

# De getemde (deep)fakes

**Deepfakes, oftewel nepvideo's gemaakt met kunstmatige intelligentie, dragen bij aan de steeds grotere stroom aan nepnieuws. Hier ligt werk voor factcheckers.**

In de brede stroom aan misleidende internettools dook ruim een jaar geleden de deepfake op. Deepfake is een techniek die gebruikmaakt van kunstmatige intelligentie om nepvideo's te maken. We zagen een bedrieglijk echte Barack Obama die zei dat hij Donald Trump 'a complete dipshit' vindt, en we lachten om acteur Nicolas Cage die opdook in films waarin hij nooit speelde. Maar er doken ook filmpjes op waarin hoofden van filmsterren waren samengevoegd met lichamen van pornoactrices, met het voorspelbare ranzige resultaat tot gevolg.

De Zwitserse Jeanine Reutemann, regisseur, mediaontwerper en onderzoeker van het Haagse Centre for Innovation, zegt dat deepfakes meer zijn dan de zoveelste techniek die bijdraagt aan fakenieuws

en 'alternatieve' data. 'Deepfakes zijn een nieuwe vorm van pesterij zoals we die niet eerder hebben meegemaakt. Ze kunnen personen, hun imago en reputatie flink beschadigen.'

## Techniek steeds beter

Het is nu vaak nog wel mogelijk om deepfakes te onderscheiden van 'echte' video's (zie kader), maar de techniek achter deepfake wordt steeds beter en het is zeker niet uitgesloten dat er een moment aanbreekt waarop echt en nep niet meer van elkaar te onderscheiden zijn. Bovendien wordt het steeds makkelijker om dit soort filmpjes te maken.

Een filmpje waarin Donald Trump de oorlog verklaart aan Noord-Korea, is makkelijk te benoemen als deepfake. De handel en wandel van de Amerikaanse president worden op de voet gevolgd door alle media en de conclusie is snel getrokken wanneer een opzienbarende uitspraak van de Amerikaanse president wordt verspreid vanaf één dubieuze website en niet via CNN of NBC wordt getoond. Maar hoe zit het met een video waarin de sheriff van een stad met rassenrellen racistische opmerkingen maakt? Of, dichter bij huis, wanneer je een filmpje ziet van je buurman die in bed ligt met iemand die niet je buurvrouw is?



**Rob Feenstra**

Projectleider/consultant bij de Universitaire Bibliotheken Leiden; heeft als aandachtsgebieden bibliotheeksystemen en de digitale bibliotheek



## Detectietechnologie

Technologiereuzen als Facebook maken gebruik van software om deepfakes op te sporen. Maar omdat de deepfakesoftware zich blijft ontwikkelen, loopt de detectietechnologie noodgewongen altijd enigszins achter de feiten aan. En wat levert het ons op? Transparantie en vertrouwen zijn volgens Reutemann de sleutelbegrippen als het gaat om het beteugelen

van deepfakes en andere bedreigingen. Maar tegelijkertijd constateert ze dat transparantie bij de grote techbedrijven ver te zoeken is en dat er bij het publiek mede daardoor weinig vertrouwen is in de media.

Het werk van factcheckers is broodnodig om deepfakes te ontmaskeren. En consumenten moeten in de eerste plaats zichzelf expliciet als factcheckers zien. Vrijwel iedereen beschouwt zichzelf als een kritisch wezen, maar we zijn van nature geneigd om iets te geloven als we het met eigen ogen zien, vooral als we het graag willen zien. De Socialistische Partij van Vlaanderen verspreidde een deepfake van Donald Trump die de Belgen aanspreekt over het klimaatakkoord. De kwaliteit van de beelden is belabberd (denk aan de Jaap Aap-filmpjes van André van Duin), maar veel mensen gingen

**'We zijn van nature geneigd om iets te geloven als we het met eigen ogen zien'**



er serieus op in omdat het filmpje hen bevestigde in hun mening.

### Bewustwording

Bewustwording van de consument is dus een belangrijk issue. Op de vraag wat we zelf kunnen doen tegen deepfakes, antwoordt Reutemann: 'Voor deepfakes geldt hetzelfde als voor andere fact-check issues: controleer altijd je bronnen. Dat betekent ook dat je je persoonlijke netwerk op digitale platforms moet checken. Wie zit er achter de nieuwsbronnen op je sociale netwerk? Moeten ze zich verantwoorden naar jou of naar anderen? Hoe onafhankelijk zijn de bronnen waar het materiaal vandaan komt? Probeer beter te begrijpen hoe audiovisueel materiaal zoals video werkt – video is gewoon een taal, een audiovisuele taal. Kijk en luister zorgvuldig naar wat iemand beweert in een video.'

De opmars van de deepfake maakt eens te meer duidelijk dat er een belangrijke dubbele taak is weggelegd voor informatiespecialisten. Natuurlijk moeten we duidelijk maken welke technieken we kunnen gebruiken om internetmateriaal te beoordelen. Maar belangrijker nog is dat we erin slagen om internetgebruikers er bewust van te maken dat ze kritisch moeten zijn. <

**'Vaak is de verhouding tussen lichaam en hoofd niet helemaal goed'**

## Hoe ontmasker je een deepfake?

- > **Rare gezichten:** als alles er normaal uitziet, maar het gezicht wasachtig is, gaat het meestal om een deepfake.
- > **Flikkeren:** bij slechte deepfakes flikkert het gezicht af en toe en is het originele gezicht soms even te zien.
- > **Lichaamsverhoudingen:** vaak is de verhouding tussen lichaam en hoofd niet helemaal goed. Dat kan zijn omdat het hoofd gewoon te groot of te klein is, maar bijvoorbeeld ook doordat het hoofd van een mager persoon is geplakt op dat van een dikkerd.
- > **Korte fragmenten:** juist omdat er veel mis kan gaan, duren deepfakes vaak maar een paar seconden.
- > **Geen geluid of slechte synchronisatie:** vaak wordt de stem nog gewoon ingesproken.
- > **Dubieuze bron:** als een filmpje afkomstig is van de NOS-website is het waarschijnlijker dat het om een 'echt' filmpje gaat dan wanneer het fragment afkomstig is van [trickypictures.com](http://trickypictures.com)
- > **Oogbewegingen:** vaak knipperen de hoofden in deepfakes niet of nauwelijks met hun ogen.
- > **Coördinatie stem/lichaam:** wanneer er bijvoorbeeld armbewegingen worden gemaakt die niet in overeenstemming zijn met wat er op dat moment gezegd wordt, zou er sprake kunnen zijn van deepfakes.
- > **Vertraagd afspelen:** op die manier is het eenvoudiger om niet kloppende veranderingen of bewegingen te signaleren.
- > **Kijk goed naar de mond:** de deepfake-software heeft nog veel moeite om lippen, tanden en tong correct af te beelden. Let op onscherpe vlekjes.