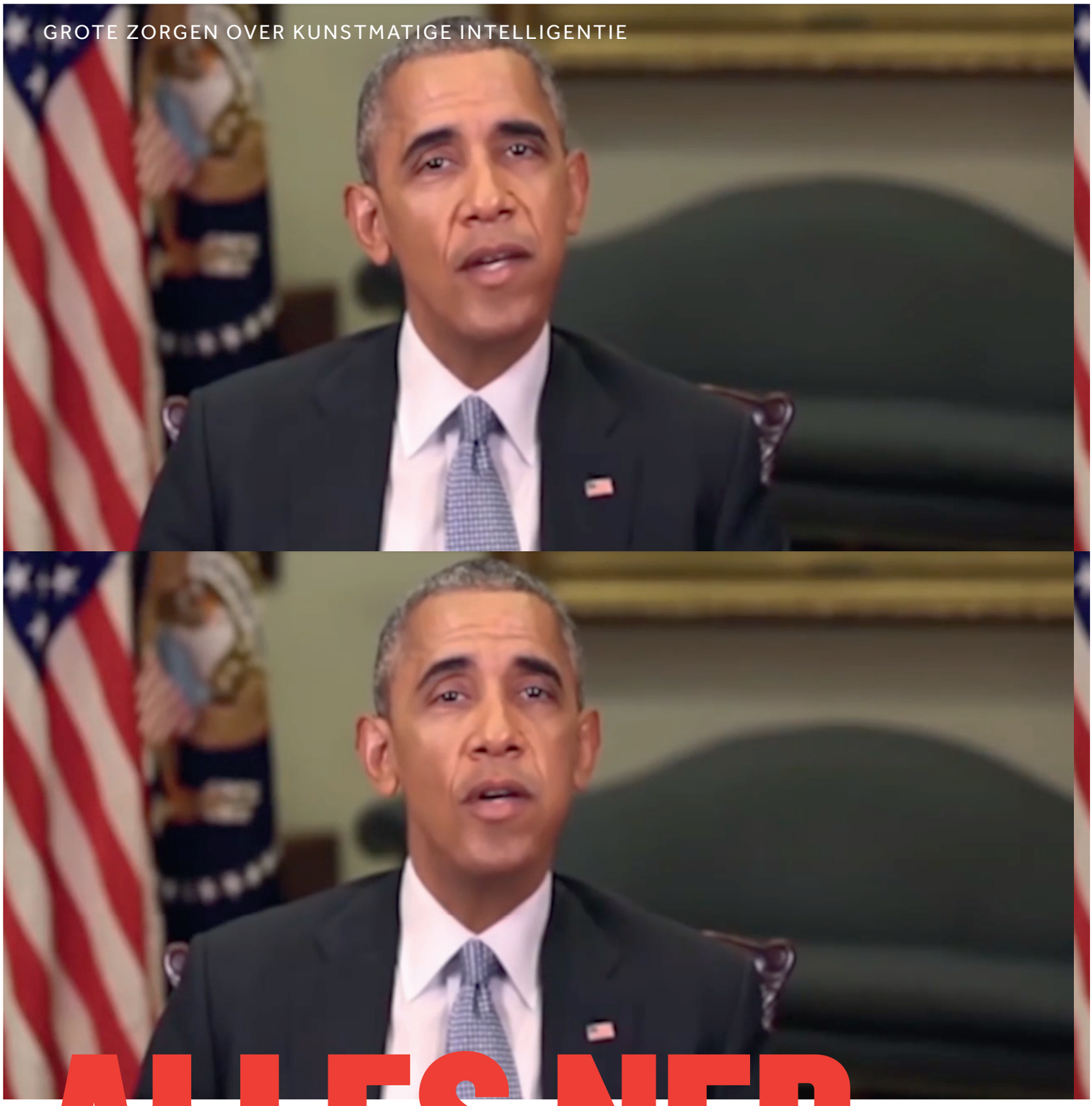
GROTE ZORGEN OVER KUNSTMATIGE INTELLIGENTIE



ALLES NEP OP HET WEB

Deepfake videobeeld van de vroegere Amerikaanse president Barack Obama.

