

Utrecht Data School onderzoekt vluchtelingendebat op Twitter

Onderzoek naar tweets met #gelukszoeker

Hoe haal je betekenis uit tweets?

De Utrecht Data School analyseert de vluchtelingencrisis op Twitter door netwerkanalyse te koppelen aan semantische indicatoren. *Paul Verhaar en Mirko Tobias Schäfer*

Schäfer

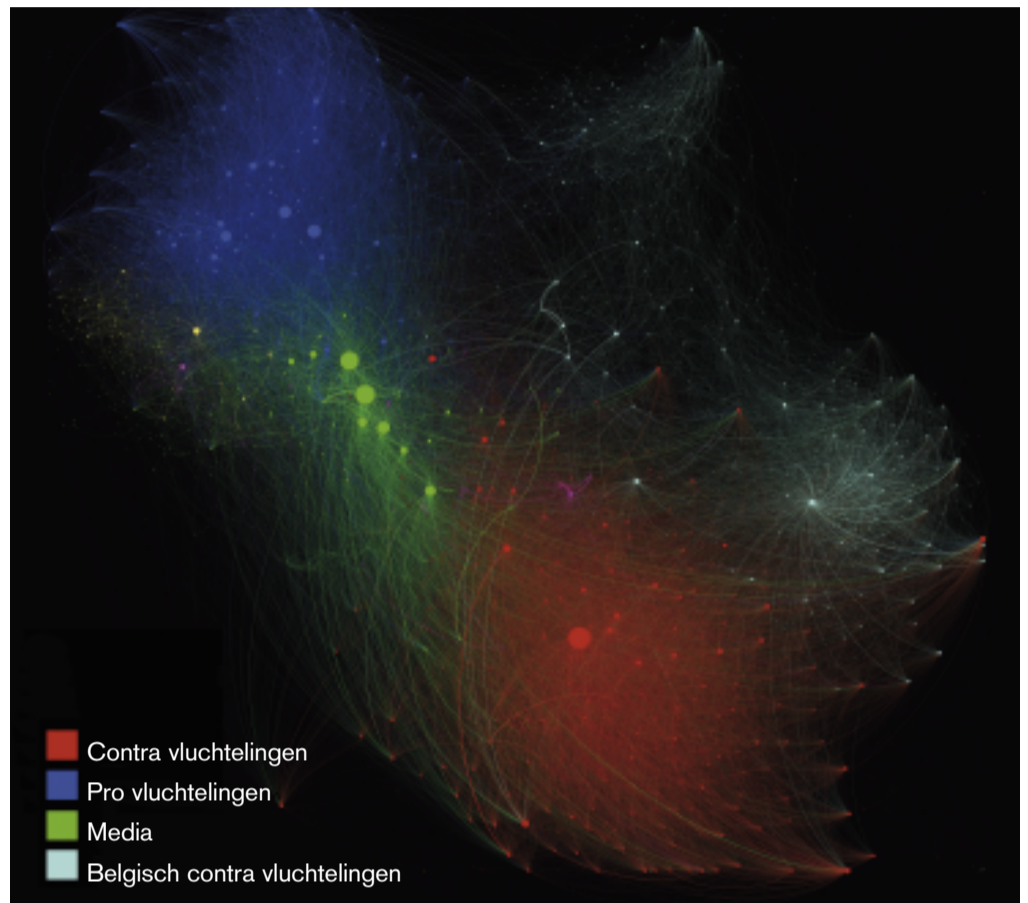
Door het toenemende aantal vluchtelingen nam ook het aantal gerelateerde berichten op social media toe. Talloze berichten werden verspreid over het internet, waarin gebruikers hun mening uitten. Een van de grootste uitdagingen is om betekenis uit deze berichten te halen.

“Ons onderzoek focust zich op het in kaart brengen van het vluchtelingendebat met behulp van netwerkanalyses en linguïstische analyses. Onze totale dataset bestaat uit 561.179 tweets, waarvan 198.100 retweets, met het woord ‘gelukszoeker’. Dit woord brengt een dubbele connotatie met zich mee en is een interessante invalshoek voor onze analyse.”

561.179 tweets

Dit Retweet-netwerk laat (re)tweets van politiek-Nederland zien. Rood staat voor politiek rechts (38% van het totaal aantal retweets), blauw staat voor politiek links (24% van het totaal aantal retweets). De visualisatie laat zien dat het debat duidelijk verdeeld is. De media staan in het midden: hun berichten worden vanuit de twee clusters geretweet. Het Twitter-account van Geert Wilders is het account met de meeste retweets binnen het rechtse cluster. Maar ook PVV-fractievoorzitter René Dercksen is een hoogactieve opinie-leider in dit gebruikerscluster. In het linkse cluster is het debat veel diverser, een grotere groep aan opinie-leiders is zichtbaar.

