

# Online discriminatie tegengaan: hoe ziet de huidige aanpak eruit en wie is aan zet?

Een rapport voor organisaties die een rol (kunnen) spelen bij het tegengaan van online discriminatie.



# Inhoud

<b>Inleiding .....</b>	<b>3</b>
<b>De huidige aanpak van online discriminatie .....</b>	<b>7</b>
Europese wetgeving en gedragscodes.....	7
Nederlandse wetgeving .....	10
Aanpak vanuit onderwijs, maatschappelijke organisaties en gebruikers .....	13
Aanpak vanuit socialemediabedrijven .....	16
<b>Wie zijn aan zet? .....</b>	<b>20</b>
<b>Bijlage: begrippenlijst .....</b>	<b>24</b>
<b>Literatuur .....</b>	<b>26</b>

# Inleiding

In dit rapport zet Kennisplatform Inclusief Samenleven (KIS) de huidige aanpak van online discriminatie in Nederland uiteen. We gaan in op de Europese en Nederlandse wet- en regelgeving die hier van toepassing op is evenals de huidige aanpak vanuit socialemediabedrijven, onderwijs en maatschappelijke organisaties. Vervolgens geven we een overzicht van Nederlandse organisaties en actoren die aan zet zijn bij de huidige en toekomstige aanpak van online discriminatie. Een [ander rapport](#) is gewijd aan de aanbevelingen voor deze organisaties. De inzichten zijn gebaseerd op de expertise en werkervaring van geïnterviewde experts en inzichten uit de internationale wetenschappelijke literatuur.

Internet en sociale media zijn niet meer weg te denken uit ons leven. In 2021 had 97 procent van de Nederlanders thuis toegang tot het internet en maakte 87,6 procent hiervan dagelijks gebruik (CBS, 2021). Deze ontwikkeling in digitaal gebruik maakt het mogelijk dat mensen hun gedachten, gevoelens en overtuigingen snel en gemakkelijk delen met miljarden andere internetgebruikers. Het zorgt ervoor dat individuen en gemeenschappen over de gehele wereld continu met elkaar in verbinding kunnen staan. De toegankelijkheid tot het internet heeft ook een keerzijde. Juist het internet, en in het bijzonder sociale media, biedt een gemakkelijke mogelijkheid om je negatief richting anderen te uiten, bijvoorbeeld in haatvolle of discriminerende reacties (Burnap & Williams, 2016; Costa & Halpern, 2019; Van Huijstee, Nieuwenhuizen, Sanders, Masson &

Van Boheemen, 2021). Zo blijkt uit een peiling van het Netwerk Mediawijsheid (2022) dat bijna twee derde van de Nederlanders wel eens ervaart dat iemand online wordt gekwetst, bijvoorbeeld in de vorm van racisme, online shaming of buitensluiten. In buurland Duitsland geeft bijna de helft van de internetgebruikers aan, voornamelijk jongeren, vaak of heel vaak racistische reacties online te zien (Landesanstalt für Medien NRW, 2018).

## Internet en haat

Het internet, en sociale media in het bijzonder, biedt een gemakkelijk podium voor het uiten van haat en discriminatie. Mensen voelen zich online onzichtbaar, beschermd en veilig. Dit kan een versterkende factor zijn voor het uiten van immoreel gedrag (Fox, Cruz & Young Lee, 2015; Klivington & Price, 2017). Online is het eenvoudig om anoniem en op afstand deel te nemen aan discussies. Er wordt (veel) minder gevreesd voor de sociale consequenties van uitlatingen, waardoor deze online vaker pesterig of beledigend worden geformuleerd (Barlett, Gentile & Chew, 2016). Ook dragen algoritmes bij aan de verspreiding van haatberichten. Zo blijkt uit onderzoek dat sociale media prioriteit geven aan expressieve, polariserende content (Matthew, Dutt, Goyal & Mukherjee, 2019). Een overvloed aan dergelijke content kan resulteren in een normverschuiving over welk gedrag online acceptabel is.

## Wat is online discriminatie?

Online discriminatie, bijvoorbeeld geuit op sociale media, kan gericht zijn tegen een persoon of tegen een groep. Individuele online discriminatie houdt in dat een persoon wordt gedemigreerd of buitengesloten op basis van kenmerken als etniciteit, religie, sekse, seksuele voorkeur of huidskleur door het gebruik van online content zoals *emoticons*, gesproken berichten, filmpjes, afbeeldingen, memes, gifjes en teksten. Het kan ook zijn dat online discriminatie niet direct gericht is op een individu, maar dit wel door een individu als discriminerend wordt ervaren. Het gaat hier bijvoorbeeld om ‘grappen’, negatieve afbeeldingen of beledigingen die worden geuit over de groep waartoe iemand zichzelf rekent (Felten & Taouanza, 2017).

Vaak wordt het onderscheid gemaakt tussen strafbare en ervaren discriminatie. Strafbare discriminatie gaat om discriminatie die is aangemerkt in het Wetboek van Strafrecht. De gronden waarop discriminatie verboden is [in artikel 137d](#), zijn: ras<sup>1</sup>, godsdienst of levensovertuiging, geslacht, hetero- of homoseksuele gerichtheid of lichamelijke, psychische of verstandelijke handicap. Bij ervaren discriminatie kan het gaan over discriminatie op gronden die niet in de wet zijn aangemerkt als discriminatiegrond, om ervaringen of gevoelens waar geen juridisch vervolg aan is gegeven, of om ervaringen die niet of moeilijk objectief vast te stellen zijn. Bijvoorbeeld wanneer discriminatie plaatsvindt in de vorm van micro-agressies, alledaagse handelingen die als discriminerend kunnen worden ervaren, of wanneer het gaat om een duidelijk gevoel wat lastig aan te tonen is met objectief bewijs (Andriessen, Hoegen Dijkhof, Van der Torre, Van den Berg, Pulles, Iedema & De Voogd-Hamelink, 2020).

1 Binnen KIS geven wij de voorkeur aan schrijven over discriminatie op grond van afkomst, huidskleur en/of religie en vermijden we de term ‘ras’ zoveel als mogelijk. Biologisch gezien bestaan er namelijk geen menselijke rassen en bestaat het alleen als een sociaal construct dat is bedacht om mensen op arbitraire gronden (zoals huidskleur) te categoriseren volgens een bepaalde rangorde. De term ras komt echter wel voor in wet- en regelgeving, waardoor we in verwijzingen toch soms schrijven over ras. In de juridische context wordt discriminatie op grond van ras ruim uitgelegd en kan het gaan over huidskleur, afkomst of nationale of etnische afstamming.

## Online discriminatie of haatspraak?

Begrippen die vaak door elkaar worden gehaald zijn online discriminatie en online haatspraak (of *hatespeech* in het Engels). Beide vormen zijn onder te brengen onder de overkoepelende term online haat, waaronder ook online bedreigingen en intimidatie zijn te vangen (Van Huijstee et al., 2021). Wat de definiëring en het verschil tussen online discriminatie en haatspraak lastiger maakt, is dat er verschillende definities in omloop zijn en er vanuit nationaal en Europees perspectief, evenals vanuit verschillende organisaties, anders naar wordt gekeken. Over het algemeen is online discriminatie smaller gedefinieerd dan haatspraak. In 2022 bracht de Raad van Europa de volgende definitie naar buiten:

*Online haatspraak zijn alle soorten uitingen die aanzetten tot geweld of geweld bevorderen, verspreiden of rechtvaardigen, haat of discriminatie tegen een persoon of een groep personen, of die hen kleineren vanwege hun daadwerkelijke of toegeschreven persoonlijke kenmerken of status zoals ‘ras’, kleur, taal, religie, nationaliteit, nationale of etnische afkomst, leeftijd, handicap, geslacht, genderidentiteit en seksuele geaardheid.*

Wanneer deze definitie wordt gevolgd, dan is online discriminatie een onderdeel van online haatspraak. Haatspraak bevat immers ook aanzetten tot geweld of geweld bevorderen, verspreiden of rechtvaardigen.

Bij online haatspraak worden soms niet-discriminerende woorden, cijfers, *memes* of *emoticons* gebruikt die een dubbele betekenis hebben en zodoende niet snel door moderatie worden verwijderd. Dit wordt ook wel *dog whistle*

genoemd; een hondenfluitje dat niet hoorbaar is voor de meeste mensen. Een *dog whistle* kan een discriminerende of haatvolle boodschap hebben. Voorbeelden hiervan zijn [Pepe the Frog](#), het woord '[jogger](#)' bedacht door extreemrechtse groepen op 4chan ter vervanging van het n-woord of '88' wat staat voor 'Heil Hitler' (twee keer de achtste letter van het alfabet, wat de H is).

In dit rapport focussen we op online discriminatie op grond van afkomst, huidskleur en religie, die online wordt geuit, bijvoorbeeld via sociale media. Soms wordt ook haatspraak of de overkoepelende term 'online haat' gebruikt wanneer dit door respondenten of in onderzoeken zo wordt beschreven.

In de [Begrippenlijst](#) wordt dieper ingegaan op de relevante definities.

## Hoe vaak komt online discriminatie voor?

Het meldpunt voor internetdiscriminatie [MiND](#) registreert meldingen over online uitingen – bijvoorbeeld op websites, sociale media en blogs – die de melder als discriminerend ervaart. Het aantal meldingen dat wordt geregistreerd, is laag. In 2021 ontving het meldpunt 339 meldingen, dat is minder dan de helft van het totaal aantal meldingen in het voorgaande jaar en het laagste aantal ooit. Een mogelijke verklaring die hiervoor wordt gegeven, is dat er, anders dan in voorgaande jaren, weinig incidenten voorkwamen die gepaard gingen met een grote publieke verontwaardiging. Wel is het zo dat er meer unieke meldingen waren in 2021 (88 procent) in vergelijking met 2020 (47 procent). In 2020 waren er meer bulkmeldingen, dus meer meldingen over één incident. In lijn met de vorige jaren gaan deze meldingen van online discriminatie voornamelijk over herkomst (35 procent). Ook kwamen er verschillende meldingen op basis van godsdienst binnen (6 procent) (Walz & Fiere, 2022).

Uit de jaarlijkse monitoring van het Centrum Informatie en Documentatie Israël ([CIDI](#)) naar antisemitisme in Nederland blijkt dat er in 2021 47 incidenten waren van antisemitische digitale privéberichten, een ruime verdubbeling van het jaar ervoor. De organisatie geeft aan dat in de monitor geen rekening is gehouden met andere antisemitische uitingen via het internet (mede) vanwege het feit dat online haatspraak te vaak voorkomt (CIDI, 2022). Het CIDI heeft eerder wel een onderzoek laten uitvoeren naar online antisemitisme. Dit onderzoek heeft 1.033 antisemitische uitingen geconstateerd (Kaal & Bouwman, 2020). Uit onderzoeken naar online haat en discriminatie, blijkt dat deze cijfers geen realistisch beeld geven over de daadwerkelijke omvang van de problematiek. Zo blijkt uit data-analyses dat er op Twitter een [duidelijke toename](#) is van Nederlands racistisch taalgebruik. Ook ontving Sylvana Simons nadat zij haar kandidaatschap bij DENK bekendmaakte circa 40.000 online haatberichten (ECRI, 2019).

## Gevolgen van online discriminatie

Discriminatie kan de gezondheid ernstig aantasten. Ook als je zelf niet gediscrimineerd wordt, maar meemaakt dat mensen die op jou lijken (vaak gediscrimineerd worden, kan dit negatieve effecten hebben. Hetzelfde geldt voor discriminerende en haatvolle reacties online. Deze kunnen ernstige gevolgen hebben zowel op de korte als de lange termijn. Denk aan negatieve invloed op de mentale gezondheid zoals stress, maar ook aan sociale uitsluiting, pesten en slechtere (werk- of school)prestaties (Ikram, 2016). Haatberichten op sociale media geven bovendien voeding voor groeiende tegenstelling tussen groepen in de samenleving, en leveren daarmee brandstof voor polarisatie in de samenleving (Verloove, Van Wonderen & Felten, 2020).

