



Universiteit
Leiden
The Netherlands

Quis custodiet ipsos custodes? Een kritisch perspectief op de rol van fact-checkers in het Unierechtelijke desinformatiebeleid

Wit, A. de

Citation

Wit, A. de. (2025). Quis custodiet ipsos custodes?: Een kritisch perspectief op de rol van fact-checkers in het Unierechtelijke desinformatiebeleid. *Mediaforum*, 2025(Studenteneditie), 36-42. Retrieved from <https://hdl.handle.net/1887/4252625>

Version: Publisher's Version

License: [Leiden University Non-exclusive license](#)

Downloaded from: <https://hdl.handle.net/1887/4252625>

Note: To cite this publication please use the final published version (if applicable).

Quis custodiet ipsos custodes?

Artikel

Een kritisch perspectief op de rol van fact-checkers in het Unierechtelijke desinformatiebeleid

Alexander de Wit*

1. Introductie

Fact-checken wordt door journalisten, overheden en academici alom geroemd als een noodzakelijk en effectief middel in de strijd tegen desinformatie.¹ De Europese Commissie (hierna: EC of Commissie) ziet het verminderen van vertrouwen in instituties, het aanwakken van extremisme en het bemoeilijken van het maken van geïnformeerde keuzes als belangrijke schadelijke effecten van desinformatie.² De schadelijke gevolgen van desinformatie doen zich met name gevoelen in de digitale publieke sfeer, op grote socialemediaplatforms. Tegen deze achtergrond is er in de afgelopen jaren in razendsnel tempo een sterk Unierechtelijk desinformatiebeleid opgetuigd. Op grond van de *Digital Services Act* (hierna: DSA) van de Europese Unie (hierna: EU of Unie), in verbinding met de *Strengthened Code of Practice on Disinformation*, zijn socialemediaplatforms ertoe gehouden desinformatie tegen te gaan. Grote online platforms zoals Facebook en Instagram verminderen de zichtbaarheid van posts die door fact-checkers als desinformatie worden bestempeld significant.³ In deze bijdrage betoog ik dat het huidige beleid – waarbij posts die als desinformatie worden aangemerkt onzichtbaar worden gemaakt – vanuit democratisch perspectief onwenselijk is. Meer specifiek is het onverenigbaar met de deliberatieve democratietheorie van de Duitse filosoof Jürgen Habermas. Zijn werk biedt uitstekende aanknopingspunten om de verhouding tussen waarheid en democratie te doordenken. Dit fundamenteel-theoretische spanningsveld concretiseert zich op zijn beurt in de problematiek rondom desinformatie.

2. Desinformatie: morele paniek of echt gevaar?

Ten tijde van de coronapandemie sprak de Directeur-Generaal van de *World Health Organization*, Tedros Ghebreyesus, van een ‘infodemie’ die nog schadelijker zou zijn dan de verspreiding van het virus zelf.⁴ In een recente toespraak op de *Copenhagen Democracy Summit* presenteerde de voorzitter van de Commissie, Ursula von der Leyen, het *European Democracy Shield*, dat als belangrijke drijfveer het terugdringen van desinformatiecampagnes heeft.⁵ Hoge Vertegenwoordiger van de EU, Josep Borrell, stelt zelfs dat desinformatie de democratie vergiftigt.⁶ Het heersende narratief is duidelijk: desinformatie vormt een groot gevaar voor de democratie. In deze bijdrage zal ik desinformatie breed definiëren. Het is onjuiste of misleidende informatie die schadelijke publieke gevolgen kan hebben.⁷ Hoewel het verspreiden van desinformatie kan overlappen met strafrechtelijke begrippen zoals aanzetten tot haat, smaad en laster, wordt het begrip hier opgevat als restcategorie van typen van informatie die niet zonder meer strafbaar gesteld zijn, maar toch als onwenselijk worden beschouwd. Men kan bij het verspreiden van desinformatie bijvoorbeeld denken aan beïnvloedingsoperaties door Rusland, dat met slinkse campagnes chaos en verwarring probeert te zaaien. Zo is het inmiddels bekende ‘*Doppelgänger*’-programma’ operationeel, waarbij Westerse websites worden gekopieerd en onder een net andere naam worden gelanceerd, gevuld met pro-Russische propaganda.⁸ Ook kan men denken aan foutieve beeldvorming over Israël en Hamas, waarin aan foto’s en video’s een andere oorsprong wordt toegedicht met het oogmerk de publieke opinie te beïnvloeden.⁹ Ten tijde van de Nederlandse parlementsverkiezingen in 2023 is er door nationale media middels fact-checks evenzeer gewezen op desinformatie, bijvoorbeeld

* Mr. drs. Alexander de Wit is onlangs afgestudeerd in de masters Encyclopedie en Filosofie van het recht (LLM) en Philosophy of Law, Governance and Politics (MA). Momenteel is hij werkzaam als docent/onderzoeker aan de Universiteit Leiden, bij de afdeling encyclopedie van de rechtswetenschap.

1 Van der Linden e.a., 2024, p. 2.

2 Europese Commissie, mededeling COM(2018)236 final, ‘Mededeling van de Commissie aan het Europees Parlement, de Raad, het Europees Economisch en Sociaal Comité en het Comité van de Regio’s. Bestrijding van online-desinformatie: een Europese benadering’, p. 1.

3 Meta stelt: “We dramatically reduce the distribution of False and Altered posts, and reduce the distribution of Partly False to a lesser extent.” Zie Meta, ‘How Fact-checking works’, *Meta Transparency Center*, 2 april 2024, beschikbaar op: <https://transparency.meta.com/en-us/features/how-fact-checking-works/>.

4 Ghebreyesus 2020, beschikbaar op: <https://www.who.int/azerbaijan/news/item/29-06-2020-working-together-to-tackle-the-infodemic->.

5 Von der Leyen 2024, *Copenhagen Democracy Summit*, p. 2-3, beschikbaar op: <https://admin.epp.eu/files/uploads/2024/05/Ursula-von-der-Leyen-Democracy-Summit-Copenhagen-Denmark-14-May-2024.pdf>.

