



Universiteit  
Leiden  
The Netherlands

## **Framing van deepfakes in de media en de casus van imitatie-Leonid Volkov**

Neurink, Yitsz

### **Citation**

Neurink, Y. (2024). *Framing van deepfakes in de media en de casus van imitatie-Leonid Volkov*.

Version: Not Applicable (or Unknown)

License: [License to inclusion and publication of a Bachelor or Master Thesis, 2023](#)

Downloaded from: <https://hdl.handle.net/1887/3714241>

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

Framing van deepfakes in de media en de casus van imitatie-Leonid  
Volkov

Masterscriptie voor Mediastudies, track Journalistiek en Nieuwe Media  
aan de Universiteit Leiden

Yitsz Neurink (s3678539)

Datum: 15 Juni 2023

Scriptiebegeleider: Alexander Pleijter

Tweede lezer: Peter Burger

## Inhoudsopgave

Abstract	pagina 3
Hoofdstuk 1: Introductie	pagina 4
Hoofdstuk 2: Theoretisch Kader	pagina 7
2.1: Deepfakes	pagina 7
2.2: Deepfakes in de media	pagina 8
2.3: Framing van nieuwe technologie	pagina 9
2.4: Afsluiting	pagina 11
Hoofdstuk 3: Methode	pagina 13
3.1: Opzet van het onderzoek	pagina 13
3.2: Inductieve frame-analyse	pagina 13
3.2.1: Proces van frame-analyse	pagina 13
3.2.2: Selectie van artikelen	pagina 15
3.3: Reconstructie berichtgeving	pagina 16
Hoofdstuk 4: Resultaten	pagina 17
4.1: Deepfakes in het Nederlandse nieuws	pagina 17
4.1.1: Algemene observaties over de artikelen	pagina 17
4.1.2: Framebundel van angstige toekomstbeelden	pagina 18
4.1.3: Framebundel van de kwetsbare samenleving	pagina 19
4.1.4: Framebundel van preventieve maatregelen	pagina 20
4.1.5: Framebundel van deepfaketoepassingen	pagina 21
4.1.6: Framebundel van beperkingen	pagina 22
4.1.7: Afsluiting van de frameanalyse	pagina 22
4.2: Reconstructie van de casus	pagina 25
4.2.1: Reconstructie door middel van inhoudsanalyse	pagina 25
4.2.2: Reconstructie door middel van interviews	pagina 29
Hoofdstuk 5: Conclusie en Discussie	pagina 31
Literatuurlijst	pagina 34
Bijlage 1: Lijst van artikelen gebruikt in frameanalyse	pagina 37
Bijlage 2: Lijst van artikelen en relevante bronnen over de zaak imitatie-Volkov	pagina 39
Bijlage 3: Transcripten interviews	pagina 43

## Abstract

Er is onderzoek gedaan naar de rol die bestaande beelden over deepfakes in de Nederlandse media speelden in berichtgeving over een specifieke casus die zich afspeelde in april 2021. Hierbij deed een lid van een Russisch bedriegersduo zich voor als de Russische oppositiepoliticus Leonid Volkov en sprak met de Vaste Kamercommissie voor buitenlandse zaken. Eerst rapporteerde de berichtgeving over deze casus dat de bedrieger gebruik maakte van deepfaketechnologie, maar dit bleek later niet zo te zijn. Deze scriptie onderzoekt hoe deze fout tot stand gekomen is en welke factoren daar een rol bij hebben gespeeld aan de hand van drie stappen. In de eerste stap is er een frameanalyse uitgevoerd over de Nederlandse berichtgeving over deepfakes in de periode 23 november 2020 tot 23 april 2021, waaruit bleek dat de framing van deepfakes in deze periode grotendeels negatief was. In de tweede en derde stap is een reconstructie van de berichtgeving over de casus uitgevoerd, respectievelijk door middel van een analyse van de berichtgeving zelf en door interviews met twee journalisten die over de casus schreven. Hieruit bleek dat meerdere factoren een rol speelden in het ontstaan van de fout, inclusief de bestaande hype over deepfaketechnologie, maar ook gebrek aan beeldmateriaal van de bedrieger en de tijdsdruk die in het algemeen binnen het journalistieke vak bestaat.

## Hoofdstuk 1: Introductie

Begin 2021, iets minder dan een jaar voordat Rusland Oekraïne zou binnenvallen, werden een aantal Europese politici benaderd door Leonid Volkov, een politicus uit de Russische oppositie en de stafchef van diens leider, Alexei Navalny. Later bleek het echter te gaan om een imitator van Volkov, die onder andere politici uit Letland, Estland, Litouwen en het Verenigd Koninkrijk om de tuin leidde. Volgens de initiële berichtgeving gebruikte de bedrieger deepfakefilters om Volkov na te bootsen (Roth, 2021).

In april 2021 had de Nederlandse vaste Kamercommissie voor Buitenlandse Zaken een live videogesprek met de imitatie-Volkov dat merkwaardig verliep. De media meldden eerst dat het ging om een deepfake-imitatie van Volkov die gebruik zou hebben gemaakt van deepfaketechnologie (Brouwers & Verhagen, 2021). Later bleek echter dat de bedrieger een lid was van het Russische satirische duo Vovan (Vladimir Kuznetsov) en Lexus (Aleksei Stolyarov). Lexus had geen deepfakefilters gebruikt om Volkov te imiteren, maar hij had zichzelf simpelweg geschminkt (Rottenberg, 2021). De initiële berichtgeving zat er dus naast over hoe Lexus zich voordeed als Leonid Volkov. Waar ging dat verhaal over deepfakes dan over?

Deepfakes zijn beeld- en geluidsmaterialen die gemaakt worden door kunstmatige intelligentie, of *artificial intelligence* (AI) in het Engels, een technologie die de laatste jaren een opkomst gemaakt heeft. Net zoals het internet de laatste decennia de samenleving doordrongen en veranderd heeft, zou AI het potentieel hebben om iets vergelijkbaars te doen (Khakurel et al., 2018). Online tools zoals *ChatGPT* en *Stable Diffusion* kwamen bijvoorbeeld op in de loop van 2022. Deze maken het gemakkelijk om snel met AI respectievelijk teksten en afbeeldingen (specifiek schilderijen en tekeningen) te maken. Dit kan alleen al door een korte prompt in te toetsen in het geval van deze specifieke tools. *ChatGPT* wordt bijvoorbeeld door leerlingen op middelbare scholen gebruikt om huiswerk te maken (Schellevis & Moerland, 2023).

Deepfakes hebben de laatste jaren ook veel discussie teweeg gebracht. Het gaat om audio, afbeeldingen of video's, vaak van personen, die gemaakt worden met AI. In de meeste gevallen worden deepfakes gebruikt om mensen dingen te laten zeggen (of doen) die ze niet gezegd (of gedaan) hebben (Weikmann & Lecheler, 2022). De naam van deepfakes komt ook van het feit dat het nepmedia ('fake') zijn die gemaakt zijn met behulp van *deep learning* ('deep'), een AI-techniek (Yadlin-Segal & Oppenheim, 2021). En net zoals *ChatGPT* en *Stable Diffusion* het eenvoudig maken om tekst en tekeningen te genereren, is het maken van deepfakes ook eenvoudiger geworden de afgelopen jaren. Waar een aantal jaar geleden er nog veel tijd, data en moeite nodig was om er één te maken, bestaan er tegenwoordig gratis smartphone apps die het eenvoudig maken (Thomson et al., 2020). Deepfakes hebben dus het potentieel om het neppen van mensen te vereenvoudigen.

Een ontwikkeling die zich de laatste jaren ook heeft voortgedaan is de opkomst van verschillende vormen van online mis- en desinformatie (vaak nepnieuws of *fake news* genoemd), die ook offline gevolgen gehad heeft. Door een online complottheorie genaamd "Pizzagate" liep in 2016 een man met een geweer een pizzeria in de Amerikaanse stad Washington binnen. Volgens de complottheorie zou deze pizzeria het hoofdkwartier zijn van een criminele organisatie geleid door Hillary Clinton die kinderen seksueel zou misbruiken. De man kwam naar de pizzeria om de theorie te bewijzen (Tandoc Jr. et al., 2017).

De ontwikkeling gaat echter niet alleen over informatie die compleet nep is, maar ook om informatie die op een misleidende manier neergezet wordt. In 2018 bleek dat het Britse bedrijf *Cambridge Analytica* gedetailleerde profielen van Facebookgebruikers had opgebouwd. Deze hielpen politieke campagnemakers om de Facebookgebruikers te bereiken met gepersonaliseerde reclame. Deze reclame bevatte vaak zeer partijdige boodschappen die inspeelden op de vooroordelen en angsten van het beoogde publiek, wat invasiever kan zijn dan informatie die compleet nep is (Heawood, 2018). De informatie die werd aangeboden was niet perse onwaar, maar hoe deze informatie neergezet werd of de framing ervan zorgde wel voor een misleidend karakter dat tot doel had om stemgedrag te beïnvloeden.

Als reactie op deze ontwikkelingen zijn initiatieven opgezet om de verspreiding van mis- en desinformatie tegen te gaan. Vanuit de journalistiek wordt bijvoorbeeld gedaan aan factchecken om het waarheidsgehalte van online gemaakte claims te meten, en dit lijkt effectief te zijn (Courchesne et al., 2021). Ook zijn er initiatieven binnen de journalistiek die zich specifiek richten op het tegengaan van nepfoto's, audio en video's, waar deepfakes onderdeel vanuit maken. Automatische tools zouden bijvoorbeeld kunnen helpen om te bepalen welke beelden echt zijn binnen de journalistiek (Katsounidou et al., 2020). Factchecken en beeldverificatie gaan vaak samen, maar redacties doen ook aan beeldverificatie om er voor te zorgen dat ze in hun artikelen geen nepbeelden plaatsen. Rond 2020 was er echter weinig onderzoek naar hoe journalisten te werk gingen met beeldverificatie. Uit het onderzoek dat wel was gedaan bleek dat dat werk vooral werd overgelaten aan individuele journalisten en dat het niet altijd goed of überhaupt niet gedaan werd. In veel gevallen kwam dit door tijdsgebrek. Ook werden bestaande tools voor beeldverificatie vaak niet gebruikt om dezelfde reden (Thomson et al., 2020).

Beeldverificatie is een relevant onderwerp voor de journalistiek in een samenleving waarin door AI gedreven mis- en desinformatie steeds vaker voorkomt. Er kan echter gevraagd worden hoe journalisten omgaan met beeldverificatie, hoe zij tegen deepfakes aankijken en hoe ze schrijven over deepfakes in hun artikelen. Het onderzoek van deze scriptie richt zich op het laatstgenoemde aspect in de context van de berichtgeving over een casus waarin aanvankelijk niet volledig duidelijk was of er deepfakes in het spel waren of niet. De specifieke casus gaat over de incidenten die aan het begin van deze scriptie genoemd werden: het Russische satirische duo dat zich voordeed als Leonid Volkov. Het gaat hier specifiek om het gesprek dat de imitatie-Volkov voerde met Nederlandse Tweede Kamerleden. Volgens de media zou Lexus deepfaketechnologie hebben gebruikt, maar dit bleek later niet zo te zijn. De vraag hier is dan hoe het kan dat de media eerst meldden dat er sprake was van een deepfake-imitatie. Hebben bestaande beelden over of het frame van deepfakes als bedreigende nieuwe technologie een rol gespeeld in deze fout?

Dit is waar het onderzoek van deze scriptie zich bezig mee houdt, aangezien beeldverificatie de laatste jaren een belangrijk deel van het journalistieke vak is geworden. De focus op deepfakes is gekozen omdat de opkomst hiervan onderdeel is van de algemene opkomst van AI en het potentieel heeft om het probleem van mis- en desinformatie te vergroten. Aan de andere kant schrijven Ternovski et al. (2022) dat volgens recente onderzoeken blijkt dat deepfakes geen of een geringe overtuigingskracht hebben. Ternovski et al. (2022) zelf onderzochten het effect van berichtgeving in de media over deepfakes die lezers aanraadden om kritisch te zijn op politieke video's. Dit zou weer kunnen leiden tot het afdoen van echte politieke video's als deepfakes. Het kan dus interessant zijn om te zien hoe Nederlandse journalisten aankijken tegen deepfakes en beeldverificatie in het kader van een specifieke casus waarin beide concepten een rol spelen. Hoe zetten Nederlandse journalisten deepfakes neer in hun artikelen? En waardoor is het fout gegaan in het geval van de casus van de

imitatie-Volkov?

Deze vragen haken in op de vraag die twee alinea's geleden gesteld werd, de hoofdvraag van deze scriptie. De vragen aan het eind van de vorige alinea kunnen als volgt op een algemenere manier gesteld worden en helpen om de hoofdvraag te beantwoorden: hoe werden deepfakes in de media geframed in het halfjaar dat vooraf ging aan de casus van de imitatie-Leonid Volkov? En welke specifieke factoren hebben geleid tot de initiële foute berichtgeving over deze casus als gekeken wordt naar de inhoud van deze artikelen en hoe de journalisten in kwestie er zelf op terugkijken? Relevante theorie over deze vragen en de hoofdvraag wordt behandeld in het hoofdstuk Theoretisch Kader, waarin deepfakes, de beeldverificatie in de media, en de theorie van framing besproken wordt. Daarna gaat het hoofdstuk Methode in op de onderzoeksmethode die gebruikt is om de deelvragen te beantwoorden. Het hoofdstuk Resultaten gaat in op de resultaten van het onderzoek. Ten slotte wordt in het hoofdstuk Conclusie en Discussie de hoofdvraag expliciet beantwoordt en de resultaten breder bediscussieerd.

## Hoofdstuk 2: Theoretisch Kader

Het doel van deze scriptie is om in kaart te brengen hoe Nederlandse journalisten aankijken tegen beeldverificatie en deepfakes en hoe dit hun berichtgeving beïnvloedt. Dit wordt gedaan aan de hand van de casus van een bedriegersduo dat in 2021 pretendeerde Leonid Volkov te zijn en met de vaste Kamercommissie voor Buitenlandse Zaken sprak. Aan dit onderzoek liggen een aantal begrippen ten grondslag die in dit hoofdstuk gedefinieerd worden in het kader van hoe ze in de scriptie gebruikt en toegepast worden. Sectie 2.1 gaat in op wat deepfakes zijn en welke verschillende soorten er bestaan. Sectie 2.2 gaat over de rol die deepfakes in de dagelijkse praktijk van de beeldverificatie spelen en visies over hoe dit in de toekomst zou kunnen veranderen. Sectie 2.3 tenslotte gaat over de theorie van het framen, die relevant is om de eerste deelvraag van deze scriptie te beantwoorden, en hoe nieuwe technologie in het bijzonder en deepfakes specifiek in de media worden geframed.

### 2.1: Deepfakes

Deepfakes zijn meestal nepvideo's die samengaan met nepaudio, maar soms ook nepfoto's. Het woord deepfake bestaat uit twee termen: 'deep' en 'fake'. 'Deep' komt van de term 'deep learning', een vorm van *machine learning*. Dat refereert naar technieken die er voor zorgen dat programma's taken op zichzelf kunnen uitvoeren en steeds beter worden in deze taken omdat ze van hun eigen ervaring kunnen leren (Anastasia et al., 2019). *Deep learning* specifiek stoelt op het gebruik van grote hoeveelheden data om te leren om taken uit te voeren (Anastasia et al., 2019). Deepfakes worden meestal gemaakt door veel echte foto's en nepfoto's in een *deep learning*-model te zetten en vervolgens dat model het verschil tussen de echte foto's en de nepfoto's te laten leren. Het doel is dan om foto's te genereren die zo goed mogelijk in de categorie van "echte foto" passen (Zheng et al., 2019).

'Fake' in het woord 'deepfake' verwijst naar het feit dat 'deepfakes' nepmedia zijn. In de literatuur over nepinformatie die online verspreid wordt, wordt er een onderscheid gemaakt tussen misinformatie en desinformatie. Het verschil tussen de twee komt neer op de gedachte van degene die de nepinformatie verspreidt: weet iemand niet beter, dan spreekt men van misinformatie, maar wordt het met opzet gedaan, dan spreekt men van desinformatie (Tandoc Jr. et al., 2017). Het hangt dus af van de context onder welk begrip deepfakes vallen. Bij de bron, als deepfakes gemaakt worden, gaat het om desinformatie, maar als ze verder verspreid worden door mensen die door de deepfakes genept zijn, gaat het om misinformatie.

Hiernaast is "deepfake" zelf ook een brede term en verwijst naar een aantal verschillende technieken. Sohrawadi et al. (2020) maken onderscheid tussen twee deepfakecategorieën: gezichtsverwisseling (*face-swapping*) en poppenspel (*puppet mastery*). Bij gezichtsverwisseling wordt het gezicht uit een bronvideo bovenop het gezicht in een doelwitvideo gezet. Poppenspel verwijst naar methoden waarbij een bronvideo gebruikt wordt om de uitdrukkingen en poses van iemand in een doelwitvideo te besturen. Hierdoor kan iemand in de bronvideo de persoon in de doelwitvideo als het ware als een pop bespelen. Farid et al. (2019) definiëren hiernaast nog twee deepfakecategorieën. Naast wat zijn gezichtsvervanging (*face replacement*) en gezichtsnaoenerij (*face re-enactment*) noemen, hebben Farid et al. (2019) het ook over het genereren van gezichten (*face generation*) en over spraaksynthese (*speech synthesis*). Bij dat eerste worden foto's van gezichten gemaakt van mensen die niet bestaan en bij dat laatste wordt iemands spraak nagebootst. Sohrawadi et al. (2020) doelen ook op spraaksynthese in hun bespreking van het poppenspel, waarbij zij stellen dat methoden die in die categorie vallen het mogelijk maken om mensen dingen te laten zeggen die ze in werkelijkheid niet gezegd hebben.



Voor het onderzoek in deze scriptie zijn de deepfakecategorieën die Sohrawadi et al. (2020) beschrijven het relevantst, want deze maken het vervalsen van video's en foto's van echt bestaande mensen mogelijk. Spraaksynthese is ook relevant, maar wordt in het onderzoek van deze scriptie gezien als onderdeel van deepfakes die binnen de categorie poppenspel vallen. Het genereren van niet-bestaande gezichten is minder relevant voor het onderwerp van deze scriptie, aangezien de mensen in dat soort deepfakes niet echt bestaan. Deze deepfakes kunnen wel een rol spelen in het verspreiden van mis- en desinformatie, bijvoorbeeld door te fungeren als profielfoto's van botaccounts die desinformatie verspreiden, maar dit fenomeen valt niet binnen de strekking van deze scriptie.

## 2.2: Deepfakes in de media

Als reactie op de opkomst van online mis- en desinformatie zijn er, zoals benoemd in de Introductie, in de journalistiek initiatieven gestart om daar tegen in te gaan. Er wordt aan beeldverificatie gedaan om te bepalen welke foto's en video's echt zijn en welke nep (of bewerkt). Deze vallen breed gezien in twee categorieën: het lezen van metadata om meer over een foto te weten te komen en het inspecteren van de foto zelf om te kijken of er mee geknoeid is (Katsaounidou et al., 2020). In de praktijk gaat het bij de tweede categorie vaak over foto's die op relatief simpele manieren bewerkt zijn zoals het weghalen van objecten door andere delen van de foto te kopiëren en over het te verwijderen deel van de foto te plakken. Deepfakes vormen dan ook een klein deel van foto's die gecheckt moeten worden in de dagelijkse praktijk van de beeldverificatie (Katsaounidou et al., 2020).

Er is in het algemeen minder onderzoek gedaan naar de dagelijkse praktijk van journalistieke beeldverificatie dan naar journalistieke verificatieprocessen over sociale media volgens Thomson et al. (2020), maar desondanks zijn er wel conclusies te trekken over beeldverificatie. In het algemeen zouden journalistieke reacties op online mis- en desinformatie bemoeilijkt worden door het feit dat nieuwsmedia economisch afhankelijk zijn van de platformen waarop die des- en misinformatie floreert, waar ook tijdsdruk bovenop komt door de eisen van de nieuwscyclus. Als het aankomt op beeldverificatie hebben journalisten ruwweg gezien twee opties: met het oog checken of er iets mis is met een afbeelding of detectiesoftware gebruiken die berust op *deep learning*-technieken. Beiden hebben volgens Thomson et al. (2020) beperkingen, want niet alle soorten gemanipuleerde afbeeldingen zijn met het oog te detecteren en detectiesoftware kan niet volledig het kritisch denkvermogen van de journalist overnemen. Over de effectiviteit van beeldverificatiemethodes in de journalistiek lijkt ook gering of geen onderzoek gedaan te zijn.

Aangezien deepfakes niet vaak in de dagelijkse praktijk voorkomen, richt veel van de wetenschappelijke literatuur die gaat over beeldverificatie zich op de relatief simpele vervalsingstechnieken en niet op deepfakes (Zheng et al., 2019). De angst die binnen de media over deepfakes bestaat gaat ook vooral over politieke deepfakes, maar er zijn ook andere domeinen waar ze al schade aanrichten. Er is bijvoorbeeld sprake van seksuele deepfakes waarbij nepporno gemaakt wordt die vervolgens de vrouwen schaadt die in deze deepfakes worden afgebeeld. De media focust zich in het algemeen veel minder op dit soort deepfakes (Gosse & Burkell, 2020). Dat er in de dagelijkse praktijk van de beeldverificatie weinig met deepfakes gedaan wordt, zou dus ook kunnen komen omdat er een focus is op een bepaald genre van deepfake: de politieke.

Ondanks het feit dat in de huidige dagelijkse praktijk van beeldverificatie nog weinig casussen te maken hebben met deepfakes, bestaat er wel een idee dat er een nieuw tijdperk aan zit te komen waarin deepfakes vaker zullen voorkomen (Katsaounidou et al., 2020; Temir, 2020). De meningen verschillen over hoe deze toekomst er uit zal gaan zien. Temir (2020) ziet dit als een *post-truth* tijdperk, waarin het onderscheid tussen waarheid en fictie niet meer of heel lastig te maken zal zijn. Silbey en Hartzog (2019) schrijven dat deepfakes geen nieuwe problemen zullen veroorzaken, maar dat ze bestaande problemen in de samenleving kunnen verergeren. De auteurs vinden dat de huidige problemen met mis- en desinformatie in de maatschappij nog niet hebben geleid tot een groot genoeg verlangen om grote hervormingen door te voeren in maatschappelijke instituten. Deepfakes zouden de problemen zo erg kunnen maken dat de samenleving als geheel hervormd zou moeten worden, omdat diepgewortelde waarheden over ons vertrouwen in wat we zien en horen onderuit gehaald zullen worden.

Aan de andere kant nuanceren, zoals in de introductie al benoemd, Ternovski et al. (2022) het effect van deepfakes. Volgens hun onderzoek en ander onderzoek dat zij citeren over het vermogen van deepfakes om mensen te neppen, is dat vermogen van deepfakes hetzelfde of iets groter dan andere vormen van online mis- en desinformatie. Als de resultaten van deze onderzoeken in acht genomen worden, dan kunnen er vraagtekens geplaatst worden bij het zwarte toekomstbeeld dat bestaat over het tijdperk van deepfakes.