## Onderzoeksvragen

Met dit onderzoek willen we de huidige aanpak van online discriminatie inzichtelijk maken. Daarnaast zetten we uiteen welke Nederlandse organisaties een rol kunnen spelen in de preventieve en repressieve aanpak van online discriminatie. Op deze manier willen we een overzicht geven van de huidige stand van zaken omtrent de aanpak van online discriminatie. Voor dit onderzoek hebben we de volgende onderzoeksvragen geformuleerd:

- Hoe ziet de aanpak en regelgeving voor strafbare vormen van online discriminatie eruit?
- Welke (Europese en Nederlandse) wetgeving is van toepassing op online discriminatie, en welke ontwikkelingen en obstakels zijn hierin te vinden voor de Nederlandse aanpak?
- Welke maatregelen nemen socialemediabedrijven, wat zijn de ontwikkelingen hierin en wat is het effect van deze aanpak?
- Welke aanpakken zijn er ontwikkeld vanuit het onderwijs, maatschappelijke organisaties en gebruikers?
- Welke organisaties/partijen zijn aan zet bij de aanpak van online discriminatie?

## Onderzoeksmethode

Om antwoord te geven op de onderzoeksvragen, hebben we een literatuurverkenning en deskresearch uitgevoerd. Hierbij hebben we gezocht in de wetenschappelijke literatuur, voornamelijk in *Google Scholar*, en in de meer 'grijze literatuur', denk hierbij aan rapporten. Ook hebben we inzichten verworven uit andere betrouwbare bronnen, zoals verslagen van de Europese Commissie, monitoringsdata en het (online) Wetboek van Strafrecht. Daarnaast hebben we interviews afgenomen met experts op dit gebied. In totaal hebben

we met 12 Nederlandse en Europese experts gesproken in 9 interviews. Deze experts zijn werkzaam bij de overheid (nationaal en Europees), een socialemediabedrijf, een Europese mensenrechtenorganisatie, politie, Openbaar Ministerie, een AI-bedrijf (kunstmatige intelligentie), een Vlaamse antidiscrimatievoorziening (ADV) en MiND. De experts zijn door de onderzoeksgroep geselecteerd op basis van hun expertise en werkervaring op het onderwerp online discriminatie en haatspraak. Ook is hierover overleg gevoerd met het ministerie van Sociale Zaken en Werkgelegenheid, dat KIS financiert. Daarnaast hebben we tijdens de interviews navraag gedaan of we een bepaald perspectief (en daarmee een bepaalde expert) misten. De interviews zijn afgenomen aan de hand van een semigestructureerd topiclijst die is ontwikkeld op basis van de literatuur. De interviews vonden zowel online (Zoom of Microsoft Teams) en fysiek plaats, werden volledig opgenomen en getranscribeerd. De data-analyse is uitgevoerd met behulp van het ondersteuningsprogramma MAXQDA. Alle experts hadden de mogelijkheid om nog feedback te geven op het rapport.

## Leeswijzer

Dit rapport bestaat uit twee delen. In het eerste deel beschrijven we de huidige aanpak van online discriminatie en de ontwikkelingen daarin. We gaan in op Europese wetgeving en gedragscodes en Nederlandse wetgeving. Vervolgens staan we stil bij de aanpak vanuit onderwijs, maatschappelijke organisaties en gebruikers evenals de aanpak vanuit socialemediabedrijven. In het tweede deel geven we een overzicht van organisaties die in Nederland een rol kunnen spelen in de aanpak van online discriminatie. In de bijlage is een begrippenlijst te vinden van begrippen die wij hanteren in dit rapport. Het rapport wordt afgesloten met de literatuurlijst.

# De huidige aanpak van online discriminatie

Hoe ziet de huidige aanpak van online discriminatie eruit? In dit deel inventariseren we de bestaande aanpak van verschillende organisaties. We gaan eerst in op de wet- en regelgeving, zowel op Europees niveau als op nationaal niveau. Vervolgens staan we stil bij de aanpak vanuit onderwijs, maatschappelijke organisaties en gebruikers. Het hoofdstuk sluiten we af met interventies die socialemediabedrijven inzetten.

## Europese wetgeving en gedragscodes

In recente jaren zijn er op Europees niveau afspraken gemaakt met socialemediabedrijven, waaronder Facebook, Twitter, YouTube, Instagram, TikTok en Snapchat, over hun deelname aan een gedragscode in het kader van het tegengaan van online haatzaaien. Als onderdeel van dit initiatief (de Code of Conduct) wordt haatzaaiende content dat gemeld is bij de platforms door zogenaamde ‘trusted flaggers’ binnen 24 uur geëvalueerd en indien gegrond gevonden, verwijderd door het desbetreffende socialemediabedrijf.<sup>2</sup>

De gedragscode stelt verschillende verzoeken aan sociale mediabedrijven:

- Het hebben van regels en normen die het aanzetten tot haatspraak verbieden en systemen en teams hebben om haatzaaiende content te evalueren en te melden.

[2\\_assessment\\_of\\_the\\_code\\_of\\_conduct\\_on\\_hate\\_speech\\_on\\_line\\_-\\_state\\_of\\_play\\_0.pdf\(europa.eu\)](#)

- Content die als haatzaaiend wordt aangegeven binnen 24 uur te evalueren en indien gegrond te verwijderen.
- Regelmatig trainingen te geven aan hun personeel met betrekking tot online haatzaaien.
- Het aangaan van samenwerkingen en opleidingsactiviteiten met het maatschappelijke middenveld om betrouwbare netwerken van verslaggevers te vergroten.
- Contactpunten vanuit de nationale autoriteiten aanwijzen om informatie ter kennisgeving te ontvangen.
- Bevorderen van transparantie zowel naar gebruikers als naar het grote publiek toe.

De Europese gedragscode wordt regelmatig geëvalueerd door een netwerk van organisaties. In 2022 heeft de [zevende monitoring](#) plaatsgevonden. Het aantal meldingen gemaakt door trusted flaggers van illegale haatspraak dat binnen 24 uur door socialemediabedrijven wordt beoordeeld, ligt op 64,4 procent en is gedaald ten opzichte van 2021 toen dit nog 81 procent was. Ook ten opzichte van 2020 ligt dit percentage lager, waar het toen nog ging om 90.4 procent. Het percentage berichten dat na de beoordeling wordt verwijderd in 2022 is ongeveer gelijk met 2021, 63,6 procent tegen 62,5 procent. Dit ligt wel lager dan in 2020 toen het percentage op 71 procent lag. De grootste reden waarom

socialemediabedrijven geen gehoor gaven aan het verwijderen van de gerapporteerde content was dat zij zelf deze content niet als illegaal zagen volgens hun eigen gedragsregels of de lokale wetgeving (Reynders, 2022). Dit geeft de complexiteit aan van illegale haatspraak.

Europese landen hebben elk een eigen wetgeving wat betreft discriminatie en haatspraak, deze zijn soms expliciet toegespitst naar online discriminatie. De verschillen tussen landen zijn vooral zichtbaar in hoe specifiek er wordt uitgesplitst welke groepen worden beschermd door de wet. In het Nederlandse Wetboek van Strafrecht wordt discriminatie op grond van ras, godsdienst, levensovertuiging, geslacht, hetero- of homoseksuele gerichtheid of lichamelijke, psychische of verstandelijke handicap genoemd. In België wordt dit verder uitgesplitst naar ras, godsdienst, levensovertuiging, geslacht, hetero- of homoseksuele gerichtheid, lichamelijke, psychische of verstandelijke handicap, kleur, nationaliteit, etnische afkomst, leeftijd, burgerlijke staat, economische staat, levensovertuiging, politieke voorkeur, taal, gezondheid, handicap, uiterlijke kenmerken en geslacht. Ook als het gaat om straffen zijn er verschillen tussen landen in Europa. Zo is de straf voor het aanzetten tot haat of discriminatie in Nederland maximaal twee jaar gevangenisstraf of een geldboete van de vierde categorie (zie [artikel 137d](#)). Ter vergelijking: in Duitsland ligt de straf voor het aanzetten tot haat, het schenden van de menselijke waardigheid of het ontkennen van de holocaust op een gevangenisstraf tussen de drie maanden en vijf jaar. Daarbij stelt The German Network Enforcement Act dat socialemediabedrijven in Duitsland ‘overduidelijke illegale’ content binnen 24 uur moeten verwijderen. In meer ambigue situaties krijgen ze hier zeven dagen voor. Mochten de platforms dit niet doen dan riskeren ze de kans op een boete die kan oplopen tot 5 miljoen euro (Rudnicki & Steiger, 2020).

## Ontwikkelingen Europese wetgeving

### Digital Services Act (DSA)

EU-burgers online beter beschermen. Dat is een van de belangrijkste doelen van de nieuwe Europese wetgeving Digital Services Act (DSA). De DSA regelt de verantwoordelijkheid en aansprakelijkheid van internetaanbieders, hostingbedrijven, online platforms, zoekmachines en marktplaatsen voor de activiteiten via hun diensten. Het idee achter de nieuwe wetgeving: ‘Alles wat offline is verboden, moet ook online worden verboden’, aldus de Deense [Christel Schaldemose](#), die verantwoordelijk is voor het wetsontwerp. Als het gaat om socialemediabedrijven, zijn zij ook in de nieuwe wetgeving niet verantwoordelijk voor de content die op hun platforms wordt gepubliceerd. De bedrijven zijn alleen verantwoordelijk voor de content waar ze vanaf weten, bijvoorbeeld doordat er een melding is gemaakt. De wetgeving heeft wel invloed op hoe de platforms met illegale content, zoals online discriminatie en haat, omgaan. De DSA biedt bijvoorbeeld extra verplichtingen aan online platforms die in de DSA als ‘zeer groot’ worden bestempeld.

De nieuwe wetgeving is 16 november 2022 in werking getreden. Bedrijven hebben tot februari 2024 de tijd om aan de nieuwe regels te voldoen. Op 17 november 2024 wordt de DSA volledig van toepassing en moeten EU-lidstaten bijvoorbeeld ook ‘digitaledienstencoördinatoren’ hebben. De DSA vervangt nationale wetgeving, ten behoeve van horizontale regelgeving voor de hele EU. Wel blijft de nationale wetgeving leidend als het gaat om wat wordt bestempeld als illegale content en de strafbaarheid ervan.

De volgende verplichtingen die de DSA stelt, zijn relevant als het gaat om de aanpak van online discriminatie:



1. Het verwijderen van informatie of gebruikersaccounts moeten digitale aanbieders gaan onderbouwen en uitleggen. Gebruikers kunnen verwijdering aanvechten.
2. Een verbod op ingebouwde ‘paden’ die bezoekers naar illegale content of desinformatie leiden.
3. Voor bijzondere omstandigheden, zoals de oorlog in Oekraïne of de coronacrisis, heeft de Europese Commissie een noodprocedure opgesteld waarbij aanvullende maatregelen van techbedrijven geëist kunnen worden. Zij worden dan verplicht om tijdelijk hun algoritmes aan te passen om de verspreiding van desinformatie (of specifieke oorlogspropaganda) tegen te houden.
4. De grootste online platforms – de sites met meer dan 45 miljoen actieve Europese gebruikers per maand, dit zijn vooral de bekende platforms als Facebook, Instagram, YouTube en TikTok – worden het strengst gereguleerd. Hiervoor gelden extra verplichtingen zoals verplichte onafhankelijke audits. Onder de extra verplichtingen valt ook een risicobeoordelingsexercitie waarbij het platform moet kijken naar in hoeverre gebruikers worden blootgesteld aan illegale goederen of content. Ook moeten zij openbaar maken hoeveel personeel zij inzetten om schadelijke content op te sporen en te verwijderen.
5. Maatschappelijke (zelf)organisaties (de ngo’s) kunnen een ‘trusted flagger’ worden. Dit houdt in dat de meldingen van illegale content met prioriteit door socialemediabedrijven worden behandeld. De organisaties kunnen die status aanvragen bij de nieuwe nationale ‘digitaaldienstencoördinatoren’.

