Rathenau Instituut

Scrollend naar de stembus

De rol van aanbevelingsalgoritmen bij inmenging in verkiezingen

Rapport

Auteurs
Luuk Ex, Quirine van Eeden, Wouter Nieuwenhuizen, Mariette van Huijstee, met medewerking van Joost Gerritsen

Illustraties
Laura Marienus/Rathenau Instituut

Foto omslag
Mensen in de rij bij een stembureau in Den Haag in 2021. Foto: Shutterstock

Bij voorkeur citeren als:
Rathenau Instituut (2026). Scrollend naar de stembus – De rol van aanbevelingsalgoritmen bij inmenging in verkiezingen. (Auteurs: Ex, L., Van Eeden, Q., Nieuwenhuizen, W., & Van Huijstee, M., m.m.v. Gerritsen, J.)

Voorwoord

In 2024 werden Roemeense verkiezingen ongeldig verklaard nadat het berichtenverkeer op social media mogelijk in het voordeel van een kandidaat was gemanipuleerd. Ook in Nederland bestaan er zorgen over beïnvloeding van verkiezingen vanuit het buitenland. De Tweede Kamercommissie voor Digitale zaken vroeg ons te onderzoeken welke rol de aanbevelingsalgoritmen van grote socialmediaplatformen kunnen spelen bij inmenging in verkiezingen.

Die algoritmen vormen een belangrijke schakel in het bepalen welke berichten mensen online te zien krijgen. De algoritmen werken daarbij als een soort megafoon. Ze bepalen welke stemmen versterkt worden en welke niet. Inhoud waar mensen naar blijven kijken en die een reactie oproept (likes, comments, shares, kijktijd) wordt versterkt. Mensen blijven dan langer op een platform en dat is goed voor de advertentieomzet.

Er zijn uiteenlopende mogelijkheden voor kwaadwillende buitenlandse actoren om slim op deze aanbevelingsalgoritmes in te spelen en zo invloed uit te oefenen op het politieke debat in Nederland. Zo kunnen zij bijvoorbeeld met honderden nepaccounts tegelijk een vergelijkbaar bericht versturen en mensen daarmee misleiden over de populariteit van de inhoud. Of ze betalen influencers om een boodschap te versterken.

Er zijn al veel verschillende maatregelen genomen om inmenging via platformen aan te pakken, zowel door de wetgever als door platformen zelf. Onze studie laat echter zien dat die slechts ten dele voldoen. Wij geven drie handelingsperspectieven om het risico op inmenging te verminderen.

De eerste gaat over het platformontwerp. Er zijn bijvoorbeeld ook andere aanbevelingsalgoritmen mogelijk die de mogelijkheden op inmenging verkleinen. Het tweede handelingsperspectief betreft het versterken van de informatiepositie van toezichthouders, onderzoekers, journalisten en politici. De derde richting gaat over hoe we de weerbaarheid van de samenleving vergroten zodat ze beter bestand is tegen inmengingspogingen, onafhankelijk van wat de platformen doen.

Wij hopen dat deze handelingsperspectieven de Tweede Kamer helpen om de risico's op inmenging in verkiezingen via socialmediaplatformen in te dammen.

Prof. dr. ir. Eefje Cuppen
Directeur Rathenau Instituut

Samenvatting

Op verzoek van de Tweede Kamercommissie voor Digitale Zaken deed het Rathenau Instituut onderzoek naar welke rol aanbevelingsalgoritmen van grote socialmediaplatformen kunnen spelen bij inmenging in verkiezingen.

We analyseerden academische literatuur, juridische literatuur en grijze literatuur zoals journalistieke bronnen, rapporten en verslagen van rechtszaken. Daarnaast hebben we vertegenwoordigers van een aantal platformen (schriftelijk) geïnterviewd. Ook spraken we beleidsmedewerkers die zich bezighouden met verkiezingen, platformbeleid en inmenging. Verder interviewden en consulteerden we tijdens een expertsessie deskundigen uit onder meer de vakgebieden politieke communicatie, informatica, rechten, en mediawetenschappen.

Inmenging
Inmenging is beïnvloeding door of namens een buitenlandse statelijke actor die strijdig is met de soevereiniteit, waarden of belangen van het beïnvloede land. De handelingen zijn dwingend, heimelijk, misleidend of corrumperend. Inmenging kan mensen tegen elkaar opzetten, de stemuitslag beïnvloeden en twijfel zaaien over het democratisch proces.

Inmenging via socialmediaplatformen
We richtten ons op een relatief nieuwe vorm die steeds laagdrempeliger en geavanceerder wordt: inmenging via de aanbevelingsalgoritmen van socialmediaplatformen.

Onder inmenging via socialmediaplatformen verstaan we de georkestreerde inzet van meerdere accounts, pagina's of platformen door buitenlandse statelijke actoren en/of hun vertegenwoordigers om de zichtbaarheid van bepaalde inhoud te vergroten of te verkleinen. Dat doen ze door in te spelen op de aanbevelingsalgoritmen van socialmediaplatformen. Betrokken actoren proberen daarbij de ontvanger van de inhoud bewust te misleiden over de identiteit, motieven, acties of connecties van de betrokken actoren en/of de populariteit van de inhoud. Doel is om uiteindelijk de publieke opinie of verkiezingen te beïnvloeden.

Aanbevelingsalgoritmen
Aanbevelingsalgoritmen zijn een verzameling aan systemen die inhoud sorteren en rangschikken op basis van data, met als doel om mogelijk interessante inhoud aan te raden aan gebruikers. Aanbevelingsalgoritmen zijn een antwoord op het probleem van informatieoverload op het internet: uit al de beschikbare inhoud wordt

geprobeerd jou iets te laten zien wat je leuk, interessant of anderszins relevant vindt.

Platformen en hun ontwerpers hebben een voorkeur voor inhoud die zorgt dat gebruikers lang op het platform blijven en het platform vaak bezoeken. Sinds 2016 stapten veel grote platformen over op het rangschikken van inhoud op basis van het verwachte engagement (likes, reacties, tijd besteed, etc.). Want hoe meer engagement, hoe hoger de advertentie-inkomsten.

Het gevolg van aanbevelingsalgoritmen gebaseerd op engagement is dat bepaalde inhoud meer zichtbaar gemaakt kan worden (*geamplificeerd,* versterkt).

Het inmengingsinstrumentarium
Opruiende of manipulatieve inhoud die wordt verspreid om kiezers te beïnvloeden, is gemaakt om engagement van mensen op te roepen, zoals liken, doorsturen of een video opnieuw kijken. Kwaadwillende (buitenlandse) actoren kunnen gebruik maken van de mogelijkheden die op engagement gebaseerde aanbevelingsalgoritmen bieden om een groot bereik te krijgen.

Dit onderzoek laat zien dat het instrumentarium voor inmengingsactiviteiten op socialmediaplatformen groot en divers is, en dat de mogelijkheden voor inmenging steeds laagdrempeliger en geavanceerder worden. Socialmediaplatformen veranderen continu, waardoor het instrumentarium ook verandert. Nu socialmediaplatformen steeds meer gebruikmaken van aanbevelingsalgoritmen, spelen kwaadwillende actoren hierop in. Ze gebruiken bijvoorbeeld hyperactieve accounts en coördineren hun berichten en hashtags.

Bestaande maatregelen voldoen slechts deels
De wetgever heeft instrumenten ontwikkeld om verschillende aspecten van inmenging aan te pakken. Voorbeelden zijn het vergroten van de verantwoordingsplicht van platformen, het ingrijpen op het ontwerp en het vergroten van transparantie over circulerende inhoud. En platformen hebben richtlijnen, c*ommunity guidelines*, over welke inhoud en activiteiten op hun platforms verboden zijn.

Experts signaleren echter dat juist eenvoudige maatregelen, zoals het verwijderen van nepaccounts, niet door platformen genomen lijken te worden. Aanbevelingsalgoritmen zouden ook kunnen worden aangepast zodat ze minder geschikt worden voor inmengingsactiviteiten.

Bovendien is het voor onafhankelijke onderzoekers en voor journalisten moeilijk om grip te krijgen op de aard, omvang en effectiviteit van inmenging via

socialmediaplatformen. Het onderzoek is namelijk afhankelijk van medewerking van de platformbedrijven. Welke inhoud er relatief meer rondgaat en wie dat bereikt, is daardoor onduidelijk. Dat maakt het moeilijk om te beoordelen of de maatregelen van platformen effectief zijn. Recente voorbeelden in Nederland en Europa maken duidelijk dat aanbevelingsalgoritmen een scala aan mogelijkheden voor inmengingsactiviteiten bieden en daarmee een risico vormen voor het online politieke debat.

Drie handelingsperspectieven om risico op inmenging te verkleinen
Om de risico's op inmengingsactiviteiten te verkleinen, geeft dit rapport handelingsperspectieven in drie richtingen: (1) platformontwerp veranderen, (2) informatiepositie versterken, (3) weerbaarheid vergroten. Een combinatie van deze handelingsperspectieven is wenselijk.

De eerste richting vraagt om aanpassingen aan aanbevelingsalgoritmen en om het inzetten op meer fundamentele veranderingen in het landschap van socialmediaplatformen. De tweede richting gaat over het versterken van de informatiepositie van toezichthouders, onderzoekers, journalisten, politici en beleidsmakers. De derde richting vraagt om blijvende bescherming van verkiezingsprocessen en huidige wetgeving en investeren in onafhankelijke onderzoekers en weerbare burgers.

De drie handelingsperspectieven:

1. **Platformontwerp veranderen**
    a. Stimuleer platformen om aanbevelingsalgoritmen aan te passen.
    b. Intervenieer in governance en verdienmodel.
    c. Zet in op decentralisering en interoperabiliteit.

2. **Informatiepositie versterken**
    a. Maak inzicht mogelijk in platformdata voor overheid, toezichthouders en onderzoekers.
    b. Institutionaliseer onafhankelijke, transparante en proactieve detectie van inmenging.

3. **Weerbaarheid vergroten**
    a. Versterk verkiezingsprocedures.
    b. Bescherm de huidige wetgeving en verhelder waar nodig.
    c. Investeer in onafhankelijke journalistiek.
    d. Investeer in technologisch burgerschap en mediawijsheid.

Inhoud

Begrippenlijst

Aanbevelingsalgoritmen: Aanbevelingsalgoritmen zijn een verzameling aan systemen die inhoud sorteren en rangschikken op basis van data, met als doel om mogelijk interessante inhoud aan te raden aan gebruikers.

Amplificatie: Voor het begrip amplificatie (letterlijk: versterking) bestaan verschillende definities. Als wij het begrip amplificatie gebruiken, doelen we op wat algoritme-onderzoekers Thorburn, Stray en Bengan 'relatieve algoritmische amplificatie' noemen: 'Een verandering in de verspreiding van inhoud onder een aanbevelingsalgoritme, in vergelijking tot een alternatief, terwijl gebruikersgedrag constant blijft' (Thorburn et al., 2023).

Deze definitie verheldert dat bij amplificatie altijd duidelijk moet worden dat het gaat om amplificatie *ten opzichte* van een alternatieve situatie, en dat je moet aannemen dat het gedrag of de voorkeuren van gebruikers hetzelfde blijft.

Contentmoderatie: Contentmoderatie is het proces waarmee bepaald wordt onder welke voorwaarden inhoud, zoals tekst, foto's en video's, wel en niet is toegestaan binnen online omgevingen (Rathenau Instituut, 2025a). Dit proces omvat de beoordeling van content gegenereerd door gebruikers op het platform.

Engagement: In dit rapport gebruiken we de definitie van algoritmenonderzoekers (Stray et al., 2024, p. 13): 'Een set van gebruikersgedragingen, die tijdens normaal gebruik van het platform ontstaan en waarvan wordt aangenomen dat ze samenhangen met waarde voor de gebruiker, het platform of andere belanghebbenden.' In de context van socialmediaplatformen wordt hiermee vaak bedoeld: liken, delen, reacties, tijd besteed, en andere interacties die aangeven hoe waardevol of relevant de content is voor de gebruiker of het platform.

Gebruikersretentie: Het bijhouden en meten van de tijd die een gebruiker doorbrengt op een socialmediaplatform. Het wordt gemeten door te kijken naar hoelang een gebruiker actief blijft tijdens een sessie, en of de gebruiker actief blijft op het platform op de lange termijn.

Inmenging: Inmenging is buitenlandse beïnvloeding door of namens een buitenlandse statelijke actor die strijdig is met de soevereiniteit, waarden of belangen van een staat. Die handelingen zijn dwingend, heimelijk, misleidend of corrumperend van aard. In de context van dit onderzoek hebben die activiteiten ten doel de verkiezingen in Nederland te beïnvloeden. Inmenging in de context van

verkiezingen betekent dat een staat activiteiten onderneemt om een verkiezing in Nederland te beïnvloeden door een deelnemende partij aan die verkiezing te bevoordelen of te benadelen.

Inmenging via social media: Onder inmenging via socialmediaplatformen verstaan we de georkestreerde inzet van meerdere accounts, pagina's of platformen door buitenlandse statelijke actoren en/of hun vertegenwoordigers om de zichtbaarheid van bepaalde inhoud te vergroten of te verkleinen. Dat doen ze door in te spelen op de aanbevelingsalgoritmen van socialmediaplatformen. Betrokken actoren proberen daarbij de ontvanger van de inhoud bewust te misleiden over de identiteit, motieven, acties of connecties van de betrokken actoren en/of de populariteit van de inhoud. Doel is om uiteindelijk de publieke opinie of verkiezingen te beïnvloeden.

Social media: In dit onderzoek hebben we in het bijzonder aandacht voor de socialmediaplatformen van Google (YouTube), Meta (Facebook en Instagram), Microsoft (LinkedIn), Snap (Snapchat), TikTok en X (voorheen Twitter), omdat zij de meeste gebruikers in Nederland hebben.

1 Inleiding

Vrije meningsvorming van burgers is een voorwaarde voor een democratie. Socialmediaplatformen spelen een steeds belangrijkere rol in de meningsvorming over politiek-maatschappelijke vraagstukken. Het zijn plekken waar politieke partijen campagne voeren en direct in contact komen met burgers. Daarnaast komen burgers onderling tot politieke uitwisseling en oordeelsvorming. De inhoud die op socialmediaplatformen circuleert is dus onderdeel van het publieke debat. Aanbevelingsalgoritmen op socialmediaplatformen vormen een belangrijke schakel in het bepalen welke berichten mensen online te zien krijgen. De WRR spreekt in dit kader van opiniemacht van platformen: socialmediaplatformen en meer specifiek hun aanbevelingsalgoritmen, hebben in toenemende mate een poortwachtersfunctie in de informatievoorziening van steeds meer Nederlanders (Wetenschappelijke Raad voor het Regeringsbeleid (WRR), 2024).

De centrale positie van socialmediaplatformen in het publieke debat brengt ook risico's met zich mee. Ze zijn niet alleen plekken voor politieke informatievoorziening, informatie-uitwisseling en campagnevoering, maar ook kanalen waar buitenlandse actoren kunnen proberen de publieke opinie te beïnvloeden. De Europese Unie heeft in het verleden vastgesteld dat buitenlandse actoren via online kanalen probeerden verkiezingen te beïnvloeden (European External Action Service, 2023).

In 2024 werd in Roemenië een verkiezingsronde ongeldig verklaard nadat berichtenverkeer op socialmediaplatformen mogelijk doelbewust in het voordeel van een kandidaat werd beïnvloed. Ook in Nederland bestaan zorgen over beïnvloeding van de verkiezingen vanuit het buitenland (AIVD, 2025; *Kamerstukken II, 36 552, nr. 12*, 2025). Ten tijde van de Tweede Kamerverkiezingen van 2025 zijn verschillende heimelijke en manipulatieve activiteiten op socialmediaplatformen waargenomen, waardoor zorgen over inmenging zijn gegroeid (HEIO Consortium et al., 2026).

Het Rathenau Instituut onderzocht op verzoek van de Tweede Kamercommissie voor Digitale Zaken hoe aanbevelingsalgoritmen op socialmediaplatformen, die voor een belangrijk deel bepalen welke berichten wij online te zien krijgen, kunnen bijdragen aan heimelijke beïnvloeding van verkiezingen vanuit het buitenland. Het resultaat van dit onderzoek is beschreven in dit rapport.

De centrale vraag in dit onderzoek is:
Welke rol spelen aanbevelingsalgoritmen op grote socialmediaplatformen bij inmenging in verkiezingen?

Deze hoofdvraag is beantwoord met behulp van een aantal deelvragen, die in opeenvolgende hoofdstukken in dit rapport behandeld worden.

De deelvragen zijn:
1. Wat is inmenging via socialmediaplatformen in de context van verkiezingen? (hoofdstuk 2)
2. Hoe werken aanbevelingsalgoritmen van grote socialmediaplatformen? (hoofdstuk 3)
3. Hoe kunnen aanbevelingsalgoritmen gebruikt worden voor inmenging in verkiezingen? (hoofdstuk 4)
4. Welke inspanningen om inmenging via aanbevelingsalgoritmen te voorkomen worden al gedaan? (hoofdstuk 5)
5. Welke handelingsopties zijn er om inmenging via aanbevelingsalgoritmen verder te voorkomen en/of aan te pakken? (hoofdstuk 6)

Om de hoofd- en deelvragen te beantwoorden is gebruik gemaakt van literatuuronderzoek, interviews, een expertsessie en een reviewproces. Een overzicht van de geraadpleegde partijen is te vinden in bijlage 1. Een uitgebreide beschrijving van de methode staat in bijlage 2. Bijlage 3 bevat extra informatie over wetgeving en beleid.

1.1 Afbakening

Dit onderzoek richt zich niet op de daadwerkelijke gevolgen van inmenging op verkiezingsuitslagen, maar op de mogelijkheden die aanbevelingsalgoritmen op social media bieden voor het bewust misleiden van het online publieke debat in een poging de publieke opinie en verkiezingen te beïnvloeden. Daarmee sluiten we aan bij een groeiende maatschappelijke en politieke zorg: hoe beschermen we vrije verkiezingen en daarmee de democratie in het digitale tijdperk?

We beschrijven de werking van aanbevelingsalgoritmen uitgebreid en diepgaand in dit rapport. Daarbij geven we ook voorbeelden van hoe aanbevelingsalgoritmen werken op huidige grote online platformen, gebaseerd op wetenschappelijk onderzoek dat gedaan is in samenwerking met platformen of met gebruikers van platformen. Een systematische vergelijking van de verschillen tussen aanbevelingsalgoritmen op grote socialmediaplatformen maakt geen deel uit van dit

onderzoek. Om inzicht te krijgen in verschillen tussen platformen is een ander type onderzoek nodig.

Dit onderzoek richt zich op de rol van aanbevelingsalgoritmen bij inmenging. Het is belangrijk om te realiseren dat er ook routes voor inmenging zijn waarbij het aanbevelingsalgoritme geen rol speelt. Denk bijvoorbeeld aan het bedreigen van een politicus via privéberichten, aan infiltratie, of aan het beïnvloeden van andere media zoals kranten of televisieprogramma's. Ook kan worden gedacht aan recente ontwikkelingen rondom chatbots. Een groeiend aantal mensen lijkt chatbots als stemhulp gebruiken. Het is mogelijk om de output van chatbots te beïnvloeden, maar dit valt buiten het aandachtsgebied van dit onderzoek.

1.2 Leeswijzer

Dit rapport beantwoordt de hoofdvraag aan de hand van bovenstaande vijf deelvragen. Elk hoofdstuk gaat in op een van deze deelvragen, waarna we in de conclusie de hoofdvraag beantwoorden.

In hoofdstuk 2 gaan we dieper in op het begrip inmenging in het algemeen en in de context van social media in het bijzonder. Ook bespreken we de mogelijke gevolgen van inmenging.

In hoofdstuk 3 bespreken we in detail wat aanbevelingsalgoritmen op social media zijn, wat het betekent dat ze inmiddels veelal gebaseerd zijn op engagement, en wat het gevolg daarvan is voor verspreiding van inhoud op social media.

In hoofdstuk 4 pakken we het instrumentarium uit dat dankzij aanbevelingsalgoritmen op social media beschikbaar is voor kwaadwillende (buitenlandse) actoren.

In hoofdstuk 5 beschrijven we op hoofdlijnen de inspanningen die momenteel al bestaan tegen inmenging via social media. We geven een overzicht van bestaande wet- en regelgeving, beschrijven platforminspanningen tegen inmenging, en gaan in op de zorgen die experts ondanks deze inspanningen nog steeds hebben over de mogelijkheden voor inmenging in onze verkiezingen.

In hoofdstuk 6 beantwoorden we onze centrale onderzoeksvraag en beschrijven we welke handelingsopties het onderzoek heeft opgeleverd om inmenging via social media verder te beperken.

2 Inmenging nader beschouwd

Wat is inmenging? Hoe ziet inmenging eruit op socialmediaplatformen in de context van verkiezingen? Hoe wordt inmenging vastgesteld, en wat zijn mogelijke gevolgen ervan? Dat zijn de vragen die we in dit hoofdstuk behandelen.

We gaan eerst in op wat buitenlandse inmenging is. Er is geen breed gedragen wetenschappelijke definitie, daarom gebruiken we een beschrijving die in de Europese praktijk vaak gebruikt wordt. Vervolgens wordt kort toegelicht waarom inmenging problematisch is vanuit internationaalrechtelijk perspectief.

Daarna bespreken we hoe inmenging vorm krijgt op social media en introduceren we een definitie voor dit type inmengingsactiviteit. We lichten toe waarom het afbakenen van dit type activiteit gevoelig ligt en laten zien hoe dit type inmengingsactiviteit eruitziet aan de hand van een casus, de Roemeense presidentsverkiezingen van 2024.

We lichten verder toe waarom het lastig is voor verschillende (publieke) actoren anders dan platformen om inmenging online te herkennen. Het belang van inzicht in de wijze waarop informatie circuleert op online platformen, illustreren we met de casus over nepaccounts uit het buitenland op X rond de recente Tweede Kamerverkiezingen.

Het proces in Nederland om inmenging vast te stellen blijkt verschillende kwetsbaarheden te kennen. Die bespreken we in de daaropvolgende paragraaf. Tot slot gaan we in op de zorgen over inmengingspogingen via social media in het verloop van verkiezingen.

2.1 Wat is inmenging?

Er is geen alom geaccepteerde wetenschappelijke definitie van inmenging (Berzina & Soula, 2020; European Commission, 2023). Ook beleidsmakers hebben moeite met het vinden van een juiste definitie, omdat het lastig is om bij zo'n definitie recht te doen aan de verschillende, constant in ontwikkeling zijnde verschijningsvormen van inmenging en tegelijkertijd de politieke rechten en vrijheden van burgers te respecteren (zie ook 2.2) (Berzina & Soula, 2020).

Een binnen de Europese praktijk veel gehanteerde definitie van buitenlandse inmenging is: 'een type buitenlandse beïnvloedingsactiviteit door of namens een buitenlandse actor op staatsniveau dat niet valt onder diplomatiek verkeer, en zich daarvan onderscheidt doordat het dwingend, heimelijk, misleidend of corrumperend te werk gaat en handelt in strijd met de soevereiniteit, waarden en belangen van de Europese Unie' (eigen vertaling) (Directorate-General for Research and Innovation, 2022, p. 2).

Denk bij dit soort activiteiten aan het onder druk zetten van politieke vertegenwoordigers, (verkapte) financieringssteun leveren, personen in strategische posities onder druk zetten, hybride aanvallen uitvoeren of het verspreiden van desinformatie (Directorate-General for Research and Innovation, 2022).

inmenging in de context van verkiezingen betekent dat een staat activiteiten onderneemt om een verkiezing in Nederland te beïnvloeden door een deelnemende partij aan die verkiezing te bevoordelen of te benadelen (Ohlin & Hollis, 2021).

Als de afzender van dergelijke inmengingsactiviteiten in verkiezingen een staat is, of door een staat wordt gesteund, kan dat een inbreuk zijn op het internationaalrechtelijk principe van non-interventie en artikel 2 lid 4 van het Handvest van de Verenigde Naties (1945; McWhinney, 1966).

Artikel 2 lid 4 schrijft voor dat VN-leden elkaars territoriale integriteit en politieke onafhankelijkheid moeten respecteren. Het verbiedt geweld te gebruiken of op andere wijze de onafhankelijkheid en integriteit van staten te ondermijnen. Het gaat dus uit van het idee dat staten soeverein zijn en zich niet mogen bemoeien met interne politieke aangelegenheden. In de context van verkiezingen is volgens de Amerikaanse rechtswetenschapper Ohlin zelfbeschikking het meest behulpzame begrip: het recht van een volk om te beslissen over zijn eigen politieke lot (Ohlin, 2017).

2.2 Wat is inmenging via social media?

In dit onderzoek richten we ons op een relatief nieuw *type* inmengingsactiviteit dat steeds laagdrempeliger wordt én tegelijkertijd steeds geavanceerder is (Justice for Prosperity, 2025; Ohlin & Hollis, 2021): inmenging via de aanbevelingsalgoritmen van socialmediaplatformen.

Onder inmenging via socialmediaplatformen verstaan we:
De georkestreerde inzet van meerdere accounts, pagina's of platformen door buitenlandse statelijke actoren en/of hun vertegenwoordigers om de zichtbaarheid van bepaalde inhoud te vergroten of te verkleinen. Dat doen ze door in te spelen op de aanbevelingsalgoritmen van socialmediaplatformen. Betrokken actoren proberen daarbij de ontvanger van de inhoud bewust te misleiden over de identiteit, motieven, acties of connecties van de betrokken actoren en/of de populariteit van de inhoud. Doel is om uiteindelijk de publieke opinie of verkiezingen te beïnvloeden.

Deze definitie is grotendeels geïnspireerd op de definitie van de politicologen Thiele en collega's.(Thiele et al., 2025)[1] Met 'bewust te misleiden' bedoelen we dat de inmengende partij de ontvanger bewust probeert te misleiden over de herkomst van de inhoud die de ontvangers te zien krijgen, door als afzender zo slecht mogelijk traceerbaar te zijn online en andere kritieke onderdelen van de operatie te verhullen.

Belangrijk is verder dat de betrokken actor probeert een misleidend beeld te geven van hoe populair bepaalde inhoud online is. Dat doet de actor door de zichtbaarheid te vergroten of te verkleinen door in te spelen op aanbevelingsalgoritmen met behulp van het instrumentarium uit hoofdstuk 4. Daarbij moet worden opgemerkt dat we niet kunnen weten hoe populair de content zou zijn geweest als de actor deze activiteit niet had verricht (zie hoofdstuk 3 over relatieve amplificatie).

Het aspect *bewuste misleiding* is een noodzakelijke voorwaarde voor het beschrijven van inmenging via social media. Dat aspect onderscheidt een inmengingsactiviteit van een doorsnee politieke campagne. Politieke partijen willen immers ook hun online zichtbaarheid op een georganiseerde manier vergroten en zullen daarbij in toenemende mate moeten leunen op de aanbevelingsalgoritmen op social media (zie hoofdstuk 3).

Een heldere definitie is van belang, want het (dis)kwalificeren van een politieke campagne als inmenging in aanloop naar de verkiezingen kan ingrijpende consequenties hebben. Grondrechten als de vrijheid van meningsuiting kunnen geschonden worden en uiteindelijk leiden tot het (averechtse) effect dat het wantrouwen in instituties vergroot wordt. Het vermogen om een campagne als inmenging te kwalificeren kan een ingrijpend instrument zijn dat uiteindelijk ook misbruikt kan worden door antidemocratische krachten.

---

1 'De georkestreerde activiteit van meerdere accounts of pagina's op socialmediaplatformen die gericht is op het beïnvloeden van de publieke opinie of het publieke debat door de zichtbaarheid van specifieke inhoud of actoren te vergroten of te verkleinen met behulp van de mogelijkheden van deze platformen en die het publiek opzettelijk misleidt over de identiteit, motieven, acties of connecties van de betrokken actoren en/of over de populariteit van de verspreide inhoud (eigen vertaling, (Thiele et al., 2025, p. 4).'

De definitie van Thiele en collega's gaat over bewust misleidende activiteiten op social media, waarbij niet relevant is door welke actor die worden uitgevoerd. Met het oog op ons onderzoek hebben we aan hun definitie toegevoegd dat bewuste misleiding plaatsvindt in de context van verkiezingen en wordt uitgevoerd namens of door buitenlandse statelijke actoren.

Experts verschillen van mening over de vraag of de focus van onderzoek naar bewuste misleiding op de rol van buitenlandse actoren moeten liggen. In de eerste plaats omdat detectie lastig is, eens temeer omdat het verhullen van hun operatie inherent is aan de inmengingsactiviteit. Het feit dat bewuste misleiding niet altijd kan worden toegeschreven aan een bepaalde buitenlandse actor, maakt bewuste misleiding bovendien niet minder problematisch. In dit onderzoek is ervoor gekozen om de focus echter op buitenlandse statelijke actoren te leggen.

Een voorbeeld van inmenging via socialmediaplatformen is te lezen in onderstaande casus over de Roemeense presidentsverkiezingen van 2024.

Kader 1 Casus inmenging via social media: Roemeense verkiezingen

In november 2024 won Calin Georgescu in de eerste ronde van de Roemeense presidentsverkiezingen. Dat was volgens velen onverwacht. De kandidaat had geen campagnebudget en was tijdens de verkiezingscampagne nauwelijks zichtbaar op televisie. Een plotselinge piek in zijn online zichtbaarheid op met name het socialmediaplatform TikTok in aanloop naar de verkiezingen, deed vermoeden dat hij een bijzonder succesvolle socialmediacampagne had opgetuigd. Of de campagne heeft geleid tot een andere verkiezingsuitslag, is lastig vast te stellen.

Eén van de belangrijkste bevindingen volgens de Roemeense veiligheidsdiensten, en later ook van enkele ngo's zoals EDMO, Digihumanism, Expert Forum en DFRlab en de Franse veiligheidsdiensten, was de inzet van een gecoördineerde campagne met als doel de zichtbaarheid van Georgescu te vergroten op onder andere TikTok. Georgescu's TikTok-account kreeg in de verkiezingsmaand 2.500% meer volgers.

De campagne gebeurde door via Telegram-kanalen duizenden Roemenen te instrueren om pro-Georgescu boodschappen te verspreiden. Een andere

---

tactiek was het massaal plaatsen van pro-Georgescu-hashtags onder berichten die al populair waren, onder meer bij posts van tegenkandidaat Lasconi. Waarschijnlijk zijn daarbij ook bots ingezet. Verder zijn ook influencers gerekruteerd om Georgescu te promoten. Influencers maakten ogenschijnlijk neutraal ogende video's met indirecte referenties aan de kandidaat en vermeldden daarbij niet dat het om bepaalde politieke content ging. De video's werden overstelpt met berichten van pro-Georgescu accounts. Mogelijke waren daar ook nepaccounts bij.

Georgescu ontkent achter de campagne te zitten, hoewel er aanwijzingen zijn die het tegendeel suggereren. De Roemeense veiligheidsdiensten stellen in hun vrijgegeven documentatie dat een buitenlandse actor erachter zit en stellen verder dat Rusland een lange geschiedenis kent van inmenging in electorale processen in andere landen.

Er zijn experts die parallellen trekken tussen de beïnvloedingscampagne voor Georgescu en die voor Stoianoglo, een Moldavische presidentskandidaat die ook steun kreeg via betaalde Roemeense influencers. Ook heeft Rusland volgens Roemeense onderzoeksjournalisten in ieder geval indirect bijgedragen aan het succes van Ruslandgezinde kandidaten door desinformatiecampagnes in aanloop naar de verkiezingen van het aan Rusland gelieerde marketingbureau AdNow.

Zie ook het artikel op de website van het Rathenau Instituut: Waarom is Roemenië een belangrijk voorbeeld rond politieke inmenging?

2.3 Mogelijke gevolgen van inmenging

Over de mogelijke consequenties van inmenging in het verloop van verkiezingen bestaan verschillende zorgen. Hieronder geven we een overzicht van zorgen genoemd in de geraadpleegde literatuur:

1. Het tegen elkaar opzetten van mensen om tegenstelling in de maatschappij te vergroten. (paragraaf 2.3.1)
2. Het veranderen van de stemuitslag. (paragraaf 2.3.2)
3. Twijfel zaaien over het democratische proces. (paragraaf 2.3.3)

2.3.1 Mensen tegen elkaar opzetten

Heeft inmenging als gevolg dat het bestaande tegenstellingen in de maatschappij versterkt en dat mensen radicaliseren? Daarover bestaan zorgen onder wetenschappers (Vliegenthart et al., 2024). Social media zouden mogelijk affectieve polarisatie veroorzaken: mensen zijn minder positief over andersdenkenden (Miltenburg et al., 2022).

In hoofdstuk 4 bespreken we verschillende manieren waarmee kwaadwillende buitenlandse actoren bewerkstelligen dat inhoud die inspeelt op emoties en grenst aan wat toegestaan is op platformen en extreme inhoud van andere gebruikers op een platform meer zichtbaarheid krijgen dan andere inhoud. Het aanbevelingsalgoritme kan als een vliegwiel werken en dit type inhoud zichtbaarder maken (in hoofdstuk 3 gaan we hier uitgebreid op in).

Mensen die inhoud op het platform bekijken kunnen hierdoor een vertekend beeld krijgen over hoe er op het platform gemiddeld gedacht wordt.

Het versterken van specifieke inhoud kan ook effecten hebben op het publieke debat dat zich afspeelt buiten de platformen. Als bijvoorbeeld in reguliere media wordt bericht over inhoud op social media krijgt de inhoud opnieuw aandacht. De specifieke zorg hierbij is dat wanneer extreme meningen die op social media veel bereik krijgen in reguliere media worden herhaald, ze als meer gemiddeld worden gezien. Dit kan leiden tot het verschuiven van de norm richting extremere standpunten. De vrees is dat deze normverschuiving effect heeft op hoe politici zich uiten. Online cultuur kan namelijk gespiegeld worden in de Tweede Kamer.(Data School Universiteit Utrecht, 2023)

2.3.2 Veranderen van de stemuitslag

Een van de zorgen over inmenging is het veranderen van de verkiezingsuitslag. In de afgelopen jaren zijn er met name zorgen geweest over de verspreiding van mis- en desinformatie via social media. Als voorbeeld wordt vaak verwezen naar de presidentsverkiezingen in de Verenigde Staten in 2016. Tijdens deze verkiezingen is waargenomen dat een deel van de politieke advertenties uit Rusland afkomstig was, en gericht was op kiezers in swingstaten waarbij het doel was meer mensen op de republikeinse kandidaat Donald Trump te laten stemmen. Overigens is er door verschillende onderzoekers gewezen op het beperkte en selectieve bereik van de verspreide berichten (met name werden Republikeinen bereikt die toch al op Trump zouden stemmen) en de aandacht tijdens de verkiezingen voor binnenlandse berichtgeving. Hierdoor wordt aangenomen dat inmenging via social

media geen effect had op kiesgedrag (Honingh & Van Ham, 2024, p. 156). Toch blijven deze zorgen bestaan omdat in toekomstige desinformatiecampagnes het anders kan uitpakken.(Ecker et al., 2024). De ongeldig verklaarde verkiezingen in Roemenië zijn hiervan een voorbeeld. Deze werden ongeldig verklaard omdat het vermoeden was dat de uitslag door inmenging via social media was veranderd (zie kader 1 over de Roemeense verkiezingen).

Vanwege het meerpartijenstelsel is vaak gedacht dat Nederland minder kwetsbaar is. Het type inmenging, zoals in de Verenigde Staten waarbij een poging wordt gedaan om een partij de grootste te maken, zou in Nederland niet werken. In Nederland doen er veel meer partijen mee om de race de grootste te worden. Bovendien moet de winnende partij in Nederland de macht delen in een coalitie. Hierdoor heeft de winnaar relatief minder macht dan in de Verenigde Staten.

Daarnaast werd aangenomen dat het Nederlands taalgebied enigszins bescherming biedt tegen inmenging, omdat minder mensen wereldwijd de taal machtig zijn en er dus ook minder mensen te vinden zijn die inmengingsactiviteiten kunnen uitvoeren (Irwin & van Holsteyn, 2021). Met de huidige vertaalsoftware is het echter voor buitenlandse actoren goed mogelijk om Nederlandse inhoud te maken. Ook is het makkelijker geworden om bestaande te interpreteren en te kiezen welke inhoud geschikt te verspreiden.

Bovendien kan het wellicht juist lonend zijn om in de Nederlandse politiek een verschuiving op zetelniveau te veroorzaken. Als bijvoorbeeld een kleinere partij gesteund wordt, kan het zetelaantal relatief snel groeien, bijvoorbeeld van een naar drie zetels. Dit levert in de Tweede Kamer relatief meer spreektijd op dan wanneer een grote partij er twee zetels bij krijgt.

