



# Online geweld tussen burgers onderling

Wat weten we al?

# 1 Inhoud

1	Inleiding	3
2	Wat is online geweld?	4
3	Omvang van strafbaar online geweld tussen burgers in Nederland	6
3.1.	Online intimidatie in Nederland	6
3.2.	Online haatzaaien in Nederland	6
3.3.	Online aanjagen van geweld en opruiing	6
3.4.	Online geweld onder jongeren	7
4	Belangrijke beïnvloedende factoren	8
4.1.	Enkele verklarende theorieën	8
5	Kenmerken van burgers betrokken bij online geweld	10
6	Impact van online geweld op individuen en de maatschappij	12
6.1.	Impact op individueel niveau	12
6.2.	Impact op maatschappelijk niveau	12
6.3.	Impact specifiek voor jongeren	13
7	Mogelijke interventies en beleidsmaatregelen	14
7.1.	Interventies gericht op individueel niveau	14
7.2.	Interventies gericht op gemeenschapsniveau	15
7.3.	Beleidsmaatregelen en wet- en regelgeving	16
7.4.	Technologische interventies	17
8	Conclusie/aanbevelingen	20
9	Bronnen	21

# 1 Inleiding

(Doods)bedreigingen via sociale media. Het online aansporen van mensen om deel te nemen aan geweld. Racistische en haatdragende berichten en campagnes die viraal gaan. Dit zijn slechts enkele voorbeelden van strafbare vormen van online geweld. Door het grote bereik, het gevoel van anoniem zijn en de directe verspreiding van online berichten, kan online geweld snel escaleren. De impact is groot: slachtoffers ervaren psychische schade, reputatieschade en angst, terwijl de samenleving kampt met polarisatie, een afname van vertrouwen in digitale platforms en een groeiend gevoel van onveiligheid online. Om dit complexe probleem aan te pakken, is meer inzicht nodig. In opdracht van het Ministerie van Justitie en Veiligheid voeren het Verwey Jonker Instituut en het Centre of Expertise Veiligheid en Veerkracht van Avans Hogeschool onderzoek uit naar **strafbare vormen van online geweld tussen burgers onderling**<sup>[1]</sup>. In dit onderzoek kijken we naar de aard en omvang, onderliggende mechanismen, motieven en kenmerken van plegers en slachtoffers en mogelijke interventies.

## Leeswijzer

Dit literatuuroverzicht is gebaseerd op een (inter)nationale literatuurstudie waarin er zowel naar wetenschappelijke als grijze literatuur is gekeken. Deze publicatie biedt inzicht in:

- Wat we verstaan onder online geweld en de juridische en maatschappelijke definities van verschillende vormen (2).
- De aard en omvang van online geweld tussen burgers; specifiek inzoomend op: bedreigingen & intimidatie, discriminatie & haatspraak<sup>[2]</sup>, kleinschalige oproepen tot geweld (3).
- Kenmerken van plegers en slachtoffers van online geweld (4).
- Beïnvloedende factoren van online geweld en inzichtvolle theorieën (5).
- Mogelijke beleidsmaatregelen en interventies (6).

1 Dit onderzoek bestaat uit de volgende onderzoeksactiviteiten: literatuuronderzoek, analyse van kwantitatieve data over online geweld onder jongeren, social media analyse, interviews met praktijk-, wetenschappelijke experts en sleutelfiguren in de aanpak, interviewgesprekken met plegers en slachtoffers, focusgroepen met professionals en een reflectiesessie over de aanbevelingen uit het rapport. Zie hier meer over het onderzoek: [Onderzoek naar strafbare vormen van online geweld - Verwey-Jonker Instituut](#).

2 Voor dit onderzoek onderscheiden we haatzaaien van bredere vormen van discriminatie en racisme. Hoewel racisme en discriminatie belangrijke vormen van online geweld zijn valt dit buiten de scope van dit onderzoek.



## 2 Wat is online geweld?

*Online geweld* verwijst naar agressief gedrag dat zich afspeelt op digitale platforms en sociale media. Dit geweld richt zich vaak op individuen of groepen door middel van bedreigingen, haatdragende inhoud en verspreiding van privé-informatie. Een duidelijk voorbeeld van online geweld is het gerichte geweld tegen minderheden zoals LHBTI+ groepen, vrouwen, mensen met een migratieachtergrond waarbij haatdragende berichten en bedreigingen (b.v. via memes, tweets en video's) online verspreid worden. Online geweld is vaak niet alleen verbaal of emotioneel, maar bevat ook het verspreiden van onjuiste informatie, het openbaar maken van persoonlijke gegevens (doxing) en zelfs het aanzetten tot fysiek geweld en opruiing (Obemaier et al 2024; Wright & Wachs 2019; Castano-Pulgarin et al 2021; KIS 2022). Online geweld kan zich uiten in verschillende vormen (Peterson & Densley 2017; Mukred et al 2014). Tabel 1 licht de verschillende vormen van online geweld verder toe.

**Tabel 1. Overzicht juridische en maatschappelijke definities en kenmerken van strafbaar online geweld**

Soort geweld	Juridische definitie (Nederland)	Bredere toelichting/ maatschappelijke kenmerken
<b>Online bedreigingen en intimidatie</b>	Bedreiging, zoals omschreven in artikel 285 van het Wetboek van Strafrecht, is een strafbaar feit waarbij iemand een ander persoon dreigt met geweld, doodslag, zware mishandeling, of een ander ernstig misdrijf. Dit geldt ongeacht de frequentie van de bedreigingen, of ze nu schriftelijk of mondeling zijn, en of ze online of offline plaatsvinden. Ook als er twijfel is over de uitvoering van het dreigement, kan dit strafbaar zijn.	Met de opkomst van het internet komen online bedreigingen vaker voor. Deze worden via digitale kanalen geuit. Dit kan zijn via sociale media, e-mail, chatplatforms, forums, of online games. De bedreigingen kunnen gericht zijn op de persoon zelf, hun naasten, of hun eigendommen en kunnen variëren van cyberpesten, dreigen met fysiek geweld tot het onthullen van persoonlijke informatie (doxing).
<i>"Als je niet stil bent, steek ik je huis in brand."</i>		
<i>"Als je geen seks met mij wilt, dwing ik je wel."</i>	Niet iedere bedreigende uitlating is echter strafbaar. Een strafbare bedreiging moet voldoen aan bepaalde vereisten:	
<i>"Ik zorg ervoor dat je kinderen iets overkomt."</i>	<ul style="list-style-type: none"> <li>• aard van de bedreiging;</li> <li>• opzet met het ontstaan van vrees;</li> <li>• kennisname van de bedreiging.</li> </ul>	

Soort geweld	Juridische definitie (Nederland)	Bredere toelichting/ maatschappelijke kenmerken
<b>Haatzaaien</b>	Haatzaaien kan worden gezien als een manier om anderen aan te zetten tot gewelddadig of ontwrichtend gedrag. Discriminatie wordt daarbij vaak ingezet als middel om haat aan te wakkeren. Het aanzetten tot haat, discriminatie of gewelddadig optreden tegen bepaalde groepen is strafbaar volgens artikel 137d van het Wetboek van Strafrecht. Dit geldt voor openbare uitlatingen, zowel mondeling als schriftelijk, en omvat ook online uitingen. De wet beschermt specifieke groepen op basis van kenmerken zoals ras, godsdienst of levensovertuiging, geslacht, hetero- of homoseksuele gerichtheid en lichamelijke, psychische of verstandelijke handicap. De Hoge Raad definieert "aanzetten" als het (trachten) bewegen van iemand in een bepaalde richting, terwijl haat wordt omschreven als diepe afkeer en vijandigheid. Discriminatie wordt gezien als elke vorm van onderscheid die gelijkheid in rechten en vrijheden aantast. Geweld wordt niet expliciet gedefinieerd, maar kan uit jurisprudentie worden afgeleid, waarbij context en toon van de uitlatingen een rol spelen.	Via memes, muziek, gebaren, en slogans, video's en berichten op verschillende platformen kunnen haatdragende ideologieën geprezen en gepromoot worden (b.v. misogynie, anti-lhbt+, antisemitisme, anti-moslim).
<i>"Daarom gaan we individuele moslims, Turken en Marokkanen goed in de gaten houden. Ook volgen we de vuiligheid die in de moskee te horen is".</i>		
<i>"Gaskamers bouwen voor moslims. Meer moslimhaat, meer haatzaaien"</i>		
<i>"Wat moet zo'n F****G n*g*r iets zeggen over ons Nederlandse traditie. Oprotten nu, voordat de echte Nederlanders dat doen met jou".</i>		
<i>"Zolang die k*t n*g*r tegen Zwarte Piet blijft strijden en ook mag aanzetten tot geweld. Mag iedereen dat."<sup>3]</sup></i>		

3 Voorbeelden gebaseerd op Nederlandse berichten op X. Deze komen voort uit de sociale media analyse 2025. Publicatie van factsheet op basis van sociale media analyse verwacht mei 2025.

Soort geweld	Juridische definitie (Nederland)	Bredere toelichting/ maatschappelijke kenmerken
<b>Oproepen tot geweld<sup>4)</sup></b>  <i>"Gooi die winkelruit in!"</i>  <i>"Sloop alles wat je ziet"</i>	Opruiing, oftewel het oproepen of aanzetten tot geweld, is strafbaar onder de artikelen 131 en 132 van het Wetboek van Strafrecht. Dit geldt ook voor online uitingen, zoals het verspreiden van opruiende teksten of afbeeldingen. Opruiing houdt in dat iemand aanspoort tot een strafbaar feit of gewelddadig optreden tegen het openbaar gezag, zonder dat het strafbare feit daadwerkelijk hoeft te worden uitgevoerd. De uitlating moet openbaar zijn en mondeling, schriftelijk of via afbeeldingen worden gedaan.	Voorbeelden van 'grootschalige' ordeverstoringen zijn: oproepen tot geweld bij demonstraties, aanzetten tot rellen, dreigen met aanslagen, haatzaaien tegen minderheden, illegale feesten en plunderingen. Dit onderzoek is gericht op kleinschalige incidenten, zoals uit de hand gelopen ruzies die ontstaan als gevolg van online berichten (b.v. drillrap, ruzies op scholen). Sociale media zijn bij uitstek geschikt voor opruiing van groepen. Middels tekstberichten, memes, video's en livestreams kunnen mensen elkaar aansporen de straat op te gaan en mee te rellen. Onder het oog van de smartphonecamera's motiveren relschoppers elkaar tot nóg extremer gedrag. De beelden worden vervolgens verspreid via platformen zoals Instagram en Snapchat.



4 Binnen dit onderzoek zal de focus liggen op kleinschalige incidenten zoals ruzies en kleinschalige rellen die ontstaan door online content.

# 3 Omvang van strafbaar online geweld tussen burgers in Nederland

Hoewel er steeds meer cijfers beschikbaar zijn over strafbaar online geweld tussen burgers in Nederland, geven deze geen volledig of zelfs consistent beeld van de werkelijke omvang van online geweld in Nederland. Ten eerste doet slechts een klein deel van de slachtoffers melding bij instanties vanwege angst voor vergelding, het gevoel dat het toch niets oplevert omdat instanties er weinig aan kunnen doen en onduidelijkheid over wat strafbaar online geweld precies inhoudt (KIS, 2022; CBS, 2022). Ten tweede lijken de meeste onderzoeken en cijfers te gaan over online intimidatie, online haatzaaien en discriminatie en niet over het online aanjagen van geweld of opruiing. Tenslotte hanteren overheidsinstanties, onderzoeksinstituten en meldpunten verschillende definities van de vormen van online geweld en gebruiken ze uiteenlopende indicatoren. Hierdoor blijft het moeilijk om een compleet en consistent overzicht te krijgen van de omvang van strafbaar online geweld in Nederland. Bovendien zijn er mogelijk cijfers beschikbaar van techbedrijven, bijvoorbeeld over het aantal meldingen van grensoverschrijdend gedrag op hun platforms. Deze gegevens worden echter vaak niet openbaar gedeeld, waardoor het moeilijk is om een vollediger en consistent beeld van online geweld te verkrijgen. Dit is een punt waarop we in de conclusie wellicht verder kunnen reflecteren.