Na jaren onderhandelen werd het Europees Parlement het eens om ontwikkeling en gebruik van kunstmatige intelligentie te reguleren. Ons kabinet liet weten niets te willen doen. Deskundigen noemen dat 'diep teleurstellend'. Want er is een toename van het aantal deepporn-incidenten en verwacht wordt dat dit ook gaat gelden voor oplichting, fraude en misleiding. De politie slaat alarm. 'Onze democratie en rechtsstaat staan op het spel.'

TEKST RICK BLOM **FOTOGRAFIE** APPLE, YOUTUBE, GETTY IMAGES, POLITIE.NL

Stel je voor: de zon schijnt, je zit buiten op een terras. Voor je op tafel ligt een schaakbord. Tegenover je zit je vader. Jullie zijn geanimeerd in gesprek en gezellig een potje aan het spelen. Alleen: je vader is allang overleden. En ook: ieder ander op dat terras ziet jouw vader niet, alleen jij. Utopisch? Niet volgens Manon den Dunnen, Strategisch Specialist Digitaal bij de politie en binnen die organisatie hét aanspreekpunt en dé kennishouder op het gebied van synthetische media: 'Er gaat veel geld naar eye-tracking, analyse en targeting om te kunnen begrijpen waar iemand naar kijkt en wat diens geestelijke en lichamelijke gesteldheid is, zodat je met gepersonaliseerde informatie getarget kan worden. Vanuit projectie in de omgeving krijg jij straks zo je eigen individuele informatielagen over die omgeving heen. Zoals met de komst van de Apple Vision Pro-brillen die vanaf begin 2024 in de Verenigde Staten verkrijgbaar zullen zijn, maar ook de slimme autoruiten waar nu aan gewerkt wordt. Daarmee zie je zowel de gewone wereld als projecties er bovenop: mixed reality dus. Zo krijgen we steeds meer personalisatie in de fysieke wereld om ons heen waardoor we steeds minder een gezamenlijk wereldbeeld hebben.'
Is dit iets dat de kwaliteit van ons leven zal verbeteren? Allesbehalve. Den Dunnen: 'Door gebruik te maken van allerlei online-tools wordt de wereld steeds meer voor ons geconstrueerd. En ook, door wie de wetten van de technologie kent, steeds meer gemanipuleerd en misvormd. Wanneer weet je wat nog echt is en wat jouw gedrag bepaalt? Dat is nu al zo. Wie heeft nog door dat als je Google Maps gebruikt je eigenlijk door een digitale bril naar de werkelijkheid kijkt?'
Het gaat verder: de algemene consensus in de wereld van het internet is dat binnen nu en twee jaar meer dan 90 procent van de informatie op het web gemanipuleerd is. Het is niet echt. Het is nep. We gaan ons wereldbeeld dus voor een groot deel laten bepalen door informatie die op wat voor manier dan ook door *artificial intelligence* (AI) gemanipuleerd of gegenereerd is. We gaan handelen naar een gecreëerde illusie. Hoe kun je je dan nog tot het echte leven en anderen – je buren, je vrienden of collega's, de overheid – verhouden? Den Dunnen: 'We laten ons het meest beïnvloeden door wat we zelf zien en horen. Door de steeds verdergaande ontwikkelingen gaan we naar een samenleving waarin je niet meer kunt vertrouwen op je eigen waarneming. Ik maak me daar grote zorgen over.'

Incidenten

De ontwikkeling van AI wordt ingezet om de kwaliteit van het leven te verbeteren. Developers schermen met nobele intenties. Wie weet is dat oprecht. Kunstmatige intelligentie zou kunnen helpen om kanker te genezen, het verkeer veiliger te maken, of de klimaatcrisis op te lossen. Maar grote zorgen zijn er ook. Oók bij

DE ALGEMENE CONSENSUS IS DAT BINNEN NU EN TWEE JAAR MEER DAN 90 PROCENT VAN DE INFORMATIE OP HET INTERNET GEMANIPULEERD IS

De Apple Vison Pro-bril komt begin 2024 op de markt. Links: Manon den Dunnen.