De Europese Commissie zal toezicht houden op de naleving van de DSA door de grootste online platforms en zoekmachines. Deze bedrijven betalen zelf via

een heffing voor dit toezicht. Als zij niet aan de nieuwe verplichtingen voldoen, kunnen ze bovendien enorme boetes krijgen (European Commission, z.d.).

Meer weten over de DSA? Ga naar de [website van de Europese Unie](#).

### Aanbeveling van de Raad van Europa

In mei 2022 heeft het Comité van Ministers bestaande uit ministers van Buitenlandse Zaken van alle EU-lidstaten een aanbeveling van de Raad van Europa – genaamd ‘Recommendation CM/REC (2022 16)’ – van de Raad van Europa aangenomen. Ter verduidelijking: dit is geen nieuwe wetgeving, maar een aanbeveling die nationale regeringen ook naast zich neer kunnen leggen. In de aanbeveling worden Europese regeringen opgeroepen om alomvattende strategieën te ontwikkelen om haatzaaiende uitlatingen te voorkomen en te bestrijden, inclusief het toepassen van een effectief wettelijk kader en de uitvoering van adequate en evenredige maatregelen. De nationale autoriteiten moeten daarbij een zorgvuldig evenwicht vinden tussen het recht op privéleven, het recht op vrijheid van meningsuiting en het verbod op discriminatie. In de richtlijnen wordt aanbevolen dat de lidstaten onderscheid maken tussen ten eerste ‘ernstigste gevallen’ van haatzaaiende uitlatingen, die strafrechtelijk moeten worden verboden, ten tweede haatzaaiende uitlatingen die onder het burgerlijk en bestuursrecht vallen, en ten slotte beledigende of schadelijke uitingen die niet ernstig genoeg zijn om legitiem te worden beperkt op grond van het Europees Verdrag voor de Rechten van de Mens, maar die desalniettemin om alternatieve interventies vragen. Ook stelt de aanbeveling dat Europese lidstaten meer moeten doen tegen de niet-illegale vormen van haatzaaien. Zo moet meer ingezet worden op bewustwording, onderwijs, het bieden van counterspeech, moet er duidelijke ondersteuningsstructuur komen

voor gedupeerden en moeten politie, justitie en andere professionals getraind worden op dit onderwerp (Council of Europe, 2022).

Meer informatie over deze aanbeveling? Ga naar [de website van de Raad van Europa](#).

### Haatzaaiende uitingen strafbaar

Vanuit de Europese Commissie is gedetecteerd dat haatzaaiende uitingen en haatmisdrijven sterk zijn toegenomen, zowel offline als online. Nu bestaat er nog geen rechtsgrondslag om dit op Europees niveau strafbaar te stellen. Op 9 december 2021 heeft de Europese Commissie een voorstel gedaan bij de Europese Raad om de huidige lijst van 'EU-misdrijven' in artikel 83, lid 1, VWEU uit te breiden met haatmisdrijven en haatzaaiende uitlatingen. Als het voorstel wordt aangenomen, zou de Europese Commissie in een tweede stap secundaire wetgeving kunnen voorstellen die de EU in staat stelt om naast racistische of xenofobe motieven ook andere vormen van haatzaaiende uitlatingen en haatmisdrijven strafbaar te stellen (European Commission, z.d.).



## Nederlandse wetgeving

Iemand die zich in Nederland online of offline gediscrimineerd voelt, kan zich melden bij een antidiscriminatievoorziening (ADV) en/of melding of aangifte doen bij de politie. Discriminatie melden kan online via het [Meldformulier Discriminatie](#) op de website van de politie. Aangifte van discriminatie moet altijd op een politiebureau gebeuren. Voor strafbare, discriminerende uitingen op het internet op basis van etniciteit, geslacht, godsdienst of levensbeschouwing, handicap of seksuele voorkeur is ook een landelijk meldpunt: MiND (Meldpunt Internet Discriminatie). Wanneer iemand advies wil om te kunnen beoordelen of het gaat om een strafbaar feit of doorverwezen wil worden naar de juiste instantie, dan kan een lokale ADV of het College voor de Rechten van de Mens (Het College) om advies worden gevraagd (Rijksoverheid, z.d.).

Daarnaast zijn er specifieke organisaties waar iemand een klacht kan indienen als er discriminatie is ervaren. Zo beoordeelt het College voor de Rechten van de Mens (Het College) discriminatieklachten. Het College en de ADV's beoordelen deze meldingen op basis van Algemene wet gelijke behandeling (AWGB). Hierbij geldt het voor discriminatie op de terreinen arbeid, goederen, vrije beroep, lidmaat vakbond of vereniging van beroepsgenoten en sociale bescherming. Online discriminatie wordt niet genoemd in deze wet (College voor de Rechten van de Mens, z.d.).

In artikel 1 van de Grondwet is vastgelegd dat discriminatie in Nederland verboden is. Discriminatie op basis van godsdienst, levensovertuiging, politieke gezindheid, ras, geslacht of op welke grond dan ook, is niet toegestaan. In Nederland is dit vastgelegd in verschillende wetten. [Artikel 266](#) van het Wetboek van Strafrecht stelt dat er een verbod is op eenvoudige belediging. Voorbeelden van eenvoudige belediging zijn het uitschelden van iemand, het opsteken van

een middelvinger of het spugen in het gezicht. Specifiek voor strafbare feiten die betrekking hebben op discriminatie, zijn de regels vastgelegd in [artikel 137c](#) tot en met 137e. Het opsporen en vervolgen van strafbare online discriminatie en haat gebeurt nog weinig. Bij de politie zijn geen agenten die zich specifiek bezighouden met online discriminatie. Bovendien wordt er weinig aangifte gedaan van online discriminatie. Uit onderzoek van RTL Nieuws (2021) komt naar voren dat daders van online haat hierdoor vaak onbestraft blijven.

De wetgeving in Nederland met betrekking tot strafrechtelijke discriminatie is vastgelegd in [artikel 137c](#) tot en met 137g en artikel 429quater van het Wetboek van Strafrecht. Daarbij zijn er bij een aangifte van online discriminatie de volgende [regels](#) vastgelegd: het OM draagt de politie op onder meer het volgende vast te leggen: de accountnaam waar de uiting mee is geplaatst, de openbaarheid van de uiting, de periode waarin de uiting online heeft gestaan, de plaatsingsdatum van de uiting, de datum waarop de uiting is vastgelegd ten behoeve van het proces-verbaal, en de samenhang en context van de uiting.

Alle aangiften die binnenkomen, worden opgepakt en beoordeeld. Dit wordt gedaan op basis van de genoemde artikelen en de jurisprudentie van Nederlandse rechtspraak en van het Europees Hof voor de Rechten van de Mens. Na iedere aangifte volgt een opsporingsonderzoek, dan een proces-verbaal en vervolgens wordt gekeken of het feit strafbaar en vervolgbaar is. Dit is bij online discriminatie complex. Zo kan het IP-adres en waar de persoon woont worden achterhaald, maar daadwerkelijk bewijzen dat die persoon het ook heeft geplaatst, is lastig. Zo kan het zijn dat er meerdere mensen op dat adres wonen en bij mobiele apparaten kan ook iemand anders toegang hebben gehad tot het apparaat. Daarnaast is het ook lastig te bewijzen bij een naam of niet iemand anders het profiel misbruikt heeft, zo licht het OM toe in een interview:

*‘(...) iemand die zegt ‘ja, ik leen mijn telefoon wel eens uit’ dan is het aan ons om te bewijzen dat die persoon het heeft gestuurd. Het is niet zo dat je per se verantwoordelijk bent voor jouw profiel. Je moet echt bewijzen wie dat bericht heeft getypt. Als iemand met het verweer komt ‘ja, maar er was iemand in mijn woning en die heeft ingelogd op mijn Facebook en heeft dat bericht geplaatst’ dan moeten wij het tegendeel kunnen aantonen.’*

Het komt ook voor dat er onder een bericht van een (prominent) persoon meer dan tweehonderd strafbare berichten staan. Hierbij is het praktisch niet mogelijk om al deze mensen op te sporen. Een manier om hiermee om te gaan is om de ergste berichten door het slachtoffer zelf te laten selecteren, geeft een respondent van het OM aan. Hierbij is transparantie volgens het OM het belangrijkste, dus hoe het OM en/of het slachtoffer tot een selectie van deze berichten is gekomen:

*‘We hebben ook wel eens criteria opgesteld en dat we dan zelf door de reacties zijn gegaan. Bijvoorbeeld mensen die vaker reageren of die bepaalde uitlatingen doen. We proberen zo te zorgen dat het enigszins objectief is en behapbaar. Maar dat blijft een uitdaging.’*

Ook wat betreft straffen zijn er uiteraard wettelijke regels, hierin wordt geen onderscheid gemaakt tussen online en offline discriminatie. De straffen op basis van artikel 137c tot en met 137g en artikel 429quater bestaan vaak uit een geldboete of een taakstraf voor een first offender. Bij recidive valt de straf hoger uit. Bij zaken waarbij er meerdere wetten naast deze [artikelen](#) worden overschreden, kan deze straf anders zijn.

## Anti-discriminatie wetgeving versus vrijheid van meningsuiting

In Nederland staat het recht op vrijheid van meningsuiting in artikel 7 van de Grondwet. Daarin staat dat niemand voorafgaand toestemming nodig heeft van de overheid om iets te zeggen of te schrijven in het openbaar. Daarnaast is de vrijheid van meningsuiting ook in verschillende internationale verdragen opgenomen, zoals in artikel 10 van het Europees Verdrag voor de Rechten van de Mens. Vrijheid van meningsuiting is een belangrijk fundament van de democratie. Mensen moeten kunnen zeggen en schrijven wat ze vinden en vrij informatie kunnen delen en ontvangen – ook online. Tegelijkertijd mag niet alles gezegd en geschreven worden. Op de [website](#) van het College voor de Rechten van de Mens staat dan ook: ‘Vrijheid van meningsuiting houdt op daar waar in het strafrecht een grens is getrokken.’ Zo is bijvoorbeeld discriminatie – mensen anders behandelen, achterstellen of uitsluiten op basis van hun persoonlijke kenmerken, zoals hun geslacht, het hebben van een beperking of hun geloof – verboden. Ook online. Vaak zijn mensen het niet eens over waar die grens ligt en uiteindelijk is het aan de rechter om te bepalen of iets onder de vrijheid van meningsuiting valt.

De Hoge Raad heeft een toetsingskader (het ‘drie stappenplan’) ontwikkeld om te kunnen beoordelen of een beperking van het recht van vrijheid van meningsuiting is toegestaan. Dit toetsingskader kent drie stappen. Te weten: is de uitlating beledigend, neemt de context waarin de uitlating is geplaatst het beledigende karakter uit de uitlating weg en indien dit het geval is, is de uitlating onnodig grievend. Er zijn verschillende rechtelijke uitspraken gedaan als het gaat om aanzetten tot haat of discriminatie. Via de zoekmachine [JURE](#) komen er 219 uitspraken naar voren. Dit is zowel online als offline.

## Online discriminatie melden

MiND (Meldpunt Internet Discriminatie) is het landelijke meldpunt voor strafbare, discriminerende uitingen op basis van ras, geslacht, godsdienst of levensbeschouwing, handicap of seksuele voorkeur op het internet. Denk hierbij aan sociale media, autonome websites en forums. Het meldpunt registreert en beoordeelt meldingen die zij ontvangen op basis van het strafrecht. Melden mag anoniem. MiND kan er als zogenoemde ‘trusted flagger’ ervoor zorgen dat een socialemediabedrijf met voorrang de melding onderzoekt. Het bedrijf beslist dan zelf of ze hier wel of niet aan voldoen. Alleen de rechter of de officier van justitie kan vervolgens een verwijderbevel geven als hier aangifte van wordt gedaan.