2.3.3 Twijfel zaaien over het democratische proces

Zelfs wanneer inmenging weinig kiezers zou overhalen om anders te stemmen, bestaan er zorgen over dat mensen minder vertrouwen krijgen in hoe eerlijk de verkiezingen zijn verlopen. Overigens wijzen cijfers over het vertrouwen van Nederlanders in het verkiezingsproces daar momenteel niet op (Garnett et al., 2025) Toch worden hierover zorgen geuit. Inmenging waarbij twijfel wordt gezaaid over een eerlijk verlopen verkiezing, kan het democratische proces in het algemeen ondermijnen (Honingh & Van Ham, 2024).

2.4 Gebrekkige grip op inmenging

Om mogelijke inmenging te signaleren, zijn er in Nederland voorafgaand aan de verkiezingen verkiezingstafels georganiseerd. Deze tafels bestaan uit ministeries, veiligheidsdiensten, het Nationaal Cybersecurity Centrum (NCSC), de Autoriteit Consument en Markt (ACM), gemeenten en de Kiesraad. Ze signaleren risico's en nemen maatregelen. De veiligheidsdiensten spelen een belangrijke rol bij het detecteren van gedrag van statelijke actoren in zoverre dat de nationale veiligheid bedreigt (*Kamerstukken II, 36 552, nr. 12*, 2025).

Waar het gaat om de circulatie van inhoud die een bedreiging vormt voor verkiezingen, kan het ministerie van Binnenlandse Zaken inmiddels haar zogeheten verkiezingsflaggerstatus inzetten. Als een platform een melding krijgt van een verkiezingsflagger, dan moet het met prioriteit beoordelen of een bericht een risico kan vormen voor de integriteit van een verkiezingsproces. Het ministerie meldt het overigens niet op eigen initiatief maar na berichtgeving in de media of na meldingen van onafhankelijke factcheckers of andere overheidsinstanties.

Het ministerie van BZK heeft bijvoorbeeld op 17 oktober 2025 een melding gedaan bij X over de verspreiding van foutieve online instructies voor GroenLinks-PvdA-stemmers om te stemmen op twee kandidaten. X stelde vervolgens dat de inhoud in lijn was met hun gebruikersvoorwaarden en heeft de inhoud laten staan. Het ministerie van BZK heeft ook Meta op de post gewezen. Dat gebeurde op 20 oktober. Meta heeft de post op 25 oktober weggehaald (*Kamerstukken II, 35 165, nr. 102*, 2026).

Doordat de circulatie van informatie zich steeds meer concentreert op een handjevol platformen waar miljarden burgers samenkomen (Rathenau Instituut, 2025c), hebben deze platformen het beste zicht op wat er online gebeurt. Ze weten het beste welke inhoud wie bereikt, van wie die afkomstig is en door wie er wordt betaald.

Die informatie is niet zomaar voor iedereen beschikbaar. Onderzoekers hebben slechts een algemeen beeld van hoe algoritmen werken. Wat gebruikers online te zien krijgen en welke rol aanbevelingsalgoritmen hierin spelen, blijft grotendeels onbekend (Cooper & Chapman, 2025).

Voor journalisten, wetenschappers, ngo's en institutionele actoren als het ministerie van BZK en de Kiesraad is het niet eenvoudig te achterhalen of er sprake is van inmengingsactiviteiten via socialmediaplatformen. De toegang tot betrouwbare en controleerbare informatie voor dergelijke partijen is echter essentieel om hun rol goed te kunnen vervullen. Een kwalificatie van inmenging in verkiezingen zonder

uitgebreide en transparante onderbouwing kan het wantrouwen in een samenleving juist vergroten. De Digitale dienstenverordening (DSA) probeert het zicht daarop voor toezichthouders en aangewezen onderzoekers te vergroten. Of, en in hoeverre, dat lukt, bespreken we in hoofdstuk 5.

Overigens is er sprake van een institutionele kwetsbaarheid die in de context van inmenging relevant kan zijn. Het vaststellen van de verkiezingsuitslag en het besluit tot een eventuele herstemming wordt namelijk gemaakt door de zittende Tweede Kamer.[2] Die heeft echter een direct belang bij de beslissing of een verkiezing opnieuw moet. Als de nieuwe vertegenwoordigers eenmaal zijn toegelaten tot de Tweede Kamer of de gemeenteraad is de uitslag onherroepelijk. Er is geen beroep bij de rechter mogelijk.

Bovendien is nergens gespecificeerd in welke gevallen de Tweede Kamer kan besluiten tot een herstemming (Raad van State, 2024; Trapman, 2024a). Jurisprudentie van het Europees Hof voor de Rechten van de Mens, een advies van de Kiesraad en een advies van de Raad van State problematiseren de huidige wijze waarop ons kiesstelsel verkiezingsgeschillen beslecht (*Mugemangango t. Belgium*, 2022; Raad van State, 2024).

Kader 2 Nepaccounts uit buitenland op X tijdens Nederlandse verkiezingen

In Nederland berichtte RTL Nieuws rondom de Tweede Kamerverkiezingen van 2025 over georkestreerde activiteiten van meerdere accounts die bewust misleidend waren. Journalisten onderzochten ruim 550 accounts die zich voor en na de Tweede Kamerverkiezingen hebben bemoeid met politieke berichten. De 550 accounts verstuurden in totaal ruim 23.000 tweets. Het grootste deel hiervan waren retweets van accounts van zeer actieve en uitgesproken Nederlanders.

De accounts hadden nepnamen en namen van bekende Nederlanders. De accounts deden zich voor als Nederlandse gebruikers door gebruik te maken van Nederlandse accountomschrijvingen en door berichten van Nederlandse

---

2 De Tweede Kamercommissie voor het Onderzoek van de Geloofsbrieven is belast met het uitbrengen van een verslag aan de Tweede Kamer over het verkiezingsverloop, de vaststelling van de uitslag en de toelating van de leden. De Kamer is formeel niet gebonden aan de bevindingen van de commissie, maar neemt in de praktijk wel haar conclusies over (Trapman, 2024b).

politieke actoren te delen. Ze werden voornamelijk aangestuurd vanuit landen in West-Afrika. Zeker 225 accounts bleken beheerd te worden vanuit Nigeria.

Doordat de nepaccounts tweets van kleinere accounts begonnen te delen leken deze mogelijk populairder dan wanneer de campagne niet actief was geweest. Mogelijk zijn deze berichten daardoor aan meer X-gebruikers getoond dan zonder de activiteiten van de nepaccounts (hoeveel gebruikers de tweets daadwerkelijk hebben gezien is alleen door het platform te zeggen).

In deze inmengingscampagne zijn drie vormen van inmenging te zien. Ten eerste het versterken van bestaande tegenstellingen. De meeste interacties van de nepaccounts waren retweets over onderwerpen waarover in Nederland politiek discussie bestaat zoals over migratie en de situatie in Gaza. Ten tweede versterkten de nepaccounts een specifiek politiek perspectief. Dit deden de nepaccounts door rechtsgeoriënteerde politici en opiniemakers te retweeten. Het doel hiervan is vermoedelijk om deze accounts een groter bereik te geven. Forum voor Democratie en PVV-leider Geert Wilders werden het meest geretweet. Ter vergelijking, de CDA en VVD kregen allebei één retweet van het netwerk. Ten derde werd er inhoud verspreid om twijfel te zaaien over het democratische proces. Het bericht dat de meeste retweets van de nepaccounts kreeg, was een bericht waarin verkiezingsfraude werd gesuggereerd.

Twee experts geven in het bericht van RTL Nieuws aan dat ze vermoeden dat Rusland betrokken is. De manier van het orkestreren van de inmengingsactiviteiten had kenmerken van een door CNN beschreven inmengingscampagne die in 2020 werd georganiseerd vanuit Ghana en Nigeria. (Clarissa Ward et al., 2020). *(Trollenlegers uit buitenland versterkten politieke en opruiende berichten rond verkiezingen, 2025)*

2.5 Conclusie

In dit hoofdstuk hebben we een algemene definitie voor inmenging geïntroduceerd. Inmenging is een buitenlandse beïnvloedingsactiviteit door of namens een buitenlandse actor op staatsniveau die niet valt onder diplomatiek verkeer, en zich daarvan onderscheidt doordat het dwingend, heimelijk, misleidend of corrumperend te werk gaat en handelt in strijd met de soevereiniteit, waarden en belangen van een staat. Buitenlandse inmenging in de context van verkiezingen betekent dat een

staat activiteiten onderneemt om een verkiezing in Nederland te beïnvloeden door een deelnemende partij aan die verkiezing te bevoordelen of te benadelen.

Als de afzender van dergelijke inmengingsactiviteiten in verkiezingen een staat is, of door een staat wordt gesteund, kan dat een inbreuk zijn op het internationaalrechtelijk principe van non-interventie en artikel 2 lid 4 van het VN-Handvest.

Onder inmenging via socialmediaplatformen verstaan we de georkestreerde inzet van meerdere accounts, pagina's of platformen door buitenlandse statelijke actoren en/of hun vertegenwoordigers om de zichtbaarheid van bepaalde inhoud te vergroten of te verkleinen. Dat doen ze door in te spelen op de aanbevelingsalgoritmen van socialmediaplatformen. Betrokken actoren proberen daarbij de ontvanger van de inhoud bewust te misleiden over de identiteit, motieven, acties of connecties van de betrokken actoren en/of de populariteit van de inhoud. Doel is om uiteindelijk de publieke opinie of verkiezingen te beïnvloeden.

Net als de algemene definitie voor inmenging, kent de definitie van inmenging op social media uitdagingen. Dat heeft te maken met het feit dat sommige componenten lastig te bewijzen zijn en dat miskwalificatie grote impact kan hebben op de vrijheid van meningsuiting.

De gevolgen van inmenging kunnen zijn dat mensen tegen elkaar worden opgezet doordat bestaande tegenstellingen in de samenleving op socialmediaplatformen worden uitvergroot, extreme meningen een groter bereik krijgen, de stemuitslag wordt beïnvloed, of dat wantrouwen in het democratisch proces wordt aangewakkerd. In het volgende hoofdstuk onderzoeken we de rol die aanbevelingsalgoritmen van socialmediaplatformen spelen bij inmengingscampagnes.

Het zicht van journalisten, beleidsmakers, onderzoekers, burgers op de informatie die op socialmediaplatformen circuleert, neemt af, terwijl het informatieverkeer zich steeds sterker op deze platformen concentreert. Het wordt daardoor lastiger om onafhankelijk inmengingsactiviteit vast te stellen en door andere partijen te laten controleren. Ook ligt er een kwetsbaarheid bij het feit dat beslissingen over de betwisting van de uitslag en het eventueel organiseren van een herstemming politiek gekleurd kunnen zijn.

3 Hoe werken aanbevelingsalgoritmen?

In dit hoofdstuk behandelen we de vraag: hoe werken aanbevelingsalgoritmen van grote socialmediaplatformen? Sinds ongeveer 2016 zijn we steeds meer gewend geraakt aan inhoud die ons aanbevolen wordt op social media, zonder dat we daar zelf actief om hebben gevraagd. Aanbevelingsalgoritmen spelen een steeds grotere rol in wat we online te zien krijgen. Daarom is het belangrijk om te begrijpen hoe aanbevelingsalgoritmen werken.

We bespreken in dit hoofdstuk wat aanbevelingsalgoritmen gebaseerd op engagement zijn, en welke alternatieve vormen van aanbevelingsalgoritmes ook bestaan.

3.1 Hoe wordt bepaald wat jij online ziet?

Op elke tijdlijn is maar beperkt plek. Platformontwerpers moeten daarom kiezen welke berichten ze in je tijdlijn laten zien. En ze maken keuzes over hoe die inhoud[3] wordt gerangschikt. Het Rathenau Instituut laat in het onderzoek Inclusief online zien hoe de visie, het verdienmodel en de eigenaarsstructuur van online platformen van grote invloed zijn op ontwerpkeuzes achter aanbevelingsalgoritmen en de manier waarop de gebruikersomgeving ontworpen is (Rathenau Instituut, 2025a). De keuze voor het advertentie-verdienmodel werkt bijvoorbeeld dataverzameling en gepersonaliseerde aanbevelingsalgoritmen in de hand.[4] Keuzes over welke berichten op jouw tijdlijn verschijnen, zijn dus geen neutrale of puur technische ontwerpkeuzes.

Socialmediaplatformen kunnen berichten op verschillende manieren rangschikken. Daarvoor maken ze gebruik van algoritmen: een set aan instructies om een bepaald probleem op te lossen, zoals bepalen welke inhoud gebruikers te zien krijgen uit heel veel mogelijke opties (Rathenau Instituut, 2022). Algoritmen hebben

3 In literatuur over aanbevelingsalgoritmen wordt vaak gesproken over 'items' in plaats van content of inhoud. Dat is omdat items breder kunnen zijn: ook gebruikersaccounts kunnen worden aanbevolen, en dat valt niet onder inhoud. In dit rapport geven we aan wanneer we inhoud bedoelen en wanneer accounts.

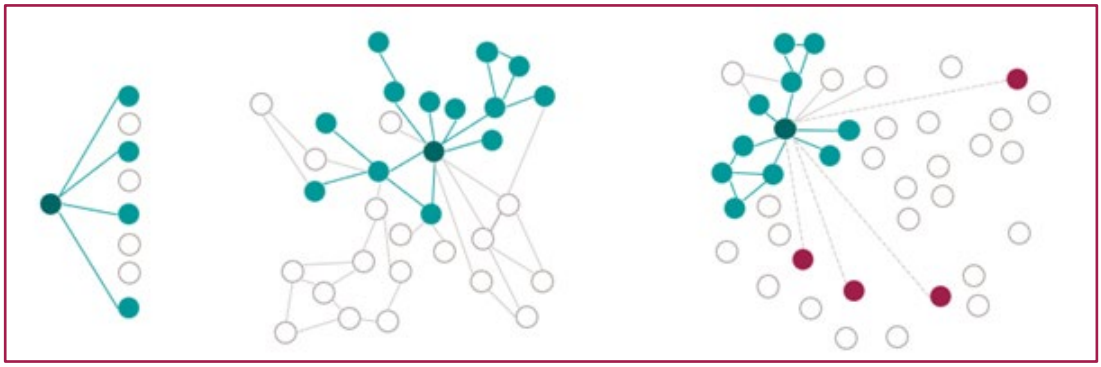
4 De keuze om gebruikers online te volgen, tracken en profileren hangt nauw samen met het dominante huidige verdienmodel van online omgevingen: het dataverdienmodel. In combinatie met het tonen van gepersonaliseerde advertenties zijn platformen geneigd om gebruikers te bewegen richting het aanmaken van accounts, downloaden van een applicatie om gebruikers in hun eigen ecosysteem te houden, en gebruikers aan te moedigen hun echte naam te gebruiken. Daarmee zijn gebruikers gemakkelijker te tracken en profileren. Zie voor meer uitleg het rapport *Inclusief online* (Rathenau Instituut, 2025a).

de reputatie moeilijk te begrijpen te zijn. Maar niet alle algoritmen zijn ingewikkeld. Het rangschikken van berichten op omgekeerd chronologische volgorde is een voorbeeld van een algoritme dat vrij gemakkelijk te doorgronden is: geef de meest recente inhoud weer van iedereen die iemand volgt, en sorteer deze inhoud chronologisch, met de nieuwste inhoud bovenaan.

3.1.1 Rangschikking op basis van abonnement, netwerk en aanbevelingsalgoritmen

We onderscheiden rangschikking op basis van drie modellen: abonnement, netwerk en aanbevelingsalgoritmen (zie ook de figuur hieronder).

Figuur 1 Rangschikking van berichten



Rangschikking op basis van abonnement (links), netwerk (midden) en aanbevelingsalgoritmen (rechts). Deze indeling is gebaseerd op onderzoek uit 2016 door onderzoekers van Stanford University en Microsoft research (Goel et al., 2016). Figuur: Laura Marienus/Rathenau Instituut

In de beginjaren van social media werd inhoud vooral gerangschikt op basis van de modellen abonnement en netwerk. Je zag berichten van accounts die je zelf volgde, in chronologische volgorde, eventueel afgewisseld met advertenties (links in figuur 1 hierboven). Of je zag inhoud die verspreid werd door mensen in jouw netwerk, bijvoorbeeld door retweets (midden).[5] Inmiddels combineren de meeste platformen rangschikking op basis van abonnement en netwerk met aanbevelingsalgoritmen, waarbij ook, of vooral, berichten van accounts van buiten je netwerk worden getoond. Hierdoor is de kans groter dat een gebruiker inhoud ziet van mensen van wie die nooit had aangegeven deze te willen volgen.

5 Tegenwoordig verplicht Europese wetgeving grote platformen ertoe om ook deze niet-gepersonaliseerde rangschikking van je tijdlijn te blijven aanbieden. Artikel 38 van de Digitaledienstenverordening (DSA) verplicht platformen om minstens één aanbevelingssysteem dat niet gebaseerd is op profilering aan te bieden aan gebruikers.

Bekende voorbeelden van rangschikking op basis van aanbevelingsalgoritmen (rechts) zijn de voor-jou-pagina's van TikTok, Instagram en X, of *Discover* op Snapchat. Rangschikking op basis van abonnement (wie je volgt) en wat mensen in je netwerk delen speelt ook op die platformen nog een rol, maar in steeds beperktere mate. TikTok is zelfs groot geworden dankzij dit model, waarbij vrijwel alle zichtbaarheid via aanbevelingsalgoritmen tot stand komt (Narayanan, 2022).

Aanbevelingsalgoritmen, maar ook verspreiding binnen een netwerk, kunnen bijdragen aan het viraal laten gaan van inhoud op socialmediaplatformen. Het begrip *viraal gaan* is geleend van hoe er gesproken wordt over virussen die zich verspreiden. Viraal gaan is ook online dus iets anders dan populair zijn. Viraliteit gaat over hoe inhoud zich via andere mensen verspreidt binnen een bepaald netwerk (Narayanan, 2022). Dat doet denken aan de modellen die wetenschappers in de coronacrisis gebruikten om te voorspellen hoe besmettingen zouden verlopen. Op dezelfde manier kan iets op het internet populair zijn, puur omdat iemand veel volgers heeft waarmee iets gemakkelijk door veel van die mensen opgepakt kan worden (verspreiding door abonnement). Dat is nog niet hetzelfde als viraal gaan (Goel et al., 2016).

Iemand die het voor elkaar krijgt om een viraal te gaan, vindt dus een groter en vooral ook nieuw publiek. In hoofdstuk 4 gaan we in op tactieken die mensen gebruiken om te proberen dit mechanisme te beïnvloeden.

3.1.2 Verschillen tussen socialmediaplatformen

Tussen grote socialmediaplatformen bestaan verschillen in hoe wordt bepaald wat jij ziet. Dat kunnen verschillen zijn in de gebruikersinterface: welke interactie wordt door het socialmediaplatform aangemoedigd of ontmoedigd? Denk bijvoorbeeld aan de introductie van frictieloos ontwerp zoals de *endless scroll* (waarbij je als gebruiker eindeloos naar beneden kunt scrollen en er nooit een einde komt aan de hoeveelheid inhoud die je kan zien), of de mogelijkheid om te reageren op openbare inhoud. Maar ook verschillen in de manier waarop inhoud wordt gerangschikt: hoe groot is de rol van inhoud uit je eigen netwerk? En op basis van welke data wordt bepaald wat je voorgeschoteld krijgt?

Met een grotere rol voor aanbevelingsalgoritmen verschuift de controle over wat je als gebruikers ziet steeds meer naar de ontwerpers van aanbevelingsalgoritmen. Niet alle online omgevingen werken op deze manier: op bijvoorbeeld sociaal netwerk Mastodon zie je nog steeds alleen berichten van mensen die je zelf volgt, of die gedeeld zijn door mensen in jouw netwerk. Een platform als Substack, waar

je je kunt abonneren op nieuwsbrieven van makers, combineert alle drie de vormen van rangschikking (McKenzie & Monga, 2024; Substack, 2025).

Tabel 1 Drie manieren van rangschikking online

	Manier 1: Rangschikking door abonnement	**Manier 2: Rangschikking door delen in netwerk**	**Manier 3: Rangschikking door aanbevelingsalgoritmen**
Wat zien gebruikers?	Alleen inhoud van kanalen die je volgt	Inhoud van accounts die je volgt, inclusief inhoud van anderen die worden gedeeld door deze accounts	Inhoud waarvan aanbevelingsalgoritmen voorspellen dat je er interactie mee aan zal gaan (bijv. een video afkijken, bericht liken of doorsturen naar iemand anders)
Wat zijn bekende voorbeelden?	Abonnement op nieuwsbrieven, podcasts of RSS-feeds	Twitter en Instagram tot 2016, Mastodon, optionele pagina 'Volgend' op TikTok, Instagram en Facebook	YouTube, LinkedIn, Pinterest, standaardoptie op Instagram, X en TikTok
Hoe krijgt content een groter bereik?	Doordat mensen actief aangeven zich erop te willen abonneren	Doordat mensen het bericht delen in hun netwerk	Doordat aanbevelingsalgoritmen voorspellen dat een bericht relevant voor je is

Tabel: Rathenau Instituut

3.2 Wat zijn aanbevelingsalgoritmen?

Aanbevelingsalgoritmen zijn een verzameling aan systemen die sorteren en rangschikken op basis van data, met als doel om mogelijk interessante inhoud aan te raden aan gebruikers. Deze aanbevelingen zijn gebaseerd op factoren als de inhoud van het bericht, je eerdere gedrag en interacties binnen je netwerk (Huszár et al., 2022b).

Aanbevelingsalgoritmen worden bijvoorbeeld gebruikt om te bepalen welke nieuwe serie jou wordt aangeraden, of welk paar schoenen lijkt op iets dat je al eerder gekocht hebt (Ricci et al., 2022). Maar ze zijn vooral bekend van socialmediaplatformen. Denk aan hoe je tijdlijn op socialmediaplatformen eruitzien: staan de nieuwste berichten bovenaan of de meest 'relevante' berichten? Het aanbieden van 'relevante' inhoud gaat altijd op basis van aanbevelingsalgoritmen. Aanbevelingsalgoritmen vervullen dus een functie in het rangschikken en sorteren

van de grote hoeveelheid inhoud die op het internet beschikbaar is: ze zijn een technisch antwoord op het probleem van informatieoverload.(Lin et al., 2024).

Let op: binnen het wetenschappelijke vakgebied van machinelearning worden de combinatie van modellen en systemen die worden gebruikt om inhoud aan te bevelen (van dataverzameling tot monitoring, filtering, rangschikking en training) vaak *recommender systems* genoemd. Een aanbevelingsalgoritme refereert dan aan de specifieke techniek die als onderdeel van een *recommender system* wordt gebruikt om voorspellingen te doen. Het zijn dus andere termen, die buiten de wetenschap vaak door elkaar worden gebruikt. In dit onderzoek gebruiken we *aanbevelingsalgoritmen* als synoniem voor *recommender system*.

Aanbevelingsalgoritmen kunnen ook voor een deel zelflerend zijn op basis van feedback. Een aanbevelingsalgoritme kan bijvoorbeeld leren van gebruikersinteractie en op basis daarvan iemand andere inhoud voorschotelen.

Het aanbevelingsalgoritme leert zelfstandig, maar heeft wel een doel dat door mensen bepaald is. Als het doel is om iemands tijd op een platform te vergroten, zal het aanbevelingsalgoritme leren van inhoud waar iemand veel tijd aan besteedt en dit type inhoud opnieuw aanraden.

Uiteindelijk raden socialmediaplatformen inhoud aan die overeenkomt met de doelen die het platform met het aanbevelingsalgoritme wil bereiken (KGI Expert Working Group on Recommender Systems, 2025).

Hoe meer tijd mensen doorbrengen op een socialmediaplatform, hoe meer advertentie-inkomsten ze genereren voor de bedrijven achter deze platformen. Het loont dus om aanbevelingsalgoritmes zo te optimaliseren dat ze het engagement van gebruikers vergroten. Engagement is een set van gebruikersgedragingen, die tijdens normaal gebruik van het platform ontstaan en waarvan wordt aangenomen dat ze samenhangen met waarde voor de gebruiker, het platform of andere belanghebbenden (Stray et al., 2024). In de context van social media wordt hiermee vaak bedoeld: likes, delen, reacties, tijd besteed, en andere interacties die aangeven hoe waardevol of relevant de content is voor de gebruiker of het platform.

Onderzoekers van het Vector Institute in Canada leggen een aanbevelingsalgoritme uit als een functie *f* die de waarde van een item *i* (een stuk content) voorspelt voor de gebruiker *u*, weerspiegeld in $\hat{f}_{ui}$. De functie *f* is vaak gebaseerd op historische data. Θ representeert de parameters van het model, gebaseerd op eerdere data (Raza et al., 2026). Zie ook figuur 2, hierna.

Figuur 2 Wiskundige weergave van een aanbevelingsalgoritme

$$\hat{r}_{ui} = f(u, i; \Theta)$$

Bron: (Raza et al., 2026).

Aanbevelingsalgoritmen dragen bij aan dat rangschikken en sorteren door de 'waarde' te voorspellen die inhoud voor een gebruiker heeft. Het doel van die voorspelling kan per platform anders zijn. Bij een webshop gaat het om de kans dat je een product in je winkelwagentje stopt of zal afrekenen. Op social media kan het gaan om de kans dat jij een post liket, of doorstuurt. Aanbevelingsalgoritmen worden op grote hoeveelheden data getraind voor die verschillende doelen.

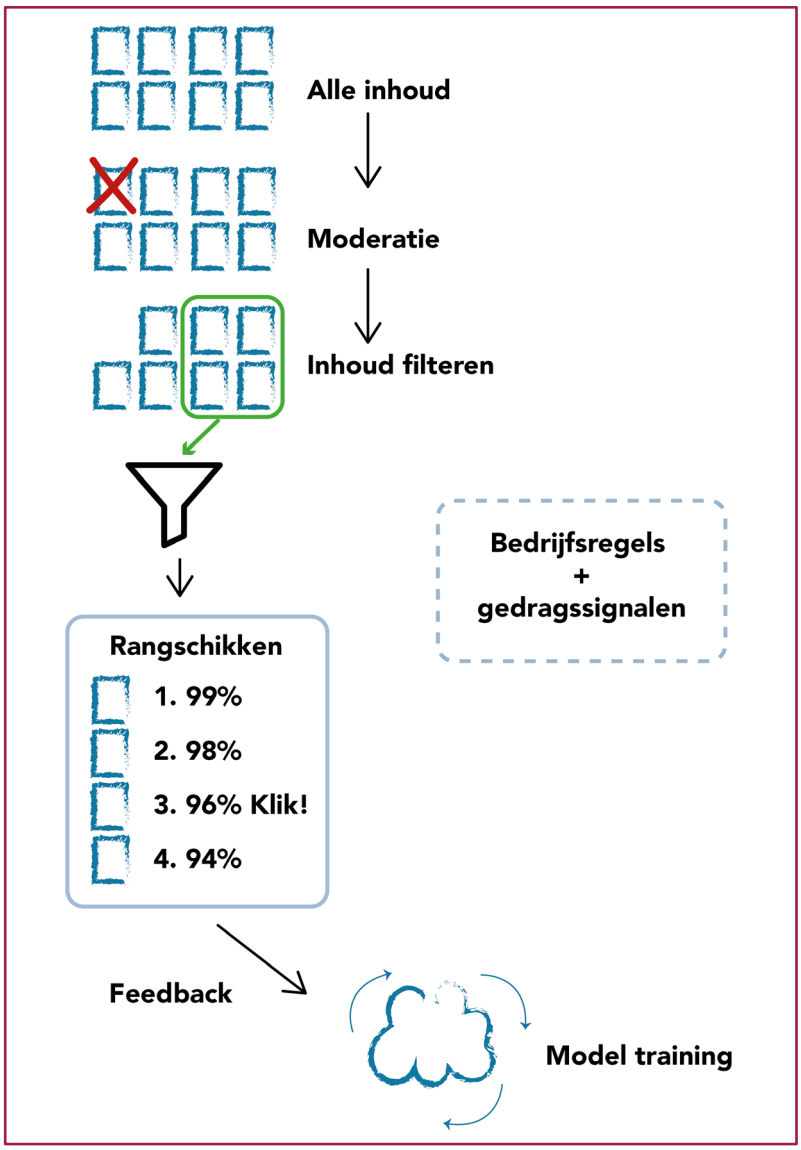
3.3 In vijf stappen naar gepersonaliseerde aanbevelingen

Voordat inhoud wordt aanbevolen aan gebruikers, worden verschillende stappen doorlopen.

We onderscheiden vijf belangrijke stappen die elkaar opvolgen:
1. Contentmoderatie
2. Aanbevelingsalgoritme trainen
3. Inhoud filteren
4. Gefilterde inhoud rangschikken
5. Leren van feedback

Deze stappen (zie ook figuur 3 en de uitleg hierna) zijn een simplificatie. In werkelijkheid verschillen aanbevelingsalgoritmen (in deze context dus vaak *recommender systems* genoemd) in aanpak en uitwerking. We laten gepersonaliseerde advertenties buiten beschouwing in dit stappenplan, omdat voor die aanbevelingen andere systemen worden gebruikt. Zie voor een uitleg voor hoe platformen tot gepersonaliseerde advertenties komen het rapport *De prijs van gratis internet* (Rathenau Instituut, 2025b).

Figuur 3 Stappenplan: van alle inhoud naar aanbevolen inhoud



Een stappenplan van hoe platformen van 'alle inhoud' naar 'aanbevolen inhoud' gaan. Te zien is hoe in iemands tijdlijn met verschillende waarschijnlijkheidsscores wordt aangegeven welke inhoud in welke volgorde verschijnt. De stap erna is feedback, die leidt tot model training. Figuur: Rathenau Instituut, gebaseerd op (Raza et al., 2026; Stray et al., 2024).

3.3.1 Stap 1: contentmoderatie

De eerste stap in het aanbevelen van inhoud is inhoudsmoderatie, ook wel contentmoderatie genoemd. Contentmoderatie is het proces waarmee binnen online omgevingen bepaald wordt onder welke voorwaarden inhoud, zoals tekst, foto's en video's, wel en niet is toegestaan (Rathenau Instituut, 2025a). Dit proces omvat de beoordeling van content gegenereerd door gebruikers op het platform.

In het rapport *Inclusief online* laat het Rathenau Instituut zien dat online omgevingen verschillende filosofieën, stijlen en waarden hanteren bij het modereren van inhoud. (Rathenau Instituut, 2025a) Platformen kiezen bijvoorbeeld voor meer efficiëntie, of juist kwaliteit. Ze zijn straffend, of meer opvoedend, zo stellen onderzoekers in mens-computerinteractie in een wetenschappelijk paper uit 2023 (Jiang et al., 2023).

3.3.2 Stap 2: modeltraining

Een belangrijke stap in het aanbevelen van inhoud is het (doorlopend) trainen van aanbevelingsmodellen. Op grote socialmediaplatformen is dit een proces waar vele werkprocessen in samenkomen en veel engineers aan samenwerken. Het trainen van een aanbevelingsmodel gebeurt op basis van grote hoeveelheden data. Zoals posts die eerder zijn geüpload op het platform, en de interactie van gebruikers met die posts. Die data worden geanalyseerd in een trainingspijplijn, waarin het model wordt getraind op grote hoeveelheden data en specifieke doelen, zoals het vinden van inhoud die op elkaar lijkt. Vervolgens worden de uitkomsten van het model gevalideerd en wordt daar weer van geleerd.

Voor het trainen van aanbevelingsmodellen wordt gebruikgemaakt van technieken als *computer vision* (het door computers interpreteren en analyseren van beeldmateriaal) en *natural language processing* (het verwerken en classificeren van tekstmateriaal).

3.3.3 Stap 3: inhoud filteren

De derde stap in het aanbevelen van inhoud is het maken van een eerste grove selectie van inhoud om te laten zien. Dat gebeurt met een *retrieval model* en deze stap wordt ook wel *filtering* genoemd. Het doel is om het kaf van het koren te scheiden en posts te filteren die een hoge kans hebben om later ook hoog gerangschikt te worden (Stray et al., 2024; TensorFlow, 2023).

We onderscheiden drie belangrijke manieren waarop aanbevelingsalgoritmen inhoud op social media filteren. De drie filtermanieren zijn:

1. Content-based filtering
2. Collaborative filtering
3. Hybride filtering

Content-based filtering

Aanbevelingsalgoritmen die op content gebaseerd zijn, bevelen inhoud aan die lijkt op wat een gebruiker in het verleden leuk heeft gevonden (Roy & Dutta, 2022). Een aanbevelingsalgoritme dat films aanraadt om te kijken, werkt op basis van vergelijkbare kenmerken van films die iemand in het verleden een positieve score heeft gegeven. Films met vergelijkbare kenmerken (regisseur, genre, thema's, etc.) zullen dan een hogere kans hebben om aanbevolen te worden.[6]

Deze manier van inhoud aanbevelen is niet afhankelijk van de interactie van of met andere gebruikers, maar gaat puur over de interactie die een gebruiker met inhoud heeft. Bijvoorbeeld door iets te liken of vaak opnieuw te kijken.

Collaborative filtering

Collaborative filtering werkt op basis van de interactie van vergelijkbare gebruikers. Jouw gebruikersprofiel (met bijvoorbeeld je leeftijd, kijkgedrag, apparaat waarop je inlogt en andere voorkeuren) wordt vergeleken met andere gebruikers. Het aanbevelingsalgoritme raadt vervolgens inhoud aan gebaseerd op de kenmerken van andere gebruikers die op jou lijken. Als je bijvoorbeeld geïnteresseerd bent in inhoud over mode en als vergelijkbare gebruikers naast mode ook veel inhoud over treinreizen bekijken, dan zal jij ook sneller inhoud over treinreizen krijgen aanbevolen. Ook als die inhoud ogenschijnlijk niks met elkaar te maken heeft.

Hybride filtering

Vaak wordt voor het aanbevelen van inhoud een combinatie van verschillende typen filtering gebruikt. Een systeem dat gebruikers met elkaar vergelijkt, leert dan bijvoorbeeld van de uitkomsten van een systeem dat typen inhoud met elkaar vergelijkt. Of het ene systeem probeert de uitkomst van de andere te verbeteren, op basis van andere data. Veel hedendaagse socialmediaplatformen gebruiken een combinatie van *content based filtering* en *collaborative filtering* samen met verschillende vormen van machinelearning, zoals *reinforcement learning* of *large language models* (Lin et al., 2024; Raza et al., 2026).

Reinforcement learning is een vorm van machinelearning waarbij een wiskundig model probeert een beloning te bereiken (bijvoorbeeld een gebruiker die een item liket) door trial-and-error en door te leren van feedback. Een *agent* raadt bijvoorbeeld video's aan een gebruiker aan met als doel dat een gebruiker de video kijkt. Elke keer dat een gebruiker de video wegklikt, leert de agent: dit was geen

6 *Content-based filtering* werkt op basis van vectorisatie: het wiskundig representeren van tekst of beeld. Met simpele vormen van vectorisatie, zoals met het wiskundige model Word2vec dat in 2012 werd ontwikkeld door Google, kan de relatie tussen woorden berekend worden. De woorden *kat* en *hond*, liggen dan wiskundig dichtbij elkaar, net als de woorden *Amsterdam* en *Nederland*. Op die manier kan inhoud gezocht worden die lijkt op wat iemand eerder al leuk heeft gevonden. (Mikolov et al., 2013).

goede aanbeveling. Na verloop van tijd zullen de aanbevelingen beter aansluiten op het doel van de agent: video's selecteren die mensen afkijken.

Soms is bijna geen data beschikbaar over de voorkeuren van een gebruiker, vooral als iemand zich voor het eerst aanmeldt op een platform, of wanneer iemand zonder account inlogt. In machinelearning wordt dit ook het wel het koudestartprobleem genoemd. Historische data ontbreken om jou 'vergelijkbare' inhoud voor te schotelen. Platformen kunnen dan andere gegevens gebruiken om jou inhoud te laten zien. Bijvoorbeeld inhoud die recent populair is geweest in jouw land of regio.

3.3.4 Stap 4: rangschikken van gefilterde inhoud

Filtering heeft in de vorige stap gezorgd voor een grove selectie van inhoud om te tonen. In stap 4 (rangschikken) draait het om het rangschikken en ordenen van inhoud in iemands tijdlijn. Dat gebeurt op basis van waarschijnlijkheid: wat is de kans dat je als gebruiker de inhoud relevant vindt? Platformen proberen verschillende vormen van interactie te voorspellen. Die voorspelling wordt bepaald door gedragssignalen, zoals of je iets een like geeft, of je een video helemaal afkijkt of doorstuurt. Gedragssignalen kunnen ook een negatieve uitwerking hebben waarvan geleerd kan worden. Laat je veel video's die op elkaar lijken links liggen? Dan krijgen die de volgende keer een lagere waarschijnlijkheidsscore.

Platformen gebruiken verschillende typen data om inhoud te rangschikken. Dat rangschikken gebeurt op grote socialmediaplatformen op basis van engagement, maar kan ook gebeuren op basis van andere signalen, zoals feedback uit gebruikerssurveys. Ook kunnen aanbevelingsalgoritmen worden getraind om te rangschikken met andere doelen dan engagement. Zoals mensen dichter bij elkaar brengen.