## 3.1. Online intimidatie in Nederland

In 2022 gaf 4% van Nederlanders (15+) aan slachtoffer te zijn geweest van online bedreigingen of intimidatie. Dit komt neer op ongeveer 600.000 mensen. Jongeren, vooral jonge vrouwen en lhbt+-personen lopen een hoger risico. Online bedreiging (1,9%), stalking (1,2%) en shamesexting (0,7%) komen het meest voor (6. Online bedreiging en intimidatie | CBS). Terwijl in 2024 10% van Nederlanders 16+ aangaven ooit slachtoffer te zijn geweest van (seksuele) intimidatie (ongeveer 760.000).

## 3.2. Online haatzaaien in Nederland

Vanwege het lage aantal meldingen en aangifte is er een incompleet beeld van de daadwerkelijke omvang van online discriminatie. De rapporten van uiteenlopende instanties, meldpunten, en onderzoeksinstituten tekenen een divers beeld. Bijvoorbeeld, CBS registreerde 535 discriminatiefeiten, waarvan 28% online plaatsvond. Dit omvatte discriminatie op sociale media, forums en andere digitale platforms. In 2022 gaf volgens CBS 2% van Nederlanders van 15 jaar en ouder aan zich in de afgelopen 12 maanden online gediscrimineerd te hebben gevoeld. Gronden van online discriminatie waren: ras/huidskleur (40%), nationaliteit (34%), godsdienst of levensbeschouwing (29%), politieke overtuiging (25%), geslacht (25%), seksuele gerichtheid (25%) (Online discriminatie | CBS). Terwijl MiND in 2023 3.305 meldingen van online discriminatie registreerde, een stijging van bijna 20% ten opzichte van 2022. Ongeveer 64% van de meldingen ging over haatdragende taal, vooral op sociale media. Discriminatie op basis van herkomst en huidskleur kwam het vaakst voor (48%), gevolgd door seksuele gerichtheid (15%) en religie (12%). Racisme en antisemitisme waren de meest voorkomende thema's (Discriminatiecijfers 2023 | Meld.Online Discriminatie | Meld.Online Discriminatie). De CIDI Monitor in 2023 meldde 245 meldingen van online antisemitisme, een stijging van 180% ten opzichte van 2022. Dit is het hoogste aantal ooit geregistreerd. Van alle antisemitische incidenten in 2023 vond 38% online plaats. De piek werd veroorzaakt door het conflict in Gaza na 7 oktober, waarbij sociale media werden overspoeld met haatdragende uitingen tegen Joden. Vooral directe bedreigingen en complottheorieën over Joden en Israël namen toe (Monitor Antisemitische Incidenten · CIDI). De daadwerkelijke omvang ligt vermoedelijk hoger vanwege onderrapportage en het wijdverspreide karakter van online haat. Zo vermeldde KIS (2022) dat ongeveer 60% van Nederlanders aangeeft ooit online discriminatie te hebben waargenomen (Online Discriminatie Rapport).

## 3.3. Online aanjagen van geweld en opruiing

Cijfers over het online aanjagen van (kleinschalig) geweld en opruiing zijn beperkt. Het Openbaar Ministerie registreerde in 2023 een stijging van discriminatiefeiten met een antisemitisch karakter, deels gerelateerd aan online uitingen tijdens voetbal gerelateerde incidenten en projecties op gebouwen die maatschappelijke onrust veroorzaakten. Een recent voorbeeld hiervan zijn de Maccabi in Amsterdam in November 2024 (Politie grijpt in na wedstrijd Ajax-Maccabi Tel Aviv in Amsterdam; Opnieuw celstraffen geëist tegen verdachten na geweld rond Ajax-Maccabi). Exacte cijfers over welk deel van de discriminatie specifiek leidde tot voetbalgerelateerde incidenten zijn niet bekend (OM: Geregistreerde discriminatiefeiten in 2023 gestegen | Nieuwsbericht | Openbaar Ministerie).

### 3.4. Online geweld onder jongeren

Aangezien over de aard en omvang van bovenstaande vormen van online geweld helaas nog niet veel data beschikbaar is, gaan we hier nog even kort in op de omvang van online geweld onder jongeren. Het Verwey-Jonker Instituut heeft een kwantitatieve verkennende analyse uitgevoerd naar online geweld onder jongeren [Online Geweld \\_Liggend](#) (2025). Hieruit komt naar voren dat 12,9% van de jongeren wel eens één of meerdere keren onderstaande online delicten heeft gepleegd:

- Versturen, delen of reposten van een intieme foto of video van iemand tegen diens wil in.
- Versturen van kwetsende berichten of reacties op sociale media over iemands etnische achtergrond, nationaliteit, religie, genderidentiteit, seksuele voorkeur, of een andere soortgelijke reden.
- Gebruiken van internet, e-mail of sociale media om anderen te bedreigen en/of te misleiden om geld te verdienen.
- Hacken van een computer of account om data te verkrijgen, dan wel te vernietigen.

Verder komt naar voren dat 26,2% van de jongeren wel eens één of meerdere keren van bovenstaande delicten slachtoffer is geweest.

Uit ander onderzoek onder online geweld onder jongeren van de Avans Hogeschool blijkt dat de meeste respondenten (74%) gemiddeld tussen de 2 en 4 uur per dag doorbrengen op sociale media ([Wat zijn behoeften van jongeren die te maken hebben met online geweld?](#)). Bij 18% loopt dit op tot meer dan 5 uur per dag. Whatsapp is het populairste medium (98% maakt hier gebruik van). Opvallend is dat bijna de helft wel eens ongewenst seksueel getinte beelden heeft ontvangen (46%). Ook online pesten (25%), online bedreiging en intimidatie (21%) en online stalking (21%) raakt een aanzienlijk deel van hen. Slachtoffers van het ontvangen van seksueel getint beeldmateriaal zijn vaker vrouw (84%) dan man (16%). Ook van de personen die aangeven online stalking te hebben meegemaakt, is 80% vrouw. Daar staat tegenover dat personen die te maken hebben met (bedreiging m.b.t.) ongewenste verspreiding van naaktfoto's of -filmpjes vaker man (63%) zijn. In de meeste situaties vond het geweld enkel online plaats, met pesten als uitzondering: hierbij was er in 47% van de gevallen ook sprake van offline geweld. 44% is wel eens getuige geweest van online geweld of is daarover door een ander in vertrouwen genomen.

Eerdere cijfers tonen aan dat bijna één op de vijf jongeren tussen de 15 en 25 jaar slachtoffer wordt van online criminaliteit. Dit is anderhalf keer zo vaak als bij andere leeftijdsgroepen. Vooral hacken, online bedreiging en intimidatie komen vaker voor bij jongeren. Ondanks een algemene daling van het aantal slachtoffers van online criminaliteit is onder jongeren een toename te zien in het aantal gevallen van online bedreiging en intimidatie.<sup>[5]</sup>



5 In Rosielle e.a., 2024: M. Akkermans e.a., Online veiligheid en criminaliteit 2022, Den Haag: Centraal Bureau voor de Statistiek 2023; M. Akkermans e.a., Veiligheidsmonitor 2023, Den Haag: Centraal Bureau voor de Statistiek 2024; S. Stoker, 'Trends in online criminaliteit onder jongeren: een vergelijking tussen 2021 en 2023', Wegwijzer Jeugd en Veiligheid 2024, [wegwijzerjeugdenveiligheid.nl/actueel/trends-inonline-criminaliteit-onder-jongeren-een-vergelijking-tussen-2021-en-2023/#:~:text=Ondanks%20een%20algehele%20daling%20in,en%20intimidatie%20en%20bij%20hacken](https://wegwijzerjeugdenveiligheid.nl/actueel/trends-inonline-criminaliteit-onder-jongeren-een-vergelijking-tussen-2021-en-2023/#:~:text=Ondanks%20een%20algehele%20daling%20in,en%20intimidatie%20en%20bij%20hacken); A. van Welzen & F. Pessers, Offlimits Onderzoek 'Onderzoek Online Lastigvallen', 2023, [https://offlimits.nl/assets/downloadable\\_files/onderzoek\\_online\\_lastigvallen.pdf](https://offlimits.nl/assets/downloadable_files/onderzoek_online_lastigvallen.pdf).

## 4 Belangrijke beïnvloedende factoren

De oorzaken van online geweld zijn complex en meervoudig. De anonimiteit van internetgebruikers maakt het gemakkelijker voor mensen om gewelddadig gedrag te vertonen wat de impact van het geweld vergroot. Dit kan tot langdurige psychologische gevolgen voor slachtoffers leiden, zoals gevoelens van angst, depressie en sociale isolatie. Naast anonimiteit draagt een gebrek aan regulering bij aan de toename van online geweld. Veel daders voelen zich onaantastbaar en gaan online verder dan ze in een face-to-face situatie zouden durven (Rathenau 2021; Greijdanus et al 2023; Mukred et al 2024).

De normalisering van agressief gedrag in online omgevingen voedt bovendien een bredere cultuur van polarisatie en extremisme. Online geweld wordt daarbij niet los gezien van bredere maatschappelijke ontwikkelingen, maar eerder gevoed en versterkt door de context waarin mensen zich bevinden. Zowel nationale als lokale kwesties kunnen spanningen in zowel de offline als online wereld aanwakkeren en leiden tot verhitte discussies waarin agressieve en extremistische uitingen toenemen. Bovendien is er sprake van een wisselwerking: maatschappelijke polarisatie kan online geweld uitlokken, terwijl online geweld op zijn beurt de spanningen in de offline wereld verder kan laten escaleren. Polarisation speelt hierin een sleutelrol (KIS 2022; Rathenau 2021; Hietanen & Eddebo; Alorainy et al 2018; Castano-Pulgarin et al 2021). Dit proces, waarbij tegenstellingen tussen groepen verscherpen en kunnen leiden tot sociale spanningen en segregatie, is niet per definitie negatief. In een democratische samenleving kan polarisatie bijdragen aan sociale verandering en politieke bewustwording. Echter, wanneer polarisatie gepaard gaat met negatieve stereotypen, angst en vijandigheid, kan dit de basis leggen voor online en offline conflicten. In Nederland spelen verschillende polariserende kwesties een rol in de dynamiek van online geweld. Discussies over bijvoorbeeld migratie, klimaatverandering, het coronavirus en culturele controverses, zoals het Zwarte Piet-debat, hebben geleid tot verhitte online interacties waarin haatdragende taal, bedreigingen en desinformatie floreren. Daarnaast leidt het groeiende wantrouwen in de overheid tot een vorm van verticale polarisatie, waarbij groepen burgers zich niet alleen tegen elkaar, maar ook tegen overheidsinstanties keren. Het Rathenau-rapport Online Ontspoord (2021) benadrukt dat sociale media een belangrijke rol spelen in deze dynamiek. Algoritmes versterken polariserende content en creëren 'echokamers' waarin gebruikers vooral worden blootgesteld aan informatie die hun bestaande opvattingen bevestigt, wat radicalisering in de hand kan werken (Rathenau 2021; Pérez-Escolar & Noguera-Vivo (2021); Mukred et al; Movisie; KIS 2022; Verwey Jonker Instituut 2023). Tabel 2 vat de onderliggende factoren die in de (inter)nationale literatuur omschreven zijn samen.

Tabel 2. Samenvatting onderliggende factoren

<b>Anonimiteit en gedeelde verantwoordelijkheid</b>	Online anonimiteit verlaagt de drempel voor agressief gedrag. Mensen voelen zich minder verantwoordelijk voor hun daden en uiten sneller haatdragende of gewelddadige uitspraken. In online gemeenschappen wordt verantwoordelijkheid vaak verspreid over een groep, waardoor individuen zich minder bezwaard voelen om zich gewelddadig te uiten.
<b>Maatschappelijke en politieke context</b>	Online geweld wordt versterkt door maatschappelijke spanningen, zoals migratie, klimaatbeleid of pandemieën. Polarisation in het offline domein kan zich vertalen naar online haat en vice versa. Overheidswantrouwen en institutionele polarisation kunnen ook bijdragen aan online agressie, waarbij groepen burgers zich niet alleen tegen elkaar maar ook tegen de overheid keren.
<b>Polarisation en echokamers</b>	Online platformen versterken polarisation door gebruikers bloot te stellen aan eenzijdige informatie binnen 'gated communities'. Dit leidt tot versterking van groepsidentiteit en vijandigheid tegenover andersdenkenden. Polarisation vergroot de kans op online geweld doordat tegenstellingen worden uitvergroot en extremistische denkbeelden sneller wortelschieten.
<b>Algoritmische versterking en de rol van sociale media</b>	Sociale media-algoritmen geven voorrang aan sensationele en controversiële content, omdat deze meer interactie genereren. Dit moedigt de verspreiding van haatzaaiende berichten en desinformatie aan. Gebruikers blijven hierdoor in een vicieuze cirkel waarin ze steeds extremere inhoud consumeren en delen, wat online agressie stimuleert.
<b>Misleidende retoriek en misinformatie</b>	De manier waarop bepaalde groepen of individuen worden geframed in online discussies kan bijdragen aan haat en geweld. Misleidende metaforen en vijandige taalgebruik dehumaniseren doelwitten, waardoor agressief gedrag tegen hen wordt genormaliseerd.

