

Twitter onder de loep

Het is verleidelijk om social media in te zetten voor onderzoeksdoeleinden, maar hoe doe je dat op een verantwoorde manier?

Lees mee met twee ervaringsdeskundigen.

Marika de Bruijne en Erica Renckens

Mick Couper,
Research Professor aan Michigan University:

‘Minder hype en meer hypothesen’

Mick Couper, Research Professor aan Michigan University en bekend surveymethodoloog, vertelt hoe social media en surveydata elkaar kunnen ondersteunen.

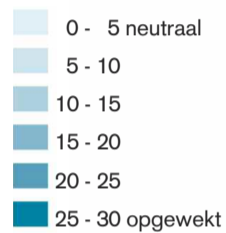
Mick Couper: “Social media en surveys kunnen worden gekoppeld op drie niveau’s: ge-aggregeerd, conceptueel of individueel. De meesten koppelen op ge-aggregeerd niveau en laten verbanden zien tussen surveyresultaten en social media.” Bekende voorbeelden zijn analyses op basis van Twitter-trends. Tijdens een griepseizoen wordt veel over griep getweet. Door te analyseren hoe vaak bepaalde woorden voorkomen, wordt een schatting gemaakt van de verspreiding van griep door het land. “Mijn zorg hierbij is dat surveys vaak worden gebruikt als maatstaf. Hierbij zijn de algoritmen voor het identificeren van trends in social media gekozen doordat ze passen bij de surveydata. Wat nodig zou zijn, is een onafhankelijke evaluatie van trends.”

Hypothesen testen

Couper vervolgt: “Met het koppelen op conceptueel niveau bedoel ik het testen van hypothesen in beide databronnen (surveys en social media). Bijvoorbeeld: is er een relatie tussen houding en gedrag hetzelfde wanneer deze wordt gemeten in een survey als wanneer deze wordt gemeten in social media? Ook kan er niet alleen naar trends van losse indicators in social media worden gekeken, maar naar associaties zoals: sturen mensen die over een bepaald onderwerp tweeten ook tweets over een ander onderwerp?”

Beperkte groep Twitteraars

Couper: “Surveyresponses zijn ook op individueel niveau te koppelen aan persoonlijke berichten op social media. Maar dat is niet makkelijk. Op basis van surveydata weten we wie Twitter-gebruikers zijn, maar niet wat ze op Twitter posten. En binnen een groep respondenten maakt vaak slechts een beperkte groep gebruik van Twitter.” Zo blijkt bijvoorbeeld uit het LISS panel van CentERdata dat de Twitter-populatie verschilt van de gehele populatie en jonger is dan gemiddeld. Verder stuurt slechts een deel van de mensen op Twitter zelf actief tweets en heeft men te maken met automatische tweets, Twitterbots, retweets en bedrijfstweets.



Trends oppikken Voor de Volkskrant bracht Tjong Kim Sang met behulp van Twiqs het humeur van de Nederlandse twitteraar per gemeente en per provincie in kaart. Hiervoor maakte hij gebruik van een eigenhandig samengestelde lijst met positieve en negatieve woorden. Deze afbeelding laat het gemiddelde sentiment van tweets per gemeente zien voor november 2013. Tjong Kim Sang liep ook hier tegen een beperking aan van het gebruik van tweets als databron: Limburg en Vlaanderen kwamen opvallend somber uit de bus - waarschijnlijk door het ontbreken van typische dialectwoorden in de lijst. bron de Volkskrant

Manieren zoeken om onderzoeksmethoden te combineren of om ze samen te gebruiken en deze nog verder te ontwikkelen, is de boodschap. “Er worden al veel studies gepubliceerd, maar bij veel daarvan heb ik mijn twijfels. Social media bieden een kans voor onderzoekers en naarmate social media analytics zich ontwikkelen, zullen we meer studies zien van goede kwaliteit.” Of, zoals Couper concludeert: “We hebben minder hype en meer hypothesen nodig.”

Het artikel *Is the Sky Falling? New Technology, Changing Media, and the Future of Surveys* uit Journal of the European Survey Research Association (volume 7, nummer 3, 2013) omschrijft meer bevindingen van Mick Couper. (MdB)

<https://ojs.ub.uni-konstanz.de/srm/article/view/5751>

10 mln voor nieuw MIT-lab

Twitter stimuleert ook zelf actief het gebruik van tweets voor wetenschappelijk onderzoek. Op 1 oktober maakte het bedrijf bekend 10 miljoen dollar te investeren in een nieuw lab aan MIT, het Laboratory of Social Machines. Onderzoekers in het lab krijgen volledige toegang tot de real-time stroom aan tweets en tot het gehele archief vanaf de begindagen van Twitter.

Op basis daarvan zullen zij nieuwe technologieën ontwikkelen die wijs kunnen worden uit de semantische en sociale patronen in de online mediaberichten op Twitter en andere media.

Erik Tjong Kim Sang,
onderzoeker bij het Meertens Instituut:

‘Interessant om trends op te pikken’

Erik Tjong Kim Sang ontwikkelde Twiqs, een zoekmachine waarmee onderzoekers, studenten en andere geïnteresseerden Nederlandse tweets kunnen analyseren.

