de opdrachtacceptatie.

Mevrouw **Simons**:

U bent niet betrokken geweest bij de opdrachtacceptatie. Weet u of er wellicht voor uw tijd een dergelijke toets is geweest?

De heer **Koemans**:

Ik stap in een lopende opdracht. Ik vertrouw op de systemen die er bij de werkgever zijn.

Mevrouw **Simons**:

Weet u of er wellicht voor uw aantreden zo'n toets is geweest, ja of nee?

De heer **Koemans**:

Een toets op het gebied van?

Mevrouw **Simons**:

Op het gebied van een normenkader of een code voor goed bestuur.

De heer **Koemans**:

Ik weet niet of dat onderdeel uitmaakte van die opdrachtacceptatie. Dat kan ik u niet zeggen.

Mevrouw **Simons**:

Wat houdt een dergelijke code in? U zegt: er wordt altijd standaard getoetst. Maar u was daar niet bij, zei u.

De heer **Koemans**:

U bedoelt de code of conduct?

Mevrouw **Simons**:

Nee. Toen ik u de vraag stelde, zei u: ik was er zelf niet bij, maar bij een opdrachtacceptatie vindt ook altijd een toets plaats aan een normenkader of een code van goed bestuur. U zei: ik was er zelf niet bij. Maar u zei tegelijkertijd ook: dat gebeurt wel altijd.

De heer **Koemans**:

Ja, daar hebben wij een systeem voor.

Mevrouw **Simons**:

Kunt u daar iets meer over vertellen? Waar kijkt zo'n code of toets dan naar?

De heer **Koemans**:

Die kijkt onder andere naar de voorwaarden waaronder de opdracht geaccepteerd kan worden, bijvoorbeeld op het gebied van onafhankelijkheid in relatie tot onze accountancycollega's, zoals ik uw collega net schetste. Er wordt gekeken naar welke andere risico's er vanuit ons bezien in een

opdracht kunnen zitten. Er is een systeem dat dat bijhoudt. Deze opdracht was door dat systeem gekomen.

Mevrouw **Simons**:

Dus die voldeed in dat opzicht aan de code?

De heer **Koemans**:

Ja, dat zou u op zich ... In die zin is dat wel een vorm van een normenkader; daar heeft u gelijk in, ja.

Mevrouw **Simons**:

Dank u wel. Zijn volgens u persoonsgegevens, zoals leeftijd, geslacht en nationaliteit, aantoonbaar noodzakelijk voor het bepalen van een verhoogd risico op een foute toeslagaanvraag of misschien zelfs toeslagfraude?

De heer **Koemans**:

De criteria waaraan getoetst werd en ook de indicatoren die wij gebruikt hebben in het statistisch model, zijn door experts van Toeslagen opgesteld.

Mevrouw **Simons**:

Zijn ze noodzakelijk, meneer Koemans?

De heer **Koemans**:

Dat is niet mijn expertise. Daar kan ik u niks over verklaren.

Mevrouw **Simons**:

U kunt niet aangeven welke kenmerken dan wel gegevens de facto noodzakelijk zijn voor een dergelijk risicoclassificatiemodel?

De heer **Koemans**:

Nee, want dat is niet mijn specialisme. Ik ben een wiskundige. Ik heb lang in data-analyse gewerkt bij een vorige werkgever. Ik ben ook als wiskundige aangenomen bij Deloitte. Dit is ingewikkelde materie waaraan u refereert. Alle wetgeving die met de toeslagen te maken heeft, is ook ingewikkeld, dus ik kan u dat niet verklaren.

Mevrouw **Simons**:

Kunt u als wiskundige iets zeggen over hoe indicatoren en risicoregels zich tot elkaar verhouden om tot een juiste analyse te komen? Moet er een bepaalde balans zijn tussen indicatoren? Kun je alle indicatoren in één model

stoppen om tot een goede uitkomst te komen? Hoe zit dat, wiskundig gezien?

De heer **Koemans**:

Het model dat uiteindelijk is opgesteld, werkte als volgt. Er zijn data verzameld die bij Toeslagen aanwezig waren. Daar zijn inderdaad indicatoren van gemaakt die mogelijk opgenomen konden worden in het model. Die zijn vervolgens op statistische relevantie getoetst op de vraag of die indicatoren individueel het onderscheid kunnen aangeven tussen de hoogrisico- en de laagrisicoaanvragen.

Mevrouw **Simons**:

En kunt u in z'n algemeenheid iets zeggen over de statistische relevantie als ik u de indicatoren leeftijd, geslacht, nationaliteit en dergelijke voorhoud?

De heer **Koemans**:

Dan is er dus ook voor die indicatoren getoetst of er sprake was van statistische relevantie in relatie tot het risicoselectiemodel.

Mevrouw **Simons**:

Kunnen we dan aannemen dat, omdat zij vervolgens gebruikt zijn, de conclusie was dat er statistische waarde was?

De heer **Koemans**:

Ja, als deze indicatoren gebruikt zijn, was dat er inderdaad blijkbaar.

Mevrouw **Simons**:

Hoe wordt dan besloten of er statistische waarde is?

De heer **Koemans**:

Daar zijn statistische toetsen voor. U zou het kunnen zien als rekenregels. Er zijn meerdere methodes, zoals statistische methodes of kunstmatige-intelligentiemethodes, die in aanmerking zouden kunnen komen. Die zijn ten opzichte van elkaar afgewogen. Er is uiteindelijk voor gekozen om dit type model te gebruiken, juist vanwege de uitlegbaarheid van het model, zodat ook achteraf nog vast is te stellen waarom een bepaalde aanvraag als hoogrisico is aangemerkt, waar dat dan door komt. Daarom is voor deze methode gekozen. De statistische relevantie zit in die methode uiteindelijk ... Doordat er matrixrekenen plaatsvindt, komen er bepaalde statistische maatstaven uit waar dat aan getoetst kan worden.

Mevrouw **Simons**:

Ik ben geen wiskundige. Ik ben er ook niet echt in onderlegd, vrees ik. Kunt u mij uitleggen of je in z'n algemeenheid kunt zeggen dat het niet uitmaakt welke indicatoren je ergens in stopt, maar dat ze altijd een relatie met of een correlatie tot elkaar zullen hebben? Ze kunnen altijd met elkaar gebruikt worden. Klopt dat?

De heer **Koemans**:

Nou, je kunt het altijd proberen, maar dat wil niet zeggen dat het altijd lukt.

Mevrouw **Simons**:

Duidelijk. We hebben het in deze parlementaire enquête al vaker gehad over de fraude die bekend is geworden als de Bulgarenfraude. Vanmorgen hebben we het daar nog over gehad met de heer Veringmeier en vorige week met onder meer de heer Weekers en de heer Blankestijn. Daaruit is naar voren gekomen dat het managementteam van de Belastingdienst op 25 april 2013 heeft besloten om naar aanleiding van deze specifieke fraude het risicoclassificatiemodel bij de Dienst Toeslagen versneld in te zetten. Het contract met Deloitte is vervolgens een dag later getekend, op 26 april 2013. Weer een dag later, op 27 april, bij de aanvang van het kroningsweekend, waarin koning Willem-Alexander tot Koning werd gekroond, werd besloten voortvarend door te pakken. Men heeft besloten het risicoclassificatiemodel in dat weekend verder te ontwikkelen, te testen en meteen daarna in gebruik te nemen. Heeft u enig idee of de ondertekening van het contract met Deloitte voor dit specifieke project een direct gevolg is geweest van de actuele gebeurtenissen, zoals de Bulgarenfraude?

De heer **Koemans**:

Ik kan daar alleen maar naar raden. Ik was er niet bij.

Mevrouw **Simons**:

Durft u te raden?

De heer **Koemans**:

Wat u zegt, klinkt mij logisch in de oren. Maar nogmaals, ik was er niet bij, dus ik weet het niet.

Mevrouw **Simons**:

Heel duidelijk. Dan gaan we naar dat weekend van 27 tot en met 30 april 2013, zoals ik al zei. Men gaat aan de slag om het risicoclassificatiemodel verder te ontwikkelen, te testen en te meten, om het daarna meteen in

gebruik te nemen. Uit ons onderzoek en uit het gesprek met de heer Veringmeier vanmorgen blijkt dat medewerkers van Deloitte hierbij aanwezig waren. Weet u wat de rol was van die medewerkers tijdens dat weekend?

De heer **Koemans**:

Dat zal op basis van het contract waaraan ik refereerde, tot stand zijn gekomen. Maar ook hierbij geldt: ik was daar niet bij.

Mevrouw **Simons**:

Weet u of dat die twee medewerkers waren waaraan u eerder refereerde toen u zei: er waren al mensen actief vóór april 2013? Zouden dat dezelfde medewerkers kunnen zijn geweest?

De heer **Koemans**:

Als ik heb aangegeven dat er al medewerkers van Deloitte aan werkten vóór april 2013, heb ik mij vergist. Volgens mij waren er medewerkers van Deloitte bij betrokken vanaf 27 april 2013, zoals u net ook zei, dus niet al vóór april.

Mevrouw **Simons**:

Goed.

De heer **Koemans**:

Waarvan akte. Terug naar uw vraag, want die heb ik volgens mij niet helemaal beantwoord.

Mevrouw **Simons**:

Mijn vraag was of u weet wat de rol van deze medewerkers is geweest. Uw antwoord daarop was dat u dat niet kunt zeggen, omdat u daar niet bij was.

De heer **Koemans**:

Dat klopt.

Mevrouw **Simons**:

Goed. U vertelde eerder ook dat er geen functioneel ontwerp was en dat er geen schriftelijke stukken waren over de voortgang van het project. Wat was de omstandigheid waarin is besloten om het model na die fraude in dat weekend versneld te ontwikkelen? Wat was de oorzaak van het ontbreken van die documenten? Is het gangbaar dat er niets wordt bijgehouden?

De heer **Koemans**:

Ook daar kan ik u niks over zeggen omdat ik er niet bij was.

Mevrouw **Simons**:

Was het in die tijd gangbaar om niet bij te houden hoe een project zich ontwikkelt? Met "in die periode" bedoel ik: rond 2013.

De heer **Koemans**:

Dat denk ik niet, nee. Maar nogmaals, ik was daar specifiek niet bij, dus ik kan dat niet beoordelen.

Mevrouw **Simons**:

Verbaast het u dat er rondom een dergelijk project en een dergelijk weekend geen documenten zijn te vinden waarop de voortgang en de ontwikkeling zijn beschreven?

De heer **Koemans**:

Als u mij dat persoonlijk vraagt, dan verbaast mij dat. Ik denk dat het zeker beter had moeten, terugkijkend vanaf nu.

Mevrouw **Simons**:

Heeft u vaker aan projecten gewerkt waar er geen verslaglegging was?

De heer **Koemans**:

Nee. Als u het zo vraagt, niet, nee ... Nee.

Mevrouw **Simons**:

Duidelijk. Was het risicoclassificatiemodel, dat met de hulp van Deloitte werd doorontwikkeld, een model dat in staat was om fraude te detecteren?

De heer **Koemans**:

In mijn ogen was het model bedoeld om hoogrisicoaanvragen te selecteren.

Mevrouw **Simons**:

Hoog risico op wat?

De heer **Koemans**:

Hoog risico op fouten of fraude.

Mevrouw **Simons**:

Fouten of fraude? Was het uw opdrachtgever, de Dienst Toeslagen, duidelijk wat het risicomodel wel en niet kon?

De heer **Koemans**:

Ja. Doelt u op een specifiek punt?

Mevrouw **Simons**:

Nou, er schijnt nogal wat misverstand te zijn over de vraag of dit model bedoeld was om fouten dan wel fraude op te sporen. Wij hebben vanmorgen nog gehoord dat het model ontwikkeld was om fouten op te sporen en dat het geenszins de bedoeling was dat het ingezet zou worden om fraude op te sporen. En dan hoor ik u zeggen: het is ontworpen en ontwikkeld om zowel fouten als fraude op te sporen.

De heer **Koemans**:

Ja, want ik denk als wiskundige dat je op basis van de data die er waren, geen onderscheid kan maken tussen fouten en fraude.

Mevrouw **Simons**:

Was dat intentioneel? Het zijn toch twee verschillende dingen. Dus het lijkt mij handig om dat onderscheid wel te kunnen maken.

De heer **Koemans**:

Ik kan me voorstellen dat het handig kan zijn om dat onderscheid te kunnen maken. Ik kan u op basis van de statistiek geen model geven dat dat kan.

Mevrouw **Simons**:

Zegt u nu: ik kan niet zeggen dat er een model bestaat dat alleen fouten opspoot; ik kan me niet voorstellen dat er een model is dat alleen fraude opspoot, maar ik kan me voorstellen dat er een model is dat zowel fouten als fraude opspoot? Wat was in uw optiek de bedoeling van dit model? Werde erom verzocht omdat er behoefte was om fraude op te sporen of werd erom verzocht omdat er behoefte was om fouten op te sporen?

De heer **Koemans**:

In mijn optiek allebei en die twee dingen zijn door het model niet uit elkaar te trekken.

Mevrouw **Simons**:

Dus dan is het ook een model om fraude op te sporen. En als ik het model gebruik, weet ik niet wat een fout dan wel fraude is.

De heer **Koemans**:

Dat klopt.