Voor de rangschikking van inhoud onderscheiden platformen kwaliteitsscores, gedragssignalen, gewichten (*weights*) en *business rules*.

Kwaliteitsscores
Platformen kunnen proberen om volgens hen laagwaardige content, zoals spam en clickbait, lager te rangschikken. Platformen gebruiken verschillende variabelen om kwaliteit mee te nemen in rangschikking. Zo betaalt Google mensen om content een rating te geven om aan te geven of een zoekresultaat kwalitatief is, en gebruikt deze informatie voor hun voorspellingen over kwaliteit van zoekresultaten (Google, 2025c). LinkedIn voorspelt wanneer posts 'kennis en advies' bevatten en gebruikt die informatie in hun aanbevelingsalgoritme (Feifer, 2023).

Platformen kunnen er bijvoorbeeld ook voor kiezen om inhoud van nieuwe accounts lagere kwaliteitsscores te geven, als uit hun risicoanalyses blijkt dat nieuwe accounts vaak kwalitatief laagwaardige inhoud delen. Een dergelijke keuze zou dan wel weer nadelig kunnen zijn voor alle nieuwe gebruikers op een platform.

Gedragssignalen en andere metrics
Gebruikers genereren bewust en onbewust een groot aantal gedragssignalen. Dat kunnen bewuste signalen zijn (zoals op like klikken) maar ook onbewuste signalen (zoals hoelang je blijft hangen op een specifieke foto). Onderzoekers van de Expert group on recommender systems van het Knight Georgetown Institute wijzen ook op het verschil tussen impulsieve signalen (een link delen zonder er eerst zelf op geklikt te hebben) en deliberatieve signalen (een lang commentaar schrijven) (Knight Georgetown Institute, 2025).

Makers van content kunnen mensen ook uitnodigen om signalen te genereren (zoals het bekende 'like and subscribe') omdat ze verwachten dat hun inhoud daardoor aan meer gebruikers zal worden aangeraden.

In tabel 2 en tabel 3 op de volgende pagina geven we voorbeelden van de voorspellingen die worden gebruikt om inhoud te rangschikken.

Ook andere informatie kan gebruikt worden om te voorspellen welke inhoud aan gebruikers aangeraden, zoals of een post een link bevat, of de populariteit van iemand in een netwerk. Zie voor deze voorbeelden tabel 3 op de volgende pagina.

Tabel 2 Voorspellingen en gedrag

Voorspelling	Gedragssignalen die de voorspelling beïnvloeden
Hoe waarschijnlijk het is dat je minder dan drie seconden van een reel bekijkt*	• Hoe vaak het bericht is overgeslagen binnen twee seconden nadat het werd geopend • Hoe vaak deze reel is bekeken met het geluid aan • Hoe vaak gebruikers in *Ontdekken* ten minste drie seconden van de reel hebben bekeken • Hoe vaak de reel is genegeerd • Hoe vaak gebruikers ten minste drie seconden van de reel hebben bekeken, overal op Instagram
De kans dat je op een van de grote, rechthoekige reels in *Ontdekken* klikt om deze op volledig scherm te bekijken op het tabblad *Reels*	• De tijd die je hebt besteed aan het bekijken van reels • Hoeveel berichten en reels je hebt gezien in *Ontdekken* • Op hoeveel grote reels je hebt geklikt in *Ontdekken* om ze op volledig scherm te bekijken op het tabblad *Reels*
Hoe waarschijnlijk het is dat je naar de volgende reel scrolt	• Het aantal berichten en reels dat je hebt gezien in *Ontdekken* • Welke berichten en reels je hebt gezien in *Ontdekken* • Hoe vaak je naar de volgende reel bent gescrold wanneer je reeksen met reels bekeek in *Ontdekken*

Voorbeelden van voorspellingen en gedragssignalen die een rol spelen bij het aanbevelen van inhoud op Instagram Reels. *Reels zijn korte video's gemaakt door andere Instagramgebruikers die gebruikers op verschillende plekken in de Instagram-app kunnen bekijken. Bron: Meta (30 juli 2025). Selectie door het Rathenau Instituut.

Tabel 3 Voorbeelden van metrics

Categorie	Metric
Plaatser van inhoud	• Centraliteit in het netwerk • Populariteit van de plaatser buiten het platform
Interactie gebruiker-plaatser	• Volgen van de plaatser • Blokkeren van de plaatser
Content	• Content bevat link • Content is nieuw • Content is kwalitatief van aard (menselijk beoordeeld)

Bron: Voorbeelden van andere typen metrics die een rol kunnen spelen in aanbevelingsalgoritmen. Gebaseerd op (Cunningham, 2023).

Gewichten (*weights*)

Platformen maken keuzes over welke gewichten (*weights*) ze bepaalde gedragssignalen toekennen. Dat betekent dat ze kiezen welke gedragssignalen ze meer of minder belangrijk maken voor de voorspelling van het aanbevelingsalgoritme. Vindt het platform dat een reactie plaatsen zwaarder weegt dan het delen van een bericht? Of dat het volledig kijken van een video belangrijker is dan het consequent kijken van alle video's van één account?

Inhoud op het platform krijgt een rankingscore op basis van de voorspelling en het bijbehorende gewicht. Daarbij geldt: score = ($gewicht_1$ x $voorspelling_1$) + ($gewicht_2$ x $voorspelling_2$) + ... + ($gewicht_n$ x $voorspelling_n$)

Tabel 4 Voorspellingen en gewichten op Twitter

Voorspelling	Gewicht
De kans dat de gebruiker deze tweet als favoriet aanmerkt	0,5
De kans dat de gebruiker de tweet retweet	1,0
De kans dat de gebruiker reageert op de tweet	13,5
De kans dat de gebruiker het auteursprofiel van de tweet opent en een tweet van de auteur liket of op een tweet reageert	12,0
De kans dat (voor een video-tweet) de gebruiker minstens de helft van de video zal afkijken	0,05
De kans dat de gebruiker reageert op een tweet en dat de auteur van de tweet engagement aangaat met die reactie	75,0
De kans dat de gebruiker zal klikken op het gesprek van deze tweet en reageert op een tweet of een tweet liket	11.0
De kans dat de gebruiker zal klikken op het gesprek van deze tweet en daar voor minstens 2 minuten blijft	10.0
De kans dat de gebruiker negatief reageert op de tweet (verzoek om 'minder te laten zien' op de tweet of auteur, of het blokkeren of muten van de auteur	-74.0
De kans dat de gebruiker zal klikken op 'rapporteer tweet'	-369.0

Alle voorspellingen en gewichten (*weights)* die in 2023 een rol speelden in het rangschikken van tweets op de voor-jou-tijdlijn van Twitter (april 2023).(Twitter-team, z.d.)

Bedrijfsregels

Niet alleen gedragssignalen en hun gewicht, maar ook bedrijfsregels kunnen de rangschikking beïnvloeden. Zo heeft X bepaald dat 50% van de gepersonaliseerde tijdlijn van gebruikers bestaat uit inhoud van accounts die je niet volgt (Twitter-team, z.d.) Dat geldt voor iedereen, ongeacht of je nu 4 of 4.000 mensen volgt.

Meta heeft hun bedrijfsregels over politieke inhoud de afgelopen jaren meermaals aangepast (Stepanov & Gupta, 2021). In 2021 begonnen ze te experimenteren met het verminderen van 'politieke inhoud' in de feeds van mensen in Canada, Brazilië en Indonesië.[7] Op basis van deze experimenten concludeerde Meta dat mensen minder politieke content in hun tijdlijn wilden zien. Om dat te bereiken, paste Meta de gewichten van comments en likes aan bij het rangschikken van politieke content. Daardoor zagen mensen minder inhoud die ze 'niet waardevol vinden', aldus Meta.

Op 7 januari 2025 kondigde Meta aan politieke content niet meer automatisch te *deranken* op Facebook, Instagram en Threads.[8] Daarvoor in de plaats kwam een optie om minder politieke content te zien. Op Facebook kunnen gebruikers ervoor kiezen helemaal geen politieke content aanbevolen te krijgen. Bovenstaande voorbeelden tonen hoeveel invloed de ontwerpkeuzes van platformen kunnen hebben op wat mensen online te zien krijgen.

3.3.5 Stap 5: leren van feedback

Door uitspraken van platformen uit het verleden weten we ook dat er voortdurend wordt gesleuteld aan welke vorm van engagement het zwaarste weegt. Zo maakte YouTube in 2012 bekend dat totale kijktijd vanaf dan belangrijker werd gewogen dan hoeveel mensen klikten op een bepaalde video (Google, 2012). YouTube zag namelijk dat veel mensen aantrekkelijke plaatjes (thumbnails) gingen gebruiken om mensen te verleiden op een video te klikken, terwijl de video dan inhoudelijk niet altijd door veel mensen afgekeken werd. YouTube probeerde dit op te lossen door hun aanbevelingsalgoritme zo aan te passen dat bekeken video's belangrijker werden dan video's waarop alleen geklikt werd.

Platformen gebruiken A/B-tests, online gecontroleerde experimenten, om het effect van nieuwe ontwerpkeuzes op groepen gebruikers te testen (Quin et al., 2024). Dat gebeurt door een hypothese te formuleren ('deze aanpassing aan het aanbevelingsalgoritme zal ervoor zorgen dat mensen vaker de app openen') en twee groepen gebruikers variant A en variant B voor te schotelen. Het zijn als het ware live-experimenten die platformen inzicht geven in het effect van hun ontwerpkeuzes op gebruikersgedrag en doelen die voor het platform belangrijk zijn.

---

7 Deze tests werden hierna verder uitgebreid naar de Verenigde Staten, Costa Rica, Zweden, Spanje en Ierland.

8 Meta definieert 'politieke inhoud' breed: het omvat inhoud rondom overheden, verkiezingen en sociale onderwerpen. Deze tests werden hierna verder uitgebreid naar de Verenigde Staten, Costa Rica, Zweden, Spanje en Ierland.

Controle over wat je te zien krijgt
Socialmediaplatformen geven verschillende expliciete en impliciete opties om invloed uit te oefenen op wat je te zien krijgt. Afhankelijk van hoeveel gewicht platformen dit geven in hun aanbevelingsalgoritme, kun je zelf accounts volgen of ontvolgen. Soms kun je ook specifieke berichten verbergen, een duimpje geven of reageren op een directe vraag of je een aanbeveling relevant vond (Stray et al., 2024). De mate van inspraak varieert wel per platform. Op Instagram wordt bijvoorbeeld actief feedback gevraagd in een pop-up in het scherm. Gebruikers kunnen hier aangeven wat ze van een bericht vinden. Wat de daadwerkelijke invloed (gewicht, *weight*) is van deze mechanismen op de voorspellingen van het aanbevelingsalgoritme, is onbekend.

Uit onderzoek blijkt dat Nederlandse jongeren tussen de 16 en 26 jaar oud hun invloed op algoritmische selectie inschatten als gering (Swart, 2021). Veel jongeren geven aan dat ze niet de moeite nemen om actief te proberen het aanbevelingsalgoritme naar hun voorkeuren te bewegen.

3.4 Waarom platformen overstapten op engagement

Sinds 2016 zijn veel grote platformen overgestapt op het rangschikken van inhoud op basis van het voorspellen van engagement. Dat deden ze naar verluid stuk voor stuk op basis van intern onderzoek dat liet zien dat de gebruikersretentie hoger was wanneer inhoud gesorteerd werd op basis van engagement (Cunningham et al., 2025). Gebruikersretentie wordt zowel gemeten op de korte termijn (hoelang een gebruiker actief blijft tijdens de sessie) en op de lange termijn (of de gebruiker later terugkeert).

Facebook en Instagram deden in 2020 een intern experiment tussen gebruikers met een semi-chronologische tijdlijn en een tijdlijn met rangschikking gebaseerd op engagement. Wat bleek? Op Facebook gingen mensen met een semi-chronologische feed 20% minder tijd besteden dan de mensen die de engagementrangschikking kregen voorgeschoteld en op Instagram 10% minder (Guess et al., 2023). Twitter vond vergelijkbare resultaten: gebruikers die sinds 2016 (als onderdeel van een wetenschappelijk experiment) een chronologische tijdlijn hielden in plaats van een tijdlijn met aanbevelingsalgoritme, bekeken gemiddeld 38% minder inhoud per dag (Bandy & Lazovich, 2023).

Wanneer werkt een aanbevelingsalgoritme nou? Vanuit het perspectief van machinelearning werkt een aanbevelingsalgoritme als het doel wordt bereikt om jou 'relevante' inhoud laten zien. Dat wordt gemeten aan de hand van gedragssignalen, zoals of je op een aanbeveling klikt. Ook als je niet blij wordt van de inhoud waar je

op klikt, of als inhoud je verdrietig of boos maakt, is het wiskundige doel van het aanbevelingsalgoritme nog steeds bereikt. Voor bouwers van aanbevelingsalgoritmen is het pas een ongewenste uitkomst van het systeem als niet de juiste dingen aanbevolen worden.

De aanname van de makers van aanbevelingsalgoritmen is vaak dat gedragssignalen zoals likes en hoe lang iemand op een platform doorbrengt een indicator zijn van wat mensen willen. Dat terwijl uit psychologisch onderzoek blijkt dat het impulsieve gedrag van mensen niet altijd een indicator is voor wat ze op de lange termijn willen. Gedragseconomen Kleinberg, Mullainatha en Raghavan zien dit als een mismatch tussen wat platformen denken te weten over gebruikers op basis van wat gebruikers doen (liken, de app openen) in plaats van wat gebruikers echt willen (zoals blij zijn) (Kleinberg et al., 2024).

Bedrijven investeren veel geld en tijd in het doorontwikkelen van machinelearningtechnieken om technisch betere voorspellingen te doen tegen zo laag mogelijke kosten (rekenkracht en energie). In een wetenschappelijk paper uit 2019 laten machinelearningonderzoekers zien hoe hun technische verbetering van een machinelearningmodel bijdroeg aan 0,37% meer engagement op YouTube (Yi et al., 2019). Op grote schaal kunnen zelfs deze ogenschijnlijk kleine verbeteringen veel geld opleveren. Meta verdiende in 2024 bijna 160 miljard dollar per jaar, grotendeels door advertentie-inkomsten. In 2016 was dit nog maar 26 miljard dollar: een stijging van 515%.

3.5 Wat is amplificatie van inhoud online?

In debatten over de impact van aanbevelingsalgoritmen op wat mensen online te zien krijgen, speelt *amplificatie* een prominente rol. Soms wordt amplificatie (letterlijk: versterking) als synoniem gebruikt voor verspreiding. Soms wordt iets als amplificatie benoemd als mensen inhoud te zien krijgen die niet overeenkomt met hun voorkeuren. Amplificatie is dan ook niet altijd een helder gedefinieerd begrip. Dus hoe bepaal je of een bepaalde post geamplificeerd is door een platform?

In ons onderzoek gaan we uit van de definitie van algoritme-onderzoekers Luke Thorburn, Jonathan Stray en Priyanjana Bengani:

Relatieve algoritmische amplificatie: Een verandering in de verspreiding van inhoud onder een aanbevelingsalgoritme, in vergelijking tot een alternatief aanbevelingsalgoritme, terwijl gebruikersgedrag constant blijft (Thorburn et al., 2023).

Deze definitie verheldert dat bij amplificatie altijd gesproken moet worden van amplificatie *ten opzichte* van een alternatieve situatie, en dat je moet aannemen dat het gedrag of de voorkeuren van gebruikers hetzelfde blijft. Als een platform anders ontworpen was geweest, was de content niet geamplificeerd geweest. Voor deze interpretatie is een *counterfactual* belangrijk, een alternatief scenario. Bijvoorbeeld de amplificatie van een bericht door een aanbevelingsalgoritme, in plaats van op een chronologische tijdlijn. Een ander voorbeeld is de amplificatie van een bericht nadat het platform hun aanbevelingsalgoritme heeft aangepast.

Ook andere vormen van rangschikking van inhoud, zoals de omgekeerd chronologische tijdlijn, zijn niet neutraal (zie paragraaf 3.1.1). In dat soort tijdlijnen krijgt content van mensen die heel vaak en veel posten bijvoorbeeld automatisch meer zichtbaarheid. Maar je kunt wel proberen te onderzoeken of bepaalde berichten meer relatieve amplificatie krijgen bij aanbevelingsalgoritmen op basis van engagement, in vergelijking met een chronologische tijdlijn. Of je kunt vergelijken wat het effect is van verschillende bedrijfsregels (zie paragraaf 3.3.4) op relatieve amplificatie.

De definitie die we in dit onderzoek gebruiken, past ook bij de complexe relatie tussen aanbevelingsalgoritmen en wat mensen te zien krijgen: verschillende factoren beïnvloeden elkaar. Gebruikers passen hun gedrag ook weer aan op ontwerpkeuzes achter het aanbevelingsalgoritme. Daardoor verandert hun interactie met inhoud. En dat kan weer leiden tot aangepaste aanbevelingsalgoritmen.

3.6 Welke inhoud krijgt meer relatieve amplificatie?

Socialmediaplatformen rangschikken inhoud gebaseerd op engagement omdat het de gebruikersretentie vergroot. Dat kan leiden tot voordelen voor gebruikers, zoals het ontdekken van niche-inhoud die je zelf niet opgezocht zou hebben. Of, vanuit de verzender kan het een breder bereik opleveren. (Narayanan, 2023). Tegelijkertijd kan die rangschikking gebaseerd op engagement ook op gespannen voet staan met publieke waarden, zoals veiligheid en de autonomie van gebruikers.

Onder wetenschappers bestaat veel discussie over de definities en effecten van zogenoemde echokamers en filterbubbels. De termen zijn vaak slecht gedefinieerd en daarom ook moeilijk empirisch te onderzoeken, zo stelt onderzoeker Alex Bruns in Internet Policy Review (Bruns, 2019). De discussie draait vaak om de vraag: verstoren aanbevelingsalgoritmen de eigen persoonlijke keuze voor de content die iemand zou willen zien? En is het dus mogelijk dat iemand door aanbevelingen allerlei inhoud ziet die die persoon in eerste instantie niet wilde zien? Filterbubbels

zouden bijvoorbeeld de *confirmation bias* (het onbewust zoeken naar informatie die je opvatting bevestigt) van mensen kunstmatig nog groter maken door de manier waarop aanbevelingsalgoritmen werken (Pariser, 2012).

Onderzoekers proberen grip te krijgen op het type inhoud en de typen mensen die meer lijken te profiteren van relatieve amplificatie dan anderen. Ze proberen bijvoorbeeld de 'eerlijkheid' van aanbevelingsalgoritmen te meten. Ook proberen ze, op basis van gebruikersdata of toegang tot platformdata, de vraag te beantwoorden welke inhoud meer relatieve amplificatie krijgt dan andere inhoud.

Om te onderzoeken welke inhoud meer relatieve amplificatie krijgt dan andere inhoud en wat daarbij de rol is van aanbevelingsalgoritmen, gebruiken wetenschappers verschillende methoden. Veelgebruikte methoden zijn datadonaties door gebruikers, het gebruik van accounts die door de onderzoekers worden aangemaakt en onderzoek met toegang platformdata. Verschillende methoden kunnen andere typen inzichten opleveren.

Met datadonaties kan het mediadieet van individuele gebruikers goed in kaart worden gebracht. Het gebruik van onderzoeksaccounts helpt wetenschappers om het aanbevelingsalgoritme te bestuderen en te kijken of aangeraden inhoud overeenkomt met gebruikerssignalen zoals gevolgde accounts. Wetenschappers wijzen erop dat het erg lastig is om causale relaties aan te wijzen tussen aanbevelingsalgoritmen en uitkomsten voor gebruikers (zoals verandering in politieke voorkeur). (Stray et al., 2024).

3.6.1 Emotionele inhoud krijgt meer relatieve amplificatie

Psychologisch onderzoek wijst uit dat content vaker viraal lijkt te gaan als het een bepaalde emotie uitlokt. Dit spoort mensen namelijk aan om een bericht te delen, te liken of erop te reageren (engagement). En dat maakt weer dat het aanbevelingsalgoritme het bericht sneller verspreidt aan mensen (Berger & Milkman, 2012). Onderzoekers psychologie van de Universiteit van Cambridge en de New York University zagen bijvoorbeeld hoe vijandige content over politieke tegenstanders significant meer kans had om gedeeld te worden op Facebook en Twitter tussen 2016 en 2020 (Rathje et al., 2021). In een ander onderzoek vonden onderzoekers dat tweets van Amerikaanse senatoren die negatieve emoties en *moral outrage* bevatten, relatief vaker gedeeld of als favoriet aangemerkt werden op Twitter tussen 2013 en 2021 (Mercadante et al., 2023).

Wetenschappers vermoeden dat aanbevelingsalgoritmen goed inspelen op menselijke bias richting zogenaamde PRIME-inhoud, waarbij de afkorting verwijst naar (Brady et al., 2023): PRrestigious In-group Moral Emotional.

In experimenten van wetenschappers met gebruikers van grote socialmediaplatformen komt ook naar voren welke inhoud amplificatie krijgt door aanbevelingsalgoritmen. Zo lieten Amerikaanse onderzoekers in 2025 zien hoe het op engagement gebaseerde aanbevelingsalgoritme van Twitter (nu X), tweets amplificeerde die negatieve emoties boosheid, verdriet en angst bevatte (Milli et al., 2025). Ook amplificeerde het aanbevelingsalgoritme tweets die ervoor zorgden dat gebruikers zich slechter gingen voelen over mensen met een andere politieke voorkeur. Daarnaast gaven mensen aan dat ze de aanbevolen politieke tweets niet waardeerden (Milli et al., 2025). De onderzoekers suggereren daarom ook dat het aanbevelingsalgoritme van Twitter op bepaalde punten niet aansloot bij gebruikersvoorkeuren.

Door socialmediaplatformen zelf wordt gewezen op borderline-inhoud. Dat is een term die gebruikt wordt om inhoud te beschrijven die in het grensgebied valt van wat wel en niet is toegestaan op het platformen.[9] De term borderline-inhoud is geen begrensde definitie, maar een parapluterm voor verschillende typen inhoud.

Platformen hebben er in het verleden zelf op gewezen dat borderline-inhoud, net als PRIME-inhoud, meer aandacht krijgt van gebruikers dan niet-borderline-inhoud en daardoor vaker wordt aangeraden door aanbevelingsalgoritmen (Gillespie, 2022; Macdonald & Vaughan, 2024).

Een intern onderzoek bij Facebook liet zien dat inhoud met het hoogste bereik, in het top 1-2% percentiel, vaker als 'slecht voor de wereld' dan als 'goed voor de wereld' werden aangemerkt in gebruikerssurveys (Haugen, 2022). Ook weten we uit onderzoek dat engagement vaak negatief gecorreleerd is met kwaliteit van inhoud. Dat betekent dat inhoud met het hoogst voorspelde engagement vaker inhoud bevat die laag scoort op kwaliteit: spam, clickbait, misleidende titels of misinformatie (Cunningham, 2023). Platformen kunnen dit effect proberen te verminderen door kwaliteitsscores toe te kennen aan inhoud (zie paragraaf 3.3.4).

9 YouTube omschreef *borderline content* vier jaar geleden in een blogpost over contentmoderatie als: '*videos that don't quite cross the line of our policies for removal but that we don't necessarily want to recommend to people*'. Meta omschrijft *borderline content* als inhoud die niet verboden is door gebruikersvoorwaarden, maar wel dicht in de buurt komt. (Macdonald en Vaughan, 2024, p. 350).

3.6.2 Populaire inhoud krijgt meer relatieve amplificatie

Onderzoekers proberen met wetenschappelijke experimenten te begrijpen wat voor soort aanbevelingsalgoritmen het meest kunnen bijdragen aan het verspreiden van bijvoorbeeld misinformatie. Fernández, Bellogin en Cantador gebruikten Twitterdata om te bestuderen of het aanpassen van algoritmen invloed heeft op de potentiële verspreiding van misinformatie (Fernández et al., 2021). Zij laten zien dat aanpassingen in de manier waarop het algoritme is ingesteld, invloed heeft op de verspreiding.

Aanbevelingsalgoritmen die grotendeels gebaseerd zijn op *collaborative filtering*, dus die inhoud aanbevelen die andere vergelijkbare gebruikers leuk vinden, hebben last van *popularity bias* (Fernández et al., 2021). Dat wil zeggen: reeds populaire of virale inhoud wordt nóg populairder. Computerwetenschappers noemen dit ook wel het the-rich-get-richer-effect (Bellogín et al., 2017).

3.6.3 Hyperactieve gebruikers krijgen meer relatieve amplificatie

Onderzoek door Papakyriakopoulos, Serrano en Hegelich laat zien dat aanbevelingsalgoritmen beïnvloed kunnen worden door hyperactieve gebruikers (Papakyriakopoulos et al., 2020a). Ze baseren hun onderzoek op politieke discussies in Facebookgroepen in 2016 en tonen hoe gebruikers die overmatig veel inhoud delen, relatief meer engagement krijgen per post dan andere gebruikers en ook overmatige invloed lijken te hebben op aanbevelingen door aanbevelingsalgoritmen. Dat komt doordat de inhoud van hyperactieve accounts overmatig kan voorkomen in de trainingsdata van aanbevelingsalgoritmen, waarmee zij ook weer toekomstige aanbevelingen voor andere gebruikers kunnen beïnvloeden. De auteurs wijzen dan ook op het ingebakken risico van rangschikking met *collaborative filtering* en met neurale netwerken.[10]

3.6.4 Rechtse politieke inhoud lijkt meer relatieve amplificatie te krijgen

Wetenschappers proberen te onderzoeken of aanbevelingsalgoritmen van platformen een bias hebben wanneer ze gebruikers politieke inhoud aanraden. Uit verschillende onderzoeken naar platform X (voorheen Twitter) blijkt dat inhoud van rechtse politici meer relatieve amplificatie krijgen. Wetenschappers Ye, Luceri en

10 In machinelearning wordt dit een *unbalanced learning problem* genoemd.

Ferrara onderzochten met 120 voor het onderzoek aangemaakte accounts de aanbevolen inhoud op X tijdens de presidentsverkiezingen in de Verenigde Staten in 2024, en vonden dat nieuwe accounts relatief vaker inhoud van *rechtsgelieerde* accounts aanbevolen kregen (Ye et al., 2025). Ze deelden de accounts in vier groepen: een groep accounts die niemand volgde, een groep die rechtse media volgde, een groep die linkse media volgde en een groep die een mix van rechtse en linkse media volgde. De accounts die niemand volgde, kregen relatief meer rechtse inhoud aanbevolen.

Ook econoom en politicoloog Germain Gauthier en zijn collega's concluderen recent dat het aanbevelingsalgoritme van X meer conservatieve inhoud aanraadt dan liberale inhoud. Ze vergeleken het met de chronologische tijdlijn van gebruikers die meededen aan het onderzoek (Gauthier et al., 2026).

De bevindingen komen overeen met een iets ouder onderzoek waarin twee miljoen Twittergebruikers uit zeven landen werden onderzocht op een bias in relatieve amplificatie van politieke inhoud (Huszár et al., 2022a). Huszár et al concludeerden dat in zes van de zeven landen mainstream-political-right-inhoud meer relatieve amplificatie kreeg door het aanbevelingsalgoritme.

In een preprint onderzoek (nog niet door andere wetenschappers gecontroleerd) onderzochten Duitse wetenschappers van de Universiteit van Potsdam meer dan 500.000 TikTok-video's die aan 78 voor het onderzoek aangemaakte accounts werden aanbevolen tijdens Duitse regionale verkiezingen in 2024 en parlementsverkiezingen van 2025 (Tjaden et al., 2025). Van de accounts zonder politieke voorkeur was de kans het grootst dat ze video's van de AfD werden aangeraden.

Ook Paul Bouchaud deed onderzoek naar de relatieve amplificatie van Europarlementariërs op Twitter. Hij vond dat rechtse content over het algemeen meer aanbevolen werd op Twitter, maar dat wanneer gecorrigeerd werd voor de politieke voorkeur van gebruikers, dit effect wegviel (Bouchaud, 2024).

3.6.5 Specifieke gebruikers kunnen meer relatieve amplificatie krijgen

Ontwerpkeuzes over aanbevelingsalgoritmen gaan deels over welke scores bepaalde inhoud toebedeeld krijgen ten opzichte van andere berichten. De keuze om berichten anders te wegen, kan ook voor individuele accounts worden gemaakt. Accounts met veel volgers krijgen dan een streepje voor op accounts met minder volgers. Elon Musk zou in 2023 deze gewichten persoonlijk hebben laten

aanpassen voor zijn eigen berichten. Uit interne memo's bleek dat hij dit deed nadat een bericht van Joe Biden meer interactie kreeg dan een bericht van hemzelf. Hij liet, naar verluidt, zijn medewerkers het algoritme zo aanpassen dat zijn berichten duizend keer meer te zien waren dan die van andere gebruikers (Schiffer, 2023).

Gebruikers speculeren op social media regelmatig over de vraag of hun inhoud anders wordt gewogen dan die van anderen. Ze beschuldigen het platform dan bijvoorbeeld van *shadowbanning*: het minder zichtbaar maken van content op social media zonder deze direct te verwijderen. Het platform houdt de inhoud beschikbaar, maar toont het niet meer in de tijdlijn van mensen. In een rechtszaak die privacy-activist Danny Mekić aanspande tegen X gaf de Amsterdamse rechtbank hem in 2024 gelijk: X had zijn berichten niet mogen weglaten uit de tijdlijn van anderen, zonder hem die beslissing uit te leggen (*10767307 CV FORM 23-13934*, 2024).

3.7 Alternatieven voor engagement

Het aanpassen van op engagement gebaseerde aanbevelingsalgoritmen om schadelijke effecten te beperken, kan op gespannen voet staan met andere bedrijfsbelangen, zoals het maximaliseren van engagement en gebruikersretentie. Desondanks worden in wetenschappelijk onderzoek en in rapporten van ngo's diverse voorstellen gedaan om op engagement gebaseerde aanbevelingsalgoritmen aan te passen om negatieve effecten te beperken. Wetenschappers doen bijvoorbeeld voorstellen voor bridging-based-rangschikking in plaats van rangschikking gebaseerd op engagement (Lasser & Poechhacker, 2025). Dat is een vorm van rangschikking gebaseerd op het bouwen van bruggen tussen mensen. In plaats van gebruikerssignalen rondom engagement mee te nemen in aanbevelingen, suggereert internetonderzoeker Aviv Ovadya om inhoud aan te bevelen aan mensen die juist *niet* op elkaar lijken (Ovadya, 2022). Inhoud die bruggen slaat tussen groepen mensen krijgt dan meer relatieve amplificatie door aanbevelingsalgoritmen.

Recent onderzoek op X laat zien dat aanpassingen aan de manieren van rangschikken ook *binnen* op engagement gebaseerde aanbevelingsalgoritmen positieve invloed kunnen hebben. Onderzoekers van Stanford University en andere Amerikaanse universiteiten lieten in 2025 zien dat het veranderen van de rangschikking van tweets, door verdeeldheid zaaiende tweets naar onderen in iemands tijdlijn te verplaatsen, zorgde voor mildere gevoelens ten opzichte van politieke tegenstanders (Piccardi et al., 2025). Deze ingreep werkte dus zonder inhoud te verwijderen, maar puur door de rangschikking te veranderen.

De onderzoeken merken wel op dat deze ingreep zorgde voor een afname (vijf minuten) van de tijd die mensen spendeerden op X. Dat zou dus haaks staan op het verdienmodel van platformen om de tijd die gebruikers doorbrengen te maximaliseren voor zo hoog mogelijke advertentie-inkomsten.

3.8 Neutrale aanbevelingsalgoritmen bestaan niet

In de voorgaande paragrafen hebben we laten zien hoeveel ontwerpkeuzes achter aanbevelingsalgoritmes schuilgaan. Zelfs de keuze voor een erg simpel sorteer- en rangschiksysteem, de chronologische feed, is geen neutrale ontwerpkeuze. Chronologische feeds die alle inhoud die je ziet, sorteren op basis van (1) accounts die je volgt en (2) rangschikken op basis van tijdstip (nieuw naar oud), zorgen voor een oververtegenwoordiging van hele actieve accounts. Volg je veel nieuwsmedia die meerdere keren per uur posten, maar ook een aantal vrienden die slechts 1 keer per week iets delen? Grote kans dat de inhoud van je vrienden maar sporadisch in je tijdlijn tegenkomt. Dat zou je kunnen oplossen door content van vrienden een ander gewicht te geven dan de content van nieuwsmedia, maar ook dat is een ontwerpkeuze.

Je kunt dus ook niet spreken van een gewoon of neutraal algoritme en een gemanipuleerde variant. Natuurlijk is het wel mogelijk voor platformen, en in mindere mate voor onderzoekers, om met behulp van A/B-tests te kijken wat de invloed is van bepaalde keuzes op wat mensen te zien krijgen. Daarom is het belangrijk om de ontwerpkeuzes bloot te leggen achter de manieren waarop mensen online inhoud zien. Een chronologische tijdlijn is in ieder geval voor gebruikers beter te doorgronden. Dat kan hun controle over wat ze te zien krijgen vergroten.

3.9 Conclusie

Omdat op elke tijdlijn maar beperkt plek is, moeten platformontwerpers keuzes maken over *welke* berichten ze in je tijdlijn laten zien. Ook maken ze keuzes over hoe die inhoud wordt gerangschikt. Dat kan op basis van wie je volgt, of wat mensen in je netwerk delen, maar ook op basis van aanbevelingsalgoritmen.

Aanbevelingsalgoritmen zijn een verzameling aan systemen die sorteren en rangschikken op basis van data, met als doel om mogelijk interessante inhoud aan te raden aan gebruikers. Aanbevelingsalgoritmen sorteren en rangschikken inhoud op basis van voorspellingen. Tegenwoordig zijn dat vaak voorspellingen op basis van engagement. Bijvoorbeeld: wat is de kans dat jij deze post liket, deze video

afkijkt of dit bericht doorstuurt naar een vriend? Aanbevelingsalgoritmen zijn een antwoord op het probleem van informatieoverload op het internet: uit al de beschikbare inhoud wordt geprobeerd jou iets te laten zien wat je leuk, interessant of anderszins relevant vindt.

Platformen maken veel verschillende keuzes in hoe ze gefilterde inhoud rangschikken. Zelfs de keuze voor een erg simpel sorteer- en rangschikkingsysteem, de chronologische feed, is geen neutrale ontwerpkeuze. Bij aanbevelingsalgoritmen gebaseerd op engagement maken ze keuzes over welke gedragssignalen ze meewegen in hun voorspellingen, en welk gewicht ze die signalen geven. Ook de manier waarop ze contentmoderatie doen en hoe ze bepalen wat de *kwaliteit* is van bepaalde inhoud, valt onder de keuzes die platformen maken. Platformen testen ook voortdurend de werking en effectiviteit van hun aanbevelingsalgoritmen. Dat doen ze onder andere om te kijken hoe ze het engagement van mensen kunnen vergroten.

De verandering naar aanbevelingsalgoritmen gebaseerd op engagement is ingegeven door het feit dat platformen ontdekten dat ze daarmee de gebruikersretentie konden verhogen.

Aanbevelingsalgoritmen kunnen bepaalde inhoud *relatief amplificeren:* dat wil zeggen meer zichtbaarheid geven ten opzichte van andere manieren van rangschikken. Aanbevelingsalgoritmen gebaseerd op engagement kunnen zorgen voor meer relatieve amplificatie van onder andere emotionele inhoud, populaire inhoud, hyperactieve gebruikers en rechtse politieke inhoud. Aanbevelingsalgoritmen zouden ook op basis van andere signalen kunnen werken dan engagement, en bijvoorbeeld inhoud belonen die bruggen slaat tussen groepen mensen.

Als je begrijpt hoe aanbevelingsalgoritmen op grote socialmediaplatformen werken, is het mogelijk om hierop in te spelen en daarmee te proberen om bepaalde inhoud een groter bereik te geven. In het volgende hoofdstuk gaan we in op de mogelijkheden voor inmengers om daarop in te spelen.

4 Instrumentarium voor inmengingsactiviteiten

In hoofdstuk 3 hebben we de werking van het aanbevelingsalgoritme beschreven. In hoofdstuk 4 verplaatsen we het perspectief naar buitenlandse actoren. Hoe spelen die in op aanbevelingsalgoritmen om vanuit het buitenland bewust te misleiden en invloed uit te oefenen op het online publieke debat?

We beschrijven hieronder het instrumentarium om bepaalde inhoud meer of minder zichtbaar te maken. Deze instrumenten zijn voor iedereen toegankelijk. Het is ook mogelijk om inmengingsactiviteiten uit het buitenland te bestellen (zie paragraaf 4.2.4 en 4.2.5). Overigens is het motief van mensen die gebruikmaken van het instrumentarium niet altijd gericht op het beïnvloeden van verkiezingen. Bewuste misleiding vanuit het buitenland kan ook een middel zijn om geld te verdienen.[11] Ook bedienen sociale bewegingen en activistengroepen zich van dit instrumentarium omdat ze als doel hebben inhoud meer of minder zichtbaar te maken (Treré & Bonini, 2024).[12]

Het instrumentarium verandert continu. Platformbedrijven voeren dagelijks aanpassingen door aan hun platformen en lanceren daarbij ook nieuwe diensten. Kwaadwillende (buitenlandse) actoren kunnen hierop inspelen. Met nieuwe diensten creëren platformen ook nieuwe manieren om invloed uit te oefenen op het online publieke debat. Voorbeelden zijn nieuwe diensten zoals livestreaming en de mogelijkheid om geld te verdienen met een socialmedia-account.