6 Borrell 2024, beschikbaar op: https://www.eeas.europa.eu/eeas/disinformation-and-foreign-interference-speech-high-representativevice-president-josep-borrell-eeas_en.

7 Dit concept sluit aan bij de door de Europese Commissie gehanteerde definitie in de *Strengthened Code of Practice on Disinformation* 2022, p. 1. Het klassieke onderscheid tussen misinformatie (het verspreiden van onjuiste informatie zonder opzet op de negatieve gevolgen) en desinformatie (verspreiding met opzet op de negatieve gevolgen) wordt hiermee in belangrijke mate genuanceerd. Dit is begrijpelijk, nu de mogelijke negatieve effecten zich ongeacht de intentie van de verspreider zullen doen gevoelen. Strafrechtelijk bezien blijft het opzetcriterium (uiteraard) wel relevant, zie daarvoor Siegel & Klos 2024, p. 690.

8 Alaphilippe e.a. 2022, ‘Doppelgänger, Media clones serving Russian Propaganda’, beschikbaar op: <https://www.disinfo.eu/doppelganger>; European Union External Action, ‘Doppelgänger strikes back: FIMI activities in the context of the EE24’, beschikbaar op: https://euvsdisinfo.eu/uploads/2024/06/EEAS-TechnicalReport-DoppelgangerEE24_June2024.pdf.

9 Kuypers 2023, ‘Factcheck-overzicht: beelden en berichten over de oorlog tussen Israël en Hamas’, beschikbaar op: <https://nieuwscheckers.nl/factcheck-overzicht-beelden-en-berichten-over-de-oorlog-tussen-israel-en-hamas/>.

ten aanzien van cijfers over migratie, klimaatverandering en regeldruk vanuit de EU.¹⁰ Het is begrijpelijk dat dergelijke vormen van desinformatie als onwenselijk worden beschouwd, nu het publiek hierdoor verkeerd geïnformeerd kan raken en zo tot oneigenlijke keuzes komt. Tegelijkertijd is het goed om de gevaren van desinformatie in perspectief te plaatsen. Bij drie veelgehoorde claims wil ik op grond van recentelijk empirisch onderzoek een stevige kanttekening plaatsen. Daarbij gaat het ten eerste om de vermeende prevalentie van desinformatie in het huidige publieke discours, ten tweede om de bijzondere overtuigingskracht die desinformatie zou hebben, en tot slot de contextuele setting waarbinnen desinformatie schadelijk zou zijn.

Ten eerste wordt vaak gedacht dat we in een ongeëvenaard tijdperk van onwaarheden leven; bovengenoemde citaten onderstrepen dat treffend. Woorden als ‘infodemie’, ‘misdroom’, ‘infocalyps’ en ‘post-truth’ zijn politiek-retorisch aantrekkelijk, maar als duiding van sociale problematiek zijn zij nogal overtrokken.¹¹ Een meta-analyse uit 2022 van tien empirische studies naar de blootstelling van desinformatie toont bijvoorbeeld aan dat het gemiddelde percentage desinformatie ten opzichte van juiste informatie zich bij Amerikaanse en Europese gebruikers van sociale media consistent tussen de 1% en 5% bevindt.¹² Ten tijde van de Europese Parlementsverkiezingen is geschat dat slechts 4% van het nieuws gedeeld op Twitter (nu: X) uit onbetrouwbare bron kwam.¹³ Perspectief is dus belangrijk. Zo circuleert er in de wetenschappelijke literatuur over desinformatie het cijfer dat negen miljoen mensen op sociale media interacteerden met liefst twintig desinformatienarratieven. Dat klinkt gigantisch, maar het is slechts 0,042% van het totale aanbod op dat platform, ervan uitgaande dat iedereen slechts één keer per week zou posten.¹⁴ Indrukwekkende cijfers over desinformatie moeten dus vooreerst in deze context worden gezien, anders ontstaat er een vertekend beeld.

Ten tweede rust het desinformatiebeleid van de EU op de gedachte dat onware dan wel misleidende informatie sneller wordt geloofd dan juiste informatie. Het toonaangevende wetenschappelijke tijdschrift *Science* publiceerde in 2018 een invloedrijk onderzoek waaruit zou volgen dat onjuiste informatie zich zes keer sneller dan juiste informatie verspreidde.¹⁵ Dit onderzoek bevatte evenwel een belangrijke *sampling bias*; alleen reeds als problematisch bestempelde en daarom gefact-checkte posts werden op hun wijdverspreidheid onderzocht.¹⁶ Veelvuldig gedeelde posts die niet beoordeeld en gefact-checkt waren – omdat zij niet problematisch geacht werden, bijvoorbeeld de foto van Messi die de wereldbeker vasthoudt – werden in het onderzoek niet meegenomen. De gehanteerde methodologie is zodoende sturend voor de uitkomst. Hetgeen men categoriseert en definieert als desinformatie is geen neutraal vertrekpunt maar vormt reeds een normatief geladen keuze. De claim in het onderzoek van Soroush Vosoughi e.a. (2018) lijkt sterk en algemeen toepasbaar, maar op grond van het onderzoek kan slechts worden gesteld dat door *fact-checkers*

als onwaar beoordeelde posts zich sneller verspreiden dan door hen als waar bestempelde posts.¹⁷

De Amerikaanse politicoloog Joseph Uscinski stelt dat onderzoeken naar desinformatie vertrekken vanuit de misvatting dat desinformatie bijzondere kwaliteiten zou bezitten op grond waarvan zij sneller wordt geloofd dan juiste informatie.¹⁸ De psychologische grondslagen van overtuigingsvorming zijn echter bij zowel juiste informatie als desinformatie steeds hetzelfde. Bepaalde vooraf vaststaande gegevens (zogenoemde *priors*) zijn daarbij bepalend. Het gaat dan om zaken als ideologie, persoonlijkheidskenmerken, achtergrondkennis, vertrouwen in instituties en psychologische fenomenen (zoals *biases*).¹⁹ Complotdenkers en anti-vaxxers gebruiken in de kern dezelfde methoden, zoals verificatie, *fact-checks* en het citeren van wetenschappelijke studies om hun standpunten te ondersteunen.²⁰ De redenen waarom mensen in desinformatie geloven, waaronder institutioneel wantrouwen, zijn daarmee relevanter dan het symptoom – het bestaan van desinformatie zelf. Daarbij zijn mensen in hun overtuigingsvorming geneigd sterk gebruik te maken van *motivated reasoning*, het fenomeen waarbij naar de eigen, al bestaande gedachte toe wordt geredeneerd.²¹ Geloof in de waarheid is daar het slachtoffer van, maar niet omdat de desinformatie zo overtuigend is. Veeleer wordt de informatie die in het voordeel van het eigen standpunt is – waar of onwaar – gebruikt, en de rest in de wind geslagen. Het politieke domein wordt dus primair bepaald door een ideeënstrijd, en slechts beperkt door een feitenstrijd.

