



Sociaal en Cultureel Planbureau

Verantwoording verwerking persoonsgegevens

Onderzoeksrapport *Coronasceptis in de media*

Het SCP deed aan de hand van verschillende methoden onderzoek naar coronascepsis in Nederland. In dit document geven we meer uitleg over hoe we persoonsgegevens hebben verwerkt voor het onderzoeksrapport *Coronascepsis in de media*.

Voor dit onderzoeksrapport verzamelden we berichten van verschillende sociale media: Twitter, Facebook, Instagram, YouTube, BitChute, 4Chan en Telegram. Het gaat om berichten op de openbare delen van de platformen en die bestaan uit geschreven tekst.

Het SCP werkte voor dit onderzoeksrapport samen met onderzoekers van de Universiteit van Amsterdam (UvA) en de Rijksuniversiteit Groningen (RUG). Onderzoekers van deze universiteiten downloadden berichten van de geselecteerde sociale media die relevant waren voor de onderzoeksvragen van dit project. Zij richtten zich hierbij op berichten die relevant waren voor de onderzoeksvragen van dit rapport:

- 1) Waar gaat coronascepsis over op sociale media?
- 2) Hoe veranderen coronasceptische ideeën op sociale media over tijd?
- 3) Wat zijn de belangrijkste verschillen in coronasceptische ideeën tussen verschillende platformen.

Om dit onderzoek uit te voeren hebben onderzoekers van de UvA en RUG via API's van platformen en inhoudsanalysetools van de universiteiten berichten verzameld. Het ging daarbij om berichten die tussen januari 2020 en april 2022 zijn geplaatst. Het downloaden van berichten vond plaats tussen september 2021 en april 2022.

Om de privacy van de gebruikers van sociale media te beschermen is er op dit project 'privacy by design' toegepast. De berichten zijn na het downloaden ontdaan van alliansen. In verband met de mogelijke (indirecte) herleidbaarheid van de socialemediaberichten, is de map waarin deze socialemediaberichten zijn opgeslagen, afgeschermd. Alleen projectmedewerkers kregen toegang tot deze map. Rechten worden bepaald op basis van betrokkenheid bij het onderzoeksproject. Het SCP levert deze bestanden ook niet aan derden. De verzamelde gegevens worden door het SCP nooit herleidbaar tot een individueel persoon naar buiten gebracht. Daarom citeren we niet uit socialemediaberichten in de onderzoeksrapporten. Geheimhoudingsafspraken met de UvA en de RUG zijn vastgelegd in een overeenkomst over de gezamenlijke verwerkingsverantwoordelijkheid.

Belang van het SCP bij deze gegevensverwerking

Het SCP is sinds 1973 een overheidsinstituut, onderdeel van het ministerie van VWS. Het SCP volgt, verklaart en verkent hoe het met de inwoners van Nederland gaat op sociaal en cultureel gebied. Het SCP heeft een onafhankelijke wetenschappelijke positie. Uit het Meerjarenplan 2021-2025 blijkt dat het SCP zich focust op politiek en institutioneel vertrouwen, legitimiteit van beleid, sociale cohesie en de rol die media daarin spelen. Sociale media vervullen een belangrijke rol in het maatschappelijke en politieke debat. Om die rol en de invloed van sociale media op politiek en institutioneel vertrouwen of sociale cohesie te onderzoeken, is het nodig om onderzoek te kunnen doen naar hoe debatten op sociale media gevoerd worden. Dit soort onderzoek kan beleid ten aanzien van bijvoorbeeld sociale mediaplatformen informeren. Het onderzoeken van socialemediaberichten doet het SCP dus voor de vervulling van een taak van algemeen belang (wetenschappelijk onderzoek). Het algemene belang van het SCP is gelegen in de noodzakelijkheid om de publieke taak uit te oefenen. Het gaat om taken die in regeling aanwijzing planbureaus zijn vastgelegd.

Doelbeperking

De verzamelde berichten van sociale media zijn alleen voor dit onderzoek gebruikt.

Gegevensbeperking

Er is alleen gezocht op coronasceptische termen. We hebben niet gezocht op personen, of andere onderwerpen die niet aan coronasceptis raken.

Juistheid

Het doel van het onderzoek was om een goed beeld te krijgen van waar coronasceptis over gaat. Gericht zoeken en juiste gegevens verwerken is ook daarvoor cruciaal. In eerste instantie lazen we veel socialemediaberichten en stelden we een lijst van coronasceptische zoektermen op. Met deze zoektermen doorzochten we de sociale media en vervolgens downloadden we van al die berichten de bijbehorende hashtags. Hashtags bestaan vaak uit woorden die de post kernachtig beschrijven. Mensen gebruiken ze om aan bestaande debatten te linken en om meer aandacht voor een bericht te genereren. Het zoeken van berichten op basis van de nieuw gevonden hashtags leverde weer nieuwe termen op om mee te zoeken. Het heen en weer gaan tussen de berichten en onze zoektermen verhoogde bij iedere stap de validiteit van de meting. Uiteindelijk deelden we de hashtags en zoektermen in een aantal categorieën in die naar ons inzicht verschillende onderwerpen in de discussies vertegenwoordigden. Bij deze selectie en indeling keken we steeds goed of de kernwoorden sceptisch van aard waren en welke indeling van zoektermen het best de verschillende onderwerpen vertegenwoordigde. Het selecteren en indelen deden we met meerdere onderzoekers; twijfelgevallen bespraken we om tot een weloverwogen indeling te komen. Van iedere zoekterm lazen we ook een aantal berichten, om te kijken of die berichten inderdaad over corona gingen, sceptisch van aard waren en het onderwerp vertegenwoordigden dat paste bij de indeling. Zo controleerden we steeds of we met de zoektermen wel vonden wat we ermee beoogden te zoeken. Accurate zoektermen vonden we van groter belang dan het ophalen van alle artikelen, waarbij er onvermijdelijk veel ruis zal zijn. Termen die we als ambigu beoordeelden, deelden we niet in.

Bewaartermijn

Alle socialemediaberichten zijn direct na de analysefase van het onderzoek verwijderd, zowel aan de kant van het SCP als aan de kant van de RUG en UvA. Normaal gesproken is het voor de replicerbaarheid van onderzoek noodzakelijk om de gegevens een tijd te bewaren. In dit geval is dat niet noodzakelijk, omdat met de zoektermen, API's en inhoudsanalysetools het onderzoek voor collega-onderzoekers te repliceren is. Het bewaren van de berichten van sociale media is daar niet voor nodig.

Integriteit en vertrouwelijkheid

De socialemediaberichten zijn gepseudonimiseerd. Dat wil zeggen dat de gebruikersnamen zijn verwijderd van de berichten, waardoor de berichten niet meer direct herleidbaar tot personen zijn. De onderzoekers van SCP, RUG en UvA hebben afspraken gemaakt over versleuteld data versturen en veilig data opslaan. Dit laatste is gebeurd in beveiligde mappen, waar alleen de onderzoekers van dit project bij kunnen. In de rapporten is niet uit de socialemediaberichten geciteerd, waardoor informatie vanuit het onderzoeksrapport niet herleidbaar is. Er wordt alleen op samenvattend niveau over content op sociale media gerapporteerd.