

CITAAT

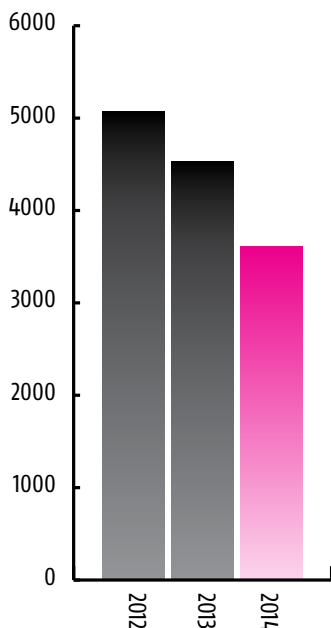
“ Amice, ik kan mijn indianenpak niet vinden. ”

Bij het uitwerken van bandjes gaat er weleens iets mis. 'Amice, ik kan me in die aanpak niet vinden,' had de advocaat in kwestie eigenlijk ingesproken (NRC Handelsblad, 14 oktober 2014).

CIJFERS

Daling straatroof

Aantal straatroven van januari tot en met augustus in 2012, 2013 en 2014.



Bron: Rijksoverheid.nl

COLUMN



Harry Veenendaal

Rutte II: de folteraar van de rechtsstaat

In de luwte van de Zwarte Piet-discussie verwordt Nederland rechtsstatelijk tot het Mordor van middelmatigheid, met Rutte II als folteraar. Grote dramatische woorden? Inderdaad.

Vicepremier Lodewijk Asscher (PvdA) wil de belangrijkste bestanden met uw privégegevens schakelen ten behoeve van 'fraudepreventie'. Denk daarbij aan arbeidsgegevens (hoe vaak bent u ziek), boetes en sancties, fiscale gegevens, gegevens van roerende en onroerende goederen, handel- en huisvestingsgegevens, identificerende gegevens, inburgerings-, nalevings- en onderwijsgegevens, pensioengegevens, re-integratiegegevens, schuldenlast-, uitkerings-, toeslagen- en subsidiegegevens, vergunningen en ontheffingen, bank- en zorgverzekeringsgegevens.

Het softwaresysteem Risico Identificatie (SyRI) gaat deze bestanden koppelen. SyRI wordt onder andere door de Nederlandse inlichtingendiensten gebruikt voor het opsporen van bijvoorbeeld leden van Al Qaida. Door de bestanden te schakelen heeft de overheid de mogelijkheid om via datamining risicoprofielen op te stellen van wie 'mogelijk' gaat frauderen. Dit houdt in dat *a priori* risicogroepen worden gedefinieerd die extra in de gaten worden gehouden zonder dat er enige aanleiding is van het plegen van een strafbaar feit. Ik hoef hier niet nader in te gaan op het juridisch vulgaire karakter van een dergelijke redenering. Dit interesseert Rutte II weinig en daarmee hebben we te maken met een *fait accompli*. Maar er bestaat een evenzo groot pro-

bleem met de uitvoering van datamining door de overheid.

Een voorbeeld: neem een database met de gegevens van alle ooievaars in een bepaalde periode in een bepaald gebied. Voeg aan deze database gegevens toe over het aantal geboren baby's in diezelfde periode in hetzelfde gebied. Op basis van datamining zal op enig moment een verband tot stand komen tussen de toe- of afname van het aantal ooievaars en de toe- of afname van het aantal baby's. Iedereen snapt dat er geen enkele causaliteit bestaat tussen gelijke toe- of afname van ooievaars en baby's terwijl de cijfers iets anders suggereren. Vervang nu ooievaars en baby's eens door handelshuis Van der Moolen en de dataminingfaciliteiten van de AEX/AFM. In het geval van het handelshuis is de handel een paar jaar geleden stilgelegd naar aanleiding van bepaalde indicatoren die het datamining-systeem aangaf. Hierdoor zou sprake zijn van handel met voorkennis. Achteraf blijkt er geen enkele causaliteit te zijn met alle gevolgen van dien.

Het invoeren van een systeem zoals SyRI is om veel redenen al problematisch, een groter probleem zijn de ambtenaren die de data interpreteren. De enigen hiertoe in staat zijn statistici, econometristen of wiskundigen. Die werken meestal bij Shell, McKinsey of andere bedrijven. Afgezien daarvan zal geen enkele statisticus zich professioneel wagen aan het gemodder van SyRI. Wat blijft over? Welwillende alfa's die wiskunde voor het laatst op de middelbare school hebben gehad. Met alle gevolgen van dien.