Mevrouw **Simons**:

Dank u wel. Dan kom ik terug bij mijn vraag. Was de uitleg die u hier nu geeft, in uw opinie ook duidelijk voor de Dienst Toeslagen? Wist de Dienst Toeslagen met welk model zij werkte, namelijk een model dat geen onderscheid kan maken tussen fouten en fraude?

De heer **Koemans**:

Ja. Daar twijfel ik niet aan.

Mevrouw **Simons**:

Hoe weet u dat zo zeker? Waarom heeft u geen twijfel?

De heer **Koemans**:

... Omdat dat in die periode – u vraagt me nu wel iets te zeggen over een periode van tien jaar terug – in mijn ogen altijd duidelijk was. Er werd op de werkvloer ook op die manier over gesproken. De aanvragen die het hoogste scoorden in de risicoselectie, hadden een hogere kans om fouten of fraude te bevatten, ja. Het was ook volstrekt duidelijk voor alle betrokkenen dat het onderscheid tussen die twee niet gemaakt kon worden op basis van de data.

Mevrouw **Simons**:

Dank u wel. Ik wil met u onder andere even spreken over de betrokkenheid van de directeur Toeslagen bij dit model. Wat is nou het product dat Deloitte leverde aan de Dienst Toeslagen? Hoe zou u dat omschrijven?

De heer **Koemans**:

Los van die twee stromen die ik uw collega al beschreef, leidde vooral die tweede stroom uiteindelijk tot een lijst met hoogrisicoaanvragen.

Mevrouw **Simons**:

De tweede stroom was de stroom met de trainingscases? Bedoelt u die?

De heer **Koemans**:

Nee, juist de andere, dus waar het model gebruikt werd voor

toeslaaanvragen voor kinderopvang dan wel huurtoeslag. Want dat waren de twee toeslagsoorten waar het voor gebruik werd. Voor de andere twee werd het niet gebruikt. Het werd gebruikt om die van een risicoscore te voorzien. Vervolgens werd het gefilterd, omdat Toeslagen wilde kijken naar de nieuwe aanvragen of de aanvragen die tot een verhoging van het toeslagbedrag leidden. De aanvragen die ook nog eens daaraan voldeden en dus een verhoogd risico hadden, werden als een lijst opgeleverd aan een toezichtteam van Toeslagen. De omvang van die lijst verschilde in principe per maand. Typisch waren het 100 tot 150, soms misschien een keer 300 casussen per toeslagsoort per maand. Ze vertelden ons dat dat werd ingericht op basis van de beschikbare capaciteit.

Mevrouw **Simons**:

Als u in max twee, drie woorden zou moeten zeggen welk product Deloitte leverde aan Dienst Toeslagen?

De heer **Koemans**:

Ik kan dat niet in twee, drie woorden. Het was een aanvulling op de twee stromen die ik aan uw collega schetste. Daarvoor was het ook nodig dat die data verzameld werden en in de juiste rij en gelid gezet werden.

Mevrouw **Simons**:

En wat heeft de Dienst Toeslagen gekocht bij Deloitte?

De heer **Koemans**:

Handjes ingehuurd. Detachering.

Mevrouw **Simons**:

Detachering. En dat ging dan ook om het zelflerend maken van dat model?

De heer **Koemans**:

Ja, ook.

Mevrouw **Simons**:

Aan wie leverde u dit product op?

De heer **Koemans**:

Wij zagen de directeur en het mt van Toeslagen als de formele opdrachtgever. In de praktijk hadden wij, had ik, de meeste afstemming met de teamleider Handhavingsregie en Intelligence.

Mevrouw **Simons**:
En dat was op dat moment?

De heer **Koemans**:
Dat was de heer Veringmeier.

Mevrouw **Simons**:
Duidelijk. Zoals u weet, hebben we hier vorige week gesproken met de heer Blankestijn. Hij vertelde al iets over zijn rol bij de selectie door het risicoclassificatiemodel. De commissie wil u hier nu ook enkele vragen over stellen. Wat was de rol van de directeur Toeslagen in het proces?

De heer **Koemans**:
Wij zagen hem als formele opdrachtgever. De lijst, die ik u schetste, van toeslagaanvragen met een verhoogd risico die aan die aanvullende criteria voldeden, werd aan de directeur, aan het mt, opgeleverd. Die was ook voorzien van de indicatoren die die maand in die versie van het model – het model werd telkens aangescherpt – gebruikt werden.

Mevrouw **Simons**:
U zegt: telkens aangescherpt. Hoe vaak bedoelt u dan?

De heer **Koemans**:
Dat deden we in principe één keer per maand. In de dagen voordat men tot uitbetaling van de toeslagen overging, hadden wij weer nieuwe trainingscases verzameld. We hadden natuurlijk ook nog oude en die werden daarmee samengevoegd. En op basis van die trainingscases werd weer een nieuw model, zo zou je kunnen zeggen, ontwikkeld. Dat is dan de O in de OTAP-straat die ik noemde. Dan was er weer een nieuw model ontwikkeld en dat werd vervolgens getest op alle nieuwe en verhoogde toeslagaanvragen. Dat werd de A van geaccepteerd of niet. Vervolgens werd het dan meestal in productie genomen.

Mevrouw **Simons**:
Wie besloot om de aanvragen van mensen die waren uitgeworpen door het systeem op te schorten?

De heer **Koemans**:
Daar besloot Toeslagen over.

Mevrouw **Simons**:

Toeslagen. Hield de directeur Toeslagen zich ook bezig met de risicoselectie zelf?

De heer **Koemans**:

In de begintijd van mijn inhuur was de directeur er zelf ook bij betrokken. Iemand anders uit het managementteam was er daarna ook wel bij betrokken. Later werden die runs, zoals wij ze noemden, die maandelijkse classificatie, de hele beoordeling daarop en het akkoord daarop ook wel gedaan door de teamleider Handhavingsregie en Intelligence.

Mevrouw **Simons**:

Ik vraag het u, hè. Waarom denkt u was er tussenkomst van de directeur Toeslagen bij het selecteren van de risico-indicatoren?

De heer **Koemans**:

Ik zou niet zeggen dat hij betrokken was bij het selecteren van de risico-indicatoren, want dat gebeurde op basis van de statistische rekenmodules die ik u schetste. Het technische team kwam met een voorstel en er was dan wel enige vorm van acceptatie. Nogmaals, in die OTAP-straat staat de A van acceptatie en die was nodig.

Mevrouw **Simons**:

Kreeg de directeur dan gewoon een lijst voorgelegd om daar zijn akkoord op te geven? Hoe ging dat in zijn werk?

De heer **Koemans**:

Formeel was daar een lijst van.

Mevrouw **Simons**:

Een formele lijst?

De heer **Koemans**:

Excuus, ik verstond u even niet goed.

Mevrouw **Simons**:

U zegt: daar was formeel een lijst van?

De heer **Koemans**:

Ja. Ik kan me niet herinneren dat wij individuele bsn-nummers aan de directeuren toonden als dat uw vraag is. Belangrijk is wel die T van testen,

want daar kwam een percentage uit, een percentage van de geselecteerde aanvragen die volgens de experts van Toeslagen een verhoogd risico hadden. Dat percentage werd er ook bij opgeleverd, zodat met dat in de hand beslist kon worden over het al dan niet accepteren van die risicorun.

Mevrouw **Simons**:

Daar wil ik later nog even bij u op terugkomen. Toen ik u vroeg of de directeur Toeslagen een lijst of iets dergelijks kreeg, associeerde u dat eigenlijk meteen met bsn-nummers.

De heer **Koemans**:

Mhm.

Mevrouw **Simons**:

U zei: het is niet zo dat de directeur dan een lijst met bsn-nummers kreeg. Waarom moest u daaraan denken? Was dat wat er te zien was bij een uitworp? Neemt u ons alstublieft even mee.

De heer **Koemans**:

Nou, uiteindelijk is die lijst eh ...

Mevrouw **Simons**:

Zou u heel kort ... Er is een uitworp op enig moment. Kunt u ons eens meenemen in dat proces? Wat gebeurt er allemaal?

De heer **Koemans**:

Nou, daar kan ik u weinig over zeggen, want wij leverden vanuit het technische team die lijst op aan dat toezichtteam van Toeslagen. Zij gingen daarmee aan de slag.

Mevrouw **Simons**:

Wat stond er op die lijst dan?

De heer **Koemans**:

In de praktijk waren dat bsn-nummers op die lijst, maar die stonden daar op omdat ... Sorry, ik moet precies zijn: het waren bsn-nummers voorzien van een toeslagsoort. Iemand kan bijvoorbeeld voor de huurtoeslag wel op die lijst staan maar voor de kinderopvangtoeslag niet. Er stonden dus toeslagsoort-bsn-combinaties op die lijst. Daar ging het toezichtteam van Toeslagen mee aan de slag.

Mevrouw **Simons**:

Dus er werd een lijst met bsn-nummers uitgeworpen uit het systeem ...

De heer **Koemans**:

Ja.

Mevrouw **Simons**:

... en daarmee ging het team van Toeslagen aan de slag. Weet u ook welk team?

De heer **Koemans**:

Nee, ik weet niet meer hoe dat team heet. Nee. Misschien als aanvulling om nog iets preciezer te zijn: in de praktijk was het technisch nodig om de bsn's te selecteren. Die bsn's stonden op die lijst omdat hun toeslagaanvraag geselecteerd was. Ik schetste u net al dat in mijn ogen de toeslagensystematiek ingewikkeld is, maar iemand kan meerdere aanvragen doen, iemand kan gedurende de maand meerdere wijzigingen indienen. Dat kan de persoon zelf doen. Er kunnen uit andere bronnen wijzigingen komen die leiden tot verandering van de toeslag.

Mevrouw **Simons**:

En dan was het bsn-nummer gekoppeld aan de aanvraag.

De heer **Koemans**:

Precies. Maar toeslagaanvragen werden geselecteerd.

Mevrouw **Simons**:

Is het daarvoor nodig – ik ga weer even naar het woord "noodzakelijk" – om dat aan de hand van een bsn-nummer te kwalificeren? Waren er andere manieren geweest om aanvragen te beoordelen, zonder bsn-nummer?

De heer **Koemans**:

Ehm ... Ja, voor het statistische model denk ik dat er ook andere mogelijkheden zijn. Je zult wel enige vorm van een unieke identifier nodig hebben, die natuurlijk uiteindelijk aan een bsn te koppelen is, omdat daar systemen voor zijn, maar technisch gesproken kan dat ook een ander soort nummer zijn. Het liefst een nummer.

Mevrouw **Simons**:

Het liefst een nummer. Zou dat een nieuw gegenereerd nummer kunnen zijn, of zegt u: het moet een bestaand nummer zijn?

De heer **Koemans**:

Met een nieuw gegenereerd nummer wordt de hele procedure natuurlijk ingewikkelder, maar dat kan zeker.

Mevrouw **Simons**:

Het had wel gekund. En die lijst werd vervolgens aan de heer Blankestijn voorgelegd. Begrijp ik dat goed?

De heer **Koemans**:

Die werd aan dat toezichtteam voorgelegd, wat ik al schetste. Ik kan me niet herinneren of we nou individuele bsn's aan de heer Blankestijn lieten zien, wel inderdaad het percentage dat ik u noemde, gecombineerd met welke risico-indicatoren specifiek bij die run daadwerkelijk gebruikt werden.

Mevrouw **Simons**:

Goed, dus dan hebben we het over de risico-indicatoren. Wij hebben begrepen dat de directeur Toeslagen zich niet bemoeide met die risico-indicatoren. Nu ben ik even verward of u nou net iets anders verklaarde. U zei eerder dat ...

De heer **Koemans**:

Ik zei u net, om precies te zijn, dat de directeur Toeslagen zich niet bemoeide met de selectie van de indicatoren.

Mevrouw **Simons**:

Duidelijk.

De heer **Koemans**:

Dat deed het technische team op basis van statistiek.

Mevrouw **Simons**:

Duidelijk. Op basis waarvan deed het technische team dat dan? Welke keuzes moesten zij maken?

De heer **Koemans**:

Even ervan uitgaande dat het type model al gekozen is, onder het model liggen trainingscases die het verschil moeten maken tussen de hoogrisico- en de laagrisicoaanvragen. Dan zijn er statistische maatstaven – die zijn ingebouwd in de software die gebruikt werd – om te beslissen welke indicatoren wel en welke niet gebruikt worden, die maand.

Mevrouw **Simons**:

Was het model in staat om alle risico-indicatoren mee te nemen of moest je daar echt eerst een selectie voor maken? Had het model het gekund met meer indicatoren of wellicht minder indicatoren?

De heer **Koemans**:

Hoe dat in z'n werk gaat, is dat er enkele tientallen indicatoren waren geprogrammeerd door het technische team die het model zou kunnen gebruiken. Vervolgens komt het model op basis van statistiek met wat statistisch optimaal is. Dan werden er typisch tien tot vijftien indicatoren gebruikt die maand.

Mevrouw **Simons**:

Dank u wel. Toonde de directeur Toeslagen verder interesse in het risicoclassificatiemodel en de doorontwikkeling daarvan, los van de indicatoren?

De heer **Koemans**:

In de zin van een geïnteresseerde opdrachtgever zeker.