### 4.1. Enkele verklarende theorieën

De literatuur heeft meerdere verklarende theorieën aangedragen voor deze onderliggende factoren. Een van de belangrijkste theorieën hiervan is de "**Online Disinhibition Effect**". Deze theorie beschrijft hoe mensen zich online anders gedragen dan in face-to-face interacties. Dit fenomeen, zoals beschreven door o.a. Suler et al (2004), verklaart waarom individuen online sneller geneigd zijn tot impulsief, agressief of extreem gedrag. De afwezigheid van fysieke aanwezigheid, anonimiteit en het asynchrone karakter van online communicatie zorgen ervoor dat mensen minder remmingen ervaren en sneller grenzen overschrijden. Waar deze ontremming soms leidt tot positieve uitkomsten, zoals openhartige zelfonthulling, kan het ook een negatieve vorm aannemen, waarbij agressie en vijandigheid worden versterkt. De invloed van deze ontremming is duidelijk zichtbaar in online geweld en het verspreiden van haatdragende en discriminerende berichten.



Onderzoek laat zien dat de mate waarin iemand zich online ontremd voelt, samenhangt met de kans om zelf agressief gedrag te vertonen. Vooral in digitale omgevingen waarin haatdragende taal genormaliseerd wordt, voelen gebruikers zich aangemoedigd om dit gedrag te kopiëren. Mensen met een hoog niveau van empathie blijken hier beter tegen bestand, terwijl personen met een sterke neiging tot negatieve ontremming eerder geneigd zijn om zich gewelddadig te uiten. Dit onderstreept hoe digitale interacties niet alleen worden gevormd door technologie en sociale context, maar ook door psychologische factoren zoals impulscontrole en inlevingsvermogen. De wisselwerking tussen online ontremming, polarisatie en sociale dynamieken maakt duidelijk hoe schadelijk gedrag zich in digitale ruimtes kan verspreiden en versterken. Online platformen, die anonimiteit en afstand creëren, verlagen de drempel voor het uiten van haat en agressie, waardoor mensen zich minder verantwoordelijk voelen voor de impact van hun woorden. Dit verklaart waarom groepen die al kwetsbaar zijn disproportioneel doelwit worden. De normalisering van online agressie en de rol van algoritmisch versterkte polarisatie dragen verder bij aan een klimaat waarin digitale vijandigheid niet alleen wordt toegestaan, maar soms zelfs wordt beloond (Suler 2004; Joinson 2007; Hollenbaugh et al 2013).

De **moral disengagement-theorie** beschrijft hoe individuen morele principes tijdelijk kunnen uitschakelen om gedrag te rechtvaardigen dat normaal gesproken als immoreel wordt beschouwd. In de context van cyberpesten en online agressie speelt deze theorie een cruciale rol. Volgens Runions en Bak (2015) kunnen kenmerken van online omgevingen, zoals anonimiteit en fysieke afstand, deactivering van morele remmingen vergemakkelijken, waardoor gebruikers eerder geneigd zijn tot cyberagressie en cyberpesten. Wright en Wachs (2019) onderzochten specifiek de relatie tussen empathie, toxische online ontremming en homofobe cyberpesterijen. Hun studie toonde aan dat een gebrek aan empathie en een hoge mate van online ontremming de kans vergroten dat iemand zich schuldig maakt aan homofobe cyberpesterijen. Dit suggereert dat morele uitschakeling, gecombineerd met verminderde empathie en verhoogde online ontremming, een belangrijke factor is in het begrijpen van dergelijk gedrag. Deze theorie benadrukt dat online omgevingen unieke omstandigheden creëren die morele uitschakeling faciliteren wat kan leiden tot een toename van online geweld.

**De situationele actie theorie (Wikström, 1998; 2006; 2012; 2020).** Crimineel gedrag komt voort uit de interactie tussen persoonlijke factoren van een individu en omgevingsfactoren waarin dit individu zich bevindt zoals blootstelling, morele normen en waarden en zelfcontrole. Uit ons eerste onderdeel van dit onderzoek, Online geweld onder jongeren ([Onderzoek naar strafbare vormen van online geweld - Verwey-Jonker Instituut](#)), komt naar voren dat er zowel overeenkomsten als verschillen zijn in online en offline delicten bij jongeren. Zo blijkt uit deze kwantitatieve

studie dat online delicten procentueel vaker voorkomen dan offline delicten. Vooral meisjes zijn hiervan het slachtoffer. Er zijn ook verschillen in de achtergrondkenmerken van plegers en slachtoffers van online en offline delicten. Offline jeugdcriminaliteit neemt in het algemeen af naarmate jongeren ouder worden en een hoger opleidingsniveau hebben. Als het gaat om het plegen van online delicten dan zien we in onze eerste verkenning dat deze vorm van online geweld inderdaad afneemt als jongeren ouder worden, maar plegers van online delicten komen voor op alle niveaus van onderwijs. Gekeken naar slachtofferschap, zien we dat slachtoffers van online delicten nauwelijks verschillen wat betreft leeftijd en dat het voorkomt op alle niveaus van het onderwijs. Een ander opmerkelijk verschil is dat er minder sprake is (geen sterk verband) van overlap tussen slachtoffer- en plegerschap als het gaat om online delicten dan offline; terwijl andere onderzoeken juist wel overlap tussen daderschap en slachtofferschap tonen (*zie hoofdstuk 5*). De rol van sociaaleconomische status lijkt voor online delicten wel vergelijkbaar als voor offline delicten. De *situationele actietheorie*, die vaak wordt gebruikt om crimineel gedrag van jongeren te verklaren, lijkt minder goed toepasbaar op online delicten. Dit onderzoek betreft een eerste verkenning en meer onderzoek is nodig om te achterhalen welke (aspecten van) theoretische modellen en factoren beter verklaren hoe en waarom jongeren online geweld plegen en ervaren.

# 5 Kenmerken van burgers betrokken bij online geweld

Uit de literatuur blijkt dat er nog weinig bekend is over de (achtergrond)kenmerken van burgers die betrokken zijn bij online geweld. Ook hier geldt dat de verschillende manieren waarop online geweld wordt gedefinieerd en gemeten, het lastig maken om een compleet en consistent beeld te krijgen. Veel onderzoeken gaan er nog steeds vanuit dat slachtoffers alleen slachtoffers zijn en daders alleen daders, terwijl de werkelijkheid vaak complexer is. De dynamiek tussen daders en slachtoffers is namelijk fluïde, vooral in de context van digitale agressie en online haatspraak.

Onderzoek laat zien dat individuen die slachtoffer zijn geweest van online agressie, later zelf daders kunnen worden (Mukred et al., 2024; Naranjo-Pou et al., 2025). Dit wijst op een cyclus waarin de rollen van slachtoffer en dader niet strikt gescheiden zijn, maar in de loop van de tijd kunnen verschuiven. Persoonlijke ervaringen met slachtofferschap kunnen leiden tot agressief gedrag, waardoor slachtoffers uiteindelijk zelf daders worden in een poging om controle te herwinnen of pijn te verwerken. De wisselende rol tussen dader en slachtoffer onderstreept het belang van het loslaten van de traditionele scheiding tussen deze twee categorieën. Het inzicht dat slachtoffers ook daders kunnen worden en omgekeerd, zoals blijkt uit meerdere studies, maakt duidelijk dat begrip van deze fluiditeit kan bijdragen aan effectievere strategieën om online agressie en haatspraak te voorkomen en te verminderen (Mukred et al., 2024; Obermaier et al., 2024). Dit suggereert ook dat persoonlijke ervaringen met slachtofferschap een risicofactor kunnen zijn voor het plegen van online geweld (Mukred et al., 2024).



## Kenmerken van 'slachtoffers'

Slachtoffers van digitale agressie en online haatspraak vertonen vaak specifieke kenmerken die hen kwetsbaarder maken voor dergelijke negatieve ervaringen. Deze kenmerken kunnen grofweg worden onderverdeeld in demografische en psychosociale factoren.

Demografische factoren zoals leeftijd, geslacht en etniciteit spelen een grote rol bij het risico om slachtoffer te worden van online geweld. Geslacht is daarbij een van de meest bepalende factoren. Vrouwen lopen bijvoorbeeld een groter risico op online stalking, seksuele intimidatie en bedreigingen met seksueel geweld. Onderzoek laat zien dat vrouwen vaker het doelwit zijn van ongewenste berichten, verspreiding van expliciete beelden zonder toestemming en bedreigingen met fysiek of seksueel geweld. Dit patroon wijst op gendergerelateerd geweld, waarbij de dader vaak probeert controle uit te oefenen of het slachtoffer te intimideren (Castellanos et al., 2023; [Online geweld onder jongeren: een eerste verkenning - Verwey-Jonker Instituut](#)). Jongens daarentegen worden vaker geconfronteerd met fysieke bedreigingen en vernedering, wat duidt op een andere dynamiek in de ervaring van online geweld (Castellanos et al., 2023).

Leeftijd speelt eveneens een belangrijke rol. Jongeren zijn vaker het doelwit van cyberpesten, haatspraak, het ongewenst delen van intieme beelden (*sexting*), en manipulatie via sociale media. Jongeren en jongvolwassenen bevinden zich in een levensfase waarin sociale interactie en acceptatie cruciaal zijn, waardoor online uitsluiting en vernedering een sterke psychologische impact kunnen hebben (Castellanos et al., 2023; [Online geweld onder jongeren: een eerste verkenning - Verwey-Jonker Instituut](#)).

Ook etniciteit en seksuele geaardheid zijn belangrijke voorspellers van online geweld. Jongeren met een migratieachtergrond zijn vaker slachtoffer van racistische haatspraak en discriminerende uitlatingen (Castellanos et al., 2023). lhbti+-personen worden daarentegen vaker geconfronteerd met homofobe of transfobe bedreigingen en uitsluiting. De anonimiteit van het internet vergroot het gemak waarmee daders zich schuldig maken aan dergelijke vormen van haatspraak, omdat ze minder snel verantwoordelijk worden gehouden voor hun acties (Mukred et al., 2024).

Sociaaleconomische status speelt eveneens een rol. Personen met een lagere sociaaleconomische status lopen meer risico op financieel gemotiveerde vormen van online geweld, zoals fraude en online oplichting. Een gebrek aan digitale geletterdheid of beperkte toegang tot hulpbronnen maakt deze groep extra kwetsbaar voor manipulatie en uitbuiting ('t Hoff-de Goede & Leukfeldt; [Online geweld onder jongeren: een eerste verkenning - Verwey-Jonker Instituut](#)).

Naast demografische factoren spelen psychosociale kenmerken een grote rol in de kwetsbaarheid voor digitale agressie. Jongeren met een laag zelfbeeld of beperkte sociale vaardigheden hebben een verhoogd risico om slachtoffer te worden (Mukred et al., 2024). Onzekerheid over de eigen identiteit en een gebrek aan sociale weerbaarheid maken het moeilijker om effectief te reageren op online aanvallen. Dit kan slachtoffers in een negatieve spiraal brengen, waarbij ze steeds kwetsbaarder worden voor herhaald slachtofferschap (Mukred et al., 2024). Ook lijkt het erop dat hoe vaker mensen zich in de online wereld begeven, hoe hoger de kans op slachtofferschap van online geweld ([Wat zijn behoeften van jongeren die te maken hebben met online geweld?](#)).