### 2.3: Framing van nieuwe technologie

In de literatuur wordt dus op bepaalde manieren geschreven over deepfakes en hetzelfde geldt voor de media. Hoe de media schrijven over deepfakes in het bijzonder en nieuwe technologie in het algemeen, kan gezien worden als framing. De theorie van het framen, die gaat over hoe frames in teksten gebruikt worden, werd geformaliseerd door Robert Entman (1993) in de jaren negentig. Volgens het literatuuronderzoek van Entman werden termen als “frame”, “framing” en “framework” vaak gebruikt in het sociaalwetenschappelijke debat van voor 1993. Papers die de termen gebruikten kwamen echter niet vaak met een rigoureuze definitie en er werd aangenomen dat lezers wel wisten wat de termen inhielden. Entman gaf een concrete definitie voor deze vage concepten, gebaseerd op hoe ze in andere papers gebruikt werden. Volgens hem staan er twee aspecten centraal bij framen: selectie (*selection*) en opvallendheid (*saliency*). Er worden bepaalde aspecten van een onderwerp geselecteerd en die worden vervolgens opvallend (of minder opvallend) gemaakt om een bepaald beeld van het onderwerp te schetsen. Grotere opvallendheid kan bijvoorbeeld bereikt worden door zaken te herhalen of door ze te associëren met bekende culturele symbolen. Frames hebben vier functies: ze definiëren problemen, zoeken oorzaken achter die problemen, geven een moreel oordeel over de problemen en oorzaken en doen aanbevelingen om de problemen op te lossen (Entman, 1993). Baldwin van Gorp (2007a) noemt deze functies *reasoning devices* en ze kunnen apart van de frames gedefinieerd worden volgens hem.

Van Gorp (2007b) pleit ook voor een constructivistische kijk op de theorie van het framen. Hierin wordt een onderscheid gemaakt tussen enerzijds de frames en anderzijds de mediaberichten en het onderwerp waar zij over berichten. Frames verwijzen hier naar impliciete culturele fenomenen die worden uitgedragen door elementen in de tekst en nodigen uit om de tekst en diens onderwerp op een bepaalde manier te lezen en te interpreteren volgens Van Gorp. Het laat de deur open om te zien hoe een onderwerp ook anders neergezet zou kunnen worden. De focus bij deze constructivistische definitie van frames ligt dan ook op het koppelen van de frames in een tekst aan bredere culturele

symbolen en verwijzingen.

Frames zijn breed toepasbaar. Entman (1993) geeft bijvoorbeeld aan dat ze gebruikt kunnen worden om publieke opinie te duiden, inhoudsanalyse uit te voeren en journalistieke objectiviteit te analyseren. De theorie van Entman is ook opgepakt in het veld van mediastudies. In experimenten zoals die gedaan door Nelson et al. (1997) en Lee et al. (2008) blijkt dat de manier waarop onderwerpen in de media geframed worden invloed heeft op hoe het publiek tegen deze onderwerpen aankijkt en er over nadenkt. Dit speelt in op de rol die de media heeft als *agenda-setters*, waarbij de media bepaalt welke onderwerpen aandacht verdienen onder het publiek (Nguyen & Herkam, 2022). Zo wordt de publieke agenda bepaald en dit wordt gestuurd door welke onderwerpen opvallend gemaakt worden in de media, aangezien de media vooral bepaalt waar over gedacht wordt en niet perse hoe daar over gedacht wordt volgens de theorie van *agenda setting* (Marks et al., 2007). Hoeveel media-aandacht een bepaald onderwerp krijgt kan bepalen of het publiek er veel over nadenkt of niet, en de framing van deze onderwerpen kan weer invloed hebben op hoe ze er over nadenken.

Als het aankomt op berichtgeving over nieuwe technologieën in de media, zijn er patronen te zien, ook al zijn de nieuwe technologieën zelf verschillend: van nanotechnologie (Weaver et al., 2009) en biotechnologie (Marks et al., 2007) tot AI (Nguyen & Herkam, 2022). De drie geciteerde papers deden kwantitatief onderzoek naar de framing van deze technologieën in Amerikaanse en Engelse kranten over langere periodes (rond de 10 jaar) en kwamen tot vergelijkbare conclusies. Uit deze papers blijkt dat de kwantiteit van berichtgeving over nieuwe technologieën eerst langzaam op gang komt, voordat het erg toeneemt, waarna het na een aantal jaar weer in elkaar zakt. Voor nanotechnologie was deze boom ongeveer in de periode van 2003 tot 2008 (Weaver et al., 2009), voor biotechnologie ongeveer in de periode 1997 tot 2001 (Marks et al., 2007) en voor AI ongeveer van 2014 tot 2018 (Nguyen & Herkam, 2022). Nguyen en Herkam karakteriseren dit als een *hype-cycle*. De focus op technologie kan ook de berichtgeving van andere onderwerpen gaan domineren, zoals het geval was bij de berichtgeving over de Steubenvilleverkrachting in 2013. Hier ging het minder om de gebeurtenis zelf en meer over hoe deze werd vastgelegd met technologie (Pennington & Birthisel, 2016).

Als naar de papers van Weaver et al. (2009) en Nguyen en Herkam (2022) gekeken wordt, kan ook een patroon gevonden worden in de framing van nieuwe technologieën. Weaver et al. (2009) kwamen er in hun analyse achter dat in het begin van hun onderzochte periode (1999 – 2008) frames die de nadruk legden op het potentieel van nanotechnologie benadrukt werden in de media. Dit gold in mindere mate voor frames die de nadruk legden op verantwoordelijkheid van actoren die beleid over nanotechnologie konden beïnvloeden (zoals de regering en de markt). Dit veranderde na verloop van tijd, want in latere jaren werden meer frames gebruikt die de nadruk legden op regulatie en de wisselwerking tussen markt en regering. Volgens Nguyen et al. (2022) was er een verschuiving in de framing van AI in de media van AI als een abstract onderzoeksobject naar AI als een concreet iets met concrete sociale, culturele, economische en politieke gevolgen. Als de *hype-cycle* wat meer begint aan te slaan, lijkt er dus een verschuiving plaats te vinden van meer hypothetische en speculerende frames naar frames die de impact van de nieuwe technologie op de samenleving benadrukken.

Dit gaat ook samen met een verschuiving van positievere frames naar negatievere frames, aangezien de media ook de impact van nieuwe technologieën op de samenleving gaat benadrukken. Dit kan gezien worden als een verschuiving van positievere naar negatievere frames, maar is echter genuanceerder. Marks et al. (2007) laten dit zien in hun analyse over de berichtgeving van biotechnologie. Volgens hun onderzoek verschilde de verandering in framing per vorm van biotechnologie. Medische biotechnologie werd naar verloop van tijd positiever geframed, terwijl

agrarische biotechnologie naar verloop van tijd juist negatiever geframed werd. Het morele oordeel dat geveld wordt in de framing hangt dus ook af van de specifieke technologie die besproken wordt.

Als het aankomt op deepfakes, is de manier waarop journalisten deze nieuwe technologie neerzetten zeer negatief. Wahl-Jorgensen en Carlson (2021) analyseerden de journalistieke discoursen van nieuwsartikelen die zij verzamelden uit de online database *Nexis UK*, waarin artikelen van kranten uit het Verenigd Koninkrijk verzameld werden. Zij deden dit voor de periode januari 2018 tot juli 2019 en concludeerden dat de gevolgen van deepfakes als dystopisch neergezet worden. Deepfakes zouden een *arch-villain* vormen die de democratie zou ondermijnen en waar de journalistiek tegenover zou moeten staan om de dystopische toekomst te voorkomen. Volgens Wahl-Jorgensen en Carlson (2021) zijn er drie discoursen waar Engelse journalisten aan deelnemen als het gaat om deepfakes:

- Angst voor de toekomst en worst-case scenario's waarin deepfakes gebruikt worden om de democratie te ondermijnen en waarin video's en foto's niet meer vertrouwd worden. Dit gaat ook over speculaties over deze toekomst, die soms zeer specifiek kunnen zijn.
- Aannames die de grondslag vormen voor de hierboven beschreven worst-case scenario's. Dit gaat vooral over het idee dat de samenleving te weinig kritisch denkvermogen en mediawijsheid heeft om niet te vallen voor deepfakes. Er wordt aangenomen dat de samenleving hier kwetsbaar voor is.
- Debatten over hoe de dystopische toekomst voorkomen kan worden door middel van betere technologie, mediawijsheid en de rol van de journalistiek hierin.

Wahl-Jorgensen en Carlson nuanceren (2021) het zwarte beeld dat journalisten in hun stukken neerzetten over deepfakes ook. Ze stellen dat deepfakes niet de eerste of de laatste technologie is of zal zijn die als een grote bedreiging voor de journalistiek wordt neergezet. In de toekomst zullen er waarschijnlijk nieuwe *arch-villains* opduiken. Het geeft echter wel een beeld van hoe journalisten deepfakes framen en lijkt aan te sluiten bij de manier waarop delen van de in deze sectie besproken literatuur ook aankijkt tegen deepfakes.

#### 2.4: Afsluiting

De relatie tussen deepfakes en de media past dus goed binnen de relatie tussen nieuwe technologie en media in het algemeen. Volgens Nguyen & Herkam (2022) lag de piek van de hype-cycle over AI in de periode tussen 2014 en 2018 en gebaseerd op dat paper en dat van Weaver et al. (2018) is de verwachting dat de framing in de berichtgeving over AI sinds die periode een nadruk legt op de impact AI op de samenleving. Dit lijkt in ieder geval te kloppen voor berichtgeving over deepfakes als gekeken wordt naar het werk van Wahl-Jorgensen en Carlson (2021), waaruit blijkt dat Britse journalisten in 2018 en 2019 pessimistische toekomstbeelden schetsten over de impact die deepfakes in de toekomst zouden hebben. Deze factoren zijn relevant voor de eerste stap van het onderzoek in deze scriptie, dat de framing van deepfakes in Nederlandse berichtgeving onderzoekt tussen november 2020 en april 2021. Hier wordt gekeken naar soorten deepfakes die relevant zijn in het kader van online mis- en desinformatie: gezichtsverwisseling en poppenspel (en de daarbij horende spraaksynthese). Immers zijn niet alle vormen van deepfake bruikbaar om desinformatie te maken.

Wat ook bleek uit het werk van Wahl-Jorgensen en Carlson (2021) is dat de angst van deepfakes vooral geworteld zit in wat ze in de toekomst kunnen gaan doen, niet perse wat ze in het heden al doen. Dit komt overeen met de dagelijkse praktijk van beeldverificatie in de journalistiek, waarin deepfakes niet veel voorkomen en het vooral gaat om het opsporen van simpelere vormen van beeldvervalsing (Katsaounidou et al., 2020). De literatuur is eveneens angstvallig over de toekomst die deepfakes met zich mee zouden kunnen brengen, maar dit wordt genuanceerd door Ternovski et al. (2022) die stellen dat het effect van deepfakes relatief beperkt is. Of de angst voor deepfakes aanwezig is in Nederlandse berichtgeving is een relevant onderwerp voor de eerste stap van dit scriptie-onderzoek en zal ook als achtergrond bieden voor de tweede stap van het onderzoek: de reconstructie van de berichtgeving over imitatie-Volkov. De methode achter deze twee stappen zal in het volgende hoofdstuk van deze scriptie behandeld worden.

## Hoofdstuk 3: Methode

### 3.1: Opzet van het onderzoek

Dit hoofdstuk legt de methode uit die gebruikt is om de twee deelvragen die in de Introductie gesteld zijn te beantwoorden: Hoe werden deepfakes in de media geframed in het halfjaar dat vooraf ging aan de casus van de imitatie-Leonid Volkov? En welke specifieke factoren hebben geleid tot de initiële foute berichtgeving over deze casus als gekeken wordt naar de inhoud van deze artikelen en hoe de journalisten in kwestie er zelf op terugkijken? Deze vragen zijn op twee verschillende manieren beantwoord, omdat beide vragen gaan over twee verschillende aspecten van het onderwerp.

Sectie 3.2 gaat in op de beantwoording van de eerste deelvraag. Specifiek gaat het om kwalitatieve inhoudsanalyse door middel van een inductief proces. Hierbij zijn de dominante frames geïdentificeerd in Nederlandse berichtgeving in het half jaar voorafgaand aan de zaak van de imitatie-Volkov. De beschrijving van dit proces, dat bestaat uit een aantal stappen gedefinieerd door Baldwin van Gorp (2007a) en Linström en Marais (2012), staat centraal in sectie 3.2.1. Er is voor deze methode gekozen omdat het een concrete lijst oplevert van de frames en *reasoning devices* die journalisten inzetten om deepfakes neer te zetten. Het doel is om context te geven voor de berichtgeving over de casus van de imitatie-Volkov zelf en de lijst die het proces heeft opgeleverd is ook gebruikt voor de beantwoording van de tweede deelvraag.

Sectie 3.3 gaat daar op in. De beantwoording van de tweede deelvraag heeft zich in twee stappen voltrokken. De eerste stap bestond uit een kwalitatieve analyse van artikelen over de imitatie-Volkov, specifiek om meer te weten te komen over hoe de fout heeft kunnen ontstaan. Hoe kwamen de artikelen bijvoorbeeld tot de conclusie dat de bedrieger deepfakefilters gebruikte? Hoe kijken artikelen van na de onthulling dat het ging om een geschminkte bedrieger terug op de claims over deepfakes? De tweede stap bestond uit interviews met twee van de journalisten die in 2021 schreven over de casus om beter te begrijpen hoe de fout tot stand kwam. Hierbij werd gevraagd naar zaken die na de reconstructie op basis van de artikelen nog onduidelijk waren en naar hoe de specifieke berichtgeving van de geïnterviewde journalisten tot stand kwam.

### 3.2: Inductieve frame-analyse

#### 3.2.1: Proces van frame-analyse

Het eerste deel van het onderzoek is gedaan door het reconstrueren van deepfakeframes in artikelen die in het halve jaar voor de casus van de imitatie-Volkov gepubliceerd werden. Hiermee is duidelijk geworden welke beelden er in 2021 heersten over deepfakes in de Nederlandse media. Dit was een inductief proces waarbij de frames van de teksten werden gereconstrueerd en er werd uitgegaan van het idee dat deze frames de enige 'ware' frames zijn (Van Gorp, 2007a). In deze scriptie is specifiek gekeken naar de framing van deepfakes en de analyse heeft zich voltrokken volgens een stappenplan dat zich ontleent aan papers van Van Gorp (2007a) en Linström en Marais (2012).

De eerste stap in het proces hield zich bezig met het kiezen van een onderwerp en het selecteren van teksten die geanalyseerd moesten worden. Dit is ook de eerste stap die Van Gorp (2007a) noemt en deze kan verder onderverdeeld worden in de eerste drie stappen van Linström en Marais (2012): het kiezen van een onderwerp/medium, een tijdsbestek en het daadwerkelijke verzamelen van artikelen. Het onderwerp van deze scriptie is al uiteengezet in de Inleiding en in de vorige alinea herhaald. Het medium is Nederlandstalige kranten en andere nieuwspublicaties, waarbij relevante artikelen uit de online database *LexisNexis* verzameld zijn. Deze database verzamelt sinds 1979 artikelen van onder andere publicaties van *DPG Media* en *Mediahuis Nederland*, de twee grootste media-uitgevers van de

Benelux. Vanwege deze omvang is de database hiervoor geschikt en de keuze voor *LexisNexis* spiegelt zich ook aan het onderzoek van Wahl-Jorgensen en Carlson (2021), die hun analyse deden op basis van artikelen uit *Nexis UK*. De relevante periode is eerder in deze scriptie al gespecificeerd als het halve jaar dat voorafging aan de datum waarop de casus begon: 23 april 2021. Specifiek gaat het dus om de periode tussen 23 november 2020 en 23 april 2021.

Hierna volgde het selecteren van de analyse-eenheid (*unit of analysis*), de vierde stap volgens Linström en Marais (2012). Zij stellen dat voor frameanalyse deze eenheid vaak nieuwsartikelen zijn uit de relevante krant(en) en de geselecteerde tijdsperiode, wat ook voor deze scriptie het geval is. Er is geanalyseerd worden hoe deepfakes geframed worden per artikel door in te gaan op de inhoud van deze artikelen. Hierna volgde het open coderen, wat volgens Van Gorp (2007a) het systematisch doornemen van alle verzamelde teksten en het inventariseren van de elementen in die teksten die verband houden met het onderzoeksonderwerp (*framing devices*) is.

Stap 5 van Linström en Marais (2012) is hiermee vergelijkbaar, maar pleit om eerst een willekeurig gekozen monster te nemen uit de geselecteerde artikelen om tot een initiële lijst van frames te komen die daarna moet worden toegepast op de andere teksten (als een soort geschiktheidstoets). Aangezien de frame-analyse in deze scriptie vooral dient om context te bieden voor de reconstructie van de berichtgeving over de casus van de imitatie-Volkov, is Van Gorps methode in deze stap van het proces gevolgd. Hierom is in vervolgstappen ook vooral zijn methode gebruikt, aangezien Van Gorps methode en die van Linström en Marais (2012) op dit punt sterk divergeren.

Na de inventarisatie van relevante tekstelementen kwam de stap van het axiale coderen. Van Gorp (2007a) definieert dit als het stapsgewijs een steeds grotere afstand nemen van de primaire bronteksten naar een zeker abstractieniveau. Het gaat er hierbij om in de inventaris van tekstelementen patronen te zoeken die opgevat kunnen worden als individuele frames. Linström en Marais (2012) stellen dat er voor ieder artikel alleen een dominant frame of een dominant frame én een secundair frame, dat het dominante frame ondersteunt, geïdentificeerd moet worden. Dit is in acht genomen bij het zoeken naar frames in de relevante artikelen. De drie categorieën die Wahl-Jorgensen en Carlson (2021) identificeerden (zie pagina's 9 en 10 van deze scriptie) zijn in deze stap als uitgangspunt genomen. Andere relevante frames in de berichtgeving zijn ook benoemd.

Hiernaast zijn ook *reasoning devices* in de artikelen geïdentificeerd die overeenkomen met de kenmerken van frames die Entman (1993) identificeerde: Hoe worden onderwerpen op verschillende manieren gedefinieerd? Welke oorzaken worden aangeduid voor een onderwerp en hoe wordt dit moreel neergezet? Wie wordt aangewezen als verantwoordelijke voor een probleem en wat zijn aangedragen oplossingen?

Hierna kwam de stap van het selectieve coderen waarbij de frames en de *reasoning devices* in een matrix gezet werden (waarbij de frames de rijen vormen en de *reasoning devices* de kolommen). De matrix werd eerst per kolom ingevuld, waarna naar logische combinaties van frames gezocht werd. Uiteindelijk zijn een aantal bundels geïdentificeerd die geen elementen delen. Achteraf werden de bundels benoemd door ze in verband te brengen met bredere culturele symbolen en andere fenomenen. Voorbeelden hiervan zijn het verhaal van Don Quichot, dat Van Gorp (2007a) expliciet noemt, of het verhaal van de race tussen de haas en de schildpad. De framebundels ontleend aan Wahl-Jorgensen en Carlson (2021) zijn vernoemd naar de aspecten die zij in hun paper bespraken.

### 3.2.2: Selectie van artikelen

Zoals gesteld in de vorige sectie zijn artikelen geselecteerd uit de online database Lexis Nexis in de periode tussen 23 november 2020 en 23 april 2021. Er is gezocht naar artikelen waar de term “deepfake” in voorkwam, wat ongeveer 2000 artikelen opleverde. Veel van deze resultaten waren uit niet-relevante landen en in niet-relevante talen. Na het geografische gebied te specificeren als “Europe” en de taal te specificeren als “dutch”, bleven er 581 artikelen over. De database had ook de optie om artikelen die vaker dan één keer voorkwamen te verwijderen, waarna er 429 artikelen overbleven. Echter, het aanzetten van deze optie verwijderde lang niet alle duplicaten. Waarschijnlijk kwam dit omdat artikelen net een andere kop hadden of net een ander woordenaantal door kleine verschillen in de inhoud. Na deze automatische selectie heeft er een handmatige selectie plaatsgevonden om de artikelen met geduplicateerde headlines er verder uit te halen. Sommige artikelen hadden echter grote verschillen in woordenaantal, dus deze variaties zijn in de uiteindelijke selectie wel opgenomen. Verder bleek uit de handmatige selectie dat er ook artikelen over de casus van imitatie-Volkov tussen de resultaten zaten en die zijn niet meegenomen voor deze stap van het onderzoek. Ook was er één Duits artikel dat ook niet is meegenomen.

Na de handmatige selectie bleven er 205 artikelen over, die allemaal gedownload zijn. Dit bleek te veel voor het onderzoek in deze scriptie, dus er is een verdere selectie gemaakt. Relevante gegevens van deze artikelen zoals de datum van publicatie en het blad waar het artikel in stond zijn vervolgens geïnventariseerd en hierop gebaseerd is de verdere selectie gemaakt. Tijdens deze verdere selectie zijn er nog wat artikelen verwijderd omdat ze dezelfde kop en inhoud bleken te hebben, waarna er 182 artikelen overbleven.

Van deze 182 is een selectie van 60 artikelen, dus ongeveer een derde van het totale aantal, gemaakt om daadwerkelijk te analyseren. Er zijn voor elke maand van de halfjaarperiode (november tot april) 10 artikelen geselecteerd waarvan er 18 uit regionale kranten en 42 uit nationale kranten kwamen. Dit omdat van de totale 182 artikelen er bleek dat ongeveer 33% van de artikelen uit regionale kranten kwam en de rest uit nationale kranten. Bij de selectie van de 60 artikelen is ook gekozen om, waar mogelijk, voor elke krant die vaak terugkwam in de 182 artikelen één artikel per maand te hebben. Het gaat hier om de volgende publicaties: *De Volkskrant*, *NRC*, *Trouw*, *De Telegraaf*, *Het Parool*, het *Reformatorisch Dagblad* en *Metro*. Dit was echter niet altijd mogelijk, omdat elk van deze kranten niet altijd een artikel had in elke maand en daarom zijn in de uiteindelijke selectie sommige kranten vaker vertegenwoordigd en andere minder. Al deze keuzes zijn gemaakt om er voor te zorgen dat er voldoende variatie was in de geanalyseerde artikelen op het gebied van publicatiedatum, regionaal versus nationaal perspectief en politiek spectrum.

Nadat de analyse van de 60 artikelen begon bleek dat sommige artikelen duplicaten van elkaar waren met koppen die net iets anders waren. Ook bleek dat sommige van de artikelen helemaal niet over deepfakes gingen. De database van *Lexis Nexis* had deze artikelen in de resultaten opgenomen omdat verwijzingen gemaakt werden naar andere artikelen die wel over deepfakes gingen in secties als “Lees Verder” en dergelijke. Door deze factoren is het aantal artikelen verder afgenomen van 60 naar 57. Artikelen die een vergelijkbare inhoud hadden maar een andere kop zijn wel in de uiteindelijke analyse opgenomen omdat de verschillende koppen iets kunnen zeggen over hoe verschillende kranten hetzelfde artikel duiden. De uitwerking van de analyse kan gevonden worden in sectie 4.1.



