

# Data scientist bij de provincie Groningen

door Willy Tadema voor De Jonge Onderzoekers, oktober 2018



**OpenStreetMap**  
The Free Wiki World Map



**THE THINGS**  
NETWORK



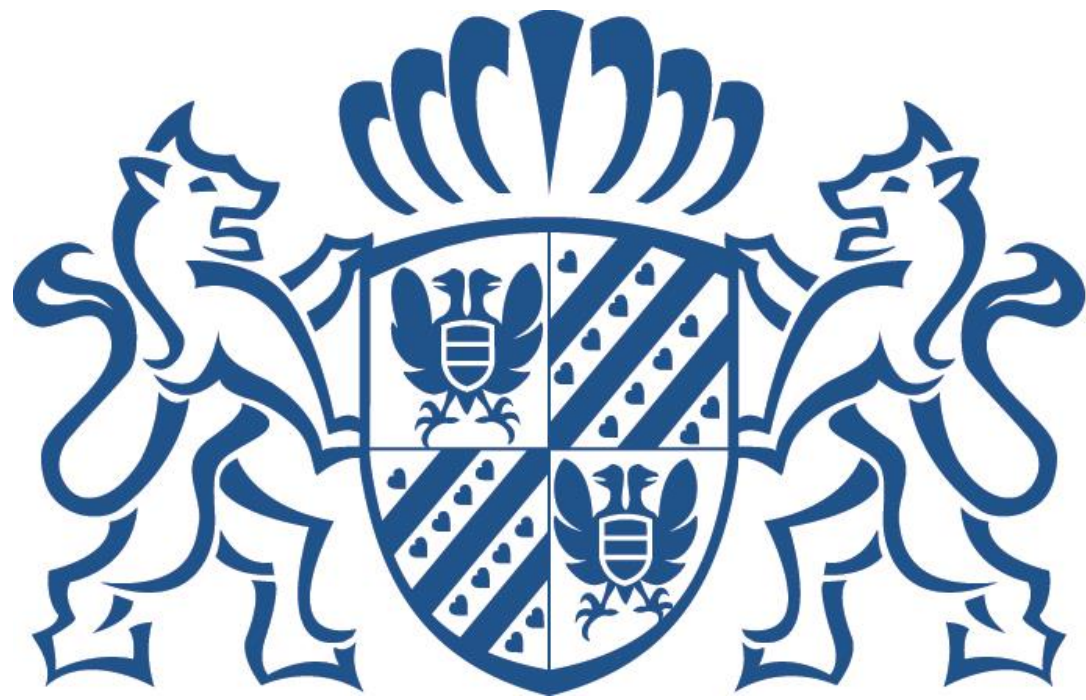
**Hack de Overheid**



**Missing  
Maps**







**provincie**  
**groningen**

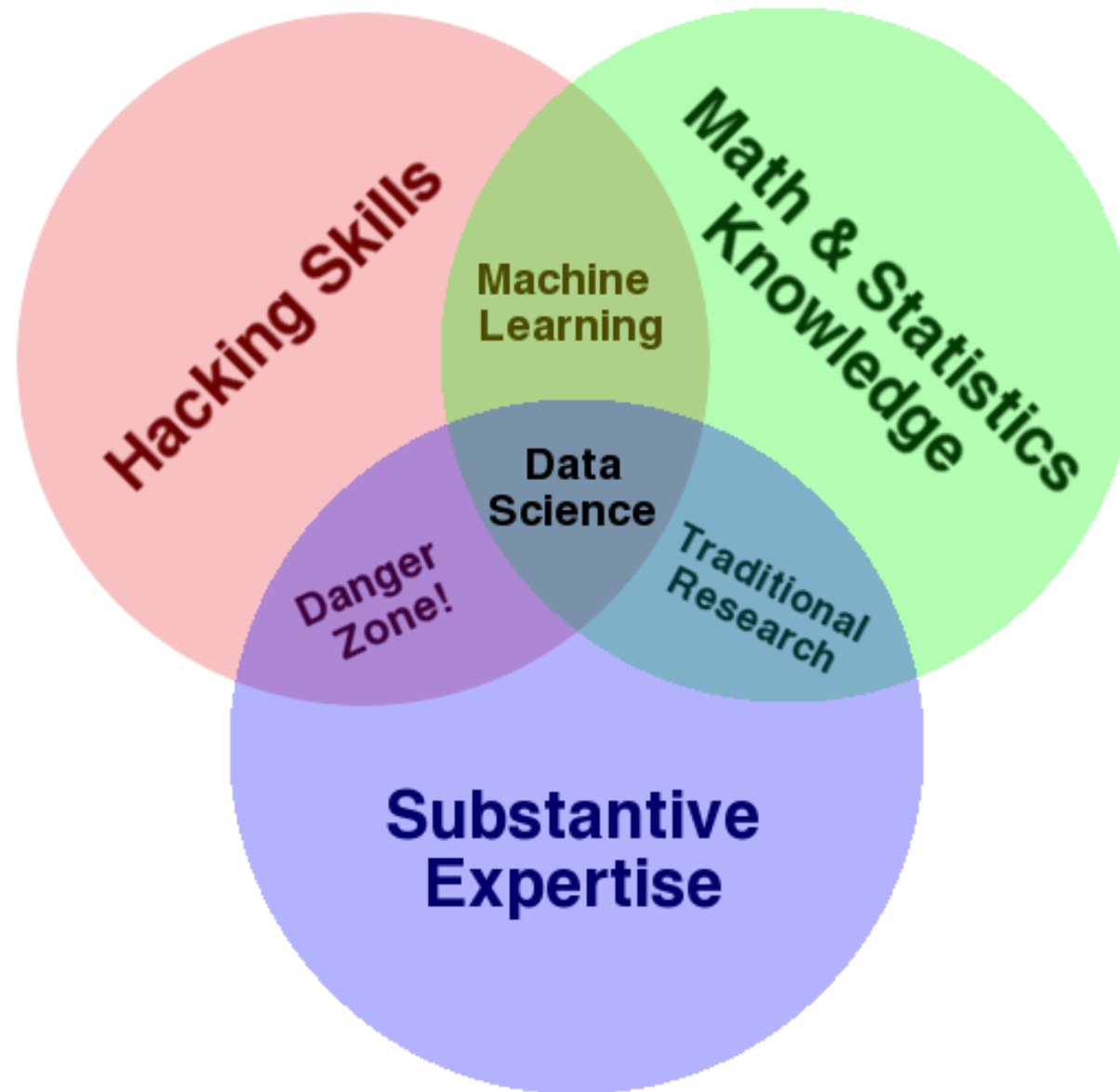
Wat is een data scientist?

# ‘Data Scientist: The Sexiest Job of the 21st Century’

‘A data scientist is someone  
who knows more statistics than a computer scientist  
and more computer science than a statistician’

‘A data scientist is  
someone who extracts insights  
from messy data’





# Team Datakracht





# Waarom is data belangrijk voor de overheid?

- besluiten op basis van juiste informatie
- verantwoording afleggen
- efficiëntere en effectievere publieksdiensten

# De 'kluutcrisis': over het Dollardkanaal dat er niet kwam

‘Andere tijden, andere zeden’. De strekking van dit gezegde ligt ten grondslag aan een van de grootste mislukte projecten van het naoorlogse Groningen: het nimmer gegraven Dollardkanaal. Plannen hiervoor ontstonden in de jaren zestig, toen druk werd gestudeerd op verbetering van de afwatering van Westerwolde.

Onderdeel van de werken was de bouw van een nieuw sluizencomplex bij de Punt van Reide, dat het water van het nieuwe kanaal moest afvoeren. In 1974 werd het kunstwerk - kosten 30 miljoen gulden - opgeleverd. Geen druppel water heeft de sluis echter door haar deuren zien vloeien. De aanleg van het Dollardkanaal was inmiddels gestuit op een onneembare hindernis: de kluut. Milieu- en natuurbeschermers, verenigd in de Werkgroep Dollard, hadden deze sympathieke vogel verheven tot boegbeeld in hun strijd tegen aantasting van het kwelderlandschap.

