costaid

Coping Strategies Against Information Disorder

Richtlijnen voor eerstelijners

Co-funded by the European Union

Gefinancierd door de Europese Unie. De hier geuite ideeën en meningen komen echter uitsluitend voor rekening van de auteur(s) en geven niet noodzakelijkerwijs die van de Europese Unie of het Europese Uitvoerende Agentschap onderwijs en cultuur (EACEA) weer. Noch de Europese Unie, noch het EACEA kan ervoor aansprakelijk worden gesteld.

Auteurs:

Eliane Smits van Waesberghe & Tim Paulusse – Verwey-Jonker Instituut (Hoofdredacteuren)

Leen D'Haenens & Joyce Vissenberg – KU Leuven

Tzvetalina Genova – International Management Institute

Wolfgang Eisenreich – Wissenschaftsinitiative Niederösterreich

Sonja Bercko Eisenreich – Integra Institute

Alenka Valjašková – QUALED

Pantelis Balaouras – Connexions

Verklaring inzake auteursrecht:

costaid
Coping Strategies Against Information Disorder

Hoofdstuk 3

Technologie & Tools

Doelgroep

Deze richtlijnen zijn gericht op zogenaamde “eerstelijners” (in het Engels: first-liners). “Eerstelijners” is een overkoepelende term voor alle mensen die in direct contact staan met mensen die kwetsbaar zijn voor informatiestoornissen, gericht op groepen in het beroepsonderwijs. Voorbeelden van mensen die onder de overkoepelende term vallen zijn: opvoeders, leraren, trainers, jeugdadviseurs, maatschappelijk werkers en jeugdwerkers. Dit is echter een onvolledige lijst. De reikwijdte van dit project omvat ook andere mensen die werkzaam zijn in de onderwijs-, sociale of gezondheidszorgsector.

3.1 Inleiding tot dit hoofdstuk

Het concept van desinformatie is in eerdere hoofdstukken besproken. Nieuws, maar ook mis- en desinformatie, wordt doorgaans verspreid door professionele nieuwsaanbieders, waaronder publieke media, commerciële nieuwsmedia, onafhankelijke professionele journalistiek of amateurgebruikers (in het geval van sociale mediaplatforms). Het is toegankelijk en beschikbaar in verschillende formaten:

- **Digitaal nieuws:** Nieuws verspreid via internetkanalen in digitaal mediaformaat (tekst, afbeeldingen, audio, video).
- **Gedrukte media:** Kranten en tijdschriften met inhoud op basis van tekst en beeld.
- **Omroep:** TV en radio met op video en audio gebaseerde inhoud.

In dit hoofdstuk richten we ons vooral op het eerste formaat, digitaal nieuws. De gepresenteerde discussies kunnen echter ook van toepassing zijn op andere formats die gebruikmaken van video, audio en/of afbeeldingen.

Digitaal nieuws is voornamelijk toegankelijk via nieuwsplatforms die worden beheerd door professionele nieuwsaanbieders. Gebruikers hebben gratis toegang tot deze nieuwsplatforms of kunnen zich erop abonneren. Deze platforms bieden webfeeds in populaire browsers en nieuwsfeeds op sociale mediaplatforms, zodat gebruikers gepersonaliseerde nieuwsupdates kunnen ontvangen.

Een nieuwsfeed is een webpagina of scherm dat regelmatig wordt bijgewerkt om het laatste nieuws of informatie weer te geven. Gepersonaliseerde nieuwsfeeds zijn diensten die zijn geïntegreerd in webbrowsers (webfeeds) en sociale mediaplatforms die nieuws leveren aan gebruikers op basis van hun persoonlijke voorkeuren. Deze voorkeuren worden bepaald door verschillende factoren, zoals het abonneren op webfeedkanalen, het bezoeken van specifieke webpagina's en meer.

Op sociale mediaplatforms ontvangen gebruikers ook gedeelde media van andere gebruikers, waaronder gevallen van misinformatie.