Mevrouw **Simons**:

Kunt u dat nog specificeren? Waar was hij dan in geïnteresseerd?

De heer **Koemans**:

Nee, dat is meer een algemeen beeld dat ik heb. Ik kan u niet zeggen dat de directeur specifiek bepaalde vragen heeft gesteld of in die zin een andere heel directe betrokkenheid had met het model. Nee. Maar daarmee neem ik niet weg dat hij een geïnteresseerde opdrachtgever was en dat dit natuurlijk een onderdeel van zijn team was waar hij uiteindelijk eindverantwoordelijk voor was. Dus een gewoon interesse, zou ik zeggen.

Mevrouw **Simons**:

Een gezonde interesse, zou u zeggen?

De heer **Koemans**:

Zoals het een directeur betaamt.

Mevrouw **Simons**:

Dank u wel. Ik geef het woord voor nu aan collega Maatoug.

Mevrouw **Maatoug**:

Meneer Koemans, we hebben het in dit verhoor nu al een paar keer over het risicoclassificatiemodel gehad. U gaf ook aan dat er verschillende typen modellen zijn. We vonden in ons onderzoek ook in de stukken dat bij een van de eerste presentaties gezegd is dat er keuze is, dat je kunt kiezen voor regressie, random forest, k-nearest neighbours en support vector machines. Wat voor type model is het uiteindelijk geworden?

De heer **Koemans**:

Logistische regressie.

Mevrouw **Maatoug**:

Logistische regressie. U gaf net in het verhoor met mevrouw Simons aan dat u bent geïnterviewd voor een PIA, ook wel gegevenseffectbeoordeling.

De heer **Koemans**:

Dat klopt.

Mevrouw **Maatoug**:

Wanneer was dat?

De heer **Koemans**:

Dat was in het voorjaar van 2014.

Mevrouw **Maatoug**:

In het voorjaar 2014. En door wie bent u toen geïnterviewd?

De heer **Koemans**:

Dat weet ik niet meer.

Mevrouw **Maatoug**:

Kunt u ons iets vertellen over dat gesprek, over wat u toen allemaal heeft meegegeven, of de hoofdpunten daaruit?

De heer **Koemans**:

Op hoofdlijnen weet ik nog dat ik erover gesproken heb en dat ik die mevrouw heb uitgelegd, hoe dat model technisch werkte. Het staat mij nog zo goed voor ogen en ik weet ook zeker dat dat gesprek heeft plaatsgevonden omdat die mevrouw mij corrigeerde op de afkorting "PIA".

Mevrouw **Maatoug**:

Wat was de correctie dan?

De heer **Koemans**:

De A kan van assessment zijn of van analyse, geloof ik.

Mevrouw **Maatoug**:

Ah!

De heer **Koemans**:

Soms onthoud je dat soort details, waardoor je je een gesprek nog herinnert.

Mevrouw **Maatoug**:

En u had dat gesprek met een mevrouw. Weet u nog bij welke afdeling of in welke hoek deze mevrouw werkte?

De heer **Koemans**:

Volgens mij was deze mevrouw niet werkzaam bij Belastingdienst/Toeslagen, maar was ze een externe partij.

Mevrouw **Maatoug**:

U gaf ook aan dat u het PIA of de GEB, waar de letters ook voor staan, niet gezien heeft. Weet u wanneer die uiteindelijk verschenen is?

De heer **Koemans**:

Nee, dat weet ik niet.

Mevrouw **Maatoug**:

Ik wil het nu met u hebben over de rol van Deloitte in de doorontwikkeling en toepassing van het risicoclassificatiemodel, waarvan we nu weten dat het een logistischeregressiemodel is, wat ook logisch is, want de uitkomstvariabele was ...

De heer **Koemans**:

Een 0 of een 1, zou je kunnen zeggen.

Mevrouw **Maatoug**:

Zo is het.

De heer **Koemans**:

U bent goed ingelezen.

Mevrouw **Maatoug**:

Hoever was de Dienst Toeslagen al gevorderd met de ontwikkeling van het risicoclassificatiemodel op het moment dat Deloitte erbij werd betrokken?

De heer **Koemans**:

Ik kan u alleen iets zeggen over de staat van het model toen ik erbij betrokken raakte. Toen was er een werkend model, medio juni 2013.

Mevrouw **Maatoug**:

Wat betekent "een werkend model"?

De heer **Koemans**:

Dat betekent dat er één systeem is, één plek zogezegd in IT-termen, waar die data aanwezig zijn, waar dat model geschat kan worden, waar dat model gemaakt kan worden en waar dat model vervolgens toegepast kan worden op alle toeslagaanvragen.

Mevrouw **Maatoug**:

Dus een werkmodel zegt niks over de kwaliteit van de schatting.

De heer **Koemans**:

Dat klopt.

Mevrouw **Maatoug**:

"Een werkend model" is dus: het staat er; het doet iets.

De heer **Koemans**:

Het doet iets. Het was ook geaccepteerd en eh ... Ja, punt. Maar het klopt wat u zegt: theoretisch kan het ook zo zijn dat er op zich een werkend model is, maar dat de kwaliteit heel laag is. Dat is theoretisch natuurlijk een mogelijkheid.

Mevrouw **Maatoug**:

Want toen u begon medio juni, hoeveel werd er toen goed verklaard?

De heer **Koemans**:

Dat waren eigenlijk ... In ieder geval vanaf dat moment waren dat hoge percentages, die door dat toezichtteam ook als hoog risico beoordeeld werden, in de zin dat 90% à 95% van de output ook als hoog risico werd aangemerkt.

Mevrouw **Maatoug**:

Ja, en nu wordt het interessant. Wij zien dat in de gegevenseffectbeoordeling van 2021 – dat is een openbare versie die wij ook hebben – wordt aangegeven dat er een aantal risico's in het model zitten. Een van de risico's is die kunnen ontstaan, is dat de verdeling in het model niet evenredig is, dat er selectiebias is. We hebben in het verhoor hiervoor gehoord dat er geen aselechte steekproeven zijn gebruikt. Wij weten ook dat er in de verhouding van de testcases meer foute cases dan goede cases waren. We weten ook dat een van de databronnen die gebruikt zijn en die ook in SAS zijn geladen, de hot/hor is. Voor de mensen die niet al langer meeluisteren naar deze verhoren: dat is "hoog bedrag". En dan staat er tussen haakjes: dus een hoog risico. Maar wij weten nu dat dat niet altijd zo was. Mijn vraag aan u is: als er een werkend model is, wat betekent dat dan? Wat zegt dat over de uitkomsten? En wat zegt het als een selectieteam zegt: hier is sprake van een hoog risico, want wij zijn uitgegaan van het feit dat het een hoog bedrag was?

De heer **Koemans**:

Dat zegt in ieder geval dat ik die redenering niet deel, want het feit dat het een hoog bedrag is, zegt in mijn ogen niks over het risico.

Mevrouw **Maatoug**:

Interessant, want was dat een van de variabelen in het model?

De heer **Koemans**:

De hoogte van het toeslagbedrag?

Mevrouw **Maatoug**:

Nee, hot/hor. Dat is een van de bronbestanden.

De heer **Koemans**:

Als u mij toestaat, spiek ik dan even in de openbaarmaking van het risicoselectiemodel. Ik weet dat niet uit mijn hoofd.

Mevrouw **Maatoug**:

Laten we het dan even wat algemener maken. U was niet betrokken bij het kroningsweekend, bij de totstandkoming. U gaf aan dat u projectmanagement deed en twee mensen aanstuurde. Wij gaan zo naar de verdere werking van het model en naar hoe dat in de praktijk ging in uw periode. Maar als een model in drie dagen ontwikkeld wordt, getest wordt en in productie gaat, kun je de datakwaliteit dan toetsen?

De heer **Koemans**:

Ja.

Mevrouw **Maatoug**:

Hoe doe je dat dan?

De heer **Koemans**:

De datakwaliteit toetsen?

Mevrouw **Maatoug**:

Mhm.

De heer **Koemans**:

Dat doe je vooral door te kijken naar de inputdata. Er is een aantal checks om te kijken of die inputdata plausibel eruitzien, of die correct zijn.

Mevrouw **Maatoug**:

Is dat gebeurd? Heeft Deloitte dat gedaan bij het ontwikkelen en testen van dat model?

De heer **Koemans**:

Ik kan alleen verklaren over de periode waar ik bij betrokken was. Toen werd de kwaliteit van de inputdata inderdaad bij elke run getoetst.

Mevrouw **Maatoug**:

Maar vanaf het moment dat het model in gebruik wordt genomen, wordt het op de hele toeslagenpopulatie toegepast. Dus terwijl men bezig gaat – daar gaan we zo heen – wordt het verbeterd. Maar het gaat even om het moment dat besloten wordt dat men het model in productie kan nemen. Heeft de collega van wie u een overdracht heeft gekregen, op dat moment de kwaliteit getoetst van de data die waren aangeleverd waarmee het model werd gevoed?

De heer **Koemans**:

Dan doelt u op het kroningsweekend?

Mevrouw **Maatoug**:

Ja.

De heer **Koemans**:

Daar was ik niet bij. Daar kan ik u niks over zeggen.

Mevrouw **Maatoug**:

Maar weet u of uw collega's dat gedaan hebben? Want u heeft een overdracht gekregen.

De heer **Koemans**:

Dat weet ik niet.

Mevrouw **Maatoug**:

Dat weet u niet. Oké. Ik noemde net al een risico uit de gegevenseffectbeoordeling uit 2021. Was u op de hoogte van het risico op selectiebias in het model?

De heer **Koemans**:

Ja, zeker. Zeker is dat een belangrijk punt om rekening mee te houden. Daarom is het inderdaad belangrijk om naar die trainingscases te kijken. Ik zoek even naar mijn voorbereiding op dit punt, want het is terecht dat u deze vraag stelt. De situatie bij Toeslagen was zo dat risicoselectie niet de enige reden was om controles bij bepaalde aanvragen te doen.

Mevrouw **Maatoug**:

Meneer Koemans, ik onderbreek u toch even. Ik vraag u niet naar de toepassing en het verdere gebruik. Het gaat ons in dit verhoor om het begrip van het model en ook het technische begrip van het model. Ik vroeg niet naar wat er daarna gebeurde. Het gaat specifiek om de manier waarop in de training van het model, in de invoering van het model, is gewaarborgd dat die selectiebias geminimaliseerd is.

De heer **Koemans**:

Ja, dat zit in de samenstelling van de trainingscases. Die kwamen, naarmate het model langer in gebruik was, voor een deel ook uit de risicoselectie zelf, maar die kwamen zeker ook uit andere controles die bij Toeslagen hadden plaatsgevonden. Daar vindt thematisch toezicht plaats. Er kwamen bijvoorbeeld tips binnen bij Toeslagen, en ook vanuit die bronnen werden er trainingscases toegevoegd.

Mevrouw **Maatoug**:

Heeft u vanuit uw technische verantwoordelijkheid in de governance aangegeven dat het nodig was om ook een aselechte steekproef te doen voor het testen om selectiebias te voorkomen?

De heer **Koemans**:

Ja.

Mevrouw **Maatoug**:

Is dat gebeurd?

De heer **Koemans**:

Wij hebben in ieder geval lijsten opgeleverd om die aselecte steekproef mogelijk te maken.

Mevrouw **Maatoug**:

En is dat gebeurd?

De heer **Koemans**:

Wij hebben daar trainingscases over teruggekoppeld gekregen, dus dan is mijn conclusie dat dat gebeurd moet zijn.

Mevrouw **Maatoug**:

Oké. U gaf net in het verhoor met mevrouw Simons aan dat er allerlei wiskundige tests gedaan kunnen worden en ook gedaan worden om de samenhang tussen variabelen te bekijken. Kunt u de commissie vertellen welke tests Deloitte heeft gedaan om te kijken hoe de indicatoren onderling samenhangen en elkaar beïnvloedden?

De heer **Koemans**:

De ingehuurd medewerkers hebben inderdaad met het technische team, dat bestond uit mensen van Deloitte en Toeslagen, die testen uitgevoerd; misschien mag ik dat er nog even aan toevoegen. Dus dat zijn de tests die de statistische relevantie van een indicator beoordelen en daarmee aangeven of een variabele nog iets extra's toevoegt ten opzichte van ... Dat werkt een beetje trapsgewijs, zou je kunnen zeggen. Trapsgewijs ... Er wordt in technische termen van een bepaald model uitgegaan. Dan wordt er gekeken of het iets toevoegt om een variabele die daar nog niet in gebruikt is, toe te voegen, ja of nee.

Mevrouw **Maatoug**:

Oké. Ik zou graag nu met u stap voor stap daardoorheen willen. U helpt ons door daar zo duidelijk en concreet mogelijk in te zijn. Allereerst, hoeveel risico-indicatoren werden er oorspronkelijk door Toeslagen verstrekt?

De heer **Koemans**:

Sorry, wat bedoelt u met "oorspronkelijk"? Bedoelt u een bepaald moment?

Mevrouw **Maatoug**:

Het liefst bij de totstandkoming van het model en anders vanaf het moment dat u er weet van had.

De heer **Koemans**:

In mijn herinnering waren er enkele tientallen indicatoren, waarbij sommige zowel voor de huurtoeslag als voor de kinderopvangtoeslag gebruikt werden. Sommige, kunt u zich voorstellen, gingen specifiek alleen over de huur of over het aantal uren kinderopvang.