Ook passen platformen de mogelijkheden op het platform aan. Bijvoorbeeld door het veranderen van de regels voor contentmoderatie. Een bekend voorbeeld van een dergelijke aanpassing is de boodschap van Mark Zuckerberg vlak na de inauguratie van de Amerikaanse president Donald Trump, dat contentmoderatieregels rondom migratie en gender veranderden om meer te voldoen aan 'mainstream discourse' (Hendrix, 2025). Deze aanpassing van de contentmoderatieregels werd publiekelijk gedeeld. Maar hoe platformen intern contentmoderatieregels aanscherpen of versoepelen, wordt vaak niet precies

11 Het online nieuwsmedium 404 Media meldde in 2025 dat verschillende accounts die zich voordeden als Amerikaanse burgers en zich mengden in het online politieke debat in de Verenigde Staten waarschijnlijk actief waren vanuit andere landen om hiermee geld te verdienen. Het nieuwsmedium refereerde naar YouTube-video's waarin mensen uitleggen hoe je geld kan verdienen met nepaccounts. (Koebler, 2025).

12 De onderzoekers geven een overzicht van netwerken van activisten en sociale bewegingen die gecoördineerd posten met als doel dat het aanbevelingsalgoritme bepaalde inhoud meer of juist minder zichtbaar maakt.

bekend gemaakt omdat platformen willen voorkomen dat kwaadwillende weten hoe ze moderatieregels moeten omzeilen (Jiang et al., 2023).

Platformen passen zich ook aan om aan wetgeving te voldoen. Hoe aanpassingen uitwerken in de praktijk is afhankelijk van hoe gebruikers omgaan met wat platformen aan mogelijkheden bieden. Instagram geeft bijvoorbeeld de mogelijkheid om aan te geven dat inhoud die wordt geüpload met AI genereerd is. Dit gebeurt lang niet voor alle met AI gegenereerde inhoud. Uit onderzoek van CampAItracker bleek bijvoorbeeld dat Nederlandse politieke partijen veel inhoud hebben verspreid zonder te melden dat het om met AI-gegeneerde berichten ging (Simon Kruschinski & Fabio Votta, 2025).

Binnen deze dynamische platformomgevingen gebruiken kwaadwillende (buitenlandse) actoren tal van tactieken tegelijk. Een illustratief voorbeeld is te vinden op de website Like.vn dat bekend werd door de veelvoud van likes die het gaf aan inhoud van het account van Frans Timmermans (zie paragraaf 4.2.5). Op de website wordt amplificatie aangeboden met erbij een melding dat het zinvol is om een 'combinatiepakket van engagement te bestellen, waaronder postlikes, accountvolgers en commentaar op verzoek'. Het beste effect zou volgens de aanbieder bereikt worden als deze 'engagement boosting' gecombineerd wordt met 'natuurlijke engagementsignalen' (like.vn, 2026).

In deze dynamische omgeving geven onafhankelijke onderzoekers zoals onderzoeksjournalisten, academici en ngo's een beeld van wat er gebeurt op socialmediaplatformen aan de hand van openbare bronnen. Gekeken kan worden naar het initiatief DISARM. Dit biedt een uitgebreide, opensourcebeschrijving van manieren van inmenging met voorbeelden uit de praktijk (DISARMFoundation, z.d.).

Bovendien zijn in onderzoeken uitgevoerd ten tijde van de Tweede Kamerverkiezingen van 2025 verschillende inmengingsactiviteiten in Nederland waargenomen (HEIO Consortium et al., 2026; Jasper Bunskoek et al., 2025; Justice for Prosperity, 2025).

De inmengingsactiviteiten die we hierna beschrijven, verdienen de aandacht van politici en beleidsmakers die de mogelijkheden van inmenging willen doorgronden en schadelijke gevolgen willen verminderen. Paragraaf 4.1 gaat in op de soorten inhoud. Paragraaf 4.2 behandelt de verschillende technieken. Paragraaf 4.3 vormt de conclusie van dit hoofdstuk.

4.1 Inhoudskeuze

Mensen die actief deelnemen aan het politieke debat op socialmediaplatformen maken een keuze over welke inhoud ze verspreiden, en welke inhoud van andere gebruikers ze willen versterken. Kwaadwillende (buitenlandse) actoren moeten dus allereerst een keuze maken over welke inhoud ze willen verspreiden. We onderscheiden drie type inhoud binnen het publieke debat op socialmediaplatformen: rechtmatige inhoud (meer in paragraaf 4.1.1), borderline-inhoud (4.1.2) en onrechtmatige inhoud (4.1.3).

Een voorbeeld van rechtmatige inhoud is bericht over op wie iemand stemt of over politiek maatschappelijk onderwerpen. Maar ook politici of politieke partijen die campagne voeren. Borderline-inhoud is inhoud die grenst aan wat niet is toegestaan op socialmediaplatformen. Van deze inhoud is bekend dat het meer aandacht krijgt waardoor aanbevelingsalgoritmen gebaseerd op engagement deze inhoud relatief meer amplificeren dan andere inhoud (zie ook paragraaf 3.6).

Dan is er nog inhoud die in strijd is met het recht: onrechtmatige en strafbare (of illegale) inhoud. Onrechtmatige inhoud is informatie, door mensen op het internet geplaatst, die in strijd is met het recht, vanwege de schadelijke gevolgen ervan en/of omdat de belangen van anderen daardoor op ernstige wijze worden aangetast (Hoboken et al., 2020, p. 11). Onrechtmatig is een term uit het civiele recht en gaat over geschillen tussen burgers en bedrijven onderling of over geschillen waarbij de overheid zich bedrijfsmatig opstelt. Strafbare inhoud gaat over iets dat niet is toegestaan op basis van regels van Wetboek van Strafrecht, zoals zware bedreigingen met een strafbaar feit zoals moord of zware mishandeling.

Kwaadwillende (buitenlandse) actoren kunnen ervoor kiezen om een van de drie typen inhoud zelf te maken of bestaande inhoud van andere gebruikers te versterken.

4.1.1 Rechtmatige inhoud

Op socialmediaplatformen vindt politiek debat plaats. Gebruikers verspreiden er inhoud over politici of politieke denkbeelden. In aanloop naar de Roemeense verkiezingen werden er bijvoorbeeld pro-Georgescu-inhoud verspreid: beelden van televisieoptredens van de presidentskandidaat.

Inmengingsactiviteiten kunnen eruit bestaan deel te nemen in dit politieke debat door berichten te verspreiden die qua inhoud niet in strijd zijn met community-richtlijnen of met wet- en regelgeving. We spreken van inmenging wanneer een

buitenlandse actor zich voordoet als een Nederlander en rechtmatige inhoud verspreid. Op die manier probeert de actor bewust te misleiden (zie ook paragraaf 4.2.3 Nepaccounts).

4.1.2 Borderline-inhoud

Borderline is een term die gebruikt wordt om inhoud te beschrijven die in het grensgebied valt van wat socialmediabedrijven wel en niet toestaan op hun platform.[13] De term borderline is geen begrensde definitie, maar een parapluterm voor verschillende typen inhoud. Bij borderline-inhoud kan worden gedacht aan: misinformatie en desinformatie, seksueel suggestieve inhoud, bloederige beelden, inhoud die verkiezingen delegitimeert, inhoud die aanzet tot zelfbeschadiging, haat verspreidt of eetstoornissen promoot (Macdonald & Vaughan, 2024, p. 351). Hieronder beschrijven we enkele voorbeelden van bordeline-inhoud die direct inspelen op het misleiden.

Kloonwebsites
Makers van een kloonwebsite zetten een webadres op met een titel die sterk lijkt op de titel van een bestaande organisatie. Zo maken ze gebruik van de bekendheid en reputatie van de organisatie. Vanuit de naam van deze titels verspreiden ze met socialmedia-accounts berichten van de kloonwebsite. Zo proberen de accounts bezoekers te verleiden het socialmediaplatform te verlaten om op de kloonwebsites terecht te komen. Hierop staan bijvoorbeeld politiek gekleurde of misleidende berichten.

Op Facebook en Twitter zijn in 2022 via botnetwerken kloonwebsites van bekende mediatitels uit Duitsland, Italië, Verenigde Koninkrijk en Spanje verspreid. Een van de doelwitten was het Duitse dagblad Bild. Het webadres van de kloonpagina begon wel met Bild, maar had een toevoeging, zoals bild.vip ('Under the Hood of a Doppelgänger – Qurium Media Foundation', 2022).

Meta beschreef in het Adversarial Threat Report 2025 activiteiten van accounts die Facebookpagina's oprichtten waarop zich voordeden als lokale mediatitels. Vanaf deze pagina's werden advertenties ingekocht gericht aan verschillende Afrikaanse landen. Meta wijst erop dat deze pagina's vermoedelijk centrale instructies kregen omdat verschillende pagina's precies, of nagenoeg dezelfde boodschap herhaalden (Meta, 2025a).

13 YouTube omschreef *borderline content* in een blogpost over contentmoderatie als: '*videos that don't quite cross the line of our policies for removal but that we don't necessarily want to recommend to people*'. Meta omschrijft borderline content als inhoud die niet verboden is door gebruikersvoorwaarden, maar wel dicht in de buurt komt. (Macdonald en Vaughan, 2024, p. 350).

Desinformatie
Naast misleiding over de identiteit van een website kan inhoud ook misleiden. Dit wordt vaak omschreven als mis- en desinformatie.[14] In hoofdstuk 2 gaven we hiervan een recent Nederlands voorbeeld: een socialmediabericht waarin werd gesteld dat kiezers die voor PvdA-GroenLinks wilden stemmen twee vakjes moesten inkleuren op het stembiljet (wat de stem ongeldig verklaart). (*Kamerstukken II, 35 165, nr. 102*, 2026)

Een ander recent beschreven voorbeeld is een socialmediabericht van een gebruiker die suggereerde dat er verkiezingsfraude is gepleegd tijdens de Tweekamerverkiezingen van 2025. Dit bericht werd gedeeld door nepaccounts die zich voordeden als Nederlanders maar die vanuit het buitenland werden beheerd (zie ook hoofdstuk 2: Casus: Nepaccounts uit buitenland op X tijdens Nederlandse verkiezingen).

4.1.3 Onrechtmatige inhoud

Politici, en dan met name vrouwelijke politici, krijgen steeds meer te maken met online haat en bedreiging (Karlijn Saris & Coen van de Venm, 2021). Door gerichte haatcampagnes kunnen commentaarsecties op socialmediaplatformen ingezet worden om een bepaald narratief over een kandidaat te verspreiden. Gerichte haatcampagnes kunnen (potentiële) kandidaten ervan weerhouden om mee te doen aan verkiezingen en ze kunnen potentiële kiezers weerhouden om op die kandidaat te stemmen. Het verspreiden van haat kan dus een effectieve manier zijn om invloed uit te oefenen op de deelname van politici en het verloop van de verkiezingen (Honingh & Van Ham, 2024, p. 161).

In 2023 is een lastercampagne vanuit De Verenigde Arabische Emiraten georganiseerd die gericht was op Nederlandse publieke figuren. (Heck & Kouwenhoven, 2023) Hierbij werden nepaccounts (zie ook paragraaf 4.2.2. Nepaccounts) gebruikt om Nederlanders te beschuldigen van het aanhangen van het Moslimbroederschap. Voormalig GroenLinks-Kamerlid Kauthar Bouchallikht en de Arnhemse burgemeester Ahmed Marcouch waren onder andere doelwitten van deze campagne.

14 Over de misinformatie en desinformatie bestaan verschillende publieke en wetenschappelijke definities (Egelhofer et al., 2020). De uitleg van desinformatie die zowel in academische literatuur als in het Nederlandse publieke debat recent meestal wordt gegeven is: de verspreiding van grotendeels of geheel (wetenschappelijk) ongefundeerde claims waarvan de verzender *weet* dat ze niet kloppen. Het doel van het maken en verspreiden kan zijn om geld te verdienen, voor vermaak, om het democratisch proces te verstoren of een combinatie hiervan.(Humprecht et al., 2020). Bij misinformatie geldt het criterium over de intentie van de verzender niet. (Wardle & Derakhshan, 2017)

Kader 3 Met AI gegenereerd tekst- en beeldmateriaal

Met generatieve AI kunnen realistisch uitziende video's worden gemaakt op basis van tekstinstructies (zie ook het rapport *Rathenau Scan Generatieve AI* (Rathenau Instituut, 2023).

Het is een nieuw type inhoud dat aangeboden wordt op verschillende socialmediaplatformen en dat op die platformen gegenereerd kan worden. Het gegenereerde materiaal kan qua inhoud rechtmatig, borderline of onrechtmatig zijn. Het kan ook als techniek worden ingezet om een groot publiek te bereiken. Socialmediaplatformen kunnen mogelijk vanuit bedrijfsoverwegingen generatieve-AI-toepassingen op het platform promoten. Aanbevelingsalgoritmen zouden deze inhoud dan aan meer gebruikers aanraden (Simon Kruschinski & Fabio Votta, 2025).

De gegenereerde foto's en video's kunnen worden ingezet om een bepaald narratief te versterken of mensen te overtuigen van de echtheid van bepaalde beelden. Het beeldmateriaal schetst een politiek droombeeld of doembeeld van de werkelijkheid. Een recent beschreven voorbeeld van met AI gegeneerd materiaal is een afbeelding waarop toenmalig partijleider Frans Timmermans van GroenLinks-PvdA werd voorgesteld met handboeien om (Feenstra & Sabel, 2025).

In aanloop naar de verkiezingen van 2025 is door onderzoekers bijgehouden welke door AI gegenereerde inhoud er werd verspreid op Facebook, Instagram, TikTok en X door Nederlandse politici, politieke partijen en geselecteerde influencers en commentatoren. Dit laat zien dat maar een deel van dit soort inhoud wordt gelabeld als gegeneerd met AI, al is dit wel verplicht volgens platformenrichtlijnen (Simon Kruschinski & Fabio Votta, 2025).

4.2 Technieken die zichtbaarheid vergroten of verkleinen

Invloed uitoefenen via socialmediaplatformen gebeurt niet alleen door een bepaald type inhoud te uploaden. De inhoud, door een kwaadwillende buitenlandse actor gemaakt of afkomstig van andere gebruikers, wordt ook verspreid. Het doel van de verspreiding is om zoveel mogelijk mensen of juist een specifieke doelgroep te bereiken. Door bij inhoud engagementsignalen af te geven, zoals inhoud liken,

delen, becommentariëren of bekijken, is het mogelijk dat het aanbevelingsalgoritme deze inhoud verder verspreid. Ook kan het gevolg zijn dat het aantal likes, of aantal keer delen de indruk wekt dat de inhoud populairder is dan als het minder likes had.

Platformen bieden verschillende mogelijkheden om inhoud te verspreiden. Zoals uitgelegd in hoofdstuk 2 spelen aanbevelingsalgoritmen een cruciale rol bij verspreiding, De volgende paragrafen gaan er dieper op in.

4.2.1 Openbaar posten en posten in groepen

Inhoud verspreiden wordt op de meeste platformen omschreven als posten. Dit kan openbaar waardoor berichten voor iedereen zichtbaar worden, maar ook binnen (besloten) groepen. In dit laatste geval krijgen alleen mensen die toegelaten zijn tot een groep of mensen die zichzelf hebben aangemeld voor een groep berichten te zien. Onder andere Facebook, Telegram en Whatsapp bieden groeps-chats aan. Mensen kunnen zich inschrijven of uitgenodigd worden voor een groep. Gebruikers die de groep beheren, kunnen zelf bepalen of het bestaan van de groep kenbaar wordt gemaakt aan alle gebruikers. Als de groep openbaar zichtbaar is, speelt het aanbevelingsalgoritme een rol omdat de groep wordt aangeraden aan gebruikers die nog niet lid zijn.

Facebookgroepen zijn een voorbeeld van een online omgeving waarbinnen berichten alleen aan groepsleden worden gestuurd. De met AI gegenereerde inhoud over Frans Timmermans werd bijvoorbeeld als eerste gedeeld in een openbaar toegankelijke Facebookgroep. Deze groep was openbaar vindbaar en had volgens de Volkskrant in juni 2025 130.000 volgers (Feenstra & Sabel, 2025).

Inmenging kan gebeuren door dit soort groepen op te zetten of eraan deel te nemen. Groepen kunnen geschikt zijn om specifieke gebruikers te bereiken, zoals bijvoorbeeld mensen die geïnteresseerd zijn in een politieke partij.

4.2.2 Streamen

Streaming is een directe vorm van communicatie via een platform waarbij een individuele gebruiker een live-videoverbinding heeft met meerdere andere gebruikers. Vergeleken met posten is deze vorm van platformcommunicatie een directe en ongefilterde manier om groepen gebruikers te bereiken. Grote platformen die streaming aanbieden zijn onder meer: TikTok, YouTube, Facebook, Instagram en Twitch.

Aanbevelingsalgoritmen spelen bij streaming een rol omdat verschillende platformen livestreams een prominente plek geven. De platformen promoten hiermee deze vorm van communicatie.

Actoren die invloed uitoefen via livestreaming kunnen handig gebruikmaken van bedrijfsregels om livestreams te promoten. Contentmoderatie in livestreams kan minder goed dan bij posten. Omdat een stream een liveverbinding heeft met gebruikers kan de inhoud niet worden beoordeeld voordat het bij het publiek terechtkomt. Het platform kan er alleen voor kiezen om een stream offline te halen.

Invloed uitoefenen op livestreams kan door reacties achter te laten in de streams of door specifieke gebruikers te belonen met een gift. Deze beloningsmechanismen kunnen een markt creëren om bepaalde type inhoud in de stream lucratief te maken (HEIO Consortium et al., 2026, p. 52; *Turning Passion to Profit: WAYS TO MAKE MONEY ON TIKTOK LIVE*, z.d.)

4.2.3 Nepaccounts

Nepaccounts worden ook wel sockpuppet-accounts genoemd, een verwijzing naar een sokpop of handpop. Nep-accounts worden aangemaakt om een verzonnen of gestolen identiteit aan te nemen (Rathenau Instituut, 2021).

Sockpuppets kunnen een negatief of positief sentiment creëren over politieke partijen, politici of politiekmaatschappelijke kwesties.

Activiteiten van dit soort accounts worden al sinds de opkomst van het internet waargenomen. Rond de Tweede Kamerverkiezingen zijn nepaccounts op TikTok en LinkedIn waargenomen die zich voordeden als het account van PVV-leider Geert Wilders (HEIO Consortium et al., 2026, p. 39). Journalisten van RTL Nieuws beschreven 550 nepaccounts die actief waren vanuit het buitenland en zich mengde in het Nederlandstalige berichtenverkeer op X en Facebook (zie hierover ook hoofdstuk 2: Casus: Nepaccounts uit buitenland op X tijdens Nederlandse verkiezingen).[15]

De verspreiding kan verder reiken dan de platformen zelf. Berichten door sockpuppet-accounts werden bijvoorbeeld in het verleden overgenomen door reguliere media, terwijl het om accounts met verzonnen namen ging die werden

15 In 2025 werd door X (voormalig Twitter) voor veel accounts aangegeven vanuit welke IP-locatie en Appstore deze accounts actief zijn (Nikita Bier [@nikitabier], 2025).

beheerd vanuit Rusland. (*Alle grote media in Noorwegen trappen in tweets van Russische trollen*, 2020).

4.2.4 Gecoördineerd posten

Een vaak beschreven manier van pogingen om online debatten te beïnvloeden, is het gecoördineerd posten van inhoud. Bij het gecoördineerd posten werken accounts samen om specifieke inhoud op een afgesproken moment te verspreiden. Deze gecoördineerde campagnes proberen om veel engagement te genereren en viraliteit te creëren met als doel dat het lijkt of de inhoud populair is (Rogers & Righetti, 2025). Recent lanceerde Israëlische bedrijven twee applicaties waarmee gebruikers in een handomdraai pro-Israëlische inhouden konden verspreiden en socialmediaplatformen kunnen overspoelen met klachten over inhoud die ongunstig is voor Israël. Met de app kon het zo worden uitgevoerd dat het inhoud leek te zijn die gebruikers zelf spontaan hadden gemaakt (Beemsterboer, 2026).

Deze manier van inmenging wordt internationaal waargenomen en is beschreven in academisch onderzoek. Politiekecommunicatiewetenschapper Thiele en collega's (Thiele et al., 2025) geven een overzicht van voorbeelden van gecoördineerd posten waarnaar uitgebreid onderzoek is gedaan. Thiele stelt dat campagnes variëren in grootschaligheid, mate van verdekt opereren en uitwerking van de inhoud. Campagnes die door de onderzoekers zijn bestempeld als inmenging variëren in schaal. Bij grootschalige inmenging wordt met behulp van veel verschillende accounts inhoud geüpload gedurende een lange periode op meerdere platformen tegelijk. Dit soort grootschalige inmengingscampagnes kunnen worden ingekocht bij aanbieders zoals het bedrijf Internet Research Agency (Thiele et al., 2025, p. 14).

Thiele en collega's beschreven ook campagnes gericht op een specifiek publiek waarbij elke keer precies hetzelfde bericht werd gekopieerd en geüpload door verschillende accounts. Een vergelijkbare poging tot beïnvloeding is recent ook waargenomen in Nederland. Verschillende accounts hebben in korte tijd identieke Nederlandstalige berichten verspreid. Deze berichten verwezen allemaal naar hetzelfde webadres waaruit af te leiden viel dat de accounts vermoedelijk werden beheerd door één individu. Deze campagne werd overigens door de onderzoekers niet toegeschreven aan een buitenlandse actor, maar aan een Nederlands account (Justice for Prosperity, 2025).

Recent lanceerde Israëlische bedrijven twee applicaties waarmee gebruikers in een handomdraai pro-Israëlische inhouden konden verspreiden en socialmediaplatformen kunnen overspoelen met klachten over inhoud die

ongunstige is voor Israël. Met de app kom het zo worden uitgevoerd dat het inhoud leek te zijn die gebruikers zelf spontaan hadden gemaakt (Beemsterboer, 2026).

4.2.5 Bot-accounts en hyperactiviteit

In de literatuur zijn verschillende beïnvloedingscampagnes beschreven waarbij duizenden botaccounts (botnetwerken) zijn ingezet om berichten met politieke inhoud meer bereik te geven. Thiele en collega's (2025) geven een overzicht van onderzoeken waarbij inmenging met gecoördineerde botnetwerken een rol hebben gespeeld.

Een vaak aangehaald voorbeeld is een vanuit Rusland aangestuurd botnetwerk dat actief was op Facebook, Reddit en Twitter in aanloop naar de presidentsverkiezingen in de Verenigde Staten in 2016. Met duizend accounts zijn over een lange periode veel verschillende berichten verspreid op verschillende platformen, waaronder 10,4 miljoen Tweets en 116.000 Instagramberichten. In de tweets werd ingegaan op politiek nieuws waarop weer gereageerd werd door andere botaccounts (Thiele et al., 2025).

Activiteiten van hyperactieve accounts zijn: delen, liken en reageren op bestaande inhoud. Het doel is om specifieke accounts of inhoud waarop gereageerd wordt meer zichtbaar te maken op platformen. Dat kan bijvoorbeeld op directe wijze door op berichten van een account van politicus te reageren. Maar uit onderzoek van RTL Nieuws blijkt dat het ook op indirecte wijze gebeurde: door te reageren op accounts die zich bezighouden met politieke issues of politieke partijen, maar zelf niet van een politieke partij zijn. Dit laatste gebeurt vermoedelijk om minder zichtbaar te werk te gaan (Jasper Bunskoek et al., 2025).

Zoals beschreven in paragraaf 3.6.3 blijkt uit onderzoek dat hyperactieve accounts succesvol kunnen zijn omdat de aanbevelingsalgoritmen activiteit belonen. Dit kan leiden tot een overrepresentatie van een bepaalt type content. Individuen of groepen georganiseerde gebruikers kunnen hierop inspelen met een hyperactief account.

Engagement op bestelling
In aanloop naar de verkiezing werd een netwerk gezien dat een groot aantal likes deed bij berichten op X van Frans Timmermans (Justice for Prosperity, 2025, p. 17). De accounts die likes achterlieten onder de berichten hadden een duidelijk signatuur in de gebruikersnaam waardoor kon worden herleid dat ze afkomstig waren van een Vietnamese aanbieder.

Op de website van de Vietnamese aanbieder valt te lezen dat er campagnes besteld kunnen worden voor socialmediaplatformen: Facebook, TikTok, Instagram, YouTube, X, Threads en LinkedIn: *'We provide packages to increase likes, follows, views, and comments on social media platforms such as Facebook, TikTok, Instagram, YouTube, etc., through an automated system. You simply need to send the link you wish to boost engagement for, and the system will process it according to your requirements within the agreed timeframe.'* (Eigen vertaling, like.vn, z.d.)[16]

4.2.6 Hashtagboosting

Veel platformen werken met hashtags. Het doel van het gebruik van hashtags is dat inhoud van berichten gelabeld wordt. Wanneer gebruikers bijvoorbeeld #tweedekamerverkiezingen aan inhoud toevoegen, geven ze aan dat het bericht gaat over dit onderwerp. Andere gebruikers kunnen door op de hashtag te klikken een overzicht krijgen van inhoud met hetzelfde label. Veel platformen geven een overzicht van de meest populair hashtags op dat moment. Inhoud met deze hashtags krijgt hierdoor een groter bereik dan minder populaire hashtags.

Tijdens de Roemeens verkiezingen is waargenomen hoe op populaire hashtags werd meegelift om inhoud van een bepaalde kandidaat meer bereik te geven.

Gecoördineerd inhoud verspreiden kan andere inhoud overstemmen en daardoor minder zichtbaar maken. Treré en collega's noemen dit *flooding*. Ze illustreren dit met een voorbeeld van gebruikers die zichzelf verenigen als K-pop-fans. Deze fanbeweging uploadde videoclips van K-pop-groepen en *memes* waarbij ze op dat moment populaire hashtags #MAGA, en #whitelivesmatter gebruikten. Het doel van het uploaden was het overstemmen van racistische en haatzaaiende berichten die ook met deze hashtags werden geüpload.

De auteurs geven ook een voorbeeld van *flooding* waarbij onduidelijk was wie er achter de activiteiten schuilgingen. In Mexico kwam in 2014 een sociale beweging op gang naar aanleiding van zes doden en 43 verdwenen studenten. Op Twitter werd de hashtag #YaMeCanse gebruikt om protesten te organiseren tegen het geweld. Daarna werden accounts actief die ruis introduceerden door met de hashtag inhoud te verspreiden in strijd met de richtlijnen van het platform (pornografisch materiaal, advertenties en gewelddadige afbeeldingen). Dit zorgde voor zoveel ruis tussen de berichten onder de hashtags dat het moeilijk werd voor

16 Vertaald met vertaalsoftware van DeepL vanuit het Vietnamees. Oorspronkelijk tekst op de website is: '*Chúng tôi cung cấp các gói tăng like, follow, view, comment cho các nền tảng mạng xã hội như Facebook, TikTok, Instagram, YouTube... thông qua hệ thống tự động. Bạn chỉ cần gửi đường link cần tăng tương tác, hệ thống sẽ tiến hành xử lý theo đúng yêu cầu trong thời gian cam kết.*'

de sociale beweging om protesten te organiseren waarna de hashtag verdween uit de trendinglijst van het platform (Treré & Bonini, 2024, p. 313).

4.2.7 Influencer-marketing

Een manier om indirect invloed uit te oefenen op het online publieke debat op socialmediaplatformen is om invloedrijke gebruikers die veel bereik hebben te betalen om inhoud te verspreiden. Een voorbeeld hiervan is het betalen van influencers.

Deze vorm van verspreiding is beschreven door Goanta van de Universiteit Utrecht, gespecialiseerd in recht, economie en beleid. Tijdens de presidentsverkiezingen in Roemenië bekeek ze influencers op TikTok die veel vergelijkbare video's bleken te maken. Het viel op dat die allemaal hun volgers vroegen om een beschrijving te geven van een ideale presidentskandidaat, terwijl kwaliteiten van kandidaat Georgescu werden genoemd. Uit een advertentie op een platform FameUp voor influencers bleek dat het om een betaalde opdracht ging die de influencers uitvoerden (*FameUp – FameUp – Over 500+ Nano and Micro Influencers Activated in 24 Hours*, z.d.).

Uit een analyse van de influenceraccounts bleek een grote verscheidenheid aan influencers: van makers van content over mode, beauty, reizen en fitness. Slechts in zeven van de veertig geanalyseerde video's gaven influencers aan dat het om betaalde marketing ging (Goanta, 2024).

Dit voorbeeld uit Roemenië laat zien dat er binnen influencermarketing soms wordt gekozen om influencers met relatief weinig volgers, maar met een specifieke doelgroep, te betalen. Dit wordt micro- of nano-influencing genoemd. Adverteerders kunnen op deze manier hun middelen gerichter inzetten. Zo voorkomen ze dat ze geld uitgeven aan een boodschap voor de verkeerde doelgroep.

Financiering van influencers kan ingezet worden zonder dat een influencer weet waar het geld vandaan komt. Op TikTok kan iemand bijvoorbeeld geld doneren aan een account zonder contact op te nemen met de influencer. Dit maakt het mogelijk om verdekt invloed uit te oefenen op de inhoud die influencers verspreiden (Goanta, 2024).

4.3 Conclusie

Platformen werken steeds meer met aanbevelingsalgoritmen om inhoud te rangschikken. Gebruikers kunnen via aanbevelingsalgoritmen makkelijker mensen bereiken die niet expliciet hebben aangegeven inhoud van hen te willen zien. De mogelijkheid om invloed uit te oefenen op welke inhoud mensen bereikt, neemt hierdoor toe. Bijvoorbeeld wanneer inhoud viraal gaat. Maar zelfs als inhoud niet viraal gaat, kan het via aanbevelingsalgoritmen een groot bereik krijgen doordat inhoud getoond wordt aan mensen die nooit hebben aangegeven de poster te volgen.

Kwaadwillende (buitenlandse) actoren hebben hierdoor een breed scala aan opties om invloed uit te oefenen op wie welke inhoud te zien krijgt. Ze kunnen ervoor kiezen om inhoud te posten en via aanbevelingsalgoritmen een zo groot mogelijke groep gebruikers te bereiken of proberen een specifieke groep gebruikers te bereiken.

Met behulp van activiteiten zoals het beheren van meerdere hyperactieve accounts, gecoördineerd posten en hashtagboosting wordt geprobeerd om een vliegwieleffect te veroorzaken.

Platformen veranderen continu. Het gevolg hiervan is dat de mogelijkheden voor inmenging ook veranderen. Daardoor is het voor onderzoekers en toezichthouders een uitdaging om up-to-date te blijven.

Het instrumentarium is beschikbaar voor iedereen. Inmengingsactiviteiten raken vaak verweven met binnenlandse activiteiten waardoor de grens tussen buitenlandse beïnvloeding en binnenlandse activiteiten vervaagd.

5 Bestaande inspanningen die inmenging via aanbevelingsalgoritmen voorkomen

In dit hoofdstuk behandelen we de vraag welke inspanningen om inmenging via aanbevelingsalgoritmen te voorkomen, al gedaan worden. We beantwoorden deze vraag aan de hand van een analyse van bestaande wet- en regelgeving. Na een beschrijving van de hoofdlijnen in wet- en regelgeving, duiden we het bestaande reguleringskader kort aan de hand van inzichten uit een door het Rathenau Instituut georganiseerde expertsessie (zie bijlage 1) en wetenschappelijk onderzoek. Daarna behandelen we in vogelvlucht de inspanningen van platformen op basis van hun eigen rapportages en (schriftelijke) interviews. Ook deze inspanningen duiden we op basis van de gehouden expertsessie en literatuur.

5.1 Bestaande wetgeving en beleid

Om de deelvraag te beantwoorden, is een analyse gedaan van de meest relevante wetgeving en beleid die betrekking heeft op inmenging en/of aanbevelingsalgoritmen. Het is belangrijk te vermelden dat de wetten ieder hun eigen reikwijdte en doelstellingen hebben en dus vanuit verschillende invalshoeken en op directe en indirecte wijze het probleem van inmenging via aanbevelingsalgoritmen benaderen.

Relevante wetten en beleid zijn:

- AI-verordening
- Algemene verordening gegevensbescherming (AVG)
- Digitaledienstenverordening (DSA) en de gedragscode tegen desinformatie
- Europese wetgeving over mediavrijheid (EMFA)
- European Democracy Shield (EDS)
- Richtlijn oneerlijke handelspraktijken
- Richtlijn audiovisuele mediadiensten
- Rijksbrede strategie voor de effectieve aanpak van desinformatie
- Verordening transparantie en gerichte politieke reclame (VPR)

De Kieswet bespreken we hier niet. Die wet is in hoofdstuk 1 kort ter sprake gekomen.

In bijlage 3 is een uitgebreide analyse te vinden. In deze paragrafen hieronder bespreken we de wetten hoofdlijnen.

Er is veel geregeld om inmengingsactiviteiten via socialmediaplatformen te beperken. Wetgevers en beleidsmakers zijn zich duidelijk bewust van het probleem en hebben daarvoor verschillende instrumenten ontwikkeld.

Waar het gaat om beleid, is er onder meer de *Rijksbrede strategie voor de effectieve aanpak van desinformatie*. Deze nationale aanpak moet de verspreiding van desinformatie op nationaal niveau tegengaan, de weerbaarheid van burgers versterken en de kennisontwikkeling op dit thema vergroten. Daarnaast staat het European Democracy Shield (EDS), bedoeld om inspanningen te ondersteunen van Europese lidstaten om hun democratie te versterken. Als onderdeel van het EDS is er een European Centre for Democratic Resilience opgericht dat tot doel heeft om samen te werken en informatie uit te wisselen op het gebied van onder andere desinformatie en buitenlandse informatiemanipulatie en beïnvloeding (Foreign Information Manipulation and Interference, FIMI).

Ten aanzien van de meest relevante wetgeving zijn een drietal rode draden te identificeren die relevant zijn bij het voorkomen en bestrijden van inmenging:
- Vergroten van de verantwoordingsplicht van platformen. (paragraaf 5.1.1)
- Ingrijpen op het ontwerp van platformen. (paragraaf 5.1.2)
- Transparantie over circulerende inhoud. (paragraaf 5.1.3)

5.1.1 Vergroten verantwoordingsplicht van platformen

Een eerste rode draad in de relevante wetgeving is de nadruk op het vergroten van de verantwoordingsplicht van platformen. Dat gebeurt met name via de digitale dienstenverordening (DSA).[17] De DSA moet gebruikers van digitale platformen beschermen tegen de verspreiding van illegale inhoud en daarbij hun grondrechten respecteren. De wet legt daarvoor verantwoordelijkheden en een verantwoordingsplicht vast voor aanbieders van intermediaire diensten, waaronder socialmediaplatformen.

17 Andere wetten, zoals de Verordening transparantie en gerichte politieke reclame (VPR) en de Algemene verordening gegevensbescherming (AVG) zorgen eveneens voor verantwoordingsplichten voor platformen. Die bespreken we kort hieronder en uitvoeriger in bijlage 3.

De DSA verplicht platformen om de verspreiding van illegale content zoveel mogelijk te beperken. Verschillende typen uitingen (zowel offline als online) zijn in Europese lidstaten, waaronder Nederland illegaal. Denk hierbij aan: discriminerende content, bedreigende content, terroristische content. Daarnaast bevat de DSA bepalingen voor klachtenprocedures en zorgvuldigheidsverplichtingen voor platformen. Bijvoorbeeld dat er adequaat gereageerd moet worden nadat illegale content is geplaatst.

Belangrijke bepalingen voor met betrekking tot inmenging zijn artikelen 34 en 35, die om een proactieve houding vraagt van zeer grote online platformen (Very Large Online Platforms, VLOPs) en zeer grote zoekmachines (Very Large Online Search Engines, VLOSEs). Daaronder vallen onder andere YouTube, Facebook, Instagram, LinkedIn, TikTok, Twitch, Google Search, Snapchat, X, WhatsApp en Bing.

Volgens de DSA moeten de zeer grote platformen en zoekmachines de systeemrisico's inventariseren en redelijke, evenredige en doeltreffende maatregelen nemen. Bijvoorbeeld om de negatieve impact op verkiezingsprocessen te voorkomen. Hierbij kan worden gedacht aan circulatie van misleidende informatie over kandidaten, stemdata of verzonnen steunbetuigingen aan kandidaten. De platformen zijn zelf vrij te bepalen hoe ze maatregelen nemen. Platformen moeten over hun inspanningen rapporteren aan de Europese Commissie.