Voorbeeld: Verbreding N33





**Corrective** Maintenance



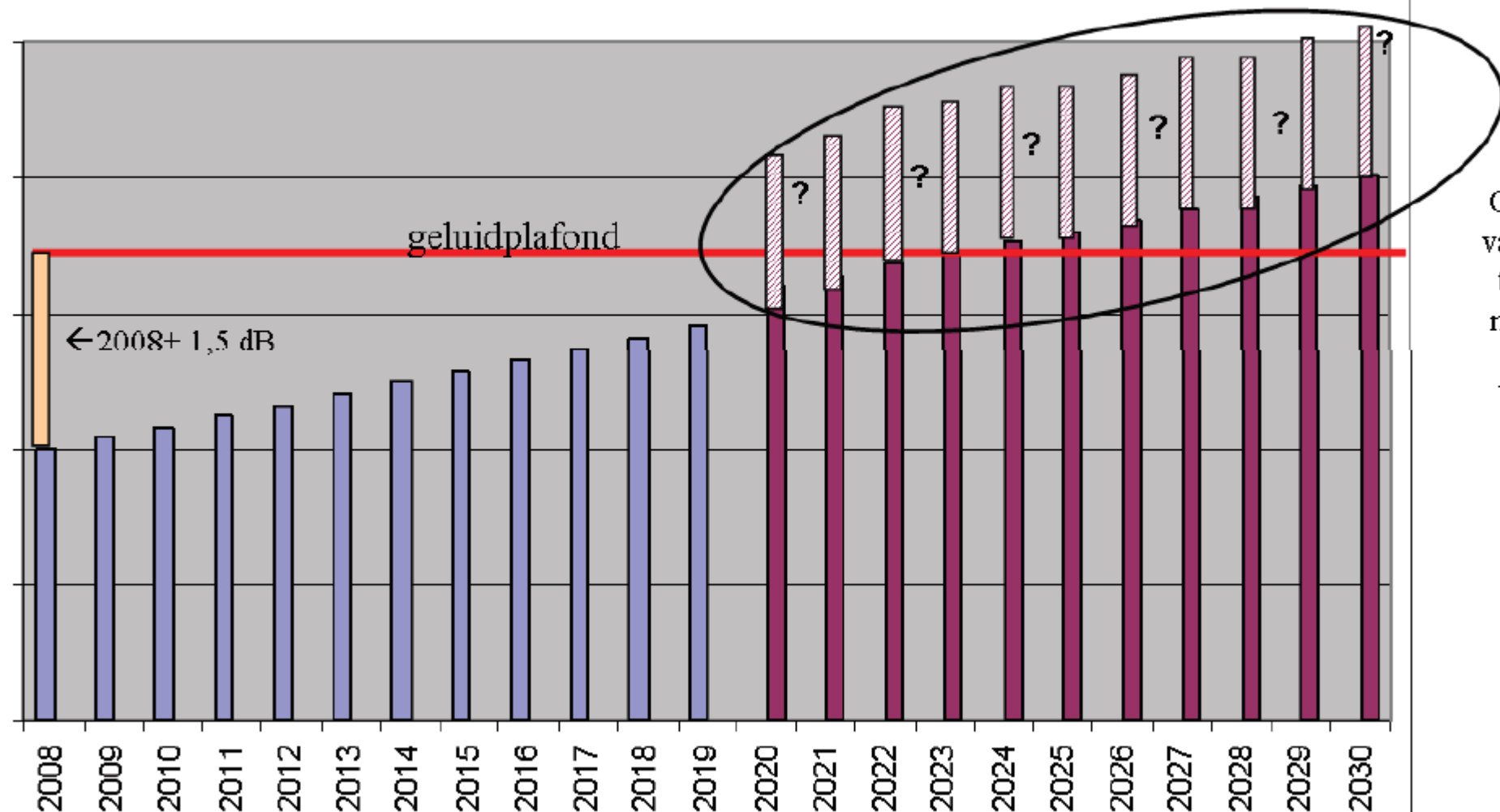
**Preventative** Maintenance



**Predictive** Maintenance



## Voordorp (met Ring)



Overschrijding  
van plafond niet  
toegestaan →  
maatregelen in  
project Ring  
Utrecht nodig



# Verkeerde aannames bij berekening geluid Lelystad

Weer & Verkeer | 16 oktober 2017 om 21:32

Delen:  

Voor de geluidsberekening in het milieueffectrapport (MER) voor luchthaven Lelystad is gebruikgemaakt van enkele verkeerde aannames. Maar demissionair staatssecretaris Sharon Dijksma (Infrastructuur en Milieu) verwacht niet dat het gevolgen zal hebben voor de gekozen vliegroutes en voor de opening van de luchthaven in 2019.





# Asbestdaken. Hoe herken je ze in luchtfoto's?

Een zoektocht naar de beste methode

November 22, 2016 - Luchtfoto, Asbestdaken, Machine Learning

Readaar detecteert asbestdaken uit luchtfoto's. Maar hoe herken je asbest? En waarom willen we dat?