Sociale isolatie en beperkte steun vanuit het sociale netwerk vergroten eveneens de kans op slachtofferschap. Jongeren die zich eenzaam voelen of weinig sociale steun ervaren, hebben minder mogelijkheden om zich te verdedigen tegen online aanvallen en zijn daardoor een gemakkelijker doelwit voor digitale agressie. Deze combinatie van persoonlijke onzekerheid en een gebrek aan sociale steun maakt slachtoffers vatbaarder voor herhaaldelijke agressie en kan leiden tot langdurige psychologische schade (Mukred et al., 2024).

Het verband tussen demografische en psychosociale kenmerken en online geweld laat zien dat slachtofferschap geen willekeurig proces is, maar deels voortkomt uit bredere sociale en structurele ongelijkheden. Door deze dynamieken beter te begrijpen, kunnen gerichte preventiestrategieën en passende hulpverlening worden ontwikkeld om slachtoffers te beschermen en online agressie terug te dringen.

### **Kenmerken van 'daders'**

De kenmerken van daders kunnen grofweg onderverdeeld worden in demografische en psychosociale factoren. Hoewel veel onderzoek zich richt op jongeren, blijkt uit recente studies dat online geweld en agressie ook onder ouderen voorkomt. Bij 66-plussers gaf 22,7% toe recent online agressief gedrag te hebben vertoond, terwijl 26,8% slachtoffer was (Pabian & Vandenbosch, 2024). Deze agressie wordt vaak verklaard door overtuigingen zoals victim blaming en de perceptie dat online gedrag minder ernstige gevolgen heeft. Bij jongvolwassenen (18-25 jaar) liggen de percentages van daderschap en slachtofferschap echter beduidend hoger (60,1% en 54,6%) (Pabian & Vandenbosch, 2024). Dit onderstreept dat daderschap niet exclusief aan jongeren is verbonden, en dat verschillende leeftijdsgroepen met uiteenlopende motieven en overtuigingen online agressie kunnen vertonen. Tegelijkertijd is het nog onvoldoende duidelijk of jongeren daadwerkelijk vaker online geweld plegen, of dat zij vooral vaker worden onderzocht. Meer inzicht is nodig in hoe generatieverschillen samenhangen met de aard van online geweld en het gebruik van specifieke platformen.

Psychologische factoren spelen een belangrijke rol bij daderschap. Daders van cybergeweld ervaren vaak verhoogde niveaus van stress, depressie en angst vergeleken met niet-betrokken leeftijdsgenoten (Sincek et al., 2017). Deze psychologische belasting kan bijdragen aan het ontstaan of voortduren van agressief gedrag online, waarbij gevoelens van machteloosheid of frustratie zich uiten in digitaal geweld. Ook een negatieve thuissituatie, zoals conflicten met ouders en een gebrek aan ouderlijke steun, verhoogt het risico op dadergedrag (Elsaesser et al., 2017). Jongeren die zich thuis onbegrepen of genegeerd voelen, zoeken soms via online agressie naar controle of erkenning.

Naast psychologische factoren spelen sociale kenmerken een rol. Jongeren die zich sociaal geïsoleerd voelen of weinig sociale steun ontvangen van ouders en leeftijdsgenoten, hebben een verhoogde kans om zich agressief te uiten op digitale platforms (Wright & Wachs, 2019). Het ontbreken van sociale steun kan ertoe leiden dat deze jongeren hun frustraties afreageren via online haatspraak of pesten. Het gevoel van anonimiteit op internet speelt hierbij een versterkende rol; daders zijn minder bang voor directe consequenties en ervaren minder empathie voor hun slachtoffers (Barlett et al., 2019). Er lijkt bovendien een verband tussen sociale dominantie en het plegen van online haatspraak (Castellanos et al., 2024). Personen met een hoge sociale dominantieoriëntatie, gekenmerkt door een streven naar superioriteit over anderen, vertonen vaker dergelijk gedrag. Dit wordt mede verklaard door een verminderd empathisch vermogen en een verhoogde morele ontkoppeling, waardoor zij minder geneigd zijn de negatieve impact van hun acties te erkennen (Pabian et al., 2015). Bovendien ontwikkelen deze daders vaker positieve attitudes ten opzichte van geweld, wat hun bereidheid om zich agressief te gedragen in een online context vergroot (Wright, 2014).

De kenmerken van daders van online agressie zijn divers en worden gevormd door een combinatie van psychologische, sociale en demografische factoren. Naast jongeren blijken ook ouderen zich schuldig te maken aan online agressie, met andere drijfveren en overtuigingen als verklaring. Deze inzichten benadrukken het belang van een brede benadering in preventie en interventie, waarbij rekening wordt gehouden met zowel leeftijdsspecifieke kenmerken als onderliggende psychologische en sociale factoren.

# 6 Impact van online geweld op individuen en de maatschappij

Online geweld heeft ingrijpende gevolgen voor zowel individuen als de samenleving, waarbij de grens tussen online en offline effecten vaak vervaagt. Slachtoffers ervaren psychologische schade, sociale isolatie en gezondheidsproblemen, terwijl de maatschappelijke effecten zich uiten in polarisatie, ondermijning van het publieke debat en verlies van vertrouwen in democratische instituties. De literatuur toont aan dat online geweld niet alleen schadelijk is voor directe slachtoffers, maar ook bijdraagt aan maatschappelijke spanningen en het functioneren van de democratie kan ondermijnen. Online geweld treft niet alleen individuele slachtoffers maar verhardt ook het publieke debat en draagt bij aan een klimaat van angst en onveiligheid, vooral binnen gemarginaliseerde groepen. In dit hoofdstuk zullen we eerst de impact op individueel niveau toelichten en daarna ingaan op de maatschappelijke gevolgen.

## 6.1. Impact op individueel niveau

De gevolgen van online geweld zijn voor individuen op psychologisch en sociaal vlak ingrijpend. Slachtoffers van online geweld ervaren vaak gevoelens van angst, stress, machteloosheid en een verlies van controle over hun eigen leven. Volgens Sineck et al (2017) zijn angststoornissen, depressie en posttraumatische stress veelvoorkomende gevolgen van online geweld, vooral wanneer het geweld langdurig aanhoudt, vooral door de vaak permanente zichtbaarheid van online berichten kan het gevoel bij slachtoffers ontstaan dat het geweld altijd door blijft gaan. Deze psychologische schade kan ook fysieke gevolgen hebben. Bijvoorbeeld, chronische stress veroorzaakt door online intimidatie kan leiden tot slaapstoornissen en verhoogde bloeddruk en andere fysieke gezondheidsproblemen (Sineck et al 2017; [Wat zijn behoeften van jongeren die te maken hebben met online geweld?](#)). Online geweld heeft dus tastbare effecten op het dagelijkse functioneren van slachtoffers.

Daarnaast toont de literatuur aan dat online geweld ook sociale gevolgen heeft voor slachtoffers. Leukfeldt et al (2018) benadrukt, bijvoorbeeld, dat online geweld het gevoel van veiligheid van de slachtoffers aantast. Slachtoffers kunnen vertrouwen verliezen in de digitale platforms en de online gemeenschappen waar ze eerst deel van uitmaakten. Het rapport van Rathenau (2021) onderstreept deze bevindingen en geeft aan dat mensen voorzichtiger worden met het uitingen van hun

meningen en het deelnemen aan het publieke debat uit angst voor vergelding. Voor sommigen kan dit tot een gevoel van sociale isolatie leiden. Dit geldt vooral voor bepaalde groepen zoals vrouwen, mensen van kleur, lhbti+ personen die onevenredig vaak het doelwit zijn van online geweld. Deze groepen ervaren vaak intersectionele vormen van geweld, waarbij verschillende vormen van discriminatie samenkomen en de impact van het geweld verergeren. Door het gevoel van onveiligheid dat hierdoor ontstaat, trekken ze zich vaker terug en beperken ze hun deelname aan online en offline sociale en politieke activiteiten (Castaño-Pulgarin et al., 2021).

## 6.2. Impact op maatschappelijk niveau

Online geweld raakt niet alleen individuele burgers, maar ondermijnt ook de sociale cohesie, het publieke debat en het vertrouwen in democratische processen. Het Rathenau Instituut (2021) wijst op het destabiliserende effect van online haat en intimidatie op het politieke klimaat. Niet alleen politici, journalisten en publieke figuren, maar ook gewone burgers worden geconfronteerd met online bedreigingen, haat en discriminatie. Dit leidt vaak tot zelfcensuur en het terugtrekken uit het publieke debat, waardoor bepaalde stemmen minder worden gehoord en het politieke discours verschuift naar extremere en populistische standpunten. Onderzoek van Hietanen & Eddebo (2022) bevestigt dit beeld en benadrukt dat haatspraak tegen minderheden en kwetsbare groepen leidt tot een versmalling van het publieke debat. Wanneer bepaalde stemmen worden onderdrukt door systematische intimidatie, ontstaat een vertekend maatschappelijk beeld waarin haatdragende en extremistische standpunten disproportioneel veel ruimte krijgen. Alorainy et al. (2018) laten zien hoe bedreigende en uitsluitende taal het 'wij-zij'-denken versterkt. Door groepen af te schilderen als vijanden of buitenstaanders wordt sociale polarisatie vergroot, wat leidt tot een diepere scheiding tussen verschillende maatschappelijke groepen. Sociale media-algoritmes spelen hierin een versterkende rol. Castaño-Pulgarin et al. (2021) wijzen erop dat platforms zoals Facebook, X en YouTube geneigd zijn polariserende en emotioneel geladen content te belonen met meer zichtbaarheid. Dit zorgt ervoor dat negatieve en haatdragende berichten sneller worden verspreid, wat bijdraagt aan de normalisering van online geweld en een verdere erosie van vertrouwen in digitale platforms. Tenslotte draagt het gebrek aan effectieve handhaving bij aan het probleem. Leukfeldt et al. (2018) stellen dat slachtoffers van online criminaliteit en geweld zich vaak niet gehoord voelen door politie en overheidsinstanties. De handhaving van digitale misdrijven verloopt complex en traag, waardoor daders zich vaak onaantastbaar voelen en slachtoffers het vertrouwen in instituties en het rechtssysteem verliezen.

### 6.3. Impact specifiek voor jongeren

Online geweld bij jongeren kan ingrijpende gevolgen hebben. Uit het onderzoek [Wat zijn behoeften van jongeren die te maken hebben met online geweld?](#) blijkt dat jongeren die online bedreiging en intimidatie meemaakten zich minder veilig voelden (56%), minder vertrouwen hebben in anderen (48%), slaapproblemen ervaren (32%) en depressieve en/of stress gerelateerde klachten (32%). Ook had een deel van hen last van schaamte (28%), zelfverwijt en schuldgevoelens (24%). Bijna een kwart rapporteerde angst- en/of paniekaanvallen (24%). Ook voelde 20% zich uitgesloten of buitengesloten. Zij hebben verschillende hulpbehoeften. Een deel van de jongeren die online bedreiging en intimidatie meemaakten had na het voorval geen behoefte aan informatie of hulp (32%), een ander deel gaf aan zich weer veilig te willen voelen (32%), te willen dat het geweld stopt (28%) of de behoefte te hebben erover te praten (20%), rust te willen en beter te willen slapen (16%) of informatie te willen ontvangen over het doen van aangifte (16%).



# 7 Mogelijke interventies en beleidsmaatregelen

Vanwege de complexiteit van strafbaar online geweld, het meerdere vormen aanneemt en het feit dat het beïnvloed wordt door diverse factoren op meerdere niveaus (van individueel tot maatschappelijk), is het noodzakelijk een breed scala aan interventies en beleidsmaatregelen in te zetten om deze problematiek aan te pakken. In dit hoofdstuk zoomen we in op interventies op verschillende niveaus:

- a. Individueel niveau
- b. Gemeenschapsniveau
- c. Beleids- en wetgevingsniveau
- d. Technische oplossingen

Ook introduceren we enkele interventies en maatregelen die in de (inter)nationale literatuur zijn beschreven en geven de bijbehorende randvoorwaarden weer.

## 7.1. Interventies gericht op individueel niveau

Hiermee bedoelen we interventies die gericht zijn op het beïnvloeden van de kennis, het gedrag en de reacties van individuen. Deze interventies beogen vooral de bewustwording en digitale weerbaarheid van individuen te vergroten.