MiND richt zich uitsluitend op groepsdiscriminatie. Individuele online discriminatie behandelen ze niet. Hierbij gaat het om discriminatie specifiek gericht op één enkel persoon. Na het ontvangen van een melding start MiND een onderzoek. Er wordt gecontroleerd of de uiting nog online staat, vervolgens wordt er aan de hand van het Wetboek van Strafrecht (artikel 137c tot en met 137e) beoordeeld of er sprake is van een strafbaar feit.

Een kwart van de meldingen die MiND ontvangt, leidt tot een verwijderingsverzoek bij de website, die in vrijwel alle gevallen wordt overgenomen. Bij de rest wordt aan de melder teruggekoppeld, zo vertelt een medewerker van MiND, dat ze de melding niet gaan behandelen, mits er een e-mailadres is achtergelaten. Als aan een (herhaald) verzoek door het socialemediaplatform geen gehoor wordt gegeven, zal MiND een aangiftedossier opstellen voor het OM. Het OM kan een strafrechtelijk onderzoek starten naar aanleiding van een aangifte (MiND, z.d.). De rechter of de officier van justitie kan vervolgens een verwijderbevel geven.

## Aanpak vanuit onderwijs, maatschappelijke organisaties en gebruikers

Niet alleen vanuit de Europese en nationale wetgeving wordt de strijd aangegaan met online discriminatie en haat. Ook het onderwijs, maatschappelijke (zelf)organisaties en (groepen) gebruikers zetten zich in voor een discriminatievrij sociale media, dit doen zij op verschillende wijzen. We onderscheiden hierbij educatieve interventies met als doel te informeren, bewustzijn en/of weerbaarheid te creëren, interventies om te monitoren, interventies van (groepen) gebruikers, campagnes en hulp aan (potentiële) gedupeerden.

### Kabinetsaanpak online discriminatie

Naar aanleiding van de [motie Van Dijk](#), die de regering oproept om te inventariseren hoe de aanpak van racisme en antisemitisme geïntensiveerd kan worden, voert het ministerie van Justitie en Veiligheid in samenwerking met de ministeries van Binnenlandse Zaken en Koninkrijksrelaties, Onderwijs, Cultuur en Wetenschap en Sociale Zaken en Werkgelegenheid een inventarisatie uit. Hiermee wordt een eerste aanzet gedaan voor een plan van aanpak online discriminatie, dat de bestrijding van online discriminatie moet versterken. Het kabinet heeft toegezegd de Tweede Kamer voor het einde van het 2022 over de verdere kabinetsambities te informeren.

### Educatieve interventies

Educatieve interventies worden ingezet om mensen, voornamelijk kinderen en jongeren, vaardigheden aan te leren om kritisch te kijken naar online content en om verontrustende, haatvolle en misleidende content te identificeren. Ook kunnen handelingsperspectieven worden aangeboden tijdens deze interven-

ties over hoe te reageren op discriminerende content (Blaya, 2019). De meeste educatieve interventies die aandacht besteden aan online discriminatie vallen onder bredere interventies die zich richten op (digitale) mediawijsheid of digitale geletterdheid. Deze interventies zijn ontwikkeld om de mogelijke schade van (sociale) media te verminderen. In de praktijk zijn deze interventies verschillend, maar in de kern richten ze zich op het overbrengen van vaardigheden en kennis die kunnen bijdragen om kritisch te kijken naar en gebruik te maken van (sociale) media. Het doel van deze interventies is de impact van (sociale) media op iemands opvattingen, attitudes, normen en gedrag te beperken en te verminderen (Jeong, Cho & Hwang, 2012).

Sinds 2007 zet de Rijksoverheid meer en meer in op het terrein van mediawijsheid. In het regeerakkoord van 2007 werd opgenomen dat er media-educatie komt voor kinderen, jongeren, ouders en scholen over hoe om te gaan met de veelheid van media-uitingen. Een jaar later werd via een kabinetsvisie ingezet op het zelfstandig kunnen omgaan met de gevaren van media door individuele burgers (Nikken & Jongmans, 2013). In ditzelfde jaar werd ook het [Netwerk Mediawijsheid](#) opgericht. Dit centrum zet zich in voor het mediawijs maken van iedereen in Nederland. Sindsdien zijn er verschillende organisaties ontstaan die zich inzetten voor deze missie. Bij het netwerk van het Netwerk Mediawijsheid zijn nu meer dan duizend organisaties aangesloten. De lobby groeit, maar mediawijsheid in het onderwijs is momenteel niet verplicht. En online discriminatie en haat is er niet een vast onderdeel van. Er zijn geen recente cijfers bekend over hoeveel scholen binnen het curriculum aandacht besteden aan mediawijsheid. De meest recente cijfers dateren uit 2013 en zijn daarmee, zeker voor dit onderwerp, sterk verouderd. Het is ook niet bekend hoeveel aandacht er bij mediawijsheid wordt gegeven aan online discriminatie of haat. Wel is er met de Week van de Mediawijsheid (zie kader volgende pagina) bij o.a. scholen

aandacht gevraagd voor het stimuleren van pro sociaal gedrag online. Ook speelde een recordaantal van ruim 180.000 leerlingen uit groep 7 en 8 in 2022 mee aan de serious game MediaMasters over de kansen en gevaren van (digitale) media (Week van de Mediawijsheid, 2022).

Uit een meta-analyse blijkt dat mediawijsheidsinterventies over het algemeen een positief effect hebben op het verminderen van riskant en zogenoemd antisociaal gedrag online, waar we online discriminatie onder kunnen scharen. Bovendien zagen onderzoekers dat bij de deelnemers aan dergelijke interventies de zelfeffectiviteit werd vergroot om zulk gedrag te voorkomen. Er werd echter geen verandering in de norm gevonden (Jeong et al., 2012). Uit een evaluatie van het lesprogramma Under Pressure van Diversion, een lesprogramma om de weerbaarheid van jongeren te vergroten tegen desinformatie en het vergroten van democratisch burgerschap, blijkt dat het zinvol is om jongeren in het voorgezet onderwijs en mbo met deze extra lessen weerbaar te maken tegen desinformatie en eenzijdige meningsvorming (Van Wonderen & Peeters-Osseyran, 2022). Dit is mogelijk ook het geval voor online discriminatie. Een eerste stap richting de preventie van online discriminatie is door het leren van jonge gebruikers hoe ze discriminerende content moeten beoordelen en wat ze hiermee kunnen doen (Blaya, 2019). Hierin kan het onderwijs en andere organisaties in het maatschappelijke domein die jongeren bereiken een rol spelen door het inzetten van deze educatieve interventies.

Uit ons onderzoek komt naar voren dat er in Nederland nauwelijks interventies zijn die zich richten op bewustwording, omgaan met en weerbaarheid tegen specifiek online discriminatie. Wel zijn er allerlei interventies – zoals lespakketten, theatervoorstellingen en films – gericht op cyberpesten. Ook komt in bestaande anti-discriminatie interventies ‘online’ vaak niet of uitsluitend als een

onderdeel aan bod. Ook zijn er in Nederland weinig weerbaarheidstrainingen over het omgaan met (offline en/of online) discriminatie, laat staan onderzoek naar de effectiviteit ervan. Een training die wel is geëvalueerd, is de training [Reageren op discrimineren](#), waarbij de ervaringen van jongeren met uitsluiting en discriminatie centraal staan. Uit de evaluatie blijkt dat de training bijdraagt aan een versterkt bewustzijn van en inzicht in manieren van omgaan met discriminatie. Een ander voorbeeld is de weerbaarheidstraining [This is me!](#). Er is ook veel en sterk wetenschappelijk bewijs voor het inzetten van peer support, het delen van ervaringen en verhalen (Omlo, 2017). Dit is onder andere toegepast bij [#DatMeenJeNiet](#), een initiatief van Movisie, Hogeschool Inholland en Diversity Media, waarbij jongeren samen in actie komen tegen online discriminatie.

### De Week van de Mediawijsheid in het onderwijs

De Week van de Mediawijsheid is een jaarlijks terugkerende campagne die mediawijsheid wil stimuleren. De week is een initiatief van het Netwerk Mediawijsheid en staat ieder jaar in het teken van een ander thema. In november 2022 had de Week het thema ‘Like en Cancel’ en ging over het stimuleren van pro sociaal gedrag online. Vanuit deze Week worden overal in Nederland diverse activiteiten georganiseerd. Meer informatie: [www.weekvandemediawijsheid.nl](http://www.weekvandemediawijsheid.nl)

### Interventies om te detecteren

Bij de aanpak van online discriminatie, is het signaleren ervan cruciaal. Behalve dat gebruikers dit (leren) signaleren, leveren bedrijven een bijdrage aan het opsporen van (potentieel) illegale discriminerende of haatvolle content. Zo is

het mogelijk om problematische berichten of beelden automatisch te detecteren. Een techniek waar platforms zelf aan bijdragen, maar er zijn ook bedrijven die zich hierin specialiseren (Van Huijstee et al., 2021). Voorbeeldprojecten hiervan zijn de Online Observatory of Hate Speech en Detect Then Act (zie kader).

### **Detect Then Act**

Het project [Detect Then Act](#) monitorde met nieuwe taaltechnologie online haat en voerde actie om omstanders aan te sporen om zich uit te spreken. In het project werden honderden socialemediagebruikers opgeleid tot weerbare medestanders die constructief konden reageren op de opgespoorde online haat via een algemeen account. Meer dan 400 socialemediagebruikers binnen Europa hadden zich hierbij aangesloten. Bij dit project waren meerdere partijen betrokken zoals burgers, ngo's, universiteiten en techbedrijven uit verschillende landen, waaronder Nederland. Het werd gesubsidieerd door de Europese Unie.

### **European Observatory of Online Hate**

Dit Europese project van Textgain in samenwerking met Dare to be Grey, Hogeschool Utrecht en PDCS voert een tweejarig onderzoek uit naar en rapporteert over de fundamentele aard van de dynamiek van online haat, hoe haat zich manifesteert, de verbanden tussen de daders en hun invloed en desinformatiestrategieën. Binnen dit project wordt een monitoringtool ontwikkeld met behulp van geavanceerde AI-tools. Meer informatie: [EOOH.eu](http://EOOH.eu)

### **Interventies van groepen gebruikers**

#DatMeenJeNiet is een voorbeeld waarbij groepen gebruikers samen in actie komen tegen online discriminatie op sociale media. In het Engels wordt dit ook wel technological placekeeping genoemd: het actief 'onderhouden' van digitale ruimtes (Van Huijstee et al., 2021). Een ander voorbeeld hiervan is het Europese project [Detect Then Act](#) (zie kader) waarbij met gebruik van taaltechnologie haat wordt opgespoord en groepen socialemediagebruikers zijn opgeleid om constructief te reageren. Ook kunnen gebruikers discriminerende berichten melden bij de desbetreffende website, bij MiND, ADV's en bij de politie. Daarnaast zijn er specifieke meldpunten zoals Meld Islamofobie of het [CIDI](#) voor antisemitisme waar mensen ook terecht kunnen voor online incidenten.