Ten derde lijkt de mate waarin desinformatie wordt geloofd samen te hangen met de algehele kwaliteit van het medialandschap. Psycholoog en cognitief wetenschapper Hugo Mercier toont dat in zijn boek *Not Born Yesterday* aan voor propagandacampagnes in Rusland en China, landen waar het veel slechter gesteld is met de kwaliteit van (online) media dan in Europa. Zelfs daar hebben propagandacampagnes maar marginaal effect; zij slagen slechts beperkt wanneer mensen dezelfde overtuiging zijn toegedaan als de propaganda, maar hebben een averechts effect op mensen die al scherp zijn op desinformatie en kritisch kijken naar propaganda.²² In een pluriform medialandschap als het onze – de gemiddelde Nederlander gebruikt nog altijd zo’n vijf nieuwsbronnen per week en slechts 11% van de Nederlanders consumeert primair nieuws via sociale media – stemmen zulke onderzoeken hoopvol.²³

Het bovengenoemde kader nuanceert zodoende de wijdverspreidheid, overtuigingskracht en effecten van desinformatie in belangrijke mate. Er bestaat zodoende nogal een groot verschil tussen de uitspraken die academisch gezien geschraagd kunnen worden over het gevaar van desinformatie en de heersende gedachte van journalisten en ‘experts’.²⁴ In mijn optiek is de retoriek over desinformatie tegen deze achtergrond een voorbeeld van wat in de sociologie bekend staat als morele paniek: een disproportionele reactie op een bepaald overtrokken schrikbeeld.²⁵

10 EU Disinfo Lab 2023, ‘Disinformation landscape in the Netherlands’, beschikbaar op: https://www.disinfo.eu/wp-content/uploads/2023/09/20230919_NL_DisinfoFS.pdf; Pointer 2023, ‘Hoeveel van onze wetten komen uit Brussel?’, beschikbaar op: <https://pointer.kro-ncrv.nl/hoeveel-van-onze-wetten-komen-uit-brussel>.

11 Hannon 2023, p. 45.

12 Acerbi, Altay & Mercier 2023 p. 2, beschikbaar op: https://misinforeview.hks.harvard.edu/wp-content/uploads/2022/01/acerbi_fighting_misinformation_20220112.pdf.

13 Ibid.

14 Altay, Berriche & Acerbi 2023, p. 3.

15 Vosoughi et al., 2018, p. 1148.

16 Ibid.

17 Altay, Berriche & Acerbi 2023, p. 4.

18 Uscinski 2023, p. 3.

19 Ibid.

20 Altay, Berriche & Acerbi 2023, p. 6.

21 In de meta-analyse van Ditto et al. 2019, p. 283 vinden we bevestigd dat *motivated reasoning* gelijkelijk verdeeld over het gehele politieke spectrum plaatsvindt.

22 Mercier 2020, p. 131-132.

23 Digital News Rapport, Commissariaat voor de Media 2023, p. 17 en p. 26.

24 Zo heeft het World Economic Forum 1500 experts (let wel: academici, ambtenaren, managers van divers pluimage) gevraagd wat de grootste korte termijn bedreiging voor de gehele wereld (!) is, en desinformatie kwam op nummer 1, zie WEF, *Global Risks Report* 2024, p. 11.

25 Cohen 2002, p. vii.

2.1 Het huidige reguleringsraamwerk: van sturend naar dwingend

Desinformatie wordt middels een samenspel van *soft law* en *hard law* instrumenten gereguleerd. De basis ervan vindt men in de *Strengthened Code of Practice on Disinformation* (hierna: de Praktijkcode). Deze bouwt voort op de eerste versie uit 2018, een instrument bedoeld voor zelfregulering dat de Europese Commissie in nauwe samenwerking met grote digitale spelers zoals Meta, Google en TikTok opgetuigd heeft.²⁶ Deze eerste Praktijkcode uit 2018 was daarmee wel gezaghebbend voor de industrie, maar niet bindend. Met de komst van de DSA van 2022 ligt dat genuanceerder.²⁷ De DSA is een bindende verordening en beoogt de regulering van lidstaten aangaande digitale diensten, waaronder die van grote online platforms, te harmoniseren.²⁸ Onderdeel daarvan is een voor het platform verplichte analyse van systeemrisico's die voortspruiten uit de aard van de aangeboden dienst (artikel 34 DSA) en de verplichte mitigatie van die risico's (artikel 35 DSA). Het verbaast niet dat een belangrijk systeemrisico van grote online platforms als Facebook of TikTok er nu juist in bestaat dat desinformatie snel en op grote schaal verspreid kan worden.²⁹ Dat kan leiden tot negatieve effecten op de informatievoorziening van het publieke debat en verkiezingsprocessen (artikel 34 lid 1 sub c DSA). Artikel 35 lid 1 sub h DSA ziet het voldoen aan gedragscodes als een belangrijke manier om hiertegen redelijke, proportionele en effectieve maatregelen te nemen. Er zijn momenteel vergoederde plannen om de Praktijkcode te converteren tot een echte gedragscode onder de DSA.³⁰ Daarmee zal de Praktijkcode ook jaarlijks onafhankelijk gecontroleerd moeten worden (artikel 37 lid 1 sub h DSA) en komt zij onder het boeteregime, waarbij het niet voldoen aan de verplichtingen kan leiden tot een boete van maximaal 6% van de wereldwijde jaaromzet (artikel 52 lid 3 DSA). Hiermee komt de regulering van desinformatie Europeesrechtelijk in een co-reguleringsregime, waarbij de Uniewetgever tezamen met private partijen actie tegen desinformatie onderneemt.