### ***Innocent bystander interventies***

Het gedrag van omstanders speelt een belangrijke rol in het voortbestaan van online geweld. Obermaier (2024) wijst erop dat jongeren en jongvolwassenen vaak terughoudend zijn om in te grijpen bij online haatzaaien, zelfs wanneer ze zich bewust zijn van de schadelijke effecten. Deze passiviteit wordt veroorzaakt door verschillende factoren, zoals angst voor vergelding, het gevoel dat hun tussenkomst weinig effect zal hebben, en de sociale norm om zich niet te mengen in online conflicten. Hierdoor blijven haatdragende berichten en online intimidatie ongestoord doorgaan, wat het probleem verder verergert. Wanneer omstanders niet ingrijpen, krijgen daders impliciet toestemming om door te gaan met hun gedrag, wat bijdraagt aan de normalisering van online geweld en het gevoel van straffeloosheid onder daders versterkt.

Zij stellen voor dat interventies worden geïmplementeerd om omstanders te motiveren om actief en direct in te grijpen bij online haatzaaien. Deze interventies zouden het bewustzijn over de negatieve impact van online geweld moeten vergroten, het versterken van de zelfverzekerdheid van mensen in hun vermogen om in te grijpen en het verminderen van angst voor vergelding. Ze stellen het ontwikkelen van online weerbaarheidstraining voor om de vaardigheden en het zelfvertrouwen van burgers om effectief om te gaan met online geweld voor. Zo'n training zou praktische handvatten moeten bieden voor het herkennen van schadelijke online gedragingen, succesvolle manieren om te reageren op online geweld en zouden digitale weerbaarheid moeten versterken. Het doel is om door middel van digitale weerbaarheidstraining en het aanmoedigen van peer-support, de sociale normen te veranderen en burgers actief te betrekken bij het tegengaan van online geweld.

Uit onderzoek naar de behoeften van online geweld onder jongeren komt naar voren dat ze uiteenlopende hulpbehoeften hebben, zoals zich weer veilig willen voelen of het geweld willen stoppen. De jongeren zijn weinig bekend met het hulpaanbod. 79% van de respondenten weet geen organisaties te noemen die informatie en/of hulp bieden aan slachtoffers van online geweld. Met name na het meemaken van online discriminatie, online stalking en de (dreiging van) verspreiding van naaktfoto's of filmpjes zonder toestemming gaat meer dan 80% ook niet op zoek naar informatie. Bij het ontvangen van ongewenste seksueel getinte afbeeldingen betreft dit zelfs 91%, bij online bedreiging 72%. Opvallend is dat bij online pesten 40% wél op zoek gaat naar informatie. Respondenten geven voor het al dan niet zoeken verschillende redenen aan. Wanneer dit gekoppeld wordt aan de gevolgen van het online geweld, valt op dat met name deelnemers die schaamte ervaren vaak niet op zoek gaan naar informatie (76%), terwijl zij dit minder vaak zelf aangeven als reden waarom ze geen informatie hebben gezocht. Als jongeren hulp zochten en ontvingen, kwam dit vaak uit informele bronnen (familie of vrienden) of iemand op school. Over de hulp van familie is 80% tevreden of zeer tevreden. Ook over de hulp van vrienden is 82% tevreden of heel tevreden. Over hulp van iemand op school is 50% tevreden of heel tevreden, 25% neutraal, 19% ontevreden en 1 respondent (6%) heel ontevreden. Dit wijst op extra aandachtspunten om de geleverde hulp vanuit de school te verbeteren. Slechts zeven deelnemers konden inzicht geven in hun ervaringen met de verschillende hulporganisaties, vier gingen in op hun ervaringen met Slachtofferhulp Nederland.<sup>[6]</sup>

Jongeren geven aan dat het onderwerp het beste kan worden gepromoot via sociale media, colleges, video's, een website of posters om de toegang tot hulp te verbeteren.<sup>[7]</sup> Jongeren ontvangen graag meer informatie over online geweld en het hulpaanbod via social media (43%)

6 Rosielle e.a., 2024.

7 [Wat zijn behoeften van jongeren die te maken hebben met online geweld?](#).

en lessen op school (38%). Ook informatieve video's op sociale media (29%), een website over 'hulp online geweld' (28%) of een online platform met ervaringsverhalen (25%) bieden mogelijkheden. Een kwart geeft de voorkeur aan flyers of posters. Contact met hulporganisaties verloopt bij voorkeur via chat (61%) of Whatsapp (47%), terwijl bij het ontvangen van hulp een persoonlijk gesprek met een professional de voorkeur heeft (72%).

## 7.2. Interventies gericht op gemeenschapsniveau

### Versterken en toegankelijker maken van publiek debat

In een samenleving waarin burgers actief deelnemen aan het publieke debat, vormt online strafbaar geweld, zoals online haatspraak en intimidatie, een serieuze bedreiging voor sociale cohesie en democratische waarden. Pérez-Escolar en Noguera-Vivo (2022) benadrukken dat online haatspraak en polarisatie hand in hand gaan en het publieke debat verstoren door het creëren van vijandigheid en angst. Zoals omschreven in hoofdstuk 4 zorgt online geweld ervoor dat bepaalde groepen systematisch worden uitgesloten van het publieke debat. Online platforms versterken deze uitsluiting door algoritmische filters en echokamers, die polarisatie bevorderen en het bereik van haatdragende boodschappen vergroten. Dit leidt ertoe dat minderheden en kwetsbare groepen zich steeds vaker terugtrekken uit het debat, waardoor hun stem verloren gaat en het maatschappelijke draagvlak voor inclusieve beleidsmaatregelen afneemt. Ook Greijdanus et al. (2023) onderstrepen dat online sociale interacties kunnen escaleren tot fysiek geweld in de publieke ruimte, waarbij negatieve dynamieken zoals polarisatie en sociale uitsluiting een grote rol spelen. Zij wijzen erop dat online haatspraak en intimidatie niet alleen schadelijk zijn voor het individu, maar ook het vertrouwen in het maatschappelijke debat ondermijnen en het risico vergroten dat conflicten zich vertalen naar de fysieke wereld. Om deze negatieve consequenties van online geweld tegen te gaan, moet het publieke debat versterkt en toegankelijker worden, zodat diverse burgers zich veiliger en gehoord voelen in de online wereld.

Het versterken van het publieke debat vraagt om een combinatie van juridische, sociale en technologische maatregelen. Pérez-Escolar en Noguera-Vivo (2022) stellen dat het creëren van veilige digitale ruimtes, waar haatspraak actief wordt aangepakt en waar respectvolle discussie wordt gestimuleerd, essentieel is voor het creëren van een gevoel van veiligheid. Dit kan door andere programma's die de deelname van vertegenwoordigde groepen in het publieke debat stimuleren, effectieve wetgeving die online haatspraak strafbaar stelt en duidelijke richtlijnen biedt voor handhaving, algoritmische transparantie om te voorkomen dat polariserende en haatdragende content door automatische systemen wordt versterkt, en educatie en bewustwording over het herkennen en melden van haatspraak. Pérez-Escolar en Noguera-Vivo (2022) pleiten voor inclusieve online platforms waar diverse stemmen worden aangemoedigd en beschermd.

Dit kan worden bereikt door stimuleringsprogramma's voor deelname van ondervertegenwoordigde groepen in het publieke debat, technologische toegankelijkheid zoals betere ondertiteling, taalopties en gebruiksvriendelijke interfaces, en betere ondersteuning voor meldsystemen zodat slachtoffers van haatspraak snel en effectief geholpen worden.

Greijdanus et al. (2023) benadrukken dat het versterken van het publieke debat ook een rol speelt bij het voorkomen en verminderen van online aangejaagd geweld. Zij wijzen op het belang van samenwerking tussen verschillende netwerkpartners, zoals gemeenten, politie, jongerenwerkers, scholen en reclassering, om negatieve online dynamieken te doorbreken. Door deze actoren actief te betrekken bij het publieke debat, ontstaat een breder maatschappelijk draagvlak voor het bestrijden van online geweld en haatspraak. Wanneer deze partijen gezamenlijk optreden, kunnen zij zorgen voor een veilige en respectvolle digitale omgeving waarin meer ruimte ontstaat voor constructieve discussie en diversiteit van stemmen. Daarnaast benadrukken Greijdanus et al. (2023) het belang van preventieve interventies gericht op het creëren van een gezonde online cultuur. Dit houdt in dat er niet alleen wordt ingegrepen bij escalaties, maar dat ook wordt ingezet op het bevorderen van respectvolle communicatie en het versterken van mediawijsheid onder burgers. Door burgers bewuster te maken van de impact van online interacties en hen de vaardigheden te geven om haatspraak en misinformatie te herkennen en aan te pakken, wordt de weerbaarheid van het publieke debat vergroot. Verder stellen zij dat het verbeteren van de toegankelijkheid van online platforms een belangrijke rol speelt. Dit kan door duidelijke richtlijnen en moderatiebeleid, zodat gebruikers weten welke gedragingen acceptabel zijn en welke niet. Door transparantie en handhaving van deze richtlijnen voelen burgers zich veiliger om deel te nemen aan het debat, zonder angst voor intimidatie of uitsluiting.

Het aanpakken van online haatspraak en het versterken van het publieke debat vereist dus een gelaagde strategie waarin regulering, technologische hervormingen en maatschappelijke betrokkenheid samenkomen. Pérez-Escolar en Noguera-Vivo (2022) richten zich hierbij op het creëren van veilige en inclusieve digitale omgevingen, terwijl Greijdanus et al. (2023) het belang van samenwerking tussen maatschappelijke actoren, relatieopbouw met jongeren en preventieve interventies benadrukken. Door deze inzichten te combineren, kan het publieke debat worden versterkt en toegankelijker worden gemaakt, waardoor meer stemmen worden gehoord en sociale cohesie en democratische waarden worden beschermd.

### 7.3. Beleidsmaatregelen en wet- en regelgeving

Mukred et al. (2024) benadrukken dat het gebrek aan duidelijke regelgeving en de beperkte bereidheid van sociale mediaplatforms om schadelijke inhoud te reguleren het probleem verergert. Zonder gerichte maatregelen op het niveau van platformregulering en handhaving blijft online geweld een structureel probleem dat zowel de sociale samenhang als het vertrouwen in democratische processen onder druk zet.

#### Digital Services Act

Één van die wetten op Europees niveau is de Digital Services Act (DSA). De DSA moet ervoor zorgen dat het internet veiliger en transparanter wordt. Sinds november 2022 wordt de wet stap voor stap ingevoerd en vanaf februari 2024 geldt hij volledig. De DSA stelt nieuwe regels voor online platforms die in de EU actief zijn, zoals sociale mediabedrijven, webwinkels, zoekmachines en cloudopslagdiensten.

Een belangrijk onderdeel van de wet is dat online platforms snel moeten ingrijpen als er illegale inhoud, zoals haatzaaiende berichten, gewelddadige video's of andere verboden inhoud, op hun websites verschijnt. De DSA geeft geen nieuwe definitie van wat illegale inhoud is, maar verplicht platforms om bestaande wetten beter te handhaven.

Ook legt de DSA online platforms op om een duidelijk meldings- en actieproces te hebben. Dit houdt bijvoorbeeld in dat ze meldknoppen voor gebruikers aanbieden, dat ze vervolgens snel moeten reageren op deze meldingen en dat ze gebruikers informeren waarom een account is verwijderd. Ook moet de gebruiker de kans krijgen om hiertegen in beroep te gaan. Daarnaast moeten grote online platforms transparant zijn over hoe hun algoritmes werken en worden advertenties strenger gereguleerd. (Europese commissie, n.d.). Omdat de DSA geldt vanaf februari 2024 moeten de individuele landen nog wel uitwerken wat de DSA voor hen individueel betekent en hoe er beleidsmatig en binnen hun eigen wet- en regelgeving handen en voeten aangegeven kan worden ([DSA: verplichtingen voor en toezicht op allergrootste online platforms ingegaan | Nieuwsbericht | Rijksoverheid.nl](#)).