### 3.3: Reconstructie berichtgeving

Het tweede deel van het onderzoek, dat de tweede deelvraag beantwoordt, is gedaan door middel van een reconstructie van berichtgeving over de casus van de imitatie-Volkov. Dit volstreekte zich in twee stappen waarbij het ging om er achter te komen waarom media eerst berichten dat de imitatie-Volkov gebruik maakte van deepfakefilters, terwijl de bedrieger zich gewoon geschminkt had.

De eerste stap was een reconstructie van de berichtgeving door middel van een casestudie met een *single-case design*. Er was hiervoor gekozen omdat de casus waar deze scriptie om draait een uniek geval is in de berichtgeving over deepfakes: een politiek relevante gebeurtenis waarbij er eerst werd gemeld dat er deepfakes in het spel waren en dat dit later niet zo bleek te zijn. Het ging er hierbij om te kijken naar wat er uniek was aan deze casus en waardoor de fout heeft kunnen plaatsvinden (Aberdeen, 2013). Het gaat hier specifiek om een chronologische ordening waarbij de berichtgeving onder de casus van de imitatie-Volkov en andere relevante gebeurtenissen op een rij gezet werden om een geschiedenis van de casus te schetsen. Het hoofddoel hier was om te verklaren waar de fout in de berichtgeving gemaakt is en hoe hier later op gereageerd werd (Aberdeen, 2013).

Hiervoor zijn zo veel mogelijk relevante artikelen verzameld die berichten over de casus (zie Bijlage 2 voor een lijst van de artikelen). Na de zoektocht is de inhoud van deze artikelen geanalyseerd om duidelijk te maken hoe de artikelen berichten over deepfakes in het kader van de casus. Waar ontstond de claim dat het ging over deepfaketechnologie? Hoe komen de artikelen tot de conclusie dat het ging over deepfaketechnologie? Welke bronnen worden voor beide zaken genoemd? Hoe evolueerde de berichtgeving toen duidelijk werd dat er geen deepfaketechnologie in het spel was? Wanneer werd dat überhaupt duidelijk? De uitwerking van de reconstructie kan gevonden worden in sectie 4.2.1 van deze scriptie.

De tweede stap was het interviewen van twee journalisten die in 2021 berichtten over de casus van de imitatie-Volkov. De twee journalisten, Laurens Verhagen van *De Volkskrant* en Mark Hendrikman van *Tweakers*, waren hiervoor uitgekozen na de voltooiing van de eerste reconstructiestap. Zij zijn uitgekozen omdat zij meerdere artikelen hadden geschreven over de casus en omdat in de eerste stap van de reconstructie bleek dat hun artikelen nieuwe informatie verschaften die daarvoor nog niet bekend was. De Volkskrantartikelen stonden aan de voorhoede van de casus en waren bepalend voor veel van de andere berichtgeving over de casus (zie sectie 4.2.1) en de Tweakersartikelen berichtten over aspecten die elders in de berichtgeving niet aan bod kwamen.

De interviews zelf waren relatief korte, semigestructureerde interviews die vooral dienden om duidelijk te krijgen hoe de geanalyseerde berichtgeving uit de eerste stap tot stand kwam en waar de fout gemaakt werd. Er waren nog een paar dingen onduidelijk over de casus na de eerste reconstructiestap en de interviews dienden om een aantal van deze gaten te vullen. Er is bijvoorbeeld gevraagd of en hoe deze journalisten aan beeldverificatie deden, hoe zij tot het onderwerp kwamen en hoe er intern is gereageerd op de fout. Beide interviews zijn opgenomen en volledig uitgeschreven in Bijlage 3. Het doel van de interviews was om breed gezien te reconstrueren hoe een aantal van de artikelen, fout en al, tot stand zijn gekomen als aanvulling op de reconstructie van de berichtgeving uit de eerste stap. De uitwerking hiervan kan gevonden worden in sectie 4.2.2 van deze scriptie.

## Hoofdstuk 4: Resultaten

Dit hoofdstuk zet de resultaten uiteen aan de hand van de onderzoeksstappen beschreven in het vorige hoofdstuk over de Methode. De resultaten van de eerste stap van het onderzoek zijn beschreven in sectie 4.1. De framebundels staan hierin beschreven om duidelijk te maken hoe deepfakes werden neergezet in de Nederlandse media in de relevante periode. De resultaten van de tweede onderzoeksstap staan beschreven in sectie 4.2, waar sectie 4.2.1 fungeert als de reconstructie van de berichtgeving en sectie 4.2.2 gaat over de interviews die de productiekant reconstrueren.

### 4.1: Deepfakes in het Nederlandse nieuws

#### 4.1.1: Algemene observaties over de artikelen

Na de selectie van artikelen beschreven in sectie 3.2.2 zijn de verdere stappen doorlopen die beschreven staan in sectie 3.2.1 om tot een aantal framebundels te komen die laten zien hoe deepfakes in de Nederlandse media worden neergezet. Er zijn in totaal vijf framebundels geïdentificeerd, waarin frames verbonden worden aan *reasoning devices*: angstig toekomstframe, kwetsbaarheidframe, preventieframe, toepassingsframe en beperkingsframe. De eerste drie van deze frames zijn ontleend aan de discoursen geïdentificeerd door Wahl-Jorgensen en Carlson (2021), die op pagina 10 van deze scriptie kort zijn beschreven. De rest van deze sectie van de scriptie beschrijft de framebundels en geeft de matrix van *reasoning devices* en frames, maar eerst zullen een aantal algemene observaties over de geanalyseerde artikelen beschreven worden in deze sectie. Een complete lijst van de geanalyseerde artikelen kan gevonden worden in Bijlage 1.

Ruwweg zijn de 57 artikelen in drie categorieën te verdelen. De eerste categorie bestaat uit artikelen die gaan over specifieke gebeurtenissen die te maken hebben met deepfakes. Denk hier bijvoorbeeld aan een *Zondag met Lubach* uitzending met een item over deepfakes, acteur Lykele Muus die dacht een deepfakevideo ontvangen te hebben of toenmalig *NPO*-voorzitter Shula Rijxman die waarschuwde voor de toenemende dreigementen jegens journalisten. De tweede categorie bestaat uit algemenere artikelen die waarschuwen voor de invloed van nepnieuws en deepfakes en dieper op één van beide onderwerpen ingaan. Hierbij worden deepfakes vaak behandeld als deel van het bredere fenomeen van nepnieuws. Deze tweede categorie bestaat voor een groot gedeelte uit opiniestukken. Soms vallen deze categorieën ook samen, waarbij auteurs een bepaalde gebeurtenis gebruiken om een algemeen verhaal over het gevaar van nepnieuws of deepfakes te vertellen. De derde en kleinste categorie bestaat tenslotte uit artikelen die verder niet over deepfakes gaan, maar waarin deepfakes in vogelvlucht genoemd worden in relatie tot een breder onderwerp.

Verder zit er in het algemeen weinig verschil in de twee dimensies waar de 57 artikelen op geselecteerd zijn: regionaal versus nationaal en de tijdsdimensie. Er zit weinig verschil in de berichtgeving over deepfakes tussen regionale kranten en nationale kranten. In tegendeel, sommige artikelen komen in beide soorten kranten voor met een andere kop, waarschijnlijk omdat de kranten in kwestie onder dezelfde uitgever vallen of omdat deze artikelen geschreven zijn door het *Algemeen Nederlands Persbureau*. Ook door de tijd heen is er weinig verschil in hoe deepfakes worden neergezet in deze artikelen: er zit geen duidelijke ontwikkeling of verandering in de berichtgeving in de onderzochte periode.

#### 4.1.2: Framebundel van angstige toekomstbeelden

De eerste framebundel, die van de angstige toekomst, is ontleend aan Wahl-Jorgensen en Carlson (2021). Binnen deze bundel wordt de zwarte toekomst beschreven waar overmatig gebruik van deepfakes toe zou kunnen leiden of er worden algemene verwijzingen gemaakt naar zulke toekomstbeelden, zonder dat deze expliciet worden ingevuld. Als gekeken wordt naar hoe deze toekomst worden neergezet in Nederlandse berichtgeving, dan komen een aantal termen terug die min of meer dezelfde lading hebben: *post-truth*, *post-reality* en de infocalyps. Deze beschrijven samenlevingen die niet meer geloven in één gedeelde waarheid of realiteit; letterlijk gesproken post-waarheid en post-realiteit.

Er zijn twee specifieke angsten die in de artikelen naar voren komen en die soms ook samen genoemd worden. De eerste is dat iedereen in zijn of haar eigen werkelijkheid gaat leven en dingen voor waar aanneemt die niet waar zijn maar wel overeenkomen met het eigen wereldbeeld. Soms worden hier ook voorbeelden van gegeven door reportages over specifieke gebeurtenissen, zoals in het artikel “Radicale Trump-fans geloven niet dat veroordelende video president echt was”. Hierin wordt genoemd dat bepaalde Trump-aanhangers Trumps afwijzing van de capitolbestormingen op 6 januari 2021 neerzetten als een deepfake. De tweede is de angst dat machtige mensen zoals politici dingen die niet goed voor hen zijn, maar wel echt gebeurt zijn, weg kunnen zetten als deepfakes. Volgens de artikelen zal door deze twee ontwikkelingen een samenleving ontstaan waarin niks meer te vertrouwen is. Deze zouden weer ontstaan doordat nepnieuws, en specifiek vaak ook deepfakes die als de apex van nepnieuws neergezet worden, de samenleving doordringen. Vandaar ook de term “infocalyps”.

Het zwarte toekomstbeeld wordt dus impliciet neergezet als een “apocalyps” en hoe deze er zal komen wordt soms uitvoerig beschreven. De framebundel van de angstige toekomst fungeert vandaar ook vooral om de problemen die deepfakes kunnen veroorzaken neer te zetten, als de framebundel wordt bekeken vanuit het perspectief van *reasoning devices*. Er wordt, in het kort, gewaarschuwd voor wat er kan gebeuren als er niks gedaan wordt. Dit kan ook gezien worden in de manier waarop de artikelen deepfakes in het algemeen soms beschrijven. Dit gebeurt vaak in de toekomstige tijd: een artikel stelt dat vertrouwen in wat echt is “over een tijdje” zal afnemen, een ander artikel schrijft dat deepfakes “in opkomst” zijn en een derde dat “er verwacht wordt” dat deepfakes de toekomst op zijn kop gaat zetten. Het sentiment lijkt te zijn dat deepfakes hun volle potentieel nog niet bereikt hebben en als dat moment er aan komt, dat de problemen dan echt beginnen. Dit komt overeen met de waarschuwendende functie die deze framebundel heeft.

Een ander aspect dat hierbij opvalt is dat in verhalen over de zwarte toekomst, nepnieuws als algemeen concept vaak wordt aangedragen als brenger van de zwarte toekomst. Deepfakes vormen vaak maar een klein onderdeel van deze verhalen. Er zijn echter wel artikelen die ingaan op hoe deepfakes specifiek een zwarte toekomst kunnen brengen, maar deze zijn in de minderheid. Dit sluit aan op de waarschuwendende functie van de framebundel, maar ook op het feit dat deepfakes gezien worden als deel van het bredere fenomeen “nepnieuws”. Deepfakes hebben hun volledige potentieel nog niet bereikt, maar als ze dat wel doen, zullen ze de apex van het nepnieuws zijn.

#### 4.1.3: Framebundel van de kwetsbare samenleving

De tweede framebundel is ook ontleend aan Wahl-Jorgensen en Carlson (2021) en betreft aannames die ten grondslag liggen aan de zwarte toekomstbeelden van de eerste framebundel. Er wordt aangenomen dat de samenleving kwetsbaar is en dat de mens niet op zou kunnen tegen technologie. Concreet gebeurt dit vooral in de vorm van angst dat er te weinig mediawijsheid in de samenleving is en dat deepfakes in de nabije toekomst niet meer van echte beelden te onderscheiden zullen zijn. In de artikelen worden met concrete casussen ook voorbeelden gegeven van de bredere framebundel zoals de casus van Lykele Muus. Centraal in dit verhaal is dat hij van iemand een video van Mark Rutte ontving waarin de minister-president hem vertelt dat hij uit moet gaan met de zender van de video. Aanvankelijk dacht Muus dat de video een deepfake was, maar na latere verificatie bleek het om een echte video te gaan. Deze casus past ook in een trend binnen deze framebundel waar Mark Rutte vaak als voorbeeld wordt gebruikt. Er wordt dan gesteld dat je hem alles kunt laten zeggen, waarschijnlijk om het gevaar van deepfakes te illustreren als ontwrichtend voor de democratie en de samenleving, aangezien Mark Rutte de politiek leider van Nederland is.

Als het aankomt op problemen in het heden is er ook sprake van een aparte categorie deepfakes, die van de deepnude. Dit zijn deepfakes van naakte mensen waar gezichten van andere mensen op geplakt zijn of beelden die gemaakt zijn door geklede beelden van mensen als het ware uit te kleden (Keultjes, 2023). Als artikelen schrijven over deze vorm van deepfake focussen zij op de vervelende gevolgen die deepnudes kunnen hebben voor de slachtoffers. Er worden hierbij vaak vergelijkingen gemaakt met wraakporno. Dit onderwerp was wat lastig te plaatsen binnen de framebundels, waarschijnlijk omdat artikelen vaker focussen op de politieke gevolgen van deepfakes dan de impact die deepfakes in het heden al hebben op vrouwen in de vorm van deepnudes, zoals Gosse en Burkell (2020) al aanstipten. Berichtgeving over deepnudes gaat dus ook niet over zwarte toekomstbeelden, maar dragen wel bij aan het idee dat deepfakes het onderscheid tussen echt en nep zullen doen vervagen. Dit doordat er wordt benadrukt dat deepnudes goed genoeg zijn om slachtoffers veel nare gevolgen te brengen. Dit grijpt weer terug op de aannames over waarom de samenleving kwetsbaar zou zijn die besproken zijn in de vorige alinea en daarom is het fenomeen van de deepnude ondergebracht bij de framebundel van de kwetsbare samenleving in deze scriptie.

Verder gaat het in de framebundel over de redenen die schuilen achter de opkomst van deepfakes (en nepnieuws). Als dit aan bod komt, worden een aantal redenen gegeven: geldzucht, machtszucht van politici, grappenmakerij, omdat mensen zich vergissen en om “te pesten” (aldus een artikel gericht op de jeugd). Soms wordt de schuld ook bij specifieke partijen zoals de Russische regering en het verdienmodel van sociale mediabedrijven gelegd. Dit gebeurt echter vooral binnen de artikelen die dieper ingaan op deepfakes of nepnieuws. Vaker leggen de artikelen de verantwoordelijkheid van de opkomst van deepfakes bij het beter worden van technologie. In deze gevallen wordt geconstateerd dat computers steeds sneller en beter worden en het steeds makkelijker wordt om dingen te doen met computers, zoals het maken van deepfakes. Deze artikelen graven niet dieper en duiden niet wat deze technologische ontwikkelingen drijft of wie er baat bij heeft.

Ook wordt er binnen deze framebundel ook een focus gelegd op hoe “nep” en “niet authentiek” deepfakes zijn, met een impliciet of soms expliciet oordeel dat dit ondergeschikt is aan de “authentieke” werkelijkheid. De term synthetische media, die wat neutraler is dan de term deepfake, wordt soms ook gebruikt maar vaak tussen aanhalingstekens (“synthetische media”), wat gedaan lijkt te worden om twijfel te zaaien over deze specifieke term. Dit alles speelt in op de aanname dat deze neppe media manipulatief zijn en het verschil tussen nep en echt doen vervagen, waarbij de term

synthetische media verdacht gemaakt moet worden om de ernst van de zaak te onderstrepen. Als gekeken wordt naar *reasoning devices*, valt deze framebundel in de hoek van het aanstippen wat de problemen zijn en wat de oorzaak is van andere problemen. De zaken die over deze framebundel besproken zijn, zoals het vervagen van het onderscheid tussen nep en echt en de redenen die schuilen achter de opkomst van deepfakes, vormen in het heden al problemen. Tegelijkertijd dienen ze ook als specifiekere verklaringen achter de zwarte toekomst die er aan zit te komen. Die framebundel gaat vooral over wat problemen in de toekomst zullen zijn en dus de problemen in het heden liggen daar aan ten grondslag als oorzaken. Beide factoren onderstrepen dat de mens en de samenleving kwetsbaar zijn tegenover technologie in het heden en de toekomst. Verder worden die huidige problemen ook verklaard binnen de framebundel van de kwetsbare samenleving, dus fungeert deze bundel verder als verklaarder. Ook fungeert de bundel om aan te stippen wie de verantwoordelijken zijn voor de problemen, wat ook hoort bij het *reasoning device* van de oorzaak achter een probleem.

Deze framebundel heeft ten slotte ook een verbinding met het *reasoning device* van het morele oordeel. Hierbij worden deepfakes als “gevaarlijk” neergezet: beschreven als “eng”, “gevaarlijk” en “echte vijanden”. Relevant is dat dit gebeurt zonder dat er verder veel context voor gegeven wordt. De schrijvers van de artikelen lijken aan te nemen dat de lezer al weet waarom deepfakes gevaarlijk zijn of dit wordt elders in de desbetreffende artikelen uitgelicht. Deepfakes worden als zeer negatief weggezet en dit heeft ook connecties met aannames over hoe kwetsbaar de samenleving is: deepfakes zijn gevaarlijk en vijanden, want de mens kan niet tegen ze op. En dit is weer een aanname die ten grondslag ligt aan angstige toekomstbeelden, aangezien dat het gevaar van deepfakes in bepaalde artikelen in de toekomstige tijd beschreven wordt.

#### 4.1.4: Framebundel van preventieve maatregelen

De derde framebundel is de laatste die ontleend is aan Wahl-Jorgensen en Carlson (2021) en gaat over discussies over maatregelen om het zwarte toekomstbeeld uit de eerste framebundel te voorkomen. Deze framebundel is dan verbonden met het *reasoning device* van hoe de aangedragen problemen in heden en toekomst opgelost en voorkomen kunnen worden. Er worden concrete maatregelen aangedragen binnen deze framebundel en hier worden ook oordelen over gedaan, aangezien het gaat over een debat. De maatregelen vallen algemeen gesproken binnen drie categorieën. De eerste pleit voor meer mediawijsheid door middel van bewustwordingscampagnes onder het publiek (sommige van de geanalyseerde artikelen hebben dit ook als functie) en goede journalistiek, die tegenover nepnieuws en deepfakes wordt neergezet in dit kader. Journalistiek zou kunnen bijdragen aan het brengen van “goede” informatie. De tweede categorie pleit voor juridische maatregelen tegen de verspreiders en producenten van deepfakes en ook voor het strafbaar stellen van het maken deepfakes in het algemeen of bepaalde vormen (zoals deepnudes) door middel van wetgeving. Deze maatregelen worden het vaakst genoemd in het kader van de Nederlandse regering, maar andere artikelen deden ook bericht over de *AI Act* van de Europese Commissie van 21 april 2021. In de laatste categorie wordt technologie als een optie aangedragen om het onderscheid te maken tussen echte beelden en deepfakes. Het artikel “Bewijs van echtheid” pleit zelfs voor een “echtheidscertificaat” op beelden dat werkt als de voormalige blauwe vinkjes op Twitter.

Over het algemeen worden al deze maatregelen in de artikelen positief beoordeeld en aangedragen als reële oplossingen voor de problemen die deepfakes met zich mee kunnen brengen. Dit is vooral te zien in de berichtgeving over de Europese *AI Act*, waarin politici aan het woord komen die stellen dat de Europese Unie de eerste in de wereld is met zulke wetgeving (wat ook in de koppen van deze artikelen staat). Andere gebieden van de wereld zouden er een voorbeeld aan kunnen nemen in het beschermen van burgers tegen te ver gaande technologie volgens deze artikelen.

Er is echter ook kritiek op de aangedragen maatregelen, alhoewel in mindere mate. Experts die aan het woord gelaten worden in de artikelen over de *AI Act* stellen ook dat de daadwerkelijke uitvoering van de wet te ingewikkeld is en dat de wet te veel uitzonderingen open laat. Ook stellen sommige artikelen dat wetgeving langzaam is en achter de feiten aanloopt, terwijl de technologie zich steeds verder en sneller ontwikkelt. Deze kritiek richt zich vooral op wetgevende maatregelen, maar ook mediawijsheid krijgt kritiek als oplossing. Zo zou het niet helpen, net zoals het internet ooit mediawijsheid zou brengen volgens het artikel “Bellingcat: wereldkampioen digitaal speuren”. Dit noemt deze gedachte “naïef”. Een laatste artikel stelt dat er in het algemeen niks te doen is aan de opkomst van deepfakes.

#### 4.1.5: Framebundel van deepfaketoepassingen

De vierde framebundel gaat over gevallen waarbij de toepassingen van deepfakes buiten het neppen van mensen en het maken van deepnudes worden geïllustreerd en besproken. Deze framebundel vormt een breuk met de eerste drie aangezien het over het algemeen veel positiever is over de technologie dan de framebundels ontleend aan Wahl-Jorgensen en Carlson (2021). Uit de artikelen komt naar voren dat de toepassingen die deepfakes kunnen hebben vrij divers zijn. Individuele artikelen focussen zich het meest op één specifieke toepassing, maar als naar de artikelen als geheel wordt gekeken, dan komt het diverser beeld naar voren, aangezien twee artikelen bijna nooit over dezelfde toepassing berichten. Voorbeelden zijn deepfakes inzetten om satire zoals die van *Zondag met Lubach* te maken, angsttherapie om mensen te helpen met zaken als hoogtevrees en het tot leven brengen van historisch beeldmateriaal.

Dit frame is vooral verbonden met de *reasoning devices* van de probleemstellingen en de probleemoplossing. Er wordt een bepaald probleem aangedragen, zoals dat mensen met irrationele angsten de behandeling voor die angsten niet meer vergoed krijgen door hun zorgverzekeringen, en deepfakes worden dan als oplossing aangedragen door middel van *immersive* angsttherapie. Op dit vlak is, zoals geconstateerd in de vorige alinea, deze framebundel vrij positief over deepfakes. De artikelen waarbij het toepassingsframe dominant is hebben het echter ook vaak over de gevaren van deepfakes. Dit gebeurt vaak kort zonder verdere uitleg en soms wordt er meer ingegaan op de gevaren, maar van belang is dat het bijna nooit ongenoemd blijft.

Ook worden deepfakes soms gebruikt om hun eigen gevaren te illustreren in satirische programma's. Deze programma's, zoals *Zondag met Lubach* of de uitzending met een deepfake-imitatie van toenmalig koningin Elizabeth II van het Britse *Channel 4*, worden als grappig beschreven in de artikelen maar tegelijkertijd illustreren de programma's zelf ook het gevaar van deepfakes door middel van uitleg. De artikelen die berichten over deze programma's noemen ook vaak de “grappige” en “gevaarlijke” kant van deepfakes in dezelfde zin. Ook zijn er twee artikelen die een tentoonstelling over deepfakes neerzetten als een “cultureel uitje”, waarbij de angst voor deepfakes een toepassing op zich is voor entertainment. Het toepassingsframe zelf is dus grotendeels positief, maar gaat vaak gepaard met andere, negatievere frames die aan een andere probleemstelling doen of een moreel oordeel geven.