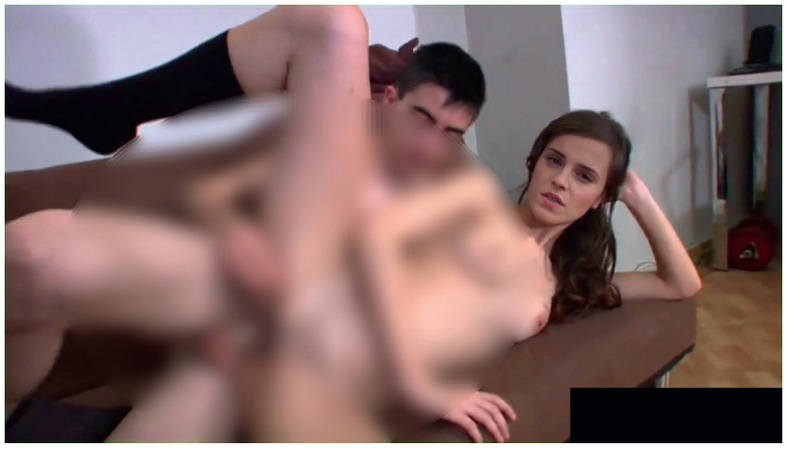
de Nederlandse politie. We worden in toenemende mate geconfronteerd met het kwaad waartoe AI in staat is: cyberaanvallen, afpersing, spoofing, phishing, poisoning, malware aanvallen, fraude, het verschaffen van een vals alibi (je kunt al een bankrekening op naam van iemand anders openen, omdat je de online-authenticatiesystemen voor de gek kunt houden), inbraak in systemen die worden beschermd door gezichtsherkenning, misleiding, haat zaaien, pornografie, virtuele aanranding, het verspreiden van misinformatie, het beïnvloeden van democratische verkiezingen en de koersen op de aandelenbeurzen. Dat laatste gebeurde onlangs nog in de Verenigde Staten toen op Twitter een nepfoto van een explosie in de buurt van het Pentagon werd geplaatst. De brandweer van Virginia meldde dat er helemaal geen explosie had plaatsgevonden bij het gebouw van het Amerikaanse ministerie van Defensie. Toch daalde de Amerikaanse S&P 500-index met 0,26 procent. Als gevolg van het verspreiden van één enkele foto. De beurs veerde pas weer op toen onomstotelijk bleek dat het beeld nep was.

Volgens de onafhankelijke Britse organisatie AIAAIC (AI, Algorithmic, and Automation Incidents and Controversies) dat incidenten in verband met ethisch misbruik van AI bijhoudt, is het aantal sinds 2012 liefst 26 keer zo groot geworden. De toename is exponentieel en wordt wel vergeleken met de curve van een hockeystick. Inmiddels zitten we in de lijn recht omhoog. Het is een beeld dat Manon den Dunnen herkent. ‘Het in kaart brengen van alles wat gaande is, kunnen wij allang niet meer met slechts een afdelinkje volgen. De volledige organisatie van de politie moet zich er bewust van zijn.’

Dat is niet voldoende. Volgens Den Dunnen komt er zoveel op de politie af, dat het onmogelijk is om alles te verwerken. ‘We kunnen nu al niet alle zaken aan die binnenkomen. Door deepfake-technologie worden delicten als sextorsion steeds makkelijker te plegen. Door Chat GPT of een soortgelijk programma kan iedereen informatie maken. Je afvragen wat daarvan nep is of echt, is niet meer relevant, want straks is álles gemanipuleerd. Voor de politie ›

maakt het onderzoeken complexer, want waar kunnen we nog op vertrouwen? We moeten meer investeren in manieren waarmee onze mensen zo snel mogelijk chocola kunnen maken van de enorme brei aan informatie die op hen afkomt. Welke vragen moeten ze stellen? Welke hulplijnen kunnen ze gebruiken? Hoe kunnen ze dat op zo'n manier doen dat ze hun beslissingen achteraf transparant kunnen uitleggen? Wat weten we wel en wat niet? De wereld is snel digitaal geworden en dat betekent dat we als politie ook midden in de digitale transformatie van ons korps zitten. Iedereen – van wijkagent en rechercheur tot en met de korpschef – zal moeten begrijpen wat de consequenties van de digitale wereld voor ons werk zijn en daar op elk niveau naar moeten kunnen handelen. Dat is geen gemakkelijke opgave.'

Twitch-ster QTCinderella in tranen nadat ze voorkomt in een deepfake pornovideo. Een lot dat ook actrice Emma Watson (onder) en acteur Tom Holland (rechtsonder) overkwam.



