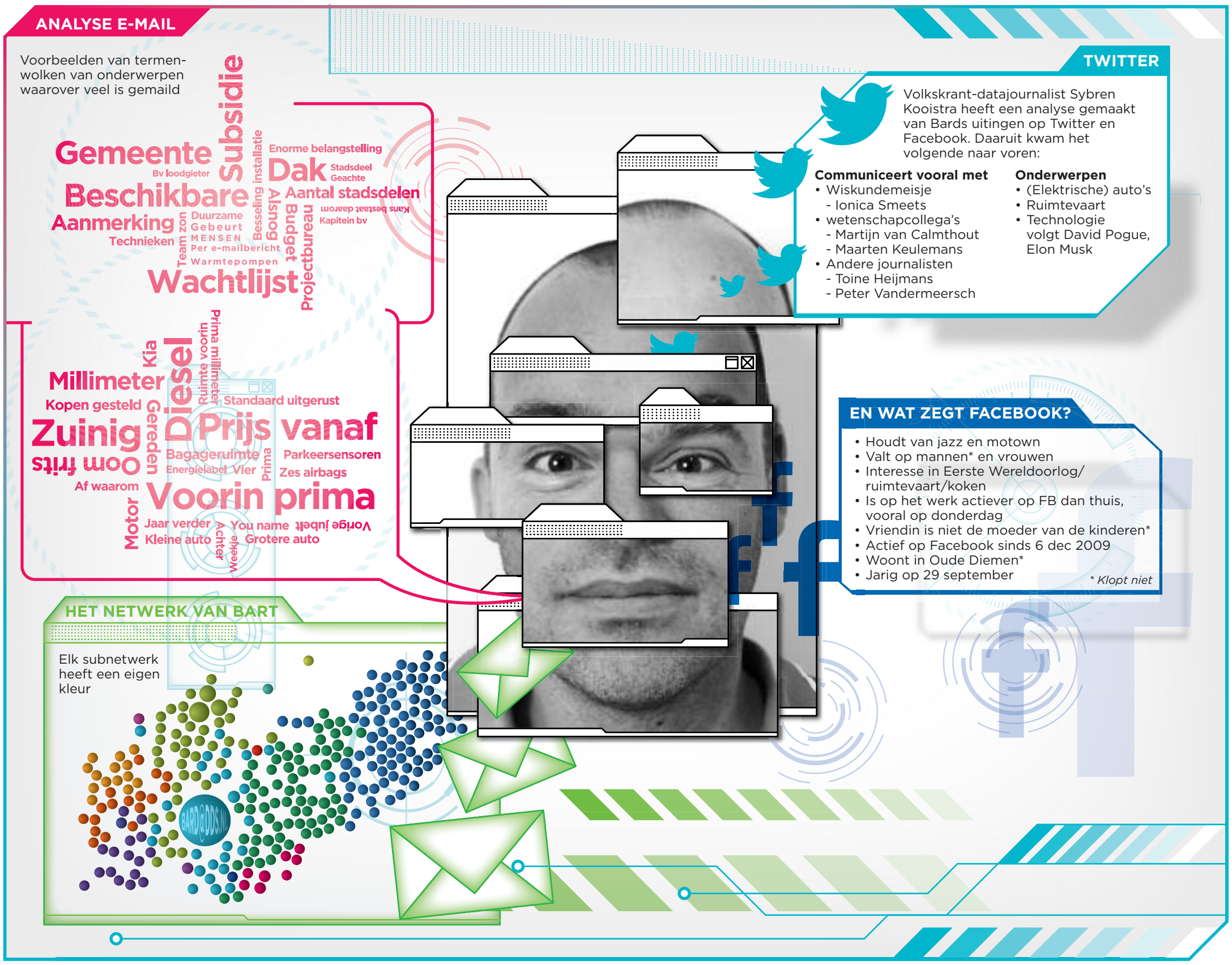
De zaak-Snowden als game

Nederlandse gamemakers hebben een – enigszins melige - online-game ontwikkeld over de zaak-Snowden. Daarbij kruipt de gamer in de huid van klokkenluider Edward Snowden, die usb-sticks laadt met gevoelige informatie, om deze uit het raam van het NSA-gebouw te werpen zonder te worden betrapt door beveiligers en camera's. bit.ly/1aPUsxn



wolk kun je heel snel belangrijke conversatie-onderwerpen ontdekken', zegt onderzoekster Verberne. Woorden die daarin veel voorkomen kun je gebruiken als 'ingang' voor verder onderzoek.

Om te voorkomen dat veelgebruikte woorden als 'het' en 'groetjes' telkens het beeld domineren, wordt alle mail eerst vergeleken met een referentie-e-mailbestand. Woorden die in beide datasets veel voorkomen, worden uitgesloten. Wat overblijft zijn als het goed is woordwolken met zeggingskracht. 'Ik zag bijvoorbeeld dat een afzender Lieve Bart in zijn aanhef zette', zegt Verberne. 'Dat is merkwaardig, omdat 'lieve' een intieme band suggereert terwijl de afzender je voornaam toch verkeerd spelt.'

Termenwolken zijn de vuurtorens van de datazee. Ze wijzen op de onderwerpen waarover gesproken wordt, maar kunnen bijvoorbeeld ook helpen bij het ontdekken van gecodeerd taalgebruik. In de drugsscene was het in de jaren negentig gebruikelijk om bij de dealer 'versnellingsbakken' te bestellen als speed bedoeld werd. Dat zou de eventueel meeluisterende politie om de tuin moeten leiden. Bij de analyse van het e-mailverkeer zou deze omschrijving er in de termenwolk zonder meer uitknallen en zou al snel duidelijk worden wat ermee bedoeld wordt.