De toepassingsframebundel is echter niet altijd positief, aangezien artikelen zich soms ook negatief uitlaten over de manier waarop deepfakes worden toegepast. Er worden kanttekeningen geplaatst bij het toepassen van de technologie, voornamelijk dat het de problemen die het pretendeert op te lossen niet (volledig) oplost of dat de technologie nog niet geavanceerd genoeg is. Bij de angsttherapie bijvoorbeeld wordt geconstateerd dat deepfakes waarschijnlijk nooit helemaal angsttherapie zullen vervangen en vooral een eerste stap in het proces zullen vormen. Een specifiek artikel, “Laat techniek niet flirten met de dood”, is zeer negatief over deepfaketoepassingen en valt buiten het algemene patroon van positieve berichtgeving. Het gaat hierbij om mensen die spreken met een rouwtherapeut die zich voordoeft als een overleden nabestaande door middel van deepfaketechnologie. In het desbetreffende artikel uit het *Reformatorisch Dagblad* wordt de nadruk gelegd hoe on-Bijbels het is en hoe het mensen voor de gek houdt. Dit is niet zozeer deel van de eerste twee framebundels, aangezien er geen zwart toekomstbeeld geschetst wordt en dat dit onderwerp ook niet past binnen de aannames over de kwetsbare samenleving onder die zwarte toekomstbeelden.

#### 4.1.6: Framebundel van beperkingen

De laatste framebundel gaat over situaties waarin de nadruk gelegd wordt op de beperkingen van deepfakes. Dit frame is het vaakst dominant in artikelen over het programma *RTL Exclusief*, waarin gebruik werd gemaakt van deepfakes om grappen te maken. Dit programma wordt in de relevante artikelen negatief neergezet en de blaam wordt deels bij deepfakes gelegd: de technologie is niet goed genoeg om grappig te zijn en is ook niet voldoende op het werk van de showmakers over te nemen (aangezien de grappen in het programma negatief beoordeeld worden). In andere artikelen worden de beperkingen van deepfakes in het algemeen geïllustreerd, bijvoorbeeld door te illustreren hoe ver ze nog afwijken van echte afbeeldingen. Deze framebundel is dus vooral verbonden met het *reasoning device* van het morele oordeel. Deepfaketechnologie is nog gebrekkig en laat te wensen over. Er kan hier ook een verbinding gemaakt worden met de kwetsbaarheidsframebundel en ook de framebundel van de toepassingen: deepfakes zijn in het heden gebrekkig en hebben een beperkte waarde als het aankomt op toepassingen, en dit is ook een aanname die ten grondslag ligt aan zwarte toekomstbeelden: deepfakes zijn nu nog geen groot gevaar, maar kunnen dat in de toekomst wel worden.

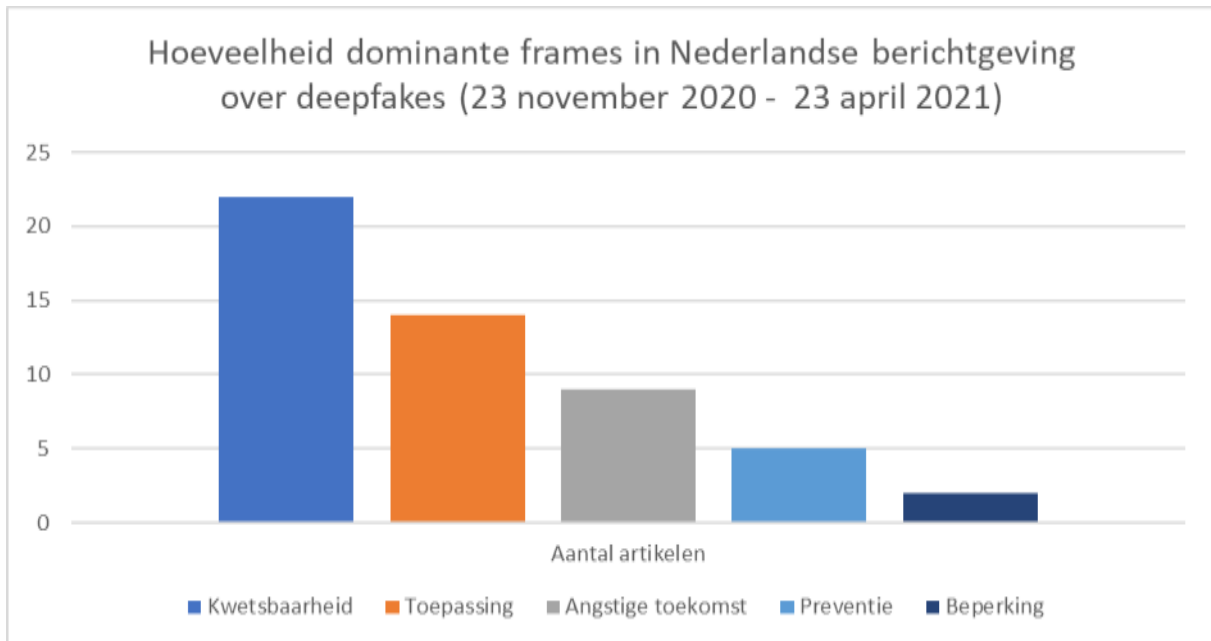
#### 4.1.7: Afsluiting van de frameanalyse

Ter afsluiting van de frameanalyse zijn Tabel 1 en Figuur 1 gemaakt, respectievelijk voor een samenvatting en een conclusie voorafgaand aan de bredere conclusie van deze scriptie in Hoofdstuk 5. Tabel 1 (zie volgende pagina) is de matrix van de frames en *reasoning devices*, waarin kort wordt aangegeven hoe deze met elkaar relateren waar dat relevant is en waarin de hoofdpunten van de analyse opgesomd worden. Figuur 1 (zie pagina 23) laat zien hoe vaak elk frame dominant is in de geanalyseerde artikelen.

	Wat is het probleem?	Wat zijn de oorzaken van het probleem?	Hoe kan het probleem opgelost worden?	Moreel oordeel
Angstige Toekomst	Deepfakes kunnen als apex van het nepnieuws een toekomst teweeg brengen waarin het onderscheid tussen nep en echt beeld niet meer te maken is.	-	-	-
Kwetsbaarheid	Deepfakes doen het onderscheid tussen echt en werkelijk vervagen, wat gevaarlijk is voor de democratie doordat ze steeds beter worden en mensen te weinig mediawijsheid hebben. Ook helpen deepfakes in het verspreiden van wraakporno.	Vooraf dat computers steeds makkelijker steeds betere deepfakes kunnen maken. Verspreiders van deepfakes kunnen vele motieven hebben, zoals lust naar geld en macht, maar ook niet beter weten.	-	Deepfakes zijn gevaarlijk, eng en vijanden van de journalistiek.
Preventie	-	-	Deepfakes kunnen op drie manieren tegengegaan worden: met mediawijsheid, met wetgeving en met technologie die deepfakes opspoor. Bij de eerste twee opties zijn echter kanttekeningen te plaatsen op het gebied van hun effectiviteit	-
Toepassing	Er bestaan diverse problemen in de samenleving.	De problemen in de samenleving hebben ieder hun eigen oorzaken.	Deepfakes kunnen dienen als oplossingen voor de problemen in de samenleving	Deepfakes zijn grappig en kunnen gebruikt worden om hilarische satire te maken, maar ze hebben ook gevaarlijke kanten.
Beperking	-	-	-	Deepfakes hebben een aantal significante beperking waardoor ze hun beoogde toepassingen (nog) niet goed kunnen uitvoeren.

Tabel 1: De matrix van framebundels en *reasoning devices*.





Figuur 1: Hoeveelheid dominante frames in Nederlandse berichtgeving over deepfakes tussen 23 november 2020 en 23 april 2021; in totaal zijn 52 van de 57 artikelen meegenomen de 5 missende artikelen zijn artikelen die identiek zijn aan andere maar een andere kop hebben.

Zoals te zien in Figuur 1 zijn twee van de frames ontleend aan Wahl-Jorgensen en Carlson (2021) vaak dominant in de geanalyseerde artikelen: die van de kwetsbare samenleving en de angstige toekomst. Dit komt ook overeen met de conclusies van Wahl-Jorgensen en Carlson (2021), die stelden dat Britse berichtgeving over deepfakes overweldigend negatief was. Wat opvalt aan Figuur 1 is echter dat het derde frame van Wahl-Jorgensen en Carlson (2021) minder vaak dominant was dan de andere twee, wat erop duidt dat de Nederlandse berichtgeving in deze periode zich meer bezig hield met het benoemen van problemen dan de oplossingen van deze problemen. In veel gevallen worden oplossingen wel benoemd, maar is dit minder opvallend gemaakt dan het benoemen van de problemen. Verder valt aan Figuur 1 ook op dat het overwegend positieve toepassingsframe ook vaak dominant is: de berichtgeving hield zich naast het benoemen van de negatieve kanten van deepfakes dus ook bezig met het tonen van hoe deepfakes in de samenleving ingezet kunnen worden (alhoewel de gevaren van deepfakes in zulke artikelen ook benoemd worden). Het laatste frame, dat vooral morele oordelen over deepfakes inhoudt, is minder vaak dominant en is dit vaak in artikelen waarin deepfakes maar kort aan bod komen.

Het beeld van de berichtgeving dat hier naar voren komt is dus minder negatief dan het beeld uit het paper van Wahl-Jorgensen en Carlson (2021), aangezien de Nederlandse journalistiek in deze periode ook de focus lag op de positievere aspecten van deepfakes. Het blijft echter wel zo dat de positievere frames maar weinig dominant zijn (14 artikelen waar het toepassingsframes dominant zijn) vergeleken met de negatievere frames (38 artikelen waar de andere, negatieve frames dominant zijn). En daar komt ook bij dat als de positieve frames dominant zijn, ze vaak gepaard gaan met negatieve frames. De berichtgeving is dus grotendeels negatief over deepfakes, maar minder negatief dan de Britse berichtgeving geanalyseerd door Wahl-Jorgensen en Carlson (2021).

## 4.2: Reconstructie van de casus

### 4.2.1: Reconstructie door middel van inhoudsanalyse

Met een duidelijk beeld van hoe deepfakes in de Nederlandse media werden neergezet in de periode voorafgaand aan de casus, kan ingegaan worden op de daadwerkelijke casus. De gebeurtenis zelf, het gesprek tussen imitatie-Volkov en de vaste Kamercommissie voor buitenlandse zaken, vond plaats op 22 april 2021 en de daaropvolgende berichtgeving ging door tot 9 juni 2021 (met een enkel relevant artikel dat in januari 2023 gepubliceerd is). Als gekeken wordt naar de literatuur over berichtgeving over AI, dan kan gezien worden dat de relevante periodes na de piek van de *hype-cycle* over het onderwerp vallen. Deze was van 2014 tot 2018, waarna de hoeveelheid berichtgeving weer afnam (Nguyen & Herkam, 2022). Zoals gesteld in Hoofdstuk 2 hoort bij deze latere fasen van berichtgeving over nieuwe technologie een focus op frames die de impact van de nieuwe technologie op de samenleving benadrukken. Dit soort frames kwamen veelvuldig terug in de frameanalyse uit sectie 4.1 en dit is dus het kader waarin de berichtgeving over de casus zich in begeeft.

Voor de reconstructie in deze sectie zijn 54 Nederlandse artikelen over de casus en relevant achtergrondmateriaal geanalyseerd. Het gaat bij dat laatste om zaken zoals een artikel over de andere gesprekken waar imitatie-Volkov bij betrokken was, twee interviews met Vovan en Lexus en opnames van de gesprekken die het duo op hun *YouTube* kanaal hadden gezet (voordat deze later verwijderd zijn). Ruwweg kunnen vier fases geïdentificeerd worden in de berichtgeving. De eerste loopt van 22 tot de middag van 24 april en het is hier waarin wordt gesteld dat imitatie-Volkov gebruik maakte van deepfaketechnologie. In de tweede fase, die van 24 april tot 27 april, ontstaat verwarring over of er daadwerkelijk sprake was van deepfaketechnologie. De derde fase begint op 27 april en hier wordt het verhaal dat imitatie-Volkov zichzelf geschminkt had meer opgepakt, maar er is nog steeds onduidelijkheid. Deze fase duurt tot 8 mei, waarna de berichtgeving stopt. Op 28 mei laait deze echt weer op wanneer Lexus en Vovan een opname van hun gesprek met de Kamerleden online zetten en het duidelijk wordt dat er nooit deepfakes in het spel waren. Deze laatste fase duurt tot 9 juni. De rest van deze sectie van de scriptie gaat hier gedetailleerder op in.

Zoals gesteld in de Introductie is de achtergrond van de casus een aantal vergelijkbare incidenten waarbij Lexus en Vovan spraken met Europese politici terwijl Lexus zich voordeed als Leonid Volkov. In het kader van deze incidenten publiceerde *The Guardian* op 22 april 2021 een artikel, waarin werd geïdentificeerd met welke politici gesproken was: Rihard Kols, toen voorzitter van de Letse Commissie van Buitenlandse Zaken, Tom Tugendhat, toen voorzitter van de vaste parlementaire commissie Buitenlandse Zaken van het Verenigd Koninkrijk, en de toenmalige minister presidenten van Litouwen en Estland. In dit artikel, dat voor de daadwerkelijke casus met de Nederlandse Kamercommissie werd gepubliceerd, werd al gesproken van deepfakes. Deze claim werd gemaakt door de echte Leonid Volkov gebaseerd op een screenshot dat Rihard Kols zogenaamd genomen had van de bedrieger waarmee hij gesproken had. *The Guardian* had ook contact opgenomen met Lexus voor dit artikel, die vaag bleef over of hij achter het gesprek zat, maar wel ontkende dat hij gebruik had gemaakt van “filters” en dat de echte Volkov “waarschijnlijk foute informatie” had.

Een dag nadat het artikel van *The Guardian* gepubliceerd werd, op 23 april, publiceerde *De Volkskrant* een artikel getiteld “Kamerleden vergaderen met deepfake-imitatie van stafchef Russische oppositieleider Navalny” waarin melding werd gedaan van een gesprek tussen de vaste Kamercommissie voor Buitenlandse Zaken en de imitatie-Volkov. Dit zou diezelfde avond aan *De Volkskrant* gemeld zijn door de griffie van de Tweede Kamer. In de kop en lead van het artikel stelde *De Volkskrant* dat de bedrieger een “deepfake-imitatie” was en in de eerste zin van het artikel zelf werd gesteld dat de inhoud van de lead bevestigd was door de voornoemde griffie. Later in het artikel werd het screenshot dat Rihard Kols op Twitter geplaatst had voorgelegd aan Jarno Duursma, een tech-trendwatcher. Hij stelde dat de imitatie van Volkov “zeer overtuigend” gedaan was, maar deed geen uitspraak over of er gebruik gemaakt was van deepfaketechnologie. Hij deed zijn uitspraak echter wel gebaseerd op de claim dat zulke technologie gebruikt is volgens het artikel. Wat ook opvalt is dat dit artikel een week later, op 30 april, geüpdatet is om duidelijk te maken dat Duursma zijn uitspraak deed op basis van het screenshot van Kols en niet op een opname van het gesprek met de vaste Kamercommissie. Het eind van het artikel had ook een sectie genaamd “wat zijn deepfakes?”, waarin de goede en slechte kanten van deepfakes belicht werden, en ook wat deepfakes zijn.

Na deze initiële publicatie van *De Volkskrant* waren er later in de nacht van 23 april en in de ochtend van 24 april andere nieuwsorganisaties die het verhaal oppakten. Het gaat hierom artikelen van *NRC*, *De Telegraaf*, *NU.nl*, het *Noordhollands Dagblad*, *Trouw*, *Tubantia*, *RTL Nieuws*, *Linda*, *NOS*, *Nieuwsuur* en *Tweakers*. In de desbetreffende artikelen werden het artikel van *De Volkskrant* en de Tweede Kamer als bronnen geciteerd voor de claim dat er deepfakes gebruikt zouden zijn. Sommige van de artikelen waarschuwden ook voor het gevaar van deepfakes, maar de meeste namen vooral het verhaal van *De Volkskrant* over. Alleen het artikel van *NRC* voegde verdere context toe aan de casus: D66-politicus Sjoerd Sjoerdsma, één van de Kamerleden die met imitatie-Volkov gesproken had, stelde dat er ook een gesprek was met een bedrieger die zich voordeed als Ivan Zhdanov, de directeur van Alexei Navalny's anticorruptiebeweging. Dit aspect van de casus werd nu en dan genoemd in de berichtgeving, maar speelt een kleine rol en werd niet in verband gebracht met het deepfake-aspect. De enige keer dat dit gebeurde is in het desbetreffende *NRC*-artikel, waarin wordt gesteld dat het onduidelijk is of de imitatie-Zhdanov gebruik had gemaakt van deepfaketechnologie.

In de middag van 24 april, iets voor drie uur, publiceerde *De Volkskrant* een vervolgartikel getiteld “Kamerleden spraken met een nepversie van stafchef van Navalny: deepfake of dubbelganger?”, waarin het verhaal over de deepfaketechnologie genuanceerd werd. De claim van de echte Leonid Volkov werd in dit artikel herhaald. Er waren ook twee andere bronnen die werden aangehaald voor deze kant van het verhaal. De eerste was een tweet van Rihard Kols van 23 april waarin hij een statement deelde dat expliciet stelde dat in de incidenten met imitatie-Volkov er gebruik gemaakt werd van “manipulated and artificial intelligence (AI) generated media”. Dit statement was ondertekent door Kols zelf, en ook zijn collega's uit Estland en Litouwen: Marko Mihkelson en Zygimantas Pavilionis. De tweede bron was een artikel van de Estse publieke omroep Eesti Rahvusringhääling, waarin gesteld werd dat er gebruik gemaakt was van deepfaketechnologie. Ook werden de claims van Duursma uit het eerdere *Volkskrant*-artikel herhaald. In het nieuwe artikel van *De Volkskrant* werd echter ook de andere mogelijkheid belicht: dat de bedrieger geen deepfaketechnologie gebruikte. Theo Gevers, hoogleraar computer vision aan de Universiteit van Amsterdam, vertegenwoordigde deze kant van het verhaal. Hij stelde dat op basis van het screenshot dat Rihard Kols deelde niet te zeggen viel of er deepfaketechnologie gebruikt was, maar dat de man in de foto wel op een echte persoon leek. Ook stelde het *Volkskrant*-artikel dat het in het algemeen lastig te zeggen was omdat er geen bewegende beelden van imitatie-Volkov beschikbaar waren.

Met dit artikel begon de tweede fase van de berichtgeving over de casus, waarin wat twijfel ontstond over het deepfakeverhaal. *Tweakers* publiceerde in diens artikel “Tweede Kamerleden hadden videogesprek met deepfake stafchef Navalny” bijvoorbeeld een update waarin de strekking van het Volkskrantartikel van 24 april samengevat werd. De *NOS* publiceerde ook het artikel “Veel onduidelijk over 'deepfake-gesprek' van Kamerleden met medewerker Navalny” waarin werd gesteld dat de Tweede Kamer niets had bevestigd over deepfakes en waarin de drie bronnen van *De Volkskrant* (echte Volkov, Baltische politici en LTV) ook werden genoemd. Ook is dit het eerste Nederlandstalige artikel waarin werd erkent dat Lexus in het initiële artikel van *The Guardian* zei dat hij geen gebruik had gemaakt van “een filter”. *Villamedia* stelde dat de echte Volkov zei dat er sprake was van deepfaketechnologie, maar ook dat er veel onduidelijkheid was, waarbij Marietje Schaake, lid van het Cyber Policy Center van Stanford University, werd aangehaald. Ten slotte publiceerde *De Volkskrant* op 25 april ook een stuk waarin Zygis Pavilionis als bron gebruikt werd. Deze Litouwse politicus stelt dat hij geen bewijs had dat er deepfaketechnologie in het spel was, maar wel dat hij sterke vermoedens had. Desondanks noemde dit artikel van *De Volkskrant* een afbeelding van imitatie-Volkov een “deepfakeversie” van de man.

Dit laatste is emblematisch voor een andere, meer voorkomende trend in deze periode. Publicaties die voorheen nog niet over de casus hadden gepubliceerd, doen dit nu wel en pushen het deepfakeverhaal. *Apparata*, *De Limburger*, *De Standaard*, *Dutch IT Channel* en het *Reformatorisch Dagblad* meldten zonder verdere nuance dat er deepfaketechnologie gebruikt was. Afwisselend werden geen bronnen of *De Volkskrant* en de Tweede Kamer als bronnen genoemd, terugverwijzend naar het Volkskrantartikel van 23 april. Ook gebruikten *Apparata*, *De Limburger* en *De Standaard* de gebeurtenis om een algemeen verhaal te vertellen over deepfakes, waarbij veelvuldig werd gewaarschuwd voor de toekomst die ze zouden kunnen brengen en werd uitgelegd waarom ze gevaarlijk zijn. Andere nieuwspublicaties vertelden een iets genuanceerder verhaal waarbij er wel werd gesteld dat er onduidelijkheid is, maar waarin ook werd geneigd naar het deepfakeverhaal. Zo stelde het *Dagblad van het Noorden* dat het “lijkt” dat imitatie-Volkov deepfaketechnologie gebruikte en Elsevier dat het “vermoedelijk” zo was. Ook een opiniestuk gepubliceerd in *De Volkskrant* stelde in de lead dat er deepfakes in het spel waren, voordat in het artikel zelf geschreven werd dat het “onduidelijk” is.

Deze periode van verwarring duurde van de middag van 25 april tot 27 april en eindigde wanneer Lexus en Vovan meer details gaven over hoe ze te werk waren gegaan. Op 26 april gaven de twee een interview aan de Baltische nieuwswebsite *Baltnews* en vertelden ze dat er geen gebruik gemaakt was van deepfaketechnologie. Ook was er in dit interview een foto gepubliceerd van Lexus als imitatie-Volkov. Journalistencollectief *Raam op Rusland* deed een dag later bericht hierover en stelde ondubbelzinnig dat er geen gebruik gemaakt was van deepfaketechnologie, ondanks wat de echte Volkov en de Baltische politici claimden. Dezelfde dag publiceerde het *Reformatorisch Dagblad* een artikel dat ging over het gevaar van deepfakes in het algemeen, waarin de casus gebruikt werd als voorbeeld. Er werd ook terugverwezen naar het Volkskrantartikel van 23 april en Jarno Duursma. Het is onduidelijk of dit stuk van het *Reformatorisch Dagblad* voor of na het *Raam op Rusland*-artikel gepubliceerd was, maar valt meer in de lijn van de berichtgeving uit de periode voor 27 april.