### **Campagnes**

De afgelopen jaren zijn campagnes gevoerd door diverse organisaties en overheden, soms samen met (groepen) burgers, om bewustwording te stimuleren. Vaak is het onderwerp van deze campagnes breder dan uitsluitend online discriminatie of wordt zowel offline als online discriminatie aangekaart. Een bekend nationaal voorbeeld van een dergelijke campagne is [#doeslief](#) van SIRE, die mensen bewust wilde maken van asociaal gedrag offline én online. Een ander voorbeeld is de campagne [Kom op Amsterdam!](#) waarbij de gemeente Amsterdam Amsterdammers oproept om zich uit te spreken tegen offline en online discriminatie. Ook zijn er vanuit bepaalde sectoren campagnes ontwikkeld tegen discriminatie, zoals [#OneLove](#) van voetbalbond KNVB. Een goed buitenlands voorbeeld is de Duitse campagne [#UniteAgainstHate](#) waarbij voetballers zich uitspreken over online haatspraak. [Slachtofferhulp Nederland](#) heeft recent ook een campagne gelanceerd om mensen bewust te maken van de (mogelijke) gevolgen van 'shame sexting'. Deze campagne zet zich niet direct

in op bewustwording van online discriminatie, maar richt zich wel op de mogelijke negatieve gevolgen van sociale media.

### Hulp aan gedupeerden

Op dit moment is er in Nederland nog weinig gespecialiseerde hulp voor gedupeerden van online discriminatie, zoals ook breder het geval is bij schadelijk of immoreel gedrag online (Van Huijstee et al., 2021). Na het meemaken van (online of offline) discriminatie, kunnen gedupeerden terecht bij Slachtofferhulp Nederland. Deze organisatie kan emotionele hulp bieden en begeleiden in een eventueel strafproces. Voor steun en advies kunnen gedupeerden ook terecht bij ADV's. Ook kunnen gedupeerden melding maken bij het eerder genoemde MiND, die niet zozeer ondersteuning biedt, maar meldingen op strafbaarheid controleert en ervoor kan zorgen dat een socialemediabedrijf met voorrang de melding onderzoekt.

### Aanpak vanuit socialemediabedrijven

Socialemediabedrijven, de bedrijven die online interacties op sociale media faciliteren, zijn volgens de huidige Europese wet- en regelgeving alleen verantwoordelijk voor de content waar zij zich bewust van zijn. Oftewel: als grensoverschrijdende content niet wordt gemeld of gemonitord, zijn ze daar op dit moment niet aansprakelijk voor. Wel hebben en nemen deze bedrijven verschillende activiteiten om online haatspraak en discriminatie tegen te gaan en te voorkomen. Hier worden vier aanpakken besproken: Communityrichtlijnen, Modereren, Deplatformen en Tools.

### Communityrichtlijnen

Sociale mediabedrijven hebben zogenaamde communityrichtlijnen. Deze richtlijnen zijn gemaakt om te zorgen voor een bepaalde gedragsnorm die wordt verwacht van de leden op de platforms om een veilige omgeving te creëren voor alle gebruikers. Deze richtlijnen variëren per platform. In de richtlijnen van [Meta](#) wordt aangegeven wat zij onder haatspraak definiëren:

*We definiëren haatdragend taalgebruik als een directe aanval op mensen (in plaats van concepten of instellingen) op basis van wat we beschermde kenmerken noemen: ras, afkomst, nationaliteit, beperking, religie, seksuele geaardheid, kaste, sekse, geslacht, genderidentiteit, of ernstige ziekten. We definiëren een aanval als gewelddadig of mensionterend taalgebruik, schadelijke stereotypen, beweringen over minderwaardigheid, uitingen van minachting, afschuw of verwerping, scheldwoorden en oproepen tot uitsluiting of segregatie.*

Vervolgens volgt een gedetailleerde lijst van content die verboden is te plaatsen op het platform. Bij soortgelijke socialemediakanalen als [Instagram](#) (onderdeel van Meta), [TikTok](#), [Twitter](#) en [YouTube](#) zien we een soortgelijk beleid met betrekking tot haat en discriminatie. De kern van deze gedragsregels is dat gebruikers van de kanalen niet mogen discrimineren of aanzetten tot haat (Blaya, 2019). Niet elk platform heeft duidelijk in de richtlijnen vermeld wat de gevolgen zijn als er een melding wordt gemaakt van haatzaaiende content, waaronder online discriminatie. Andere platforms doen dit wel, zoals Twitter. Twitter benoemt, anno oktober 2022, expliciet dat een van de gevolgen van haatzaaien kan zijn dat het account van deze persoon volledig wordt verwijderd van het socialemediaplatform (Twitter, z.d.).



Communityrichtlijnen ontwikkelen zich over de jaren, zo blijkt uit het interview met Meta. Zo zorgen verschillende maatschappelijke ontwikkelingen ervoor dat er meer druk komt te staan om de richtlijnen aan te scherpen. Denk bijvoorbeeld aan de toename van nepnieuws op sociale media (Van Huijstee et al., 2021).

Gebruiksvoorwaarden en gedragscodes worden doorgaans geëvalueerd als weinig effectieve middelen om online gedrag in goede banen te leiden en daarmee o.a. online discriminatie te voorkomen of te verminderen (Van Huijstee et al., 2021). De richtlijnen worden veelal niet gelezen en worden niet als gebruiksvriendelijk ervaren. Zo kost het een gemiddelde gebruiker van Instagram 31 minuten, van Facebook 34 minuten en van Twitter 46 minuten om deze gedragsregels door te lezen. Daarnaast zijn deze gedragsregels ook niet voor alle leeftijden verplicht (Costa & Halpern, 2019; Van Huijstee et al., 2021). Een onderzoek laat zien dat wanneer gebruikers de mogelijkheid hebben om de gedragsregels te accepteren zonder deze te hoeven lezen, de meeste gebruikers de gedragsregels niet lezen (Steinfeld, 2016).



## Modereren

Socialemediabedrijven kunnen de content op hun platforms modereren. Bij moderatie gaat het om het controleren en beoordelen van content of het gepast is om op het platform te staan. Dit proces van moderatie kan in verschillende fasen plaatsvinden: voordat het online komt te staan (pre-moderatie) of nadat het is gepubliceerd (post-moderatie) (Roberts, 2016; Veglis, 2014). Wanneer de gemodereerde content de gedragsregels van een platform overtreedt, kan de content worden verwijderd of geblokkeerd. Ook is het mogelijk dat de platforms overgaan tot hardere maatregelen door accounts te schorsen of te verbannen of het blokkeren van een IP-adres (Blaya, 2019; Rogers, 2020; Van Huijstee et al., 2021). Uit het interview met Meta komt naar voren dat door stevige inzet op kunstmatige intelligentie (AI) in 2021 bijna 96 procent van de online haatspraak proactief werd gevonden en tegengehouden op Facebook, voor Instagram gold dit voor bijna 92 procent. Dat is een flinke toename: in 2018 werd via AI rond de 25 procent proactief tegengehouden op Facebook en ruim 40 procent op Instagram (zie [hier](#)). Het grootste gedeelte van de content wordt gemodereerd via automatische detectie, slechts een klein deel van de online haatspraak wordt gerapporteerd door gebruikers.

Moderatie kan uitgevoerd worden door mensen, waarbij mensen de content beoordelen of het geschikt is om online te (blijven) staan. De kosten hiervan zijn echter hoog. Zo zijn er bij Facebook, YouTube en Twitter dagelijks duizenden mensen actief met het modereren van de content. Maar dit zijn er niet genoeg. Bij Facebook bijvoorbeeld zijn er op zijn minst dertigduizend moderatoren nodig om alle content te beoordelen. Bovendien is het niet altijd duidelijk voor de menselijke moderatoren welke online content verwijderd of geblokkeerd moet worden. Subjectieve beoordeling, zoals de (sociale) context, de normen en waarden van de moderator kunnen hierdoor invloed hebben op

welke content er online blijft staan (Barret, 2020; Papegnies, Labatut, Dufour & Linarès, 2018). Zo laat een onderzoek zien dat 28 moderatoren 2.400 verschillende labels ophangen aan een bepaalde tweet (Hettiachchi & Goncalves, 2019). Bovendien zijn er verhalen bekend van de erbarmelijke werkomstandigheden van menselijke moderatoren, zoals het onderbetaald krijgen, een (te) hoge werkdruk en de negatieve (mentale) gevolgen van het analyseren en beoordelen van de content. Deze moderatoren krijgen namelijk, onder meer, racistische, homofobe en vrouwonvriendelijke content onder ogen (Arsht & Etcovitch, 2018).

Het modereren van content door menselijk handelen heeft dus zo zijn haken en ogen. Veel socialemediabedrijven zetten daarom in op hybride moderatie. Hierbij wordt menselijke moderatie gecombineerd met moderatie door middel van automatische procesessen. Deze automatische processen filteren en categoriseren de online content. Dit zorgt ervoor dat er in korte tijd veel content kan worden geanalyseerd en waar nodig worden geblokkeerd of verwijderd (Blaya, 2019; Vidgen & Derczynski, 2020). Ook moderatie via automatische processen verloopt niet geheel vlekkeloos. De automatische processen richten zich voornamelijk op expliciete uitingen van discriminatie en haat. Meer verborgen vormen, impliciete discriminatie, ontglipt aan deze automatische processen (Salminen, Almerexhi, Milenković, Jung, An, Kwak & Jansen, 2018). Bij impliciete vormen van (online) discriminatie gaat het om gedrag dat niet per se discriminerend bedoeld is, maar wel zo wordt ervaren (Briels, Felten, Taouanza & Walz, 2018). Ook blijken specifieke vormen van online discriminatie zoals seksisme lastig te detecteren voor automatische systemen – onder andere doordat seksisme vaak niet als discriminatie of hatespeech wordt gezien door de mensen die dit handmatig labelen – in tegenstelling tot racisme en lhbt+ discriminatie (Davidson, Warmesley, Macy & Weber, 2017). Bredere

categorisatie van online haatspraak en discriminatie in automatische detectie kan ervoor zorgen dat meer content als discriminerend wordt gezien, maar die feitelijk niet discriminerend zijn. Dit zijn vals positieve berichten. Hiertegenover staan de vals negatieve berichten. Dit zijn berichten die discriminerend zijn maar niet als zodanig worden gezien, wat een gevolg kan zijn van een te smalle categorisatie van online discriminerende content (Jurgens, Chandrasekharan & Hemphill, 2019). Daarnaast nemen automatische processen voornamelijk scheldwoorden als uitgangspunt om te bepalen of content discriminerend is. Maar scheldwoorden worden vaak spontaan gebruikt en zijn onderdeel van informeel taalgebruik voor veel mensen (Pamungkas, Basile & Patti, 2020). Ook hoeven woorden niet per se discriminerend bedoeld te zijn, maar worden ze wel gebruikt om iemand te discrimineren en kunnen ze in bepaalde contexten wel degelijk strafbaar zijn. Denk hierbij aan woorden als ‘homo’ en ‘jood’. Een ander mogelijk gevaar dat komt kijken bij moderatie is dat uitingen gericht aan specifieke groepen sneller en/of meer verwijderd worden dan gericht aan andere groepen.

### Deplatformen

Wanneer discriminerende content is gesignaleerd door de socialemedia-bedrijven kunnen zij gebruikers verbannen of schorsen. Dit wordt ook wel *deplatforming* genoemd. Bij deplatformen gaat het om het sluiten, verbannen of blokkeren van een account dat de gedragsregels van een platform heeft overtreden (Radsch, 2021). Ook kunnen socialemediabedrijven over gaan tot blacklisting. Dit is het blokkeren van een IP-adres waardoor er vanaf dit adres geen toegang meer is tot een bepaalde website (Van Huijstee et al., 2021).