Gepersonaliseerd nieuws verschijnt op de browser of het sociale mediaplatform van een gebruiker via diensten en sociale media algoritmen. Gebruikers zijn zich mogelijk niet volledig bewust van de werking van deze algoritmen, omdat het onduidelijk blijft of de selectie van nieuws uitsluitend is gebaseerd op de voorkeuren van de gebruiker of op andere criteria. Sociale mediaplatforms presenteren bijvoorbeeld nieuws op basis van de nieuwsaanbieders die een gebruiker volgt, de leesgewoonten van zijn vrienden of recent aangeklikte artikelen, wat een soort voorkeur kan vormen. Bijgevolg wordt een lijst met nieuws gefilterd, wat betekent dat niet al het nieuws wordt weergegeven, maar eerder het nieuws dat het algoritme het meest interessant vindt voor de gebruiker. Velen beweren dat dit een informatiebel creëert, waarin nieuws wordt geselecteerd door een algoritme en mogelijk geen nieuws bevat waarin een gebruiker geïnteresseerd is, maar dat het algoritme niet opneemt. Daarom is het belangrijk om niet alleen te vertrouwen op persoonlijke nieuwsfeeds, maar om professionele platforms van nieuwsaanbieders te bezoeken die

worden vertrouwd. Het is altijd belangrijk om de geloofwaardigheid en nauwkeurigheid van nieuws kritisch te beoordelen, ongeacht het formaat of de leveringsmethode.

3.2 Zoekmachines & algoritmen

Veel websites op internet streven ernaar om hun gebruikers betrokken te houden en hun tijd op het platform te maximaliseren. Dit geldt met name voor sociale mediaplatforms, die verschillende strategieën toepassen om de gebruikerservaring te verbeteren. Een van die strategieën is gebruikers content tonen die aansluit bij hun interesses. Websites kunnen echter geen gedachten lezen om de voorkeuren van gebruikers te kennen. Om dit aan te pakken, worden algoritmen gebruikt om gebruikersgegevens te analyseren en gepersonaliseerde inhoud te leveren.

Algoritmen zijn complexe formules die de inhoudsvoorkeuren van een individu observeren, meten en berekenen. Het kan hierbij gaan om factoren zoals de kijktijd van de gebruiker voor specifieke soorten video's, de tijd die hij besteedt aan een bepaalde post of engagementacties zoals het achterlaten van likes of reacties. Door deze gegevens te analyseren, bepalen algoritmes welk type inhoud gebruikers langer betrokken houdt. Deze informatie wordt vervolgens gebruikt om soortgelijke inhoud te cureren en aan te bevelen aan de gebruiker.

Hoewel deze aanpak logisch en onschuldig lijkt, zijn er potentiële nadelen aan algoritmisch gebruik. De voorkeur van het algoritme voor inhoud die de gebruiker interessant vindt, kan filterbubbels creëren, waar gebruikers alleen worden blootgesteld aan specifieke standpunten. Filterbubbels beperken de diversiteit van de inhoud en kunnen leiden tot echokamers of bestaande echokamers versterken.

Naast filterbubbels kunnen algoritmes de extremiteit van inhoud versterken door steeds meer niche-, rand- en extreme posts aan te bevelen. Het doel is om gebruikers betrokken te houden, maar dit kan ertoe leiden dat gebruikers terechtkomen in online ruimtes zonder verschillende standpunten genaamd echokamers. Dit proces waarin een persoon begint in een informatierijke omgeving begint maar eindigt in een informatiearme echokamer, wordt een “rabbit hole” of "konijnenhol" genoemd.

Filterbubbels, konijnenholen en echokamers stellen gebruikers bloot aan radicale inhoud en de daarmee verbonden gemeenschappen. Deze online ruimtes vormen een vruchtbare bodem voor de ontwikkeling, groei, vervorming en verspreiding van misinformatie en desinformatie.

Naarmate gebruikers het konijnenhol doorlopen, worden extreme discussiepunten en foutieve informatie genormaliseerd. Deze normalisering vergemakkelijkt verder de verspreiding en acceptatie van mis- en desinformatie uit radicale bronnen.