We zijn een startup gespecialiseerd in datamining uit luchtfoto's. Met die achtergrond krijg je de gekste vragen: "zou jullie trampolines in achtertuinen in kaart kunnen brengen", of "witte muren voor graffiti artiesten". Er was één vraag die ons bleef achtervolgen: **"zouden jullie ook geautomatiseerd asbestdaken kunnen herkennen?"**

Asbestdaken zijn vanaf 2024 [verboden](#) door de overheid. Het bleek daarom al snel dat er een hoop reuring is op dit gebied. Van asbestsaneerders die gouden tijden zien aankomen, adviseurs die asbest herkennen aan het printje op het linoleum tot gemeenten die nu al aan de slag moeten om dit verbod in 2024 te kunnen handhaven. Gemeenten die vaak niet weten hoe groot het probleem eigenlijk is, of waar die daken precies liggen.



# Fraude bij registratie van kalveren

Landbouw

Bij 7.700 bedrijven is er mogelijk gefraudeerd, schrijft minister Schouten aan de Kamer.

✎ Esther Rosenberg, Joep Dohmen & Floor Bouma · 23 januari 2018



Foto: Jerry Lampen/ANP

Duizenden melk- en rundveebedrijven frauderen mogelijk met de registratie van kalveren. Door hun administratie te vervalsen, houden boeren meer koeien dan wettelijk is toegestaan. Minister Schouten (Landbouw, CU) meldt in een brief aan de Tweede Kamer hard tegen de fraude te zullen optreden. Dinsdagochtend had ze hierover overleg met de sector.

De fraude kwam aan het licht bij een controle door de Rijksdienst voor Ondernemend Nederland (RVO) en de Nederlandse Voedsel- en Warenautoriteit (NVWA). De diensten onderzochten de registratie van pasgeboren kalfjes. Bij 7.700 van de in totaal 23.200 veebedrijven wezen de gegevens op mogelijke fraude. Daarop inspecteerde de NVWA gisteren 93 bedrijven. Bij 45 van die inspecties bleek de situatie in de stal af te wijken van de registratie. Het vermoeden van fraude is in die gevallen bevestigd, aldus Schouten.

# Translink deelde reisgegevens studenten met DUO voor fraudebestrijding

Translink, het bedrijf achter de ov-chipkaart, gaf reisgegevens door aan DUO. De dienst vraagt om deze gegevens om studenten op te sporen die frauderen met een uitwonende beurs. Dat gebeurde zonder tussenkomst van een officier van justitie, wat wel zou moeten.

Tweakers.net, 22 augustus 2017

Translink zegt dat het inderdaad aan 'vijf à tien verzoeken per week' gehoor gaf en gegevens deelde met DUO. De Volkskrant [schrijft](#) dat dit doorging totdat een van de studenten een rechtszaak begon, omdat de privacyregels overtreden zouden zijn. De rechter gaf de student gelijk en legde hem geen boete op. DUO wil tegen het vonnis in beroep gaan. De organisatie vindt dat er voor het opvragen een juridische basis aanwezig is, namelijk de Algemene Wet Bestuursrecht.

≡ Menu

nrc.nl

abonneer

## DUO mag ov-gegevens studenten opvragen

Het doorgeven van ov-gegevens botst weliswaar met privacy, maar dat mag om fraude op te sporen.

✎ Liza van Lonkhuyzen © 5 februari 2018





# Politie gaat misdaad voorspellen met nieuw systeem

MA 15 MEI, 13:36 BINNENLAND



Fietsen bij een studentencomplex in Utrecht ANP

De Nationale Politie gaat aan de slag met een systeem dat op basis van nauwkeurige criminaliteitsgegevens misdaden voorspelt. Met dit Criminaliteits Anticipatie Systeem (CAS) is al geëxperimenteerd in onder andere Amsterdam en Den Haag.



## Nu én in de toekomst schoolverlaten voorkomen

**Rutger Kerkhoff**

Data Scientist

[Toon alle berichten van Rutger Kerkhoff →](#)

Jaarlijks komen ongeveer 25.000 jongeren zonder afgeronde opleiding of startkwalificatie op straat te staan. Dit heeft vaak grote gevolgen voor de rest van hun leven: van langdurige werkloosheid tot sociale problematiek en criminaliteit. [Het ministerie van Onderwijs, Cultuur en Wetenschap stelt voor 2021 het doel dit aantal tot maximaal 20.000 voortijdig schoolverlaters \(vsv'ers\) terug te brengen.](#) Dit betekent dat RMC regio's niet achterover kunnen leunen en moeten blijven zoeken naar mogelijkheden om het aantal VSV'ers met nog eens 20% te verminderen.