De algoritmen kunnen ook zien of de toon van de berichten positief of negatief is, door te kijken of woorden die als negatief worden beschouwd vaak voor komen. Zo ontdekte Sappelli dat ik kennelijk 'een issue heb gehad' met de buren over poezenoverlast. Inderdaad: we hadden tot voor kort een kat, net als enkele andere buren, en een van deze katten poepte soms op het gazon van een van de buren, waar kleine kinderen spelen. Dat leidde enkele jaren terug tot wat geagiteerd mailverkeer. Ik was het alweer vergeten, mijn mailbox niet.

De grote vraag is natuurlijk wat de onderzoekers met al hun graafwerk van me te weten zijn gekomen. Hebben ze een beeld van wie ik ben? Hoogleraar Kraaij gaat er even goed voor zitten. Ze hebben achterhaald dat ik op IJburg woon (inclusief adres), dat mijn vriendin psycholoog is en in de verslavingszorg werkt (ook bij welke instelling). Ze weten zelfs op welke dagen ze werkt. Ze weten dat ik een zoon Teun heb, op welke school hij zit en waar hij zwemles heeft. Ze weten dat ik een kat heb en bovengemiddeld geïnteresseerd ben in Mercedes-Benz (tot ongeveer een jaar geleden reden we in een klassieke Mercedes en ik was lange tijd actief in een club).

Dat is een behoorlijk compleet gezinsplaatje, al is een belangrijk aspect gemist: er is nog een tweede zoon. Die is wellicht over het hoofd gezien door zijn bijzondere naam: Dieks. Ooit een populaire voornaam, inmiddels hevig in onbruik geraakt; er zijn nog maar 19 Dieksen in Nederland. 'We hebben hem voorbij zien komen, maar hem niet als zoon opgemerkt', zegt Verberne. Waarmee blijkt dat de analyse van de data mensenwerk blijft, waarbij vergissingen mogelijk zijn.

'Verder niks spannends?', vraag ik. Geen minnaressen? Geen extreme politieke voorkeuren? Geen zaken waar een overheid op zou aanslaan? 'Nee, niets van dat alles', zegt Verberne. 'Als we je uitgaande mail hadden kunnen bekijken, hadden we misschien meer gevonden. Wat je zelf verstuurt, zegt meer over je dan wat je ontvangt.' Maar die data is niet beschikbaar. De conclusie is helder. 'Wat er niet in zit, komt er ook niet uit', zegt Kraaij. 'Dit is kennelijk wie je bent.'

Ik voel een lichte teleurstelling dat in mijn leven de VvE kennelijk zo'n grote rol speelt. Maar er is ook opluchting, omdat het nu officieel vaststaat: ik ben braaf geweest.

MAG DE OVERHEID (OF JE BAAS) ALLES GEBRUIKEN WAT OPENBAAR IS?

De overheid mag niet zomaar doen wat de onderzoekers van de Radboud Universiteit en TNO hebben gedaan: iemands mailbox lezen. 'Ook de politie heeft natuurlijk iemands wachtwoord nodig om in de mail te kunnen komen, en toestemming van de rechter-commissaris voor een e-mailtap', zegt Mark van Staalduinen, die voor TNO de politie adviseert over digital profiling, het maken van profielen op basis van sporen die iemand digitaal nalaat.

Door de onthullingen over Prism, het snuffelnetwerk dat de Amerikaanse veiligheidsdienst NSA gebruikt om de gangen van burgers na te gaan op onder meer sociale netwerken, is ook in Nederland de discussie opgelaaid over wat de recherche en veiligheidsdiensten mogen inzien. Een van de vragen is of gebruik mag worden gemaakt van openbare bronnen als Twitter en de publiek toegankelijke data op Facebook. 'Je kunt argumenteren dat mensen die informatie daar nooit hebben achtergelaten om de politie eventueel van dienst te zijn', zegt Van Staalduinen. Aan de andere kant verlangt de samenleving dat de overheid ingrijpt als iemand op een openbaar prikbord verkondigt dat hij een slachting wil aanrichten op een middelbare school, stelt hij. 'In dat spanningsveld moeten we opereren. En waar de grens precies ligt, zal gaandeweg duidelijker worden.'

De techniek die hier gebruikt is om de mailbox van de verslaggever te analyseren, is niet alleen nuttig bij opsporingsdoeleinden, ze kan ook helpen het welzijn van bijvoorbeeld werknemers te verbeteren. Hoogleraar Wessel Kraaij werkt aan de Radboud Universiteit aan het project Swell, waarbij hij vergelijkbare methoden toepast om ervoor te zorgen dat werknemers bijvoorbeeld niet overbelast raken. 'Je kunt bijvoorbeeld iemands mail scannen en dan ontdekken dat hij steeds meer woorden gebruikt die duiden op stress', zegt Kraaij. In zo'n geval zou de software, met wat de hoogleraar 'zachte signalen' noemt, een werknemer daarop kunnen wijzen en hem adviseren de e-mail per project te lezen in plaats van op volgorde van ontvangst.

De software zou ook binnenkomende mail kunnen filteren op relevantie. 'Als je op je werk zit is een bericht van de Vereniging van Eigenaren minder relevant, dat kan zo'n systeem dan even elders parkeren. Terwijl een mailtje van iemand die je net hebt geïnterviewd wel doorkomt', aldus Kraaij. Doel van het project is om op langere termijn gezondere en dus productievere werknemers te krijgen. Ook hier vormt privacy een heet hangijzer. 'Mensen willen niet dat de manager kan meekijken om te zien hoe gestrest iemand is. Men wil de baas blijven over de eigen data.'