#### Wetgeving in het buitenland

Sommige landen hebben al specifieke wetten die gericht zijn op het bestrijden van online geweld. Zo is er in het Verenigd Koninkrijk de **Malicious Communications Act** uit 1988, die communicatie strafbaar stelt als deze bedoeld is om angst aan te jagen, schade te berokkenen of iemand lastig te vallen. Dit geldt voor berichten die grof beledigend, bedreigend of lasterlijk zijn en die via verschillende kanalen worden verzonden, zoals brieven, e-mails, sms'jes en sociale media (UK

Government, 1988). Sinds de invoering is het online landschap sterk veranderd, waardoor de wet is aangepast om ook digitale bedreigingen en online haatspraak te bestrijden. Daarnaast zijn er ook andere juridische kaders, zoals de recent ingevoerde **Online Safety Act** (UK Government, 2023). Deze wet zorgt voor extra maatregelen om gebruikers online te beschermen, vergelijkbaar met de DSA, maar dan gericht op het Verenigd Koninkrijk. Belangrijke verschillen zijn dat de Online Safety Act meer een focus heeft op het beschermen van kinderen en zich ook richt op schadelijke inhoud, terwijl de DSA zich beperkt tot illegale inhoud.

Duitsland heeft in 2017 de **Netzwerkdurchsetzungsgesetz** (Netwerkhandhavingswet; afgekort NetzDG). Deze wet werd ingevoerd om te strijd tegen illegale inhoud op sociale media te versterken. Het doel van de wet is om online platforms verplichten sneller op te treden tegen haatzaaiende berichten, extremisme en ander illegale content (Bundesregierung, 2022). De wet is vergelijkbaar met de DSA, alleen gericht op Duitsland, en was één van de eerste grote wetten die platforms aansprakelijk stelde voor de inhoud die zij hosten.

#### Leeftijdsgrens voor sociale media

In veel landen is er indirecte wetgeving die een minimumleeftijd vaststelt voor het gebruik van sociale mediaplatforms. In Europa wordt dit geregeld in de General Data Protection Regulation (GDPR), die bepaalt dat kinderen onder de 16 jaar alleen een sociale media-account aan mogen maken met ouderlijke toestemming (European Commission, 2018), waarbij sommige landen een leeftijd van 13 aanhouden. Desondanks kunnen techbedrijven zelf bepalen hoe dat wordt nageleefd. Livingstone et al. (2013) hebben het gebruik van sociale media onder minderjarigen in relatie tot het toenmalige Europese beleid onderzocht in 2013 en vonden dat maar liefst 38% van de kinderen tussen 9 en 12 jaar sociale media gebruikt. Dit wijst erop dat het beleid omtrent leeftijdsrestricties weinig effectief zijn en makkelijk worden omzeild, vaak met medeweten van of ondanks het toezicht van ouders.

Door sommige landen wordt een minimumleeftijd voor sociale media besproken. In Nederland heeft de Tweede Kamer recent een motie aangenomen die pleit voor een minimumleeftijd van 15 jaar voor het gebruik van sociale media (Tweede Kamer der Staten-Generaal, 2025). Algemene kritiek op deze maatregel is dat leeftijdsverificatie makkelijk te omzeilen is, door bijvoorbeeld een VPN te gebruiken of simpelweg een andere geboortedatum in te vullen. Desondanks adviseren Livingstone et al. wel om leeftijdsgrenzen te behouden. Ze pleitten onder andere ervoor om veiligheidsvoorzieningen van platforms toegankelijker te maken en ouderlijk toezicht te verbeteren. Meer kwalitatief onderzoek naar de wisselwerking tussen de wensen van kinderen, ouderlijk toezicht en technische mogelijkheden is daarvoor nodig.



## Handhaving

Overheidsinstanties spelen een essentiële rol in de handhaving van wet- en regelgeving met betrekking tot online geweld, zowel ten aanzien van burgers als techbedrijven. Echter tonen studies aan dat veel overheidsinstanties onvoldoende kennis, middelen en technische capaciteit hebben om effectief op te treden tegen online geweld (Bantema et al 2018; Bartelds et al 2023). Bantema et al. (2018) benadrukken dat de politie en andere overheidsinstanties onvoldoende zijn toegerust om effectief op te treden in de online wereld, met name door een gebrek aan technische middelen en specifieke kennis. Bartelds et al. (2023) wijzen op de beperkte capaciteit van handhavers om digitale geweldsdelicten adequaat te monitoren en af te handelen, wat vaak resulteert in het uitblijven van vervolging. De Vries en Bantema (2021) voegen hieraan toe dat het moeilijk is voor overheden om effectief te reageren op online bedreigingen, omdat er onvoldoende samenwerking en kennis is binnen de lokale en nationale overheidsstructuren. Dit tekort aan middelen en expertise maakt het moeilijk om de nodige stappen te zetten voor een efficiënte handhaving van wetten die online geweld moeten aanpakken. De Vries en Bantema (2021) benadrukken de noodzaak van lokale bestuurlijke instrumenten, zoals digitaal gebiedsverbod, om daders van online geweld effectief te beboeten en verantwoordelijk te houden. Bij gebrek aan een adequaat juridische structuur en middelen kunnen overheidsinstanties niet altijd snel reageren, waardoor slachtoffers van online geweld vaak machteloos blijven. Bovendien is de samenwerking met techbedrijven essentieel om een rechtvaardige en efficiënte handhaving te garanderen, aangezien platforms hun verantwoordelijkheid moeten nemen in het modereren van schadelijke content. Echter, de beperkte samenwerking en technologische competentie bij overheden maakt het moeilijk om effectief op te treden (Bantema et al., 2018; Bartelds et al., 2023). Tenslotte tonen Perez-Escolar & Noguera-Vivo (2021) het belang van verbeterde handhaving op de interactie tussen de online en offline wereld, gezien het de impact dat online haatzaaien en discriminatie hebben op de bredere, fysieke, maatschappij. Dit vereist niet alleen technische oplossingen, maar ook bestuurlijke en juridische aanpassingen. Bantema (2022) en Vuik & Bantema (2021) wijzen op de noodzaak van lokale bestuurlijke instrumenten, zoals digitaal gebiedsverbod, om aanvallers verantwoordelijk te houden voor hun online gedragingen.

Tenslotte maakt het internationale karakter van online geweld handhaving bijzonder complex. Het Rathenau Instituut (2021) wijst erop dat grensoverschrijdende jurisdictie de handhaving bemoeilijkt, doordat onrechtmatig gedrag vaak onbestraft blijft wanneer daders of techbedrijven buiten Nederland actief zijn. Dit gebrek aan effectieve handhaving draagt bij aan een gevoel van machteloosheid bij slachtoffers. Een sterkere samenwerking tussen overheden, zowel nationaal als internationaal, is essentieel om online geweld beter aan te pakken. Het ontwikkelen van grensoverschrijdende regelgeving kan hierbij helpen door de handhaving te versterken en de effectiviteit

van online geweldpreventie te vergroten. De rol van techbedrijven is daarbij van groot belang. Als beheerders van online platformen hebben zij de verantwoordelijkheid om schadelijke content te modereren en te verwijderen. De samenwerking tussen overheden en techbedrijven is cruciaal, aangezien techbedrijven vaak de tools en middelen hebben om schadelijke content snel te identificeren en te blokkeren. Echter, de onvoldoende regulering en inconsistentie in de naleving van wetten door techbedrijven maakt effectieve handhaving moeilijk. Internationale samenwerking en uniforme wetgeving zijn noodzakelijk om techbedrijven aansprakelijk te stellen en ervoor te zorgen dat zij daadwerkelijk hun verantwoordelijkheid nemen in het modereren van schadelijke inhoud.

## 7.4. Technologische interventies

Zoals hierboven duidelijk omschreven speelt online geweld af op het snijvlak van technologie, de samenleving en wetgeving, daarmee spelen techbedrijven een essentiële rol in het ontwikkelen en implementeren van technologische oplossingen om online geweld aan te pakken. Daarmee hebben techbedrijven ook een sleutelrol in het versterken van privaat-publieke samenwerkingen om online geweld structureel aan te pakken.

### **Privaat-publieke samenwerking versterken tussen online platforms, overheidsinstanties en maatschappelijke organisaties**

Een gecoördineerd aanpak tussen overheidsinstanties, maatschappelijke organisaties en handhavers noodzakelijk om het probleem effectief te bestrijden (Bartelds et al., 2023; Greijdanus et al., 2023). De laatste jaren wordt de belangrijke rol van techbedrijven ook erkend en wordt er gekeken hoe de samenwerking tussen overheden en techbedrijven via Privaat-publieke samenwerkingen vorm kunnen krijgen (Greiman, 2015). Een van de belangrijkste bijdragen van techbedrijven aan deze samenwerking ligt bij het delen van data en inzichten over online geweld. Door trends, patronen en signalen van online geweld tijdig te analyseren en deze informatie te delen met overheidsinstanties en maatschappelijke organisaties, kunnen vroege interventies worden ingezet. Zo stelt het monitoren van escalaties in online discussies of het snel signaleren van haatdragende content de politie en andere handhavingsinstanties in staat om gericht en tijdig in te grijpen. Daarnaast kunnen techbedrijven een actieve rol spelen in het verbeteren van de meldinfrastructuur. In samenwerking met meldpunten zoals MiND en overheidsinstanties zoals de politie en het Openbaar Ministerie, kunnen platforms meldingssystemen toegankelijker en effectiever maken. Het automatisch doorgeleiden van relevante meldingen naar handhavingsinstanties kan de reactietijd verkorten en slachtoffers sneller en beter ondersteunen. Techbedrijven kunnen ook bijdragen aan het versterken van educatie- en preventieprogramma's. Door samen te werken met scholen, jongerenwerkers en maatschappelijke organisaties, kunnen platforms voorlichting geven

over veilig online gedrag en de gevolgen van online geweld. Zo kunnen sociale mediabedrijven jongeren trainen in digitale weerbaarheid en het herkennen van desinformatie en online intimidatie. Door deze combinatie van technologische oplossingen, maatschappelijke betrokkenheid en juridische ondersteuning kunnen techbedrijven niet alleen de impact van online geweld beperken, maar ook bijdragen aan het creëren van een veiliger en inclusiever online klimaat. Een strategische, gezamenlijke aanpak biedt de beste kans om online geweld effectief terug te dringen en het vertrouwen in digitale platforms te herstellen (Bartelds et al., 2023; Greijdanus et al., 2023).

### **Moderatie**

De meeste online platforms hebben hun eigen huisregels voor contentmoderatie, maar de handhaving hiervan blijft een grote uitdaging. Er zijn significante verschillen tussen platforms in hoe ze balans zoeken tussen het beschermen van gebruikers en het waarborgen van de vrijheid van meningsuiting. De regels variëren van het verbod op haatzaaien en intimidatie tot het reguleren van valse informatie. Elk platform heeft specifieke richtlijnen, afhankelijk van hun doelgroepen en het soort inhoud dat wordt gedeeld. Ook de manier waarop de platformen zich positioneren in de markt lijkt de huisregels die ze hanteren te beïnvloeden. Zo omschrijft TikTok het platform als puur een entertainment platform en zijn de huisregels gericht op het creëren van een 'veilige en warme omgeving' waarin creativiteit centraal staat ([Community Guidelines | TikTok](#)). Terwijl X zich meer positioneert op het idee van het beschermen van vrijheid van meningsuiting ([Rules and policies](#)). Terwijl Meta recent aanpassingen heeft doorgevoerd in de huisregels waardoor er nu minder wordt gefocust op misinformatie en factchecking ([The Code of Conduct | Meta](#); [Meta gets rid of fact checkers and says it will reduce 'censorship' | CNN Business](#)). Tenslotte lijken de huisregels van Nu.nl het meest uitgebreid waarbij er een focus lijkt te liggen op de journalistieke waarden van het platform ([Wil je meepraten op NUjij? Dit zijn onze huisregels | NUjij | NU.nl](#)).

De variatie in huisregels wordt beïnvloed door culturele verschillen, wetgeving en zakelijke belangen. Op basis van deze huisregels gebruiken sociale mediaplatforms een combinatie van geautomatiseerde systemen en menselijke moderators om inhoud te controleren en ervoor te zorgen dat deze voldoet aan de richtlijnen. Geautomatiseerde systemen scannen berichten op schadelijke inhoud, zoals haatzaaien en intimidatie, en verwijderen of markeren deze wanneer nodig. Hoewel automatisering helpt om problematische inhoud snel te identificeren, zijn menselijke moderators (zoals psychologen of journalisten) vaak noodzakelijk voor complexere gevallen. De algoritmes die door verschillende platforms worden gebruikt, verschillen van elkaar, evenals de mate waarin ze menselijke moderators inzetten.