3. Fact-checken in de hoofdrol

Het belangrijkste wapen uit de Praktijkcode is de inzet van onafhankelijke fact-checks in het contentmoderatiebeleid van platforms.³¹ Om duidelijk te maken hoe dat in de praktijk vormgegeven wordt, is het nuttig om het *fact-check*beleid van Meta (Facebook, Instagram) onder de loep te nemen. Meta contracteert – net zoals alle andere grote online platforms – alleen gecertificeerde institutionele *fact-checkers*. Deze zijn

aangesloten bij Europese en internationale *fact-check* organisaties zoals de *European Fact-Checking Standards Network*, *International Fact-Checking Network* of allebei. In de woorden van Meta: “We don’t think a private company like Meta should be deciding what’s true or false, which is exactly why we have a global network of fact-checking partners (...).”³² *Fact-checkers* verkrijgen hiermee een belangrijke machtspositie; platforms maken namelijk op grote schaal gebruik van hun diensten. Uit eigen rapportage blijkt dat Meta in de tweede helft van 2023 140.000 *fact-checks* van 26 *fact-check* organisaties gebruikt heeft om de zichtbaarheid van maar liefst 67.000.000 *posts* binnen de EU te reduceren.³³ Hoewel *posts* zodoende niet worden verwijderd, is het effect wel dat veel minder gebruikers de gecensureerde bijdragen te zien krijgen.³⁴ In die zin vervullen *fact-checkers* een poortwachtersrol op sociale media: hun bijdragen worden benut om te bepalen welke uitspraken prominent (bij waar geachte bijdragen) en minder prominent (bij onwaar geachte bijdragen) weergegeven worden. Desinformatie wordt bovendien voorzien van een waarschuwingslabel.³⁵ Voorts is dit beleid – zeker als de Praktijkcode de status van een DSA-gedragscode verwerft – juridisch bindend.³⁶ Elk groot online socialemediaplatform kan de verspreiding van desinformatie immers tot haar systeemrisico's rekenen, waardoor het moeilijk voorstelbaar is hoe platforms zonder *fact-checks* aan hun mitigatieverplichtingen zouden kunnen voldoen. Dat de EC hier strikt op zal toezien blijkt uit een tweet van Eurocommissaris Thierry Breton gericht aan platform X: “you can run, but you can’t hide.”³⁷ De ratio van het huidige beleid is simpel: onwaarheden moeten worden weersproken met de waarheid.³⁸ In een mededeling aan het Europees Parlement (hierna: ‘EP’) stelt de EC dat onze open democratische samenlevingen afhankelijk zijn van het publieke debat en een goed geïnformeerd publiek.³⁹ Dat publiek staat immers aan de basis van het beleids- en wetgevingsproces, doordat zij input levert ten aanzien van wat er moet gebeuren. Dit geschiedt niet alleen middels verkiezingen, maar in toenemende mate ook door het kenbaar maken van wensen via sociale media.⁴⁰ In de volgende paragrafen zal ik de grondslagen en gevolgen van dit beleid kritisch belichten vanuit de deliberatieve democratietheorie van Habermas.

4. Normatieve reflectie: waarheid en deliberatieve democratie

Het huidige EU-beleid ten aanzien van desinformatie heeft zich in rap tempo ontwikkeld. De inmiddels alomtegenwoordige praktijk van het *fact-checken* is met existentiële claims over de

26 Mededeling Com(2018)236, ‘Mededeling van de Commissie aan het Europees Parlement, de Raad, het Europees Economisch en Sociaal Comité en het Comité van de Regio's. Bestrijding van online-desinformatie: een Europese benadering’, p. 7.

27 Verordening (EU) 2022/2065 (DSA).

28 DSA, preamble lid 2.

29 Wikforss 2023, p. 98.

30 Mededeling Com(2021)262 final, ‘Mededeling van de Commissie aan het Europees Parlement, de Raad, het Europees Economisch en Sociaal Comité en het Comité van de regio's, Richtsnoeren van de Europese Commissie ter versterking van de praktijkcode betreffende desinformatie’, p. 3; Zoals bevestigd in de derde vergadering van de European Board of Digital Services, 25 april 2024. <https://digital-strategy.ec.europa.eu/en/news/third-meeting-european-board-digital-services>, p. 3.

31 Praktijkcode paragraaf VII, sub c.

32 Facebook, 1 juni 2021, beschikbaar op: <https://www.facebook.com/formedia/blog/third-party-fact-checking-how-it-works>.

33 Rapport inzake de Praktijkcode van Meta tot en met maart 2024, p. 3. (Hierna: rapport Meta 2024).

34 Ook generatieve kunstmatige intelligentie wordt ingezet die op grote schaal automatisch bijdragen van gebruikers cureert, zie rapport Meta 2024, p. 147.

35 Zie Meta, ‘How Fact-checking works’, *Meta Transparency Center*, 2 april 2024, beschikbaar op: <https://transparency.meta.com/en-us/features/how-fact-checking-works/>.

36 DSA overweging 104, zie ook Husovec 2024, p. 447.

37 Thierry Breton, tweet 26 mei 2023, beschikbaar op: <https://x.com/ThierryBreton/status/1662194595755704321>. X heeft zich in 2023 teruggetrokken uit de Praktijkcode. In augustus 2024 stuurde Breton voorts zelfstandig een uitgebreide waarschuwingsbrief waarin hij X's baas Elon Musk er indringend aan herinnert dat de DSA ook voor X geldt. Zie Thierry Breton, tweet 12 augustus 2024, beschikbaar op: <https://x.com/ThierryBreton/status/1823033048109367549>.

38 European Commission, ‘European elections: EU institutions prepared to counter disinformation’, 6 juni 2024, beschikbaar op: <https://digital-strategy.ec.europa.eu/en/news/european-elections-eu-institutions-prepared-counter-disinformation>.