Na de publicatie van het *Raam op Rusland*-artikel vond er een bredere omslag plaats in de berichtgeving, waarin het verhaal van het deepfakegebruik als minder waarschijnlijk werd gezien dan dat van Lexus en Vovan. Een artikel getiteld “Spraken Kamerleden met een deepfake of met een komiek in een badjas?” van *NRC* vertelde het deepfakeverhaal en het verhaal van Lexus en Vovan, maar concludeerde dat het lastig was om te zeggen wat er echt gebeurt was omdat er geen opnames waren van de gesprekken. Verder leek de foto van Lexus uit het *Baltnews*artikel niet op het screenshot

dat Rihard Kols gedeeld had volgens dit artikel. Een ander *NRC*-artikel, gepubliceerd op dezelfde dag, 28 april, getiteld “‘Het is moeilijk een goede deepfake-video te maken’” ging dieper in op de kwestie en kwam tot dezelfde conclusie dat er onduidelijkheid was. Dit artikel stelde echter ook dat er te weinig terughoudendheid was in de berichtgeving volgend op het originele artikel van *The Guardian*. Volgens Jan Broersen, hoogleraar AI aan de Universiteit Utrecht, die te woord kwam in dit artikel, kwam dit doordat er een *hype* bestaat rond deepfaketechnologie.

Artikelen van andere nieuwsorganisaties van na 28 april kwamen tot vergelijkbare conclusies over of er deepfakefilters gebruikt waren: er was onduidelijkheid en iedereen dacht dat er deepfakes in het spel waren, maar het verhaal van Vovan en Lexus werd als een betere verklaring gezien. Specifiek werd het artikel van *Baltnews* geciteerd, terwijl *Raam op Rusland* nergens genoemd werd. Artikelen van het *FD*, het *Reformatorisch Dagblad*, *Tweakers* en *Tubantia* vallen binnen deze trend. Een derde artikel van *NRC* uit deze periode, een opiniestuk, ageerde ook tegen het gebruik van deepfaketechnologie: er was nog steeds onduidelijk wat er was gebeurd, maar er waren politici succesvol beetgenomen en dit werd gebruikt als inleiding voor een stuk dat stelt dat deepfakes de liberale democratie kunnen ondermijnen. Of er daadwerkelijk geen deepfakes in het spel waren of niet, de technologie blijft gevaarlijk volgens dit artikel. Niet alle berichtgeving volgde deze trend echter: een laatste artikel van *De Volkskrant*, gepubliceerd op 5 mei en van humoristische aard, stelde dat er wel deepfaketechnologie gebruikt is.

Na 5 mei stopte de berichtgeving over de casus voor een aantal weken en eindigde de derde periode. Echter was op 30 april een artikel gepubliceerd in *The Verge* waarin Lexus en Vovan meer vertelden over hoe ze te werk gegaan waren. Een dag daarvoor hadden de twee een opname gepubliceerd op *YouTube* van één van hun gesprekken dat zij met Oekraïense politici gevoerd hadden en deze video stond ook in het artikel van *The Verge*. Ook wvws in dit artikel de bron onthuld van het screenshot dat Rihard Kols aan het begin van de berichtgeving deelde op Twitter: een interview uit 2018 tussen de echte Volkov en de organisatie *After Empire*. Vovan en Lexus hadden een foto uit dat interview gebruikt als profielplaatje aldus het artikel. *Tweakers* deed bericht over dit artikel van *The Verge* op 1 mei en noemde beide aspecten. De opname werd niet gebruikt door andere Nederlandse media voor verdere beeldverificatie, ondanks dat eerder gesteld was dat het mogelijk zou zijn om te bepalen of de bedrieger deepfakefilters gebruikt had zonder bewegende beelden. Ook het feit dat het screenshot dat Kols deelde niet van een bedrieger was, kwam niet aan bod in latere artikelen.

De vierde en laatste periode in de berichtgeving begon op 27 mei, wanneer Lexus en Vovan een opname van hun gesprek met de Nederlandse Kamercommissie op *YouTube* zetten. Waar de opname van het gesprek met de Oekraïense politici nauwelijks media-aandacht kreeg in Nederland, daar kreeg deze opname dat wel. Het eerste bericht, van 28 mei, is wederom van *De Volkskrant* in het artikel “Deepfake? Nee, gewoon een slecht geschminkte grapjurk in een badjas” waarin definitief gesteld werd dat er geen deepfaketechnologie in het spel was. Dit werd gezegd door Theo Gevers, die ook in het *Volkskrant*-artikel van 24 april aan het woord kwam, op basis van de opname. Het artikel stelde verder dat de initiële beweringen over deepfakes gevoed werden door de echte Volkov en Baltische politici, maar de invloed van het *Volkskrant*-artikel van 23 april of andere Nederlandse berichtgeving werd hierbij niet genoemd. Wel werd er gelinkt naar het *Volkskrant*-artikel van 24 april, dat stelt dat het onduidelijk was of er deepfaketechnologie in het spel was.

Artikelen van andere nieuwsorganisaties die op 28 mei en in de dagen daarna gepubliceerd werden volgden min of meer dit stramien. Er werd gesproken over de *YouTube* opname of de afronding van het onderzoek van de Tweede Kamer naar het incident en dat het definitief ging om een geschminkte man, niet iemand die deepfakefilters gebruikte. Ook werd gesteld dat er in initiële berichtgeving gesproken werd van deepfaketechnologie, maar er werd niet ingegaan op wie schuld had aan deze fout of überhaupt waarom deze gemaakt was. Het gaat hier om publicaties van *NRC*, *NU.nl*, *Tubantia*, het *Friesch Dagblad*, *Het Parool*, het *Reformatorisch Dagblad* en *Trouw*. Alleen de artikelen van *NU.nl* en *Trouw* uit deze periode noemen hun eerdere berichtgeving over de casus en dat deze deel was van de initiële “sensatie”. Een aantal andere artikelen uit deze periode schreven ook over de opname van Lexus en Vovan, maar noemen het deepfake aspect niet.

Na 1 juni stopte de berichtgeving over de casus min of meer. Een week later, op 9 juni, noemden een aantal artikelen de casus nog wel in verband van berichtgeving over een gesprek tussen Nederlandse Kamerleden en de echte Leonid Volkov, maar dit gebeurde in vogelvlucht. Dat er deepfakes in het spel waren bij deze casus bleef echter misinformatie die later rond bleef waren, zoals blijkt uit een artikel van *VPNgids* dat op 20 januari 2023, bijna twee jaar na de casus, gepubliceerd werd en het meest recente artikel is dat de casus noemt. Dit artikel ging in op wat deepfakes waren en waarom ze gevaarlijk waren. De casus werd hier als voorbeeld gebruikt. Deepfaketechnologie zou succesvol gebruikt zijn in de casus en als bron werd het eerste Volkskrantartikel van 23 april gebruikt.

#### 4.2.2: Reconstructie door middel van interviews

Voor aanvullende context zijn, zoals genoemd in sectie 3.2 van deze scriptie, twee interviews gehouden om wat onduidelijkheden uit sectie 4.2.1 aan te vullen en de productiekant van de berichtgeving te belichten. De inzichten uit de interviews zullen in deze sectie ruwweg chronologisch gepresenteerd worden zoals de reconstructie in sectie 4.2.1.

Uit het interview met de heer Verhagen bleek dat de initiële claims over deepfakes inderdaad afkomstig waren van buitenlandse partijen zoals *The Guardian*, verschillende Baltische politici en de echte Leonid Volkov. Hetzelfde geldt voor het verhaal van de imitatie-Volkov die in andere landen politici nepte, maar het verhaal van de Nederlandse casus kwam van de griffie als aangegeven in het Volkskrantartikel van 23 april 2021. De Kamercommissie en de griffie van de Tweede Kamer gebruikten de term deepfake niet in hun contact met *De Volkskrant*. De heer Hendrikman, die destijds werkte als freelancer voor *Tweakers*, schreef de dag daarna een artikel over het gesprek tussen de Nederlandse Kamerleden en de Russische bedriegers, maar waar hij het verhaal vandaan had is niet helemaal duidelijk meer. Hij noemt het Volkskrantartikel in zijn artikel, maar tijdens het interview maakte hij duidelijk dat het mogelijk was dat een gebruiker van de Tweakerswebsite een tip had achter gelaten over de gebeurtenis. Bij het artikel stond echter geen gebruikersnaam van een potentiële tipgever en Hendrikman wist niet meer zeker waar hij het verhaal het eerst vandaan had gehaald. Hendrikman had zelf ook niet aan beeldverificatie gedaan.

Ergens op 23 of 24 april en waarschijnlijk na de publicatie van het eerste Volkskrantartikel belde de heer Verhagen met Theo Gevers over de casus. Gevers was sceptischer over het deepfakeverhaal dan de bronnen die geraadpleegd werden voor het Volkskrantartikel van 23 april. Door dit gesprek, waardoor twijfel ontstond bij Verhagen over het deepfakeverhaal, en omdat hij vond dat in het eerste artikel te stellig gesproken werd over deepfaketechnologie, schreef Verhagen het Volkskrantartikel van 24 april 2021. Op basis van dit artikel schreef Hendrikman in zijn artikel van 24 april op 25 april een update waarin het Volkskrantartikel van 24 april kort werd samengevat. Dit was omdat Hendrikman ook vond dat hij te stellig was in zijn initiële artikel en dit wou nuanceren. Achteraf vond hij ook dat de bronnen die *De Volkskrant* citeerde in het artikel van 23 april niet genoeg kennis hadden over deepfakes en dat hij dit had moeten zien. Een andere factor die hij noemt is dat hij in zijn tijd als freelancer een bepaalde kwantiteit aan artikelen moest leveren en daardoor de kwaliteit van elk artikel niet even hoog kon zijn. Een week later, op 30 april, werd de bewoording van het Volkskrantartikel van 23 april ook aangepast op expliciet verzoek van Jarno Duursma, die vond dat hij in het artikel te stellig was.

Op 1 mei kwam er meer duidelijkheid over de situatie met de Russische bedriegers toen *The Verge* een artikel publiceerde waarin Lexus en Vovan werden geïnterviewd, als gesteld in sectie 4.2.1. Hendrikman greep dit artikel aan om zijn eigen artikel voor *Tweakers* te schrijven als een rectificatie van zijn artikel van 24 april. Normaal zou Hendrikman hier voor *Tweakers* geen artikel over schrijven omdat het “techhaakje” ontbrak. Lexus en Vovan hadden immers geen deepfaketechnologie maar schmink gebruikt. Het is in dit geval echer wel gedaan als rectificatie voor het eerdere artikel. Iets vergelijkbaars geldt voor de gedachte achter het Volkskrantartikel van 28 mei, waarin geschreven werd over de video met het gesprek tussen de Kamercommissie en de imitatie-Volkov dat op *YouTube* gezet was de vorige dag. Er was door *De Volkskrant* niet geschreven over de video van het gesprek dat Lexus en Vovan voerden met Oekraïense politici uit het artikel van *The Verge* omdat Verhagen (en vermoedelijk zijn medeauteur Arnout Brouwers) wouden wachten op beeld van het Nederlandse gesprek. Dit om zeker te weten dat daar geen sprake was van gebruik van deepfaketechnologie en daarom is voor extra beeldverificatie de video ook voorgelegd aan Theo Gevers. Hendrikman schreef niet over het Nederlandse gesprek omdat de rectificatie al had plaatsgevonden met zijn artikel van 1 mei in zijn ogen en omdat de publicatie van de video op donderdag 27 mei buiten zijn weekenddienst viel. Ook ontbrak het “techhaakje” bij dit verhaal.

Verhagen en Hendrikman zien hun latere artikelen over de casus dus als retracties of rectificaties van de fout. Verhagen trad echter ook in detail over waarom het originele Volkskrantartikel van 23 april niet is aangepast. Volgens hem zou dat niet te doen zijn want dan moet el nieuwsartikel aangepast worden en dan is het beter om follow-ups te schrijven omdat nieuws zich altijd ontwikkelt. Ook vond hij het principieel niet goed om artikelen achteraf aan te passen. Ten slotte weet Hendrikman niet of er bij *Tweakers* intern beleid is veranderd door de fouten die gemaakt zijn in de berichtgeving over deze casus, omdat hij toen freelancer was en niet meedraaide met het interne beleid.

## Hoofdstuk 5: Conclusie en discussie

Ter conclusie, hebben bestaande beelden over of het frame van deepfakes als bedreigende nieuwe technologie een rol gespeeld in de initiële foute berichtgeving over de casus van de Kamercommissie die met de imitatie-Volkov sprak? Om te beginnen, wat waren die beelden precies? Zoals geïllustreerd in de frameanalyse was de berichtgeving over deepfakes in het algemeen overwegend negatief in het half jaar voorafgaand aan het begin van de casus. Er werd vaak gewaarschuwd voor het gevaar van deepfakes, wat overeenkomt met de bevindingen van Wahl-Jorgensen en Carlson (2021) over Britse media. Echter zetten Nederlandse media deepfakes ook positief in het licht, maar ook in die gevallen werd het gevaar van deepfakes gevaar expliciet of impliciet aangestipt.

Het is echter onwaarschijnlijk dat deze trends in de Nederlandse berichtgeving over deepfakes aan de wieg stonden van de claim dat de imitatie-Volkov van deepfaketechnologie gebruikt maakte. Toen de bedrieger met de Nederlandse Kamerleden sprak waren er al een aantal incidenten geweest waar verschillende Europese media over berichtgeven hadden. In deze buitenlandse berichtgeving was er al sprake van claims over deepfakes en deze werden overgenomen door Nederlandse media toen deze berichtten over de incidenten voorafgaand aan de casus. Uit het interview met de heer Verhagen blijkt ook dat deze buitenlandse berichtgeving een belangrijke bron vormde voor de initiële berichtgeving van *De Volkskrant* over deze casus, waar andere Nederlandse nieuwsmidia zich vervolgens weer op baseerden. De griffie van de Tweede Kamer en de Kamercommissie spraken nooit van deepfakes. De buitenlandse berichtgeving baseerde zich op de echte Leonid Volkov, die speculeerde dat de bedriegers van deepfaketechnologie gebruik hadden gemaakt gebaseerd op een screenshot dat Rihard Kols van de bedrieger gemaakt waar hij mee gesproken zou hebben. Dat Volkov verbaasd was dat het screenshot zo erg op hem leek is echter niet verwonderlijk omdat het screenshot niet van de bedrieger was, maar van de echte Volkov uit een interview uit 2018. Dit screenshot werd als de bedrieger aangezien omdat Lexus en Vovan het gebruikten als het profielplaatje voor het Zoom-account van de imitatie-Volkov volgens *The Verge*.

Aan de andere kant hebben de bestaande beelden en hun bijbehorende frames in de Nederlandse media waarschijnlijk wel een rol gespeeld in het in stand houden van de deepfakeclaims. De casus vormde immers een mooi voorbeeld van dat waar zulke bestaande beelden voor waarschuwden: het verschil tussen nep en echt zal straks niet meer te zien zijn en zelfs politici kunnen gefopt worden. In de berichtgeving over de casus werd het gesprek met de imitatie-Volkov dan ook meerdere malen gebruikt als een voorbeeld in of een inleiding van een algemeen verhaal over het gevaar van deepfakes. Ook in het eerste Volkskrantartikel over de casus zelf is er een stuk over “wat zijn deepfakes?” waarin een dergelijk verhaal verteld wordt. Dat de *hype* rond deepfakes in deze initiële berichtgeving een rol speelde wordt in latere fasen van de berichtgeving ook expliciet benoemd als een factor.

De bestaande beelden over deepfakes waren echter niet de enige factor bij het in stand houden van de deepfakeclaims. Een andere prominente factor is het feit dat er lang geen bewegend beeldmateriaal van de bedrieger beschikbaar was, waardoor er geen beeldverificatie kon plaatsvinden. Het enige screenshot dat zogenaamd van de bedrieger was, bleek later van de echte Volkov te zijn en daardoor aan de wieg te staan van de deepfakeclaims. Echter, toen er eind april 2021 wel beeld naar buiten kwam van de imitatie-Volkov op YouTube en in een artikel van *The Verge*, werd hier door de Nederlandse media nauwelijks over geschreven: alleen *Tweakers* noemde het in één artikel. Dit gebeurde iets meer dan een week na het begin van de berichtgeving van de casus, die op dat moment aan het afbouwen was, waardoor dit waarschijnlijk niet breder opgepakt werd door de Nederlandse media. Het feit dat het verhaal dat imitatie-Volkov gebruik maakte van deepfakes, ondanks dat hier



twijfel over ontstaan was in de voorafgaande week, goed paste binnen bestaande verhalen over deepfakes en dat het nieuwe beeldmateriaal van de imitatie-Volkov niet van het gesprek met de Nederlandse Kamerleden was speelden waarschijnlijk een rol hierin. Deze laatste factor werd ook bevestigd door de heer Verhagen in het interview dat met hem gevoerd is.

Een laatste aspect dat relevant is, zijn de reacties van media nadat er beeldmateriaal kwam van de bedrieger die met de Nederlandse Kamerleden gesproken had. De berichtgeving was er unaniem over dat de bedrieger geen deepfaketechnologie gebruikt had, maar schmink, per eerdere verklaringen van Lexus en Vovan. Er werd ook teruggeblikt op de initiële berichtgeving nadat dit beeldmateriaal naar boven kwam, maar er werd niet diep op ingegaan. In het merendeel van de artikelen uit deze periode wordt er geconstateerd dat er foutieve berichtgeving was en dat dit deel was van sensatie en *hype*, maar er wordt niet gevraagd wie hier schuld aan heeft of waarom het überhaupt gebeurt is. De artikelen die dit wel doen, zoals eerder genoemd, leggen de schuld bij de *hype* rond deepfaketechnologie. Ook zijn er geen retracties of correcties geplaatst bij de artikelen uit de eerste dagen van de casus, volgens de heren Verhagen en Hendrikman omdat de follow-up artikelen dienen als rectificaties van eerdere artikelen. De initiële artikelen staan echter nog wel online, waardoor de misinformatie over de casus twee jaar later nog steeds rondwaart zoals te zien valt in het artikel van *VPNGids*. Ook noemen alleen een artikel van *NU.nl* en een artikel van *Trouw* expliciet dat hun eerdere berichtgeving onderdeel uitmaakte van deze initiële foutieve berichtgeving. Het tweede artikel van *Tweakers* doet dit op een meer impliciete manier.

Ten slotte is van belang dat ook nadat het beeldmateriaal van het het gesprek tussen de imitatie-Volkov en de Nederlandse Kamerleden naar buiten kwam, er nog steeds gebruik werd gemaakt van de casus om een verhaal te vertellen over het gevaar van deepfakes. Deze keer ging het er in de berichtgeving over dat deepfakes gevaarlijk zijn omdat ze onduidelijk maken wat echt en nep is: er wordt impliciet gesteld dat de media gefopt zijn doordat er elders claims gemaakt werden over het gebruik van deepfaketechnologie. Lexus en Vovan staken hier ook wat de draak mee, aangezien zij hun YouTubevideo van het gesprek dat ze hadden met de Nederlandse Kamercommissie “Deepfake-show” noemden. Bestaande beelden over deepfakes hebben dus op verschillende manieren een rol gespeeld in het in stand houden van de fout, samen met het feit dat er lang geen beeldmateriaal beschikbaar was.

Er zijn echter nog een aantal zaken waar verder op ingegaan zou kunnen worden in vervolgonderzoek. De onderzochte casus blijft bijvoorbeeld deel van een reeks incidenten waarbij Lexus en Vovan ook met buitenlandse politici spraken. Dit aspect is waar nodig voor het context geven voor de Nederlandse berichtgeving in deze scriptie besproken, maar de andere incidenten zelf en de berichtgeving over die incidenten hebben geen aandacht gekregen. Het zou interessant kunnen zijn om te onderzoeken hoe de berichtgeving over deepfakeclaims in buitenlandse media zich ontwikkelde en of de ontwikkelingen vergelijkbaar zijn met die in de Nederlandse berichtgeving. Daarbij zou ook gekeken kunnen worden of de conclusies uit deze scriptie breder gelden, aangezien een zwaktepunt is dat het verloop van de in deze scriptie onderzochte casus niet perse representatief is voor andere, vergelijkbare casussen.

Verder richt deze scriptie zich vooral op inhoudsanalyse van nieuwsartikelen en hoewel er

journalisten zijn geïnterviewd, is dit vooral gedaan vanuit het oogpunt om de inhoudsanalyse aan te vullen. Het zou interessant kunnen zijn om meer journalisten te interviewen om meer inzicht te krijgen over hoe de berichtgeving evolueerde aan de productiekant. Uit het interview van de heer Hendrikman bleek bijvoorbeeld dat aan zijn kant de fout deels veroorzaakt werd omdat hij een bepaalde kwantiteit aan artikelen moest produceren, wat inhaakt op wat Thomson et al. (2020) schreven over de relatie tussen gebrek aan beeldverificatie en de eisen van het vak die gevolg zijn van zaken als de *24-hour-newscycle*. De heer Verhagen noemde dit aspect echter niet, dus het kan interessant zijn om te kijken hoe wijdverbreid dit is bijvoorbeeld. Hierbij moet wel genoemd worden dat de heren Verhagen en Hendrikman zich bepaalde zaken niet meer herinnerden omdat de berichtgeving twee jaar geleden plaatsvond. Dit is een zwaktepunt in deze scriptie omdat bepaalde zaken niet met volledige zekerheid gezegd kunnen worden en zal dat ook zijn van vervolgonderzoek dat zich richt op het houden van interviews.

Er lagen in ieder geval meerdere factoren ter grondslag aan de fout die gemaakt is in de berichtgeving over deze casus. Bestaande beelden over deepfakes speelden een rol, maar ook het gebrek aan opnames van de gesprekken met de bedrieger, structurele werkelijkheden binnen de journalistiek zoals de druk om een bepaalde kwantiteit aan artikelen te leveren en vertrouwen van journalisten in hun collega-journalisten dat zij dingen goed checken (zoals Verhagen en Hendrikman beiden aangeven in interviews). Er zijn hierbij, zoals aangegeven in de vorige twee alinea's, verschillende aspecten die verder belicht zouden kunnen worden in vervolgonderzoek, binnen de Nederlandse casus maar ook binnen andere, vergelijkbare casussen in binnen- en buitenland.

## Literatuurlijst

Aberdeen, T. (2013). Yin, R. K. (2009). Case study research: Design and methods (4th Ed.). Thousand Oaks, CA: Sage. The Canadian Journal of Action Research, 14(1), 69–71.

<https://doi.org/10.33524/cjar.v14i1.73>

Brouwers, A. & Verhagen, L. (2021, 23 april). Kamerleden vergaderen met deepfake-imitatie van stafchef Russische oppositieleider Navalny. *De Volkskrant*. Geraadpleegd op 20 maart 2023, van <https://www.volkskrant.nl/nieuws-achtergrond/kamerleden-vergaderen-met-deepfake-imitatie-van-stafchef-russische-oppositieleider-navalny~b04b5322/>

Centre for Data Ethics and Innovation. (2019). Deepfakes and Audio-visual Disinformation.

Courchesne, L., Ilhardt, J., & Shapiro, J. N. (2021). Review of social science research on the impact of countermeasures against influence operations. *Harvard Kennedy School Misinformation Review*.

<https://doi.org/10.37016/mr-2020-79>

Entman, R. M. (1993). Framing: Toward Clarification of a Fractured Paradigm. *Journal of Communication*, 43(4), 51–58. <https://doi.org/10.1111/j.1460-2466.1993.tb01304.x>

Farid, H., Davies, A., Webb, L., Wolf, C., Hwang, T., Zucconi & A., Lyu, S. (2019). *Deepfakes and audio-visual disinformation*. London: Centre for Data Ethics and Innovation. Geraadpleegd op 20 maart 2023, van <https://www.gov.uk/government/publications/cdei-publishes-its-first-series-of-three-snapshot-papers-ethical-issues-in-ai/snapshot-paper-deepfakes-and-audiovisual-disinformation>

Gosse, C. E., & Burkell, J. (2020). Politics and porn: how news media characterizes problems presented by deepfakes. *Critical Studies in Media Communication*, 37(5), 497–511.