Mevrouw **Maatoug**:

Kunt u enkele voorbeelden geven?

De heer **Koemans**:

Van die indicatoren?

Mevrouw **Maatoug**:

Ja.

De heer **Koemans**:

Bij de kinderopvangtoeslag was de afstand tussen het woonadres van de aanvrager en de locatie van kinderopvang er een, weet ik bijvoorbeeld nog. Daar is naar gekeken en eh ... Punt.

Mevrouw **Maatoug**:

Kunt u nog een voorbeeld geven?

De heer **Koemans**:

Bij de huurtoeslag bijvoorbeeld het percentage van het inkomen dat aan huur uitgegeven werd.

Mevrouw **Maatoug**:

Even voor de duidelijkheid, toen u medio juni begon, was er geen verslaglegging over welke indicatoren of variabelen er uit een bepaalde brainstorm waren gekomen of wat de logica was toentertijd. Dat was er allemaal niet. Het enige wat u had, was toegang tot een technische applicatie waarmee u kon zien welke indicatoren erin zaten. Vat ik dat juist samen?

De heer **Koemans**:

Ja. En in die technische applicatie zie je natuurlijk ook hoe het geprogrammeerd is. Dus daaruit kon ik in technische termen afleiden hoe dat werkte.

Mevrouw **Maatoug**:

Ja, maar daar staat geen log in dat zegt "we kijken naar afstand, want ..." of "het aantal kinderen is relevant, want ..."

De heer **Koemans**:

Dat klopt.

Mevrouw **Maatoug**:

... of "basis expert judgement, want ...". Die gegevens had u allemaal niet.

De heer **Koemans**:

Nee, de reden achter de risico-indicatoren, als ik dat zo mag vertalen, stond niet op papier. Althans, ik heb dat niet gezien.

Mevrouw **Maatoug**:

Toen kwam de doorontwikkeling. Welke risico-indicatoren zijn er in de fase dat het model werd verbeterd, aan toegevoegd door Deloitte, door Toeslagen of in de samenwerking?

De heer **Koemans**:

Ruwweg is er in dat jaar weleens een enkele keer een risico-indicator toegevoegd. Dat gebeurde dan omdat een onterechte aanvraag op een andere manier naar boven was gekomen die niet door het risicomodel onderkend was. Dan werd er weleens gevraagd: klopt het nou dat het risicomodel dit niet onderkend heeft; is daar technisch niet iets fout gegaan? Nee, daar was technisch niks fout gegaan, maar het model was nog niet in staat om die aanvraag te onderkennen. Ik herinner me dat er daarom weleens een indicator aan de set is toegevoegd. Ik kan u niet zeggen welke dat dan precies was. Dat weet ik niet meer.

Mevrouw **Maatoug**:

Oké. Is er een indicator toegevoegd die opeens heel veel verklaarde, waardoor de verklarende variantie van het model ineens echt vooruitsprong?

De heer **Koemans**:

Nee, een enorme sprong maken kon eigenlijk niet meer, want zoals ik al schetste, had het model al een hoog percentage van 90%, 95%.

Mevrouw **Maatoug**:

Maar aan het begin niet, toch?

De heer **Koemans**:

In het begin, toen ik erbij betrokken raakte, zeker wel.

Mevrouw **Maatoug**:

Oké. Weet u nog wat het percentage aan het begin was?

De heer **Koemans**:

Doelt u op het kroningsweekend?

Mevrouw **Maatoug**:

Ja.

De heer **Koemans**:

Dan kan ik alleen maar refereren aan het verhoor dat u gehad heeft met de heer Blankestijn.

Mevrouw **Maatoug**:

Oké. Maar die informatie werd dus niet aan u overgedragen door uw collega toen u begon, van: hé, hier is een model; de verklaarde variantie is nu X.

De heer **Koemans**:

Nee.

Mevrouw **Maatoug**:

Oké. Dan die tests. Welke test heeft u bij die doorontwikkeling gedaan om te kijken hoe de indicatoren elkaar onderling beïnvloedden, of doet het model dat zelf door hoe de programmering geschreven is?

De heer **Koemans**:

Sorry. Dat zijn details die toen zeker in de systemen van Toeslagen aanwezig waren, maar die ik nu niet meer zo uit mijn hoofd weet. Ik schetste net ook dat er nooit een model is geweest, dat er nooit data zijn geweest, op de Deloitte systemen, dus ik heb dat niet kunnen terugzoeken.

Mevrouw **Maatoug**:

Ik ga deze vraag toch anders formuleren. Er is een model. Je bent bezig het model beter te laten werken. Er gaan indicatoren in en uit. Het wisselt. Elke maand wordt dezelfde basisset indicatoren meegenomen ...

De heer **Koemans**:

Het was ook niet zo dat het elke maand enorm verschilde, hè, maar een klein beetje wel.

Mevrouw **Maatoug**:

Nee, het ging om een paar wisselingen. Er kwam af en toe wat bij. Dat heeft u net allemaal heel duidelijk verklaard. Is er in de periode dat u begon, medio juni, op een andere manier, descriptief of in het model zelf, gekeken hoe de variabelen samenhangen? U zei eerder: ik heb wiskunde gestudeerd; er zijn heel veel tests die je daarvoor doet. Welke tests werden er in de techniek van dit model gedaan?

De heer **Koemans**:

Bedoelt u dan bij het maken van het model?

Mevrouw **Maatoug**:

Ja, bij het maken, maar ook weer elke keer bij de herschatting.

De heer **Koemans**:

Ja. Ik weet niet meer precies welke statistische tests daarbij gebruikt zijn. Dat zit in die systemen. Dat zit in die logs, maar ik kan me dat uit mijn hoofd niet herinneren.

Mevrouw **Maatoug**:

Klopt het dat dit in de governance onder de verantwoordelijkheid van Deloitte viel?

De heer **Koemans**:

Ik zie dit in ieder geval als de verantwoordelijkheid van het technische team dat daar werkte, bestaande uit mensen van Belastingdienst/Toeslagen en van Deloitte.

Mevrouw **Maatoug**:

Is er ooit een variabele uit gehaald omdat bleek dat die variabele eigenlijk hetzelfde verklaarde als een andere variabele, omdat die twee in interactie

waren of, met een lelijk Engels woord, een spurious relatie hadden? Is dat weleens gebeurd in de doorontwikkelingsfase?

De heer **Koemans**:

Ongetwijfeld. Nee, met zekerheid kan ik zeggen dat er ook zeker variabelen niet meer gebruikt zijn. Er werden dus zeker niet alleen variabelen aan toegevoegd. Er werden ook variabelen in een volgende run niet meer gebruikt, maar zonder de redenering. Want u zei: er is er eentje uit gehaald, want enzovoort. Die redenering hoort er niet bij in mijn ogen. Dat is niet wat de statistische informatie aangeeft.

Mevrouw **Maatoug**:

Oké. Het zijn geen detailvragen, maar ik concludeer dat u ons hier geen concrete informatie over kunt geven. Was het de bedoeling om de risico-indicatoren in één keer gelijktijdig te wegen?

De heer **Koemans**:

Ja.

Mevrouw **Maatoug**:

Waarom?

De heer **Koemans**:

Omdat ik van Toeslagen begreep dat de methodiek daarvoor juist anders was, dat er af en toe inderdaad naar één specifieke indicator gekeken werd, en dat iedereen die voldeed aan het criterium dat voor die indicator gold, extra gecontroleerd zou moeten worden. Dat leidde tot ongewenste situaties, want dat model – dat zou je ook een vorm van een model kunnen noemen, alleen heel grofmazig – leidde niet tot de gewenste uitkomsten. Misschien mag ik u daar een voorbeeld van geven?

Mevrouw **Maatoug**:

Tuurlijk.

De heer **Koemans**:

Ik begreep dat er eerder gekeken is naar aanvragen voor kinderopvangtoeslag waarbij het kind dat naar de opvang ging, niet op hetzelfde adres woonde als de persoon die hem aanvroeg. Dan denkt u waarschijnlijk in eerste instantie dat dat opvallend zou kunnen zijn – daar was een soort van model voor gemaakt; daar was naar gekeken met een data-analyse – tot men er bij de controle achter kwam dat er zoiets bestaat

als co-ouderschap, waarbij de ouders uit elkaar zijn, het kind bij de ene ouder woont volgens de Gemeentelijke Basisadministratie, maar de andere ouder de toeslag aanvraagt. Het is niet gewenst om daar controles op te doen. Juist daarom was men op zoek naar een model dat alle variabelen ineens kon wegen.

Mevrouw **Maatoug**:

Dit was in het model daarvoor, dus voor de methode die door Deloitte is ingevoerd. Haal je dit soort dingen eruit in een pilotfase?

De heer **Koemans**:

Nee, dit was meer een exemplarisch voorbeeld van iets wat mij in die periode verteld is maar ...

Mevrouw **Maatoug**:

Maar in algemene zin, u doet vaak projecten. Je denkt: dit is logisch op basis van expert judgement. Je laadt de data. Je hebt de data. Je gaat het uitwerken en dan kom je er bij het gebruik achter dat ergens wél een goede verklaring voor is. Haal je dat soort dingen er normaal in een pilotfase uit?

De heer **Koemans**:

O zo, in algemene zin. Ja.

Mevrouw **Maatoug**:

Oké. We hebben het al een beetje aangeraakt, maar hoeveel indicatoren zijn er daadwerkelijk in het model opgenomen? Weet u dat?

De heer **Koemans**:

Typisch werden er tussen de tien en vijftien echt gebruikt, door het model statistisch relevant geacht. Het model had nog meer variabelen tot zijn beschikking.

Mevrouw **Maatoug**:

Bleven de gegevens die geen statistische relevantie hadden, in het model of werden die permanent verwijderd?

De heer **Koemans**:

Nee, de redenering was dat de data die geen statistische relevantie hadden, bij de volgende run met net andere trainingscases, weer wel relevant konden zijn.

Mevrouw **Maatoug**:

Oké. Dus die bleven dan in het model.

De heer **Koemans**:

Die data bleven ...

Mevrouw **Maatoug**:

Er werd nooit wat afgevoerd?

De heer **Koemans**:

Nou, die data bleven behouden; de indicator werd dus níét gebruikt in het model.

Mevrouw **Maatoug**:

Oké. Maar er werden geen indicatoren afgevoerd als na zoveel runs bleek dat ze niks deden.

De heer **Koemans**:

Nee, ik kan me niet herinneren dat wij indicatoren hebben afgevoerd uit die lijst van tientallen variabelen.

Mevrouw **Maatoug**:

Het is een beetje kort door de bocht, maar ik ga de vraag toch stellen om het gewoon duidelijk te krijgen. Aan hoeveel indicatoren moet de aanvrager tegelijkertijd voldoen om een voorafgaande controle te krijgen, om bovenaan die lijst te komen?

De heer **Koemans**:

Dat is zo in het algemeen niet te zeggen, maar typisch zagen wij dat mensen dan gelijktijdig aan acht, tien of twaalf risicofactoren voldeden. Maar u schetste al de situatie dat de bovenkant van de hoogrisicolijst, als ik het even zo mag noemen, opgeleverd werd aan het toezichtteam. De omvang van die lijst werd bepaald door de capaciteit die Toeslagen daarvoor beschikbaar had.

Mevrouw **Maatoug**:

Hoe worden de risico-indicatoren gewogen en gecombineerd om de totale risicoscore te berekenen? Kunt u dat in simpele taal vertellen?

De heer **Koemans**:

Dat is ingewikkeld uit te leggen. Dat zit in die logistische regressie, maar u

kunt het zien als een soort rekenmodule. Misschien kunt u zich een hele grote Excelsheet voorstellen met honderdduizenden regels en dus die tien tot vijftien variabelen in kolommen weergegeven. Daarin doe je dus per aanvraag een bepaalde weging van die vijftien risicofactoren. Sommige kunnen voor die aanvraag bijvoorbeeld de risicofactor nul hebben. Die draagt niks bij. Of een factor draagt wel wat bij, namelijk een verhoging van het risico. Bepaalde factoren kunnen ook tot een verlaging van het risico leiden. Er zijn zeker ook factoren waarbij er, als daaraan voldaan werd, sprake was van een lager risico.

Mevrouw **Maatoug**:

Heeft u een voorbeeld? Wanneer had je een verlaagd risico op een fout?

De heer **Koemans**:

Ehm ... Nee, dat weet ik zo niet uit mijn hoofd. Dat is echt informatie die in dat model zit, die ik niet heb kunnen terugzoeken en niet meer weet van tien jaar geleden. Nee.

Mevrouw **Maatoug**:

Oké. Daar kwam dan een risicoscore uit. Dat was een getal. Kunt u de range van dat getal even vertellen?

De heer **Koemans**:

Dat loopt van 0 tot 1.

Mevrouw **Maatoug**:

Van 0 tot 1.

De heer **Koemans**:

Van 0, inclusief 0, tot en met 1.

Mevrouw **Maatoug**:

Ja, tot en met 1. En dan heb ik 0,49. Wat betekent dat? Kunt u dat uitleggen?

De heer **Koemans**:

Nee, dat getal heeft niet zozeer een interpretatie. Dat moet vooral gezien worden in de onderlinge volgorde van de aanvragen.

