

SCHRIFTELIJKE VRAAG

nr. 10
van **DANIËLLE VANWESENBEECK**
datum: 5 oktober 2018

aan **PHILIPPE MUYTERS**
VLAAMS MINISTER VAN WERK, ECONOMIE, INNOVATIE EN SPORT

Kmo-portefeuille - Uitrol automatisch fraudedetectiesysteem

In antwoord op mijn vraag om uitleg nr. 2099 gaf de minister op 14 juni 2018 in de commissie Economie een overzicht van de stand van zaken inzake de ontwikkeling van een automatisch fraudedetectiesysteem voor de kmo-portefeuille. Begin mei werd een ontwikkelaar aangesteld. Een eerste versie zou eind september operationeel zijn.

1. Wat is de huidige stand van zaken? Welke stappen werden inmiddels gezet? Is inmiddels een systeem operationeel?
2. Kan de minister de concrete werking en functionaliteiten toelichten? Welke controles gebeuren door het systeem? Op welke parameters wordt gecontroleerd? Op welk niveau wordt gecontroleerd (dienstverleners, aanvragers)?
3. Welke beperkingen zijn er nog? Welke verdere stappen inzake ontwikkeling zijn nog nodig?

ANTWOORD

op vraag nr. 10 van 5 oktober 2018

van **DANIËLLE VANWESENBEECK**

1. De eerste versie van het systeem werd opgeleverd op 15 oktober 2018 en wordt momenteel intensief getest. Data van de kmo-portefeuilledatabank worden automatisch opgeladen in het systeem en gemodelleerd om gestandaardiseerde analyses mogelijk te maken. Die analyses starten automatisch en leveren een 'frauderisico-score' op voor elke subsidieaanvraag.
2. Men kan in het systeem aangeven voor welke periode men de aanvragen wenst te onderzoeken (dit kan bijvoorbeeld de voorbije maand zijn). Om tot de frauderisico-score te komen worden twee types van fraude opsporingstechnieken gecombineerd:
 - **Anomaliedetectie:** alle dienstverleners en dienstverkrijgers met mekaar vergeleken op basis van een set van indicatoren (bv. aantal projecten, gemiddelde gevraagde subsidie, omvang subsidie versus omzet, tijd tussen erkenning en eerste aanvraag etc.). Vervolgens worden de dienstverleners onderling vergeleken en in clusters gegroepeerd. Idem voor de dienstverkrijgers. Dienstverleners/-verkrijgers die statistisch afwijken van andere dienstverleners/-verkrijgers krijgen een 'anomaliescore'. Hoe groter de afwijking, hoe hoger de score. De uiteindelijke anomaliescore van een project is de som van de anomaliescores van de dienstverlener en de dienstverkrijger.
 - **Netwerkanalyse:** ook hier wordt een score toegekend aan elke dienstverlener en elke dienstverkrijger. De uiteindelijke netwerkscore per project is de som van de netwerkscores van de dienstverlener en de dienstverkrijger. De netwerkscore per dienstverlener en dienstverkrijger is ook een gecombineerde score:
 - In de eerste plaats wordt onderzocht of dienstverleners en dienstverkrijgers verbonden zijn met elkaar. Hiervoor worden alle 'relationele data' die we ter beschikking hebben van de dienstverkrijger zoals adres, rekeningnummer, telefoonnummer, ... vergeleken met de data van de dienstverlener. Hoe meer overeenstemming er gevonden wordt, hoe hoger de netwerkscore. Dienstverkrijgers die bijvoorbeeld op hetzelfde adres zitten als een dienstverlener krijgen hierdoor een hogere netwerkscore toegekend.
 - Daarnaast wordt ook een indicatie gegeven van de grootte van 'het netwerk' van ondernemingen rond de dienstverleners en de dienstverkrijgers. Hiervoor wordt het fourhubs algoritme gehanteerd.

Het resultaat is een rapport met informatie over het project, de frauderisico-score (en de individuele scores zoals hierboven vermeld om interpretatie van de frauderisico-score mogelijk te maken). De projecten worden gerangschikt volgens frauderisico-score.

3. Er worden onderzocht hoe de gegevens voor analyse uitgebreid kunnen worden, enerzijds met andere databanken en anderzijds door uitbreiding van de gegevens in de kmo-portefeuilledatabank. Anderzijds wordt gewerkt aan een verfijning van de indicatoren op basis van de eerste testen.