Deepporn

De toename aan criminaliteit zal vooral komen door de eenvoud van de tools die er zijn, weet Den Dunnen. Iedereen kan die inzetten en met één druk op de knop veel mensen bereiken. Op dit moment is het vooral deepporn dat met AI-technieken de meeste slachtoffers maakt: gezichten van vooral vrouwen, maar ook steeds meer mannen, die in pornofilms worden gegenereerd met als doel iemand af te persen of in een kwaad daglicht te zetten, om een reputatie te schaden en zo een heel leven kapot te maken. Den Dunnen: 'Het kan iedereen overkomen die zijn gezicht online heeft staan en wie heeft dat niet?'

Hoe vaak deepporn of andere vormen van deepfakes precies voorkomen, weet de politie niet. Bij criminele activiteiten als afpersing registreert ze niet op welke manier dat gebeurt. Den Dunnen: 'Ik weet van gesprekken met bijvoorbeeld het meldpunt kinderporno wel dat ze meer meldingen zien waarbij gebruik is gemaakt van deepfakes.'

Handhaving ondertussen is niet of nauwelijks mogelijk. Den Dunnen: 'We hebben er de mensen en middelen niet voor. We weten ook dat de Autoriteit Persoonsgegevens ver achterloopt met de behandeling van casussen. Als individu

Straks volstaan opnames van een bodycam niet meer als afdoende bewijs. Links: Bart van der Sloot.

kan ik daar niet terecht met een klacht over Facebook, TikTok of een willekeurig Nederlands bedrijf. Er is gewoon te weinig mankracht. Daarbij: de deepfake detectietools die er zijn, werken onvoldoende. Nu er meer incidenten bijkomen, is het onmogelijk dit met opsporing en vervolging op te lossen. Er moet echt aan de voorkant op gehandhaafd worden. Dat hebben we in Nederland nu niet goed geregeld.'

Gevaren

Bart van der Sloot is associate professor aan het Tilburg Institute for Law, Technology, and Society en specialist op het gebied van technologie en recht, privacy en Big Data. Dat de politie zegt dat je met opsporing de strijd tegen malafide deepfakes niet kunt winnen, verbaast hem niets. Op verzoek van het Wetenschappelijk Onderzoeks- en Documentatiecentrum van het ministerie van Justitie en Veiligheid schreef Van der Sloot twee jaar geleden samen met collega's van de Universiteit Tilburg een rapport

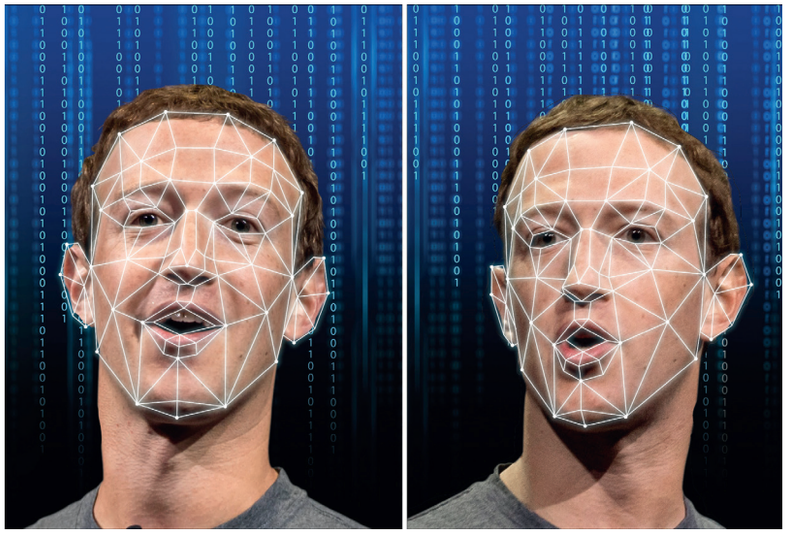
'ALS POLITIE GENEREREN WE OOK ZELF INFORMATIE WAARVAN DE AUTHENTICITEIT STEEDS VAKER BETWIST ZAL WORDEN IN DE RECHTSZAAL. DAT ONDERMIJNT DE BEWIJSLAST'

over de regulering van deepfakes. Het ministerie vroeg het onderzoek aan omdat er in de Tweede Kamer destijds al grote zorgen waren over de snelle verspreiding van deepfake-technologie en de maatschappelijke gevolgen daarvan. In hun rapport deden de onderzoekers het kabinet een aantal concrete suggesties om de ontwikkelingen in goede banen te leiden. Dat ging van het verbieden van AI-technologie voor de consumentenmarkt tot een brede publiekscampagne over de gevaren van deepfakes. Pas enkele weken geleden, twee jaar na het verschijnen van het rapport, kwam het inmiddels gevallen kabinet in een Kamerbrief met een reactie. De brief stelt dat het recht goed is toegerust om deepfake-technologieën te adresseren en erkent dat handhaving pas aan de orde komt als het kwaad al is geschied. Maar het was voor het kabinet geen reden om regels en wetten op te stellen om op voorhand de risico's te minimaliseren.