39 Mededeling Com(2018)236, ‘Mededeling van de Commissie aan het Europees Parlement, de Raad, het Europees Economisch en Sociaal Comité en het Comité van de Regio's. Bestrijding van online-desinformatie: een Europese benadering’, p. 1.

40 Mededeling Com(2018)236, ‘Mededeling van de Commissie aan het Europees Parlement, de Raad, het Europees Economisch en Sociaal Comité en het Comité van de Regio's. Bestrijding van online-desinformatie: een Europese benadering’, p. 2.

gevaaren van desinformatie gerechtvaardigd. Wat tot op heden evenwel ontbreekt is een grondige normatieve reflectie op de wenselijkheid en verenigbaarheid van het huidige beleid in het licht van onze democratische principes. In paragraaf 2 heb ik het gevaar van desinformatie al naar realistischer proporties proberen terug te brengen. In deze en volgende paragraaf ga ik in op de huidige inzet van *fact-checks* als belangrijkste wapen in de strijd tegen desinformatie en de kritiek die daarop te leveren is. De deliberatieve democratietheorie van de Duitse filosoof en kritisch socioloog Jürgen Habermas (g. 1929) vormt daartoe het normatieve kader. Zijn werk biedt uitstekende aanknopingspunten om de verhouding tussen waarheid en democratie te doordenken. Habermas' theorie geeft aan hoe conflicterende waarheidsclaims die aan de basis van complexe wetgeving staan vanuit democratisch oogpunt toch legitiem kunnen zijn. Ik zal betogen dat de manier waarop waarheidsclaims van *fact-checkers* worden gebruikt niet aan de eis van democratische legitimiteit kunnen voldoen, en de rechtvaardiging van daarin gepresenteerde waarheidsclaims zelfs bemoeilijkt. Steeds toon ik met concrete voorbeelden aan hoe de theoretische bezwaren zich ook daadwerkelijk in de praktijk hebben verwezenlijkt.

De deliberatieve democratietheorie behelst eerst en vooral een theorie over normatieve politieke legitimiteit. Het geeft aan onder welke voorwaarden er in een democratie op gezaghebbende wijze macht mag worden uitgeoefend over de bevolking. Daartoe formuleert Habermas het vermaarde Democratisch Principe: "alleen die wetten kunnen legitimiteit claimen die de toestemming van alle burgers genieten in een discursief wetgevingsproces dat zelf ook legitiem geconstrueerd is."⁴¹ Dit citaat is complex en zal ik in verschillende delen toelichten. Met een discours bedoelt Habermas een geïdealiseerde situatie waarin participanten over en weer delibereren over de redenen die aan hun standpunten ten grondslag liggen. In een politieke deliberatie kan elke burger zodoende een authentieke keuze maken, die gevrijwaard is van allerlei feitelijk bestaande illegitieme machtsstructuren zoals corruptie, discriminatie en schimmige lobbypraktijken.⁴² Zijn theorie vormt hiermee een belangrijke correctie op de zogeheten aggregatieve democratiemodellen, die legitimiteit simpelweg gelijkstellen met de meerderheidsbeslissing van het volk. Tegen die achtergrond heeft Habermas het discours wel een *Ideal Speech Situation* genoemd, waarin men alleen nog overtuigd raakt van de "dwangloze dwang van het betere argument."⁴³ Een dergelijke situatie bestaat alleen als er sprake is van gelijkheid van spreken, inclusieve deelneming en geen interne of externe dwang.⁴⁴ Met andere woorden: met het wegvallen van illegitieme machtsstructuren houden alleen de overtuigingskracht en de kwaliteit van de argumenten stand. Habermas is zich terdege bewust van de weerbarstige realiteit waarin deze eisen nooit volledig verwerkelijk kunnen worden.⁴⁵ Zolang wij als burgers en gesprekspartners serieus nader tot elkaar willen komen, moeten wij dergelijke criteria echter wel als uitgangspunt nemen.⁴⁶

Habermas reconstrueert democratische legitimiteit hiermee in louter procedurele termen. Dat betekent dat wetgeving legitiem is omdat bepaalde procedures zijn gevolgd. De procedure begint met de veronderstelling van de *Ideal Speech Situation*, waarin zich een communicatieve rationaliteit vrij van macht ontvouwt. In een diverse en pluriforme democratie als de onze is het aantrekkelijk om een dergelijk procedurele conceptie van legitimiteit te hanteren. Zo wordt voorkomen dat er bij voorbaat inhoudelijke arbitraire criteria moeten worden opgesteld om te bepalen wat nu het juiste of beste beleid is.⁴⁷ Complexe sociale vraagstukken zoals klimaatverandering en migratie raken ons allemaal, en niemand heeft daartoe de volledige waarheid in pacht. Hoewel wetenschappelijke kennis van experts ons kan informeren, kan zij niet legitimeren wat wij moeten doen. Feiten stellen immers geen normen. Tegelijkertijd heeft een beroep op een Godheid of Ideeënwereld in ons zogeheten post-metafysisch tijdperk zijn normatieve status verloren, waardoor de geldigheid van een waarheidsaanspraak uitsluitend nog besloten kan liggen in de intersubjectieve deliberatie van participanten.⁴⁸ De enige manier waarop politieke legitimiteit heden ten dage gerechtvaardigd kan worden, is zodoende via het eerder besproken discours.⁴⁹ Een inhoudelijke, epistemische eis zoals 'alleen ware uitspraken mogen ingebracht worden' is onverenigbaar met Habermas' project, precies nu die waarheid niet reeds voor het discours kan worden vastgesteld.⁵⁰ Sterker nog, als we een bepaalde externe epistemische toets zouden inbouwen in Habermas' theorie – waarmee de waarheid van uitspraken getoetst zou kunnen worden – dan ondergraaft dat zelfs de procedurele legitimiteit die via het discours kan worden bewerkstelligd. Het geven van redenen door participanten zou dan simpelweg worden vervangen door een finaal beroep op een bepaalde autoriteit, zoals expert-organen of kennisinstituten (denk aan het *Intergovernmental Panel on Climate Change*, of IPCC, inzake klimaatverandering).⁵¹