Er loopt momenteel onderzoek door de Europese Commissie naar de werking van de aanbevelingssystemen van X (European Commission, 2025a; O'Carroll, 2025). De Europese Commissie doet ook onderzoek naar TikTok's aanbevelingsalgoritmen. Ten eerste in relatie tot 'gecoördineerde inauthentieke manipulatie of geautomatiseerde exploitatie van de dienst'. En ten tweede naar TikTok's beleid op het gebied van betaalde politieke content en politieke advertenties (European Commission, 2025e).

In de DSA worden voorstellen gedaan om systeemrisico's van platformen in te perken. Denk aan het aanpassen van aanbevelingsalgoritmen om verspreiding van misleidende of schadelijke content tijdelijk te beperken. Ook doet de DSA een aantal voorstellen. Bijvoorbeeld om virale trends op andere platformen te monitoren, moderatieteams in te zetten die bekend zijn met lokale taal en cultuur, en met AI gegenereerde content in politieke advertenties tijdelijk te verbieden.

De Europese Commissie heeft daarbij richtsnoeren opgesteld voor platformen specifiek voor verkiezingsprocessen. Zoals het wegnemen van financiële prikkels voor platformen en influencers om desinformatie, haat en extremistische content via

advertenties te verspreiden (demonetisatie) (European Commission, 2024a). Ook schrijven deze richtsnoeren voor om maatregelen te nemen tegen botnetwerken, 'inauthentieke accounts' of ander misleidend gebruik van de dienst (anti-manipulatie).

Aanbieders van platformen moeten in duidelijke, begrijpelijke taal de belangrijkste parameters noemen die in hun aanbevelingssystemen worden gebruikt. Het is optioneel voor gebruikers om deze parameters te kunnen wijzigen of te beïnvloeden. Gebruikers moeten volgens de DSA naast een tijdlijn samengesteld door een aanbevelingsalgoritmen ook standaard een feed aangeboden krijgen met content van alleen gevolgde accounts (een chronologische tijdlijn).[18]

Daarnaast dienen platformen en zoekmachines de aangewezen toezichthouders en onderzoekers inzicht te geven in het functioneren van deze platformen. Toezichthouders mogen data opvragen om naleving van de DSA te beoordelen, en platformen moeten op verzoek van toezichthouders uitleg geven over de werking van hun aanbevelingsalgoritmen. Ook moeten aangewezen onderzoekers toegang krijgen tot data om systeemrisico's te onderzoeken. Platformen mogen een verzoek van onderzoekers alleen afwijzen als zij de data niet hebben of als toegang tot ernstige beveiligings- of vertrouwelijkheidsrisico's zou leiden.[19]

Tot slot is er de Gedragscode desinformatie, die gezien kan worden als een verdere uitwerking van de regels die in de DSA staan. De ondertekenaars van de gedragscode moeten zich bijvoorbeeld inspannen om ongeoorloofd manipulatief gedrag en praktijken te beperken. Hieronder vallen het gebruik van bots, valse accounts en gecoördineerd inauthentiek gedrag en 'gebruikersgedrag' gericht op het kunstmatig vergroten van het bereik of waargenomen publieke steun voor desinformatie. Daarbij wordt verwezen naar een omvangrijke, openbare en regelmatige bijgewerkte database met *tactics, techniques and procedures* van DISARMFoundation (DISARMFoundation, z.d.).

18 De ngo Bits of Freedom heeft een kort geding tegen Meta gewonnen waarbij het eiste dat gebruikers de chronologische tijdlijn als standaard kunnen instellen, in plaats van de algoritmische tijdlijn (*Meta moet (voorlopig) chronologische tijdlijn als blijvende voorkeursoptie aanbieden*, 2026).

19 X is recentelijk beboet voor het schenden van transparantieverplichtingen, het belemmerde onder meer de toegang tot data voor onderzoekers en gaf onvoldoende inzicht in advertenties (European Commission, 2025d). De Europese Commissie heeft ook gesteld in preliminaire bevindingen dat Meta en TikTok niet goed genoeg toegang gaven tot onderzoeksdata. Het onderzoek loopt nog(European Commission, 2025c). Ook loopt er nog een ander onderzoek tegen Meta vanwege de slechte toegang tot onderzoeksdata en het offline halen van CrowdTangle, dat bedoeld was om verkiezingen te monitoren (European Commission, 2024b).

5.1.2 Ingrijpen op het ontwerp van platformen

Er is wetgeving die beoogt in te grijpen op het ontwerp van platformen. Zo verplicht de DSA de zeer grote platformen en zoekmachines om de risico's van verslavend ontwerp te beperken. Een verslavend ontwerp is bijvoorbeeld een zeer gepersonaliseerd aanbevelingsalgoritme. Het aanpakken van verslavend ontwerp is in de context van dit onderzoek relevant in zoverre het gaat om het ingrijpen op de op engagement gebaseerde aanbevelingsalgoritmen waar kwaadwillende actoren op in kunnen spelen.

De voorlopige bevinding van de Europese Commissie is dat TikTok onvoldoende maatregelen heeft genomen om verslavende aspecten van het ontwerp te mitigeren (European Commission, 2026).[20] Ook is er een procedure van de Commissie gestart naar Meta's Instagram en Facebook omdat het ontwerp te verslavend zou zijn voor minderjarigen (European Commission, 2024c).

Volgens de AI-verordening kan verslavend ontwerp in een AI-systeem zelfs worden opgevat als een verboden praktijk als hiermee misbruik wordt gemaakt van de leeftijd, handicap, of socio-economische achtergrond van gebruikers. Ook volgt uit de AVG dat het verzamelen en verwerken van iemands gegevens met het doel de gebruiker verslaafd te maken, waarschijnlijk niet verenigbaar is met het behoorlijkheidsprincipe. Het demissionair kabinet heeft aangegeven de zaken in de hierboven genoemde alinea van de Europese Commissie met aandacht te volgen en anders in te zetten op de Digital Fairness Act, waar het reguleren van verslavend ontwerp van socialmediaplatformen centraal staat (*Kamerstukken II, 26 643, nr. 1391*, 2025). Ook in het coalitieakkoord is aangekondigd verslavende algoritmen aan te pakken (D66, VVD en CDA, 2026).

Verder is er aandacht voor manipulatief ontwerp. De DSA verbiedt manipulatieve technieken. Mogelijk valt verslavend ontwerp ook onder manipulatieve technieken, maar dat moeten bovengenoemde zaken uitwijzen (*Kamerstukken II, 26 643, nr. 1391*, 2025). X heeft recent een boete van de Commissie gekregen voor het misleiden van gebruikers met blauwe vinkjes (European Commission, 2025d).

Er is een verbod op AI-systemen waarbij manipulatieve technieken worden gebruikt. Of socialmediaplatformen onder manipulatieve systemen vallen, staat ter discussie. Ten tijde van schrijven loopt er een rechtszaak van Stichting SOMI tegen TikTok met de aanklacht dat het platform jongeren gericht manipuleert, en voor het misbruiken van gevoelige persoonlijke gegevens, die door het platform worden

20 Er is ook nog een (inmiddels afgesloten) onderzoek geweest van de Europese Commissie, onder andere naar de verslavende componenten van het TikTok Lite Awards Programma, welke TikTok inmiddels heeft teruggetrokken (European Commission, 2024d).

gebruikt om het eigen aanbevelingsalgoritme aan te sturen (*Nederlandse claimstichting begint rechtszaken tegen TikTok en X (voorheen Twitter)*, 2025).

Op grond van de AVG mag de verwerking van persoonsgegevens niet onevenredig, onverwacht of misleidend zijn (behoorlijkheidsprincipe). Dat zijn volgens de wet manipulatieve technieken. De onderzoeksdienst van het Europarlement, de EPRS, verwacht dat het aanbevelingssysteem van bijvoorbeeld TikTok, waarbij gedragsdata van mensen worden gebruikt – volgens de onderzoeksdienst – om mensen te manipuleren, onverenigbaar is met het behoorlijksheidsprincipe van de AVG (European Parliamentary Research Service, 2025).

5.1.3 Transparantie over circulerende inhoud

Een derde rode draad, naast het vergroten van de verantwoording en het ingrijpen op het ontwerp, is dat de wetgever de transparantie van online inhoud wil verhogen. De wetten stellen vooral transparantie-eisen aan de *makers* van inhoud, maar *platformen zelf* moeten ook voldoen aan transparantieverplichtingen. De wetten kunnen derden in staat stellen (burgers, onderzoekers, overheden) om onderzoek te doen, bijvoorbeeld naar inmengingsactiviteiten. Zo moeten onder de AI-verordening en de DSA-verordening bijvoorbeeld AI-afbeeldingen, die nauwelijks van echt te onderscheiden zijn, gelabeld worden.

Ook is rekening gehouden met de rol van influencers zoals vloggers en bloggers die een vergoeding krijgen voor (politieke) boodschappen. De Europese wetgeving voor mediavrijheid (EMFA) stelt onder meer vast wat een mediadienst is. Mogelijk vallen beroepsmatige influencers daar ook onder. Zeer grote platformen en zoekmachines moeten onder de EMFA een functionaliteit inbouwen waarmee aanbieders van mediadiensten een reeks verklaringen kunnen afleggen: dat ze media-aanbieder zijn, dat zij redactioneel onafhankelijk zijn van lidstaten, politieke partijen en derde landen, dat zij voldoen aan regelgevende voorschriften, en dat zij geen AI-gegenereerde content aanbieden zonder menselijke toetsing.

Verder is er de Richtlijn oneerlijke handelspraktijken die eisen stelt aan influencers dat ze transparant zijn over de reclame die ze maken en de voordelen die ze ermee genieten. Het inzetten van nepaccounts voor neplikes en nepvolgers is niet toegestaan. De Richtlijn audiovisuele diensten schrijft ook voor dat misleiding, agressieve praktijken en de inzet van subliminale technieken (verborgen boodschappen) van iedereen die reclame verspreidt, waaronder ook in video's en livestreams, niet toegestaan is.

Onder de noemer van het vergroten van transparantie van circulerende inhoud is er verder de Verordening transparantie en gerichte politieke reclame (VPR). Het doel van deze wet is om burgers te helpen gemakkelijker politieke reclame online te herkennen. Inmenging en online manipulatie is een van de aanleidingen geweest om de VPR op te stellen.

De VPR stelt transparantie-eisen aan politieke reclame en regels aan het gebruik van zogeheten targeting en optimalisatietechnieken voor de verspreiding van politieke reclame. De zeer grote platformen en zoekmachines hebben onder de VPR ook een paar specifieke verplichtingen, vooral gericht op publieke transparantie over de reclames en het snel reageren op verzoeken van bijvoorbeeld melders en toezichthouders. De VPR stelt dat vanaf drie maanden voordat de verkiezingen plaatsvinden, politieke reclamediensten alleen mogen worden verleend aan opdrachtgevers die een EU-burger zijn of een bepaalde band hebben met de EU.

Onder de DSA moeten de zeer grote platformen en zoekmachines registers aanleggen met informatie over de reclame (zoals beoogde doelgroep, adverteerder) die op hun platformen geplaatst wordt. Die registers moeten publiek toegankelijk zijn.

5.1.4 Versoepeling op komst?

Ten tijde van schrijven van dit rapport is een zogenaamde Digitale Omnibus voorgesteld door de Europese Commissie met als doel regels voor bedrijven te versimpelen (*Digital Omnibus Proposal COM/2025/837*, 2025). Hoewel bij het bedrijfsleven behoefte is aan minder regeldruk, bestaan er zorgen, waaronder bij het Nederlandse demissionaire kabinet, dat dit tot onwenselijke versoepeling leidt (*Kamerstukken II, 22 112, nr. 4223*, 2025, nr. 4223 2025).

Een van de mogelijke aanpassingen in dit verband is het uitstellen van de implementatie van vereisten voor hoogrisico-systemen en de mogelijkheid voor aanbieders om zichzelf vrij te stellen van de verplichtingen die gekoppeld zijn aan een hoogrisico-systeem, zonder dat aan te geven in een daarvoor bestemde publieke database. Daardoor zouden AI-systemen die gebruikt kunnen worden voor het beïnvloeden van verkiezingen onder de radar van de toezichthouder kunnen blijven.

Een andere aanpassing in dit verband is dat de definitie van persoonsgegevens mogelijk minder breed opgevat hoeft te worden. Daardoor zouden partijen zich minder snel aan de AVG hoeven te houden en hoeven manipulerende partijen zich

in zulke gevallen ook niet meer aan het behoorlijkheidsprincipe te houden. Dit omdat bepaalde gegevens niet meer onder de definitie van persoonsgegevens vallen. Verder zou het trainen en exploiteren van AI-systemen met persoonsgegevens (ook bijzondere gegevens zoals politieke informatie) worden toegestaan zonder dat daar toestemming voor nodig is.

Kader 4 Geopolitieke context: platformregulering ligt onder vuur

Sinds het aantreden van de tweede regering Trump in de Verenigde Staten, ligt de Europese platformregulering onder vuur. President Trump, vicepresident Vance, en recent ook de juridische commissie van het Amerikaanse House of Representatieves, hebben stevige kritiek op Europese regelgeving, met name op de DSA. Zij beschuldigen de Europese Commissie, de ACM (toezichthouder van de DSA in Nederland), en enkele experts die wij hebben gesproken voor dit rapport (omdat ze aanwezig waren op een rondetafel over inmenging via socialmediaplatformen, georganiseerd door de ACM) van censuur, zien de Europese Commissie als de partij die door middel van de DSA controle heeft over het publieke debat in Amerika, en dreigen met stappen in de handelsrelatie of militaire samenwerking als Europa haar greep op platformen verstevigt.

Dit Amerikaanse perspectief staat haaks op het perspectief vanuit Europa dat Amerikaanse platformen juist een grote stempel drukken op het publieke debat in Europa, dat zij verantwoordelijkheid dragen voor de negatieve effecten van hun diensten, en dat daarom regels nodig zijn. Daarnaast gaat het ook voorbij aan het feit dat de DSA heel bewust juist niet voorschrijft wat er wel of niet gezegd mag worden online (dat doen nationale wetten, door bijvoorbeeld discriminatie, haatspraak of *deepnudes* te verbieden), maar platformen een verantwoordingsplicht geeft om duidelijk te maken hoe ze wetten naleven, hoe ze moeten omgaan met meldingen van gebruikers en systeemrisico's mitigeren. Met deze ontwikkelingen is platformregulering in het hart van de geopolitieke machtsstrijd terechtgekomen. (Christina Lu, 2026; Committee on the Judiciary of the U.S. House of Representatives, 2026; *Statement by the European Commission on the U.S. Decision to Impose Travel Restrictions on Certain EU Individuals*, 2025)

5.1.5 Experts: Wetgeving lijkt niet overal in te grijpen op instrumentarium inmenging en afhankelijkheid bereidwilligheid platformen is groot

Hoewel veel is geregeld, ziet het ernaar uit dat wetgeving onvoldoende ingrijpt op het instrumentarium van inmengende actoren ten aanzien van het gebruik van AI. Opvallend is dat AI-systemen die worden gebruikt voor het beïnvloeden van de uitslag van een verkiezing of referendum of van stemgedrag, volgens de AI-verordening niet verboden zijn, maar in de hoogrisico-categorie vallen. Zeker jongere generaties maken volgens experts veelvuldig gebruik van deze systemen voor het nemen van alledaagse beslissingen. Kiesgerechtigden gebruikten volgens de Autoriteit Persoonsgegevens in toenemende mate AI-chatbots voor hun stemkeuze in aanloop naar de vorige verkiezingen (Autoriteit Persoonsgegevens, 2025).

Verder waarschuwen experts voor de opkomst van AI-gegeneerde en gepersonaliseerde nieuwsoverzichten (Schneier & Sanders, 2025). Het is nog niet duidelijk onder welke categorie deze overzichten vallen in de AI-verordening, omdat de verordening – naast hoogrisico-AI-systemen – de *modellen* reguleert die gebruikt worden voor generatieve AI-tools, niet zozeer de *tools* zelf. Het is niet uit te sluiten dat die tools hoogrisico-AI-systemen kunnen zijn. Een expert wees tenslotte op het feit dat meer aandacht moet komen voor inmenging via chatbots, die op dit moment ook beperkt gereguleerd zijn in de verordening. Het moet de gebruiker van een chatbot volgens de wet nu alleen duidelijk zijn dat deze met een chatbot praat.

Bij het succes van wetgeving op het gebied van inmenging speelt de bereidwilligheid van platformen een belangrijke rol. Zoals beschreven in dit hoofdstuk hebben toezichthouders een groot aantal onderzoeken lopen.

Daarnaast is er vanuit experts kritiek op de inspanningen van Google en Meta in aanloop naar de Verordening betreffende transparantie en gerichte politieke reclame. Met de komst van deze verordening hebben Google en Meta aangegeven geen politieke reclame meer op hun platformen te faciliteren, onder meer omdat het volgens beide bedrijven tot grote complexiteit en juridische onzekerheden voor adverteerders en platformen in de EU zou leiden (*Google stopt vanaf september met politieke advertenties*, 2025). Ook heeft Google diens zeven jarig archief met Europese politieke reclames verwijderd, wat het inzicht voor onderzoekers in inmenging verder bemoeilijkt (404 Media, 2025). Meta stelde daarnaast dat het ook geen 'maatschappelijke kwesties' meer zou toestaan.

Experts stellen dat deze keuzes impact hebben op de vrijheid van meningsuiting. Ook denken ze dat het niet toestaan van politieke advertenties politici en anderen aanzet om explosieve inhoud te verspreiden om zo via aanbevelingsalgoritmen hun doelgroepen te bereiken (Eijsvoogel, 2025; *Midden in campagnetijd ineens geen politieke reclame meer op Facebook en Instagram*, 2025). Het doel van deze wet, burgers gemakkelijker politieke reclame online laten herkennen, lijkt door de keuzes van deze platformen daardoor juist verder uit het zicht.

5.2 Inspanningen van socialmediaplatformen

Socialmediaplatformen spannen zich op verschillende manieren in om inmenging in verkiezingen te voorkomen. In deze paragraaf bespreken we *community guidelines* en contentmoderatie.

Voor dit onderzoek benaderden we Meta (Facebook en Instagram), TikTok, Snap (Snapchat), Google (YouTube), Microsoft (LinkedIn) en X voor online interviews (zie voor de methodologische verantwoording bijlage 1 en 2). Van hen reageerden Meta en TikTok schriftelijk. Met Snap hebben we een mondeling interview gehouden. Google en Microsoft hebben gekozen om niet in te gaan op ons verzoek en X heeft niet gereageerd.

Voor deze paragraaf baseren we ons naast de interviews en schriftelijke antwoorden ook op openbare informatie van socialmediaplatformen.

5.2.1 Community guidelines

Socialmediaplatformen hanteren *community guidelines* waarin staat wat wel en niet toegestaan is op platformen (zie paragraaf 3.3.1). Die richtlijnen gaan zowel over inhoud als over activiteiten, zoals manieren van verspreiding. Platformen zijn onder artikel 14 van de DSA ook verplicht om hun eigen richtlijnen te handhaven.

Community guidelines dienen als basis voor hoe gemodereerd wordt in online omgevingen. Het zijn de gedragsregels of huisregels die bepalen welke gedragingen wel en niet zijn getolereerd. Hierin staan bijvoorbeeld regels opgesteld die aangeven welke taal en afbeeldingen gebruikers wel en niet mogen gebruiken op een platform. Veelgebruikte online omgevingen hanteren bijna altijd een vorm van gecentraliseerde contentmoderatie, waarbij platformen de regels bepalen en kiezen hoe ze die regels handhaven (Rathenau Instituut, 2025a).

Contentmoderatie kan leiden tot het verwijderen van inhoud van het platform, maar ook tot het beperken van het bereik van een post op basis van de inhoud. Platformen hanteren hier verschillende regels voor. Snap hanteert bijvoorbeeld strengere regels voor inhoud die door hun aanbevelingsalgoritmen aanbevolen kan worden, dan voor inhoud die door mensen in hun eigen netwerk verspreid kan worden (Snapchat, z.d.). Meta's Community Standards specificeren de content die is verboden als het gaat om geweld of haatdragend gedrag. Daarnaast geeft Meta aan dat het bereik van bepaalde vormen van gewelddadige inhoud naar mensen boven de achttien beperkt wordt en een waarschuwing wordt getoond. Meta plaatst gebruikers met een tieneraccount automatisch in de strengste instelling wat betreft gevoelige inhoud (Meta, z.d.).

Alle grote socialmediaplatformen hebben richtlijnen die betrekking hebben op het inmengingsinstrumentarium (hoofdstuk 4): het delen van misinformatie, niet gelabelde AI-gegenereerde inhoud, haatzaaiende content, het gebruik van nepaccounts. Om weerstand te bieden tegen inmengingsactiviteiten hebben platformen dus expliciet beleid.

De zeer grote online platformen (VLOPs) zijn verplicht door de DSA om jaarlijks in transparantierapportages te melden hoe ze contentmoderatie hebben uitgevoerd (zie tabel 5 hieronder). In deze rapportages staan onder meer de aantallen berichten die zijn verwijderd omdat die in strijd waren met de richtlijnen.

Om een beeld te krijgen van openbare bronnen waarin platformen rapporteren over inspanningen, geven wij in tabel 5 (hierna) verwijzingen naar transparantierapporten en richtlijnen die gaan over het in hoofdstuk 3 beschreven inmengingsinstrumentarium.

Tabel 5 Overzicht van door platformen gerapporteerde inspanningen om inmenging tegen te gaan

Organisatie	*Community guidelines*	Specifieke maatregelen tegen inmenging in verkiezingen
Google (YouTube)	Transparantierapport, algemene richtlijnen, en specifiek: beleid tegen misleidende informatie over verkiezingen en beleid tegen valse betrokkenheid (engagement) en beleid tegen nabootsing van identiteit	Google schreef over voorbereiding verkiezingen Verenigde Staten
Meta (Facebook en Instagram)	Transparantierapport, algemene richtlijnen en specifiek beleid tegen haatzaaiend gedrag, beleid tegen niet-authentiek gedrag en beleid gericht tegen desinformatie	Election integrity reports blog over global elections meta platforms met preventing foreign interference
Microsoft (LinkedIn)	Transparantierapport, algemene richtlijnen, specifiek richtlijnen over wees betrouwbaar, richtlijnen over onjuiste of misleidende content	Niet bekend
Snap (Snapchat)	Transparantierapport, algemene richtlijnen, specifiek richtlijnen tegen schadelijke, valse of misleidende praktijken en beleid tegen haatdragende content'	Reflections on Dutch elections
TikTok	Transparantierapport, algemene richtlijnen, specifiek beleid tegen integrity and authenticity, waaronder Edited Media and AI-Generated Content (AIGC)	Global elections hub, blog over het beschermen van de integriteit van de Tweede Kamerverkiezingen
X (voorheen Twitter)	Transparantierapport, algemene richtlijnen, specifiek richtlijnen over authenticiteit en richtlijnen over hateful conduct	Niet bekend

De tabel geeft een niet-uitputtend overzicht van openbare documenten waarin platformenbedrijven schrijven over inspanningen tegen inmengingen in verkiezingen via hun socialmediaplatformen. Samenstelling: Rathenau Instituut

5.2.2 Contentmoderatie

Socialmediaplatformen gebruiken een combinatie van menselijke en automatische contentmoderatie om proactief schadelijke content te detecteren. Voor sommige content geldt dat het voor platformen snel duidelijk is of iets gemanipuleerd, schadelijk of onjuist is. Dit gebeurt automatisch met behulp van AI-systemen en/of met behulp van contentreviewers. Voor een deel van de content is het noodzakelijk

om kennis te hebben van de lokale taal en cultuur om die te kunnen beoordelen. In tabel 6 geven we een overzicht van het aantal Nederlandssprekende contentmoderatoren in dienst van grote online platformen, zoals door de platformen zelf gerapporteerd.

Tabel 6 Nederlandssprekende contentmoderatoren volgens zelfrapportages

Organisatie	Aantal Nederlandssprekende contentmoderatoren	Aantal gebruikers in Nederland
Facebook en Instagram (Meta)	111	Facebook 10.300.000 Instagram 11.900.000
LinkedIn (Microsoft)	8	5.300.000
Snap (Snapchat)	23	6.148.873
TikTok	100	6.700.000
X (voorheen Twitter)	0	8.242.130
YouTube (Google)	75	*42.100.000

Nederlandssprekende contentmoderatoren en gebruikersaantallen in Nederland per 1 augustus 2025 volgens platformrapportages. * Google differentieert geen unieke maandelijke gebruikers, dus dit getal is moeilijk te vergelijken. Bron: (Google, 2025a, 2025b; LinkedIn, 2025; Meta, 2025b, 2025c; Snap, 2025a; TikTok, 2025; X, z.d.).

In de (schriftelijke) interviews met Meta, TikTok en Snap gaven zij voorbeelden van initiatieven die ze hebben genomen in aanloop naar de Tweede Kamerverkiezingen van 2025. Zonder volledig te zijn, geven we hier een indruk van het type inspanningen dat deze platformen zeggen te leveren.

Handhaving eigen richtlijnen

TikTok geeft aan 'meer dan 4.000 video's' te hebben verwijderd die in strijd zijn met richtlijnen over 'civic integrity, misinformation and AI-generated content' voorafgaand en na de Nederlandse Tweede Kamerverkiezingen van 2025. TikTok geeft aan dat meer dan 96% van die video's zijn verwijderd voordat iemand ze rapporteerde. Ook zegt TikTok meer dan 1.000 accounts te hebben verwijderd vanwege impersonatie van Nederlandse verkiezingskandidaten en bewindspersonen.

TikTok noemt ook initiatieven die ze hebben genomen specifiek gericht op influencermarketing. TikTok geeft aan dat specifiek politieke geldinzamelingsacties en politieke fondsenwerving verboden zijn op het platform. Dit houdt ook in dat betaalde advertenties gerelateerd aan politiek of verkiezingen verboden zijn en dat makers (creators) hier ook niet voor betaald mogen worden.

Snap geeft aan dat tijdens de periode vooraf, tijdens en na afloop van Nederlandse Tweede Kamerverkiezingen van 2025 een handvol posts zijn verwijderd naar aanleiding van hun beleid over het verspreiden van valse informatie. Bij de gerapporteerde inhoud (16 stuks) was 'een AI gegenereerde post, en eentje die een poging tot humor was'. Voor de verwijderde berichten geldt dat deze 'even (beperkt) te zien zijn geweest, en daarna verwijderd'. Snap geeft aan dat slechts in één geval sprake is geweest van politieke desinformatie, en dat deze post is verwijderd voordat een groter publiek kon worden bereikt.

Dagelijks neemt Snap 'een sample van Snaps die langs alle moderatie zijn gekomen, om te kijken wat daaraan te zien is en wat er speelt en of hun beleid heeft gewerkt.'

Snap geeft aan op Spotlight politieke- en nieuwscontent van uitgevers waarmee Snap samenwerkt, en van geverifieerde 'Snap Stars' toe te staan. Alleen die content kan naar een groter publiek worden verspreid. Snap noemt dit selectie aan de poort: alleen bepaalde actoren mogen communiceren over verkiezingen.

Snap geeft aan dat inhoud niet misleidend of incorrect mag zijn: het platform doet niet aan 'een beetje desinformatie'. Snap zegt verder bij desinformatie niet aan labelen of downranken te doen.

Nepaccounts en gecoördineerde acties
Platformen noemen verschillende initiatieven gericht op het tegenwerken van nepaccounts. TikTok schrijft aan te dringen op verificatie van accounts om gebruikers beter informatie te geven over wie er achter het account zit, zodat de betrouwbaarheid kan worden ingeschat.

Meta gebruikt geautomatiseerde systemen om 'inauthentiek gedrag' en nepaccounts op te sporen. Meta schrijft in haar antwoorden op onze vragen dat detectietechnologie dagelijks 'miljoenen pogingen tot het creëren van nepaccounts blokkeert en blokkeert miljoenen aangemaakte nepaccounts binnen minuten na het aanmaken'. Verder schrijft Meta ook: 'tussen juli en september 2025 hebben we 698 miljoen nep-accounts verwijderd, waarvan 99,9% door ons werd gevonden en aangepakt voordat gebruikers ze konden melden'.

Snap geeft aan dat ze nog nooit gecoördineerde acties hebben gezien. Snap denkt dat de huidige maatregelen dermate goed zijn dat gecoördineerde acties niet op hun platform plaats kunnen vinden.

Maatregelen rondom verkiezingen
TikTok en Meta laten weten contentmoderatoren en teams te trainen gericht rondom de verkiezingen. TikTok nodigt sprekers uit met expertise in 'election integrity, media literacy en disinformation' om interne presentaties te gegeven voorafgaande aan verkiezingen. Meta schrijft: 'Er zijn verschillende multidisciplinaire teams binnen Meta die zich richten zich op verkiezing-gerelateerde onderwerpen en samen voorbereidingen te treffen om specifieke risico's rond de verkiezingen te identificeren, beoordelen en te mitigeren. Nieuwe risico's worden ook gemonitord en beoordeeld, zodat relevante mitigaties ook toegepast kunnen worden.'

Meta werkt samen met factcheckingpartners Agence France Press (AFP) en Deutsche Presse-Agentur. Deze partners beoordelen en classificeren virale misinformatie, waaronder het label 'Aangepast' voor content die is bewerkt op een manier die mensen kan misleiden, bijvoorbeeld door het gebruik van AI. Meta geeft aan dat content die het beleid van Meta schendt, wordt verwijderd.

Meta geeft verder aan gebruikers te wijzen op betrouwbare informatie over de datum van de verkiezingen en mensen te verwijzen naar de website van de Rijksoverheid met meer informatie over de verkiezingen.

Snap geeft aan dat ze 'een heel cross functional team hebben met verschillende disciplines, die bij elkaar komt in geval van crisissen.' Daarnaast geeft Snap aan dat ze 'reguliere policy briefings maken voor hun moderatieteams, waar ook ontwikkelingen rondom verkiezingen in kunnen staan'.

5.2.3 Experts: meer transparantie nodig voor onafhankelijke waarnemers

Wat gebruikers online te zien krijgen en welke rol aanbevelingsalgoritmen hierin spelen, blijft grotendeels onbekend voor onafhankelijk waarnemers (Cooper & Chapman, 2025). Meer transparantie is nodig om te onderzoeken hoe aanbevelingsalgoritmen in de praktijk uitwerken.

Deelnemers aan onze expertsessie geven aan onder andere de volgende informatie te missen om goed onderzoek te kunnen doen naar risico's op inmenging, manipulatie en misleiding op socialmediaplatformen:

- Informatie over het bereik van (politieke) inhoud. Wie heeft de inhoud allemaal gezien? Ook voordat deze verwijderd is door een platform.
- Informatie over hun monitoringsprocessen: hoe sporen platformen manipulatieve netwerken van accounts op en hoe weten we of dit voldoende werkt?
- Informatie over herkomst van accounts.
- Exactere informatie over het bereik en de afzender van advertenties.
- Inzicht in de werking van aanbevelingsalgoritmes en de gewichten en bedrijfsregels achter de aanbevelingen.

Ook geven de experts aan dat de toegang voor onderzoekers tot platformen, een verplichting onder de DSA, te wensen over laat. Verzoeken worden nu vaak afgewezen en het is vaak onduidelijk over welke informatie platformen beschikken die opgevraagd zouden kunnen worden (zie ook paragraaf 5.1).

5.2.4 Experts: handhaving *community guidelines* schiet te kort

De handhaving van *community guidelines* schiet te kort. Dit concluderen onderzoekers van het Hybrid Election Integrity Observatory, waarin Post-X Society, AI Forensics, Trollrensics, de Universiteit van Amsterdam en Justice for Prosperity samenwerken. Het onderzoekscollectief betoogt dat contentmoderatie tijdens de Nederlandse Tweede Kamerverkiezingen van 2025 'onvoldoende bescherming' bood. Zij baseren zich hierbij op het niet opvolgen van meldingen van illegale inhoud door platformen en het niet verwijderen van inhoud die tegen de *community guidelines* van platformen in ging. De onderzoekers betogen dat de *community guidelines* van platformen 'vooral voor pr-doeleinden bestaan' en 'niet bijdragen aan veiligheid van gebruikers' (HEIO Consortium et al., 2026, p. 59).

Ook in andere landen bestaan signalen dat contentmoderatie in verkiezingstijd tekortschiet. Het Fimi Defenders for Election Integrity (FDEI) consortium, waarin tien organisaties in een consortium samenwerken, schrijft in hun verkiezingsrapportage over de Tsjechische parlementsverkiezingen dat TikTok 187.000 nepaccounts, 2,9 miljoen neplikes en 2 miljoen nepvolgers heeft verwijderd voorafgaand aan de Tsjechische parlementsverkiezingen (GLOBSEC et al., 2025). Desondanks lukte het een netwerk van pro-Russische TikTok-accounts toch om tussen de 5 en 9 miljoen views te behalen, zonder dat deze offline werden gehaald.

5.3 Conclusie

Er is veel geregeld om het risico op inmenging via socialmediaplatformen te voorkomen. De wetgever is zich duidelijk bewust van het probleem en heeft verschillende instrumenten ontwikkeld om inmenging te voorkomen. De wetgever richt zich daarbij tot haar invloedssfeer (influencers, platformen) en niet op inmengende actoren zelf. Rode draden in het reguleringskader zijn erop gericht de verantwoordingsplicht van platformen vergroten, in te grijpen op het ontwerp van platformen en meer transparantie te realiseren over circulerende inhoud.

Grote online platformen leveren met behulp van richtlijnen en contentmoderatie inspanningen om inmenging in verkiezingen via hun platformen te voorkomen. Zoals in paragraaf 5.1.1 wordt uitgelegd, verplicht de DSA hen ook tot het nemen van maatregelen rondom de bescherming van de integriteit van verkiezingen.

Desondanks laten bevindingen van ngo's en wetenschappers, zelfs met beperkte toegang tot platformdata, zien dat platformen er niet altijd in slagen om inmengingsactiviteiten te voorkomen of het beschikbare instrumentarium te verminderen. Denk aan gecoördineerde acties (likes, of het delen van posts), het delen van inhoud die tegen de gedragsregels van platformen ingaat, het laat ingrijpen op meldingen en het mede mogelijk maken van de verspreiding van illegale inhoud. De experts die wij raadpleegden in onze expertsessie (zie bijlage 1 en 2) signaleren ook dat de risico's op inmenging via aanbevelingsalgoritmen groot en divers zijn, en dat meer maatregelen nodig én mogelijk zijn.

Het volgende hoofdstuk biedt handelingsperspectieven voor politiek en beleid om de risico's op inmenging via aanbevelingsalgoritmen van socialmediaplatformen verder te beperken.

6 Conclusie en handelingsperspectieven

In dit hoofdstuk beantwoorden we eerst de centrale vraag in dit onderzoek: welke rol spelen aanbevelingsalgoritmen op grote socialmediaplatformen bij inmenging in verkiezingen? Dit doen we op basis van alle voorgaande hoofdstukken. Vervolgens gaan we in op de laatste deelvraag: welke (beleids)maatregelen zijn mogelijk om inmenging via aanbevelingsalgoritmen te voorkomen en/of aan te pakken?

6.1 Conclusie

Onderzoeksvraag: welke rol spelen aanbevelingsalgoritmen op grote socialmediaplatformen bij inmenging in verkiezingen?

Kwaadwillende buitenlandse statelijke actoren kunnen inspelen op aanbevelingsalgoritmen van sociale media door de zichtbaarheid van bepaalde inhoud te vergroten of te verkleinen. Er is sprake van inmenging wanneer deze actoren proberen de ontvanger bewust te misleiden over de inhoud en/of de populariteit daarvan. Doel is om uiteindelijk de publieke opinie of verkiezingen te beïnvloeden. Ze weten daarbij het socialmedialandschap te benutten.

Socialmediaplatformen zijn steeds sterker gaan leunen op engagement gebaseerde aanbevelingsalgoritmen, en hebben zo een online omgeving gecreëerd waar inhoud die inspeelt op emotie, misleiding en opruiing een groot bereik kan krijgen. Kwaadwillende buitenlandse actoren spelen hierop in door dit type inhoud te maken én te proberen om dit type inhoud een groter bereik te geven.

Technisch kunnen aanbevelingsalgoritmen anders worden ingesteld, maar fundamentele aanpassingen gaan in tegen het verdienmodel van platformbedrijven: platformen hebben baat bij op engagement gebaseerde aanbevelingsalgoritmen, omdat het de tijd die gebruikers doorbrengen op een platform vergroot. Het instrumentarium voor inmengingsactiviteiten zit zo dus feitelijk ingebouwd in de platformen.

Er is veel geregeld om het risico op inmenging via socialmediaplatformen te voorkomen of te bestrijden. De wetgever is zich duidelijk bewust van het probleem, en heeft verschillende instrumenten ontwikkeld om verschillende aspecten van inmenging aan te pakken. Ook platformen hebben regels opgesteld. Toch wijzen

experts nog op de noodzaak voor aanvullende maatregelen en handhaving. Om risico te verkleinen, geven we handelingsperspectieven in drie richtingen: platformontwerp veranderen, informatiepositie versterken en weerbaarheid vergroten. Daarover gaat de volgende paragraaf.