Een empirische studie laat zien dat het sluiten, verbannen of blokkeren van gebruikers die veel invloed hebben op het platform ervoor zorgt dat de over-

schrijdende content op termijn verdwijnt, ook onder de volgers. Ook de volgers van deze verbannen gebruiker plaatsen minder content op het platform na het deplatformen (Jhaver, Boylston, Yank & Bruckman, 2021). Een andere studie laat min of meer hetzelfde zien: door het verbannen van bepaalde online communities van de sociale nieuwswebsite Reddit, zogenaamde subreddits, nam de prevalentie van haatspraak onder de leden van deze subreddits af, ook wanneer zij naar andere subreddits trokken. Maar ze zagen ook dat veel leden van de verbannen subreddits vertrokken naar andere socialemediaplatforms waar hun gedrag hetzelfde bleef (Chandrasekharan, Pavalanathan, Srinivasan, GLynn, Eisenstein & Gilbert, 2017). Het onderzoek van Ali, Saeed, Aldreabi, Blackburn, De Cristofaro, Zannettou en Stringhuni (2021) toont dit ook aan. Wanneer gebruikers worden verbannen of geschorst van platforms zoals Twitter en Reddit vanwege het gebruik van haatspraak en het overtreden van de gedragsregels, zien de onderzoekers dat deze gebruikers vaak opduiken bij andere platforms, zoals sociaalnetwerksite Gab, bekend van het relatief grote aandeel rechts-conservatieve gebruikers. Op Gab is (nog) minder controle op de content en het gedrag van de gebruikers. Bovendien zien de onderzoekers dat de gebruikers die zich verplaatsen naar een ander platform nog giftiger gedrag vertonen. Het aantal mensen dat het bereikt met dit gedrag is wel aanzienlijk kleiner, vanwege de omvang van het platform. Het deplatformen van Twitter wordt ook door sommige gebruikers, op bijvoorbeeld Gab, gebruikt als iets om mee te pronken. Dit doen ze bijvoorbeeld door te benoemen in hun profiel beschrijving dat ze zijn verwijderd van Twitter (Ali et al., 2021).

Het deplatformen van gebruikers omdat ze zich schuldig hebben gemaakt aan online discriminatie zorgt er mogelijk voor dat deze gebruikers zich niet meer schuldig hieraan kunnen maken op dit platform. Maar het wil niet zeggen dat online discriminatie stopt. Het verplaatst zich (soms) naar andere platforms met minder strikte moderatie en richtlijnen. Dit is wat de experts in het onderzoek van Van Huijstee et al. (2021) ook aangeven. Ook meerdere experts die wij hebben gesproken, bevestigen dit beeld.

### Tools voor gebruikers

Verschillende socialemediabedrijven hebben tools ontwikkeld die (mede) gericht zijn op het tegengaan van online haat en discriminatie. Deze tools geven bijvoorbeeld een waarschuwing af aan iemand die bepaalde content wil plaatsen dat dit mogelijk aanstootgevend kan zijn. Ook kunnen gebruikers bepaalde tools inzetten om bijvoorbeeld bepaalde woorden te weren uit hun privéberichten die zij ontvangen of aangeven wie je wel of niet een bericht kan sturen op sociale media. Zie [hier](#) voor een overzicht van verschillende tools ontwikkeld door Meta.



# Wie zijn aan zet?

Online discriminatie is een complex probleem en de aanpak ervan ligt niet bij één maar bij verschillende organisaties, zo blijkt uit zowel de literatuur als de interviews met experts. Op basis hiervan zien wij dat onderstaande organisaties en actoren aan zet zijn. We richten ons hierbij niet op wat individuele personen kunnen doen wanneer zij discriminatie ervaren of hiervan getuige zijn. We beschrijven de organisaties en actoren van ‘groot’ naar ‘klein’.

## Rijksoverheid

De landelijke overheid heeft een cruciale rol bij het coördineren van de aanpak van online discriminatie in Nederland, zowel in preventieve als repressieve zin. Hieronder beschrijven we de rol van de verschillende ministeries in de huidige aanpak van online discriminatie:

- *Het ministerie van Onderwijs, Cultuur en Wetenschap*: het voorkomen van (online) gender- en lhbt+ discriminatie en de bevordering van mediawijsheid.
- *Het ministerie van Sociale Zaken en Werkgelegenheid*: het vergroten van kennis over (online) discriminatie bij de diverse hieronder genoemde stakeholders, het doen van onderzoek en het financieren van interventies.
- *Het ministerie van Binnenlandse Zaken en Koninkrijksrelaties* (directie constitutionele zaken en wetgeving): het monitoren of wetgeving afdoende is om (online) discriminatie effectief aan te pakken.

- *Het ministerie van Justitie en Veiligheid*: bundelen van krachten en kennis over het aanpakken van strafbare (online) discriminatie en de monitoring hiervan.
- *Het ministerie van Volksgezondheid, Welzijn en Sport*: zet zich in tegen (online) discriminatie in de sport en in de (gezondheids)zorg en onbelemmerde participatie van mensen met een beperking.
- *Het ministerie van Buitenlandse Zaken*: speelt een rol in het afstemmen en aanscherpen van Europese samenwerkingen in de aanpak van online discriminatie.
- *Het ministerie van Economische Zaken en Klimaat*: in dit ministerie is de samenwerking met socialemidiabedrijven belegd.

## De Nationaal Coördinator tegen Discriminatie en Racisme (NCDR)

De Nationaal Coördinator tegen Discriminatie en Racisme verbindt, jaagt aan en treedt op als waakhond als het gaat om de bestrijding van discriminatie en racisme. De NCDR valt onder de verantwoordelijkheid van de minister van Binnenlandse Zaken en Koninkrijksrelaties en werkt samen met diverse ministeries. In hun [Nationaal Programma tegen Discriminatie en Racisme 2022](#) vraagt de NCDR onder andere aandacht voor meer bekendheid van de meldpunten voor online discriminatie. Ook gaat de NCDR verkennen hoe een kennisbank kan worden opgezet om ‘good’ en ‘bad practices’ samen te brengen en te

ontsluiten. Ook wordt nader onderzocht of peer-to-peer educatie kan worden ingezet om online discriminatie en online haatspraak tegen te gaan.

### **Nationaal Coördinator Antisemitismebestrijding (NCAB)**

De Nationaal Coördinator Antisemitismebestrijding (NCAB) geeft gevraagd en ongevraagd advies over de aanpak van antisemitisme: discriminatie, bedreiging of intimidatie van de Joodse gemeenschap in Nederland. Ook signaleert de NCAB ontwikkelingen op het gebied van antisemitisme. Zo geeft de NCAB aan dat [online antisemitisme](#) te vaak zonder gevolgen blijft en dat het makkelijker moet worden om deze vorm van Jodenhaat te bestraffen.

### **Onderwijs**

Het onderwijs, zowel primair als voortgezet, speelt een belangrijke rol bij in de preventieve aanpak van online discriminatie. Enerzijds gaat het om educatie rondom mediawijsheid en online actief zijn, anderzijds over voorlichting en educatie over (online en offline) discriminatie. [Stichting School & Veiligheid](#), die scholen ondersteunt bij het bevorderen van een sociaal veilig klimaat, kan hierbij een ondersteunende rol spelen.

### **Openbaar Ministerie (OM)**

Het opsporingsonderzoek van de politie vindt plaats onder verantwoordelijkheid van de officier van justitie. Discriminatie, zowel online als offline, is een van de prioriteiten van het OM, naast onder andere mensenhandel, moord, huiselijk geweld en kindercriminaliteit. Dit betekent dat elke aangifte van discriminatie aan het OM moet worden voorgelegd en dat het OM altijd persberichten uitstuurt over uitspraken en geseponeerde zaken waarbij discriminatie een rol speelt. Medewerkers van het OM die vragen hebben over discriminatiezaken, kunnen terecht bij het Landelijk Expertisecentrum Discriminatie (LECD), het

kenniscentrum van het OM als het om discriminatiezaken gaat. Het LECD heeft als doel het optimaliseren van de strafrechtelijke handhaving op het gebied van discriminatie, dat wil zeggen standpuntbepaling, advisering, beleidsvorming, opsporing, vervolging en rapportage.

### **De rechtspraak**

Als het Openbaar Ministerie een zaak voor de rechter brengt, doet de rechter uitspraak over hetgeen online heeft plaatsgevonden. Over de effectiviteit van strafbaarstelling van discriminatie denken rechtswetenschappers verschillend, wel blijkt dat dit een belangrijke signaleringsfunctie heeft richting de samenleving (Grijsen, 2013).

### **Politie**

De politie heeft de formele taak om (online) discriminatie te onderzoeken en waar mogelijk aan te pakken. Wanneer personen online discriminatie ervaren of wanneer zij dit hebben gezien of gehoord, kunnen zij hiervan een melding of aangifte doen bij de politie. De politie neemt de aangifte op en gaat deze onderzoeken. Uitgangspunt is dat bij (online) discriminatiefeiten opsporingsonderzoek plaatsvindt. Het onderzoek wordt uitgevoerd door de politie, onder leiding van de officier van justitie. Het onderzoek bestaat onder andere uit het verzamelen van bewijs, het opsporen van verdachte(n) en het opstellen van een proces-verbaal. In het geval van online discriminatie is de politie gehouden aan het vastleggen van specifieke zaken zoals accountnaam, plaatsingsdatum en de periode waarin de uiting online heeft gestaan. Het proces-verbaal wordt vervolgens naar het Openbaar Ministerie gestuurd. Binnen de politie bestaat sinds een jaar het Expertise Centrum Aanpak Discriminatie (ECAD-P), een functie die momenteel nog wordt ingericht. Deze landelijke kennisfunctie heeft als doelen onder meer: advisering, operationele ondersteuning en rapportage.

## Socialemediabedrijven

Socialemediabedrijven kunnen een rol spelen bij zowel de preventieve als repressieve aanpak van online discriminatie. Een deel van hen is gebonden aan de Europese Code of Conduct, waarin wordt aangegeven dat platforms haatdragende uitingen die bij hen worden gemeld binnen 24 uur moeten verwijderen. De nieuwe Europese wetgeving Digital Services Act stelt ook extra verplichtingen aan platforms als het gaat om (de transparantie over) hun aanpak van online haat. De huidige aanpak van socialemediabedrijven verschilt. Meer informatie hierover is opgenomen in het stuk [aanpak vanuit socialemediabedrijven](#).

## Gemeenten

Gemeenten spelen een belangrijke rol in de aanpak van online discriminatie. Gemeenten zijn wettelijk verplicht om een onafhankelijk meldpunt in te richten, of dit beleggen bij de antidiscriminatievoorziening van de gemeenten waar inwoners een melding kunnen maken van discriminatie (Kros, Broekroelofs, Felten & Does, 2022) Vanuit het kabinet worden gemeenten gestimuleerd om antidiscriminatiebeleid te ontwikkelen en wordt aanbevolen om zowel in te zetten op apart beleid als een koppeling met bestaand beleid. Een beleid dat zichtbaar en merkbaar is voor de burger (NCDR, 2022). Daarnaast voeren gemeenten verschillend beleid dat bij kan dragen aan de aanpak van (online) discriminatie, denk hierbij aan emancipatie- en/of diversiteitsbeleid (Walz, 2017). Echter, uit de nog te publiceren monitoring van het lokaal antidiscriminatiebeleid 2022 van gemeenten blijkt dat gemeenten (nog) weinig aandacht besteden aan de aanpak van online discriminatie. Daarnaast wordt aangegeven dat gemeenten op dit thema behoefte hebben aan ondersteuning en heerst de wens hier meer op te doen (Kros et al., 2022).

## Antidiscriminatievoorzieningen

Elke gemeente is verplicht om zorg te dragen voor toegang tot een antidiscriminatievoorziening (ADV). ADV's zijn de spin in het web tussen gemeenten en de strafrechtketen en andere instanties (NCDR, 2022). De taak van een ADV is om meldingen van discriminatie te registreren en de melders te ondersteunen. Zij ontvangen ook meldingen van discriminatie op het internet, waaronder sociale media. Hier vind je antidiscriminatievoorzieningen.