Door de invloed van algoritmen op gebruikerservaringen te begrijpen, kunnen we de risico's van filterbubbels, echokamers en konijnenholen beter begrijpen. Deze kennis is essentieel voor het navigeren door het online landschap en het aanpakken van de uitdagingen die informatieverstoring met zich meebrengt.

3.3 Online strategieën van desinformatie en extremistische organisaties

In de vorige hoofdstukken hebben we onderzocht hoe informatiestoornis wordt gecreëerd en in stand wordt gehouden binnen echokamers en filterbubbels. Maar hoe bereikt mis- en desinformatie mensen buiten deze ruimtes?

The meest gebruikte methoden

Mis- en desinformatie kan via talloze vormen van communicatie worden verspreid. Ze worden echter meestal verspreid via verschillende kanalen, zoals sociale mediaplatforms, websites, e-mail en mond-tot-mondreclame. De meest voorkomende methoden waarmee informatiestoornis wordt verspreid, zijn nauw verwant aan de zeven categorieën van problematische inhoud die worden besproken in *Hoofdstuk 1: 'nepnieuws' begrijpen*:

- **Satire of parodie:** Sommige mis- en desinformatie wordt gemaakt voor amusementsdoeleinden of satire, maar kan verkeerd worden opgevat als echt nieuws. Satirische websites of accounts op sociale media kunnen humoristische of overdreven verhalen publiceren, maar lezers die zich niet bewust zijn van hun satirische aard kunnen deze verwarren met feitelijke informatie.
- **Clickbait:** artikelen die verkeerde of misleidende informatie bevatten, maken vaak gebruik van sensationele of misleidende koppen om de aandacht te trekken en meer klikken of weergaven te genereren. Het doel is om de nieuwsgierigheid of emoties van mensen uit te buiten om verkeer naar een website te leiden en inkomsten te genereren via advertenties.
- **Onjuiste voorstelling van zaken:** Hierbij gaat het om het verdraaien of verkeerd weergeven van feitelijk nieuws door selectief feiten te presenteren of cruciale informatie weg te laten. Het kan gaan om het uit de context halen van uitspraken, het veranderen van afbeeldingen of video's of het verdraaien van de betekenis van een verhaal zodat het in een bepaald verhaal past.
- **Imitatie:** mis- en desinformatie kan ook bestaan uit het imiteren van gerenommeerde nieuwsbronnen of publieke figuren om foutieve informatie geloofwaardig te maken. Dit kan worden gedaan door het creëren van valse websites of sociale media-accounts die legitieme bronnen nabootsen en lezers laten geloven dat de informatie betrouwbaar is.
- **Politieke manipulatie:** Informatiestoornis wordt soms gecreëerd of verspreid met de bedoeling de publieke opinie of verkiezingen te beïnvloeden. Dit kan het verspreiden van foutieve informatie over politieke kandidaten inhouden, het manipuleren van het publieke sentiment of het uitbuiten van bestaande vooroordelen en verdeeldheid in de samenleving.
- **Fabricage:** mis- en desinformatie kan volledig verzonnen zijn, zonder basis in de werkelijkheid. Het gaat om het creëren van fictieve verhalen, citaten of gebeurtenissen om lezers of kijkers te misleiden.

- **Versterking via sociale media:** Sociale mediaplatforms spelen een belangrijke rol bij de verspreiding van mis- en desinformatie. Foutieve of verzonnen verhalen kunnen snel viraal gaan doordat gebruikers ze delen en opnieuw plaatsen, vaak zonder de juistheid van de informatie te verifiëren. De algoritmes die worden gebruikt door sociale mediaplatforms kunnen ook bijdragen aan de versterking door inhoud te promoten op basis van betrokkenheid in plaats van nauwkeurigheid.

Er zijn ook andere veelgebruikte methodes om mis- en desinformatie te verspreiden die in meer detail moeten worden besproken: trolling, doxxing en mainstreaming. Dit gebeurt in de volgende twee subhoofdstukken.