Recent onderzoek wijst uit dat de variatie in moderatiesystemen leidt tot onduidelijkheden en inconsistenties. Geautomatiseerde systemen hebben moeite met het detecteren van contextuele nuances. De definities van haatzaaien variëren, wat de effectiviteit van moderatie bemoeilijkt (Hietanen & Eddebo, 2022). Online interventies voor het verminderen van haatzaaien en cyberhaat leveren vaak niet de gewenste resultaten op, voornamelijk door de uitdagingen bij het begrijpen van de context waarin bepaalde uitingen worden gedaan (Windisch et al., 2022). Algoritmes hebben beperkingen bij het herkennen van "othering"-taal, die vaak wordt gebruikt in haatzaaien (Alorainy et al., 2018). Haatzaaien draagt bij aan polarisatie in participatieve samenlevingen, wat benadrukt hoe belangrijk het is om niet alleen de inhoud, maar ook de bredere sociale dynamiek te begrijpen (Perez-Escolar & Noguera-Vivo, 2021).

Het verbeteren van meldsystemen voor gebruikers is essentieel om een effectievere moderatie te waarborgen. Gebruikers moeten in staat zijn om gemakkelijk schadelijke content te melden, en deze meldingen zouden snel en transparant moeten worden behandeld. Daarnaast benadrukt Perez-Escolar & Noguera-Vivo (2021) het belang van algoritmische transparantie, waarbij platforms duidelijker moeten zijn over hoe hun algoritmes werken, welke gegevens ze gebruiken en welke beslissingen ze nemen. Dit zou niet alleen de effectiviteit van de moderatie verbeteren, maar ook zorgen voor meer vertrouwen van de gebruikers in het systeem en een eerlijker proces voor iedereen.

### **Online paspoort**

Een online paspoort is een concept waarbij elke internetgebruiker een vaste digitale identiteit heeft. Het komt erop neer dat men online slechts één account of profiel kan aanmaken, gekoppeld aan een geverifieerd identiteitsbewijs. Daarmee wordt het gebruik van anonieme of meerdere profielen sterk beperkt. Bedoeling is dat online wangedrag zoals haatberichten en bedreigingen zo lastiger wordt, omdat gebruikers niet langer ongestraft onder verschillende namen kunnen opereren. Deze aanpak wordt voorgesteld als een manier om de verantwoordelijkheid van internetgebruikers te vergroten en zo online geweld tegen te gaan. Tegelijkertijd roept het veel vragen op over de veiligheid, gegevensbescherming en het grondrecht van vrijheid van meningsuiting (bitsoffreedom, 2019; privacyweb, z.d.).

#### *Redirect method*

Door de Redirect Methode worden gebruikers die zoeken naar haatdragende of opruiende content omgeleid en doorverwezen naar educatieve en/of de-escalerende informatie en minder polariserende perspectieven. Deze technologische interventie gebruikt zoekalgoritmes en advertentie-technologie om gebruikers die zoeken naar schadelijke content (zoals extremisme of haatzaaiende

uitingen) om te leiden naar alternatieve, constructieve boodschappen. Zoekwoorden die verband houden met gewelddadige of schadelijke overtuigingen activeren advertenties op zoekmachines en sociale media. Gebruikers die op deze advertenties klikken, worden doorgeleid naar inhoud die deze overtuigingen uitdaagt of nuanceert. De methode maakt gebruik van bestaande content en algoritmen om gerichte interventies mogelijk te maken. Het RAND-onderzoek over de Redirect Method toont aan dat de advertenties effectief zijn in het bereiken van de doelgroep, met klikpercentages die overeenkomen met industriestandaarden. Gebruikers besteedden tijd aan het bekijken van het materiaal, maar er is tot op heden nog weinig bewijs dat dit ook daadwerkelijk leidt tot gedragsverandering. Het RAND rapport geeft aan dat er meer onderzoek nodig is naar de effectiviteit van de Redirect Method vooral op de lange termijn. Bovendien beveelt het rapport aan om ex-extremisten te betrekken bij het creëren van content en bij verder onderzoek (Helmus & Klein, 2018); ["Redirect Method" Yields Valuable Insights for Countering Online Extremism - VOX - Pol](#)).

## 8 Conclusie/aanbevelingen

De online wereld biedt ruimte voor creativiteit, verbinding, en kennisdeling en dialoog, maar tegelijkertijd is het een platform waar haat, intimidatie en polarisatie vrij spel kunnen hebben. In onze steeds meer gedigitaliseerde wereld heeft online geweld een diepgaande impact, niet alleen op de individuen die het direct treft, maar ook op de bredere samenleving. Het is een vorm van geweld die vaak onzichtbaar blijft voor de buitenwereld, maar die serieuze gevolgen heeft voor de betrokkenen. Van bedreigingen en intimidatie tot haatzaaien en opruiing – de online wereld blijkt ook een broedplaats voor verschillende vormen van schadelijk gedrag. De gevolgen zijn verstrekkend: het ondermijnt het gevoel van veiligheid, breekt sociale cohesie af en zorgt voor een samenleving waarin mensen zich steeds minder vrij voelen om hun mening te uiten of deel te nemen aan online discussies. Voor dit rapport hebben we een exploratieve literatuurstudie uitgevoerd van de (inter)nationale wetenschappelijke en grijze literatuur om de complexiteit van strafbaar online geweld tussen burgers in Nederland te verkennen. Hierbij hebben we ons vooral gericht op de aard en omvang van online geweld, de onderliggende beïnvloedende factoren, de kenmerken van betrokken burgers, en de mogelijke interventies en beleidsmaatregelen die effectief kunnen bijdragen aan het aanpakken van online geweld.

### **Een complexe werkelijkheid**

Online geweld in Nederland kent vele vormen, van bedreigingen en intimidatie tot haatzaaien en opruiing. Hoewel er steeds meer gegevens beschikbaar zijn, blijft het lastig om de werkelijke omvang van het probleem volledig in kaart te brengen. Dit komt deels door onderrapportage en de verschillende manieren waarop online geweld gedefinieerd wordt. Factoren zoals anonimiteit op sociale media, maatschappelijke polarisatie en de rol van algoritmes die schadelijke content versterken, spelen een belangrijke rol in de verspreiding van online geweld. Theorieën zoals het Online Disinhibition Effect en Moral Disengagement helpen ons te begrijpen waarom mensen in de digitale ruimte gedrag vertonen dat ze in de fysieke wereld vaak zouden vermijden.

Slachtoffers van online geweld zijn vaak jongeren, vrouwen, lhbt+ personen en mensen met een migratieachtergrond. De plegers van online geweld vertonen vaak tekenen van stress, depressie of sociale isolatie. Er is daarnaast een dynamiek tussen plegers en slachtoffers: soms worden slachtoffers zelf pleger. Dit toont aan hoe complex en gelaagd het probleem is. Het benadrukt dat er niet één enkel type slachtoffer of pleger is, en dat de dynamiek van online geweld niet altijd eenduidig is.

### **De weg vooruit: gezamenlijke verantwoordelijkheid en samenwerking**

Om online geweld effectief aan te pakken, is een integrale aanpak nodig. Dit betekent dat we niet alleen moeten kijken naar technologische oplossingen, maar ook naar onderwijs, wetgeving en gemeenschapsgerichte interventies. We hebben gezamenlijke inspanningen van overheden, tech-bedrijven en maatschappelijke organisaties nodig om de negatieve dynamieken die online geweld in de hand werken te doorbreken. Het verbeteren van de handhaving van bestaande wetgeving en het bevorderen van internationale samenwerking zijn eveneens belangrijke pijlers in het bestrijden van dit probleem. Daarnaast moet de digitale weerbaarheid van burgers versterkt worden, met speciale aandacht voor jongeren. Door hen bewust te maken van de gevaren van online geweld en hen vaardigheden aan te leren om zich in de digitale wereld veilig te bewegen, kunnen we voorkomen dat zij slachtoffers worden. Maar ook volwassenen moeten hun digitale vaardigheden en kennis verbeteren om effectief tegen online geweld op te treden.

Online geweld is een probleem dat de hele samenleving raakt. Het heeft niet alleen invloed op de slachtoffers zelf, maar ook op de samenleving als geheel. Dit rapport benadrukt de urgentie van een gecoördineerde aanpak waarbij verschillende interventies, wetgeving en technologische oplossingen elkaar aanvullen. Alleen door samenwerking en gezamenlijke verantwoordelijkheid kunnen we een veilige digitale omgeving creëren voor iedereen. Het is van groot belang dat we de dialoog blijven aangaan, nieuwe initiatieven blijven ontwikkelen en ons blijven inzetten voor een samenleving waarin respect en veiligheid centraal staan – zowel offline als online.

## 9 Bronnen

Alorainy, W., Burnap, P., Liu, H., & Williams, M. L. (2019). "The enemy among us" detecting cyber hate speech with threats-based othering language embeddings. *ACM Transactions on the Web (TWEB)*, 13(3), 1-26.

Bantema. (2022). Bestuurlijke mogelijkheden en uitdagingen bij de aanpak van online aangejaagde ordeverstoringen.

Bartelds, A., Postma, L., de Vries, S., Greijdanus, H., & Bantema, W. (2023). Preventie van online aangejaagd geweld: Een praktijkverkenning naar de online werkwijze van jongerenwerkers en politie in Amsterdam Zuid-Oost. [Preventie van online aangejaagd geweld: Een praktijkverkenning naar de online werkwijze van jongerenwerkers en politie in Amsterdam Zuid-Oost - the University of Groningen research portal](#).

Bitsoffreedom. (2019, 4 januari). *Anonimiteit moet een optie blijven*. Bits of Freedom. <https://www.bitsoffreedom.nl/2019/01/04/anonimiteit-moet-een-optie-blijven/#:~:text=D66%20en%20VVD%20lanceerden%20het,Zo%E2%80%99n%20online%20paspoort%20werd>.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77-101.

Brown, A. (2018). What is so special about online (as compared to offline) hate speech?. *Ethnicities*, 18(3), 297-326.

Bundesregierung. (2022). *Gesetz zur Verbesserung der Rechtsdurchsetzung in sozialen Netzwerken (NetzDG)*. Geraadpleegd van [https://www.bundesjustizamt.de/SharedDocs/Downloads/DE/Hasskriminalitaet/20220721\\_NetzDG.pdf?\\_\\_blob=publicationFile&v=2](https://www.bundesjustizamt.de/SharedDocs/Downloads/DE/Hasskriminalitaet/20220721_NetzDG.pdf?__blob=publicationFile&v=2).

Castañó-Pulgarín, S. A., Suárez-Betancur, N., Vega, L. M. T., & López, H. M. H. (2021). Internet, social media and online hate speech. Systematic review. *Aggression and violent behavior*, 58, 101608.

Castellanos, M., Wettstein, A., Wachs, S., Kansok-Dusche, J., Ballaschk, C., Krause, N., & Bilz, L. (2023, April). Hate speech in adolescents: A binational study on prevalence and demographic differences. In *Frontiers in Education* (Vol. 8, p. 1076249). Frontiers Media SA.

CBS (2023) Online veiligheid en criminaliteit 2022. [8. Online discriminatie | CBS](#).

Centrum Informatie en Documentatie Israel. Monitor antisemitische incidenten. [Monitor Antisemitische Incidenten · CIDI](#).

Chetty, N., & Alathur, S. (2018). Hate speech review in the context of online social networks. *Aggression and violent behavior*, 40, 108-118.

De Vries & Bantema. (2021). De lokale driehoek moet ook online weten wat er speelt.

European Commission. (2018). *General Data Protection Regulation (GDPR)*. Geraadpleegd van [https://ec.europa.eu/info/law/law-topic/data-protection\\_en](https://ec.europa.eu/info/law/law-topic/data-protection_en).

European Commission. (n.d.). *A safer & fairer online environment*. Shaping Europe's digital future. <https://digital-strategy.ec.europa.eu/en/factpages/safer-fairer-online-environment>.

Greijdanus, H., Postmes, T., Bartelds, A., Postma, L., de Vries, S., & Bantema, W. (2023). Interventies in de cyclus van online aangejaagd geweld: Inzichten uit een literatuurreview. [Interventies in de cyclus van online aangejaagd geweld: Inzichten uit een literatuurreview - the University of Groningen research portal](#).

Greiman, V. A. (2015). Public/Private Partnerships in Cyberspace: Building a Sustainable Collaboration. *Journal of Information Warfare*, 14(3), 30-42.