Integendeel, de minister schreef: 'Strafbare content kan uiteraard ook bestreden worden door middel van opsporing en vervolging.' Maar dat is geen oplossing, zegt Den Dunnen van de politie: 'Want dan zijn mensen al slachtoffer geworden.'

Bart van der Sloot noemt de kabinetsreactie zorgelijk: ‘Het is allereerst onbegrijpelijk dat het zo lang moest duren voor het met een reactie op ons rapport kwam. Het is dat jullie mij mailden, anders had ik niet eens geweten dat die Kamerbrief er lag. Ze hadden die net zo goed niet kunnen schrijven, want er staat eigenlijk in dat ze niks doen. De uitkomsten en aanbevelingen van het rapport waren twee jaar geleden al urgent, en dat is alleen maar meer geworden. Ik maak me grote zorgen over de cultuur binnen het ministerie van Justitie. Het vraagt rapporten aan, wij wetenschappers leveren die braaf in, en vervolgens verdwijnt het allemaal in een diepe la. Het ligt niet aan de Kamerleden, die zitten erbovenop, het ligt eerder aan een enorme onwil bij de ambtenaren daar om hier iets aan te doen en een minister die kennelijk niet in staat was die mensen in beweging te krijgen.’

Een initiatief dat een begin lijkt te maken van een oplossing, komt van de politie. In januari dit jaar organiseerde ze samen met een aantal ministeries, het Rathenau Instituut, het Electronic Commerce Platform Nederland en het Nederland Forensisch Instituut een rijksbrede conferentie over gemanipuleerde media om kennis te delen over de stand van zaken, de uitdagingen en een mogelijke aanpak. Zo was er een gesprek over de impact van misinformatie op de democratie en over de herkomst van informatie op het internet. Daarbij gaf de politie praktische adviezen om die herkomst in kaart te brengen. ‘Internationaal zijn er al initiatieven om authenticiteitskenmerken toe te voegen, bijvoorbeeld op het moment dat je met een camera een beeld maakt,’ zegt Den Dunnen. ‘Hier zijn onder andere de BBC en Microsoft bij betrokken. Google gaat dergelijke technologie ook toepassen in haar zoekmachine. Je weet dan nog steeds niet of de inhoud waar is, maar je weet in elk geval waar het vandaan komt. Als ik de NOS vertrouw, weet ik dat ik een artikel lees dat origineel door de NOS gemaakt is, en of het daarna misschien nog door een ander is gemanipuleerd. Het is belangrijk dat onderscheid te kennen. Als politie genereren we ook zelf informatie waarvan de authenticiteit steeds vaker betwist zal worden in de rechtszaal. Dat ondermijnt de bewijslast. Aan de bodycams die agenten dragen, zijn daarom nu al een aantal authenticiteitskenmerken toegevoegd, zodat we precies registreren wie wanneer iets op camera vastlegt. Dat zijn ontwikkelingen die kunnen helpen.’



AI-beeld van Facebook-oprichter Mark Zuckerberg.

Opnieuw vertraging

Tijden de conferentie was er veel draagvlak voor maatregelen bij betrokken partijen, zegt Den Dunnen. ‘Maar er is nog geen tastbaar resultaat. Kennelijk is het lastig om vervolgens tot concrete actie te komen.’