<https://doi.org/10.1080/15295036.2020.1832697>

Heawood, J. (2018). Pseudo-public political speech: Democratic implications of the Cambridge Analytica scandal. *Information Polity*, 23(4), 429–434. <https://doi.org/10.3233/ip-180009>

Katsaounidou, A., Gardikiotis, A., Tsipas, N., & Dimoulas, C. (2020). News authentication and tampered images: evaluating the photo-truth impact through image verification algorithms. *Heliyon*, 6(12), e05808. <https://doi.org/10.1016/j.heliyon.2020.e05808>

Keultjes, H., (2020, 25 november). CDA en GroenLinks: Maak verspreiden gemanipuleerde naaktbeelden strafbaar. *Het Parool*. Geraadpleegd op 29 mei 2023, van <https://www.parool.nl/nederland/cda-en-groenlinks-maak-verspreiden-gemanipuleerde-naaktbeelden-strafbaar~b5b779f5/>

Khakurel, J., Penzenstadler, B., Porras, J., Knutas, A., & Zhang, W. (2018). The Rise of Artificial Intelligence under the Lens of Sustainability. *Technologies (Basel)*, 6(4), 100.

<https://doi.org/10.3390/technologies6040100>

Lee, N., McLeod, D. M., & Shah, D. V. (2008). Framing Policy Debates. *Communication Research*, 35(5), 695–718. <https://doi.org/10.1177/0093650208321792>

- Linström, M. A., & Marais, W. (2012). Qualitative news frame analysis: a methodology. *Communitas*, 17, 21–38. <https://africaneditors.org/journal/COM/full-text-pdf/25829-37533>
- Marks, L. A., Kalaitzandonakes, N., Wilkins, L., & Zakharova, L. (2007). Mass media framing of biotechnology news. *Public Understanding of Science*, 16(2), 183–203. <https://doi.org/10.1177/0963662506065054>
- Nelson, T. R., Clawson, R. A., & Oxley, Z. M. (1997). Media Framing of a Civil Liberties Conflict and Its Effect on Tolerance. *American Political Science Review*, 91(3), 567–583. <https://doi.org/10.2307/2952075>
- Nguyen, D., & Hekman, E. (2022c). The news framing of artificial intelligence: a critical exploration of how media discourses make sense of automation. *AI & Society*. <https://doi.org/10.1007/s00146-022-01511-1>
- Pennington, R., & Birthisel, J. (2016). When new media make news: Framing technology and sexual assault in the Steubenville rape case. *New Media & Society*, 18(11), 2435–2451. <https://doi.org/10.1177/1461444815612407>
- Roth, A. (2021, 23 april). European MPs targeted by deepfake video calls imitating Russian opposition. *The Guardian*. Geraadpleegd op 20 maart 2023, van <https://www.theguardian.com/world/2021/apr/22/european-mps-targeted-by-deepfake-video-calls-imitating-russian-opposition>
- Rottenberg, H. (2021, 27 april). Russisch bedriegersduo bediende zich van schmink, geen deepfake - Raam op Rusland. *Raam Op Rusland*. Geraadpleegd op 20 maart 2023, van <https://raamoprusland.nl/actueel/1859-russisch-bedriegersduo-bediende-zich-van-schmink-geen-deepfake>
- Schellevis, J. & Moerland, S. (2023, 16 januari). ChatGPT glipt langs docenten: “Ik gebruik het om snel huiswerk te maken.” *NOS.nl*. Geraadpleegd op 20 maart 2023, van <https://nos.nl/artikel/2460020-chatgpt-glipt-langs-docenten-ik-gebruik-het-om-snel-huiswerk-te-maken>
- Silbey, J. M., & Hartzog, W. (2019). The Upside of Deep Fakes. *Maryland Law Review*, 78(4), 960. <https://digitalcommons.law.umaryland.edu/cgi/viewcontent.cgi?article=3837&context=mlr>
- Sohrawardi, S. J., Seng, S., Chintha, A., Thai, B., Hickerson, A., Ptucha, R., & Wright, M. (2020, March). Defaking DeepFakes: Understanding journalists’ needs for DeepFake detection. In *Proceedings of the Computation+ Journalism 2020 Conference, Northeastern University, Boston, MA, USA* (Vol. 21).
- Tandoc Jr., E. C., Lim, Z. W., & Ling, R. (2017). Defining “Fake News.” *Digital Journalism*, 6(2), 137–153. <https://doi.org/10.1080/21670811.2017.1360143>
- Temir, E. (2020). Deepfake: New Era in The Age of Disinformation & End of Reliable Journalism. *DergiPark (Istanbul University)*. <https://doi.org/10.18094/josc.68533>
- Ternovski, J., Kalla, J., & Aronow, P. M. (2022). Negative Consequences of Informing Voters about

Deepfakes: Evidence from Two Survey Experiments. *Journal of Online Trust and Safety*, 1(2). <https://doi.org/10.54501/jots.v1i2.28>

Thomson, T., Angus, D., Dootson, P., Hurcombe, E., & Smith, A. (2020). Visual Mis/disinformation in Journalism and Public Communications: Current Verification Practices, Challenges, and Future Opportunities. *Journalism Practice*, 16(5), 938–962. <https://doi.org/10.1080/17512786.2020.1832139>

Van Gorp, B. (2007a). Het reconstrueren van frames via inductieve inhoudsanalyse: uitgangspunten en procedures. *Kwalon*, 12(2).

Van Gorp, B. (2007b). The constructionist approach to framing: Bringing culture back in. *Hansen, a. (Ed.), Mass Communication Research Methods*.

Wahl-Jorgensen, K., & Carlson, M. (2021). Conjecturing Fearful Futures: Journalistic Discourses on Deepfakes. *Journalism Practice*, 15(6), 803–820. <https://doi.org/10.1080/17512786.2021.1908838>

Weaver, D., Lively, E., & Bimber, B. (2009). Searching for a Frame. *Science Communication*, 31(2), 139–166. <https://doi.org/10.1177/1075547009340345>

Weikmann, T., & Lecheler, S. (2022). Visual disinformation in a digital age: A literature synthesis and research agenda. *New Media & Society*, 146144482211416. <https://doi.org/10.1177/14614448221141648>

Yadlin-Segal, A., & Oppenheim, Y. (2021). Whose dystopia is it anyway? Deepfakes and social media regulation. *Convergence*, 27(1), 36–51. <https://doi.org/10.1177/1354856520923963>

Zheng, L., Zhang, Y., & Thing, V. L. L. (2019). A survey on image tampering and its detection in real-world photos. *Journal of Visual Communication and Image Representation*, 58, 380–399. <https://doi.org/10.1016/j.jvcir.2018.12.022>

## Bijlage 1: Lijst van artikelen gebruikt in frameanalyse

1. Laat techniek niet flirten met de dood (23-11-2020; Reformatorisch Dagblad)
2. Oud nep, nieuw nep en gek nep (23-11-2020; NRC)
3. De infocalyps (23-11-2020; De Volkskrant)
4. 'Zondag met Lubach' hoeft echt niet te stoppen (24-11-2020; Trouw)
5. 'Nepnaaktfilmpjes zijn net zo fout als wraakporno' (25-11-2020; AD/De Dordtenaar)
6. CDA en GroenLinks: Maak verspreiden gemanipuleerde naaktbeelden strafbaar (25-11-2020; Het Parool)
7. 'Maak verspreiden gemanipuleerde naaktbeelden strafbaar (25-11-2020; PZC)
8. 'Straf verspreiders van gemanipuleerde naaktbeelden' (25-11-2020; De Telegraaf)
9. Minister wil verspreiding gemanipuleerde naaktfilmpjes aanpakken (26-11-2020; Noordhollands Dagblad)
10. Complottheorieën zijn de zichtbare verschijnselen van verslaving (28-11-2020; Trouw)
11. Expert waarschuwt voor 'infocalyps': nepnieuws is oorlog tegen de samenleving (9-12-2020; De Stentor)
12. Sprankje licht (12-12-2020; Reformatorisch Dagblad)
13. Op de wereld om te leren (17-12-2020; Leeuwarder Courant)
14. Een zombie kan leren om je aan te kijken (22-12-2020; NRC)
15. Channel 4 zendt 'deepfake' kerstboodschap van Elizabeth uit (24-12-2020; De Telegraaf)
16. Britse mediawaakhond krijgt 200 klachten over deepfake-toespraak koningin (29-12-2020; NRC)
17. Channel 4 krijgt ruim 200 klachten over deepfake Queen Elizabeth: 'Respectloos' (29-12-2020; Metro)
18. Techblog - Britten klagen over deepfake-toespraak koningin Elizabeth: 'Slechte smaak' (30-12-2020; De Volkskrant)
19. Angela de Jong kritisch over RTL Exclusief: 'Oudejaarsshow Martijn Koning was grappiger' (31-12-2020; Brabants Dagblad)
20. Matthijs van Nieuwkerk: Satirisch schouderduwtje hoort erbij (31-12-2020; De Telegraaf)
21. Het doorbraakjaar van de deepfake (2-1-2021; Dagblad van het Noorden)
22. Technologie maakt het de leugen steeds makkelijker (2-1-2021; Trouw)
23. Radicale Trump-fans geloven niet dat veroordelende video president echt was (8-1-2021; NRC)
24. NPO-baas Shula Rijxman: 'Aanvallen op journalistiek gevaarlijk en onacceptabel' (11-1-2021; Het Parool)
25. Shula Rijxman: 'Aanvallen op journalistiek gevaarlijk en onacceptabel' (11-1-2021; De Gelderlander)
26. 'Bedreigen van pers is niet te tolereren' (12-1-2021; Het Parool)
27. Laserstralen (21-1-2021; De Gelderlander)
28. De deepfakes van Cruijff bij Beau waren extreem ongemakkelijk (25-1-2021; Het Parool)
29. Mediawaakhond doet niets met klachten deepfake-video Queen (26-1-2021; De Telegraaf)
30. 'Pak pesten met nepnaaktfoto aan' Roep LR om Rotterdamse campagne (4-2-2021; De Telegraaf)
31. Gerri Baudet presenteert GVD-journaal in hilarische deepfake Lubach (8-2-2021; Metro)
32. Kijkcijferrecord voor Zondag met Lubach met 'broer' van Thierry Baudet (8-2-2021; Brabants Dagblad)
33. Dankzij Gerri Baudet kan niemand meer om de feiten heen (8-2-2021; Trouw)
34. IJzeilen en de Elfstedentocht van 1933: vooroorlogse pret op Nederlands natuurijs (9-2-2021; Trouw)
35. Thierry Baudet wordt meelijwekkend, maar hij heeft wel twaalf boeken geschreven (9-2-2021; De Volkskrant)
36. Waarheid of waanzin: nepnieuws doorzien (17-2-2021; Reformatorisch Dagblad)
37. Bellingcat: wereldkampioen digitaal speuren (19-2-2021; NRC)
38. Hannover Messe Challenge, technische krachtmeting van UT, gaat dit jaar online: opgave

- scholierenteam is nog mogelijk (26-2-2021; Tubantia)
39. Spookmodel (6-3-2021; De Telegraaf)
  40. Wat is 'echt' en wat is 'nep'? (6-3-2021; Friesch Dagblad)
  41. Een voorproefje van de 'infocalyps': het Harry en Meghan-interview zal een deepfake worden genoemd (8-3-2021; Trouw)
  42. Hilde komt tot leven met kunstmatige intelligentie (10-3-2021; Noordhollands Dagblad)
  43. Desinformatie bedreigt onze democratie (15-3-2021; Trouw)
  44. Het overkomt niet alleen jongeren (17-3-2021; BN/De Stem)
  45. De beste culturele tips voor deze lockdownweek (18-3-2021; Het Parool)
  46. Door de aangescherpte coronamaatregelen zijn culturele instellingen tot en met 30 maart 2021 gesloten. Ontdek hier wat er de komende periode te doen is. We proberen in onze agenda zo actueel mogelijk te zijn, maar check ook altijd de websites van de culturele instellingen. (24-3-2021; De Havenloods Alexander)
  47. Wat is dat, nepnieuws? (25-3-2021; NRC)
  48. Bewijs van echtheid (27-3-2021; De Volkskrant)
  49. Angsten bestrijden met deepfake artificial intelligence Nepnieuws (2-4-2021; Leidsch Dagblad)
  50. Krakende delen in het huis van de democratie (3-4-2021; Elsevier Weekblad)
  51. Kernenergie: wat als er weer oorlog komt? (19-4-2021; Leeuwarder Courant)
  52. Europa komt als eerste op de wereld met regels voor kunstmatige intelligentie. 'Dit wordt een feest voor juristen' (21-4-2021; De Volkskrant)
  53. Biometrie-experts en nieuw onderzoek laten zien hoe gedragsbiometrie betere gegevensbescherming biedt die aan strenge regelgeving voldoet (21-4-2021; Business Wire Nederland)
  54. EU komt als eerste ter wereld met regels voor kunstmatige intelligentie (22-4-2021; De Volkskrant)
  55. Nepmensen creëren is niet democratisch (22-4-2021; Trouw)
  56. Deepfake-video van Rutte, die een acteur aanraadt om op date te gaan, blijkt écht (23-4-2021; Eindhovens Dagblad)
  57. Door schrijver gepubliceerde en verwijderde video van Rutte blijkt niet nep (23-4-2021; NRC)

## Bijlage 2: Lijst van artikelen en relevante bronnen over de zaak imitatie-Volkov

1. Леонид Волков: Мы должны быть самыми сильными (17-9-2018, (geen tijd); YouTube-kanaal After Empire После Империи, [https://www.youtube.com/watch?v=Bp4\\_yz741kc](https://www.youtube.com/watch?v=Bp4_yz741kc))
2. Tweet van Rihard Kols over de casus (22-4-2021, 10:56; Twitter-account van Rihard Kols, <https://twitter.com/RihardsKols/status/1385155498979348481>)
3. European MPs targeted by deepfake video calls imitating Russian opposition (22-4-2021, 18:02; The Guardian, <https://www.theguardian.com/world/2021/apr/22/european-mps-targeted-by-deepfake-video-calls-imitating-russian-opposition>)
4. Facebook-post van Leonid Volkov (22-4-2021, (geen tijd); Facebook van Leonid Volkov, <https://www.facebook.com/leonid.m.volkov/posts/4004434886245781>)
5. Europese parlementsleden slachtoffer van nepvergadering met Russische oppositie (23-4-2021, 13:06; De Volkskrant, <https://www.volkskrant.nl/nieuws-achtergrond/europese-parlementsleden-slachtoffer-van-nepvergaderingen-met-russische-oppositie~b1fa1b95/>)
6. Tweet van Rihard Kols met ondertekend stuk over de casus (23-4-2021, 14:48; Twitter-account van Rihard Kols, <https://twitter.com/RihardsKols/status/1385576305056534533?s=20>)
7. Venemaa "naljamehed" esinesid Balti poliitikutele Navalnõi lähikondsena (23-4-2021, 14:53; Eesti Rahvusringhääling, <https://www.err.ee/1608188794/venemaa-naljamehed-esinesid-balti-poliitikutele-navalnoi-lahikondsena>)
8. Kamerleden vergaderen met deepfake-imitatie van stafchef Russische oppositieleider Navalny, (23-4-2021, 20:56; De Volkskrant, <https://www.volkskrant.nl/nieuws-achtergrond/kamerleden-vergaderen-met-deepfake-imitatie-van-stafchef-russische-oppositieleider-navalny~b04b5322/>)
9. Nederlandse Kamerleden voerden gesprek met imitatie-stafchef Navalny (23-4-2021, 20:58; NRC, <https://www.nrc.nl/nieuws/2021/04/23/ook-nederlandse-kamerleden-voerden-gesprek-met-imitatie-stafchef-navalny-a4041164>)
10. Kamerleden spraken bedrieger in plaats van medewerker Navalny (23-4-2021, 21:09; De Telegraaf, <https://www.telegraaf.nl/nieuws/22641751/kamerleden-spraken-bedrieger-in-plaats-van-medewerker-navalny>)
11. Kamerleden vergaderen met imitatie van Navalny's stafchef (23-4-2021, 22:11; NU.nl, <https://www.nu.nl/tech/6129559/kamerleden-vergaderden-met-imitatie-van-navalnys-stafchef.html>)
12. Kamerleden spraken bedrieger in plaats van medewerker Navalni (23-4-2021, 22:45; Noord Hollands Dagblad, [https://m.noordhollandsdagblad.nl/cnt/DMF20210423\\_38900084?origin=app&utm\\_source=google&utm\\_medium=organic](https://m.noordhollandsdagblad.nl/cnt/DMF20210423_38900084?origin=app&utm_source=google&utm_medium=organic))
13. Kamerleden spraken bedrieger in plaats van medewerker Navalny. Hoe werkt de deepfake-techniek die hierachter steekt? (23-4-2021, 22:56; Trouw, <https://www.trouw.nl/binnenland/kamerleden-spraken-bedrieger-in-plaats-van-medewerker-navalny-hoe-werkt-de-deepfake-techniek-die-hierachter-steekt~bffc158b/>)
14. Kamerleden bedrogen met deepfake-imitatie van naaste medewerker Navalny (23-4-2021, 22:56; Tubantia, <https://www.tubantia.nl/politiek/kamerleden-bedrogen-met-deepfake-imitatie-van-naaste-medewerker-navalny~a63eff35/>)
15. Tweede Kamer had videogesprek met deepfake-imitatie van stafchef Navalny (23-4-2021,



- 23:09; RTL Nieuws, <https://www.rtlnieuws.nl/nieuws/nederland/artikel/5227214/tweede-kamer-navalny-deepfake-volkov>)
16. Nederlandse Kamerleden misleid door gesprek met nepmedewerker Navalny (23-4-2021, 23:17; NOS, <https://nos.nl/artikel/2377987-nederlandse-kamerleden-misleid-door-gesprek-met-nepmedewerker-navalny>)
  17. Kamerleden vergaderen met deepfake-imitatie stafchef Navalny (24-4-21, 08:17; Linda, <https://www.linda.nl/nieuws/politiek/kamerleden-vergaderen-deepfake-imitatie-stafchef-navalny/>)
  18. Kamerleden spraken met nepmedewerker Navalny (24-4-2021, 9:56, Nieuwsuur, <https://twitter.com/Nieuwsuur/status/1386046440795934720>)
  19. 'Tweede Kamerleden hadden videogesprek met deepfake stafchef Navalny' (24-4-2021, 11:05; Tweakers, <https://tweakers.net/nieuws/180898/tweede-kamerleden-hadden-videogesprek-met-deepfake-stafchef-navalny.html>)
  20. Kamerleden spraken met een nepversie van stafchef van Navalny: deepfake of dubbelganger?(24-4-21, 14:57, De Volkskrant, <https://www.volkskrant.nl/nieuws-achtergrond/kamerleden-spraken-met-een-nepversie-van-stafchef-van-navalny-deepfake-of-dubbelganger~b57ab1ab/>)
  21. De uitzending van 24 april: Huisartsen woedend op RIVM-vaccinatiebaas / Druk in Amphia-ziekenhuis onverminderd hoog / Kamerleden praten met nepmedewerker Navalny / Hoe kan Kamer slagkracht vergroten? (24-4-2021, 16:04; NOS/Nieuwsuur, <https://nos.nl/nieuwsuur/artikel/2378045-de-uitzending-van-24-april-huisartsen-woedend-op-rivm-vaccinatiebaas-druk-in-amphia-ziekenhuis-onverminderd-hoog-kamerleden-praten-met-nepmedewerker-navalny-hoe-kan-kamer-slagkracht-vergroten>)
  22. Veel onduidelijk over 'deepfake-gesprek' van Kamerleden met medewerker Navalny (24-4-2021, 18:00; NOS, <https://nos.nl/artikel/2378060-veel-onduidelijk-over-deepfake-gesprek-van-kamerleden-met-medewerker-navalny>)
  23. 'Prima dat Kamerleden in imitatie tuinen' (24-4-2021, 19:10; Dagblad van het Noorden, <https://dvh.nl/groningen/Prima-dat-Kamerleden-in-imitatie-tuinen-zegt-Groningse-deepfake-deskundige-Jarno-Duursma-26781789.html>)
  24. Tweede Kamer gefopt met deepfake tijdens Zoom-gesprek (24-4-2021, (geen tijd); Apparata, <https://www.apparata.nl/nieuws/tweede-kamer-gefopt-met-deepfake-stafchef-navalny-52373/>)
  25. SGP: man is niet 'hoofd vrouw' (24-4-2021, (geen tijd); Elsevier Weekblad)
  26. Het gevaar van 'deepfakes': Het begon met porno, maar ook Kamerleden trappen erin (25-4-2021, 07:17; De Limburger, [https://www.limburger.nl/cnt/dmf20210425\\_91962168](https://www.limburger.nl/cnt/dmf20210425_91962168))
  27. 'Nederlands gesprek met fake-Navalnyteam onderdeel van grote Russische desinformatie campagne' (25-4-2021, 18:39; De Telegraaf, <https://www.volkskrant.nl/nieuws-achtergrond/nederlands-gesprek-met-fake-navalnyteam-onderdeel-van-grote-russische-desinformatiecampagne~b7b1e3f3/>)
  28. Tweede Kamer had videogesprek met deepfake-imitatie van Russische stafchef Navalny (26-4-2021, 11:31; Villamedia, <https://www.villamedia.nl/artikel/tweede-kamer-had-videogesprek-met-deepfake-imitatie-van-russische-stafchef-navalny>)
  29. Не дипфейк, а грим. Розыгрыш латвийских журналистов – дело рук пранкеров Лексуса и Вована (26-4-2021, 14:36; Baltnews,