Trolling & doxxing

Een veelgebruikte strategie is trolling. Trolling wordt gedefinieerd als het opzettelijke gebruik van onbeleefdheid, agressie, bedrog en manipulatie in online communicatie om conflicten of amusement uit te lokken. Trollen lokken online conflicten uit door te misleiden, te manipuleren of agressief te zijn. Ze laten gesprekken ontsporen voor hun eigen plezier en doen zo in feite aan digitaal pesten.

Op kleine schaal kan trolling relatief onschuldig lijken, wat alleen maar leidt tot ergernis. Maar wanneer georganiseerde groepen van trollen een specifiek doel hebben, kan deze ergernis snel veranderen in een desinformatie-epidemie. Een voorbeeld hiervan is het Russische gebruik van sociale media trollen als 'wapen'. Rusland zette een groot netwerk van trollen in om wereldwijd desinformatie te verspreiden in meerdere talen, met als doel de online dialoog rondom Rusland te controleren. Deze trollen verspreidden niet alleen foutieve informatie, maar richtten zich ook op andere sociale mediagebruikers wanneer deze berichten deelden die afweken van het verhaal dat de trollen wilden promoten. Als gevolg daarvan onthielden veel gebruikers van sociale media zich ervan om over Rusland te discussiëren, waardoor de trollen met hun desinformatie het verhaal konden beheersen.

Doxxing, een andere vorm van internetpesten, houdt in dat persoonlijke informatie of de identiteit van personen online wordt onthuld zonder hun toestemming. Doxxing is niet een vorm van desinformatie, maar van malinformatie: juiste informatie die ingezet wordt om schade aan te richten. Hoewel deze tactiek dus niet specifiek betrekking heeft op het verspreiden van desinformatie, is het een andere strategie die internet trollen gebruiken om het verhaal over een bepaald onderwerp te controleren. Doxxing kan worden gebruikt om gebruikers van sociale media te intimideren en hun bereidheid te onderdrukken om inhoud te plaatsen die indruist tegen het verhaal van de trol.

Inzicht in de impact van trolling en doxxing is cruciaal voor het herkennen van de verschillende tactieken die worden gebruikt om online verhalen te manipuleren en te controleren. Deze strategieën dragen niet alleen bij aan de verspreiding van mis- en desinformatie, maar vormen ook een uitdaging bij het bevorderen van een open en geïnformeerde digitale omgeving.

Mainstreaming

Een belangrijke strategie om desinformatie en extremistische informatie te verspreiden is door deze te normaliseren of te "mainstreamen". Blootstelling speelt een cruciale rol in dit proces. Blootstelling aan mis- en desinformatie kan leiden tot hardnekkige misvattingen van mensen over de betreffende onderwerpen, waardoor de onjuiste ideeën in hun hoofden worden genormaliseerd. Deze blootstelling kan op verschillende manieren gebeuren.

Influencers

Een veelvoorkomende vorm van blootstelling is verspreiding van persoon tot persoon, waarbij individuen informatie delen met anderen. Dit kan gebeuren via persoonlijke interacties of op grotere schaal met influencers op sociale media. Influencers, die een aanzienlijk bereik hebben over verschillende groepen, kunnen onbewust of opzettelijk onjuiste informatie delen, waardoor een groot aantal individuen wordt beïnvloed. Dergelijke wijdverspreide blootstelling leidt tot de normalisatie van verkeerde informatie onder diverse doelgroepen.

Ironie, satire & memes

Extremistische individuen en organisaties gebruiken vaak humor, satire en ironie om hun ideeën te verspreiden.

Satire kan een krachtig middel zijn om onderdrukkende ideologieën uit te dagen, verhalen te veranderen of niche-standpunten te normaliseren binnen de mainstream. Op het gebied van desinformatie wordt satire op verschillende manieren gebruikt. Parodiewebsites zoals The Onion of De Speld publiceren niet-feitelijke inhoud met humor als einddoel, zonder de bedoeling om het publiek te misleiden. Bepaalde individuen en groepen gebruiken satire en ironie echter met kwade bedoelingen om reguliere journalistiek en wetenschap in diskrediet te brengen of extremistische ideeën en desinformatie te promoten. Door gebruik te maken van satire en humor wordt dergelijke inhoud toegankelijker en acceptabeler in het politieke discours, waardoor het een breder publiek bereikt.