Mevrouw **Maatoug**:

Oké. Dus het is een ranking.

De heer **Koemans**:

Ja.

Mevrouw **Maatoug**:

Maar ik wil de zin opschrijven: de risicoscore is 0,49. Hoe moet ik dat getal dan interpreteren?

De heer **Koemans**:

Niet. ... Dat moet u vooral interpreteren als: die risicoscore was blijkbaar lager dan ...

Mevrouw **Maatoug**:

Dus het is toch een interpretatie, dus ...

De heer **Koemans**:

... de volgende aanvrager die ...

Mevrouw **Maatoug**:

... als het getal hoger was, was er een verhoogde kans op een fout of ...

De heer **Koemans**:

O zo. Oké. Ja, hoe dichterbij 1 zit, hoe hoger het risico. Dat wel.

Mevrouw **Maatoug**:

Oké. Dan betekent het getal gelukkig toch nog wat. Dus hoe hoger bij de 1 ...
Maakt u 'm even af; ik onderbrak u.

De heer **Koemans**:

Hoe hoger het risico. Althans ...

Mevrouw **Maatoug**:

Hoe hoger het risico. Hoe lager of negatief ... Want u gaf aan dat je een negatieve score kon hebben. Of dat kon niet. Het was 0.

De heer **Koemans**:

Nee, een indicator kan tot een verlaging ...

Mevrouw **Maatoug**:

Verlaagd risico.

De heer **Koemans**:

... van de risicoscore leiden, maar door de berekenmethodiek – dat zit in die statistische methodiek – kan je nooit onder 0 uitkomen.

Mevrouw **Maatoug**:

Ja, en als het 0 is of heel dicht bij de 0 zit, dan ...

De heer **Koemans**:

Dan is er niet of nauwelijks sprake van risico. Althans, dat is wat het model beoogt te doen.

Mevrouw **Maatoug**:

Dus even voor de simpele mensen onder ons: hoe hoger het getal, hoe hoger het risico. Dat is wel de interpretatie die we aan het getal kunnen geven.

De heer **Koemans**:

Ja, in die zin wel.

Mevrouw **Maatoug**:

Oké. Is er een systeem om de betrouwbaarheid of validiteit van de risico-indicatoren te meten en te evalueren?

De heer **Koemans**:

Ja, dat zit in de twee dingen die ik schetste. Dat is enerzijds de statistische toets bij het opstellen van het model, bij het schatten van het model. Anderzijds is dat bij de outputcontrole het percentage dat door de experts van Toeslagen ook als hoog risico werd beoordeeld.

Mevrouw **Maatoug**:

Betekent dat dan dat de cases die u gebruikte om het model te laten leren in de feedbackloop die u schetste, met de twee functies die u had, dus essentieel zijn?

De heer **Koemans**:

Zeker.

Mevrouw **Maatoug**:

Want aan de hand daarvan traint het model zichzelf om het risico op een fout beter te verklaren.

De heer **Koemans**:

Ja.

Mevrouw **Maatoug**:

Dus die selectie is heel belangrijk. Wie maakte die selectie ook alweer?

De heer **Koemans**:

Het technische team kreeg dat aangeleverd door Toeslagen.

Mevrouw **Maatoug**:

Deloitte bemoeide zich niet met die selectie?

De heer **Koemans**:

Wij hebben wel ... Ik heb ... Ik kan in ieder geval voor mezelf spreken. Ik heb er heel vaak op gehamerd dat we voldoende kwalitatief goede trainingscases zouden ontvangen, juist om een goede mix van trainingscases te hebben.

Mevrouw **Maatoug**:

Maar dan gaan we even terug naar de governance. Het aanleveren van cases, van gevallen van mensen of individuele zaken die gebruikt werden om het model te trainen, was een taak die in de verdeling bij de Dienst Toeslagen lag.

De heer **Koemans**:

Mhm.

Mevrouw **Maatoug**:

Het technisch draaien van het model, dus dat het runt, lag bij Deloitte.

De heer **Koemans**:

Nee, dat lag bij het technische team. Wij werden daarvoor ingehuurd.

Mevrouw **Maatoug**:

Ja, dat lag bij het technische team, waarbij men afhankelijk was van de inhuur van Deloitte. Ja, juist?

De heer **Koemans**:

Ja.

Mevrouw **Maatoug**:

U heeft erop gehamerd, maar heeft u kunnen beoordelen of de kwaliteit van

de cases die gebruikt werden om het model te laten trainen en te leren, voldoende was, of de cases voldeden aan hetgeen waar u op hamerde?

De heer **Koemans**:

Daar is vooral dat uitkomstpercentage, dat door hun experts als hoog risico beoordeeld werd, het belangrijkste criterium voor.

Mevrouw **Maatoug**:

Dus dat is de beoordeling van mensen zelf, en die beoordeling speelde ook een rol in de cases waarmee het model gevoed wordt.

De heer **Koemans**:

Dat is inderdaad een van de stromen die leiden tot trainingscases. Maar ik schetste u eerder al dat er meer zijn.

Mevrouw **Maatoug**:

Er was niet nog een aparte groep mensen met Chinese walls die beoordelingen deden los van de werkstromen die er al waren?

De heer **Koemans**:

Ik begrijp niet helemaal waartussen u die Chinese walls wilt zetten. Ik begrijp niet waar u op doelt.

Mevrouw **Maatoug**:

Ik wil helemaal geen walls zetten. Wat ik probeer te beoordelen of wat ik probeer te vragen – laat ik het nauwkeuriger zeggen – is het volgende. U heeft duidelijk in het verhoor gezegd dat u cases kreeg die door Toeslagen werden aangeleverd.

De heer **Koemans**:

Ja.

Mevrouw **Maatoug**:

Er was ook een beoordeling door experts of er een fout was gemaakt in die zaak waar een mens achter zit, ja of nee.

De heer **Koemans**:

Ja.

Mevrouw **Maatoug**:

Diezelfde beoordeling speelt een rol bij de zaken die u terugkrijgt. Er zijn

meerdere stromen. Dat heeft u heel duidelijk verwoord. Werd het model nou ook getraind met beoordelingen los van de beoordelingen die in de praktijk plaatsvonden, dus dan wel aselect dan wel door een andere groep mensen?

De heer **Koemans**:

Ja. Het model werd ook met andere trainingscases gevoed, namelijk uit die aselecte steekproeven die ik u schetste, en vanuit andere controles die hadden plaatsvonden. Dat waren controles die niet op basis van de risicoselectie tot stand gekomen waren maar op basis van andere mechanismen. Maar er werd wel altijd – want dat was nou juist de voorwaarde – met een menselijk oog naar gekeken om zeker te weten dat we met zekerheid konden uitgaan van het label dat erop geplakt werd, dus het label van hoog risico of laag risico of van een goede of een foute casus. Want heel veel wordt automatisch afgedaan. Daar wilden we niet van uitgaan.

Mevrouw **Maatoug**:

Oké. Duidelijk. Dus u was afhankelijk van de beoordeling vanuit de Dienst Toeslagen en heeft benadrukt dat het belangrijk was dat die beoordeling juist was.

De heer **Koemans**:

Ja.

Mevrouw **Maatoug**:

Is rekening gehouden met eventuele ongewenste gevolgen van het model?

De heer **Koemans**:

Waar doelt u dan specifiek op?

Mevrouw **Maatoug**:

Dat vraag ik aan u.

De heer **Koemans**:

Ik schetste aan het begin bij de taakverdeling dat de afstemming rond het model bij Toeslagen lag.

Mevrouw **Maatoug**:

Is er rekening gehouden met een mogelijk discriminatoir onderscheid tussen groepen op basis van persoonskenmerken, zoals leeftijd, sekse of nationaliteit?

De heer **Koemans**:

Ja, dat waren variabelen. De zaken die u noemt, waren inderdaad onderdeel van het model. Daar is blijkbaar door Toeslagen een afweging in gemaakt. Op het moment dat ik bij het model betrokken raakte, waren die indicatoren al in het model aanwezig.

Mevrouw **Maatoug**:

Ja, en daar bent u gewoon mee verder gaan werken?

De heer **Koemans**:

Ja.

Mevrouw **Maatoug**:

Deloitte heeft van de Belastingdienst de opdracht gekregen om een zelflerend model te creëren. U gaf dat net al aan in het verhoor met mevrouw Simons. Uit onze gegevens blijkt dat het model – in woorden die ook in dit verhoor naar voren komen – steeds "in dezelfde hoek van de vijver vist". Hoe verklaart u dat?

De heer **Koemans**:

Dat weet ik niet.

Mevrouw **Maatoug**:

We hebben net een hele interactie gehad over die selectiebias. Toen had u daar een hele duidelijke uitleg over. U heeft in 2014 een gesprek gehad over het PIA. Meneer Koemans, ik ga het u toch vragen. Bij het model bestond dat risico van die vijver, dat vervolgens ook door collega's van Toeslagen is erkend en bevestigd, waardoor het daarna ook weer verder is doorontwikkeld. Hoe kan het dat dat model steeds in dezelfde hoek van de vijver vist, dat dat risico vergroot is?

De heer **Koemans**:

Oprecht, ik zou niet weten hoe de situatie die u schetst, er heeft kunnen zijn. Ik denk daarom dat het juist zo belangrijk is dat u als commissie het beeld dat u schetst, onderzoekt. Dit is typisch een vraag die op basis van de data die Toeslagen bezit, heel goed te beantwoorden is. Maar ik heb die data niet.

Mevrouw **Maatoug**:

Wat betekent zelflerend precies? Wat hield zo'n model in? Wat was het zelflerende element in het model dat u als Deloitte gemaakt heeft voor Toeslagen?

De heer **Koemans**:

Het belangrijkste aspect was die weging ineens, waarbij een aanvraag dus in één keer van een risicoscore werd voorzien op basis van statistiek, zodat op basis van statistiek beslist kon worden ...

Mevrouw **Maatoug**:

Waarom is dat zelflerend?

De heer **Koemans**:

Misschien mag ik het afzetten tegenover een situatie die niet zelflerend is? Wat niet zelflerend is, is dat een mens bedenkt: iemand meer dan 15 km bij de kinderopvang vandaan woont, dan vind ik dat een verhoogd risico. Hier is juist op basis van statistiek, op basis van die goede en foute casussen, beslist hoe die verdeling in elkaar zit en wat op basis van statistiek verstandig is om te doen.

Mevrouw **Maatoug**:

Oké. Toch even, dus je gaat op basis van expert judgement, op basis van de kennis van mensen in de praktijk, variabelen bedenken. Die data verzamel je. Op basis van die data maak je een model. En dat model laat je dan statistisch de beste relatieordening maken om de beste kansinschatting te maken dat een fout heeft plaatsgevonden van 0 tot 1. Dat heeft u net heel duidelijk verklaard. Nou ja, de hypothese is niet vastgelegd, het is niet opgeschreven, maar dan heeft er toch wel degelijk eerst menselijk denken plaatsgevonden voordat het model gevormd werd?

De heer **Koemans**:

Zeker heeft er menselijk denken plaatsgevonden, maar er zijn twee cruciale punten, denk ik, die het verschil maken. Dat is dat het model op basis van statistiek ook voorstellen doet om bepaalde indicatoren niet te gebruiken. Als de experts bij de opzet van het model hebben aangegeven dat iets een factor zou kunnen zijn maar dat niet wordt bewezen door de statistiek, dan wordt die factor, die indicator, dus niet gebruikt. En bij een aantal van de variabelen, ook nog gecombineerd met een bepaalde grens daarin – ik schetste u het voorbeeld van de afstand tot de kinderopvang – zou een expert kunnen bedenken: ik vind 15 km wel ver. Ja, dat is nogal arbitrair. Als wiskundige vind ik dat je beter op basis van de statistiek kan besluiten welke afstand vaker tot een aanvraag met een verhoogd risico leidt en welke niet.

Mevrouw **Maatoug**:

Dus het zelflerende zit 'm in het feit dat het model soms een variabele wel meeneemt en soms niet?

De heer **Koemans**:

Ja, ook.

Mevrouw **Maatoug**:

En de aftopping van de grens van sommige variabelen?

De heer **Koemans**:

Ja. En het zit 'm er ook in – daar hadden we het hiervoor over – dat de casussen die voor controle geselecteerd en die gecontroleerd zijn, ook weer als voeding voor het model kunnen dienen.

Mevrouw **Maatoug**:

Maar dat is niet zelflerend. Dat is gewoon het model voeden met nieuwe testcases. Dat is toch niet zelflerend, meneer Koemans?

De heer **Koemans**:

Volgens mij is dat een semantische discussie: noemen we dat dan zelflerend of niet?

Mevrouw **Maatoug**:

Oké. Laten we dat niet doen. Wat was de noodzaak om het model zelflerend te maken? Het was onderdeel van de opdracht, maar waarom was dat belangrijk?

De heer **Koemans**:

In het voorbeeld van co-ouderschap had Toeslagen een methodiek die te veel eendimensionaal keek. Dit was een voorbeeld daarvan. Dat was niet gewenst. Daarom was het juist van belang om die variabelen gezamenlijk te wegen.

Mevrouw **Maatoug**:

U gaf net het voorbeeld van afstand. Daar zat de variabele afstand tot de opvanglocatie dan in, toch?