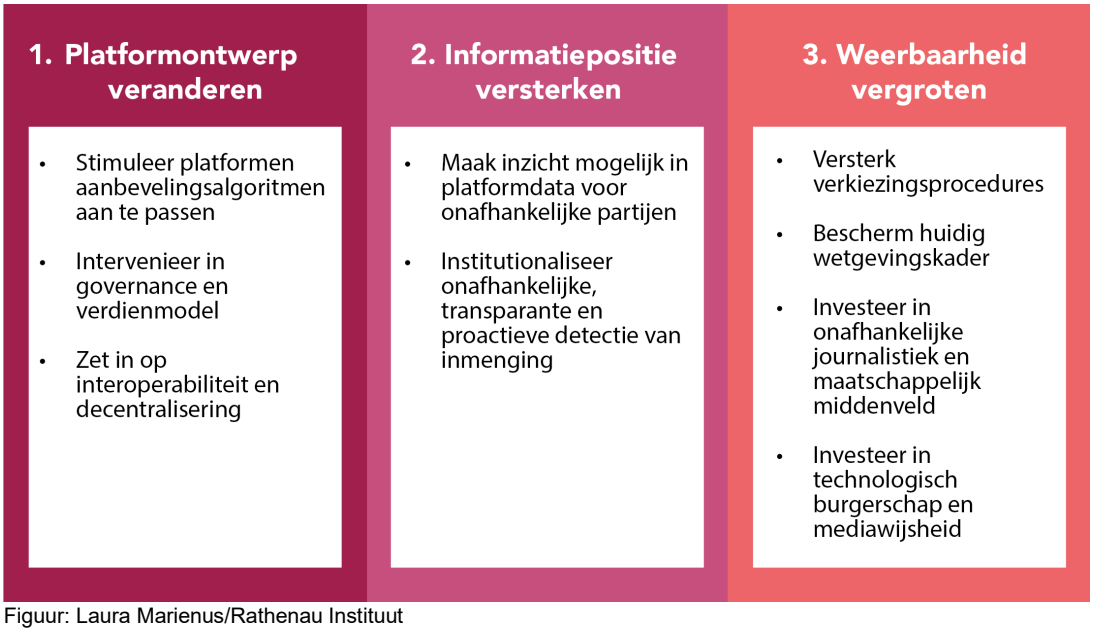
6.2 Handelingsperspectieven

We structureren de handelingsperspectieven aan de hand van drie richtingen die uit ons literatuuronderzoek, interviews, expertsessie en analyse daarvan naar voren komen

De drie richtingen zijn:
1. Platformontwerp veranderen. (paragraaf 6.2.1)
2. Informatiepositie versterken. (paragraaf 6.2.2)
3. Weerbaarheid vergroten. (paragraaf 6.2.3)

In figuur 4 vatten we de handelingsperspectieven samen. In de paragrafen hierna lichten we ze verder toe.

Figuur 4 Handelingsperspectieven om inmenging te verminderen



Figuur: Laura Marienus/Rathenau Instituut

6.2.1 Platformontwerp veranderen

Ontwerpkeuzes die platformen maken over de inrichting van hun aanbevelingsalgoritme, contentmoderatie en platform, hebben veel invloed op wat mensen online zien en kunnen verspreiden. Platformen maken die ontwerpkeuzes onder andere gebaseerd op hun visie, governance en verdienmodel (Rathenau Instituut, 2025a). De manier waarop aanbevelingsalgoritmen ontworpen zijn, biedt kwaadwillende actoren een instrumentarium om in te mengen in ons online publieke debat. Overheden kunnen daarom overwegen om platformen te stimuleren of verplichten hun platformontwerp aan te passen. Het coalitieakkoord 2026-2030 lijkt aan te sturen op aanpassingen, want hierin staat: 'Verslavende, polariserende en antidemocratische algoritmes worden verboden' (D66, VVD en CDA, 2026).

Bij de handelingsperspectieven maken we onderscheid in het ingrijpen in aanbevelingsalgoritmen en in het ingrijpen in governance en verdienmodellen.

Stimuleer platformen om aanbevelingsalgoritmen aan te passen
Overheden en toezichthouders kunnen socialmediaplatformen bevragen op de maatregelen die zij nemen om de risico's die voortvloeien uit hun aanbevelingssystemen te beperken. De risico-rapportages die de grote platformen onder de DSA moeten maken, bieden daarvoor aanknopingspunten.

We hebben een tabel opgesteld van mogelijke interventies in aanbevelingsalgoritmen die platformen kunnen doen (zie tabel 7 hierna). Sommige interventies kunnen al deels door platformen zijn doorgevoerd. Andere zijn alleen nog getoetst in wetenschappelijke experimenten of geopperd door experts.

De tabel laat zien dat er aan aanbevelingsalgoritmen nog erg veel te veranderen valt met mogelijke positieve effecten voor gebruikers, het publieke debat en het afnemen van het risico op inmenging in verkiezingen.

Het parlement kan platformen vragen om van al deze interventies aan te geven of ze de suggesties van experts hebben overwogen en getest en wat hun afwegingen zijn om de aanpassing niet door te voeren. Het aandringen op of verplichten tot aanpassingen om de zeggenschap van gebruikers te vergroten kan ook worden meegenomen in toekomstige wetgeving, zoals de Europese Digital Fairness Act waarvoor de onderhandelingen nu bezig zijn.

Tabel 7 Suggesties uit onderzoek voor mogelijke aanpassingen aan aanbevelingsalgoritmen

Aanpassing	Wat is het?	Voorbeelden
Rangschikking gebaseerd op het overbruggen van groepen mensen (Lasser & Poechhacker, 2025)	Platformen ontwerpen hun aanbevelingsalgoritmen met andere doelen, zoals het vergroten van vertrouwen tussen groepen mensen, in plaats van het doel van zoveel mogelijk engagemen.t	Recent wetenschappelijk onderzoek naar *briging based ranking.*
Circuit breakers (Snap, 2025b) (Pathak & Spezzano, 2024)	Extra controle voordat informatie viral kan gaan.	Snap geeft aan dat inhoud eerst door menselijke moderatoren wordt bekeken voordat het aangeraden wordt aan een groot publiek.
Kwaliteitsscores (Cunningham, 2023)	Content met een lagere mate van kwaliteit, zoals met generatieve AI-gegenereerde beelden, krijgt minder gewicht door het aanbevelingsalgoritme.	
Diversificatie-algoritmen (Bengani, 2023)	Diversificatie-algoritmen kunnen bij het filteren en rangschikken van inhoud zorgen voor een diverse set aan inhoud.	Platformen geven aan dit te doen, maar hoe zwaar diversiteit weegt in de aanbevelingen is onbekend. Diversificatie van inhoud kan op veel verschillende manieren: op onderwerp, geografische afkomst, bron, type media en populariteit.
Gebruikersfeedback op inhoud zwaarder laten wegen (Cunningham et al., 2025; Stray et al., 2024)	Door gebruikers explicieter te bevragen, bijvoorbeeld door gebruikerssurveys, krijgen platformen meer inzicht in de voorkeuren van gebruikers. Dat terwijl aanbevelingsalgoritmes nu vaak uitgaan van afgeleide voorkeuren, zoals likes en kijktijd.	Wanneer mensen veel vergelijkbare inhoud te zien krijgen, zeker in verkiezingstijd, zouden platformen gebruikers actief kunnen bevragen of deze inhoud waardevol voor hen is en wat de inhoud doet met hun attitudes over verkiezingen.
Downranking van inhoud (Piccardi et al., 2025)	In plaats van inhoud te verwijderen, kunnen platformen borderline-inhoud of antidemocratische inhoud minder gewicht geven in hun aanbevelingsalgoritmen.	Een recent experiment van onderzoekers laat zien dat het lager rangschikken van *antidemocratische* en *polariserende* berichten op X een positief effect had op hoe gebruikers denken over mensen met andere politieke overtuigingen.

Aanpassing	Wat is het?	Voorbeelden
Alle inhoud (ook comments) handmatig goedkeuren (Ribeiro et al., 2023)	Door comments en inhoud handmatig goed te keuren, lijken comments die niet aan de regels te voldoen, te verminderen. Dat kan de kwaliteit van het gesprek op social media vergroten.	In Facebookgroepen bestaat de mogelijkheid voor moderatoren om alle comments handmatig goed te keuren.
Community notes (Chuai et al., 2024)	Door de contextuele informatie die gebruikers geven bij een post te gebruiken als signaal voor het aanbevelingsalgoritme, zou de verspreiding van misleidende inhoud kunnen afnemen. Meer onderzoek hiernaar is nodig.	Onderzoek laat gemengde resultaten zien van het effect van *community notes* op de verspreiding van informatie op Twitter (tegenwoordig X).
Bereik aanpassen op basis van inschatting betrouwbaarheid account (Fernández et al., 2021; Stray et al., 2024)	In de periode voorafgaand en tijdens verkiezingen zouden platformen de gewichten van hun aanbevelingsalgoritme kunnen aanpassen, op basis van intern onderzoek naar accounts. Accounts met een lage betrouwbaarheidsscore zouden minder aanbevolen kunnen worden.	Nieuwe accounts en accounts die in de periode vlak voor een verkiezing actief worden, zouden minder aanbevolen kunnen worden in verkiezingstijd, of pas na een menselijke check op het type content dat wordt gedeeld.
Gebruikers informeren na interactie met inhoud (Truong et al., 2024)	Platformen kunnen gebruikers die veel engagement aangaan met accounts die *lage kwaliteit* of verwijderde inhoud delen waarschuwen en helpen misleidende inhoud te herkennen.	

Tabel samengesteld door Rathenau Instituut

Bij ontwerpaanpassingen moet vaak een afweging worden gemaakt tussen voordelen en nadelen. Ook aanpassingen kunnen effect hebben op fundamentele rechten van gebruikers, of ze kunnen de economische belangen van platformen schaden (Lubin et al., 2024).

Neem bijvoorbeeld het beperken van de zichtbaarheid van nieuwe accounts. Dit kan effectief zijn om het moeilijker te maken om in aanloop naar verkiezingen met een netwerk van nepaccounts bewust te misleiden. Het nadeel hiervan is dat *alle* nieuwe accounts minder bereik krijgen.

Of neem het sturen op minder op engagement gebaseerde aanbevelingsalgoritmen waardoor hyperactieve accounts minder bereik krijgen, maar gebruikers ook minder tijd doorbrengen op socialmediaplatformen.

Daarom is het nodig effectiviteit en impact van maatregelen te kunnen monitoren, zodat die ook kunnen worden bijgestuurd. Socialmediaplatformen lenen zich gelukkig goed voor deze vorm van op onderzoek gestoelde regulering, omdat platformen in potentie goede manieren hebben om het effect van maatregelen te testen (Lubin et al., 2024).

Hoe onderstaande voorgestelde aanpassingen zouden uitwerken en of voordelen opwegen tegen eventuele negatieve bijeffecten, is onbekend. Onderzoekers kunnen momenteel namelijk niet meekijken bij platformen. Zie hier ook de wenselijkheid om de informatiepositie van onder andere onderzoekers ten opzichte van platformen te versterken (paragraaf 6.2.2).

Intervenieer in governance en verdienmodel

De schaal waarop de huidige online platformen opereren is enorm. De eigenaarsstructuur en governance (hoe een online omgeving wordt bestuurd) van grote online platformen zorgt er dan ook voor dat de beslissingen van een kleine groep bedrijven en hun eigenaren grote invloed hebben op wat mensen wereldwijd online zien en kunnen delen (Rathenau Instituut, 2025a). Zij hebben macht over welke typen inhoud relatief versterkt wordt ten opzichte van andere inhoud, en hun ontwerpkeuzes scheppen mogelijkheden of beperkingen tot inmenging op hun platformen.

In de begindagen van het internet was de macht om online omgevingen in te richten anders verdeeld dan nu. In zijn boek *The Modem World: A Prehistory of Social Media* beschrijft Kevin Driscoll bijvoorbeeld hoe bulletinboardsystemen op een aantal manieren anders waren dan hoe we nu het internet gebruiken (Driscoll, 2022). Online omgevingen waren vaak lokaal, met veel invloed van de gemeenschap die ze gebruikten.

Inmiddels ontvangen grote platformbedrijven veruit het grootste deel van de advertentie-inkomsten waar ook journalistiek mee wordt gefinancierd. Hierdoor verschraalt het journalistieke aanbod (Digital News Report Nederland 2025, 2025, p. 40). Dat terwijl de democratie juist gebaat is bij een pluriforme en vrije informatievoorziening.

Een vergaande route is het ingrijpen op het advertentiegedreven verdienmodel van socialmediaplatformen. Dat verdienmodel is een belangrijke reden voor de opkomst van op engagement gebaseerde aanbevelingsalgoritmen. Hoe langer mensen op een platform blijven, hoe meer advertentie-inkomsten er worden verdiend. Ingrijpen in het verdienmodel heeft uiteraard ook nadelen, zoals een stijging van kosten die

aan de gebruikers wordt doorgerekend. Toch zou deze afweging op tafel moeten liggen. De Digital Fairness Act biedt hiervoor aanknopingspunten.[21]

Zet in op decentralisering en interoperabiliteit
De overheid kan een pluriform en decentraal landschap van online omgevingen stimuleren. Bij decentrale omgevingen ligt de controle niet meer bij één partij, maar ligt deze in handen van verschillende (kleinere) partijen. Dit geeft meer controle en autonomie aan gebruikers. Het rapport *Inclusief online* (Rathenau Instituut, 2025a, p. 101) gaat dieper in op het stimuleren van alternatieven voor dominante socialmediaplatformen.

Alternatieve platformen geven geen garantie tegen inmenging. Een pluriform en decentraal landschap van online omgevingen versterkt echter wel de onderhandelingsruimte van de overheid ten opzichte van de huidige dominante platformen. Ook vergroten alternatieven de keuzevrijheid voor gebruikers. Bovendien verkleinen ze de impact van één of enkele platformen.

Socialmediaplatformen beheren nu het hele proces voor het verspreiden en zien van inhoud op hun platformen. Wat als meerdere partijen in dit proces betrokken kunnen zijn? Mensen zouden dan verschillende aanbevelingsalgoritmen van andere partijen kunnen kiezen. Of verschillende vormen van contentmoderatie. Wetenschappers noemen dit idee *middlewares*: het openbreken van socialmediaplatformen zodat verschillende partijen ook een rol kunnen spelen in de governance van platformen (Hogg & DiResta, 2024).

Door andere partijen ook aanbevelingsalgoritmen te laten aanbieden, zouden gebruikers van Instagram, TikTok of Snap bijvoorbeeld kunnen kiezen hun tijdlijn te laten cureren door een landelijke krant, hun lokale sportvereniging of een andere groep mensen die ze vertrouwen. Die kunnen dan hun eigen parameters bepalen: welk type inhoud krijgt bij ons voorrang?

Hetzelfde kan gelden voor contentmoderatie die nu gecentraliseerd door grote platformen wordt gedaan. Contentmoderatie kan ook meer door gemeenschappen zelf gedaan worden. Dat kan de legitimiteit en de kwaliteit van beslissingen over inhoud vergroten (Zuckerman, 2023). Het introduceren van *middlewares* zou ook nieuwe verdienmodellen kunnen opleveren voor partijen buiten de grote gecentraliseerde platformen.

21 Zie ook de inbreng van het Rathenau Instituut van 23 oktober 2025 bij de Europese Commissie over de Digital Fairness Act: https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/14622-Digital-Fairness-Act/F33095380_en

Het verdient de aanbeveling om in Europees verband uit te zoeken hoe en of *middlewares*, kunnen zorgen voor betere bescherming tegen pogingen tot inmenging via social media. Verschillende organisaties, waaronder de economische denktank Bruegel, raden aan om interoperabiliteitseisen uit te breiden om burgers gemakkelijker te laten overstappen (Scott Morton, 2024).[22]

6.2.2 Informatiepositie versterken

Uit ons onderzoek blijkt dat de informatiepositie van toezichthouders, onderzoekers, journalisten, politici en beleidsmakers ten aanzien van platformen te wensen over laat. Andere partijen dan platformen (zoals de overheid, de wetenschap of de journalistiek) kunnen onvoldoende onafhankelijk vaststellen of er sprake is geweest van inmenging tijdens de verkiezingen. Hoe meer kennis beschikbaar is voor externe deskundigen, hoe beter en tijdiger inmenging gedetecteerd kan worden. Dergelijk publiek toezicht is tenslotte ook essentieel omdat de kwalificatie van inmenging gemakkelijk misbruikt kan worden voor politieke doeleinden (zie hoofdstuk 2).

Binnen de richting 'informatiepositie versterken' geven we inzicht in het type informatie dat externe deskundigen van platformen nodig hebben om grip te krijgen op inmengingsactiviteiten en stellen we een onafhankelijk detectie-orgaan voor dat zelf ook proactief inmengingsactiviteiten kan aanpakken.

Maak inzicht mogelijk in platformdata voor overheid, toezichthouders en onderzoekers
Meer inzicht in de wijze waarop inhoud in onze tijdlijn verschijnt, is nodig. Denk aan (precisering van) het contentmoderatiebeleid, de kwaliteitsscores die aan inhoud wordt gegeven, bedrijfsregels voor rangschikken van inhoud en de gewichten die platformen in hun aanbevelingsalgoritmen toepassen (zie hoofdstuk 3). Met behulp van dit type informatie kunnen externe partijen beter meekijken met de maatregelen die platformen daadwerkelijk nemen en in hoeverre deze effectief zijn. Hetzelfde geldt voor de inhoud die op platformen wordt geplaatst. Platformen spelen, samen met de afzenders, ook een rol in het transparant maken van dergelijke informatie. Verschillende maatregelen kunnen de informatiepositie van beleidsmakers, politici, toezichthouders, onderzoekers en journalisten verbeteren.

Volgens de DSA moeten platformen hun inhoud toegankelijk maken voor toezichthouders en onafhankelijke onderzoekers. De mate van toegang tot

22 Concreet valt te overwegen om in de herziening van de Digitale marktenverordening artikel 7 te verbreden naar andere online omgevingen.

systemen verschilt, waarbij toezichthouders de grootste toegang krijgen. Experts geven aan dat de informatie die platformen tot nu toe delen nog te beperkt is om goed grip te krijgen op de uitwerking van aanbevelingsalgoritmen en inmengingsactiviteiten (zie hoofdstuk 5).

We doen hieronder voorstellen die de informatiepositie van derden zouden kunnen vergroten en die bijdragen aan de verwezenlijking van de doelen van verschillende bepalingen uit de DSA, zoals artikel 34, 35, 37 en 40.

Ook het coalitieakkoord kan aanknopingspunten geven om platformen te laten aantonen dat hun platformen niet bijdragen aan inmenging in verkiezingen, het verspreiden van misleidende content en het zaaien van twijfel over democratische processen.

Realtimesteekproeven
Inzicht in de rol en werking van aanbevelingsalgoritmen helpt om een goed beeld te krijgen van systeemrisico's. Dat kan door realtimesteekproeven te publiceren van openbare inhoud die het meest wordt verspreid en inhoud die het meeste engagement genereert. Inzicht kan ook worden vergroot door in realtime representatieve steekproeven te verrichten van de openbare inhoud die wordt geconsumeerd tijdens een typische gebruikerssessie op het platform op een willekeurig moment (Cooper & Chapman, 2025).

Verspreiding van inhoud
Inzicht in het aantal mensen en de typen groepen waarbij reeds verwijderde inhoud terecht is gekomen, kan helpen.

Ontwerp van aanbevelingsalgoritmen
Inzicht in de technische specificaties van de architectuur van de aanbevelingsalgoritmen (*recommender systems*) van platformen is belangrijk (Panoptykon et al., 2024). Denk aan:

a. Wat is het type algoritme en wat zijn de hyperparameters? Wat is de grootte van een gebruikt neuraal netwerk?
b. Wat is de inputdata en het relatieve belang van verschillende typen inputdata? Voor lineaire regressies: wat zijn de gewichten?
c. Hoe wordt de inputdata gepresenteerd en wat is de logica achter gebruikte *embedding spaces*?
d. Wat is de *loss function* van gebruikte modellen?
e. Wordt trainingsdata door mensen gelabeld? Welke partijen zijn op welke manier betrokken in dit proces van labelen?
f. Welke inspanningen worden geleverd om voor mensen begrijpelijke machinelearningsystemen te bouwen? Wat zijn de resultaten van de ontwikkelde algoritmische-interpretatietools?

Daarnaast is het voor een goed begrip van de precieze werking van aanbevelingsalgoritmen van specifieke platformen nodig om inzicht te hebben in de 'afweging' die platformen maken in het ontwerp van hun aanbevelingsalgoritmen, inclusief de data zij als input gebruiken en de gewichten die zij toekennen bij hun aanbevelingssystemen (KGI Expert Working Group on Recommender Systems, 2025). Ook inzicht in de maatstaven die platformen gebruiken om het ontwerp van hun aanbevelingssystemen te evalueren zou behulpzaam zijn (KGI Expert Working Group on Recommender Systems, 2025).

Resultaten van interne A/B-testen van platformen
Inzicht in de effecten van (constante) wijzigingen aan het ontwerp van aanbevelingssystemen kan verkregen worden door A/B-testen. Veel platformen gebruiken voor A/B-testen een holdout-groep, een controlegroep van gebruikers die niet wordt blootgesteld aan de wijzigingen. Langdurige holdout-experimenten geven veel inzicht in het effect van ontwerpkeuzes op gebruikers en zouden op geaggregeerde wijze gepubliceerd kunnen worden. Een verplichting hiertoe zou platformen ook kunnen prikkelen om hun ontwerp beter aan te laten sluiten op behoeften van gebruikers (Cooper & Chapman, 2025).

Inspanningen rond verkiezingstijd
Inzicht in toegankelijke informatie over platforminspanningen rondom verkiezingstijd, bijvoorbeeld via pre- en postelectorale evaluatie, kan helpen bij het versterken van de informatiepositie.

Standaardisatie van toegang en rapportages
Het is tenslotte van belang om in te zetten op standaardisatie voor toegang voor onderzoekers en toezichthouders alsook standaardisatie van identificatie, mitigatie en rapportage van systeemrisico's. Effectieve risicobeperking vereist transparante meetcriteria en evaluatiemethoden, ondersteund met datagedreven onderbouwing die het mogelijk maken om de risico's en effectiviteit van maatregelen in de loop der tijd voor buitenstaanders te beoordelen. Toezichthouders kunnen de zeer grote platformen en zoekmachines volgens de DSA aanspreken op de kwaliteit van risicoanalyses en inzetten op een lerende praktijk.

Institutionaliseer onafhankelijke, transparante en proactieve detectie van inmenging
Inmenging vraagt om meer institutionele inbedding van detectie en bestrijding van inmengingsactiviteiten. Voor het melden van inhoud die de verkiezingen zouden kunnen schaden heeft de overheid een *verkiezingsflaggerstatus*. Maar er is ook behoefte aan meer zicht op inmengingsactiviteiten op sociale media vanuit de overheid. Van belang is daarbij dat er duidelijke waarborgen zijn tegen politieke invloed en censuur.

Hiervoor geven we twee handelingsopties mee. De eerste suggestie is dat de overheid meer signaleringsmogelijkheden zou kunnen krijgen. De overheid heeft *verkiezingsflaggerstatus* bij de grote platformen (zie hoofdstuk 2), die ze bewust terughoudend inzet. Maar het signaleren van content mag alleen op aanbeveling van derden, niet op basis van eigen onderzoek. Het valt te overwegen de overheid ook meer mogelijkheden te geven om proactief inhoud te signaleren die de verkiezingen ondermijnt. Het is zaak de juiste balans te vinden tussen het verruimen van handelingsmogelijkheden én het bewaken van de vrijheid van meningsuiting. Dat is in lijn met ons Bericht aan het parlement *Desinformatie bestrijden, censuur vermijden* (Rathenau Instituut, 2020).

De tweede suggestie is dat er ingezet kan worden op de ontwikkeling van een onafhankelijk orgaan dat inmenging kan signaleren. Laat het orgaan nauw samenwerken met onafhankelijke experts zoals desinformatie-experts, platformonderzoekers, computerwetenschappers, cybersecurity-experts en juristen. Ten tijde van schrijven van dit rapport loopt bij het Ministerie van Binnenlandse Zaken en Koninkrijksrelaties een verkenning voor de oprichting van een desinformatiedetectieorgaan. Dat zou hierin kunnen voorzien.

Inmengende actoren zijn inventief. Inmengingsactiviteiten zijn altijd in beweging en springen in op nieuwe ontwikkelingen, zoals de opkomst van chatbots, personalisatie van nieuws en microbetalingen aan influencers via nauwelijks te modereren livestreams. Er is dus blijvend onderzoek en monitoring nodig.

Voor de monitoring en detectie van inmengingsactiviteiten op socialmediaplatformen wordt sterk geleund op wetenschappers en ngo's. Het valt te overwegen dit soort partijen structureel te financieren om dit onderzoek systematisch uit te voeren.

6.2.3 Weerbaarheid vergroten

De handelingsperspectieven hierboven zijn enerzijds gericht op platformen en anderzijds op het verbeteren van de informatiepositie over inmenging via platformen. Deze handelingsperspectieven zijn sterk afhankelijk van handhaving en de medewerking van socialmediabedrijven. Het is in de context van inmenging via platformen ook zinvol te kijken naar interventies die niet afhankelijk zijn van socialmediaplatformen en die de weerbaarheid van de samenleving als geheel vergroten.

Hieronder gaan we in op verschillende manieren om de samenleving weerbaarder te maken tegen inmengingsactiviteiten.

Versterk verkiezingsprocedures
Inmengingspogingen kunnen leiden tot onenigheden over de verkiezingsuitslag. Dat maakt het belang van beoordeling over het eerlijk verloop van verkiezingen, die niet kwetsbaar zijn voor politieke invloed, essentieel.

Neem daarom maatregelen om de onafhankelijkheid bij besluiten over verkiezingsgeschillen te versterken. Zorg ervoor dat er duidelijke criteria zijn voor het overgaan tot hertelling en herstemming. Onderzoek of de termijn voor het vaststellen van onregelmatigheden tijdens verkiezingen kan worden verruimd aangezien inmenging via social media tijd kost om te bewijzen. Evalueer systematisch in hoeverre verkiezingsprocedures adequaat zijn toegerust op inmengingsactiviteiten en stel deze zo nodig bij.

Bescherm de huidige wetgeving en verhelder waar nodig
Heldere normstelling, consistente toepassing van bestaande wetgeving en effectief toezicht zijn nodig om inmengingsactiviteiten het hoofd te bieden.

Opvallend is dat AI-systemen die worden gebruikt voor het beïnvloeden van de uitslag van een verkiezing, referendum of stemgedrag niet verboden zijn, maar in de hoog-risico-categorie vallen. Het is momenteel niet duidelijk wat precies wordt bedoeld met AI-systemen. Zorg ervoor dat toezichthouders voldoende capaciteit hebben om te verduidelijken welke AI-systemen onder de categorie verboden en

onder de categorie hoog risico van de AI-verordening vallen. Ook kan uitgezocht worden in hoeverre dergelijke systemen verenigbaar zijn met de AVG en VPR.

Ondersteun onderzoekers en maatschappelijke organisaties en toezichthouders met voorbeelden en casestudy's, bijvoorbeeld over de wijze waarop AI-systemen stemgedrag kunnen beïnvloeden en maatschappelijke schade kunnen aanrichten, ter ondersteuning van de richtlijnen die nu worden opgesteld. Ontwikkel en test templates voor risicobeoordelingen van het gebruik van AI bij verkiezingen.

Stuur daarnaast aan op, of onderzoek de mogelijkheden voor, een moratorium op het gebruik van AI-systemen in verkiezingscampagnes om een beter inzicht te krijgen in de maatschappelijke en politieke impact en onderzoek.

Versterk ook het toezicht op de AI-verordening, DSA, AVG en VPR met het oog op inmenging. Zet in op snelle uitvoerbaarheid van de VPR in Nederland en zorg er ook voor dat platformen zich houden aan eigen *community guidelines* omtrent betaalde politieke inhoud. Zoek daarbij samenwerking op met andere Europese toezichthouders.

Investeer in onafhankelijke journalistiek
Investeer in een pluriform medialandschap, inclusief een publieke omroep, om verschraling van de onafhankelijke informatievoorziening tegen te gaan. Onafhankelijke journalistiek vervult een signaalfunctie in de samenleving en is een van de controlerende machten in een democratie. Journalistiek en onderzoek functioneren onafhankelijk van overheid en bedrijfsleven en kunnen misstanden in de samenleving aankaarten. Onafhankelijk onderzoek vervult hiermee een belangrijke rol in het informatie-ecosysteem.

Investeer in technologisch burgerschap en mediawijsheid
De DSA heeft de positie van de gebruikers ten opzichte van grote socialmediaplatformen verbeterd. Het is echter de vraag of gebruikers hiervan op de hoogte zijn. Digitale rechten gaan onder andere over toegankelijkheid voor gebruikers om melding te maken van inhoud die in strijd is met platformregels. En ook over het recht om te weten wat er met een melding gebeurt. De overheid zou bewustwording bij burgers over hun digitale rechten kunnen bevorderen.[23]

Als gebruikers van socialmediaplatformen zich realiseren wat de mogelijkheden voor inmenging zijn, zou het kunnen dat inmenging minder effectief wordt. Mensen herkennen mogelijk de inmengingsactiviteiten. Waarschuwingen voor inmenging kunnen er echter ook toe leiden dat mensen alle informatie gaan wantrouwen of dat

23 Zie bijvoorbeeld deze campagne van Bits of Freedom om bewustzijn van digitale rechten te vergroten: https://www.jouwplatformrechten.nl/

mensen minder vertrouwen krijgen over de eerlijkheid van verkiezingen (Altay et al., 2025).

Onderwijs over hoe een open democratie werkt, kan helpen de democratie te versterken. Hoe wordt er in Nederland gezorgd dat verkiezingen eerlijk verlopen? Welke waarborgen kennen journalistieke redacties? En waar is juiste informatie over verkiezingen? En hoe komen burgers tot politieke oordeelsvorming? Deze vragen kunnen onderwerpen van gesprek leveren waarbij de werking van een democratische samenleving centraal staat. Daarmee kunnen ze de weerbaarheid tegen inmengingsactiviteiten vergroten.

Literatuurlijst

404 Media. (2025, 30 september). *Google Just Removed Seven Years of Political Advertising History from 27 Countries*. 404 Media. https://www.404media.co/google-ad-transparency-european-union

10767307 CV FORM 23-13934, ECLI:NL:RBAMS:2024:3980 (Rb. Amsterdam 5 juli 2024). https://deeplink.rechtspraak.nl/uitspraak?id=ECLI:NL:RBAMS:2024:3980

AIVD. (2025). *Jaarverslag 2024*. https://www.aivd.nl/documenten/jaarverslagen/2025/04/24/jaarverslag-2024

Alle grote media in Noorwegen trappen in tweets van Russische trollen. (2020, 3 maart). https://nos.nl/artikel/2325674-alle-grote-media-in-noorwegen-trappen-in-tweets-van-russische-trollen

Altay, S., Hoes, E., & Wojcieszak, M. (2025). Following news on social media boosts knowledge, belief accuracy and trust. *Nature Human Behaviour*, *9*(9), 1833-1842. https://doi.org/10.1038/s41562-025-02205-6

Autoriteit Persoonsgegevens. (2024, 3 september). *AP legt Clearview boete op voor illegale dataverzameling voor gezichtsherkenning*. https://www.autoriteitpersoonsgegevens.nl/actueel/ap-legt-clearview-boete-op-voor-illegale-dataverzameling-voor-gezichtsherkenning

Autoriteit Persoonsgegevens. (2025, 21 oktober). *AP waarschuwt: chatbots geven vertekend stemadvies*. https://www.autoriteitpersoonsgegevens.nl/actueel/ap-waarschuwt-chatbots-geven-vertekend-stemadvies

Bandy, J., & Lazovich, T. (2023). Exposure to Marginally Abusive Content on Twitter. *Proceedings of the International AAAI Conference on Web and Social Media*, *17*, 24-33. https://doi.org/10.1609/icwsm.v17i1.22123

Beemsterboer, T. B. R. P. T., & Beemsterboer, T. B. R. P. T. (2026, 17 februari). Israël ziet TikTok als het belangrijkste geopolitieke strijdperk van deze tijd. *NRC*. https://www.nrc.nl/nieuws/2026/02/17/israel-ziet-tiktok-als-het-belangrijkste-geopolitieke-strijdperk-van-deze-tijd-a4919546

Bellogín, A., Castells, P., & Cantador, I. (2017). Statistical biases in Information Retrieval metrics for recommender systems. *Information Retrieval Journal*, *20*(6), 606-634. https://doi.org/10.1007/s10791-017-9312-z

Bengani, P. (2023, 17 november). What is Media Diversity and Do Recommender Systems Have It? *Understanding Recommenders*. https://medium.com/understanding-recommenders/what-is-media-diversity-and-do-recommender-systems-have-it-b2c86be9f08f

Berger, J., & Milkman, K. L. (2012). What Makes Online Content Viral? *Journal of Marketing Research*, *49*(2), 192-205. https://doi.org/10.1509/jmr.10.0353

Berzina, K., & Soula, E. (2020). *Conceptualizing Foreign Interference in Europe*. Alliance for securing democracy.

Bouchaud, P. (2024). Algorithmic Amplification of Politics and Engagement Maximization on Social Media. In H. Cherifi, L. M. Rocha, C. Cherifi, & M. Donduran (Red.), *Complex Networks & Their Applications XII* (pp. 131-142). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-53503-1_11

Brady, W. J., Jackson, J. C., Lindström, B., & Crockett, M. J. (2023). Algorithm-mediated social learning in online social networks. *Trends in Cognitive Sciences*, *27*(10), 947-960. https://doi.org/10.1016/j.tics.2023.06.008

Bruns, A. (2019). Filter bubble. *Internet Policy Review*, *8*(4). https://doi.org/10.14763/2019.4.1426

Bunskoek, J., van den Berg, E., & Stoker, H. (2025, 25 november). *Trollenlegers uit buitenland versterkten politieke en opruiende berichten rond verkiezingen*. RTL.nl.

Chuai, Y., Tian, H., Pröllochs, N., & Lenzini, G. (2024). Did the Roll-Out of Community Notes Reduce Engagement With Misinformation on X/Twitter? *Proceedings of the ACM on Human-Computer Interaction*, *8*(CSCW2), 1-52.

Cooper, A., & Chapman, P. (2025, 19 mei). *Making Recommender Systems Work for People: Turning the DSA's Potential into Practice - DSA Observatory*. https://dsa-observatory.eu/2025/05/19/making-recommender-systems-work-for-people/

Cunningham, T. (2023, 8 mei). *Ranking by Engagement | Tom Cunningham – Tom Cunningham*. https://tecunningham.github.io/posts/2023-04-28-ranking-by-engagement.html

Cunningham, T., Pandey, S., Sigerson, L., Stray, J., Allen, J., Barrilleaux, B., Iyer, R., Kothari, M., Rezaei, B., Kairam, S., & Milli, S. (2025). Ranking by engagement and non-engagement signals: Learnings from industry. *Annals of the New York Academy of Sciences*, *1551*(1), 19-32. https://doi.org/10.1111/nyas.15399

D66, VVD en CDA. (2026). *Aan de slag. Bouwen aan een beter Nederland*. https://www.kabinetsformatie2025.nl/documenten/2026/01/30/aan-de-slag---coalitieakkoord-2026-2030

Data School Universiteit Utrecht. (2023). *Spelen met vuur - Hoe de wisselwerking tussen de Tweede Kamer en sociale media woede aanwakkert.* Universiteit Utrecht. https://dataschool.nl/2023/10/06/spelen-met-vuur/

Digital News Report Nederland 2025 (Digital News Report, p. 52). (2025). Commissariaat voor de Media. https://www.cvdm.nl/wp-content/uploads/2025/06/Digital-News-Report-2025-1.pdf

Directorate-General for Research and Innovation. (2022). *Tackling R&I foreign interference: staff working document*. Publications Office of the European Union. https://data.europa.eu/doi/10.2777/513746

DISARMFoundation. (z.d.). *01_AMITT_TTP_Guide.pdf*. GitHub; DISARMFoundation. Geraadpleegd 22 januari 2026, van https://github.com/DISARMFoundation/DISARMframeworks/blob/20ac4ea378578e1fecf4c9014f525bfc27a9b66c/generated_pages/techniques/T0104.001.md

Driscoll, K. (2022). *The Modem World: A Prehistory of Social Media* (First Edition). Yale University Press.

Ecker, U., Roozenbeek, J., & Lewandowsky, S. (2024). *Misinformation remains a threat to democracy*.