- [https://web.archive.org/web/20210426170041/https://baltnews.ee/Russia\\_West/20210426/1019466689/Ne-dipfeyk-a-grim-Rozygrysh-latviyskikh-zhurnalistov--delo-ruk-prankerov-Leksusa-i-Vovana.html](https://web.archive.org/web/20210426170041/https://baltnews.ee/Russia_West/20210426/1019466689/Ne-dipfeyk-a-grim-Rozygrysh-latviyskikh-zhurnalistov--delo-ruk-prankerov-Leksusa-i-Vovana.html))
30. Nederlandse parlementariërs spraken met digitaal vervalste Rus (26-4-2021, 15:25; Standaard, [https://www.standaard.be/cnt/dmf20210425\\_97936061](https://www.standaard.be/cnt/dmf20210425_97936061))
  31. Opinie: Je kunt dus doodleuk als deepfake met de Kamer vergaderen, en dat is géén grap (26-4-2021, 16:12; De Volkskrant, <https://www.volkskrant.nl/columns-opinie/opinie-je-kunt-dus-doodleuk-als-deepfake-met-de-kamer-vergaderen-en-dat-is-geen-grap~b0662c22/>)
  32. Commissie van Tweede Kamer in gesprek met deepfake van stafchef Navalny (26-4-2021, (geen tijd); Dutch IT Channel, <https://dutchitchannel.nl/672059/commissie-van-tweede-kamer-spreekt-met-deepfake-van-stafchef-navalny.html>)
  33. Kamerleden gefopt met deepfake (26-4-2021, (geen tijd); Reformatorisch Dagblad)
  34. Ook Navalny's netwerk op de korrel (26-4-2021, (geen tijd); Trouw)
  35. Russisch bedriegersduo bediende zich van schmink, geen deepfake (27-4-2021, (geen tijd); Raam op Rusland, <https://raamoprusland.nl/actueel/1859-russisch-bedriegersduo-bediende-zich-van-schmink-geen-deepfake>)
  36. Je kunt niet meer geloven wat je ziet (27-4-2021, (geen tijd); Reformatorisch Dagblad)
  37. Spraken Kamerleden met een deepfake of met een komiek in een badjas? (28-4-2021, 20:39, NRC)
  38. 'Het is moeilijk een goede deepfake-video te maken' (28-4-2021, (geen tijd); NRC, <https://www.nrc.nl/nieuws/2021/04/28/het-is-moeilijk-een-goede-deepfake-video-te-maken-a4041727>)
  39. Deepfake? 'Vraag of hij zich snel wil omdraaien' (29-4-2021, 16:27; FD, <https://fd.nl/ondernemen/1381939/deepfake-vraag-of-hij-zich-snel-wil-omdraaien>)
  40. Nepvideo stafchef Navalni geen deepfake (29-4-2021, (geen tijd); Reformatorisch Dagblad)
  41. Видео-пранк с Верховной Радой и Слугой Народа (29-4-2021, (geen tijd); YouTube-kanaal Vovan222prank, <https://web.archive.org/web/20210914065346/https://www.youtube.com/watch?v=MykU8zozL7g>)
  42. 'Deepfake' that supposedly fooled European politicians was just a look-alike, say pranksters (30-4-2021, 18:10; The Verge, <https://www.theverge.com/2021/4/30/22407264/deepfake-european-politicians-leonid-volkov-vovan-lexus>)
  43. Deepfake ondermijnt liberale democratie ; Column (30-4-2021, (geen tijd); NRC, <https://www.nrc.nl/nieuws/2021/04/30/deepfake-ondermijnt-liberale-democratie-a4041871>)
  44. Duo achter imitatie stafchef Navalny: 'het was geen deepfake maar dubbelganger' (1-4-2021, 13:01; Tweakers, <https://tweakers.net/nieuws/181136/duo-achter-imitatie-stafchef-navalny-het-was-geen-deepfake-maar-dubbelganger.html>)
  45. Russen hielden Tweede Kamer voor de gek met schmink: 'Dit is niet grappig' (4-5-2021, 15:55, Tubantia, <https://www.tubantia.nl/tech/russen-hielden-tweede-kamer-voor-de-gek-met-schmink-dit-is-niet-grappig~af5991d9/>)
  46. Geen kat, wel digibeet (5-5-2021, (geen tijd); De Volkskrant)
  47. Op het tweede gezicht (8-5-2021, (geen tijd); De Volkskrant)
  48. Video-prank with the Parliament of the Netherlands (Eng)(27-5-2022, (geen tijd); Youtube-

- kanaal VovanPrank222,  
<https://web.archive.org/web/20210529192106/https://www.youtube.com/watch?v=rmeQkGNO2Zs>)
49. Deepfake? Nee, gewoon een slecht geschminkte grapjurk in een badjas (28-5-2021, 11:31; De Volkskrant, <https://www.volkskrant.nl/nieuws-achtergrond/deepfake-nee-gewoon-een-slecht-geschminkte-grapjurk-in-een-badjas~b2866b50/>)
  50. Kamervoorzitter Bergkamp: gesprek met nep-Volkov had voorkomen kunnen worden (28-5-2021, 14:15; NRC, <https://www.nrc.nl/nieuws/2021/05/28/kamervoorzitter-bergkamp-gesprek-met-nep-volkov-had-voorkomen-kunnen-worden-a4045278>)
  51. 'Medewerker Navalny' die met Kamerleden sprak geen deepfake maar acteur (28-5-2021, 14:23; NU.nl, <https://www.nu.nl/tech/6136120/medewerker-navalny-die-met-kamerleden-sprak-geen-deepfake-maar-acteur.html>)
  52. Video van door Russen voor de gek gehouden Kamerleden: 'Sturen jullie bitcoins?' (28-5-2021, 14:44; Het Parool, <https://www.parool.nl/nederland/video-van-door-russen-voor-de-gek-gehouden-kamerleden-sturen-jullie-bitcoins~b140a28f/>)
  53. In badjas (28-5-2021, 15:26; De Volkskrant, <https://www.volkskrant.nl/columns-opinie/in-badjas~b9988f30/>)
  54. Nu ook video van voor de gek gehouden Kamerleden: 'Sturen jullie bitcoins?' (28-5-2021, (geen tijd); Tubantia, <https://www.tubantia.nl/politiek/nu-ook-video-van-voor-de-gek-gehouden-kamerleden-sturen-jullie-bitcoins~a140a28f/>)
  55. Gesprek met Russische bedrieger online gezet (29-5-2021, (geen tijd); Friesch Dagblad)
  56. [Niet het ministerie van Volksgezondheid, Welzijn e...]\* (29-5-2021, (geen tijd); Nederlands Dagblad)
  57. Nep? Echt? Zet het maar... (29-5-2021, (geen tijd); Het Parool)
  58. Russische komieken leiden Kamer om tuin (31-5-2021, (geen tijd); Reformatorisch Dagblad)
  59. Pranks met politici: van een 'medewerker' van Navalny in badjas tot de vrijlating van 'terrorist' Jablabla/Kamerleden merken wel wanneer ze voor de gek worden gehouden. Of niet? (1-6-2021, (geen tijd); Trouw, <https://www.trouw.nl/politiek/pranks-met-politici-van-een-medewerker-van-navalny-in-badjas-tot-de-vrijlating-van-terrorist-jablabla~bd247eef/>)
  60. Na bedrieger spreekt Kamer met echte medewerker van Navalni (7-6-2021, 10:10; Noord Hollands Dagblad, [https://www.noordhollandsdagblad.nl/cnt/dmf20210607\\_53230646?utm\\_source=google&utm\\_medium=organic](https://www.noordhollandsdagblad.nl/cnt/dmf20210607_53230646?utm_source=google&utm_medium=organic))
  61. Gemist? Grootschalige Whatsappfraude en één op de tien studentes slachtoffer van verkrachting (9-6-2021, 10:00; AD/Algemeen Dagblad, <https://www.ad.nl/home/gemist-grootschalige-whatsappfraude-en-een-op-de-tien-studentes-slachtoffer-van-verkrachting~ae894ad2/>)
  62. Nep-Volkov misleidde niet alleen de Tweede Kamer: 'Er zijn honderden pogingen geweest' (9-6-2021, 15:39; Het Parool, <https://www.parool.nl/nederland/nep-volkov-misleidde-niet-alleen-de-tweede-kamer-er-zijn-honderden-pogingen-geweest~b943ae6c/>)
  63. Echte Navalny-medewerker tegen Tweede Kamer: Er waren 12 succesvolle gesprekken met nep-Volkov (9-6-2021, 16:03; De Gelderlander, <https://www.gelderlander.nl/politiek/echte-navalny-medewerker-tegen-tweede-kamer-er-waren-12-succesvolle-gesprekken-met-nep-volkov~a943ae6c/>)

64. Wat zijn deepfakes en hoe gevaarlijk zijn ze? (20-1-2023, (geen tijd); VPNgids, <https://www.vpngids.nl/veilig-internet/cybercrime/wat-zijn-deepfakes/>)

## Bijlage 3: Transcripten Interviews

Laurens Verhagen (12-6-2023, 13:30)

Interviewer: “Ok, want u had geschreven over deze casus met deze imitatie-Volkov.”

Geïnterviewde: “Ja, in 2021 was dat.”

Interviewer: “Ja. Ik heb hier wat artikelen open staan. Dan vroeg ik me af, even kijken, hoe kwam u initieel te weten over deze casus?”

Geïnterviewde: “Uhm, poeh. Uhm, volgens mij had de internationale media, The Guardian, er over geschreven. Euhm, ik denk dat dat de aanleiding was en ik heb het samen met een collega geschreven.”

Interviewer: “Ja.”

Geïnterviewde: “En die heeft ook vanuit Den Haag ook wat gehoord en dat hebben we samen geschreven.”

Interviewer: “Ja, want de griffie van de Tweede Kamer wordt als bron genoemd in het artikel. ”

Geïnterviewde: “Ja, ja.”

Interviewer: “En hoe is dat contact precies verlopen? Want ik vroeg mij af, even kijken, in de header staat dat, euhm, er een gesprek gevoerd is met een deepfake-imitatie. Is dat ook specifiek bevestigd door de griffie? Of weet u dat niet?”

Geïnterviewde: “Euhm, nee, dat was niet- was volgens mij- even er bij pakken. Want dat was de verwarring. Kijk, de verwarring was sowieso dat het in de eerste instantie eigenlijk iedereen er van overtuigd was dat het een deepfake was en dat bleek uiteindelijk niet zo te zijn. Hé, dat is een beetje de kern van waar het over ging.”

Interviewer: “Ja, daar gaat mijn scriptie over, ja.”

Geïnterviewde: “Euhm. Die euhm... Vaste Kamercommissie die heeft zover ik me dat kan herinneren het woord “deepfake” niet genoemd.”

Interviewer: “Ja.”

Geïnterviewde: “Euhm, volgens mij hadden ze dat in algemene bewoordingen iets over imitaties of zoiets. En euh, even kijken of ik dat, euh, oh ja. Nou ja: Vrijdag weigerden verschillende leden van de commissie aanvankelijk te bevestigen of ze met het échte team hadden gesproken. Vrijdagavond volgde de verklaring en die luidde toen: “Inmiddels heeft de commissie bevestiging gekregen dat het gesprek niet gevoerd is met Volkov zelf, maar met iemand die zich als hem voordeed.” [Geïnterviewde citeert hier een deel van het Volkskrantartikel “Kamerleden vergaderen met deepfake-imitatie van stafchef Russische oppositieleider Navalny”] Nou ja, dat is in hele algemene bewoordingen. Dus [onverstaanbaar] geen deepfake. Waarom wij in eerste instantie dat wel zo hadden opgeschreven, euh, was omdat euh het team van Volkov zelf dat had gezegd. Zij waren er zelf van overtuigd dat het een deepfake was.”

Interviewer: “Ja, ok. Jullie hadden ook- even kijken. Dan- dus dat was inderdaad initieel die fout. Even kijken, want een dag later heeft u zelf een artikel geschreven, alleen, volgens mij, over deze

casus en dan, want het is wel wat-”

Geïnterviewde: “Ja, toen en toen...”

Interviewer: “Ja.”

Geïnterviewde: “Precies, omdat er verwarring was. Dat was van 24 april: “Kamerleden spraken met een nepversie van stafchef van Navalny: deepfake of dubbelganger?” Dat is dan inderdaad de volle worp. Euhm... nou ja daar staat ook nog een keer: “‘Welkom in het deepfake-tijdperk’, stelt Volkov zelf.” Euhm, Richard- Rihard Kols, die tweet. Dat is aan de hand van het statement van, nou ja, daar hebben ze het ook over AI. Ik heb toen, even kijken, Janno Duursma gesproken, de AI-expert, die ervan overtuigd was ook dat het een deepfake was. Guardian. Nou ja, kortom, heel veel mensen. Alleen, gedurende een dag later kwam toch de twijfels en toen heb ik ook met Theo Gevers gesproken, die dat formuleert eigenlijk, die twijfel formuleert van nou ja, heel lastig om iets te zeggen zonder die video-opname. En dat was natuurlijk ook het lastige, dat wij zelf die video-opname ook niet hadden, en die nuanceerde dat vooral. Belangrijk, later bleek ook wel dat die gelijk had.”

Interviewer: “Ja, dat klopt. Waar kwam die twijfel dan specifiek vandaan? Weet u dat nog of?”

Geïnterviewde: “Ik denk via- via hem.”

Interviewer: “Theo Gevers.”

Geïnterviewde: “Ik weet het niet meer precies. Ik denk dat ik hem daarna- dat gedurende de dag hem heb gesproken hierover. Kan ik me voorstellen, want kijk als zoiets gebeurt, zit er ook altijd enige tijddruk achter. Van, euh, ok: dit is er aan de hand en je belt experts en niet elke expert is gelijk beschikbaar.”

Interviewer: “Ah, ok, ja.”

Geïnterviewde: “Die laatste aan heb ik dan Theo Gevers gesproken. Ja, hoe die twijfel dan precies ontstaat weet ik niet. Misschien bij mezelf... euh... of... ja, geen idee. Hoe dan ook, ik heb Theo Gevers toen gebeld. Die uh... nou ja, van wat ik al zei. Van ja dit is... nou ja, we weten het niet dus voorzichtige formulering, maar vermoedelijk is het niet het geval. En die neem ik dan net wat serieuzer, omdat hij... nou ja, echte expert is.”

Interviewer: “Ja, hij is wat- hij weet wat meer dan meneer Duursma.”

Geïnterviewde: “Ja, hij is hoogleraar computer vision, ja. Hij zit er op, zeg maar.”

Interviewer: “Ja, ok, en dan vroeg ik me af- u kon echt alleen afgaan op dat screenshot van Rihard Kols.”

Geïnterviewde: “Ja, meer hadden we niet. Ja, klopt. Enne euhm... dat maakt het ook lastig en, ik weet niet wat je vervolgvraag is maar als ie zou luiden: had je het iets anders gedaan? Dan is het antwoord natuurlijk ja. Dan had het in het eerste stuk- had het iets voorzichtiger moeten staan. Euh... dat het... nu stond het er heel stellig dat het een deepfake was.”

Interviewer: “Ja, en dat... Ja, ok, dat was-”

Geïnterviewde: “Nou ja, dat vind ik zelf hoor te stellig. Vandaar ook toen dat stuk er na. Om dat te nuanceren. ”

Interviewer: “Is dat ook waarom er een update in staat? In het artikel: update 30 april. ”

Geïnterviewde: “Even kijken hoor.”

Interviewer: “Helemaal op het eind.”

Geïnterviewde: “Jjjjjj”

Interviewer: “Dat van Duursma.”

Geïnterviewde: “Oh daar. Geëxpliciteerd dat uitspraken niet deed op basis van een video, maar op basis van de screenshots die... [Geïnterviewde citeert hier een deel van een update die staat aan het eind van het Volkskrantartikel “Kamerleden vergaderen met deepfake-imitatie van stafchef Russische oppositieleider Navalny”] Poeh, ja, dat was expliciet een verzoek van hemzelf. Want hij voelde zich natuurlijk ook wat ongemakkelijk daarbij, omdat hij heel stellig was. Ook van dit is een next-level deepfake. Een dag daarna bleek dat het niet zo was. En het lijkt belangrijk dat dit er nog even bij stond, nou ja.”

Interviewer: “Ik kan me die verwarring ook wel voorstellen, want toen ik deze casus heb onderzocht bleek dat dit een foto is van de echte Volkov uit een interview uit 2018, ja.”

Geïnterviewde: “Klopt. Ja, ja. Dus iedereen- weet je, iedereen moet op basis van artikelen en uitspraken en andere media doet dan uitspraken. En dan is ook Theo Gevers, prettig die academicus heeft dan zoiets van nou ja, daar moet je heel voorzichtig mee zijn, als je die video niet hebt gezien, dan kan je daar dan ook niks over zeggen.”

Interviewer: “Ja ok, en wanneer dus- wanneer bleek volgens u echt dat het niet ging om een deepfake hierbij, denkt u?”

Geïnterviewde: “Oeh, dat weet ik niet meer, want ik- ik weet niet of dat, euhm... of we daar over hebben geschreven. Dat weet ik eerlijk gezegd niet meer. Ik kan mij herinneren dat vrij kort daarna, uh nou, dat hebben we wel geschreven denk ik. Want volgens mij heeft dat team dat zelf onthuld dat dat ze gewoon- want toen kwamen die beelden alsnog vrij.”

Interviewer: “Ja ja ja, klopt. 28 mei-artikel: geschminkte grapjurk, badjas.”

Geïnterviewde: “Ja, precies.”

Interviewer: “En even kijken, zou u dan, ja even kijken, toen dit bleek, hoe zijn dan correcties doorgevoerd intern? Wat was de reactie hierop intern?”

Geïnterviewde: “Sorry, wat zeg je?”

Interviewer: “Euhm, wat was de reactie, ja, intern bij De Volkskrant toen dit bleek laten we maar zeggen?”

Geïnterviewde: “Euh, volgens mij hebben wij daar toen ook nog een stuk over geschreven. Even kijken hoor.”

Interviewer: “Stuk...”

Geïnterviewde: “Ja, onze zoekfunctie is een beetje kapot.”

Interviewer: “[Gelach] ja, nee ik had...”

Geïnterviewde: “Leonid Volkov... één moment. Ik kan me herinneren dat we er verschillende stukken over hebben geschreven.”

Interviewer: “Ja, even kijken.”

Geïnterviewde: “Euhm, even kijken... 28 mei: “Deepfake? Nee, gewoon een slecht geschminkte grapjurk in een badjas. Deepfake? Nee, toch een gek in een badjas. Zoveel is nu definitief duidelijk.” [Geïnterviewde citeert hier een deel van het Volkskrantartikel “Deepfake? Nee, gewoon een slecht geschminkte grapjurk in een badjas.”] Nou ja, toen werd dat filmpje pas euh... vrijgegeven. Dus toen was het ook duidelijk van dat, ja, hoe reageer je daar op? Ja, zo dus, door nog een update even een stuk er achter te schrijven.”

Interviewer: “Ok, en zou u dan ook-”

Geïnterviewde: “En toen heb ik dat ook die video, nou ja, dat je dan echt zeker zeker wil weten, die video nog even aan Theo Gevers voorgelegd zodat ie echt kon kijken een een definitief oordeel kon vellen.”

Interviewer: “Ja ja.”

Geïnterviewde: “Euh... waarmee dat, zijn eerdere vermoedens uit- van een maand eerder werden bevestigd.”

Interviewer: “Ja ok, nee, dat lijkt me dan wel logisch en zou u dan nog- waarom is er geen laten we maar zeggen retractie geplaatst bij het initiële artikel?”

Geïnterviewde: “Nee, dat dat, je bedoelt een verwijzing van het eerdere artikel-”

Interviewer: “Ja, bij dat eerste artikel van 23 april. Zo van: het is geen deepfake of...”

Geïnterviewde: “Ja, dat doen we eigenlijk nooit want dan blijf je bezig, weet je. Nieuws ontwikkelt zich en dat zou je dan bij elk stuk moeten doen. Je schrijft natuurlijk heel vaak dingen die uiteindelijk anders blijken te zijn. Je doet dat nooit met terugwerkende kracht. Ja, idealiter zou je dat natuurlijk wel hebben dat je daarmee dan weer verwijst naar de ontwikkelingen later. Een ander voorbeeld: de Nord Stream pijpleiding die is opgeblazen. Mededeelden media in eerste instantie... waarschijnlijk ook van nou werk van Russen, nu blijkt het weer iets genuanceerder te liggen. Ja, dan ga je niet bij oude artikelen- dat is ook in praktische zin niet te doen.”

Interviewer: “Ja, tijdsgebrek.”

Geïnterviewde: “Ja, nou ja, dat niet alleen. Dat het... dat het ook een soort principieel punt is. Ja, dingen die je hebt gepubliceerd, die moet je niet achteraf weer gaan aanpassen, lijkt mij.”

Interviewer: “Ok, dan had ik nog een vraag, even kijken, want ik merkte ook dat- jullie hebben ook op 28 april [sic; moet 28 mei zijn] dat artikel gepubliceerd over die grapjurk dus, over Lexus en Vovan. Maar ik kwam ook tegen dat al- dat ze al eerder beeld hadden vrijgegeven, maar dan van hun gesprek met Oekraïense politici. Hadden jullie dat gezien?”

Geïnterviewde: “Euh...”

Interviewer: “Een paar weken terug dan.”



Geïnterviewde: “Dat weet ik niet. Hebben we dat toen...? Ik kan met niet herinneren dat we dat ook hebben gebracht, want we wilden wachten op de beelden van de Nederlandse Kamercommissie. Want dat was de insteek.”

Interviewer: “Ok dus-”

Geïnterviewde: “Dus want anders blijf- blijf je bij vermoedens en we wilden liever wachten tot we het echt, echt 100% zeker wilden- wisten.”

Interviewer: “Ok, nou dan dat is dan wel wat ik wou weten denk ik.”

Geïnterviewde: “Ok, nou, ik hoop dat je er wat aan hebt.”

Interviewer: “Ja bedankt en als u nog iets kwijt wil, zeg het maar, of iets toe te voegen hebt.”

Geïnterviewde: “Nee hoor, dat is prima zo.”

Interviewer: ”Nou, bedankt voor uw tijd in ieder geval.”

Geïnterviewde: ”Ja, en success.”

Interviewer: “Ja, doei.”

Geïnterviewde: “Ok, oi.’

Mark Hendrikman (14-6-2023, 12:30)

Interviewer: “Ja, misschien kunt u zich even kort voorstellen om te beginnen.”

Geïnterviewde: “Ja, zeker. Ik ben Mark Hendrikman. Ik ben nieuwsredacteur voor Tweakers, tegenwoordig in vaste dienst in de tijd van, euhm. Ik ben sinds een jaar in vaste dienst. Daarvoor ben ik heel lang nieuwsfreelancer geweest voor Tweakers. En ten tijde van dit Volkskrantstuk, of eigenlijk het deepfakeverhaal waar we het over gaan hebben, was ik ook freelancer bij Tweakers.”

Interviewer: “Ja, ok. En dan wou ik eigenlijk... Ik heb nog een paar vragen, maar ik wou vragen van... Hoe kwam u eigenlijk tot het verhaal? Is dit dan iets wat u uit de Volkskrant gehaald heeft? Of is dat dan-”

Geïnterviewde: “Ik geloof... Ja, dat is dus achteraf lastig terug te halen. Maar in een algemene zin... We houden... Euhm... Ja. Andere nieuwsmedia houden we natuurlijk in de gaten. Of zij verhalen publiceren die ook voor ons interessant zijn. Tegelijkertijd kunnen lezers ons tips sturen van: hey dit kan interessant zijn. En De Volkskrant volg ik niet zo op de voet tijdens mijn eigen nieuwsdiensten, omdat De Volkskrant doorgaans nieuwsverhalen publiceert die ook voor ons interessant zijn. Dus als ik d'r op terugkijk, dan vermoed ik dat dit een tip van een lezer is geweest.”

Interviewer: “Ok, dus van een lezer. En dan van: dit is interessant en dan meer vanuit- want ik merk dat veel andere berichtgeving dan inging op... Maar ja, dan...”

Geïnterviewde: “Hoewel, om er nog even op terug te komen... Wij noteren ook de tipgever als er een tipgever is en die wenst genoteerd te worden. Dat is hier niet het geval. Dus dan weet ik het niet meer hoe we erop zijn gekomen.”

Interviewer: “Ok, dus misschien een anonieme tipgever dan?”