## Maatschappelijke organisaties

Maatschappelijke organisaties, waaronder het jongerenwerk, hebben vaak direct contact met verschillende groepen in de maatschappij. Zodoende kunnen zij een rol spelen bij de preventieve aanpak van online discriminatie. Kennis over online discriminatie onder sociaal werkers (zoals jongerenwerkers) zou ervoor kunnen zorgen dat zij online discriminatie beter kunnen signaleren en ter sprake kan brengen onder hun doelgroepen. Op deze manier kan er gewerkt worden aan meer kennis over het onderwerp en meldingsbereidheid. In [dit rapport](#) van KIS (Van Beek, Sheombar, Van Hinsberg, Koch & Hermes, 2019) is meer te lezen over online jongerenwerk.

## Het College voor de Rechten van de Mens

Het College voor de Rechten van de Mens (het College) is een zelfstandig bestuursorgaan dat discriminatiekwesties beoordeelt aan de hand van de Algemene wet gelijke behandeling (AWGB). Ondanks dat online discriminatie vaak niet gaat over situaties van ongelijke behandeling, maar om uitingen van discriminatie, kan het College een rol spelen in de voorlichting, onderzoek en advies over online discriminatie. Het College heeft namelijk ook een beleid en onderzoeksbureau dat zich in de breedte richt op de bevordering van mensenrechten in Nederland. Het College valt onder het beheerdepartement van het

ministerie van Justitie en Veiligheid, deze verzorgt de basisfinanciering. [Lees meer over het College.](#)

### Meldpunt internetdiscriminatie (MiND)

MiND is het landelijke meldpunt voor strafbare, discriminerende uitingen op internet. Zij richten zich uitsluitend op discriminerende uitingen tegen een groep en niet tegen een individu. Mensen die online discriminatie ervaren of getuige hiervan zijn, kunnen een melding hiervan doen bij MiND. Het meldpunt voert vervolgens een juridische toetsing uit om te bepalen of er sprake is van een strafbare discriminatoire uiting. Indien dit het geval is, stuurt het meldpunt een verwijderingsverzoek naar het betreffende platform. MiND is een zogenoemde ‘trusted flagger’, dit houdt in dat een socialemediabedrijf met voorrang de melding van MiND onderzoekt. Het bedrijf beslist dan zelf of ze hier wel of niet aan voldoen. MiND is in 2013 opgericht op initiatief van het ministerie van Justitie en Veiligheid en het ministerie van Sociale Zaken en Werkgelegenheid. Het ministerie van Justitie en Veiligheid geeft jaarlijks subsidie aan MiND. [Lees meer over MiND.](#)

### INACH

Het International Network Against Cyber Hate (INACH) is een internationale structuur waarin organisaties kennis uitwisselen, waaronder over online discriminatie en haatspraak. Sinds 2016 houdt het INACH toezicht op de socialemediabedrijven die de Europese Code of Conduct hebben ondertekend en verzamelen en registreren zij de reactietijden na meldingen, de feedbackkwaliteit en verwijderingspercentages van de grootste platforms. Ook heeft het INACH o.a. een online trainingsaanbod die meer kennis biedt over mediageletterdheid, nepnieuws, cyberhaat en hoe je tegenspraak kunt bieden. [Lees meer over INACH.](#)

### Belangen- en zelforganisaties

Er zijn tal van organisaties die kennis en ervaring hebben op het terrein van (online en offline) discriminatie en die overheden en maatschappelijke organisaties kunnen adviseren en/of een intermediaire rol kunnen vervullen tussen specifieke (gemarginaliseerde) groepen en overheden of andere organisaties. Ook bieden deze zelf- en belangenorganisaties ondersteuning bij de aangifte voor specifieke vormen van discriminatie. Een aantal voorbeelden van belangen- en zelforganisaties: Centrum voor Informatie en Documentatie Israël (CIDI), Collectief Tegen Islamofobie & Discriminatie, Black Pride, Asian Raisins, IZI Solutions, stichting Maruf, S.P.E.A.K. en Nederland Wordt Beter. Ook zijn er diverse zelforganisaties die mensen met een specifieke migratieachtergrond vertegenwoordigen, bijvoorbeeld Samenwerkingsverband Marokkaanse Nederlanders (SMN) en het Inspraakorgaan Turken (IOT). In de bijlage van de Movisie-publicatie [Geen ruimte voor discriminatie](#) vind je meer zelf- en belangenorganisaties die zich bezighouden met de aanpak van discriminatie.

### Slachtofferhulp Nederland

Online discriminatie kan ingrijpend zijn voor degene die dit meemaakt. Slachtofferhulp Nederland biedt nazorg. Dit doen ze op emotioneel vlak, maar ook op het gebied van het meld- en strafproces. Zo bieden ze (o.a.) hulp bij de voorbereiding van een aangifte. Daarnaast geven ze advies over hoe je bewijsmateriaal kunt verzamelen van online discriminatie. Op de website van Slachtofferhulp Nederland staan tips hoe je kunt omgaan met online discriminatie, door bijvoorbeeld hulp en steun in de omgeving te vragen. [Lees meer over Slachtofferhulp Nederland.](#)

# Bijlage: begrippenlijst

## Discriminatie

Discriminatie is ‘het nadelig behandelen van personen omdat zij behoren, of gerekend worden, tot een bepaalde groepering’ (Andriessen, Nievers, Faulk, & Dagevos, 2010). In de Nederlandse wetgeving is aangegeven wanneer en op welke gronden omstrede behandeling verboden is en wanneer iets discriminatie is. Discriminatie volgens de wet kent verschillende vormen. Het kan gaan over ongelijke behandeling/uitsluitingsdelicten of over negatieve bejegening/discriminerende uitingen. Naast de wettelijke definiëring die betrekking heeft op objectief vastgestelde discriminatie, kan discriminatie betrekking hebben op een breder scala aan ervaringen van uitsluiting, achterstelling en/of onderdrukking.

## Strafbare discriminatie

Strafbare discriminatie gaat om discriminatie die is aangemerkt in het Wetboek van Strafrecht. De gronden waarop discriminatie verboden is, zijn: ras, godsdienst of levensovertuiging, geslacht, hetero- of homoseksuele gerichtheid of lichamelijke, psychische of verstandelijke handicap. Binnen het strafrecht zijn er twee vormen van discriminatie: uitsluitingsdelicten en uitingsdelicten. Soms zit er overlap tussen discriminatie volgens het Wetboek van Strafrecht en de AWGB. Strafbare discriminatie kan vervolgd worden door het Openbaar Ministerie en bestraft met een geldboete, taakstraf of gevangenisstraf. Binnen het Strafrecht is er geen onderscheid tussen online of offline discriminatie.

## Ervaren discriminatie

Naast de nauwe, juridische definiëring, kan discriminatie ook betrekking hebben op een breder palet aan ervaringen van uitsluiting. Het kan gaan over discriminatie op gronden die niet in de wet zijn aangemerkt als discriminatiegrond, of om ervaringen of gevoelens waar geen juridisch vervolg aan is gegeven en als zodanig zijn vastgesteld, of om ervaringen die niet of moeilijk objectief vast te stellen zijn, zoals wanneer discriminatie plaatsvindt in de vorm van micro-agressies, alledaagse handelingen die als discriminatie kunnen worden ervaren, of wanneer het gaat om een duidelijk gevoel wat lastig aan te tonen is met objectief bewijs (Andriessen et al., 2020).

## Online discriminatie

Bij online discriminatie gaat het om discriminerende uitingen die online op internet plaatsvinden, bijvoorbeeld op sociale media. Online discriminatie kan geuit worden richting een persoon vanwege bijvoorbeeld diens (vermeende) herkomst, maar het kan ook gaan om groepsbelediging. Binnen het Strafrecht gaat het om een zogenaamd uitingsdelict.

Individuele online discriminatie betekent dat een persoon wordt gedenigreerd of buitengesloten op basis van kenmerken als etniciteit, religie, sekse, seksuele voorkeur of huidskleur door het gebruik van online content zoals emoti-



cons, gesproken berichten, filmpjes, afbeeldingen, memes, gifjes en teksten. Wanneer een individu online discriminatie ervaart kan deze persoon hiervan aangifte doen.

Het kan ook zijn dat online discriminatie niet direct gericht is op een individu, maar dit wel door een individu als discriminerend wordt ervaren. Het gaat hier om 'grappen', negatieve afbeeldingen of beledigingen die worden geuit over de groep waartoe iemand zichzelf rekent (Felten & Taouanza, 2017).

## (Online) haatspraak (hatespeech)

Een specifieke vorm van online haat is haatspraak (of hatespeech of haatzaaien). Er zijn verschillende definities van haatspraak in omloop. In 2022 bracht het Comité van Ministers van de Raad van Europa de volgende definitie naar buiten: haatspraak zijn alle soorten uitingen die aanzetten tot geweld of geweld bevorderen, verspreiden of rechtvaardigen, haat of discriminatie tegen een persoon of een groep personen, of die hen kleineren vanwege hun daadwerkelijke of toegeschreven persoonlijke kenmerken of status zoals 'ras', kleur, taal, religie, nationaliteit, nationale of etnische afkomst, leeftijd, handicap, geslacht, genderidentiteit en seksuele geaardheid. In Nederland is discriminatie strafbaar gesteld in artikel 137c, 137d en 137e, haatspraak kan daar onderdeel van uitmaken, maar soms valt het onder vrijheid van meningsuiting.

Door sommige groeperingen worden er ook bepaalde taal en *emojicons* gebruikt die een dubbele betekenis hebben. Wat een normale of nietszeggende uitspraak is voor veel mensen, heeft dan voor een specifieke groep extra lading. Dit wordt ook wel *dog whistle* genoemd; een hondenfluitje dat niet hoorbaar is voor de meeste mensen. Een *dog whistle* kan ook een discriminerende of haat-

volle boodschap hebben, denk aan [Pepe the Frog](#) of '88' wat staat voor 'Heil Hitler' (twee keer de achtste letter van het alfabet, de H).

## Ongelijke behandeling

Wanneer een persoon – direct of indirect – ongelijk behandeld wordt, kan die persoon een beroep doen op de Algemene wet gelijke behandeling (AWGB). Binnen de AWGB is het verboden om onderscheid te maken op grond van godsdienst, levensovertuiging, politieke gezindheid, ras, geslacht, nationaliteit, hetero- of homoseksuele gerichtheid of burgerlijke staat. In bijzondere gevallen mag er wel onderscheid gemaakt worden op een van deze gronden, dan is hiervoor een objectieve rechtvaardiging nodig. Het College voor de Rechten van de Mens (het College) kan oordelen of er sprake is van onrechtmatige ongelijke behandeling. Het College kan echter geen straf toekennen aan de discriminerende partij.



# Literatuur

Ali, S., Saeed, M. H., Aldreabi, E., Blackburn, J., De Cristofaro, E., Zannettou, S., & Stringhini, G. (2021). Understanding the Effect of Deplatforming on Social Networks. *13th ACM Web Science Conference 2021*.

Andriessen, I., Hoegen Dijkhof, J., Van der Torre, A., Van den Berg, E., Pulles, I., Iedema, J., & De Voogd-Hamelink, M. (2020). *Ervaren discriminatie in Nederland II*. Den Haag: Sociaal en Cultureel Planbureau.

Arsht, A., & Etcovitch, D. (2018). The human cost of online content moderation. *Harvard Law Review Online, Harvard University, Cambridge, MA, USA*.

Barlett, C. P., Gentile, D. A., & Chew, C. (2016). Predicting cyberbullying from anonymity. *Psychology of Popular Media Culture, 5(2)*, 171.

Barret, P. M. (2020). *Social media giants? A call to end outsourcing*. Center for Business and Human Rights.

Blaya, C. (2019). Cyberhate: A review and content analysis of intervention strategies. *Aggression and Violent Behavior, 45*, 163-172.

Briels, B., Felten, H., Taouanza, I., & Walz, G. (2018). *Handreiking antidiscriminatiebeleid voor gemeenten*. Utrecht: Movisie.

Burnap, P., & Williams, M. L. (2016). Us and them: identifying cyber hate on Twitter across multiple protected characteristics. *EPJ Data Science, 5(11)*, 1-15.