Egelhofer, J. L., Aaldering, L., Eberl, J.-M., Galyga, S., & Lecheler, S. (2020). From Novelty to Normalization? How Journalists Use the Term "Fake News" in their Reporting. *Journalism Studies*, *21*(10), 1323-1343. https://doi.org/10.1080/1461670X.2020.1745667

Eijsvoogel, J. (2025, 2 oktober). Meta wekt verwarring en woede bij overheid en ngo's met advertentieverbod voor 'maatschappelijke kwesties'. *NRC*. https://www.nrc.nl/nieuws/2025/10/02/meta-wekt-verwarring-en-woede-bij-overheid-en-ngos-met-advertentieverbod-voor-maatschappelijke-kwesties-a4908280

EU in touch with Telegram as it nears criterion for EU tech rules. (2024, 28 mei). Reuters. https://www.reuters.com/technology/eu-touch-with-telegram-it-nears-criterion-eu-tech-rules-2024-05-28

European Commission. (2023). *Commission Staff Working Document – Executive Summary of the Impact Assessment Report* (Commission Staff Working Document (SWD) SWD/2023/663 final). European Commission. https://eur-lex.europa.eu/eli/impact_assessment/52023SC0663/EN

European Commission. (2024a). *Guidelines on mitigating systemic risks for electoral processes pursuant to Article 35(3) of Regulation (EU) 2022/2065* ((C/2024/3014)). http://data.europa.eu/eli/C/2024/3014/oj

European Commission. (2024b, 30 april). *Commission opens formal proceedings against Facebook and Instagram under the Digital Services Act*. https://ec.europa.eu/commission/presscorner/detail/en/ip_24_2373

European Commission. (2024c, 16 mei). *Commission opens formal proceedings against Meta*. https://ec.europa.eu/commission/presscorner/detail/en/ip_24_2664

European Commission. (2024d, 5 augustus). *TikTok commits to permanently withdraw TikTok Lite Rewards*. https://ec.europa.eu/commission/presscorner/detail/en/ip_24_4161

European Commission. (2025a, 17 januari). *Commission addresses additional investigatory measures to X in the ongoing proceedings under the Digital Services Act*. https://digital-strategy.ec.europa.eu/en/news/commission-addresses-additional-investigatory-measures-x-ongoing-proceedings-under-digital-services

European Commission. (2025b). *Commission Guidelines on prohibited artificial intelligence practices established by Regulation (EU) 2024/1689 (AI Act)* (C(2025) 5052 final). European Commission. https://digital-strategy.ec.europa.eu/en/library/commission-publishes-guidelines-prohibited-artificial-intelligence-ai-practices-defined-ai-act

European Commission. (2025c, 24 oktober). *Commission preliminarily finds TikTok and Meta in breach of their transparency obligations* [Text]. https://ec.europa.eu/commission/presscorner/detail/en/ip_25_2503

European Commission. (2025d, 5 december). *Commission fines X €120 million under the Digital Services Act*. https://digital-strategy.ec.europa.eu/en/news/commission-fines-x-eu120-million-under-digital-services-act

European Commission. (2025e, 17 december). *Commission opens formal proceedings against TikTok under DSA* [European Commission]. https://ec.europa.eu/commission/presscorner/detail/en/ip_24_6487

European Commission. (2026, 6 februari). *Commission preliminarily finds TikTok's addictive design in breach of the Digital Services Act*. https://ec.europa.eu/commission/presscorner/detail/en/ip_26_312

European External Action Service. (2023). *1st EEAS Report on Foreign Information Manipulation and Interference Threats. Towards a Framework of Networked Defence*. European External Action Service. https://www.eeas.europa.eu/sites/default/files/documents/2023/EEAS-DataTeam-ThreatReport-2023..pdf

European Parliamentary Research Service. (2025). *TikTok and EU regulation: Legal challenges and cross-jurisdictional insights*.

FameUp – FameUp – Over 500+ nano and micro influencers activated in 24 hours. (z.d.). FameUp. Geraadpleegd 4 oktober 2025, van https://fameup.com/en/

Feenstra, W., & Sabel, P. (2025, 27 oktober). *Twee PVV-Kamerleden vallen met nepbeelden anoniem Timmermans aan, GroenLinks-PvdA doet aangifte*. de Volkskrant. https://www.volkskrant.nl/politiek/twee-pvv-kamerleden-vallen-met-nepbeelden-anoniem-timmermans-aan-groenlinks-pvda-doet-aangifte~b6958ea1/

Feifer, J. (2023, 26 juni). LinkedIn Changed Its Algorithms — Here's How Your Posts Will Get More Attention Now. *Entrepreneur*. https://www.entrepreneur.com/science-technology/linkedin-changed-its-algorithms-heres-how-your-posts/454728

Fernández, M., Bellogín, A., & Cantador, I. (2021). *Analysing the Effect of Recommendation Algorithms on the Amplification of Misinformation* (arXiv:2103.14748). arXiv. https://doi.org/10.48550/arXiv.2103.14748

Garnett, H. A., James, T. S., & Caal-Lam, S. (2025). *Electoral Integrity Global Report 2025* (PEI 11.0). Electoral Integrity PRoject. https://ueaeprints.uea.ac.uk/id/eprint/99835/1/Year_in_Elections_PEI_11_Report_FINAL.pdf

Gauthier, G., Hodler, R., Widmer, P., & Zhuravskaya, E. (2026). The political effects of X's feed algorithm. *Nature*. https://doi.org/10.1038/s41586-026-10098-2

Gillespie, T. (2022). Do Not Recommend? Reduction as a Form of Content Moderation. *Social Media + Society*, *8*(3), 20563051221117552. https://doi.org/10.1177/20563051221117552

Gleicher, N. (2019, 6 mei). Removing More Coordinated Inauthentic Behavior From Russia. *Meta Newsroom*.

GLOBSEC, Debunk.org, Institute for Strategic Dialogue, DEN Institute, & Alliance4Europe. (2025). *Election report. Assessment of Foreign Information Manipulation and Interference in the 2025 Czech Parliamentary Election*. https://fimi-isac.org/wp-content/uploads/2025/12/FRT-24_Globsec_Czech-Election-Report_12122025.pdf

Goanta, C. (2024, 17 november). Influencers & elections: An Eastern European campaign blueprint [Researchblog]. *Human Ads ERC*. https://humanads.eu/influencers-and-elections-an-eastern-european-campaign-blueprint

Goel, S., Anderson, A., Hofman, J., & Watts, D. J. (2016). The Structural Virality of Online Diffusion. *Management Science*, *62*(1), 180-196. https://doi.org/10.1287/mnsc.2015.2158

Google. (2012, 10 augustus). *YouTube Now: Why We Focus on Watch Time*. Blog.Youtube. https://blog.youtube/news-and-events/youtube-now-why-we-focus-on-watch-time/

Google. (2025a). *Information about Monthly Active Recipients under the Digital Services Act (EU)*. https://storage.googleapis.com/transparencyreport/report-downloads/pdf-report-24_2025-1-1_2025-6-30_en_v1.pdf

Google. (2025b). *EU Digital Services Act (EU DSA) Biannual VLOSE/VLOP Transparency Report*. https://storage.googleapis.com/transparencyreport/report-downloads/pdf-report-27_2025-1-1_2025-6-30_en_v1.pdf

Google. (2025c). *Search quality evaluator guidelines*. https://static.googleusercontent.com/media/guidelines.raterhub.com/en//searchqualityevaluatorguidelines.pdf

Google stopt vanaf september met politieke advertenties. (2025, 6 augustus). Tweakers. https://tweakers.net/nieuws/237800/google-stopt-vanaf-september-met-politieke-advertenties.html

Guess, A. M., Malhotra, N., Pan, J., Barberá, P., Allcott, H., Brown, T., Crespo-Tenorio, A., Dimmery, D., Freelon, D., Gentzkow, M., González-Bailón, S., Kennedy, E., Kim, Y. M., Lazer, D., Moehler, D., Nyhan, B., Rivera, C. V., Settle, J., Thomas, D. R., … Tucker, J. A. (2023). How do social media feed algorithms affect attitudes and behavior in an election campaign? *Science*, *381*(6656), 398-404. https://doi.org/10.1126/science.abp9364

Handvest van de Verenigde Naties, (1945). https://wetten.overheid.nl/BWBV0004143/1973-09-24

Haugen, F. [producer]. (2022). *FBarchive Document odoc2025266473w32 (Green Edition)*. Cambridge, MA: FBarchive [distributor]. https://fbarchive.org/user/doc/odoc2025266473w32

Heck, W., & Kouwenhoven, A. (2023, 7 juli). Dit schimmige bedrijf vernietigde reputaties van Europese moslims. *NRC*. https://www.nrc.nl/nieuws/2023/07/07/dit-schimmige-bedrijf-vernietigde-met-succes-de-reputaties-van-europese-moslims-a4169074

HEIO Consortium, Post-X Society, AI Forensics, Trollrensics, University of Amsterdam, & Justice for Prosperity. (2026). *Hybrid Election Integrity Observatory: Final Report - Monitoring the Dutch Parliamentary Elections, October 29th 2025*. HEIO Consortium. https://www.heio.nl/wp-content/uploads/2026/01/260115-HEIO-Final-Report.pdf

Hendrix, J. (2025, 7 januari). *Transcript: Mark Zuckerberg Announces Major Changes to Meta's Content Moderation Policies and Operations*. Tech Policy Press. https://techpolicy.press/transcript-mark-zuckerberg-announces-major-changes-to-metas-content-moderation-policies-and-operations

Hoboken, Joris van, Naomi Appelman, Anna van Duin, e.a. 2020. *WODC-onderzoek: Voorziening voor verzoeken tot snelle verwijdering van onrechtmatige online content*. Instituut voor Informatierecht, Universiteit van Amsterdam. https://pure.uva.nl/ws/files/87684586/3108_volledige_tekst_tcm28_464976.pdf.

Hogg, L., & DiResta, R. (2024). *Shaping the Future of Social Media with Middleware*. Georgetown University. https://doi.org/10.57928/NENS-7F25

Honingh, M., & van Ham, C. (2024). *Verkenning en verdieping democratische erosie en respons in Nederland* (Sustainable democracy). Ministerie van Binnenlandse Zaken en Koninkrijksrelaties. https://www.kennisopenbaarbestuur.nl/site/binaries/site-content/collections/documents/2024/04/12/verkenning-en-verdieping-democratische-erosie-en-respons-in-nederland/Verkenning+en+verdieping+democratische+erosie+in+Nederland+-+april++2024+finaal.pdf

Humprecht, E., Esser, F., & Van Aelst, P. (2020). Resilience to Online Disinformation: A Framework for Cross-National Comparative Research. *The International Journal of Press/Politics*, *25*(3), 493-516. https://doi.org/10.1177/1940161219900126

Huszár, F., Ktena, S. I., O'Brien, C., Belli, L., Schlaikjer, A., & Hardt, M. (2022a). Algorithmic amplification of politics on Twitter. *Proceedings of the National Academy of Sciences*, *119*(1), e2025334119. https://doi.org/10.1073/pnas.2025334119

Huszár, F., Ktena, S. I., O'Brien, C., Belli, L., Schlaikjer, A., & Hardt, M. (2022b). Algorithmic Amplification of Politics on Twitter. *Proceedings of the National Academy of Sciences*, *119*(1), e2025334119. https://doi.org/10.1073/pnas.2025334119

Irwin, G. A., & van Holsteyn, J. J. M. (2021). Keeping Our Feet Dry: Impediments to Foreign Interference in Elections in the Netherlands. *Election Law Journal: Rules, Politics, and Policy*, *20*(1), 54-69. https://doi.org/10.1089/elj.2020.0654

Jiang, J. A., Nie, P., Brubaker, J. R., & Fiesler, C. (2023). A Trade-off-centered Framework of Content Moderation. *ACM Trans. Comput.-Hum. Interact.*, *30*(1), 3:1-3:34. https://doi.org/10.1145/3534929

Justice for Prosperity. (2025). *Van botnetwerk tot beeldvorming - Beinvloedingspogingen rond de Nederlandse verkiezingen 2025* (p. 26). Justice for Prosperity. https://justiceforprosperity.org/wp-content/uploads/2026/01/Van-botnet-tot-beeldvorming-JfP-extern-1.pdf

Kamerstukken II, 35 165, nr. 102. (2026, 9 januari). Overheid.nl. https://zoek.officielebekendmakingen.nl/kst-35165-102.html

Kamerstukken II, 36 742, nr 8. (2025, 20 juli). Overheid.nl. https://zoek.officielebekendmakingen.nl/dossier/36742

Kamerstukken II, 22 112, nr. 4223. (2025, 12 december). https://zoek.officielebekendmakingen.nl/kst-22112-4223.html

Kamerstukken II, 26 643, nr. 1391. (2025, 4 september). Overheid.nl. https://zoek.officielebekendmakingen.nl/kst-26643-1391.html#ID-1211802-d36e130

Kamerstukken II, 26 643, nr. 1400. (2025, 24 september). Overheid.nl. https://www.tweedekamer.nl/kamerstukken/kamervragen/detail?id=2025D42593&did=2025D42593

Kamerstukken II, 30 821, nr. 230. (2024). Overheid.nl. https://zoek.officielebekendmakingen.nl/kst-30821-230

Kamerstukken II, 36 552, nr. 12. (2025, 4 juli). Overheid.nl. https://zoek.officielebekendmakingen.nl/kst-36552-12.html

KGI Expert Working Group on Recommender Systems. (2025). *Better Feeds: Algorithms That Put People First A How-To Guide for Platforms and Policymakers*. https://kgi.georgetown.edu/wp-content/uploads/2025/02/Better-Feeds_-Algorithms-That-Put-People-First.pdf

Kleinberg, J., Mullainathan, S., & Raghavan, M. (2024). The Challenge of Understanding What Users Want: Inconsistent Preferences and Engagement Optimization. *Management Science*, *70*(9), 6336-6355. https://doi.org/10.1287/mnsc.2022.03683

Knight Georgetown Institute. (2025). *Taxonomy of User Signals*.

Koebler, J. (2025, 24 november). *America's Polarization Has Become the World's Side Hustle*. 404 Media. https://www.404media.co/americas-polarization-has-become-the-worlds-side-hustle/

Kruschinski, S., & Votta, F. (2025). *CampAIn tracker Dashbord*. CampAIgntracker. https://www.campaigntracker.nl/

Lasser, J., & Poechhacker, N. (2025). Designing social media content recommendation algorithms for societal good. *Annals of the New York Academy of Sciences*, *1548*(1), 20-28. https://doi.org/10.1111/nyas.15359

like.vn. (z.d.). *Câu hỏi thường gặp (FAQ) - like.vn*. Like.vn ❤️ - SMM Panel Services - Cheapest API Provider - Viet Nam. Geraadpleegd 3 februari 2026, van https://like.vn/cau-hoi-thuong-gap

like.vn. (2026, 3 februari). *Mua like Instagram giá rẻ ❤Dịch vụ tăng, buff like IG giá rẻ (Like Thật)*. Like.vn ❤️ - SMM Panel Services - Cheapest API Provider - Viet Nam. https://like.vn/tang-like-instagram

Lin, Y., Liu, Y., Lin, F., Zou, L., Wu, P., Zeng, W., Chen, H., & Miao, C. (2024). A Survey on Reinforcement Learning for Recommender Systems. *IEEE Transactions on Neural Networks and Learning Systems*, *35*(10), 13164-13184. https://doi.org/10.1109/TNNLS.2023.3280161

LinkedIn. (2025). *Transparantierapport Wet inzake digitale diensten - augustus 2025*. https://content.linkedin.com/content/dam/help/tns/en/August-2025-Digital-Services-Act-Transparency-Report.pdf

Lubin, N., Mayberry, K., Moses, D., Revel, M., Thorburn, L., & West, A. (2024). Mapping the space of social media regulation. *MIT Science Policy Review*, *5*, 108-133. https://doi.org/10.38105/spr.sqyw1u2jf7

Macdonald, S., & Vaughan, K. (2024). Moderating borderline content while respecting fundamental values. *Policy & Internet*, *16*(2), 347-361. https://doi.org/10.1002/poi3.376

McKenzie, H., & Monga, S. (2024, 22 februari). Upgrading Substack's recommendation network [Substack newsletter]. *On Substack*. https://on.substack.com/p/substacks-recommendations-network

McWhinney, E. (1966). *Declaration on the Inadmissibility of Intervention in the Domestic Affairs of States and the Protection of their Independence and Sovereignty*. https://www.cambridge.org/core/product/identifier/S0002930000219969/type/journal_article

Mercadante, E. J., Tracy, J. L., & Götz, F. M. (2023). Greed communication predicts the approval and reach of US senators' tweets. *Proceedings of the National Academy of Sciences*, *120*(11), e2218680120. https://doi.org/10.1073/pnas.2218680120

Meta. (z.d.). *Community Standards | Transparency Center*. Geraadpleegd 15 januari 2026, van https://transparency.meta.com/policies/community-standards

Meta. (2025a). *Adversarial Threat Report Q2/Q3* (p. 45). Meta. https://dn720004.ca.archive.org/0/items/a-blueprint-for-content-governance-and-enforcement-facebook/A%20Blueprint%20for%20Content%20Governance%20and%20Enforcement%20_%20Facebook.pdf

Meta. (2025b). *Regulation (EU) 2022/2065 Digital Services Act Transparency Report for Facebook*.

Meta. (2025c). *Regulation (EU) 2022/2065 Digital Services Act Transparency Report for Instagram*.

Meta moet (voorlopig) chronologische tijdlijn als blijvende voorkeursoptie aanbieden. (2026, 22 januari). Mr. Online. https://www.mr-online.nl/meta-moet-voorlopig-chronologische-tijdlijn-als-blijvende-voorkeursoptie-aanbieden/

Midden in campagnetijd ineens geen politieke reclame meer op Facebook en Instagram: 'Tech bepaalt spelregels in onze verkiezingen'. (2025, 1 oktober). EenVandaag. https://eenvandaag.avrotros.nl/artikelen/midden-in-campagnetijd-ineens-geen-politieke-reclame-meer-op-facebook-en-instagram-tech-bepaalt-spelregels-in-onze-verkiezingen-161498

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). *Efficient Estimation of Word Representations in Vector Space* (arXiv:1301.3781). arXiv. https://doi.org/10.48550/arXiv.1301.3781

Milli, S., Carroll, M., Wang, Y., Pandey, S., Zhao, S., & Dragan, A. D. (2025). Engagement, user satisfaction, and the amplification of divisive content on social media. *PNAS Nexus*, *4*(3), pgaf062. https://doi.org/10.1093/pnasnexus/pgaf062

Miltenburg, E., Geurkink, B., Tunderman, S., Beekers, D., & den Ridder, J. (2022). *Burgerperspectieven 2022 | bericht 2*. SCP. https://www.scp.nl/documenten/2022/12/29/continu-onderzoek-burgerperspectieven---bericht-2-2022

Ministerie van BZK. (2025, 10 november). *Uitvoeringswet verordening transparantie en gerichte politieke reclame*. https://wetgevingskalender.overheid.nl/Regeling/WGK026963

Mugemangango t. Belgium, Nr. 310/15 (Europees Hof voor de Rechten van de Mens 10 juli 2022). https://hudoc.echr.coe.int/fre#{%22itemid%22:[%22002-12906%22]}

Narayanan, A. (2022, 15 december). TikTok's Secret Sauce. *Knight First Amendment Institute*. http://knightcolumbia.org/blog/tiktoks-secret-sauce

Narayanan, A. (2023). *Understanding Social Media Recommendation Algorithms*. Knight First Amendment Institute.

Nederlandse claimstichting begint rechtszaken tegen TikTok en X (voorheen Twitter). (2025, 5 februari). https://www.njb.nl/nieuws/nederlandse-claimstichting-begint-rechtszaken-tegen-tiktok-en-x-voorheen-twitter/

O'Carroll, L. (2025, 17 januari). EU asks X for internal documents about algorithms as it steps up investigation. *The Guardian*. https://www.theguardian.com/technology/2025/jan/17/eu-asks-x-for-internal-documents-about-algorithms-as-it-steps-up-investigation

Ohlin, J. D. (2017). Did Russian Cyber Interference in the 2016 Election Violate International Law? *Texas Law Review*. https://texaslawreview.org/russian-cyber-interference-2016-election-violate-international-law/

Ohlin, J. D., & Hollis, D. B. (Red.). (2021). *Defending Democracies: Combating Foreign Election Interference in a Digital Age* (1ste dr.). Oxford University PressNew York. https://doi.org/10.1093/oso/9780197556979.001.0001

O'Sullivan, D. (2019, 6 mei). *Be careful who you accept: Russian fake accounts targeted Facebook Groups | CNN Business*. CNN. https://www.cnn.com/2019/05/06/tech/facebook-groups-russia-fake-accounts

Ovadya, A. (2022). *Bridging-Based Ranking. How Platform Recommendation Systems Might Reduce Division and Strengthen Democracy*. https://www.belfercenter.org/sites/default/files/pantheon_files/files/publication/TAPP-Aviv_BridgingBasedRanking_FINAL_220518_0.pdf

Papakyriakopoulos, O., Serrano, J. C. M., & Hegelich, S. (2020a). Political communication on social media: A tale of hyperactive users and bias in recommender systems. *Online Social Networks and Media*, *15*, 100058. https://doi.org/10.1016/j.osnem.2019.100058

Papakyriakopoulos, O., Serrano, J. C. M., & Hegelich, S. (2020b). Political communication on social media: A tale of hyperactive users and bias in recommender systems. *Online Social Networks and Media*, *15*, 100058. https://doi.org/10.1016/j.osnem.2019.100058

Pariser, E. (2012). *The filter bubble: what the Internet is hiding from you*. Penguin books.

Pathak, R., & Spezzano, F. (2024). An empirical analysis of intervention strategies' effectiveness for countering misinformation amplification by recommendation algorithms. *European Conference on Information Retrieval*, 285-301.

Piccardi, T., Saveski, M., Jia, C., Hancock, J., Tsai, J. L., & Bernstein, M. S. (2025). Reranking partisan animosity in algorithmic social media feeds alters affective polarization. *Science*, *390*(6776), eadu5584. https://doi.org/10.1126/science.adu5584

Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL amending Regulations (EU) 2016/679, (EU) 2018/1724, (EU) 2018/1725, (EU) 2023/2854 and Directives 2002/58/EC, (EU) 2022/2555 and (EU) 2022/2557 as regards the simplification of the digital legislative framework, and repealing Regulations (EU) 2018/1807, (EU) 2019/1150, (EU) 2022/868, and Directive (EU) 2019/1024 (Digital Omnibus), (2025). https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52025PC0837

Quin, F., Weyns, D., Galster, M., & Silva, C. C. (2024). A/B testing: A systematic literature review. *Journal of Systems and Software*, *211*, 112011. https://doi.org/10.1016/j.jss.2024.112011

Raad van State. (2024). *Voorlichting over de inrichting van het systeem inzake de geschilbeslechting in het verkiezingsproces.* (W04.24.00045/I). https://open.overheid.nl/documenten/9b0b1cd5-64bb-4945-8000-2f0dcfa73562/file

Raad van State. (2025). *Uitvoeringswet verordening transparantie en gerichte politieke reclame.* https://www.raadvanstate.nl/adviezen/@152389/w04-25-00196/

Rathenau Instituut. (2020). *Desinformatie bestrijden, censuur vermijden*. https://www.rathenau.nl/nl/berichten-aan-het-parlement/desinformatie-bestrijden-censuur-vermijden

Rathenau Instituut. (2021). *Online ontspoord – Een verkenning van schadelijk en immoreel gedrag op het internet in Nederland*. (auteurs: Van Huijstee, M., Nieuwenhuizen, W., Sanders, M., Masson, E., & Van Boheemen, P.) https://www.rathenau.nl/nl/digitalisering/online-ontspoord

Rathenau Instituut. (2022). *Algoritmes afwegen – Verkenning naar maatregelen ter bescherming van mensenrechten bij profilering in de uitvoering.* (auteurs: Hamer J., Lemmens, A., & Kool, L.) https://www.rathenau.nl/nl/digitalisering/algoritmes-afwegen

Rathenau Instituut. (2023). *Rathenau Scan Generatieve AI*. (auteurs: Kool, L., Hamer, J., Hijstek, B., Van Eeden, Q., & Das, D.) https://www.rathenau.nl/nl/digitalisering/generatieve-ai

Rathenau Instituut. (2025a). *Inclusief online – Naar een ontwerp van apps en omgevingen waar mensen vrij en veilig kunnen zijn*. (auteurs: Nieuwenhuizen, W., Nieuwenhuis, T., Van Eeden, Q., & Van Huijstee, M.) https://www.rathenau.nl/nl/digitalisering/naar-een-nieuwe-verhouding-tot-technologiebedrijven/inclusief-online

Rathenau Instituut. (2025b). *De prijs van gratis internet – Richtingen voor toekomstig online-trackingbeleid.* (auteurs: Das, D., Wals, F., Hijstek, B., Lagendijk, V., & Kool, L., m.m.v. Gerritsen, J. & Nieuwenhuizen, W.) https://www.rathenau.nl/nl/digitalisering/naar-een-menswaardige-digitale-technologie/de-prijs-van-gratis-internet

Rathenau Instituut. (2025c). *Achter de macht van big tech – Verklarende factoren voor digitale macht.* (auteurs: Van Eeden, Q., Karstens, B., Roolvink, S., & Van Huijstee, M.) https://www.rathenau.nl/nl/digitalisering/naar-een-nieuwe-verhouding-tot-technologiebedrijven/achter-de-macht-van-big-tech

Rathje, S., Van Bavel, J. J., & van der Linden, S. (2021). Out-group animosity drives engagement on social media. *Proceedings of the National Academy of Sciences*, *118*(26), e2024292118. https://doi.org/10.1073/pnas.2024292118

Raza, S., Rahman, M., Kamawal, S., Toroghi, A., Raval, A., Navah, F., & Kazemeini, A. (2026). A comprehensive review of recommender systems: Transitioning from theory to practice. *Computer Science Review*, *59*, 100849. https://doi.org/10.1016/j.cosrev.2025.100849

Ribeiro, M. H., Veselovsky, V., & West, R. (2023). The Amplification Paradox in Recommender Systems. *Proceedings of the International AAAI Conference on Web and Social Media*, *17*, 1138-1142. https://doi.org/10.1609/icwsm.v17i1.22223

Ricci, F., Rokach, L., & Shapira, B. (Red.). (2022). *Recommender Systems Handbook*. Springer US. https://doi.org/10.1007/978-1-0716-2197-4

Rogers, R., & Righetti, N. (2025). Coordinated inauthentic behaviour on Facebook? A typology of manufactured attention. *Platforms & Society*. https://doi.org/10.1177/29768624251369784

Roy, D., & Dutta, M. (2022). A systematic review and research perspective on recommender systems. *Journal of Big Data*, *9*(1), 59. https://doi.org/10.1186/s40537-022-00592-5

Saris, K., & van de Ven, C. (2021, 3 maart). *Misogynie als politiek wapen*. De Groene Amsterdammer. https://www.groene.nl/artikel/misogynie-als-politiek-wapen

Schiffer, Z. (2023, 15 februari). *Yes, Elon Musk created a special system for showing you all his tweets first*. The Verge. https://www.theverge.com/2023/2/14/23600358/elon-musk-tweets-algorithm-changes-twitter

Schneier, B., & Sanders, N. E. (2025). *Rewiring Democracy: How AI Will Transform Our Politics, Government, and Citizenship*. MIT Press.
's time for the European Union to rethink personal social networking [Bruegel]. Bruegel. https://www.bruegel.org/policy-brief/its-time-european-union-rethink-personal-social-networking

Snap. (2025a, 29 augustus). *European Union Transparency | Snapchat Transparency*. https://values.snap.com/privacy/transparency/european-union-h1-2025

Snap. (2025b december). *Snapchat Moderation, Enforcement, and Appeals | Community Guidelines Explainer*. https://values.snap.com/privacy/transparency/community-guidelines/moderation

Snap. (z.d.). *Content Guidelines for Recommendation Eligibility*. Geraadpleegd 15 januari 2026.

Stepanov, A. & Gupta, A. (2021, 10 februari). Political Content in Feeds. *Meta Newsroom*. https://about.fb.com/news/2021/02/reducing-political-content-in-news-feed/

Stray, J., Halevy, A., Assar, P., Hadfield-Menell, D., Boutilier, C., Ashar, A., Bakalar, C., Beattie, L., Ekstrand, M., Leibowicz, C., Moon Sehat, C., Johansen, S., Kerlin, L., Vickrey, D., Singh, S., Vrijenhoek, S., Zhang, A., Andrus, M., Helberger, N., … Vasan, N. (2024). Building Human Values into Recommender Systems: An Interdisciplinary Synthesis. *ACM Trans. Recomm. Syst.*, *2*(3), 20:1-20:57. https://doi.org/10.1145/3632297

Substack, O. (2025, 27 oktober). Demystifying the feed [Substack newsletter]. *On Substack*. https://on.substack.com/p/demystifying-the-feed

Swart, J. (2021). Experiencing Algorithms: How Young People Understand, Feel About, and Engage With Algorithmic News Selection on Social Media. *Social Media + Society*, *7*(2), 20563051211008828. https://doi.org/10.1177/20563051211008828

TensorFlow. (2023, 27 mei). *Recommending movies: retrieval | TensorFlow Recommenders*. Recommending Movies: Retrieval. https://www.tensorflow.org/recommenders/examples/basic_retrieval

Thiele, D., Milzner, M., Heft, A., Gong, B., & Pfetsch, B. (2025). Attributing coordinated social media manipulation: A theoretical model and typology. *New Media & Society*, 14614448251350100. https://doi.org/10.1177/14614448251350100

Thorburn, L., Stray, J., & Bengani, P. (2023, 7 mei). Making Amplification Measurable. *Understanding Recommenders*. https://medium.com/understanding-recommenders/making-amplification-measurable-2be548e5986c

TikTok. (2025). *TikTok DSA Transparency Report (January - June 2025)*. https://sf16-va.tiktokcdn.com/obj/eden-va2/zayvwlY_fjulyhwzuhy[/ljhwZthlaukjlkulzlp/DSA_H1_2025/TikTok-DSATransparencyReport-January-June-2025.pdf

Tjaden, J., Wolfgram, J., Philipp, A., Weißmann, S., Bobzien, L., Kohler, U., & Verwiebe, R. (2025). *Does the TikTok feed lean right? Exposure to Political Party Content among non-partisan users during regional and federal elections in Germany* (7vdex_v1). SocArXiv. https://doi.org/10.31235/osf.io/7vdex_v1

Trapman, L. (2024a). Een nieuw stelsel voor de beslechting van verkiezingsgeschillen. *Nederlands Juristenblad*.

Trapman, L. (2024b). *Kiesrecht, verkiezingen en verkiezingscampagnes*. Radboud Universiteit Nijmegen.

Treré, E., & Bonini, T. (2024). Amplification, evasion, hijacking: algorithms as repertoire for social movements and the struggle for visibility. *Social Movement Studies*, *23*(3), 303-319. https://doi.org/10.1080/14742837.2022.2143345

Truong, B. T., Lou, X., Flammini, A., & Menczer, F. (2024). Quantifying the vulnerabilities of the online public square to adversarial manipulation tactics. *PNAS Nexus*, *3*(7), pgae258. https://doi.org/10.1093/pnasnexus/pgae258

Turning Passion to Profit: WAYS TO MAKE MONEY ON TIKTOK LIVE. (z.d.). Geraadpleegd 13 februari 2026, van https://www.tiktok.com/live/creators/en-US/article/turning-passion-to-profit-ways-to-make-money-on-tiktok-live_en-US

Twitter-team. (z.d.). *twitter/the-algorithm: Source code for the X Recommendation Algorithm*. GitHub. Geraadpleegd 16 januari 2026, van https://github.com/twitter/the-algorithm

Under the hood of a Doppelgänger – Qurium Media Foundation. (2022, 27 september). *Qurium*. https://www.qurium.org/alerts/under-the-hood-of-a-doppelganger/

Vliegenthart, R., Kruikemeier, S., Turkenburg, E., & Hamilton, A. (2024). *Literatuurscan sociale media en democratie met speciale aandacht voor anonimiteit*.

Wardle, C., & Derakhshan, H. (2017). *INFORMATION DISORDER : Toward an interdisciplinary framework for research and policy making* (p. 107). https://rm.coe.int/information-disorder-toward-an-interdisciplinary-framework-for-researc/168076277c

Wetenschappelijke Raad voor het Regeringsbeleid (WRR). (2024). *Aandacht voor media. Naar nieuwe waarborgen voor hun democratische functies* (Nr. 111). Wetenschappelijke Raad voor het Regeringsbeleid (WRR). https://www.wrr.nl/publicaties/rapporten/2024/10/03/aandacht-voor-media

X. (z.d.). *DSA Transparency Report - October 2025*. Geraadpleegd 12 februari 2026, van https://transparency.x.com/dsa-transparency-report-2025-october.html

X v Russmedia Digital SRL and Inform Media Press SRL, ECLI:EU:C:2025:935 (ECJ 2 december 2025). https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:62023CJ0492

Ye, J., Luceri, L., & Ferrara, E. (2025). Auditing Political Exposure Bias: Algorithmic Amplification on Twitter/X During the 2024 U.S. Presidential Election. *Proceedings of the 2025 ACM Conference on Fairness, Accountability, and Transparency*, 2349-2362. https://doi.org/10.1145/3715275.3732159

Yi, X., Yang, J., Hong, L., Cheng, D. Z., Heldt, L., Kumthekar, A., Zhao, Z., Wei, L., & Chi, E. (2019). Sampling-bias-corrected neural modeling for large corpus item recommendations. *Proceedings of the 13th ACM Conference on Recommender Systems*, 269-277. https://doi.org/10.1145/3298689.3346996

Zuckerman, E. (2023, 25 oktober). Let the community work it out: A throwback to early internet days could fix social media's crisis of legitimacy. *Nieman Lab*. https://www.niemanlab.org/2023/10/let-the-community-work-it-out-a-throwback-to-early-internet-days-could-fix-social-medias-crisis-of-legitimacy/

Bijlage 1: geraadpleegde partijen

Interviews
- Ministerie van Binnenlandse Zaken en Koninkrijksrelaties
- Autoriteit Consument en Markt
- Catalina Goanta (Universiteit Utrecht)
- John Albert (Universiteit van Amsterdam)
- Snap (Snapchat)

Schriftelijke interviews
- Meta (Instagram en Facebook)
- TikTok

Expertsessie
- Robert van der Noordaa (Trollrensics)
- Richard Odekerker (Trollrensics)
- Fabio Votta (Universiteit van Amsterdam)
- Jelle Postma (Justice for Property)
- Pieter van Boheemen (Post-X Society)
- Sanne Kruikemeier (Wageningen University & Research)
- Varvara Boboc (Universiteit van Amsterdam)
- Eva Hofman (De Groene Amsterdammer)
- Evangelos Kanoulas (Universiteit van Amsterdam)
- Harrie Oosterhuis (Radboud Universiteit)

Bijlage 2: methodologische verantwoording

Dit rapport is geschreven aan de hand van een hoofdvraag die is opgesplitst in vijf deelvragen. Elke deelvraag is beantwoord aan de hand van verschillende onderzoekmethoden. In deze bijlage is de dataverzamelingsmethode per deelvraag gespecificeerd.

De onderzoeksvraag van dit rapport is:
Welke rol spelen aanbevelingsalgoritmen op grote socialmediaplatformen bij inmenging in verkiezingen?

Om de hoofdvraag te beantwoorden zijn deelvragen opgesplitst door verschillende concepten uit de hoofdvraag te vatten in een deelvraag:

1. Wat is inmenging via socialmediaplatformen in de context van verkiezingen?
2. Hoe werken aanbevelingsalgoritmen van grote socialmediaplatformen?
3. Hoe kunnen aanbevelingsalgoritmen gebruikt worden voor inmenging in verkiezingen?
4. Welke inspanningen om inmenging via aanbevelingsalgoritmen te voorkomen worden al gedaan?
5. Welke handelingsopties zijn er om inmenging via aanbevelingsalgoritmen verder te voorkomen en/of aan te pakken?

In de tabel hierna staan de deelvragen waarmee we hebben gewerkt tijdens het verzamelen en analyseren van data. Naast de deelvragen staan subdeelvragen. Dit zijn vragen die een focus geven aan de onderzoekers tijdens de analyse van de data.