Geïnterviewde: “Heel soms geven lezers een tip en dan zeggen ze ik wil niet dat mijn username bij een artikel komt te staan. Maar dat is eigenlijk heel erg zeldzaam. Dus ik acht de kans groter dat we via andere wegen dit Volkskrantartikel ergens hebben gevonden. ”

Interviewer: “Ja, want ik zat me ook afvragen... Want, even kijken, jullie als techwebsite focussen natuurlijk meer op dat tech-aspect. Hé, want ik merk dat heel veel van de eerste berichtgeving, die eerste dag, dat dat meer ging op het politieke aspect. Is dat dan ook van: hé, interessant, deepfakeverhaal, we hebben daar zo naar gekeken. Dat is waarom het verhaal opgepakt is, meer dan die politieke implicaties?”

Geïnterviewde: “Ik denk dat je kan spreken van twee derde de technologie en één derde het politieke element. Maar dat het verhaal het niet overleefd had zonder dat politieke element. Toch wel. Want... Kijk, ten tijde van de publicatie van dit verhaal kennen we deepfakes al. Misschien niet *real-time*, maar in ieder geval achteraf geproduceerd. Dus dat was an sich niet interessant genoeg geweest. Maar hier werd echt een Kamercommissie, een Nederlandse Kamercommissie, werd beduust. Dus, ok, ik zei twee derde en één derde. Laten we er fifty-fifty van maken.”

Interviewer: “Ok, fifty-fifty, want de Tweede Kamer...”

Geïnterviewde: “Dat is toch wel de Tweede Kamer?”

Interviewer: “Ja. Nee, ik heb de beelden ook gezien. Het is wel een beetje...”

Geïnterviewde: “Ja?”

Interviewer: “Ja, nee, ze hadden het... Even kijken, het is inmiddels weer gedeletet, maar het staat wel op het internetarchief. Die twee... de Russen hadden het gepubliceerd, maar het account is weggehaald volgens mij, later. Ok. Even kijken, en ik vroeg me af, hebben jullie ook aan beeldchecking gedaan? Want even kijken, ik weet dat er in de tijd van dat eerste artikel een- één screenshot was. Hebben jullie daar nog iets mee gedaan verder? Voor het opschrijven of...”

Geïnterviewde: “Nee.”

Interviewer: “Nee?”

Geïnterviewde: “Nee, echt helemaal niet?”

Geïnterviewde: “Dat was... Iets van die in die trand was wel goed geweest om te doen, achteraf. Maar ik geloof dat dit verhaal in een weekend uitkwam. En daar heb ik dan nieuwsdienst en dan werk je in je uppie. En dit verontschuldigt mij helemaal niet, want ik heb hier gewoon een uitgleider gemaakt. Maar, euhm... je voelt je niet zo vrij, of ik voelde me op dat moment niet zo vrij om dieper in een verhaal te duiken. Omdat, je wil kwaliteit leveren, natuurlijk, maar je wil ook een bepaalde kwantiteit leveren. Als jij in zo'n weekend nieuwsdienst één verhaal publiceert en dat... de kwaliteit daarvan is geweldig, dan wordt je er nog steeds wel aangekeken van: eh dat is wel maar één verhaal. Dus er is een bepaalde, soort van ideale balans tussen kwaliteit en kwantiteit.”

Interviewer: “Wat zou dan de kwantiteit zijn? Hoeveel artikelen zit jij dan aan te denken?”

Geïnterviewde: “Als je er één per uur produceert, dan is dat... dan is dat goed. Nogmaals, een nieuwsdienstachtergrond verhaal is hartstikke anders. En daar kun je wel een beetje van afwijken, stel je werkt vier uur en je publiceert drie artikelen. Weet je, dat is ook ok als je daar een reden voor hebt, omdat je in één verhaal echt dieper bent gedoken. Ik heb dat op dit moment niet gedaan. Ik heb wel, nogmaals zonder te zeggen, het is allemaal de schuld van De Volkskrant. Ik heb de uitspraken van De Volkskrant en de mensen die in De Volkskrant geciteerd worden, die heb ik wel vertrouwd. En achteraf, onderbreek me als je meer structuur wil hoor, in dit verhaal, maar achteraf, want ik heb het zitten lezen. Achteraf kan je wel zeggen, euhm, de mensen die het woord deepfake gebruiken, die weten niet zo goed waar ze het over hebben. En dat had ik moeten zien. Weet je, er is een- er wordt een trendwatcher aangehaald in dat eerste verhaal en die experimenteert met deepfakes. Ja, dat is niet echt goed genoeg. En dan heb je die politici, zoals die stafchef die zelf geciteerd wordt, die hebben het over, oh dit is de deepfake, welkom in het tijdperk van de deepfake. Dit zijn politici en als het professionele politici zijn, dan mag je er echt van uitgaan dat hun technologische kennis, in 2021, niet goed genoeg is om die inschatting te maken.”

Interviewer: “Ok, ja, dat klopt ook, want even kijken, toen ik meneer Verhagen sprak, die even kijken, want jullie hebben ook een update gepubliceerd op 30 april, die bij het eerste artikel staat, onderaan volgens mij.”

Geïnterviewde: “Ja, de eerste update gekregen en de follow-up. ”

Interviewer: “Ja, want dat komt omdat er in De Volkskrant een follow-up gepubliceerd is en meneer Verhagen zei inderdaad ook zo van, ja, we hebben dus die hoogleraar computer vision gesproken, die Theo Gevers, want er was inderdaad ook al wel wat twijfel bij hun, ja ook. Dus dat was wel... ”

Geïnterviewde: “Bij de universiteit?”

Interviewer: “Nee, bij de Volkskrant. Die had hij dan later op de dag misschien gebeld, meneer Verhagen. Het is twee jaar geleden, maar goed. Ok, dus dat klopt wel, denk ik. Dus dat dan, ja, even kijken. Even kijken, hebben jullie later, ja even kijken, want je hebt twee artikelen hierover geschreven, dit eerste artikel en dan nog later één. Even kijken, dat gaat dan- dat is dan op een artikel

van The Verge.”

Geïnterviewde: “Klopt.”

Interviewer: “Kijk, wat mij opviel is dat Tweakers is volgens mij de enige publicatie die iets schrijft over dit artikel van The Verge. Dus dat, even kijken, heeft u dan misschien ook een idee van hoe dat komt of hoe zijn jullie bij dit artikel gekomen dan?”

Geïnterviewde: “Nou, deze is ook van mijn hand, heb ik gezien. Dat extra artikel op basis van The Verge. Euhm... als ik me goed herinner, maar ik ben hier ook eventjes aan het teruglezen. In die update die ik schrijf, wordt niet ondubbelzinnig gezegd van: joh, het is ook geen deepfake. Of we weten het definitief niet. Er wordt gezegd: ja, het kan. Zij komen tot een conclusie, maar eigenlijk hebben we ook niet de smoking gun. Het is een beetje onzeker. Ik denk dat de strekking van de update meer is. We waren stellig en we hadden eigenlijk niet stellig moeten zijn. Dus we zijn onstellig. En voor de rest staat het verhaal nog steeds zoals het staat. En euhm... dat euhm... ja, je kan dit later artikel op basis van The Verge, kan je een beetje zien als een echte rectificatie. Op basis van: kijk, dit verhaal van The Verge. Als wij dat eerste verhaal niet hadden gepubliceerd, dan hadden we het tweede verhaal ook niet gepubliceerd. Omdat, ok, goed, dat is dus die man in schmink, dat past niet bij tweakers. Weg d'r mee. Maar ik vond het destijds wel belangrijk om er eigenlijk een rectificatie... Ik gebruik het woord rectificatie niet, maar het fungeert eigenlijk als een rectificatie.”

Interviewer: “Ja, ok, dus dat is het idee erachter dan? Of dan...”

Geïnterviewde: “Toch wel, weet je, het gaat over mijn gedachtgang twee jaar terug, dus dat is lastig terug te halen, maar ik ken mezelf ook een beetje. En ik zou het wel belangrijk vinden om *de record straight* te zetten als dat zo gaat zoals het ging. En ik geloof dat ik ook echt heb- ergens heb geschreven: meerdere media, waaronder Tweakers, die zeiden dat het een deepfake was, of was dat in de update? In ieder geval, om een beetje context te geven, lezers kunnen bij ons reageren op artikelen. En het zijn best wel, weet je, al die techneuten, dat zijn best wel slimme lui. En ze zijn best wel, ze hebben best wel een hoge standaard. Ze zijn best wel kritisch. En het is op het internet anoniem, dus dan gaan ze er ook nog eens extra hard in. En als je bij Tweakers werkt, dan word je ook wel een beetje geconditioneerd van... Oeh, je moet het goed doen, hoor. Want als je het niet goed doet, dan hang je wel.”

Interviewer: “Dan krijg je de volle laag. ”

Geïnterviewde: “Ja, ja. We hebben, zeg maar, intern zeggen wij ook, de lezers zijn kritisch. Dus op het moment dat in de reacties onder een nieuwsartikel of een achtergrondartikel of een review, jij niet klachten krijgt, die echt aan jou geadresseerd zijn, maar de comments gaan alleen maar over de inhoud, dan heb je het hartstikke goed gedaan. Dus je stelt het een beetje bij, want het is gewoon zo zeldzaam dat een lezer zegt, he, goed artikel, wat fijn dat je dat hebt geschreven. Dat wordt niet echt gedaan. ”

Interviewer: “Nee, ok.”

Geïnterviewde: “Ze zijn eerder geneigd om te zeggen, flutartikel, nu kan ik maar beter bij De Volkskrant gaan werken. Nee, sorry, dat... De Telegraaf, sorry. Ik wilde niet De Volkskrant afbranden. Telegraafjournalist, dat worden we wel eens gezien.”

Interviewer: “Ja, de Telegraaf-reputatie. Ja, ok, even kijken, want bij dit artikel was er ook beeld van... even kijken... van dat gesprek uit Oekraïne.”

Geïnterviewde: “Hadden zij een beeld?”

Interviewer: “Ja, nee, The Verge hadden ze een... Ja, even kijken, want het staat ook volgens mij in het tweede artikel van... Ze hadden een... Ja, in ieder geval een afbeeld geproduceerd van een gesprek dat ze hadden gevoerd met Oekraïense politici toen. Als ik mij goed herinner.”

Geïnterviewde: “Ik zie hier staan dat die groep... een video op YouTube heeft gezet.”

Interviewer: “Ja, ja”

Geïnterviewde: “Die is unavailable. Op archive.org kan je die nog vinden, zeg je? ”

Interviewer: “Ja, volgens mij wel. Volgens mij is dat allemaal nog wel te vinden. Maar, even kijken, hadden jullie daar destijds iets mee willen doen met beeldverificatie of ook niet echt? Zo van, om te kijken: he nee, dit is geen deepfake, dat is gewoon geschminkt.”

Geïnterviewde: “Als ik het helemaal opnieuw zou kunnen doen, bedoel je?”

Interviewer: “Nou, even kijken, of hebben jullie daar in de tijd overwegingen over gedaan?”

Geïnterviewde: “Nee.”

Interviewer: “Nee?”

Geïnterviewde: “Ik moet er wel bij zeggen, ik was freelancer in der tijd. Dan word je echt niet meegenomen in redactiebeleid formuleren. Dus het zou kunnen dat- dat zeg maar de vaste redactie wel gesprekken had over: joh, wat doen wij met deepfakes in ons nieuwsproces? Maar euhm daar was ik dan weer niet onderdeel van. Dat kan ik eigenlijk niet zeggen.”

Interviewer: “Ok, dus dat... Nee, ok, maar- en u zelf heeft er ook niet echt iets mee bedacht van...”

Geïnterviewde: “Nee.”

Interviewer: “Nee, ok, maar...”

Geïnterviewde: “Nee, als je freelancer bent dan sta je natuurlijk ook ietsje verder weg d'r van. Dus er was meer een gevoel van, als ik voldoen aan mijn plichten, dan vind ik het ook goed. ”

Interviewer: “Ja, even kijken. En, even kijken, ik merkte ook dat een paar weken later hebben ze dus een filmpje gepubliceerd van hun gesprek met de Nederlandse... Op 28 zijn daar veel nieuwsartikelen over geweest. Tweakers heeft daar niet over gepubliceerd. ”

Geïnterviewde: “28 mei, ok.”

Interviewer: “Is daar nog een reden voor geweest? Of rond...”

Geïnterviewde: “Kan je me wat meer vertellen over die video's?”

Interviewer: “Even kijken. Ja, volgens mij hebben de bedriegers, Lexus en Vovan, die hebben volgens mij gewoon het merendeel van hun gesprek op hun YouTubekanaal gezet toen. En ja even kijken, heel veel- heel veel publicatie ging daar toen over heen. De Volkskrant berichtte toen ook weer van: bedrieger, nee, grapjas in een badj- grapjurk in een badjas. En ze hebben daar inderdaad ook weer die Theo Gevers op losgelaten. En die zegt van, nee, dit is geen deepfake.”

Geïnterviewde: “ Ja, ja, ja.”

Interviewer: “Dus ik vroeg me af van, he, er is hier veel over bericht, maar het heeft Tweakers...”

Geïnterviewde: “Euhm, ok, ja, meerdere mogelijke factoren. Ik werkte indertijd in de weekenden. Dus als zoiets door de weeks uitkomt, ja, en tegen de tijd dat het dan daarna weekend is, trein gemist, bedankt. Euhm, maar stel dat dit wel in mijn weekend langs zou komen. Dan denk ik nog steeds dat er, ja, dan weet ik wel zeker dat ik er niet over zou schrijven, want de aap uit de mauw. Het is geen deepfake. Daarvoor is het, daarmee is het technologie haakje dat wij vereisen voor nieuws, is weg. We hebben al hardop in een artikel eigenlijk gerectificeerd en geconcludeerd: geen deepfake. Dan is het afgelopen. Weet je dan. Ik snap dat het dan nog interessant is voor bijvoorbeeld De Volkskrant en allerlei andere nieuwsmedia, maar dan houdt het voor ons op.”

Interviewer: “Ja, ok, want ik vroeg me dat af, want ja, even kijken... Oh gut, het schiet me nu even weg... Euhm.”

Geïnterviewde: “Doe rustig aan.”

Interviewer: “Ja, even kijken. Oh, wat had ik nou zeggen? Euhm, ok, want kijk, waarom ik u ook wou interviewen. Ik vond juist dat de berichtgeving van Tweakers d'r wel goed, *in-depth* over was, laten we maar zeggen, dus ik wou ook eens kijken van: als ik kijk naar andere nieuwsmedia over deze casus, er werd- er waren veel lagen in en volgens mij waren jullie ook de eerste die echt over dat artikel van The Verge hebben bericht en dus ik had het idee van als andere nieuwsmedia dat misschien al opgepakt, dan was deze fout misschien een paar weken eerder gerectificeerd geweest dan tot iedereen op de Nederlandse berichtgeving wachtte.”

Geïnterviewde: “Oh, dus wat ik jou hoor zeggen is: wij deden die rectificatie op eenmalige basis van de Verge. En pas later, op basis van andere ontwikkelingen, kwamen andere nieuwsmedia naar buiten en zeiden: ze hebben geen deepfake by the way.”

Interviewer: “Ja, dus dat is iets...”

Geïnterviewde: “Nou, dat is toch nog iets goeds.”

Beide: [Gelach]

Interviewer: “Nee, maar dat bedacht me ook wel.”

Geïnterviewde: “Dat is wel apart.”

Interviewer: “Meneer Verhagen, die zei dan van: kijk, ze hebben dat toen niet gedaan, want het is niet van Nederland, het is Oekraïne. Het is... ze hebben echt gewacht op dat Nederlandse beeld, want een eerdere berichtgeving ging ook over het Nederlandse gesprek. Dus ja...”

Geïnterviewde: “Dus hun gedachtegang was, oh, de gesprekken tussen zeg maar de Baltische Staten en Oekraïne, dat dat dan geen deepfake was. Maar misschien was het gesprek met de Nederlandse Kamercommissie wel een deepfake en daarom...”

Interviewer: “Ja, kijk, ze wouden het graag helemaal zeker zeker weten, zei... zei uh....”

Geïnterviewde: “Ok, ja, goed, nou ja. Ik denk dat het hele thema van dit verhaal is dat er waarden zitten in dingen zeker weten. Tegelijk denk ik, als deze Russen de Baltische Staten en Oekraïne foppen met schmink, waarom zouden ze dan een deepfake van stal halen voor Nederland?”

Interviewer: “Ja, even kijken. Ik weet eigenlijk niet hoe dat zit in het Verenigd Koninkrijk, want ze

hebben het Verenigd Koninkrijk ook genept. Maar ik heb daar niet echt... Dat komt niet heel erg terug in de Nederlandse berichtgeving, maar daar heb ik ook naar gekeken. Ja. Ja, even kijken. Want wat ik ook dus bijvoorbeeld... Even kijken, jullie zijn volgens mij... Tweakers zijn ook de enige die volgens mij echt berichtgeven dat dat eerste screenshot, gedeeld door Rihard Koolstad, dat dat van de echte Volkov is.”

Geïnterviewde: “Deze? [geïnterviewede verwijst hier naar een afbeelding die de echte Volkov vergelijkt met de zogenaamde deepfake-imitatie uit het Tweakersartikel “Duo achter imitatie stafchef Navalny: 'het was geen deepfake maar dubbelganger'”]

Interviewer: “Ja, die rechtse. Dat is uit een interview in 2018, maar dat kwam in een andere berichtgeving ook niet terug. Want Lexus en Vovan hadden dat gebruikt voor- als profielfoto, volgens mij.”

Geïnterviewde: “Ja, en in de follow-up van de Volkskrant schreven ze dat wel, toch? Of niet? Nee, dat was ook The Verge.”

Interviewer: “Ja, nee.”

Geïnterviewde: “Ik weet niet meer wie het zei.”

Interviewer: “The Verge zei van dit een echte foto is van de echte Volkov. En Tweakers, u heeft daar nog geschreven van, ja, dit is een foto van de echte Volkov .”

Geïnterviewde: “Deze hier.”

Interviewer: “Dus ja. Ja, even kijken, ik weet niet. Het staat... Maar dat is in het... twee... volgens mij is dat in het “duo achter imitatie”-artikel [interviewer verwijst hier naar het Tweakersartikel “Duo achter imitatie stafchef Navalny: 'het was geen deepfake maar dubbelganger'”].”

Geïnterviewde: “Oh, dat is dan deze.”

Interviewer: “ Ja, dat staat hem ergens. Ja, hier. Video uit 2018 [Interviewer verwijst hier naar het videointerview tussen *After Empire* en de echte Leonid Volkov].”

Geïnterviewde: “Die fake screenshot van vorige week, een video uit 2018, die is ongetwijfeld ook niet meer. Oh, jawel.”

Interviewer: “Want die is...”

Geïnterviewde: “Oh, die is niet van die gasten zelf.”

Interviewer: “Nee. Nee, nee, nee. Dus...”

Geïnterviewde: “Nou ja, dat is hem, dat zien we meteen. Ja, dat heb ik er ook- dat heb ik hier ook gezet. Die screenshot, die... Ja, daarin had ik overigens wel slag om de arm of een beetje. Door te zeggen, dit beeld hebben we van De Volkskrant. En De Volkskrant zegt: via Facebook. Via Facebook is het, daarom heb ik het ook tussen aanhalingstekens gezet. Dat is hartstikke vaag.”

Interviewer: “Ja, nee, maar nee, even kijken. Ik snap wel, want volgens mij is deze foto... Euh... Want even kijken, die Baltische politici heeft deze gepost op Twitter.”

Geïnterviewde: “Ok.”

Interviewer: “En volgens mij heeft de echte Volkov deze in een Facebook post ook gebruikt. Dus dat is misschien waar ze het vandaan hebben, want Volkov zegt op Facebook van: welkom in het deepfaketijdperk.”

Geïnterviewde: “ Ja, ja, ja.”

Interviewer: “Maar dan in het Lets of in het Russisch.”

Geïnterviewde: “Ja, wat ze daar ook spreken inderdaad.”

Interviewer: “Ja, ok. Maar goed, ok. Ja, even kijken dan. Ja, even kijken dan ben ik er wel zo'n beetje, denk ik. Heeft u verder nog iets toe te voegen? Of iets dat u wil zeggen?”

Geïnterviewde: “ Nou ja, ik denk... Ik denk in de kern... Wat ik net eigenlijk zei, in de kern gaat het gewoon om dingen absoluut zeker weten. En ik denk dat er binnen de journalistiek wel, deels terecht, deels onterecht, een neiging bestaat om je collega's of je concullega's gewoon te vertrouwen in wat ze zeggen. Weet je, dit eerste stuk van de Volkskrant, daar wordt in de headline gezegd. Nee, niet in de headline.”

Interviewer: “In de header.”

Geïnterviewde: “In de lead, zeg maar, wat je zegt. Het gaat om een deepfake-imitatie. Dat wordt als voldoende feit gebracht. De Volkskrant heeft wel een reputatie van een bepaalde betrouwbaarheid. En weet je, op het moment dat je zo'n nieuwsdienst hebt en je moet door, dan is er de neiging om te zeggen... Het is De Volkskrant, ik vertrouw De Volkskrant. En nogmaals, helemaal niet om ze af te branden, maar euhm... ze zijn zeker te vertrouwen, daar sta ik nog steeds achter, maar ze zijn niet onfeilbaar. Ik denk zeker op het moment dat het een beetje buiten de comfortzone van de mainstream media gaat, zoals een deepfake, dat je dan echt als Tweakersredacteur even mag zeggen, maar ho even, weten jullie wel goed wat een deepfake is? En ja, weet je, als... ik denk dat. Je hebt De Volkskrant één, Tweakers twee, misschien dat een derde nieuwsmedium. Een RTL Nieuws, ik heb geen idee hoor wat zij er over hebben geschreven, daar naar kijkt. Die denkt dan: nou dat zijn twee nieuwsmedia die dit zo brengen. Dat maakt het voor hun gevoel nog misschien nog betrouwbaarder.”

Interviewer: “Nee, nee. Ja, dat zie je wel, want in deze casus, de eerste dag laten we maar zeggen is- je ziet dus dan dat eerste Volkskrantartikel en dan zijn er heel veel andere nieuwsmedia die dan... die dan binnen een paar uur. of zelfs nog voordat hun rectificatie-artikel van de Volkskrant gepubliceerd is, die dan van: ja, De Volkskrant zegt dit en het dan... ja het is een beetje overge- ja, overgeschreven wil ik niet zeggen, maar het wordt wel genoemd: Volkskrant meldt en dan... nou ja, ik denk dat dat wel klopt als ik zo kijk naar de berichtgeving, maar goed.”

Geïnterviewde: “Tot op zekere hoogte is nieuws echt overschrijven hoor. Want als een concurrent eerder is met een verhaal en dat verhaal ligt ook in jouw straatje en zij hebben het echt van begin tot eind hebben ze het goed uitgezocht en jij kan daar niks meer aan toevoegen, dan ben je het in je eigen woorden aan het overschrijven. Euhm... en, ja, zo gaan die dingen gewoon. Maar ik had er zelf beter naar moeten kijken. Ik had zelf moeten beseffen van: joh, deze mensen weten helemaal niet wat een deepfake is.