Om deze nieuwe doelstelling te halen is een effectieve aanpak nodig. [In samenwerking met RMC-regio Zuid-Holland Zuid en Amsterdam heeft Ynformed een model ontwikkeld dat de kans dat iemand VSV'er wordt vroegtijdig herkent.](#) Dit model hebben we gerealiseerd met zeer generieke schoolgerelateerde DUO-data. Hierdoor bevat het model geen persoonlijke kenmerken, maar bijvoorbeeld informatie over het aantal opleidingen, de duur hiervan en aan welke school. De generieke aanpak maakt het mogelijk om te schalen naar andere RMC regio's.



# Case: vroegtijdig signaleren van kindermishandeling

POSTED ON APRIL 25, 2017 BY WOUTER VAN DER BIJ

25  
apr

In 2015 hebben we in opdracht van de GGD Amsterdam onderzoek gedaan naar tekstanalyse op de open tekstvelden in het Digitaal Dossier Jeugdgezondheid kinderen. Centrale vraag was hoe ongestructureerde data uit open teksten beter kunnen worden gemonteerd en verbeteren van de kindergezondheid. Kan de data kindermishandeling vroegtijdig te signaleren?

Naast alle informatie die gestructureerd wordt uitgevraagd tijdens consultaties is er ook de mogelijkheid om in eigen woorden bevindingen te registreren. Ze maken daar veel gebruik van, vooral op het moment dat er vermoedens bestaan van mogelijke kindermishandeling, maar deze niet 'hard' kunnen worden gemaakt. De grote hoeveelheden ongestructureerde tekstdata uit de dossiers van kinderen hebben we geanalyseerd met behulp van textmining. Met succes.

Als een computerprogramma naar zoiets gevoelig zoekt als wie z'n kind mishandelt, zal de behoefte groot zijn om te begrijpen hoe die berekening dan werkt. Maar daar inzicht in krijgen is grotendeels een illusie, legt Tim Paauw uit. Hij was bij het bedrijf Ynformed een van de ontwikkelaars van de software. Paauw: „De computer gebruikt zo'n 1.000 variabelen, en komt via een soort beslisboom van honderden vertakkingen tot een uitkomst. Het menselijk brein kan zo'n model nauwelijks doorgronden.” Om te beginnen

NRC, 23 maart 2018



20-05-2015 | door: **Wouter Hoeffnagel**



## Belastingdienst zet in op big data met 1.500 data-analisten

**De Belastingdienst voert een grote reorganisatie door en wil zich meer gaan richten op data-analyses. Circa 5.000 bestaande functies verdwijnen bij de dienst, terwijl 1.500 nieuwe data-analisten worden aangetrokken.**

Bronnen melden tegenover de NOS dat ICT-ontwikkelingen ervoor zorgen dat de nadruk bij de Belastingdienst steeds meer komt te liggen op data-analyse. Hierdoor zouden minder medewerkers nodig zijn. Daarnaast is er ook behoefte aan data-analisten. Dit zou dan ook de reden zijn dat de reorganisatie wordt doorgevoerd. Werknemers van de Belastingdienst die hun baan verliezen worden binnen of buiten de overheid aan een andere baan geholpen.

Door meer in te zetten op data-analyses moet belastingontduiking en fraude beter aangepakt kunnen worden, terwijl belastingaanslagen nauwkeuriger worden. Big data moet hierdoor de overheid jaarlijks honderden miljoenen tot zelfs miljarden euro's opleveren.



PRIVACY

## Waarom de Belastingdienst jouw Spotify-playlist wil hebben

BK

BONNY KRAMER  
Apr 3 2017, 1:24pm

Zwenne zou er een praatje houden over privacy, maar was net wat te vroeg en schoof aan bij een presentatie van een data-analist. Die vertelde vol enthousiasme hoe de belastingmoraal bij burgers kon worden gemeten. Zwenne: "Volgens deze data-analist kan dat door bijvoorbeeld naar iemands Spotify-playlist te kijken. Hij zou het volstrekt begrijpelijk vinden als de Belastingdienst voortaan afspeellijsten van muziek zou opvragen." Aan de hand van de muziekvoorkeuren kan worden ingeschat in hoeverre iemand geneigd is zijn aangiftebiljet zorgvuldig in te vullen of niet. "Dit was geen grap, die man was bloedserieus. Mensen die naar Bach luisteren, zijn brave burgers."