Wat vermag waarheid dan in een deliberatieve democratie? Ook hier geldt weer: waarheidsclaims kunnen alleen discours-theoretisch gerechtvaardigd worden. Neem als voorbeeld klimaatbeleid, bij uitstek een complex sociaal vraagstuk waarin feit en norm door elkaar lopen. Wat de waarheidsclaims hier *democratische* legitimiteit verschafft is niet de epistemische autoriteit van experts. Zij vormen wellicht een belangrijke bron van informatie, maar kunnen ons niet dicteren wat daarmee te doen, hoe die informatie te wegen en te implementeren in het beleid. Tegelijkertijd is het te verwachten dat burgers fundamenteel tegenstrijdige opvattingen zullen hebben over de oplossingen voor een bepaald probleem. Wat er vervolgens moet gebeuren als wij als redelijk handelende burgers elkaar willen overtuigen is het volgende. Ten eerste moeten burgers de criteria van de *Ideal Speech Situation* als regulatief ideaal veronderstellen.⁵² Voorts moeten wij expliciteren dat wij waarheidsclaims doen, en die vervolgens als feilbaar hanteren. Een participant die het discours betreedt en de onomstotelijke waarheid in pacht denkt te hebben, staat niet serieus open voor andere visies en opvattingen. Idealiter ont-

41 Habermas 1996, p. 110.

42 Habermas 1996, p. 305-306.

43 Habermas 1996, p. 306.

44 Habermas 1996, p. 305.

45 Habermas noemt dat in een recent interview uit 2023 zelfs een hardnekkig misverstand, zie Habermas 2023.

46 Chambers 2019, p. 95; Habermas 2023.

47 Hannon 2024, p. 4. Zie voor een voorbeeld van zo een criterium Peter 2021, p. 150, 'een politieke bijdrage is geldig dan en alleen dan als deze waar is.' Peter stelt dat deze norm arbitrair en te veelzijdig is. Wie bepaalt vooraf wat waar is? Bovendien kan later onjuiste informatie alsnog een belangrijke (e.g. agenda-setting of kritische) democratische rol vervullen

48 Habermas 1992, p. 74.

49 Habermas 1996, p. 121.

50 Zie voor de onverenigbaarheid van dergelijke epistemische normen in deliberatieve theorieën: Peter 2021.

51 Wellmer 2004, p. 198.

52 Er zijn veel experimenten gedaan met zogeheten *mini-publics*, waarin kleine, diverse groepen burgers met elkaar delibereren. Resultaten stemmen positief. Burgers raken beter geïnformeerd en zijn begripvoller naar elkaar, maar de mate waarin generalisatie naar de 'echte' wereld mogelijk is blijft voorwerp van discussie, zie: Chambers 2021, p. 152 en Hannon 2024, p. 12-13.

staat er na verloop van tijd een rationele consensus. Deze consensus vormt het enige criterium voor wat wij als waar mogen beschouwen in een democratie. Uiteraard kan de ‘echte’ waarheid daarmee alsnog anders zijn – we zijn immers feilbaar – maar zoals gezegd bestaat er geen Archimedisch nulpunt van waaruit we dat zouden kunnen toetsen. Mensen zijn geen goden, en het middels dialoog nader tot elkaar komen is het enige wat wij kunnen doen.⁵³ Een feitelijke claim (antropogene klimaatverandering bestaat) of een morele claim (wij zouden moeten stoppen met het eten van vlees) mogen dan nog zo wetenschappelijk verantwoord zijn, hun democratische geldigheid bestaat pas bij het bereiken van intersubjectieve consensus.⁵⁴

4.1 Deliberatieve democratie en institutioneel design: two-track democracy

Zoals gezegd houdt Habermas zich niet alleen bezig met filosofie, maar werkt hij zijn theorie ook op meer praktisch-sociologische wijze uit. Zijn conceptie van een ‘two-track democracy’ is daarbij in het bijzonder relevant.⁵⁵ Het reeds besproken discours vormt namelijk het sluitstuk van de deliberatieve democratie, maar voordat argumenten daar gewogen worden, is hetgeen opkomt in de publieke sfeer een belangrijk eerste peilmoment van wat er leeft in de maatschappij. De publieke sfeer kan breed worden gedefinieerd als het informele domein waarin politiek wordt gecommuniceerd.⁵⁶ Het gaat daarbij dus niet om formeel-democratische dialogen zoals in het parlement gevoerd worden, maar juist om een vrijer, procedureel ongeregeld domein waarin opinies en standpunten vrijelijk gevormd moeten kunnen worden. Ook socialemediaplatforms kunnen bij uitstek worden gezien als een plaats waar de (digitale) publieke sfeer vorm krijgt.⁵⁷ Hier gelden de discours-theoretische voorwaarden nog niet, omdat deze context van ontdekking juist zo cruciaal is om de onbemiddelde en ongezoeten meningen van burgers te vergaren. Habermas spreekt zelfs van een “wild geheel dat elke vorm van organisatie weerstaat.”⁵⁸ Tezamen hebben de publieke sfeer (de eerste *track*) en het discours (de tweede *track*) een ‘waarheidspotentieel’.⁵⁹ Het discours fungeert dan uiteindelijk als een soort rationele trechter, die belangrijke sociale thema’s uit de kakofonische vergaarbak van meningen en standpunten die in de publieke sfeer tot wasdom gekomen zijn destilleert.