“De netwerkanalyse biedt veel potentie voor lin-



Onze totale dataset bestaat uit 561.179 tweets, waarvan 198.100 retweets, met #gelukszoeker. De blauwe en rode nodes geven 38% en 24% weer van het totaal aantal retweets. Het debat is verdeeld; het netwerk laat een duidelijke polarisatie zien. Ook de positie van de media in de discussie is duidelijk. Bovendien valt op dat het rechtse cluster relatief klein is. Het politiek linkse kamp laat veel verschillende opinieleiders zien, in het rechtste kamp domineert de leider van de PVV. credits UDS

guïstische analyse. Tot op heden zijn er nog geen onderzoeken geweest die netwerkanalyses en linguïstische analyses samenvoegen. Een netwerk kan dienen als basis voor taalanalyse waarin taalpatronen onderzocht en gekoppeld kunnen worden aan de uitkomsten van het netwerk. *Natural language processing* maakt het mogelijk om dieper naar taalrelaties binnen netwerken te kijken met behulp van computationele technieken. De

kracht ligt in de combinatie: nieuwe mogelijkheden om taal online beter te begrijpen en beter te kunnen *minen*. Door deze en andere onderzoekservaringen met grote corpora, zoals twitterberichten, hopen we op termijn specifieke linguïstieke kenmerken van extremistische posities te kunnen traceren, onafhankelijk van netwerkanalyse.”

dataschool.nl

Wint u de Dataprijs '16?

Ook dit jaar wordt de Nederlandse Dataprijs uitgereikt. Een prijs voor een onderzoeker of onderzoeksgroep die extra bijdraagt aan de wetenschap door onderzoeksdata beschikbaar te maken voor aanvullend of nieuw onderzoek.

Heidi Berkhout

Bent u of kent u zo'n onderzoeker met een goed voorbeeld van gedeelde data? U kunt uzelf of een ander via researchdata.nl nomineren. De winnaars van de voorgaande edities zijn in ieder geval enthousiast. Martine de Bruin, winnaar Dataprijs 2014: “De Dataprijs is een bevestiging dat we op het goede spoor zitten” en Marijke Dekker, winnaar Dataprijs 2014: “Door het winnen van de Dataprijs kunnen we nu ook een paar grotere, al langer gewenste verbeterlagen maken.” De prijsuitreiking is op 9 november bij NWO in Den Haag. Ook aanmelden kan online.

researchdata.nl/diensten/dataprijs

Kijk voor meer informatie over de prijs, zoals de drie categorieën, de prijzenpot en de criteria op de website van RDNL. credits RDNL

248 levensverhalen van vluchtelingen

De Stichting Bevordering Maatschappelijke Participatie (BMP) legt verhalen met vluchtelingen van toen en nu vast in het Oral History project ‘Ongekend Bijzonder’. Deze persoonlijke getuigenissen verdienen het om voor de toekomst bewaard te worden. De focus van de interviews ligt op het opbouwen van een nieuw leven in de vier grote steden. Enkele tientallen interviews zijn beschikbaar via *Interview Ongekend Bijzonder foto St. BMP EASY*, het online archiveringssysteem van DANS. Een overzicht hiervan met links naar enkele interviews staat in de thematische collectie: Ongekend Bijzonder - Oral History project. (HB)

ongekendbijzonder.nl
DOI: 10.17026/dans-zbd-b88x

E-DATA & RESEARCH

Jaargang 10 | nummer 3

Nieuwsbrief over data en onderzoek in de alfa- en gamma-wetenschappen.

E-data & Research verschijnt drie keer per jaar en wordt mogelijk gemaakt door: CentERdata, CLARIAH, DANS, Huygens ING, de Koninklijke Bibliotheek en het RIVM.

10 JAAR E-DATA

INHOUD

2 Verslagen van events in Gehoord en bijgewoond

3 Metadata van Clusius' brieven in eLaborate

3 RIS Radboud helpt bij duurzaam archiveren

4 & 5 Mijlpaal! Alweer 10 jaar E-data & Research



4 & 5 'Zorgen om de toekomst van data en software'

6 Alliantie strijd voor Text & Data Mining

7 Deze datasets zijn ook sinds kort beschikbaar

8 Patiëntenvolgsysteem: data online beschikbaar

8 Ewoud Sanders: een overdaad aan bronnen



Scan deze QR code met een smartphone om de website van E-data te bezoeken. www.edata.nl