Extremistische inhoud spreekt jongeren vaak aan als een vorm van vermaak of het zoeken naar sensatie. Jonge mensen, gedreven door een zoektocht naar betekenis, neigen naar intense en nieuwe ervaringen, waardoor ze vatbaarder zijn voor extremistische ideeën en de bijbehorende desinformatie.

Memes (veelgedeelde stukken humoristische culturele inhoud) dienen als een andere manier om extremistische ideologieën te verspreiden. Memes zijn er in verschillende

formaten, waaronder afbeeldingen, video's, audioclips, emoji's en symbolen. Hoewel memes zelf niet per definitie schadelijk zijn, gebruiken extremisten ze om hun ideeën te normaliseren. Het speelse karakter van memes stelt extremisten in staat om de schadelijkheid van hun berichten te verhullen, te ontkrachten of te ontkennen. Deze “scherpe” of provocerende inhoud wordt acceptabeler, en wanneer ze geconfronteerd worden met beschuldigingen van seksisme, racisme of vreemdelingenhaat, kunnen de makers het gemakkelijk afdoen als “gewoon een grap”. Deze vervaging van de grenzen tussen speelse ondeugendheid en problematische inhoud zorgt voor dubbelzinnigheid, waardoor het moeilijk is om onschuldige grappen te onderscheiden van extremistische boodschappen. Pepe the Frog, een stripfiguur op internet die oorspronkelijk was gemaakt als een onschuldige grap, werd toegeëigend door online witte supremacisten. Dit veroorzaakte verwarring bij internetgebruikers, omdat extremistische iteraties van deze meme vermengd raakten met de onschadelijke. De normalisering van extremistische inhoud vindt plaats naarmate meer mensen aan deze berichten worden blootgesteld, waardoor de scheidslijn tussen wat acceptabel is en wat niet, vervaagt.

3.4 Gemanipuleerde content

Technisch gezien bestaat alle informatie of "nieuws" in de media uit een combinatie van tekst, beeld, audio en video. Het cruciale aspect is echter het vaststellen van de authenticiteit van de informatie. Het is belangrijk op te merken dat desinformatie vaak echte foto's gebruikt, maar het verhaal manipuleert en de werkelijke feiten verdraait.

In het verleden werd in het algemeen aangenomen dat iedereen een tekst kon schrijven en dus niet altijd betrouwbaar waren, terwijl foto's, audio en video min of meer als authentiek werden beschouwd en voor aanpassingen professionele vaardigheden nodig waren. Maar met de recente technologische vooruitgang kunnen zelfs afbeeldingen, audio- en videobestanden worden gemanipuleerd. Dit kan worden gedaan door professionals of door toepassingen die gebruik maken van kunstmatige intelligentiesystemen, zoals **deepfake technologie**. Daarom is het cruciaal om te onderscheiden of een audio of video echt is opgenomen door een microfoon of videocamera, of dat het een resultaat is van bewerking door experts of kunstmatige intelligentie (AI) systemen (generatieve AI en synthetische media: stemklonen, deepfake video's). Daarnaast moet het technisch mogelijk zijn om de oorspronkelijke bron, producent of uitgever van een afbeelding, audio- of videobron te identificeren. De reden hiervoor is dat bronnen talloze keren kunnen worden gedeeld, gekopieerd of opnieuw gedistribueerd via het internet en sociale media. Daarom kan het voor regelmatige gebruikers een uitdaging zijn om de oorspronkelijke bron en producent te identificeren, zelfs als ze vermoeden dat het bericht verkeerde informatie bevat.