5. Het huidige beleid getoetst aan de theorie

Institutionele *fact-checkers* verkrijgen een geprivilegieerde poortwachtersrol die in beide *tracks* van de democratie niet past. Op het niveau van de digitale publieke sfeer hebben we gezien dat het van belang is een zo groot mogelijke hoeveelheid aan ongefilterde meningen en standpunten te verzamelen. Het is immers van belang dat eenieder de sociale vraagstukken die zij belangrijk vindt kan agenderen. Daarbij hoeft de geuite kritiek niet waar te zijn, maar kan zij ook onderdeel zijn van identiteits-expressieve taal, waarin burgers hun positie duidelijk maken aan de hand van claims die in beginsel losstaan van

de waarheid.⁶⁰ Het huidige beleid heeft tot gevolg dat in beginsel elke feitelijke claim in de digitale publieke sfeer zou kunnen worden onderworpen aan een *fact-check*. Het problematische aspect is voornamelijk gelegen in het feit dat *posts* die als onwaar of misleidend worden bestempeld vervolgens bijna niet meer te zien zijn. Zo krijgen we een digitale publieke sfeer waarin de interpretatie van de feitelijke werkelijkheid door *fact-checkers* de norm wordt. Maar wie controleert hier de *fact-checkers*? Door geautomatiseerde vormen van *fact-checken* is het niet duidelijk welke bijdrage om welke reden minder zichtbaar is. Bovendien blijkt uit onderzoek dat verschillende *fact-checkers* het vaak met elkaar oneens zijn over beoordelingen van bijdragen.⁶¹ Een concrete verwezenlijking van deze problematiek zagen we in coronatijd, toen er veel discussie ontstond over de oorsprong van het virus. De zogeheten ‘*lab leak theory*’ – de theorie dat het virus uit een lab in Wuhan zou komen – werd door *fact-checkers* al snel als desinformatie bestempeld. Een tijd later kwamen er toch signalen die in ieder geval een serieus onderzoek door wetenschappers hiernaar legitimeerden.⁶²

Het gaat in bovenstaand voorbeeld niet om de uiteindelijke waarheid van deze theorie, maar om het bij voorbaat uitsluiten van belangrijke kritieken, die middels het huidige beleid aan de publieke sfeer worden onttrokken. De taakopvatting van *fact-checkers* is evenwel dat zij bringers van ‘objectieve’ waarheden zijn, vrij van potentiële vooroordelen en normatieve keuzes.⁶³ Tegelijkertijd blijkt uit interviews dat *fact-checkers* zichzelf niet zien als louter controleurs van informatie, maar ook als actieve *influencers* in de samenleving.⁶⁴ Dat een dergelijke rolperceptie belemmerend kan werken voor de publieke sfeer illustreert hoogleraar sociale epistemologie Jeroen de Ridder treffend in zijn recente oratie.⁶⁵ Hij noemt daar het voorbeeld van de website *UnHerd*, die kritische beschouwingen van verscheidene auteurs post met als doel tegen de kudde in te denken.⁶⁶ Door de *Global Disinformation Index* (GDI)⁶⁷ – een groep journalisten, *fact-checkers* en deskundigen die bedrijven een score toekent – had *UnHerd* evenwel een lage score toegekend gekregen. Die score had vergaande gevolgen voor hun advertentie-inkomsten. De Ridder stelt dat de reden van het toekennen van de lage score gelegen bleek te zijn in het feit dat de feministische filosofe Kathleen Stock kritische beschouwingen over bepaalde aspecten van transgenderwetgeving op het platform plaatste. Het gaat mij – net zoals de Ridder ook opmerkt – niet om wie hier *inhoudelijk* gelijk heeft. Wel is het kwalijk dat een kritisch platform door een grotendeels achter de schermen opererende organisatie wordt gedelegeitimeerd door van het verspreiden van desinformatie beschuldigd te worden. Zeker nu desinformatie vaak wordt verbonden aan Russische propaganda en vergaande vormen van informatie-beïnvloeding (zie par. 2) is het belangrijk om in te zien wat dergelijke verwijten voor gevolgen kunnen hebben.⁶⁸ Het zou democratisch gezien veel zuiverder zijn als het publiek hierover zelf, in alle openheid, kan delibereren, en eigenstandig tot een uiteindelijke beoordeling van de waarheidsclaims van *UnHerd* komt. Die rol moeten wij gelet op Habermas’ raamwerk niet toekennen aan een private organisatie zoals de GDI.

53 Habermas 1992 p. 36.

54 Habermas 1996, p. 111; Held 2008, p. 6.

55 Habermas 1996, p. 307.

56 Chambers 2021, p. 154.

57 Habermas 2023.

58 Habermas 1996, p. 307.

59 Habermas 1996, p. 308.

60 Hannon 2023, p. 42.

61 Uscinski 2023, p. 7.

62 Ibid.

63 Bengtsson & Schousboe 2024, p. 3.

64 Micallef e.a. 2022, p. 27.

65 De Ridder 2024 (oratie VU), beschikbaar op: https://www.youtube.com/watch?v=GfzQv7ouB_c.

66 <https://unherd.com/us/>.

67 <https://www.disinformationindex.org/about>.

68 In de wens desinformatie zoveel mogelijk te bestrijden lijkt dit punt nog wel eens ondergesneeuwd te raken.

Zouden *fact-checks* dan een geprivilegieerde rol kunnen vervullen in het discours zelf? Ook hier geldt het antwoord ontkennend. De procedurele aard van democratische legitimiteit verzet zich tegen het toekennen van een geprivilegieerde status aan bepaalde epistemische autoriteiten. In het huidige beleid fungeren *fact-checks* namelijk niet simpelweg als een bijdrage van een participant, maar verkrijgen zij een bijzondere status in het contentmoderatiebeleid doordat zij een-op-een worden overgenomen door sociale media die er vervolgens de zichtbaarheid van andere bijdragen mee reduceren. Zo wordt reeds voor de deliberatie en discursieve rechtvaardiging al verondersteld wat waardevolle en niet-waardevolle feitelijke bijdragen zijn. Gelet op bovengenoemde bezwaren die kleven aan de huidige *fact-check* praktijk is dat ernstig, nu het discours hiermee kwetsbaarder wordt voor de partijdigheid en feilbaarheid van *fact-checks*, zonder dat op transparante wijze inzichtelijk is hoe en in welke mate zij precies in het beleid worden gebruikt.⁶⁹