Nu het kabinet demissionair is en dus alleen nog lopende zaken behandelt, kan het bovendien nog wel even duren voor er ook maar iets op dit dossier gebeurt. Waarschijnlijk zal dat niet eerder zijn dan na nieuwe verkiezingen en een kabinetsformatie. Van der Sloot: ‘De val van het kabinet zal opnieuw vertraging opleveren tenzij de kamer de demissionaire minister aanspoort om hier nu eindelijk snel iets aan te doen. Maar dan nog is het mijn verwachting dat er niet al te veel zal gebeuren. De kennelij-

‘HET ZOU HELPEN ALS WE ALS EUROPA EEN VIRTUELE GRENS INRICHTEN, WAARBIJ SOCIALE PLATFORMEN ZICH AAN ONZE REGELS MOETEN HOUDEN’

ke onwil bij de ambtenaren op het ministerie om hier mee aan de slag te gaan, zal de komende tijd denk ik niet veranderen.'
De oorverdovende stilte van de afgelopen jaren betekent volgens Van der Sloot dat alle problemen die hij en al die organisaties waaronder de politie signaleren, niet worden aangepakt en gewoon kunnen blijven bestaan: 'Voor de bedrijven die AI-tools ontwikkelen is dat heel fijn. Die gaan lekker hun gang. Voor de burger is het nogal cru, bijvoorbeeld als je als vrouw slachtoffer wordt van deepporn. Of wanneer democratische verkiezingen worden beïnvloed door deepfakes. Het is net als met de aardbevingsproblematiek in Groningen en de toeslagaffaire bij de belastingdienst: er zijn jarenlang signalen dat het fout gaat, maar de politiek doet niks.'

Zwak pakket

In het Europees Parlement was een paar weken geleden eensgezindheid over hoe de regels voor kunstmatige intelligentie eruit zouden kunnen zien. Over die regelgeving wordt nog onderhandeld, maar vooralsnog is het volgens Van der Sloot allemaal onvoldoende. 'Wat er nu ligt, is een zwak pakket. De deepfake-bepaling daarin zegt bijvoorbeeld dat je transparant moet zijn of iets met deepfake-technologie is gemaakt. Hoe zie je dat voor je? Gaan mensen zich daaraan houden? Ik zie niet hoe dat effect gaat hebben. Hier ga je de strijd niet mee winnen. Het is *too little, too late*.'
Het is de vraag of de overheid in Nederland of Europa überhaupt in staat is om met regulering de ontwikkeling van AI-technologie te sturen. Den Dunnen: 'Er zijn er allerlei platformen, zoals Telegram waar ook ongelooflijk veel rotzooi op verschijnt, die we niet eens kunnen aanpakken, omdat ze zich buiten ons rechtsgebied bevinden. Als je ziet dat er nog steeds pro-anorexia-platforms zijn op TikTok en hoe jonge meiden daar nog altijd mee getarget worden, dat is vreselijk. Nu laten we dat allemaal toe. Misschien zou het helpen als we als Europa een soort virtuele grens inrichten, waarbij sociale platformen zich aan onze regels moeten houden. Doen ze dat niet, dan moeten we drempels inbouwen waardoor gebruikers in elk geval weten dat zo'n platform buiten onze veilige omgeving opereert. Zo kunnen we ervoor zorgen dat minder mensen van dat soort kanalen gebruikmaken.'
Niet alleen deepporn is een groot en nauwelijks te bestrijden probleem. Dat geldt net zo goed voor oplichting en afpersing. Den Dunnen: 'We hadden al vriend in nood-fraude. Iemand deed zich voor als familielid en stuurde een app of sms met een verzoek om financiële hulp. Nu kun je op je telefoon de naam zien van het familielid dat jou belt en vervolgens een gekloonde stem van die persoon horen die je vertelt jou nodig te hebben. Natuurlijk denk je dan een bekende aan de lijn te hebben die om hulp vraagt. Het gebeurt nu ook met zogenoemde CEO-fraude, waarbij de stem van een directeur van een bedrijf wordt gekloond om geld los te krijgen. Op die manier werd twee jaar terug al in de Verenigde Arabische Emiraten een bedrag van 35 miljoen dollar buitgemaakt. Criminelen kloonden er de stem van een bankdirecteur die een medewerker opdroeg dat bedrag over te maken, zogenaamd in verband met de overname van een bedrijf. Precies die techniek om stemmen te kunnen klonen, is nu ook in het Nederlands beschikbaar.'
Door de verbeteringen van dit soort technieken is er zelfs intern bij de politie geen onvoorwaardelijk vertrouwen meer. Den Dunnen: 'Het advies is: als je gebeld wordt door een collega, weet dan dat het die collega niet hoeft te zijn. Geef niet zomaar vertrouwelijke informatie. Bel terug om te checken of je die collega daadwerkelijk aan de lijn had. Wees zorgvuldig. Dat is de cultuuromslag die we moeten hebben.'
De effectiviteit van AI-tools is veel groter dan manieren voor oplichting en afpersing in de analoge wereld. Dat heeft volgens Den Dunnen niet alleen impact in de persoonlijke sfeer, maar ook verregaande maatschappelijke consequenties: 'Met Chat GPT kan ik met één ›