Centraal Bureau voor de Statistiek. (2021, 15 oktober). *Internettoegang en internetactiviteiten; persoonskenmerken* [Dataset]. [Verkegen van Internettoegang en internetactiviteiten; persoonskenmerken \(cbs.nl\)](https://www.cbs.nl/en-gb/verkegen-van-internettoegang-en-internetactiviteiten-persoonskenmerken)

Chandrasekharan, E., Pavalanathan, U., Srinivasan, A., Glynn, A., Eisenstein, J., & Gilbert, E. (2017). You Can't Stay Here. *Proceedings of the ACM on Human-Computer Interaction, 1(CSCW)*, 1-22.

CIDI. (2022). *Monitor antisemitische incidenten 2021*. Den Haag: CIDI.

College voor de Rechten voor de Mens. (z.d.). *Wat is discriminatie?* Geraadpleegd op 21 september 2022, van [Wat is discriminatie? | Mensenrechten voor jou | College voor de Rechten van de Mens](https://www.vrechten.nl/wat-is-discriminatie/).

Costa, E., & Halpern, D. (2019). *The behavioural science of online harm and manipulation. And what to do about it: an explanatory paper to spark ideas and debate*. The Behavioural Insights Team.

Council of Europe. (2022). *Combating hate speech*. Council of Europe.

Davidson, T., Warmsley, D., Macy, M., & Weber, I. (2017). Automated hate speech detection and the problem of offensive language. In *Proceedings of the international AAAI conference on web and social media*, 11(1), 512-515.

ECRI. (2019). *ECRI-rapport over Nederland (vijfde monitoringcyclus)*. Strasbourg: ECRI/Council of Europe.

European Commission. (z.d.). *The Digital Service Act package*. Geraadpleegd op 23 november 2022, van [The Digital Services Act package | Shaping Europe's digital future \(europa.eu\)](#).

European Commission. (z.d.). *STOPHATE: The legal and policy framework in the EU*. Geraadpleegd op 17 november 2022, van [Combating hate speech and hate crime | European Commission \(europa.eu\)](#).

Felten, H., & Taouanza, I. (2017). *Online discriminatie aanpakken: wat werkt? Een eerste verkennend onderzoek*. Utrecht: Movisie.

Fox, J., Cruz, C., Young Lee, J. (2015). Perpetuating online sexism offline: Anonymity, interactivity, and the effects of sexist hashtags on social media. *Computers in Human Behavior*, 52, 436-442.

Grijsen, C. (2013). *De handhaving van discriminatiewetgeving in de politiepraktijk*. Den Haag: Boom Juridische Uitgevers.

Hettiachchi, D., & Goncalves, J. (2019). Towards Effective Crowd-Powered Online Content Moderation. *Proceedings of the 31st Australian Conference on Human-Computer-Interaction*.

Ikram, M. U. Z. (2016). *Social determinants of ethnic minority health in Europe*. [Thesis, fully internal, Universiteit van Amsterdam].

Jeong, S. H., Cho, H., & Hwang, Y. (2012). Media literacy interventions: A Meta-analytic review. *Journal of Communication*, 62(3), 454-472.

Jhaver, S., Boylston, C., Yang, D., & Bruckman, A. (2021). Evaluating the effectiveness of deplatforming as a moderation strategy on Twitter. *ACM Hum.-Comput. Interact*, 5(CSCW2), 1-30.

Jurgens, D., Chandrasekharan., & Hemphill, L. (2019). A just and comprehensive strategy for using NLP to address online abuse. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*.

Kaal, M., & Bouwman, J. (2020). *Antisemitisme op Twitter, websites, blogs en fora*. Amsterdam: Kantar.

Kilvington, D., & Price, J. (2017). Tackling Social Media Abuse? Critically Assessing English Football's Response to Online Racism. *Communication & Sport*, 7(1), 1-16.

Kros, K., Broekroelofs, R., Felten, H., & Does, S. (2022). *Monitor lokaal antidiscriminatiebeleid 2022*. Utrecht: Kennisplatform Inclusief Samenleven. Rapport in voorbereiding.

Landesanstalt für Medien NRW (2018). *Ergebnisbericht der forsa-Befragung 2018 zur Wahrnehmung von Hassrede*. NRW.

Matthew, B., Dutt, R., Goyal, P., & Mukherjee, A. (2019, juni). Spread of hate speech in online social media. In *Proceedings of the 10th ACM conference on web science*, 173-182.

MiND. (z.d.). Over MiND. Meldpunt Internet Discriminatie. Geraadpleegd op 3 oktober 2022, van [MiND | meldpunt internet discriminatie \(mindnederland.nl\)](https://mindnederland.nl).

NCDR. (2022). *Nationaal Programma tegen Discriminatie en Racisme 2022*. Den Haag: Nationaal Coördinator tegen Discriminatie en Racisme/Ministerie van Binnenlandse Zaken en Koninkrijksrelaties.

Netwerk Mediawijsheid. (2021). *Monitor Mediagebruik kinderen 7 – 12 jaar*. Hilversum: Netwerk Mediawijsheid.

Nikken, P., & Jongmans, L. (2013). *Mediawijsheid: Informatie en inspiratie voor gemeentelijk beleid*. Utrecht: Nederlands Jeugdinstituut.

Olmo, J. (2017). *Evaluatie training: Reageren op discriminatie*. De Meern: Bureau Olmo.

Pamungkas, E. W., Basile, V., & Patti, V. (2020). Do you really want to hurt me? predicting abusive swearing in social media. In *The 12th Language Resources and Evaluation Conference*. European Language Resources Association, 6237-6246.

Papegnies, E., Labatut, V., Dufour, R., & Linarès, G. (2018). Impact of Content Features for Automatic Online Abuse Detection. *Lecture Notes in Computer Science*, 404-419.

Radsch, C. (2021). Deplatforming/De-platforming. In *Belli, L., Zingales, N., & Curzi, Y. (Red.), Glossary of Platform Law and Policy Terms* (pp. 109-112). Rio de Janeiro. Brazilië: FGV Direito Rio. Verkregen van [0. MIOLO Glossary of Platform Law digital.pdf \(fgv.br\)](https://www.fgv.br/direito/rio/glossary-of-platform-law-digital.pdf).

Reynders, D. (2022). *Countering illegal hate speech online; 7th evaluation of the Code of Conduct*. European Commission.

Rijksoverheid. (z.d.). *Wat kan ik doen als ik gediscrimineerd word?* Geraadpleegd op 10 oktober 2022, van [Wat kan ik doen als ik gediscrimineerd word? | Rijksoverheid.nl](https://www.rijksoverheid.nl/onderwerpen/wat-kan-ik-doen-als-ik-gediscrimineerd-word).

Roberts, S. T. (2016). Commercial content moderation: Digital laborers' dirty work. *Media Studies Publications*, 12.

Rogers, R. (2020). Deplatforming: Following extreme internet celebrities to Telegram and alternative social media. *European Journal of Communication*, 0(0), 1-17.

RTL Nieuws (2021, 29 maart). *Racisme en haat op social media onbestraft, Kamer wil pakkans vergroten*. Geraadpleegd op 1 oktober 2022, van [Racisme en haat op social media onbestraft, Kamer wil pakkans vergroten | RTL Nieuws](https://www.rtlnieuws.nl/nieuws/2021/03/29/racisme-en-haat-op-social-media-onbestraft-kamer-wil-pakkans-vergroten).

Rudnicki, K., & Steiger, S. (2020). *Online hate speech – Introduction into motivational causes, effects and regulatory contexts*. Detect Then Act.

Salminen, J., Almerexhi, H., Milenković, M., Jung, S., An, J., Kwak, H., & Jansen, B. J. (2018). Anatomy of online hate: developing a taxonomy and machine

learning models for identifying and classifying hate in online news media. In *Proceedings of the Twelfth International AAAAI Conference on Web and Social Media (ICWSM)*.

Steinfeld, N. (2016). "I agree to the terms and conditions": (How) do users read privacy policies online? An eye-tracking experiment. *Computers in Human Behavior*, 55, 992-1000.

Twitter. (z.d.). *Beleid ten aanzien van haatdragend gedrag*. Geraadpleegd op 20 oktober 2022, van [Twitter-beleid ten aanzien van haatdragend gedrag | Twitter-hulpcentrum](#).

Van Huijstee, M., Van Nieuwenhuizen, W., Sanders, M., Masson, E., & Van Boheemen, P. (2021). *Online ontspoord: Een verkenning van schadelijk en immoreel gedrag op het internet in Nederland*. Den Haag: Rathenau Instituut.

Van Wonderen, R., & Peeters-Osseyan, M. (2022). *Werken aan weerbaarheid tegen desinformatie en eenzijdige meningsvorming. Evaluatie lesprogramma Under Pressure*. Utrecht: Verwey-Jonker Instituut.

Veglis, A. (2014). Moderation Techniques for Social Media Content. *Lecture Notes in Computer Science*, 137-148.

Verloove, J., Van Wonderen, R., & Felten, H. (2020). *Theorieën en aanpakken van polarisatie. Uitgebreide rapportage*. Utrecht: Kennisplatform Integratie & Samenleving.

Vidgen, B., & Derczynski, L. (2020). Directions in abusive language training data, a systematic review: Garbage in, garbage out. *PLoS ONE*, 15(12), 1-32.

Walz, G. (2017). *Lokaal antidiscriminatiebeleid. Het perspectief van Nederlandse gemeenten*. Utrecht: Movisie.

Walz, G., & Fiere, B. (2022). *Discriminatiecijfers in 2021. Een rapport over registraties van discriminatie-incidenten door de politie, en meldingen bij antidiscriminatievoorzieningen en andere organisaties in Nederland*. Rotterdam/Den Haag: Art. 1.

Week van de Mediawijsheid (2022, 16 november). De Week van de Mediawijsheid zit erop! Dit waren de highlights. Geraadpleegd op 9 december 2022, van *De Week van de Mediawijsheid zit erop! Dit waren de highlights*.

## Colofon

Auteurs: Joline Verloove, Koen Kros, Arwen Hoogenbosch en Serena Does  
Met medewerking van: Leyla Reches, René Broekroelofs en Hanneke Felten  
Met dank aan: alle geïnterviewde experts  
Financier: Ministerie van Sociale Zaken en Werkgelegenheid  
Vormgeving: Design Effects, Willem Plaisier

Uitgave: Kennisplatform Inclusief Samenleven  
p/a Kromme Nieuwegracht 6  
3512 HG Utrecht  
T (030) 230 3260

De publicatie kan gedownload worden via de website van het Kennisplatform Inclusief Samenleven.

© Kennisplatform Inclusief Samenleven, Utrecht 2022.

Het auteursrecht van deze publicatie berust bij Movisie. Gedeeltelijke overname van teksten is toegestaan, mits daarbij de bron wordt vermeld.  
The copyright of this publication rests with Movisie. Partial reproduction of the text is allowed, on condition that the source is mentioned.

## Kennisplatform Inclusief Samenleven

Kennisplatform Inclusief Samenleven doet onderzoek, adviseert en biedt praktische tips en instrumenten over vraagstukken rond integratie, migratie en diversiteit. Daarnaast staat het platform open voor vragen, signalen en meningen en formuleert daar naar beste vermogen een antwoord op.

Deze kennisuitwisseling is bedoeld om een fundamentele bijdrage te leveren aan een pluriforme en stabiele samenleving. Blijf op de hoogte van alle projecten, vragen en antwoorden en andere kennisuitwisseling via [www.kis.nl](http://www.kis.nl), de nieuwsbrief, Twitter en LinkedIn.

Kennisplatform Inclusief Samenleven is een programma van het Verwey-Jonker Instituut en Movisie  
T 030 230 32 60  
E [info@kis.nl](mailto:info@kis.nl)  
I [www.kis.nl](http://www.kis.nl)