De heer **Koemans**:

Ja.

Mevrouw **Maatoug**:

Was dat een dynamische variabele of was dat er maar eentje? Was de afstand in Rotterdam hetzelfde als bijvoorbeeld in Oost-Groningen? Dus hanteerde het model gewoon één aftoppingsgrens?

De heer **Koemans**:

Voor heel het land, bedoelt u?

Mevrouw **Maatoug**:

Ja.

De heer **Koemans**:

In mijn herinnering voor het hele land inderdaad.

Mevrouw **Maatoug**:

Dat deed het zelflerende model?

De heer **Koemans**:

Ja.

Mevrouw **Maatoug**:

Dus de afstand in Rotterdam was hetzelfde als die in Oude Pekela?

De heer **Koemans**:

Ja.

Mevrouw **Maatoug**:

Het model bepaalde vooraf welke aanvragen eventueel risicovol waren. Betekent dit dat het risicoclassificatiemodel ook een voorspellend model was?

De heer **Koemans**:

In die zin dat het in strikte zin een voorspelling doet over het risico, ja. Het geeft geen absolute uitkomst, zo van: dit is per se een hoog risico en dit is zeker geen risico. Nee. Het is een statistisch model.

Mevrouw **Maatoug**:

Het doet een voorspelling van de kans dat er een fout is gemaakt. Zeg ik dat goed?

De heer **Koemans**:

Ehm, ja.

Mevrouw **Maatoug**:

Klopt het dat het risicoclassificatiemodel alleen werd ingezet op aanvragen voor een verhoging en voor nieuwe aanvragen?

De heer **Koemans**:

In strikte zin werd het model toegepast op alle aanvragen en vervolgens gefilterd op de aanvragen die dan wel een nieuwe aanvraag waren, dan wel tot een verhoging van het toeslagbedrag leidden.

Mevrouw **Maatoug**:

Oké. Dus geen aanpassingen op toeslagen die al ...

De heer **Koemans**:

Of aanvragen die gewoon doorliepen en waar bijvoorbeeld die maand geen wijzigingen in waren. Die zouden in die methodiek niet op die lijst belanden die naar het toezichtteam gaat. Ook wijzigingen die binnengekomen waren en tot een verlaging van toeslagbedrag leidden, kwamen door de filtering niet op de lijst die aan het toezichtteam werd opgeleverd.

Mevrouw **Maatoug**:

In welke digitaaldataomgeving werd het risicoclassificatiemodel gebruikt en welke soort risico's werden geanalyseerd?

De heer **Koemans**:

In de SAS-software.

Mevrouw **Maatoug**:

Waar werd naar gekeken, naar de aanvraag of naar de aanvrager?

De heer **Koemans**:

De aanvraag.

Mevrouw **Maatoug**:

De aanvraag. Betekent dat dat iemand meerdere malen in het risicoclassificatiemodel kon zitten, als er meerdere aanvragen waren gedaan of als er een wijziging was gemaakt?

De heer **Koemans**:

Nee, dat betekent dat iemands aanvraag zowel voor de huurtoeslag als apart voor de kinderopvangtoeslag door het risicomodel ingeschat kan zijn als risico. Dat kan.

Mevrouw **Maatoug**:

Oké. En als ik twee aanvragen heb gedaan, zitten ze er twee keer in.

De heer **Koemans**:

Dat is een beetje technisch. Ik denk niet dat u twee aanvragen kunt doen. U kunt één aanvraag doen en u kunt wijzigingen van die aanvraag indienen. Zo zie ik het. Ik weet niet of dat precies in lijn is met de Toeslagenwet en met hoe het primaire systeem van toeslagen werkt. In het risicoclassificatiemodel kon een persoon altijd maximaal één aanvraag voor de huurtoeslag hebben, en ook maximaal één aanvraag voor de kinderopvangtoeslag.

Mevrouw **Maatoug**:

We hebben het hierover al gehad, maar ik ga het u nog een keer vragen. Als het niet lukt, mag u het in ingewikkeldere taal doen, maar kunt u zo veel mogelijk in simpele taal uitleggen hoe werd bepaald welke combinatie van indicatoren risicovol is en hoe de risicoscore door het model werd berekend?

De heer **Koemans**:

Je zou kunnen zeggen dat dat dus in die rekenmodule van SAS zit, in de statistische maatstaven.

Mevrouw **Maatoug**:

Ja, maar meneer Koemans, ik vraag u hóé dat gebeurde. Kunt u dát vertellen?

De heer **Koemans**:

Bedoelt u achter de schermen, in de statistische software?

Mevrouw **Maatoug**:

Ja.

De heer **Koemans**:

Dat is uiteindelijk matrixrekenen. Een matrix is een soort ... Ik gaf net een Excelsheet als voorbeeld. Er zijn methodes om matrixen met elkaar te vermenigvuldigen. Daar komen maatstaven uit. Die zeggen iets over de vraag of het statistisch al dan niet relevant is om die variabelen te gaan gebruiken. Op basis daarvan doet dit model dat en wordt uiteindelijk het model samengesteld.

Mevrouw **Maatoug**:

We hebben het al gehad over het zelflerende element van het model en de

opdracht. Kunt u in algemene zin vertellen in hoeverre het risicoclassificatiemodel in staat was om te leren en zich aan te passen aan veranderende omstandigheden?

De heer **Koemans**:

Ja. Dat deden wij ... In algemene zin werden de modellen, voor huur- en kinderopvangtoeslag apart, een keer per maand opnieuw geschat. In die zin werd het model elke keer opnieuw geschat en lichtelijk aangepast.

Mevrouw **Maatoug**:

Meneer Koemans, ik ga u vragen ... Ik ben me er heel erg van bewust dat ik die vraag al een paar keer heb gesteld, maar u helpt ons door daar zo duidelijk mogelijk in te zijn. Misschien helpt het als u de matrix van Excel aan de hand van een voorbeeld doet. Dan maken we dus samen even de reis van: je bent een aanvraag en gaat door de verschillende fases heen. Dan kunnen wij wat beter snappen, zo goed als mogelijk, hoe dat nou gaat "onder de motorkap". Volgens mij noemde u dat net zo.

De heer **Koemans**:

Ja, achter de schermen in de statistische software.

Mevrouw **Maatoug**:

Of "achter de schermen", ja.

De heer **Koemans**:

Wilt u dan dat ik ervan uitga dat het model al geschat is en dus toegepast wordt op een toeslagaanvraag? Bedoelt u dat?

Mevrouw **Maatoug**:

Ja.

De heer **Koemans**:

Dan zou het zo kunnen zijn dat ... Misschien kunnen we even uitgaan van een vrij eenvoudig model met maar drie variabelen bijvoorbeeld ...

Mevrouw **Maatoug**:

Nee, ik wil dit graag doen aan de hand van het risicoclassificatiemodel. Ik snap dat u niet uit uw hoofd alle variabelen langsgaat, maar wilt u het alstublieft zo goed als mogelijk ...

De heer **Koemans**:

Dan is bijvoorbeeld eh ...

Mevrouw **Maatoug**:

U mag ook de papieren erbij pakken en wat voorbeelden geven.

De heer **Koemans**:

U zou het zo kunnen zien dat het toepassen van dat model op al die aanvragen eigenlijk gewoon begint bij de bovenste regel van het databestand waar het op toegepast moet worden. Daar staat bijvoorbeeld burgerservicenummertje 1.001. Daar staat bij risicofactor 1 een bepaalde waarde. Het model hangt daar een bepaald gewicht aan. Het maakt die vermenigvuldiging. Dat is even een tussenresultaat. Dan gaat het naar de volgende variabele. Er staat een waarde voor die variabele. Het model weet welke weging het daaraan moet geven. Het telt dat op bij het tussenresultaat van daarvoor. Het gaat naar de derde, naar de vierde ...

Mevrouw **Maatoug**:

Dus "bepaald inkomen, heeft kinderen, oppervlakte van de woning bij de huurtoeslag, en afstand tot aan de opvang", om een paar voorbeelden te noemen.

De heer **Koemans**:

Ja, dat zou zo kunnen, even los van de vraag of dat de relevante variabelen waren in die maand. Inderdaad. Vervolgens moet er nog een wiskundige transformatie plaatsvinden, een berekening om tot dat getal tussen 0 en 1 te komen.

Mevrouw **Maatoug**:

Ja. En dat gewicht – je gaat bij de eerste langs, dan ga je naar de volgende en dan telt het op – is daarvoor al vastgesteld?

De heer **Koemans**:

Juist.

Mevrouw **Maatoug**:

En dat gewicht veranderde per maand?

De heer **Koemans**:

Ja.

Mevrouw **Maatoug**:

Maar niet in hele grote mate, zei u, want de wisselingen waren heel beperkt.

De heer **Koemans**:

Dat klopt.

Mevrouw **Maatoug**:

Kunt u aangeven welke variabelen het meeste gewicht gaven?

De heer **Koemans**:

Nee, dat weet ik niet.

Mevrouw **Maatoug**:

Kunt u dat even opzoeken, meneer Koemans?

De heer **Koemans**:

Nee, dat kan ik niet opzoeken, want in de openbaarmaking van het risicomodel staat alleen welke factoren er gebruikt zijn, maar niet welke weging die precies gehad hebben in elke afzonderlijke maand.

Mevrouw **Maatoug**:

Oké. En ook niet eh ... Want u gaf eerder in dit verhoor aan dat de vaststelling en dat er meer is vastgelegd, datgene wat was vastgelegd en waar verslag van was gedaan ... Dat was in de ontwikkelfase ... dat dat wel had plaatsgevonden.

De heer **Koemans**:

Sorry, wat is uw vraag?

Mevrouw **Maatoug**:

Eerder in het verhoor heeft u aangegeven dat er wel verslaglegging heeft plaatsgevonden, maar eigenlijk alleen aan het begin. Die informatie heeft u dus ook niet ingezien ter voorbereiding van dit openbare verhoor onder ede.

De heer **Koemans**:

Nee.

Mevrouw **Maatoug**:

Oké. Welke gegevensbronnen werden gebruikt om het model te voeden?

De heer **Koemans**:

Dat waren gegevens die Toeslagen al bezat. Die waren al in de systemen aanwezig. Die werden verzameld vanuit verschillende plekken binnen de Belastingdienst.

Mevrouw **Maatoug**:

Kunt u daar wat voorbeelden van noemen?

De heer **Koemans**:

Vanuit de aanvraag zelf komt er natuurlijk informatie over het aantal kinderen en over de uurprijs bij de kinderopvangtoeslag. Maar vanuit de blauwe kant van de Belastingdienst is het bijvoorbeeld van belang wat voor inkomen iemand heeft gehad in het jaar daarvoor. Ik weet zo even niet meer uit mijn hoofd of het inkomen van het jaar daarvoor ook in dat risicomodel zat, maar inkomen is iets wat ook in het model meegewogen is, zeker.

Mevrouw **Maatoug**:

En uitgesloten adressen, ouder kind, bezwaren, hot/hor, dat ik eerder noemde, kot-beschikkingen, conceptbeschikkingen, meldingen FRS en onbekende aanvragers waren ook een paar databronnen. Hoe werden de gegevens ...

De heer **Koemans**:

Als ik u even mag onderbreken: ik twijfel of alle variabelen die u noemt, ook onderdeel van het model waren.

Mevrouw **Maatoug**:

Maar u heeft net al een aantal keren heel duidelijk verklaard dat niet alles gebruikt werd, dat er in het model ook dingen zaten die niet altijd meegenomen werden in een run en dat er elke maand andere variabelen waren. Dat heeft u heel duidelijk gemaakt, meneer Koemans.

De heer **Koemans**:

Ja.

Mevrouw **Maatoug**:

Hoe werden de betrouwbaarheid en relevantie van de gegevens uit deze bronnen geëvalueerd?

De heer **Koemans**:

Door het statistische model.

Mevrouw **Maatoug**:

Werd er gebruikgemaakt van kunstmatige intelligentie om het zelflerende vermogen van het risicoclassificatiemodel te verbeteren?

De heer **Koemans**:

In mijn ogen is logistische regressie geen kunstmatige intelligentie, maar daar valt over te discussiëren.

Mevrouw **Maatoug**:

Heel duidelijk, het antwoord is nee.

De heer **Koemans**:

In mijn ogen is het nee.

Mevrouw **Maatoug**:

Werden er statistische analyses uitgevoerd om de effectiviteit van het zelflerende vermogen van het model te meten en te evalueren?

De heer **Koemans**:

Kunt u dat nog een keer herhalen?

Mevrouw **Maatoug**:

Of er statistische analyses werden uitgevoerd om de effectiviteit van het zelflerende vermogen van het model te meten en te evalueren. U gaf net aan dat er twee elementen zijn waardoor het zelflerend is. De eerste is dat er testcases zijn. De tweede is dat bepaalde grenzen worden aangegeven. Is weleens geëvalueerd hoe goed die beperkte elementen van het zelflerende waren?

De heer **Koemans**:

In mijn herinnering niet anders dan wat ik u al schetste. Punt.

Mevrouw **Maatoug**:

Waren er beperkingen of risico's verbonden aan het zelflerende en voorspellende vermogen van het model?

De heer **Koemans**:

Zeker heeft het gebruik van een statisch model beperkingen.