Een bekend trucje om via een appje geld af te troggelen.

druk op de knop eenzelfde boodschap genereren in wat voor taal dan ook. Ik kan dat vervolgens verpakken op een manier waarop die door allerlei soorten platformen gepubliceerd kan worden, aansluitend op hun specifieke doelgroep. Dat betekent dat de mogelijkheden tot beïnvloeding gigantisch toenemen. Zeker als je dat combineert met de werking van sociale media: al die nieuwsfeeds en tijdlijnen die prioriteren op basis van engagement. Als je een boodschap op zoveel mogelijk plekken verspreid, zal die door veel mensen worden gezien. Als die er al niet op reageren, kun je ook Chat GPT comments laten geven of laten retweeten. De engagement wordt zo enorm en zo'n verhaal komt dus bovenaan ieders tijdlijn. Dat betekent dat de boodschap zal landen, want als we iets herhaald zien, dan werken onze hersenen zo dat er een zaadje wordt geplant. Ook al weten we dat de inhoud niet klopt. Als iedereen zegt dat mijn blauwe koffiebeker rood is, zal ik in eerste instantie tegensputteren, maar het uiteindelijk beamen. Zo werkt onze psyche. Het betekent dat het steeds moeilijker wordt voor mensen om te bepalen wat betrouwbaar is en wat niet. Dit soort beïnvloeding is een wapen. Een wapen bovendien dat zomaar door iedereen gebruikt kan worden.'

WAT KUN JE ZELF DOEN?

Als je denkt slachtoffer te zijn van oplichting of afpersing, dan adviseert de politie:

- Ontwikkel gezond wantrouwen, we weten nu dat alles gemanipuleerd kan worden, dus het is niet gek als je in gesprek met iemand wat dingen wil controleren. Pas op met wat je zelf deelt.
- Wees je ervan bewust dat stem, gezicht en afzender (degene die jou mailt/belt) makkelijk gemanipuleerd kunnen worden.
- Herken rode vlaggen; emotie, tijdsdruk, iets verschrikkelijks of iets dat te mooi is om waar te zijn. Of als er een onomkeerbare handeling wordt gevraagd (overmaken geld, delen persoonlijke of vertrouwelijke informatie).
- Doe als hier sprake van is een extra check.
- Bel zelf terug op (of text naar) het nummer dat je zelf kent en intypt of uit je contactpersonen selecteert.
- Heb je geen direct nummer, probeer het dan via de centrale van de organisatie of een vriend/collega en zoek zelf dat nummer op voor je belt.
- Lukt dat niet, beschouw dit als extra rode vlag!
- Vraag iets dat alleen de ander kan weten, dat is niet de naam van je hond of verjaardag – geen weetjes – maar vraag naar iets dat je samen beleefd hebt. Van alle andere informatie moet je ervan uitgaan dat die wel ergens op het internet te vinden is (of dat de oplichter een kennis is die het kan weten).
- Er is niks mis met ophangen, oplichters maken gebruik van jouw behoefte beleefd te blijven.