# Fiscus vist vergeefs naar museumkaart

privacy

Willem Schoonen • Uit de krant van gisteren

Museumkaart weigert de gegevens van een kaartgebruiker te laten zien aan de Belastingdienst. Die wil achterhalen of die kaarthouder in een bepaalde periode in Nederland verbleef en dus belastingplichtig was. Vandaag dient een kort geding van de Belastingdienst tegen de Museumvereniging, die de kaart uitgeeft, om inzage alsnog af te dwingen.

De Museumvereniging zegt grote en principiële bezwaren te hebben tegen het delen van bezoekgegevens. Zij zou daartoe alleen in uitzonderlijke gevallen bereid zijn, bijvoorbeeld wanneer de nationale veiligheid in het geding is. "Dat is bij dit verzoek van de Belastingdienst niet het geval."

Bovendien geeft inzage in die gegevens helemaal geen zekerheid over de bewegingen van de houder, omdat in de onderzochte periode nog geen pasfoto op de kaart stond. Die kon dus ook door een ander worden gebruikt.







**Platooning rukt op in de logistieke wereld. Maar wat is platooning eigenlijk. De techniek nader uitgelegd.**

Bij **truck platooning** zijn vrachtauto's elektronisch aan elkaar gekoppeld waarbij de voorste truck de snelheid en route bepaalt. De andere trucks volgen automatisch zonder dat daar de hulp van de chauffeur voor nodig is.

### Voordelen platooning

Truck platooning moet zorgen voor meer ruimte op de weg dus een betere doorstroming van het verkeer. Door vrachtauto's dichter op elkaar te laten rijden zijn ook brandstof en CO<sub>2</sub> besparingen tot 10 procent mogelijk. Dat is niet alleen goed voor het terugdringen van de uitstoot van CO<sub>2</sub>, maar het kan ook interessante besparingen opleveren voor het bedrijfsleven. Truck platooning zorgt daarnaast ook voor een veiliger wegtransport, is de overtuiging van de initiatiefnemers.

*Open moet het zijn!*

# Universiteiten kampen met leegstroom ai-onderzoekers

Bestuurders spreken hun zorgen uit

28 augustus 2018 11:48 | [Zamire Willems](#)

Topic Onderwijs



**Steeds meer toponderzoekers bij universiteiten worden weggekaapt door buitenlandse techbedrijven. De slag om talent in het vakgebied van kunstmatige intelligentie (artificial intelligence, ai) begint ten koste te gaan van de kennisontwikkeling in Nederland en de rest van Europa met slechter onderzoek en onderwijs als gevolg. Hiervoor waarschuwen vier Nederlandse universiteiten bij een rondgang van NRC.**

De omvang van de zogenoemde 'braindrain', de leegloop van getalenteerde studenten en academisch personeel bij universiteiten, is moeilijk te achterhalen. Er zou niet centraal worden bijgehouden waar en bij wie vertrekkende talenten hun

carrière voortzetten. Ook omdat ai een verzamelterm is van onder meer machine learning, deep learning, computer vision en neurale netwerken, waar legio bedrijven en functies onder vallen.

Volgens de universiteiten zijn het voornamelijk bedrijven als Google, Amazon, Facebook, Microsoft, IBM, Nvidia, [Bosch](#) en Volkswagen, naast ook buitenlandse topuniversiteiten, die Nederlandse ai-talenten binnenhalen door het bieden van jaarsalarissen van meer dan 100.000 euro en inhoudelijk sterke onderzoeksplekken; grote techbedrijven beschikken nu eenmaal over zeeën van data dan universiteiten. Iets waar onderzoekers nog breder, dieper en intensiever mee aan de slag kunnen, aldus het artikel in NRC.

# Meer weten?

- [data.overheid.nl](http://data.overheid.nl)
- [pdok.nl](http://pdok.nl)
- Mustread: ‘Weapons of Math Destruction’ van Cathy O’Neil
- Doe mee met een hackathon,  
bijvoorbeeld via <http://openstate.eu> of Kennisnetwerk Data Science Noord-Nederland
- [databeheer@provinciegroningen.nl](mailto:databeheer@provinciegroningen.nl)
- [github.com/friesewoudloper](https://github.com/friesewoudloper)
- [twitter.com/friesewoudloper](https://twitter.com/friesewoudloper)