Helmus, T. C., & Klein, K. (2018). *Assessing outcomes of online campaigns countering violent extremism: A case study of the Redirect Method* (RR-2813-GNF). RAND Corporation. [https://www.rand.org/pubs/research\\_reports/RR2813.html](https://www.rand.org/pubs/research_reports/RR2813.html).

Hietanen, M., & Eddebo, J. (2023). Towards a definition of hate speech—With a focus on online contexts. *Journal of Communication Inquiry*, 47(4), 440-458.

Hollenbaugh, E. E., & Everett, M. K. (2013). The effects of anonymity on self-disclosure in blogs: An application of the online disinhibition effect. *Journal of Computer-Mediated Communication*, 18(3), 283-302.

Jahan, M. S., & Oussalah, M. (2023). A systematic review of hate speech automatic detection using natural language processing. *Neurocomputing*, 546, 126232.

Joinson, A. N. (2007). Disinhibition and the Internet. In *Psychology and the Internet* (pp. 91-106). Academic Press.

Khangura S., Konnyu K., Cushman R., et al. Evidence summaries: the evolution of a rapid review approach. *Systematic Reviews*. 2012; 1: 10.

Kholmatov, A., & Yanikoglu, B. (2005). Identity authentication using improved online signature verification method. *Pattern recognition letters*, 26(15), 2400-2408.

Kennisplatform Inclusief Samenleven (2022) 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.

Lewis, R., Rowe, M., & Wiper, C. (2017). Online abuse of feminists as an emerging form of violence against women and girls. *British journal of criminology*, 57(6), 1462-1481.

Madriaza, P., Hassan, G., Brouillette-Alarie, S., Mouchingam, A. N., Durocher-Corfa, L., Borokhovski, E., ... & Paillé, S. (2025). Exposure to hate in online and traditional media: A systematic review and meta-analysis of the impact of this exposure on individuals and communities. *Campbell Systematic Reviews*, 21(1), e70018.

Meld Online Discriminatie (2024). Landelijk rapport discriminatiecijfers 2023. [Discriminatiecijfers 2023](#) | [Meld.Online Discriminatie](#) | [Meld.Online Discriminatie](#).

Mkhize, S., Nunlall, R., & Gopal, N. (2020). An examination of social media as a platform for cyber-violence against the LGBT+ population. *Agenda*, 34(1), 23-33.

Mukred, M., Mokhtar, U. A., Moafa, F. A., Gumaei, A., Sadiq, A. S., & Al-Othmani, A. (2024). The roots of digital aggression: Exploring cyber-violence through a systematic literature review. *International Journal of Information Management Data Insights*, 4(2), 100281.

Naranjo-Pou S, Zych I. Association between victimization and perpetration of cyberhate: The moderating role of social dominance orientation. *J Adolesc*. 2025 Feb;97(2):479-488. doi: 10.1002/jad.12433.

Nederlands Studiecentrum Criminaliteit en Rechtshandhaving (2018) [Slachtofferschap van online criminaliteit. Een onderzoek naar behoeften, gevolgen en verantwoordelijkheden na slachtofferschap van cybercrime en gedigitaliseerde criminaliteit.](#)

Obermaier, M. (2024). Youth on standby? Explaining adolescent and young adult bystanders' intervention against online hate speech. *New Media & Society*, 26(8), 4785-4807.

Olteanu, A., Castillo, C., Boy, J., & Varshney, K. (2018, June). The effect of extremist violence on hateful speech online. In *Proceedings of the international AAAI conference on web and social media* (Vol. 12, No. 1).

Openbaar Ministerie (2024) Strafbare discriminatie in beeld. [OM: Geregistreeerde discriminatiefeiten in 2023 gestegen | Nieuwsbericht | Openbaar Ministerie.](#)

Pabian, S., & Vandebosch, H. (2024). An Examination of (Proximal Determinants of) Online Aggression Perpetration Among Adults Aged 19 to 79. *Deviant Behavior*, 46(1), 83-100.

Pashang, S., & Khanlou, N. (2022). The changing face of gender-based cyber violence during the COVID-19 Pandemic: Unmasking virtual identities. *Journal of concurrent disorders*, 4(3), 52.

Pérez-Escobar, M., & Noguera-Vivo, J. M. (2022). *Hate speech and polarization in participatory society* (p. 118). Taylor & Francis.

Peterson, J., & Densley, J. (2017). Cyber violence: What do we know and where do we go from here?. *Aggression and violent behavior*, 34, 193-200.

Posetti, J., Shabbir, N., Douglas, O., & Gardiner, B. (2022). The chilling: A global study of online violence against women journalists.

Privacy-web. (z.d.). *Het debat over de wenselijkheid van EU-regelgeving voor online anonimiteit*. Privacy-web. <https://privacy-web.nl/nieuws/het-debat-over-de-wenselijkheid-van-eu-regelgeving-voor-online-anonimiteit/#:~:text=Diverse%20politieke%20partijen%20zetten%20zich,gebruikers%20te%20registreren%20en%20deze>.

Rathenau Instituut (2021) *Online ontspoord*. [Online ontspoord | Rathenau Instituut](#).

Roks, R., & Van den Broek, J. B. A. (2020). Cappen voor Clout? Een verkennend onderzoek naar Rotterdamse jongeren, drill en geweld in het digitale tijdperk.

Runions, K. C., & Bak, M. (2015). Online moral disengagement, cyberbullying, and cyber-aggression. *Cyberpsychology, behavior, and social networking*, 18(7), 400-405.

Saul J.E., Willis C.D., Bitz J., et al. A time-responsive tool for informing policy making: rapid realist review. *Implementation Science*. 2013; 8: 103.

Šincek, D., Duvnjak, I., & Milić, M. (2017). Psychological outcomes of cyber-violence on victims, perpetrators and perpetrators/victims. *Hrvatska revija za rehabilitacijska istraživanja*, 53(2), 98-110.

Suler, J. (2004). The online disinhibition effect. *Cyberpsychology & behavior*, 7(3), 321-326.

't Hoff-de Goede & Leukfeldt (2024). Slachtofferschap van cybercriminaliteit.

Tweede Kamer der Staten-Generaal. (2025, 20 februari). *Motie van het lid Van der Werf c.s. over gedifferentieerde leeftijdsgrenzen voor socialmediaplatforms (Kamerstuk 26643, nr. 1299)*. Geraadpleegd op 24 maart 2025, van <https://www.tweedekamer.nl/kamerstukken/moties/detail?id=2025D07255&did=2025D07255>.

UK Government. (1988). *Malicious Communications Act 1988*.

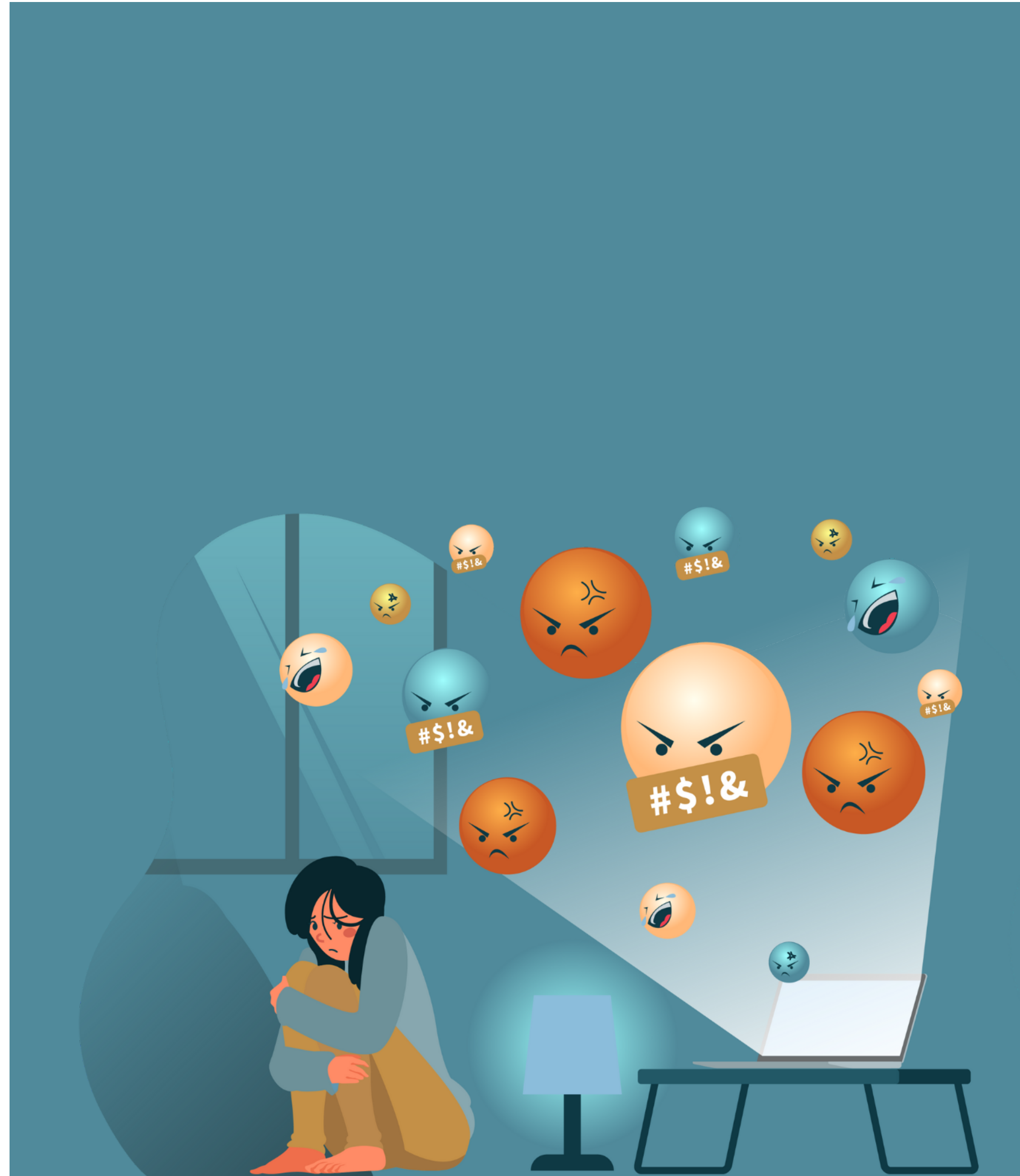
UK Government. (2023). *Online Safety Act explainer*. Geraadpleegd op 24 maart 2025 van <https://www.gov.uk/government/publications/online-safety-act-explainer/online-safety-act-explainer>.

Verwey Jonker Instituut (2023). "Rechtsextremisme op sociale mediaplatforms". Hier beschikbaar: [Aanbevelingsalgoritmen spelen beperkte rol bij rechtsextremistische radicalisering - Verwey-Jonker Instituut](#).

Vuik G., Bantema W. (2021) Het digitaal gebiedsverbod als gemeentelijk instrument tegen online aangejaagde ordeverstoringen. *Bestuurswetenschappen* 75 (3).

Windisch, S., Wiedlitzka, S., Olaghere, A., & Jenaway, E. (2022). Online interventions for reducing hate speech and cyberhate: A systematic review. *Campbell systematic reviews*, 18(2), e1243.

Wright, M. F., & Wachs, S. (2021). Does empathy and toxic online disinhibition moderate the longitudinal association between witnessing and perpetrating homophobic cyberbullying?. *International journal of bullying prevention*, 3, 66-74.



## COLOFON

Auteurs: Drs. E. de Weger  
Dr. M. Odekerken  
D. Woertink, MSc  
Illustratie Copilot, Microsoft  
Uitgave Verwey-Jonker Instituut 2025  
Giessenplein 59-C  
3522 KE Utrecht  
T (030) 230 07 99  
E [secr@verwey-jonker.nl](mailto:secr@verwey-jonker.nl)  
I [www.verwey-jonker.nl](http://www.verwey-jonker.nl)

De publicatie kan gedownload worden via onze website:  
<http://www.verwey-jonker.nl>

ISBN 978-94-6409-373-5  
© Verwey-Jonker Instituut, Utrecht, mei 2025.

Het auteursrecht van deze publicatie berust bij het Verwey-Jonker Instituut.  
Gedeeltelijke overname van teksten is toegestaan, mits daarbij de bron wordt vermeld.

The copyright of this publication rests with the Verwey-Jonker Institute. Partial reproduction of the text is allowed, on condition that the source is mentioned.