



Overzicht vooronderzoek Vaccinatieberichten

We hebben een query gerund op Coosto met de woorden 'vaccin*' en 'inenting'. Dit leverde 441.000 berichten op van 1 jan 2020 tot nu. De wordcloud geeft een overzicht.

Hieruit hebben we een random sample van 10.000 berichten getrokken en op verschillende manieren geclusterd (K-means en LDA). Voor de features hebben we gekeken naar woorden, bi en trigrammen en wordembeddings. Uiteindelijk lijkt het erop dat de combinatie van uni- en bigrammen het beste te werken. LDA geeft aan dat er 8 'clusters' te vinden zijn (kunnen er ook 9 of 10 zijn). Hieronder is een overzicht gegeven van de 20 meest voorkomende woorden en de omvang van de 8 clusters. Besef dat de duiding een 'moving target' is (dit is hoe we nu denken dat de berichten gegroepeerd kunnen worden, maar dat kan zomaar weer anders zijn). Het sentiment van de berichten is geen goede maat om als indicator voor de groepen (of het totaal) te gebruiken omdat dit niet goed te bepalen is op basis van de woorden in de berichten. Hiervoor moet je modellen gaan maken.

Topic 1 (2949; 29,5%):

willen, laat, doen, krijgen, trump, snel, ben, na, hun, weet, waar, eerste, nemen, wie, griep, echt, onze, iedereen, eens, want

Interpretatie na lezen berichten: 'Griep' cluster. Bevat vooral niet-covid berichten

Topic #2 (1383; 13,8%):

rusland, rivm, pas, 2021, heel, hugodejonge, persconferentie, eerste, nederlandse, poetin, markt, minpres, 19, ziet, experts, viroloog, veiligheid, begin, rutte, staatsgeheim

Interpretatie na lezen berichten: 'Politiek' cluster. Berichten over niet transparante beslissingen.

Topic #3 (1305; 13,1%):

gates, bill, bill gates, nieuwe, miljoen, 000, geld, 19, virussen, grieprik, staat, mag, verplichting, tijd, wereld, aantal, leiden, doden, beschikbaar, besmettingen

Interpretatie na lezen berichten: 'Bill Gates' cluster. Vooral berichten over behandelingen in een internationale context.

Topic #4 (1289; 12,9%):

hugo, jonge, jullie, maken, grote, verplichte, zonder, youtube, vragen, pandemie, af, covid19nl, via youtube, verplicht, einde, coronamaatregelen, nooit, 19, genoeg, poliovirus

Interpretatie na lezen berichten: 'Wantrouw' cluster. Vooral berichten in een nationale context.

Topic #5(1069; 10,7%):

mijn, test, neemt, kinderen, nederland, doet, werken, europese, heb, china, astrazeneca, eu, kabinet, 20, commissie, vaccindeaf, zegt, nunl, europese commissie, via nunl

Interpretatie na lezen berichten: 'Snel vaccin' cluster, gaat over de vaccindeaf en de risico's van het snel ontwikkelen van een vaccin. Met name voor kinderen.

Topic #6 (966; 9,7%):

who, youtube, polio, via youtube, 19, onderzoek, afrika, themesdna, via themesdna, vaccinatieplicht, vaccine, gates, lees, vanaf, spoedwet, bill, uitbraak, bill gates, impact, ingrediënten

Interpretatie na lezen berichten: 'Tegen' cluster. Berichten die verwijzen naar voorbeelden waar het fout ging.

Topic #7 (625; 6,3%):

astrazeneca, nederland, proefpersoon, telegraaf, kreeg, via telegraaf, ontsteking, oxford, ruggenmerg, bijwerkingen, ontsteking ruggenmerg, problemen, gekocht, proefpersoon kreeg, kreeg ontsteking, eind, nederland gekocht, problemen nederland, gekocht proefpersoon, 2020
Interpretatie na lezen berichten: 'Bijwerkingen' cluster. Voornamelijk berichten over de bijwerkingen van het covid vaccin.

Topic #8 (414; 4.1%):

nos, medicijn, hcq, blijkt, enkele, elkaar, sinds, af, via nos, hoeven, onderzoeken, dagen, mogen, hydroxychloroquine, pharma, vroeg, inzetten, treatment, early, early treatment

Interpretatie na lezen berichten: 'Alternatieven voor vaccineren' cluster, Gaat voornamelijk over andere snelle alternatieve aanpakken (early treatment).