Mevrouw **Maatoug**:

Hoe werden risico's gemitigeerd?

De heer **Koemans**:

Doelt u dan op risico's van het gebruik van het model? Waar doelt u precies op?

Mevrouw **Maatoug**:

Risico's verbonden aan het zelflerende en voorspellende vermogen.

De heer **Koemans**:

Hoofdzakelijk door ervoor te zorgen dat de trainingscases uit verschillende bronnen kwamen en door het risico door een menselijk oog te laten beoordelen.

Mevrouw **Maatoug**:

Wie was er verantwoordelijk voor de werking van het model? Wie hield er toezicht op het zelflerende proces?

De heer **Koemans**:

Ehm ... In technische zin was de technische dienst er voor het zelflerende karakter en voor het maandelijks schatten van het model. Is dat een antwoord op uw vraag?

Mevrouw **Maatoug**:

Ik stelde een vraag en u gaf een antwoord, en als dat antwoord is, is dat het antwoord. U heeft nu een paar keer het technische element benadrukt. Vervolgens kwamen we op de taakverdeling. Is die hele harde scheiding tussen dat technische element en de andere elementen gebruikelijk?

De heer **Koemans**:

Ja, dat is wel gebruikelijk, in die zin dat ik het vanuit mijn technische expertise ook als mijn verantwoordelijkheid voel om het overleg aan te gaan met mensen met andere taken en waar mogelijk te helpen, uiteraard.

Mevrouw **Maatoug**:

U bent manager analytics. Daarvoor was u senior manager bij Deloitte. Is het, om een kwalitatief hoogstaand model af te leveren, een randvoorwaarde om ook van de materie op de hoogte te zijn?

De heer **Koemans**:

Dan bedoelt u specifiek op het gebied van, in dit geval, toeslagen?

Mevrouw **Maatoug**:

Ik bedoel het in algemene zin.

De heer **Koemans**:

In algemene zin denk ik dat het van belang is dat die expertise aanwezig is in het projectteam, maar theoretisch kan je prima een statistisch model maken op een onderwerp waar de techneut an sich geen verstand van heeft, als er maar voor gezorgd wordt dat het projectteam als geheel die expertise wel heeft. Samenwerking, elkaar aanvullen.

Mevrouw **Maatoug**:

Uit ons onderzoek is naar voren gekomen dat u ook betrokken bent geweest bij de ontwikkeling van de risicoscan verblijf in het buitenland van het UWV. Deze scan is van latere datum, namelijk van 2017. Kunt u de commissie vertellen wat de verschillen en de overeenkomsten zijn tussen de risicoscan van het UWV en het risicoclassificatiemodel van Toeslagen?

De heer **Koemans**:

Ehm ... Ik ben inderdaad ook bij het UWV-project betrokken geweest, helemaal in de beginfase van dat project, maar niet meer bij het technisch werkend krijgen van dat model. Ik heb daar ook niet de data van gezien, dus ik vind het lastig om de verschillen en de overeenkomsten te benoemen.

Mevrouw **Maatoug**:

Maar dit zijn allebei projecten ... U geeft heel duidelijk aan dat u verschillende rollen hebt gehad en dat het ging om verschillende fases. Dat is heel duidelijk, meneer Koemans. Maar kunt u in algemene zin iets zeggen over verschillen en overeenkomsten?

De heer **Koemans**:

Ik denk dat de overeenkomst is dat er zo'n model gebruikt is. Verschillen zullen er te over zijn omdat er heel andere data gebruikt worden, maar dat is een beetje een open deur.

Mevrouw **Maatoug**:

Ik geef het woord aan mevrouw Simons.

Mevrouw **Simons**:

Dank, collega. Meneer Koemans, ik heb tot slot nog enkele vragen aan u over het maandelijks gebruik van het risicoclassificatiemodel door Toeslagen.

Kreeg u weleens feedback van de Dienst Toeslagen over hoe het model maandelijks werd toegepast en met welk resultaat?

De heer **Koemans**:

Zeker. In dat percentage, dat ook door de experts van Toeslagen als hoog risico werd beoordeeld, kwam dat zeker terug, ja.

Mevrouw **Simons**:

Kreeg u die feedback op schrift, via de e-mail of ...

De heer **Koemans**:

Mondeling.

Mevrouw **Simons**:

Mondeling. Ook daar zijn geen rapporten, verslagen, notulen en dat soort dingen van?

De heer **Koemans**:

Nee.

Mevrouw **Simons**:

U wilde er al iets over vertellen. Hoe luidde die feedback? U kreeg weleens feedback. Hoe luidde die?

De heer **Koemans**:

Positief. Het percentage van 90% tot 95% werd als hoog ervaren. Daar was men tevreden mee, zeker.

Mevrouw **Simons**:

Van wie kreeg u die feedback?

De heer **Koemans**:

Formeel van de opdrachtgever, maar in de praktijk van de teamleider Handhavingsregie en ook van andere betrokkenen, die dan met dat percentage naar ons terugkwamen, als het ware.

Mevrouw **Simons**:

Nou hebben we ook begrepen dat ... Corrigeert u mij als ik het verkeerd zeg. U heeft er ook al even aan geraakt in het gesprek met mijn collega Maatoug. U zegt: 95% van wat het model uitwierp, werd bevestigd als zijnde correct door medewerkers van de Dienst Toeslagen.

De heer **Koemans**:

Ja, correct in de zin dat het hoge risico ook door hen werd onderkend.

Mevrouw **Simons**:

Nou was dat min of meer, al dan niet in zijn geheel, dezelfde groep mensen die in eerste instantie betrokken waren bij het selecteren van de risico-indicatoren. Dus er is ervarings...

De heer **Koemans**:

Ik denk het niet.

Mevrouw **Simons**:

Dat denkt u niet?

De heer **Koemans**:

Dat denk ik niet. Nou ... Ik sta onder ede; het formele antwoord is: ik was niet bij dat beruchte kroningsweekend, dus ik weet niet precies hoe de indicatoren tot stand zijn gekomen en welke mensen het idee hebben gehad om specifieke risico-indicatoren in het model op te nemen.

Mevrouw **Simons**:

Ik hoor een komma, meneer Koemans.

De heer **Koemans**:

Ja, en toen was ik uw vraag even kwijt ... U vroeg of dit dezelfde mensen waren.

Mevrouw **Simons**:

Mijn vraag was eigenlijk dit. Er werd een percentage gevallen, casussen, uitgeworpen.

De heer **Koemans**:

Laten we het zo even noemen, ja.

Mevrouw **Simons**:

Dan gingen de medewerkers van de Dienst Toeslagen checken of zij het, op basis van hun expertise en jarenlange ervaring binnen de dienst, eens waren met het feit dat deze casussen werden uitgeworpen als zijnde een verhoogd risico.

De heer **Koemans**:

Ja.

Mevrouw **Simons**:

U zegt dat u er in dat weekend niet bij was, maar het heeft er alle schijn van dat in dat weekend ook indicatoren zijn aangedragen door mensen van de Dienst Toeslagen met jarenlange ervaring met het behandelen van casussen.

De heer **Koemans**:

Dat klinkt logisch als u dat zo zegt, maar nogmaals, ik was er niet bij. Ik weet ook niet of het dezelfde mensen waren die de beoordeling deden die later tot die 90% tot 95% leidde. Dat weet ik niet.

Mevrouw **Simons**:

We kunnen hooguit stellen dat het ook mensen waren die jarenlange ervaring hadden in de uitvoering bij de Dienst Toeslagen. Of dat dezelfde mensen waren, weten we niet. Daarvan is niets gedocumenteerd. Maar we weten wel dat er mensen aanwezig waren die ... U gebruikte zonet niet het woord "ervaringsdeskundige" maar een ander woord.

De heer **Koemans**:

"Experts", denk ik.

Mevrouw **Simons**:

Het had met ervaring te maken. Weet u dat nog?

De heer **Koemans**:

Nee.

Mevrouw **Simons**:

Mensen met ervaring in de uitvoering bij Toeslagen. Die krijgen een uitworp van een statistisch risicoclassificatiemodel en gaan dan op basis van hun ervaring kijken of zij vinden dat dat klopt, dat het inderdaad casussen betreft met een hoog risico op fouten en/of ... U zegt: en fouten en fraude.

De heer **Koemans**:

Ja. Dat heb ik begrepen. Op basis van die beoordeling werden er vragen gesteld aan de aanvrager en werden er eventueel bewijsstukken opgevraagd.

Mevrouw **Simons**:

Wat er ook gebeurde, meneer Koemans, is dat ze bij u terugkwamen en zeiden: wij zien dat het goed is, wij zien dat 95% van deze gevallen inderdaad een hoog risico betreft. Wat er ook gebeurde, meneer Koemans, is dat deze data vervolgens weer werden teruggevoerd in het model.

De heer **Koemans**:

Maar pas na een grondigere beoordeling, inclusief het contact opnemen met de burger en die bewijsstukken. Zo begreep ik dat.

Mevrouw **Simons**:

Daarna, zegt u, werd het pas weer teruggegeven aan het model.

De heer **Koemans**:

Ja, zeker. Want dat percentage was op de dag van de risicorun van die maand meteen nodig. Dat percentage was nodig om te beslissen over het al dan niet accepteren van dat model.

Mevrouw **Simons**:

Hoe ging het dan terug het model in? Als zijnde: "Inderdaad, dit was een hoog risico. Het klopt. Fijn dat we het gecheckt hebben. Dank u wel."? Wat ging er dan terug het model in?

De heer **Koemans**:

Na het contact met de burger, zo begrijp ik, kon dat in het model ingevoerd worden, in de zin van een nieuwe goed- dan wel foutcasus.

Mevrouw **Simons**:

Dus het model werpt hoogrisicocasussen uit. Die worden door mensen gecheckt. De burger wordt gebeld. En wat doet dan deze medewerker nadat die de burger heeft gebeld en heeft geconstateerd dat er misschien al dan niet fouten zijn gemaakt?

De heer **Koemans**:

Dan komt hij tot een finaal oordeel over die toeslag aanvraag en dat oordeel, voorzien van dat aanvraagnummer, wordt teruggevoerd in het model.

Mevrouw **Simons**:

Wat voegt hij dan toen aan het model, los van de casus? Dan moet hij ook toevoegen "deze was uitgeworpen als zijnde fout, maar na een telefoontje is

gebleken van niet" of "deze is uitgeworpen als zijnde fout en is inderdaad fout gebleken".

De heer **Koemans**:

Ja. Dat is dat label waar ik op doelde. Bij elke trainingscasus wordt ook aangegeven of het een goed- of juist een foutcasus is.

Mevrouw **Simons**:

Wat vindt u van deze werkwijze? Is dat statistisch gezien, wetenschappelijk gezien een betrouwbare manier van werken?

De heer **Koemans**:

Ik sta daarachter, ja.

Mevrouw **Simons**:

Ja?

De heer **Koemans**:

Ja.

Mevrouw **Simons**:

Ja is ook een antwoord, hè.

De heer **Koemans**:

Ja.

Mevrouw **Simons**:

Er zijn geen goede of foute antwoorden, slechts uw antwoorden.

Het eindproduct van het risicoclassificatiemodel was dus een lijst met de meest risicovolle aanvragen van die maand. U vertelde net dat er gebeld werd. Hoe heeft u zich ervan vergewist dat op deze manier een vervolg werd gegeven aan die uitwerpen?

De heer **Koemans**:

Niet, behalve in de zin dat er casussen terug bij ons kwamen als trainingscases, maar ook daar is niet ...

Mevrouw **Simons**:

U weet niet precies op welke manier er vervolg aan werd gegeven door

medewerkers van de Dienst Toeslagen met betrekking tot die risicovolle aanvragen?

De heer **Koemans**:

Nee, nee, nee. Dat was echt duidelijk niet mijn expertise. Dat was duidelijk in de taakverdeling niet bij mij belegd.

Mevrouw **Simons**:

Weet u of er weleens aselechte controles hebben plaatsgevonden na de ingebruikname van het model?

De heer **Koemans**:

Ik weet dat wij aselechte steekproeven hebben opgeleverd en een deel daarvan ook teruggekoppeld hebben gekregen als trainingscases.

Mevrouw **Simons**:

Heeft u daarover gerapporteerd? Is dat beschreven? Zijn daar documenten van?

De heer **Koemans**:

Nee.

Mevrouw **Simons**:

In hoeverre, is mijn vraag aan u tot slot, was het risicoclassificatiemodel daadwerkelijk zelflerend? We hebben net geconstateerd in uw gesprek mijn collega dat het niet zozeer zelflerend was, maar u vertelt mij nu ook net dat na een menselijke interventie informatie wordt teruggegeven aan het model. Is het dan zelflerend of is het model ook afhankelijk van wat er dan weer teruggegeven wordt aan dat model en hoe, door menselijke tussenkomst?

De heer **Koemans**:

Daar zat zeker een menselijke tussenkomst in. Absoluut. Ja.

Mevrouw **Simons**:

In hoeverre is het model dan zelflerend?

De heer **Koemans**:

Dat vind ik een ingewikkelde vraag. Op de punten die ik beschreef, is het zelflerend. Of hét model daarmee dan zelflerend is, vind ik een ingewikkelde vraag.