Gebruikers in staat stellen onderscheid te maken tussen echt en verzonnen nieuws vereist verschillende stappen. Meer informatie hierover is te vinden in *Hoofdstuk 2: Handelingen & Vaardigheden*. Het is aan te raden om dit hoofdstuk te lezen om de ingewikkelde details achter het herkennen van foutieve informatie te leren. Hier volgt echter een korte, zeer vereenvoudigde samenvatting:

- **Stap 1: Gebruikers bewust maken:** Gebruikers moeten zich ervan bewust zijn dat nieuws vervalst kan zijn. Het uitvoeren van bewustmakingsactiviteiten is cruciaal om gebruikers te informeren over wat verzonnen nieuws is en hoe ze zichzelf kunnen beschermen tegen de gevolgen ervan.
- **Stap 2: Controleer de betrouwbaarheid van de uitgever:** Het toegenomen bewustzijn van vervalst nieuws zorgt ervoor dat gebruikers de betrouwbaarheid van nieuwsbronnen en uitgevers in twijfel trekken. Het is essentieel om de mediavorm in overweging te nemen, of het nu een tv-kanaal, tijdschrift, krant (online of gedrukt) of een sociaal mediaplatform is. Mediakanalen die het gemakkelijk maken om nieuws te delen of opnieuw te verspreiden, kunnen minder betrouwbaar zijn. Omgekeerd zijn mediakanalen die bronidentificatie en -verificatie vergemakkelijken doorgaans betrouwbaarder.

Als het gaat om nieuws dat via internet wordt verspreid, moeten dienstverleners zoals nieuwsfeeds en sociale medianetwerken gebruikmaken van opkomende technologieën om de betrouwbaarheid van de bron te verifiëren en de oorspronkelijke uitgever en bron te traceren. Blockchaintechnologie is zo'n technologie die deze inspanningen kan vergemakkelijken.

Door deze stappen te volgen en gebruik te maken van technologie, kunnen gebruikers in staat worden gesteld om door het digitale landschap te navigeren, echt nieuws van verkeerde informatie te onderscheiden en weloverwogen beslissingen te nemen over de informatie die ze tegenkomen.

Inzicht in deepfakes: synthetische mediamanipulatie

Deepfakes, zoals gedefinieerd door het Cambridge Dictionary, zijn *"video- of geluidsopnames die iemands gezicht of stem vervangen door die van iemand anders, op een manier die echt lijkt"*.

In het artikel "Deepfake uitgelegd" uit 2020 vermeldt de schrijfster Meredith Somers dat *"(een) deepfake verwijst naar een specifiek soort synthetische media waarbij een persoon in een afbeelding of video wordt verwisseld met de gelijkenis van een andere persoon"*. Verder wordt uitgelegd dat *"de term 'deepfake' eind 2017 voor het eerst werd bedacht door een Reddit-gebruiker met dezelfde naam. Deze gebruiker creëerde een ruimte op de online nieuws- en aggregatiesite, waar ze pornografische video's deelden die gebruik maakten van open-source face-swapping technologie."*

Deepfakes hebben toepassingen gevonden in verschillende sectoren en zijn gebruikt voor verschillende doeleinden. Enkele opmerkelijke voorbeelden zijn:

- **Chantage:** Deepfakes kunnen worden gebruikt om vals belastend materiaal te genereren, wat kan leiden tot chantage. Omdat het steeds moeilijker wordt om deepfakes van echte inhoud te onderscheiden, kunnen slachtoffers van echte chantage bovendien beweren dat het bewijs nep is, waardoor ze het aannemelijk kunnen ontkennen.
- **Pornografie:** Deepfake pornografie heeft een grote bekendheid gekregen op het internet. Een rapport van de Nederlandse cybersecurity startup Deeptrace schatte dat ongeveer 96% van alle online deepfakes pornografisch was.
- **Politiek:** Deepfakes zijn gebruikt om bekende politici in video's in een verkeerd daglicht te stellen, desinformatie te verspreiden en de publieke perceptie te manipuleren. Voorbeelden zijn deepfakes met Barack Obama, Donald Trump, Volodymyr Zelenskyy en Vladimir Poetin.
- **Acteren/Films:** Er wordt gespeculeerd over het gebruik van deepfakes voor het creëren van digitale acteurs in toekomstige films. Hoewel digitaal geconstrueerde of veranderde mensen al eerder te zien waren in films, zouden deepfakes kunnen bijdragen aan nieuwe ontwikkelingen op dit gebied.
- **Sociale media:** Deepfakes worden gebruikt door gebruikers op verschillende sociale mediaplatforms. Individuen vervangen gezichten in populaire film- of seriescènes door hun eigen gezichten en maken zo gepersonaliseerde video's. Platforms als Facebook hebben maatregelen genomen om deepfakes te detecteren en als nep te markeren, waardoor ze minder prioriteit krijgen in de feeds van gebruikers.