6. Conclusie

De conclusie is eenvoudig: het huidige desinformatiebeleid staat op gespannen voet met de deliberatieve democratietheorie van Habermas. De poortwachtersrol van *fact-checkers* leidt ertoe dat zij in belangrijke mate bepalen welke bijdragen in de digitale publieke sfeer als desinformatie gelden, en daarmee *a contrario* wat als waarheid te gelden heeft. In de publieke sfeer belemmert dit de vrije ontwikkeling van een robuust debat over de wenselijkheid van sociale vraagstukken. Meer fundamenteel ondergraaft de huidige geprivilegieerde epistemische positie van *fact-checkers* de procedurele uitgangspunten waarop het discours gestoeld is. Nog voordat alle participanten een rationele consensus kunnen bereiken over de validiteit van waarheidsclaims, worden deze immers door een kleine groep aan het discours onttrokken. De procedurele aard van de legitimiteit in de deliberatieve democratie wordt daarmee miskend en omzeild. Tegen deze achtergrond kom ik tot de slotsom dat *fact-checks* niet meer gebruikt zouden moeten worden om de zichtbaarheid van bijdragen te verminderen. Wel zouden *fact-checkers* een meer open en adviserende rol op zich kunnen nemen. Transparantie over de rol en invloed van dergelijke *fact-checks* is dan echter wel geboden, evenals het inzicht dat ook *fact-checks* feilbaar zijn. In bredere zin noopt het vraagstuk van *fact-checken* tot een nadere doordenking van de verhouding die waarheid en democratie met elkaar hebben. In hoeverre mag de waarheid democratisch debat opschorten? Is het legitiem om aan de epistemische autoriteit van experts meer waarde toe te kennen dan aan bijdragen van 'gewone' burgers? Hoe zorgen we ervoor dat 'de feiten' niet gebruikt worden om contingent-ideologische posities als natuurlijk noodzakelijk te presenteren? Dergelijke vragen zullen in toenemende mate relevant worden nu de Europese Unie voortvarend gewag maakt met haar desinformatiebeleid.

BIBLIOGRAFIE

Literatuur

Acerbi, Altay & Mercier 2023

A. Acerbi, S. Altay & H. Mercier, 'Research Note: Fighting Misinformation or fighting for information', *Harvard Kennedy School Misinformation Review* 2023, p. 1-15.

Altay, Berriche & Acerbi 2023

S. Altay, M. Berriche & A. Acerbi, 'Misinformation on Misinformation: Conceptual and Methodological Challenges', *Soc. Med + Soc* 2023, p. 1-13.

Bengtsson & Schousboe 2024

M. Bengtsson & S. Schousboe, 'And That's a Fact: A Rhetorical Perspective on the Role of Fact-Checkers', *Journal, Pract.* 2024, p. 1-19.

Chambers 2019

S. Chambers, 'Deliberative Democracy', in: A. Allen & E. Mendieta (eds.), *The Cambridge Habermas Lexicon*, Cambridge: Cambridge University Press 2019, p. 94-97.

Chambers 2021

S. Chambers, 'Truth, Deliberative Democracy, and the Virtues of Accuracy: Is Fake News Destroying the Public Sphere?', *Political Studies* 2021, p. 147-163.

Cohen 2002

S. Cohen, *Folk Devils and Moral Panics. The creation of the Mods and Rockers*, Routledge: New York 2002.

Ditto et al. 2019

P.H. Ditto et al., 'At Least Bias is Bipartisan: A Meta-Analytic Comparison of Partisan Bias in Liberals and Conservatives', *Perspect. Psychol. Sci.* 2019, p. 273-291.

Habermas 1992

J. Habermas, *Postmetaphysical Thinking. Philosophical Essays*, Cambridge: MIT Press 1992.

Habermas 1996

J. Habermas, *Between facts and Norms. Contributions to a Discourse Theory of Law and Democracy*, Cambridge MA: The MIT Press 1996.

Habermas 2023

J. Habermas, *A New Structural Transformation of the Public Sphere and Deliberative Politics*, Cambridge: Polity Press 2023.

Hannon 2023

M. Hannon, 'The Politics of Post-Truth', *Crit. Rev.* 2023/35, p. 40-62.

Hannon 2024

M. Hannon, 'Betting Democracy on Epistemology', *Episteme* 2024 (komend), p. 1-26.

Held 2008

D. Held, *Models of Democracy*, Stanford: Stanford University Press 2008.

⁶⁹ In deze bijdrage worden overige epistemologische en methodologische bezwaren die aan de huidige *fact-check* praktijk kleven buiten beschouwing gelaten. Zie daarvoor Uscinski 2015 en Bengtsson & Schousboe 2024.

Husovec 2024

M. Husovec, *Principles of the Digital Services Act*, Oxford: Oxford University press 2024.

Mercier 2020

H. Mercier, *Not Born Yesterday. The science of Who we trust and What we believe*. Princeton: Princeton University Press 2023.

Micallef 2022

N. Micallef e.a., 'True or False: Studying the Work Practices of Professional Fact-Checkers', *Proc. ACM Hum.-Comput. Interact* 2022/127, p. 1-44.

Peter 2021

F. Peter, 'Epistemic Norms of Political Deliberation', in: M. Hannon & J. de Ridder (red.), *The Routledge Handbook of Political Epistemology*, New York: Routledge 2021, p. 395-405.

Siegel & Klos 2024

I.D. Siegel & M. Klos, 'Straffe desinformatiewetgeving? Bedenkingen vanuit Franse en Europese Unie-wetgeving' *AA* 2024, p. 683-691.

Uscinski 2015

J.E. Uscinski, 'The Epistemology of Fact-Checking (Is Still Naïve): Rejoinder to Amazeen', *Critical Review* 2015, p. 243-252.

Uscinski 2023

J.E. Uscinski, 'What Are We Doing When We Research Misinformation', *Political Epistemology* 2023, p. 2-14.

Van der Linden e.a. 2024

S. van der Linden e.a., 'Liars know when they are lying: differentiating disinformation from disagreement', *Humanit. Soc.* 2024/986, p. 1-14.

Wellmer 2024

A. Wellmer, 'The Debate About Truth, Pragmatism without Regulative Ideas' in: D. Freudlieb e.a. (red.), *Critical Theory after Habermas. Encounters and Departures* Leiden: Brill 2004, p. 181-211.

Wikforss 2023

A. Wikforss, 'The Dangers of Disinformation' in: H. Samaržija & Q. Cassam (red.), *The Epistemology of Democracy*, Oxfordshire: Routledge 2023, p. 90-112.