Mevrouw **Simons**:

Deloitte is aangetrokken om een bestaand model zelflerend te maken. Hoe is dat gelukt? Is het model zelflerend volgens u, meneer Koemans?

De heer **Koemans**:

In mijn ogen is het zelflerend.

Mevrouw **Simons**:

Op welke manier is het zelflerend? Wat leert het zelf?

De heer **Koemans**:

Een van de twee cruciale factoren is dat het model op basis van statistiek beslist of voorstelt eigenlijk – er zit namelijk nog een acceptatie tussen – op basis van statistiek kan bepalen welke factoren op welke manier van belang zijn. En van de uitkomst kan vervolgens via die menselijke tussenkomst – het klopt helemaal wat u zegt – uiteindelijk weer geleerd worden, doordat dat een nieuwe trainingscase wordt.

Mevrouw **Simons**:

Ligt het aan mij, meneer Koemans, of zijn dat beide geen voorbeelden van een zelflerend proces?

De heer **Koemans**:

In mijn ogen zijn dat voorbeelden van een zelflerend proces.

Mevrouw **Simons**:

Dus door menselijke tussenkomst wordt er nieuwe informatie toegevoegd en u zegt: dan is het zelflerend.

De heer **Koemans**:

Dat is wat wij daarmee bedoelden. Volgens mij was dat ook het doel van de opdracht. Dat is het zelflerende karakter.

Mevrouw **Simons**:

Oké. Dan ga ik u vragen: was het model ook geschikt om ingezet te worden op de manier zoals het door de Dienst Toeslagen is ingezet, met de kennis van vandaag?

De heer **Koemans**:

Ehm ... Ik kan als getuige hier niet verklaren over hoe Toeslagen dat model allemaal precies heeft ingezet. Ik weet wel dat we gedurende de opdracht

gesignaleerd hebben dat het model ook gebruikt werd voor processen buiten wat ik u schetste ...

Mevrouw **Simons**:

Kunt u heel specifiek zijn? Wat schetste u op welk moment? U zegt: buiten wat ik u schetste.

De heer **Koemans**:

Buiten het opleveren van die lijst aan dat toezichtteam en dat percentage dat daarop terugkwam, begrepen wij in de loop van die opdracht dat het ook op andere plekken gebruikt werd, op andere manieren. We hebben er ook voor gewaarschuwd dat het model daar niet geschikt voor was.

Mevrouw **Simons**:

Dat het daar niet geschikt voor is. Is het verantwoord – u heeft daar al iets over gezegd – om de pilotfase over te slaan bij dergelijke grote processen?

De heer **Koemans**:

Dat hangt in mijn ogen helemaal af van de T van testen in die OTAP-straat en hoe daarnaar gekeken wordt.

Mevrouw **Simons**:

En in dit specifieke geval?

De heer **Koemans**:

Weet ik de resultaten van het kroningsweekend niet, dus daar kan ik niet over oordelen.

Mevrouw **Simons**:

Heeft u enig idee of het feit dat de pilotfase is overgeslagen, enige invloed heeft gehad? Heeft u in een later stadium, terwijl u met het project bezig was, gemerkt dat er misschien op sommige plekken dingen anders hadden moeten of gekund die wellicht boven waren gekomen bij een pilot?

De heer **Koemans**:

Ehm ... Nee, dat kan ik me niet herinneren. Ik zie het bovendien ook niet als dat een pilotfase overgeslagen zou zijn. Nee.

Mevrouw **Simons**:

Nee. Die is volgens u wel uitgevoerd, maar binnen dat weekend van 27 tot en met 30 april?

De heer **Koemans**:
Zo begrijp ik dat.

Mevrouw **Simons**:
Toen is er getest?

De heer **Koemans**:
Wij testten in ieder geval in iedere maandelijks run inderdaad. Dan werd er getest en werd er beslist over acceptatie van dat model of niet.

Mevrouw **Simons**:
Als u vandaag weer een dergelijke opdracht zou krijgen, zou u dan akkoord gaan met het snel ontwikkelen, testen en in werking zetten binnen drie dagen?

De heer **Koemans**:
Met de kennis van nu is dat inderdaad lastig.

Mevrouw **Simons**:
Zou u dat nu, op dit moment – het is een gesloten vraag: ja of nee – acceptabel vinden?

De heer **Koemans**:
Dat hangt van de toepassing van het model af.

Mevrouw **Simons**:
En in mijn vraag zei ik: een dergelijk model met een dergelijke toepassing, zoals dit model door de Dienst Toeslagen van de Nederlandse Belastingdienst is geïmplementeerd. Zou u het op dit moment verstandig vinden om dat te doen volgens hetzelfde proces als nu is gebeurd?

De heer **Koemans**:
Nee, dat weet ik echt niet.

Mevrouw **Simons**:
Zou u het op dit moment bij een dergelijke opdracht van een dergelijke opdrachtgever weer acceptabel vinden als daar geen enkele vorm van verslaglegging plaatsvindt?

De heer **Koemans**:
Nee.

Mevrouw **Simons**:

Wat is het grootste bezwaar, wat is uw grootste bezwaar, tegen het feit dat er geen enkele verslaglegging is van dit hele proces?

De heer **Koemans**:

Dat dat geen professionele benadering is van zo'n project en dat daarmee ook beslissingen die genomen zijn, niet terug te zoeken zijn.

Mevrouw **Simons**:

Vindt u dat dat een verantwoordelijkheid was of in de toekomst zal moeten zijn van een organisatie als Deloitte?

De heer **Koemans**:

Zeker. Ook. Ja, daar ligt absoluut een gezamenlijke verantwoordelijkheid. Die projecten doen wij ook heel anders.

Mevrouw **Simons**:

Ik wil toch nog – dat is mijn allerlaatste vraag – weten waarom u in de loop van dat project, ook al bent u later ingestapt, op geen enkel moment vanuit diezelfde verantwoordelijkheid het gevoel hebt gehad om dingen toch maar eens te gaan documenteren.

De heer **Koemans**:

Nou, bepaalde dingen zijn wel degelijk gedocumenteerd, ook in de loop van het project, maar deze punten, die acceptatiebeslissingen, niet. Maar de controles op de inputdata bijvoorbeeld en de vraag welke indicatoren precies in welke maand gebruikt werden, zijn wel gedocumenteerd. Het is dus niet zo dat er helemaal niets gedocumenteerd werd, maar een aantal cruciale dingen zeker niet en dat had wel moeten.

Mevrouw **Simons**:

Dank u wel. Ik geef het woord aan onze voorzitter.

De **voorzitter**:

Dat betekent dat we bijna aan het einde van dit verhoor zijn gekomen. Ik heb toch nog één vraag. U zei net dat u gewaarschuwd had ... Nee, eerst het volgende. U gaf aan dat het model ook gebruikt werd op andere plekken dan waar het voor bedoeld was. Kunt u daar iets over vertellen?

De heer **Koemans**:

Ik weet niet meer precies op welke plekken, maar ik heb dat wel in de

stukken teruggelezen. Ik heb ook gelezen dat we daarvoor gewaarschuwd hebben, inderdaad.

De **voorzitter**:

Kunt u dat iets specifieker maken? Waar heeft u voor gewaarschuwd?

De heer **Koemans**:

Waar hebben we voor gewaarschuwd? Ik kan u citeren uit de stukken. Daarin staat: "We herkennen dat de risicoscores in de organisatie een steeds groter belang krijgen, bijvoorbeeld in de green lane maar ook in ad-hoetoezichtacties. Dit is een vooruitstrevende ontwikkeling, maar vereist volgens ons – daar staat nog een spelfout – ook een model dat uitstekend, in plaats van goed, gefundeerd en volledig up-to-date is. Bovendien zien wij een beter begrip van de betekenis van de scores binnen de organisatie als essentiële voorwaarde." Dit is tegen de achtergrond dat we in diezelfde periode ook constateerden dat het model, toen Toeslagen het zelf in gebruik had, een aantal maanden niet meer opnieuw getraind was.

De **voorzitter**:

Oké. Waar leest u dit uit voor?

De heer **Koemans**:

Dit lees ik voor uit een stuk van het allerlaatste contact dat ik in dit project met Toeslagen heb gehad.

De **voorzitter**:

Heeft dat een naam? Is dat een formeel rapport of was het een memo van u? Wat was het?

De heer **Koemans**:

Dat stuk heet overzicht Update risicoselectie – overzicht werkzaamheden, Deloitte 4 januari 2016. Wij werden formeel ingehuurd tot en met december 2015. Dit stuk is nog verstuurd op de eerste werkdag van het nieuwe jaar.

De **voorzitter**:

Aan wie was dat gericht?

De heer **Koemans**:

Aan de teamleider Handhavingsregie en Intelligence.

De **voorzitter**:

U zei dat dat daarna werd gestuurd. U weet dus niet meer wat daarmee gebeurd is, of ze hebben geluisterd naar de waarschuwingen die u hebt gegeven?

De heer **Koemans**:

Nee, daar kan ik u niks over zeggen.

De **voorzitter**:

Daar kunt u niks over zeggen of dat wilt u niet?

De heer **Koemans**:

O. Dat weet ik niet.

De **voorzitter**:

Oké. Dan heb ik tot slot nog een andere vraag. Mevrouw Simons vroeg het eigenlijk ook al. Op een gegeven moment tijdens het verhoor gaf u als antwoord op de vraag bij wie de verantwoordelijkheid ligt bij ongewenste gevolgen: de ongewenste gevolgen vanuit het model liggen bij Toeslagen. Toch? Of heb ik u verkeerd geciteerd?

De heer **Koemans**:

Formeel is dat het antwoord, denk ik, ja, ja, ja, in de zin van: natuurlijk voelden wij de verplichting om als projectteam gezamenlijk, waar wij een onderdeel van waren, een goed project neer te zetten. Je hard maken voor voldoende trainingscases zie ik bijvoorbeeld als onderdeel van de morele verantwoordelijkheid om het project als geheel tot een succes te maken.

De **voorzitter**:

Uw organisatie maar ook vele andere organisaties worden ingehuurd. Die krijgen dan een bepaalde opdracht, ondersteunen en geven adviezen, maar zijn dan in the end niet eindverantwoordelijk als er negatieve gevolgen of zelfs destructieve gevolgen uit voortkomen. Is er nou een discussie gaande over de vraag of dat juist wel een verantwoordelijkheid zou moeten zijn van verschillende corporate organisaties die adviezen geven?

De heer **Koemans**:

Zoals bij mijn werkgever, bedoelt u?

De **voorzitter**:

Uw werkgever, ja.

De heer **Koemans**:

Ik weet dat wij inderdaad een voorkeur hebben voor meer verantwoordelijkheid in projecten met een volledig team van verschillende expertises. Met verschillende expertisegebieden in zo'n project werkzaam zijn heeft zeker onze voorkeur. Ook in dit geval, bij Toeslagen, heeft dat onze voorkeur gehad, begrijp ik van collega's. Dat is ook aangeboden. Toeslagen heeft ervoor gekozen om alleen de expertise op het technische stuk van ons in te huren.

De **voorzitter**:

Wat zijn daar de implicaties van? Wat zegt u nu eigenlijk? U zegt: wij hebben die expertise, die hebben we aangeboden, maar Toeslagen heeft gezegd dat dat niet hoeft, dat ze alleen de techniek doen. Waarom is het belangrijk dat je alle expertise aan boord hebt?

De heer **Koemans**:

Om uiteindelijk met het hele projectteam ervoor te zorgen dat het hele project een succes wordt. Zeker.

De **voorzitter**:

En dus ook om negatieve gevolgen te voorkomen.

De heer **Koemans**:

Absoluut.

De **voorzitter**:

Weet u nog wie bij Toeslagen heeft gezegd om maar wat minder adviseurs van Deloitte binnen te halen?

De heer **Koemans**:

Nee, dat weet ik niet, want dat is niet aan mij gezegd. Ik heb dit van collega's begrepen.

De **voorzitter**:

Weet u nog wat uw collega heeft gezegd over waarom dat zo was?

De heer **Koemans**:

Nee.

De **voorzitter**:

Los van het gegeven om meer mensen binnen te halen, was mijn vraag wat

breder, namelijk: moet je er wat steviger in zitten als het gaat om ongewenste gevolgen, of blijft het alleen bij waarschuwen en houdt het daarna op omdat dat geen onderdeel van de opdracht was? Wat vindt u daarvan als hierop terugkijkt? Is er, los van alle gespecificeerde antwoorden die u heeft gegeven, iets wat je hieruit zou kunnen leren?

De heer **Koemans**:

En los van het feit dat het in onze organisatie vooral de verantwoordelijkheid is van de partner of director die de opdracht verkoopt en in die zin aanneemt. Maar als persoon, als u mij toestaat, vind ik het zeker een goede ontwikkeling als je die verantwoordelijkheid als bedrijf meer neemt. Als ik hierop terugkijk als burger, als u mij toestaat, als geïnteresseerde burger die dit dossier ook volgt, dan is het natuurlijk verschrikkelijk wat de slachtoffers van de toelagenaffaire is overkomen. Dat had nooit mogen gebeuren. Juist daarom is het ook zo belangrijk dat u dit als commissie onderzoekt.

De **voorzitter**:

Dank u wel voor dit verhoor.

Sluiting 16.02 uur.