Opsporen en beperken van deepfakes: technologische benaderingen

Onderzoekers zijn actief bezig met het verkennen van methoden om het probleem van deepfake audio en video te detecteren en aan te pakken. Er worden verschillende benaderingen nagestreefd:

- **Algoritmische detectie:** Eén benadering bestaat uit het ontwikkelen van algoritmen die gemanipuleerde inhoud kunnen identificeren. Deze algoritmen analyseren verschillende visuele en auditieve signalen om inconsistenties of anomalieën te detecteren die duiden op de aanwezigheid van een deepfake. Door gebruik te maken van machinaal leren en kunstmatige intelligentietechnieken kunnen deze algoritmen hun detectiecapaciteit na verloop van tijd verbeteren.
- **Blockchaintechnologie:** Een andere techniek stelt voor om blockchaintechnologie te gebruiken om de bron van media te verifiëren. De blockchain is een digitaal grootboek dat transacties over een netwerk van computers vastlegt op een veilige, transparante en fraudebestendige manier. Het maakt gebruik van decentralisatie en cryptografie om vertrouwen te garanderen zonder dat er een centrale autoriteit nodig is. In dit scenario moeten video's worden geverifieerd door een blockchain-grootboek voordat ze worden weergegeven op sociale mediaplatforms. Door ervoor te zorgen dat alleen video's van betrouwbare bronnen worden goedgekeurd, kan de verspreiding van potentieel schadelijke deepfake media worden beperkt.
- **Digitale handtekeningen:** Sommigen stellen voor om alle video's en beelden die met camera's zijn gemaakt, inclusief smartphonecamera's, digitaal te ondertekenen als middel om deepfakes tegen te gaan. Dit zou inhouden dat er unieke digitale handtekeningen worden toegekend aan elk stukje media, waardoor elke foto of video kan worden getraceerd naar de oorspronkelijke eigenaar. Hoewel deze aanpak kan helpen bij het traceren van de herkomst van content, zijn er zorgen over mogelijk misbruik, zoals het viseren van dissidenten of het schenden van privacy.

3.5 Verwijzingen

Aro, J. (2016). The Cyberspace War: Propaganda and Trolling as Warfare Tools. *European View*, *15*(1), 121–132. https://doi.org/10.1007/s12290-016-0395-5

Cambridge English Dictionary: Meanings & Definitions. (2023). https://dictionary.cambridge.org/dictionary/english

Daniels, J. (2018). The Algorithmic Rise of the "Alt-Right." *Contexts*, *17*(1), 60–65. https://doi.org/10.1177/1536504218766547

Egelhofer, J. L., & Lecheler, S. (2019). Fake news as a two-dimensional phenomenon: a framework and research agenda. *Annals of the International Communication Association*, *43*(2), 97–116. https://doi.org/10.1080/23808985.2019.1602782

Greene. (2019). "Deplorable" Satire: Alt-Right Memes, White Genocide Tweets, and Redpilling Normies. *Studies in American Humor*, *5*(1), 31–69. https://doi.org/10.5325/studamerhumor.5.1.0031