Tabel 8 Overzicht van dataverzamelingsmethode

Deelvraag	Subdeelvragen	Dataverzamelingsmethode
1. Wat is inmenging via socialmedia-platformen in de context van verkiezingen?	• Hoe wordt in literatuur en beleid invulling gegeven aan begrip inmenging? Wat wordt er gezegd over o.a.: actor (ie publiek/privaat) en diens herkomst (ie buitenlands/ binnenlands), instrument (ie algoritmen, desinformatiecampag-nes), doelstelling (verkiezingen, ondermijning rechtsorde) • Hoe zijn onderlinge verschillen van betekenis te verklaren? • Welke risico's van inmenging worden genoemd in academische en grijze literatuur m.b.t. aanbevelingsalgo-ritmen van social media?	• Deskresearch academische en grijze literatuur over: afbakening van begrip 'inmenging' en voorbeelden waarbij inmenging tijdens verkiezingen wordt gesproken, zoals Roemeense presidentsverkiezingen. • Interview met juristen en beleidsmakers over definitie inmenging. • Expertsessie over definitie inmenging. • Deskresearch academische en grijze literatuur over: zorgen, genoemd in academische literatuur, Kamerbrieven, journalistieke bronnen en rapportages ngo's. • Expertsessie ter validatie en identificatie van risico's en zorgen.
2. Hoe werken aanbevelings-algoritmen van grote socialmedia-platformen?	• Wat is een aanbevelingsalgoritme? • Wat is een op engagement gebaseerd algoritme? • Wat zijn verschillen tussen platformen? • Wat is de rol van filterbubbels, echokamers in de werking van aanbevelingsalgoritmen? • Wat is amplificatie? • Welke inhoud wordt relatief vaker aanbevolen?	• Deskresearch academische en grijze literatuur over: de technische en sociale aspecten van aanbevelingsalgoritmen op socialmediaplatformen. • Expertsessie ter validatie over de werking van aanbevelingsalgoritmen. • Interviews met platformen en openbare bronnen over de werking van aanbevelingsalgoritmen.

Deelvraag	Subdeelvragen	Dataverzamelingsmethode
3. Hoe kunnen aanbevelings-algoritmen gebruikt worden voor inmenging in verkiezingen?	• Wat zijn beschreven tools en technieken die voor actoren beschikbaar zijn voor meer zichtbaarheid op social media voor een politieke partij, kandidaat of issue? • Welke van deze beschreven tools en technieken zijn: geaccepteerd, niet illegaal maar wel onethisch, illegaal?	• Deskresearch academische literatuur en grijze literatuur over: de rol van aanbevelingsalgoritmen bij de verspreiding van politieke inhoud. • Interviews platformen over de rol van aanbevelingsalgoritmen bij viraliteit in campagnetijd. • Expertsessie ter validatie. • Analyse wetgeving door huisjurist.
4. Welke inspanningen om inmenging via aanbevelings-algoritmen te voorkomen worden al gedaan?	• Welke beleidsinstrumenten en wetgeving in NL en Europa heeft betrekking op inmenging, verkiezingen en/of aanbevelingsalgoritmen? • Op welke manieren pogen deze instrumenten iets aan inmengingsactiviteiten via social media te doen? • Welke inspanningen zeggen platformen te leveren? • Wat vinden experts van de effectiviteit van deze inspanningen?	• Deskresearch academische, grijze literatuur en relevante wet- en regelgeving over onder andere: a. De Digital Services Act en *Guidelines for the mitigation of systemic risks online for elections.* b. De Nederlandse gedragscode Transparantie Online Politieke Advertenties. c. *Regulation (EU) 2024/900 on the transparency and targeting of political advertising.* d. De Nederlandse kieswet. • Analyse bovenstaande wetgeving door onze vaste juridische expert. • Expertsessie
5. Welke handelings-perspectieven zijn er om inmenging via aanbevelingsalgo-ritmen te voorkomen en/of aan te pakken?	• Zijn bestaande inspanningen afdoende om risico's te beperken? • Zo niet, wat is er dan nog nodig/ mogelijk?	• Wetenschappelijke en grijze literatuur over inmenging via social media en manieren om dit te bestrijden. • Expertsessie ter identificatie en validatie van handelingsperspectieven om inmengingsinstrumentarium te beperken.

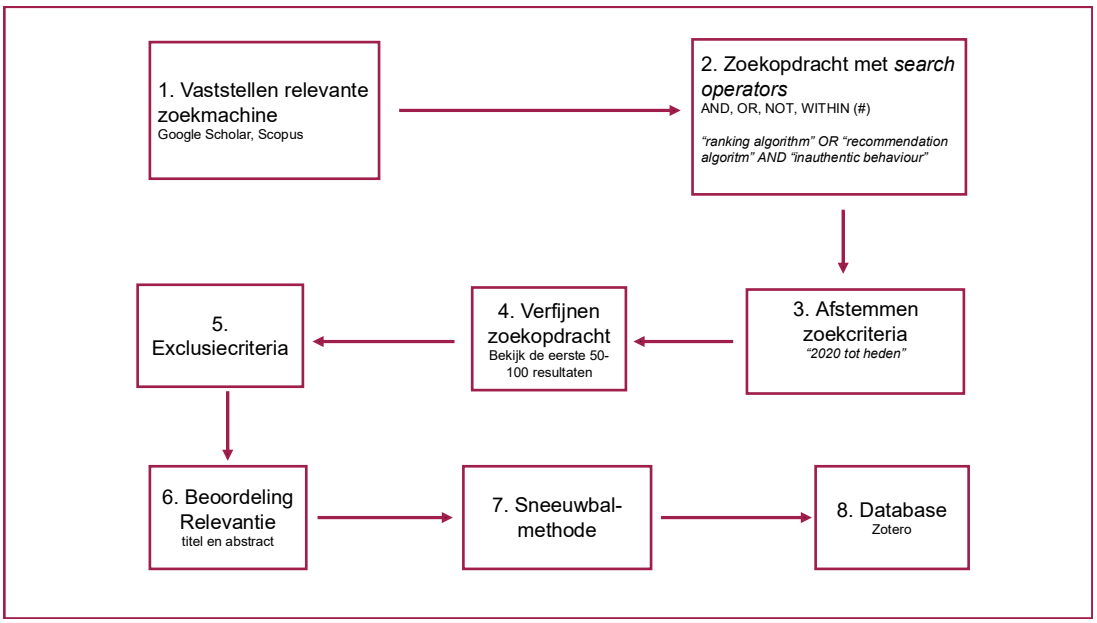
Bron: Rathenau Instituut

Dataverzameling
Tijdens de deskresearch analyseerden we verschillende bronnen: academische literatuur, grijze literatuur (waaronder journalistieke bronnen en rapporten van ngo's en onderzoeksinstituten) en beleidsbronnen en juridische bronnen. Voor het aanleggen van een database met relevante literatuur volgden we het stappenplan in figuur 4 (zie hieronder).

We zijn gestart met het gebruiken van twee zoekmachines: Google Scholar en Scopus. Google Scholar gebruikten we in eerste plaats om de zoekopdrachten mee op te stellen (stap 2). Per deelvraag ontwikkelden een of enkele zoekopdrachten, verdeeld over de verschillende onderzoekers.

Nadat de zoekopdracht per deelvraag waren opgesteld en de zoekcriteria waren vastgesteld (stappen 2 en 3) bekeken we de eerste vijftig tot honderd zoekresultaten in Scopus (stap 4) en voeren we de volgende stappen uit met deze zoekmachine. Per deelvraag selecteren we maximaal twintig bronnen. Deze selectie vormde de database met bronnen waarmee we de deelvragen beantwoorden.

Figuur 5 Schematische weergave van de dataverzameling



Bron: Rathenau Instituut

Zoterodatabase met bronnen
Voor het beheer van de database gebruikten we Zotero. Nadat de database was gevuld met relevante literatuur, verdeelden de onderzoekers de literatuur per deelvraag en lazen dit deel van de bronnen uit de database. De onderzoekers

labelen de vijf tot tien meest relevante bronnen als 'sleutelpublicaties'. Deze publicaties werden door alle onderzoekers gelezen.

Expertinterviews
De onderzoekers hielden interviews met experts voor het beantwoorden van deelvragen 1 tot en met 4. De losse interviews zijn uitgevoerd met als doel om:

- te oriënteren op het onderwerp;
- voorlopige bevindingen te toetsen bij experts;
- bestaande bronnen en hypothese in het vakgebied te lokaliseren;
- in kaart te brengen wat lopend onderzoek is en waarover nog gepubliceerd zal worden;
- verder te vragen op reeds gepubliceerd onderzoek uitgevoerd door experts.

We volgden tijdens deze interviews geen vaste vragenlijst. We legden wel een of meerdere van de deelvragen voor tijdens het interview.

Expertsessie
We organiseerden een expertsessie ter validatie en verrijking van voorlopige bevindingen en identificatie van passende beleidsmaatregelen. De sessie droeg bij aan het beantwoorden van deelvragen 1 tot en met 5.

Selectie experts
De sessie is georganiseerd met experts uit verschillende vakgebieden zoals politicologie, communicatiewetenschappen, mediawetenschappen en cybersecurity. Ook waren er onderzoeksjournalisten en ngo's aanwezig die specifiek publiceren over inauthentiek gedrag op social media.

De selectie van de experts is gedaan op basis van het bestaande netwerk van het Rathenau Instituut. Dit werd aangevuld met experts die naar voren kwamen uit de geanalyseerde bronnen van de deskresearch en door te vragen aan experts of ze mensen konden aanraden die relevante kennis hebben van het onderwerp.

Dataverzameling tijdens expertsessie en verwerking
Tijdens de expertsessie zijn voorlopige bevindingen gepresenteerd aan de groep experts. In break-outsessies zijn de verschillende deelvragen besproken. Tijdens de expertsessie zijn aantekeningen gemaakt door de onderzoekers.

Platforminterviews
Vertegenwoordigers van platformen zijn geïnterviewd voor het beantwoorden van deelvragen 2 en 4. Aan platformen zijn vragen voorgelegd over de werking van de aanbevelingsalgoritmen op hun platformen en hun inspanningen en resultaten om inmenging via aanbevelingsalgoritmen in verkiezingstijd te voorkomen.

Voorafgaand aan de interviews is deskresearch uitgevoerd naar platformbeleid en naar DSA-auditreports. De platformen die we hebben verzocht om mee te werken voor een interview zijn gekozen op basis van aantallen gebruikers in Nederland YouTube (Google), Facebook en Instagram (Meta), LinkedIn (Microsoft), TikTok, Snapchat (Snap) en X.[24] Op advies van de expertsessie is Telegram toegevoegd aan deze lijst.

Platformen die hebben meegewerkt aan het onderzoek zijn Meta, TikTok en Snap. Meta en TikTok hebben schriftelijk gereageerd op vooraf opgestelde vragen. Met Snap is een interview gehouden. Google YouTube en Microsoft (LinkedIn) hebben aangegeven niet in te willen gaan op ons verzoek voor een interview. X en Telegram hebben niet gereageerd op een verzoek tot een interview.

Tijdens de interviews hebben we een semigestructureerde vragenlijst gevolgd met enkele vragen gericht aan individuele platformen. Voor de vragenlijst hebben we onder andere geput uit een publicatie van verschillende maatschappelijke organisaties met vragen die de Europese Commissie zou kunnen stellen aan grote online platformen.[25]

Interne kwaliteitscontrole
We hebben een uitvoerige interne kwaliteitscontrole doorlopen door het conceptrapport voor te leggen, achtereenvolgens aan: coördinator, interne reviewer (niet bij het onderzoek betrokken), MT-contact, en een communicatiemedewerker.

24 Hiervoor houden we de gegevens aan van marketingbureau NewCom over aantallen gebruikers per platform: https://www.newcom.nl/nationale-social-media-onderzoek-basis-2025/

25 Fundacja Panoptykon, Irish Council for Civil Liberties, & People vs Bigtech. (2024). Fixing Recommender Systems (Briefing Note). https://peoplevsbig.tech/wp-content/uploads/2024/05/Fixing-Recommender-Systems_Panoptykon_PeopleVsBigTech.pdf

Bijlage 3: wetgeving en beleid

In deze bijlage bespreken we de meest relevante wetgeving. Het is belangrijk te vermelden dat de wetten ieder hun eigen reikwijdte en doelstellingen hebben en dus vanuit verschillende invalshoeken en op directe en indirecte wijze het probleem van inmenging via aanbevelingsalgoritmen benaderen. Het zijn wetten die naar gelang het type vraagstuk (of instrument uit de toolbox) van toepassing kunnen zijn.

De meest relevante wetten zijn de Verordening betreffende transparantie en gerichte politieke reclame (VPR), de digitale dienstenverordening, de AI-verordening, de wetgeving over mediavrijheid, de richtlijn Oneerlijke handelspraktijken, de richtlijn Audiovisuele mediadiensten en de Algemene verordening gegevensbescherming (AVG).

De *Rijksbrede strategie voor de effectieve aanpak van desinformatie* en het European Democracy Shield (EDS) zijn relevante beleidsinitiatieven. De rijksbrede strategie bespreken we apart, in tegenstelling tot het EDS. Een groot aantal van de EDS-maatregelen zijn onder te brengen onder hieronder besproken wetgeving en worden dus daar genoemd.

De Wet op de politieke partijen (Wpp) bespreken we niet in deze bijlage. De Wet op de politieke partijen zou eisen stellen aan politieke advertenties en microtargeting, maar die zijn uit het wetsvoorstel gehaald omdat deze nu geregeld zijn in de Europese Verordening betreffende transparantie en gerichte politieke reclame (*Kamerstukken II, 36 742, nr 8*, 2025). Deze verordening wordt op nationaal niveau nader uitgevoerd met de Uitvoeringswet verordening transparantie en gerichte politieke reclame. Het wetgevingstraject van de uitvoeringswet is nog niet afgerond, al zijn de relevante toezichthouders (Commissariaat voor de Media, ACM en AP) wel al aangewezen (Ministerie van BZK, 2025).

Rijksbrede strategie voor de effectieve aanpak van desinformatie
De *Rijksbrede strategie voor de effectieve aanpak van desinformatie* is bedoeld om desinformatie aan te pakken die volgens het kabinet een groot risico vormt voor het vrije en open debat (*Kamerstukken II, 30 821, nr. 230*, 2024). De strategie bestaat uit drie sporen: 1) maatregelen om verspreiders en verspreiding van desinformatie aan te pakken, 2) maatregelen om weerbaarheid van burgers te versterken, en 3) kennisontwikkeling van de problematiek en effectieve aanpak van desinformatie.

Onder deze strategie kan het Ministerie van BZK in bijzondere gevallen haar *verkiezingsflaggerstatus* inzetten (zie hoofdstuk 1). Ook werkt het kabinet nauw

samen met andere Europese lidstaten en instellingen om signalen tegen ongewenste beïnvloeding te detecteren en maatregelen te nemen. Denk daarbij aan het Europese Rapid Alert System (RAS) en het European Cooperation Network on Elections (ook genoemd in het EDS). Momenteel onderzoekt de minister hoe die beter zicht kan krijgen op de verspreiding van desinformatie en buitenlandse informatiemanipulatie en beïnvloeding (Foreign Information Manipulation and Interference, FIMI, zie hoofdstuk 1) (*Kamerstukken II, 36 552, nr. 12*, 2025).

Verordening transparantie en gerichte politieke reclame (VPR)
Het doel van deze wet is om burgers te helpen gemakkelijker politieke reclame online te herkennen. De verordening stelt transparantie-eisen aan politieke reclame en regels aan het gebruik van targeting en optimalisatietechnieken voor politieke reclame. Inmenging en online manipulatie is een van de aanleidingen geweest om de Verordening betreffende transparantie en gerichte politieke reclame op te stellen.

Optimalisatietechnieken zijn technieken die worden gebruikt om bijvoorbeeld het bereik van een reclameboodschap te vergroten en zo een specifieke groep personen te bereiken. Bij zowel targeting als optimalisatietechnieken is het relevant dat persoonsgegevens worden verwerkt. Targeting en optimalisatie is in principe toegestaan, maar vereist onder meer uitdrukkelijke toestemming van de persoon die onderwerp is van de techniek. Daarnaast mogen de persoonsgegevens van de betrokkene (zoals de socialmediagebruiker) alleen door de verwerkingsverantwoordelijke (zoals de inmengende partij) worden gebruikt als deze laatste partij de gegevens zelf heeft verzameld. Indirecte verzameling via databrokers is dus uit den boze. Politieke targeting van zeventienjarigen of kinderen die jonger zijn en targeting met bijzondere persoonsgegevens (zoals informatie over politieke voorkeuren) is niet toegestaan.

Het is niet waarschijnlijk dat inmengende partijen om toestemming vragen of dat deze partij zou erkennen iets met een inmengingsactiviteit te maken te hebben. Bovendien verlopen dit soort campagnes vaak via databrokers, en indirecte dataverzameling via dit soort partijen is niet toegestaan (Raad van State, 2025). Daarnaast beschikken databrokers in veel gevallen over data waar gebruikers geen expliciete toestemming voor gegeven hebben. Daarmee zijn dit type beïnvloedingscampagnes in de praktijk niet snel rechtmatig en kunnen ook andere betrokken partijen daarvoor verantwoordelijk worden gehouden dan alleen de inmengende partij. Bijvoorbeeld omdat een influencer niet transparant is, of omdat een socialmediaplatform in een bepaald geval zelf als aanbieder van politieke reclamediensten optreedt.

Deze wet stelt daarnaast transparantie-eisen aan iedereen die politieke reclame maakt, dus ook influencers, bloggers, vloggers etc. De wet gaat niet over de inhoud van politieke reclames. Influencers moeten volgens de VPR bij politieke reclame daarnaast nagaan of de opdrachtgever niet afkomstig is uit een verboden derde land (zie hieronder), informatie over deze opdracht aanleveren bij een Europees register en de post voorzien van een transparantieverklaring.

Met politieke reclame wordt gedoeld op een boodschap (inclusief promotie, verspreiding etc. daarvan) die tegen een vergoeding, in eigen naam of in het kader van een politieke reclamecampagne wordt overgebracht door, voor of namens een politieke actor en die van invloed kan zijn of bedoeld is om het resultaat van een verkiezing, stemgedrag of een regelgevingsproces te beïnvloeden. Volgens de richtsnoeren van de Europese Commissie is de ontmoediging om te stemmen ook een vorm van politieke reclame. Politieke actoren zijn niet alleen politieke partijen, maar kunnen ook personen of organisaties zijn die politieke doelstellingen promoten.

Een voorbeeld van politieke reclame kan zijn dat een politieke actor een netwerk van influencers betaalt om soortgelijke boodschappen te posten, waardoor indirect reclame voor hem wordt gemaakt. Daarbij valt te denken aan de inzet van influencers in aanloop naar de Roemeense verkiezingen die geen expliciet politieke uitspraken deden, maar wel vergelijkbare inhoud deelde die indirect verwees naar de waarden die een bepaalde politieke kandidaat uitdroeg.

Ook platformen kunnen een verantwoordelijkheid hebben onder de VPR. Dit speelt zodra zij een vergoeding krijgen voor de verspreiding van politieke reclame. Bijvoorbeeld als een platform tegen betaling een politieke 'reclamepost' van een influencer bevordert door deze prominent te tonen, en de influencer voor die post betaald krijgt van een inmengende partij. In zo'n geval biedt het platform een politieke reclamedienst aan en wordt hij geacht een uitgever van politieke reclame in de zin van de VPR te zijn, zo volgt uit de richtsnoeren van de Europese Commissie. De zeer grote platformen en de zeer grote zoekmachines hebben een paar specifieke verplichtingen, vooral gericht op publieke transparantie over de reclames en het snel reageren op verzoeken van bijvoorbeeld melders en toezichthouders. Bijvoorbeeld om informatie te verstrekken of onrechtmatige reclame offline te halen.

De VPR stelt dat drie maanden voordat de verkiezingen plaatsvinden, politieke reclamediensten alleen mogen worden verleend aan opdrachtgevers die een EU-burger zijn of een bepaalde band hebben met de EU. Nederland kan hier strengere eisen aan stellen. De Europese Commissie zal in het kader van het European

Democracy Shield deze en andere relevante EU-regels bij een vrijwillig netwerk van influencers onder de aandacht brengen en normontwikkeling op stimuleren.

De digitale dienstenverordening (DSA) en gedragscode desinformatie
De DSA moet de gebruikers van digitale platformen beschermen tegen verspreiding van illegale inhoud en de grondrechten van gebruikers respecteren. De wet legt daarvoor verantwoordelijkheden en een verantwoordingsplicht vast voor aanbieders van intermediaire diensten, zoals Instagram, Snapchat en Telegram.

Wat de DSA doet is platformen een verplichting geven om verspreiding van illegale content zoveel mogelijk te beperken. Bedrijven modereren op inhoud op basis van Europese en nationale wetgeving. Verschillende typen uitingen (of die nou offline of offline zijn) zijn in Europese lidstaten waaronder Nederland illegaal, zoals discriminerende content, bedreigende content en terroristische content. Daarnaast bevat de DSA bepalingen voor klachtenprocedures en zorgvuldigheidsverplichtingen voor platformen, bijvoorbeeld door adequaat te reageren nadat illegale content is geplaatst.

Voor zeer grote platformen en zoekmachines gelden strengere regels. Deze platformen en zoekmachines zijn onder meer YouTube, Facebook, Instagram, LinkedIn, TikTok, Twitch, Google Search, Snapchat, X, WhatsApp en Bing. Telegram hoort er vooralsnog niet bij, maar zit waarschijnlijk wel dicht tegen het aantal benodigde gebruikers aan (*EU in Touch with Telegram as It Nears Criterion for EU Tech Rules*, 2024).

Belangrijke bepalingen voor ons onderzoek zijn artikel 34 en 35, die om een proactieve houding vraagt van platformen. Volgens de DSA moeten zeer grote platformen en zoekmachines systeemrisico's inventariseren en redelijke, evenredige en doeltreffende maatregelen nemen, bijvoorbeeld om de negatieve impact op verkiezingsprocessen te voorkomen (denk aan circulatie van valse informatie over kandidaten, stemdata of verzonnen steunbetuigingen aan kandidaten). De platformen zijn zelf vrij te bepalen wat deze maatregelen dat zijn.

Er is een reeks van risicobeperkende maatregelen genoemd in de DSA, zoals het aanpassen van algoritmen en aanbevelingssystemen om verspreiding van misleidende of schadelijke content tijdelijk te beperken, virale trends op andere platformen te monitoren, moderatieteams bekend met lokale taal en cultuur in te zetten of AI-gegenereerde content in politieke advertenties tijdelijk te verbieden. Platformen moeten over hun inspanningen rapporteren aan de Europese Commissie.

Er loopt momenteel onderzoek naar de werking van X's aanbevelingssystemen door de Europese Commissie (European Commission, 2025a). De Europese Commissie doet ook onderzoek naar TikTok's algoritmen (in relatie tot 'gecoördineerde inauthentieke manipulatie of geautomatiseerde exploitatie van de dienst') en hun beleid op het gebied van betaalde politieke content en politieke advertenties (European Commission, 2025e).

Er zijn ook nog Commissie-richtsnoeren voor platformen specifiek voor verkiezingsprocessen, zoals het wegnemen van financiële prikkels voor platformen en influencers om desinformatie, haat en extremistische content via advertenties te verspreiden (demonetisatie) (European Commission, 2024a). Ook schrijven de richtsnoeren voor om maatregelen te nemen tegen botnetwerken, 'inauthentieke accounts of ander misleidend gebruik van de dienst (anti-manipulatie). Samen met toezichthouders en Europees samenwerkingsnetwerk voor verkiezingen actualiseert de Europese Commissie deze richtsnoeren onder het EDS.

De Europese Commissie kan tijdelijke maatregelen opleggen aan een platform volgens het crisis-response-mechanisme in geval van een 'crisis', zoals het beperken van de verspreiding van bepaalde informatie. De Europese Commissie kan ook vrijwillige crisisprotocollen opzetten. Onder het beleidsinitiatief European Democracy Shield (EDS) is de Europese Commissie bezig om samen met DSA-toezichthouders om dergelijke protocollen op te stellen.

Daarnaast moeten de zeer grote platformen en zoekmachines registers aanleggen met informatie over de reclame (zoals beoogde doelgroep, adverteerder) die op hun platformen geplaatst wordt. Die registers moeten publiek toegankelijk zijn. Ook moeten aanbieders van online platformen maatregelen nemen tegen ongeoorloofde misleiding (*dark patterns* en *deceptive design*).

Aanbieders van platformen moeten in duidelijke, begrijpelijke taal de belangrijkste parameters noemen die in hun aanbevelingssystemen worden gebruikt. Het is optioneel voor gebruikers om deze parameters te kunnen wijzigen of te beïnvloeden. Gebruikers moeten volgens de DSA naast een algoritmische feed ook standaard een feed aangeboden krijgen met content van alleen gevolgde accounts (een chronologische tijdlijn). De ngo Bits of Freedom heeft een rechtszaak tegen Meta gewonnen waarbij het eiste dat gebruikers de chronologische tijdlijn als standaard kunnen instellen, in plaats van de algoritmische tijdlijn (*Meta moet (voorlopig) chronologische tijdlijn als blijvende voorkeursoptie aanbieden*, 2026).

Platformen en zoekmachines dienen aangewezen toezichthouders en onderzoekers inzicht te geven in het functioneren van deze platformen. Toezichthouders mogen data opvragen om naleving van de DSA te beoordelen en

platformen moeten op verzoek van toezichthouders uitleg geven over de werking van hun algoritmen. Ook moeten aangewezen onderzoekers toegang krijgen tot data om systeemrisico's te onderzoeken. Platformen mogen een verzoek van onderzoekers alleen afwijzen als zij de data niet hebben of als toegang tot ernstige beveiligings- of vertrouwelijkheidsrisico's zou leiden.[26]

Verder is er de Gedragscode desinformatie die volgens de DSA bij niet-systematische naleving aanleiding kan geven voor onderzoek door de Europese Commissie. De gedragscode kan gezien worden als een verdere uitwerking van de regels die in de DSA staan. De ondertekenaars van de gedragscode moeten zich bijvoorbeeld inspannen om ongeoorloofd manipulatief gedrag en praktijken te beperken. Daaronder valt het gebruik van bots, valse accounts en gecoördineerd inauthentiek gedrag, alsook 'gebruikersgedrag' gericht op het kunstmatig vergroten van het bereik of waargenomen publieke steun voor desinformatie. Daarbij wordt verwezen naar een regelmatige bijgewerkte database met 'tactics, techniques and procedures' (DISARMFoundation, z.d.).

De DSA verplicht de zeer grote platformen en zoekmachines tenslotte om de risico's van verslavend ontwerp te mitigeren, zoals een zeer gepersonaliseerd aanbevelingssysteem. De voorlopige bevinding van de Europese Commissie is dat TikTok onvoldoende maatregelen heeft genomen om verslavende aspecten van het ontwerp te mitigeren (European Commission, 2026). Ook is er een procedure van de Europese Commissie gestart naar Meta's Instagram en Facebook omdat het ontwerp te verslavend zou zijn voor minderjarigen (European Commission, 2024c).

AI-verordening

Het doel van de AI-verordening is dat AI-systemen die organisaties binnen de EU op te markt brengen of gebruiken, veilig zijn, de gezondheid niet schaden en fundamentele rechten respecteren. Zo is er een verbod op AI-systemen waarbij manipulatieve technieken worden gebruikt. Hierbij is het niet vereist dat diegene die het systeem op de markt brengt, ook intentie heeft schade te berokkenen, aldus de (concept)-richtsnoeren van de Europese Commissie (European Commission, 2025b).

Of socialmediaplatformen en in het bijzonder TikTok, onder manipulatieve systemen vallen, staat ter discussie. Ten tijde van schrijven loopt er een rechtszaak van Stichting SOMI tegen TikTok met de aanklacht dat het platform jongeren

26 X is recentelijk beboet voor het schenden van transparantieverplichtingen. Het belemmerde onder meer de toegang tot data voor onderzoekers en gaf onvoldoende inzicht in advertenties (European Commission, 2025d). De Europese Commissie heeft ook gesteld in preliminaire bevindingen dat Meta en TikTok niet goed genoeg toegang gaven tot onderzoeksdata. Het onderzoek loopt nog.(European Commission, 2025c). Ook loopt er nog een ander onderzoek tegen Meta vanwege de slechte toegang tot onderzoeksdata en het offline halen van CrowdTangle, dat bedoeld was om verkiezingen te monitoren (European Commission, 2024b).

gericht manipuleert en voor het misbruiken van gevoelige persoonlijke gegevens, die door het platform worden gebruikt om het eigen aanbevelingsalgoritme aan te sturen (*Nederlandse claimstichting begint rechtszaken tegen TikTok en X (voorheen Twitter)*, 2025). Het wordt interessant te horen of de rechter de aanbevelingssystemen van TikTok inderdaad als manipulatief zal aanmerken, en of dit platform of onderdelen daarvan dan dus mogelijk al verboden zijn onder de AI-verordening.

Volgens de AI-verordening kan verslavend ontwerp in een AI-systeem ook worden opgevat als een verboden praktijk als hiermee misbruik wordt gemaakt van de leeftijd, handicap, of socio-economische achtergrond van gebruikers.

Opvallend is dat AI-systemen die worden gebruikt voor het beïnvloeden van de uitslag van een verkiezing of referendum of van stemgedrag niet verboden zijn, maar in de hoog-risico-categorie vallen. De ontwikkelaar moet dan bijvoorbeeld de technische documentatie en gebruiksaanwijzing uitleggen hoe grondrechtenrisico's voldoende zijn afgeschermd. Het is niet duidelijk waar AI-chatbots geïntegreerd in socialmediaplatformen die desinformatie verspreiden onder vallen, net als gepersonaliseerde, AI-gegenereerde nieuwsoverzichten.

Een andere relevante bepaling is het vereiste dat AI-gegenereerde content die nauwelijks van echt te onderscheiden is (deepfakes), volgens de AI-verordening door de gebruiker gelabeld moet worden als 'door AI-gegenereerd' of woorden van gelijke strekking. In het kader van het EDS worden aanvullende richtsnoeren opgesteld voor het gebruik van AI in verkiezingsprocessen door politieke partijen en andere relevante actoren.

Europese wetgeving voor mediavrijheid (EMFA)
Deze wet moet de vrijheid en pluriformiteit van de media in de EU beschermen en het vrije verkeer van diensten stimuleren. De EMFA stelt onder meer vast wat een mediadienst is, en daaronder kunnen beroepsmatige influencers vallen die tegen een vergoeding content plaatsen.

Derde landen (niet-EU landen) kunnen volgens de EMFA van negatieve invloed zijn op de diensten. Daarom moeten mediadiensten informatie delen over eigenaarschap en de advertentiegelden die ze van derde landen ontvangen zouden kunnen hebben.

Ook de toegang tot mediadiensten van buiten de EU wordt gereguleerd bij een bedreiging van of ernstig risico voor de openbare veiligheid. Europese Raad voor mediadiensten die door deze EMFA in het leven wordt geroepen, adviseert

hierover. In de Raad zitten het Commissariaat voor de Media en de equivalenten daarvan in andere Europese lidstaten.

Ook de zeer grote online platformen hebben een verantwoordelijkheid. Zij moeten op basis van de EMFA een functionaliteit aanbieden waarmee gebruikers eenvoudig te weten kunnen komen dat ze met een mediadienst te maken hebben en dat de dienst kan verklaren redactioneel onafhankelijk te zijn van politieke partijen of derde landen. Ook moeten mediadiensten via de functionaliteit kunnen verklaren dat ze geen AI-inhoud verstrekken zonder dat die onderworpen is aan menselijke toetsing.

De Raad zal volgens de EMFA regelmatig een gestructureerde dialoog organiseren tussen de zeer grote online platformen, van aanbieders van mediadiensten en vertegenwoordigers van het maatschappelijk middenveld. Doel hiervan is onder andere om toe te zien toe te zien op de naleving van zelfregulerende initiatieven, waaronder ook de eerdergenoemde Gedragscode desinformatie. Ook de Europese Commissie moet ervoor zorgen dat de interne markt voor mediadiensten onafhankelijk en doorlopend wordt gemonitord op mediaconcentratie en op risico's voor buitenlandse informatiemanipulatie en inmenging.

Richtlijn oneerlijke handelspraktijken
De Richtlijn beoogt onder andere oneerlijke handelspraktijken tegen te gaan, zoals agressieve marketing of onware commerciële uitingen. Het gaat om content met een commerciële lading, omdat ze als doel hebben iets te verkopen (dienst, product, etc.) aan consumenten. Influencers maken vaak reclame. Het is van belang dat de reclame als zodanig zichtbaar is. De influencer moet dus transparant zijn over geldstromen en/of voordelen die die geniet. Het gebruiken van neplikes en nepvolgers is niet toegestaan. In Nederland is de *Unfair commercial practices directive* (UCPD) geïmplementeerd in het Burgerlijk Wetboek.

Richtlijn audiovisuele mediadiensten
De Mediawet implementeert de richtlijn en gaat over mediadiensten. Zogeheten video-uploaders (influencers, vloggers, content creators, streamers) vallen daar ook onder. Ook video-uploaders moeten transparant zijn over reclame. Verder moeten ze zich aansluiten bij de Stichting Reclame Code. De code schrijft voor dat misleiding en agressieve praktijken en in de inzet van subliminale technieken verboden zijn. Ook mag een reclame niet appelleren aan gevoelens van angst of bijgelovigheid.

Videoplatformdiensten als YouTube, Snapchat en TikTok, moeten maatregelen nemen om reclame en de achterliggende partij kenbaar te maken. Daarnaast heeft Snap, die de gedragscode Desinformatie niet heeft ondertekend, een aparte

gedragscode ondertekend met Commissariaat voor de Media. De gedragscode legt onder andere vast welke maatregelen Snap moet nemen tegen schadelijke content en bescherming.

Algemene verordening gegevensbescherming (AVG)
De AVG moet de persoonsgegevens van iedereen op EU/EER-grondgebied beschermen. Persoonsgegevens die openlijk online beschikbaar zijn, vallen ook onder de AVG. Net als afgeleide gegevens, dat zijn gegevens op basis van iemands like- of share-gedrag. Gegevens over politieke voorkeur kennen extra bescherming als bijzondere persoonsgegevens, daarvoor is uitdrukkelijke toestemming nodig van de partij die de gegevens wil verwerken.

Alle partijen betrokken bij de verwerking van persoonsgegevens moeten zich als verwerkingsverantwoordelijken aan de wet houden. Een dergelijke partij hoeft zelf geen gegevens te verzamelen. Meerdere partijen kunnen verwerkingsverantwoordelijke zijn. Een inmengende actor die Nederlanders online probeert te targeten, moet zich dus aan de AVG houden, evenals alle partijen die deze campagne mogelijk maken en verwerkingsverantwoordelijke zijn.

Alle verwerkingsverantwoordelijken, waaronder ook platformen kunnen vallen, moeten transparant zijn over de wijze waarop en waarom ze gegevens verwerken en op verzoek van desbetreffende persoon inzage geven.

De AVG biedt bescherming tegen praktijken waarbij persoonsgegevens worden verzameld, zoals *profiling* en *targeting* op social media en in chatbots in andere *Large Language Models* in zoverre persoonsgegevens worden verzameld in de context van dit onderzoek. Gegevens mogen namelijk niet verzameld worden als ze onevenredig nadelig, onverwacht of misleidend zijn voor betrokken personen (behoorlijkheidsprincipe). De onderzoeksdienst van het Europarlement, de EPRS, verwacht dat het aanbevelingssysteem van bijvoorbeeld TikTok, waarbij gedragsdata van mensen worden gebruikt volgens de onderzoeksdienst om gebruikers te manipuleren, onverenigbaar is met het behoorlijksheidsprincipe van de AVG (European Parliamentary Research Service, 2025).

Platformen moeten volgens een recente uitspraak van het Hof van Justitie van de Europese Unie daarnaast proactief informatie scannen op bijvoorbeeld het gebruik van bijzondere gegevens voor een advertentie en direct ingrijpen bij een illegale publicatie van die gegevens. Normaal gesproken hebben platformen geen verplichting om proactief te handelen, en geven ze al dan niet opvolging aan een klacht (notice-and-action-mechanisme), maar de gevolgen van deze uitspraak lijken daar wel op (*X v Russmedia Digital SRL and Inform Media Press SRL*, 2025).

Bovendien mogen persoonsgegevens niet buiten de EU terechtkomen in specifieke landen als Rusland en China, zonder dat er aan bepaalde juridische voorwaarden is voldaan. Het is echter niet eenvoudig niet-EU partijen hieraan te houden. Het Amerikaanse Clearview dat boetes en maatregelen heeft opgelegd gekregen van verschillende Europese toezichthouders weigert te betalen of zich aan de maatregelen te houden. De AP onderzoekt momenteel of het de Clearview-bestuurders individueel kan aanpakken (Autoriteit Persoonsgegevens, 2024).

Open access

Het Rathenau Instituut heeft een openaccessbeleid. Rapporten, achtergrondstudies, wetenschappelijke artikelen, software worden vrij beschikbaar gepubliceerd. Onderzoeksgegevens komen beschikbaar met inachtneming van wettelijke bepalingen en ethische normen voor onderzoek over rechten van derden, privacy, en auteursrecht.

Contactgegevens

Anna van Saksenlaan 51
Postbus 95366
2509 CJ Den Haag
070-342 15 42
info@rathenau.nl
www.rathenau.nl

Bestuur van het Rathenau Instituut

Het Rathenau Instituut stimuleert de publieke en politieke meningsvorming over de maatschappelijke aspecten van wetenschap en technologie. We doen onderzoek en organiseren het debat over wetenschap, innovatie en nieuwe technologieën.

40 JAAR Rathenau Instituut
Onderzoek & dialoog | Wetenschap, technologie en innovatie