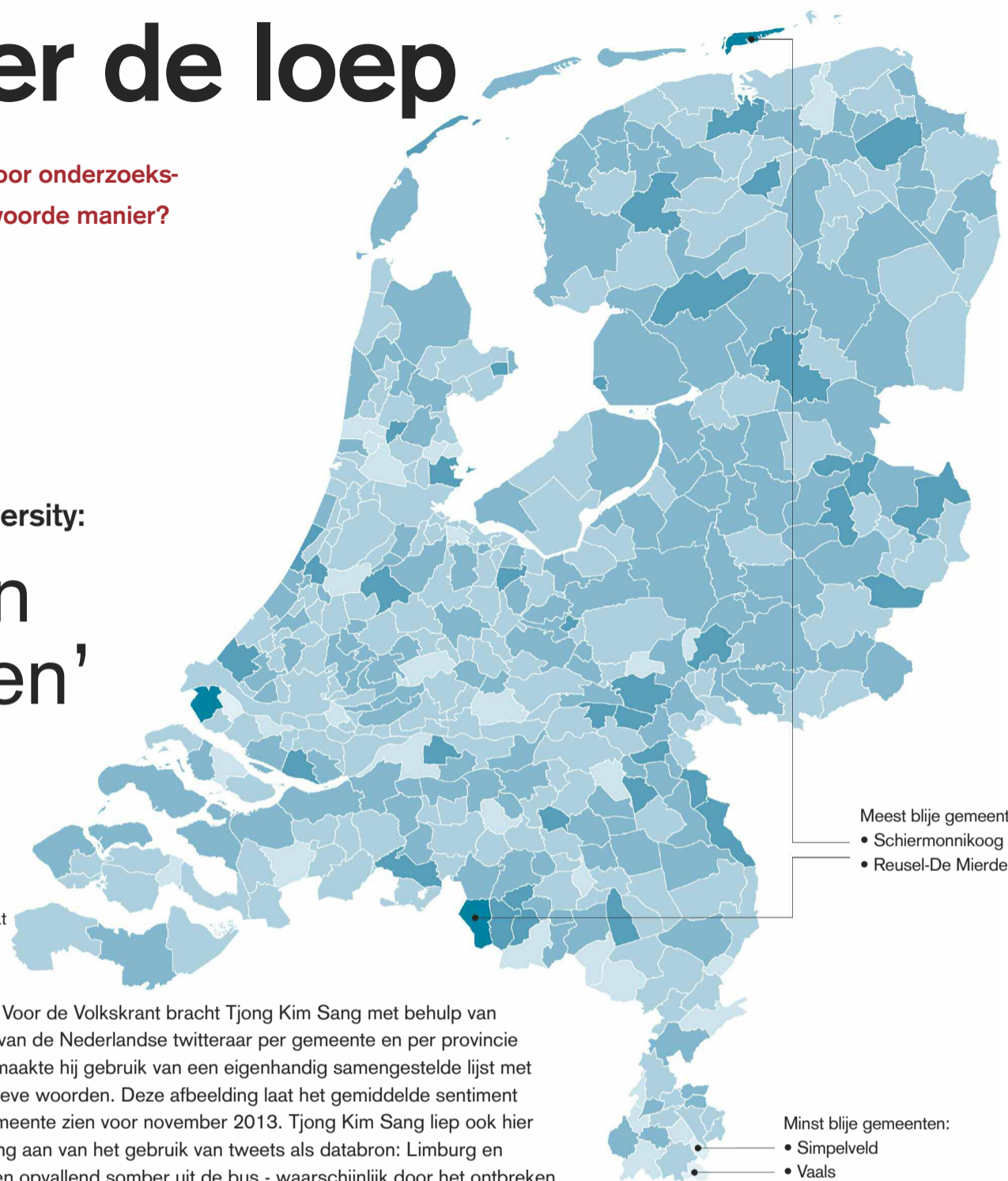
Twiqs is het resultaat van het TwiNL, een project van het Netherlands eScience Center. “Het hoofddoel van dit project was om Nederlandse tweets te verzamelen en deze op een gemakkelijk doorzoekbare manier aan te bieden aan andere onderzoekers,” aldus Tjong Kim Sang, als onderzoeker verbonden aan het Meertens Instituut. Tweets bevatten naast de feitelijke inhoud ook metadata die waardevol kunnen zijn voor onderzoek, zoals dag en tijdstip van versturen, informatie over de afzender en hashtags, die een aanwijzing kunnen geven over het onderwerp.

Twitter-data voor onderzoek

“Zelf hebben we ook gezocht naar informatie die uit deze tweets te halen was, bijvoorbeeld een voorspelling van een verkiezingsuitslag. De resultaten daarvan waren acceptabel, maar tweets zijn geen goede afspiegeling van de

Nederlandse bevolking.” In 2013 hadden ongeveer 3,3 miljoen Nederlanders een Twitter-account, van wie minder dan de helft dagelijks tuitert. De gebruikers zijn overwegend jong, hoogopgeleid en van het mannelijk geslacht. Deze niet-representatieve steekproef uit de Nederlandse bevolking is volgens Tjong Kim Sang niet het enige probleem bij het gebruik van twitter-data voor onderzoek: “Als je inzoomt op een bepaald onderwerp op Twitter, blijven er meestal weinig bruikbare tweets over.”

De onderzoeker raadt zijn collega’s daarom niet direct aan om twitter te gebruiken als bron van data voor wetenschappelijk onderzoek. “Social media lijken mij niet geschikt om statistisch relevante informatie over de hele bevolking uit af te leiden,” zegt Tjong Kim Sang. “Ze zijn wel interessant om trends op te pikken. Het Instituut voor Nederlandse Lexicologie heeft bijvoorbeeld tweets gebruikt om nieuwe Nederlandse woorden te vinden. Dit levert geen compleet plaatje op – niet alle nieuwe woorden worden gebruikt op Twitter – maar wel diverse interessante vondsten.” (ER) twiqs.nl



Meest blij gemeenten:
• Schiermonnikoog
• Reusel-De Mierden

Minst blij gemeenten:
• Simpelveld
• Vaals