Hardaker, C. (2013). "Uh. . . . not to be nitpicky,,,,,but. . .the past tense of drag is dragged, not drug." *Journal of Language Aggression and Conflict*, *1*(1), 58–86. https://doi.org/10.1075/jlac.1.1.04har

IED. (2022, August 23). *How Do Social Media Algorithms Work*. Institute of Entrepreneurship Development. https://ied.eu/blog/technology-blog/how-do-social-media-algorithms-work/

Johnson, D., & Johnson, A. (2023, June 15). What are deepfakes? How fake AI-powered audio and video warps our perception of reality. *Business Insider*. https://www.businessinsider.com/guides/tech/what-is-deepfake?international=true&r=US&IR=T

Levy, G., & Razin, R. (2019). Echo Chambers and Their Effects on Economic and Political Outcomes. *Annual Review of Economics*, *11*, 303–328. https://doi.org/10.1146/annurev-economics-080218-030343

Lewis, B., & Marwick, A. E. (2017). Media Manipulation and Disinformation Online. *Data & Society Research Institute*. https://www.posiel.com/wp-content/uploads/2016/08/Media-Manipulation-and-Disinformation-Online-1.pdf

McNealy, J. (2015). Readers react negatively to disclosure of poster's identity. *Newspaper Research Journal*, *38*(3). https://doi.org/10.1177/0739532917722977

Munn, L. (2019). Alt-right pipeline: Individual journeys to extremism online. *First Monday*. https://doi.org/10.5210/fm.v24i6.10108

Sample, I. (2020, January 13). What are deepfakes – and how can you spot them? *The Guardian*. https://www.theguardian.com/technology/2020/jan/13/what-are-deepfakes-and-how-can-you-spot-them

Schumpe, B. M., Bélanger, J. J., Moyano, M., & Nisa, C. F. (2020). The role of sensation seeking in political violence: An extension of the Significance Quest Theory. *Journal of Personality and Social Psychology*, *118*(4), 743–761. https://doi.org/10.1037/pspp0000223

Seth, S. (2023, September 11). The World's Top 10 News Media Companies. *Investopedia*. https://www.investopedia.com/stock-analysis/021815/worlds-top-ten-news-companies-nws-gci-trco-nyt.aspx

Somers, M. (2020, July 21). Deepfakes, explained. *MIT Sloan*. https://mitsloan.mit.edu/ideas-made-to-matter/deepfakes-explained

Tandoc, E. C., Lim, Z. W., & Ling, R. (2017). Defining "Fake news": A Typology of Scholarly Definitions. *Digital Journalism*, *6*(2), 137–153. https://doi.org/10.1080/21670811.2017.1360143

Van Puffelen, M. (2021). Rechtsextremisme: Geweld met een rechtsextremistisch motie. In *DSP-groep*. DSP-groep. https://www.dsp-groep.nl/wp-content/uploads/18MP-Rechtsextremisme-DSP-2021.pdf

Van Wonderen, R. (2023). *Rechts-extremistische Radicalisering op Sociale Media Platformen*. Verwey-Jonker Instituut.

Van Wonderen, R. (2023). *Richtlijn / onderbouwing Radicalisering*. Verwey-Jonker Instituut.

Van Wonderen, R. & Peeters, M. (2021). *Werken aan weerbaarheid tegen desinformatie en eenzijdige meningsvorming. Evaluatie lesprogramma Under Pressure.* Utrecht: Verwey-Jonker Instituut. https://www.verwey-jonker.nl/wp-content/uploads/2022/04/120550_Werken-aan-weerbaarheid-tegen-desinformatie-eenzijdige-meningsvorming.pdf.

Wasike, B. (2022). When the influencer says jump! How influencer signaling affects engagement with COVID-19 misinformation. *Social Science & Medicine*, *315*, 115497. https://doi.org/10.1016/j.socscimed.2022.115497

Wolfowicz, M., Weisburd, D., & Hasisi, B. (2021). Examining the interactive effects of the filter bubble and the echo chamber on radicalization. *Journal of Experimental Criminology*, *19*(1), 119–141. https://doi.org/10.1007/s11292-021-09471-0