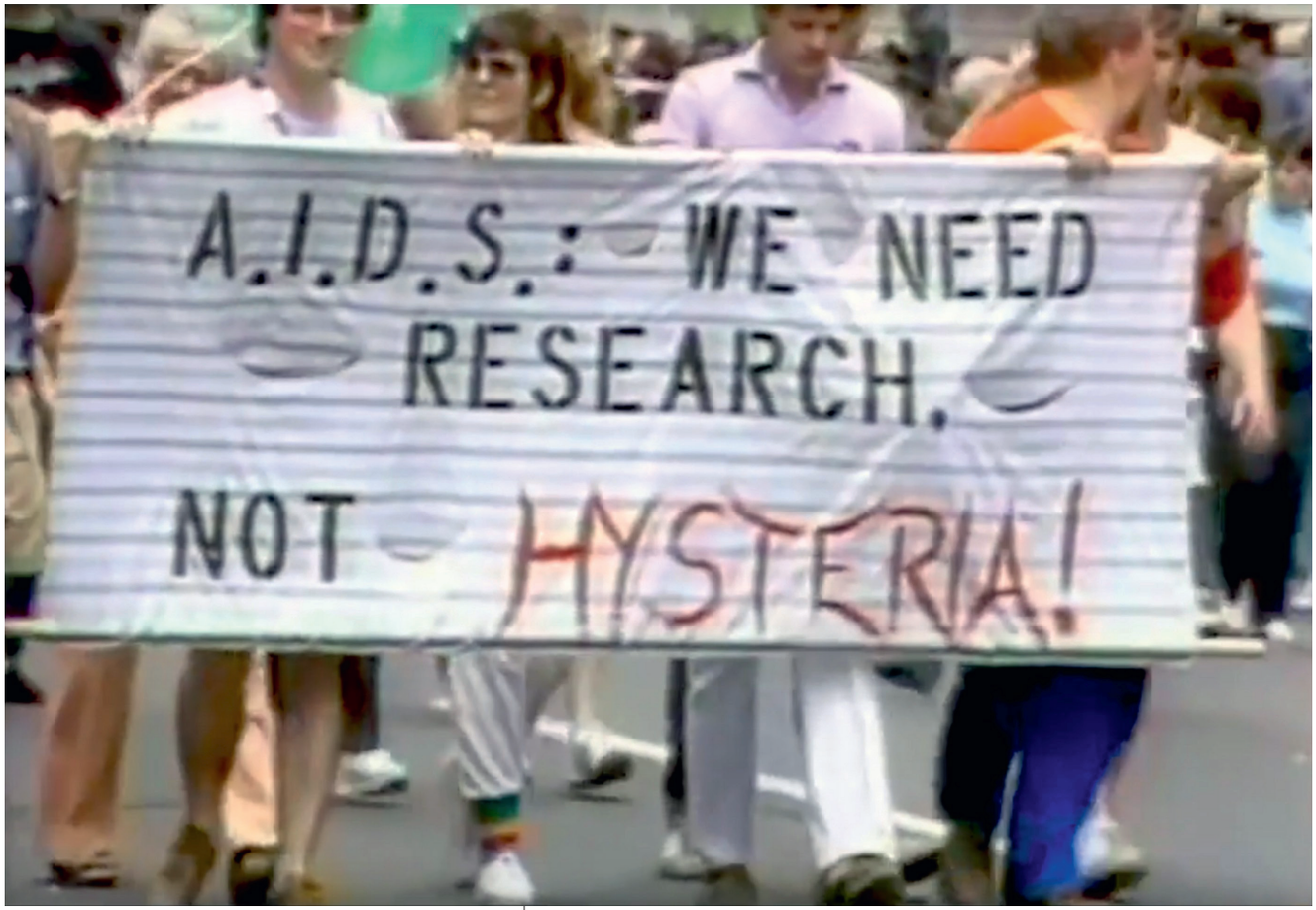
Is het tij nog te keren? Den Dunnen: 'De ontwikkelingen gaan door. De krachten die daar achter zitten zijn immens. Het gaat over macht, beïnvloeding en manipulatie. Ook mensen waarbij je het nu niet verwacht, zullen op dwaalsporen worden gebracht. Dat merkte ik tijdens de coronapandemie ook. Kijk je naar de cijfers over desinformatie in Nederland dan zie je dat de meerderheid van de mensen gelukkig nog steeds meer dan één informatiebron gebruikt. Het is nog altijd een relatief kleine groep die escaleert. Dat neemt niet weg dat het vertrouwen in de overheid afneemt. Door alle misinformatie die via AI gegenereerd gaat worden, zullen grotere groepen mensen kunnen radicaliseren.'

Wantrouwen

Is het een zichzelf versterkend mechanisme? Als het wantrouwen onderling groter wordt, zullen mensen dan nog makkelijker te manipuleren zijn? Den Dunnen: 'Je wordt gevoed in je wantrouwen. Een voorbeeld: toen een agent van de Russische KGB overstapte naar het Westen, deed hij een boekje open over hun desinformatiebeleid. Dat ging niet over het overtuigen van anderen met een nepverhaal. Hun doel was mensen zo onzeker te maken over de informatie die ze krijgen, dat ze niet meer de beslissingen konden nemen die in hun eigen belang waren. Er zijn dus overheden, maar wellicht ook anderen, die zoveel onzekerheid en paranoia willen creëren, dat je niet meer in staat bent de juiste dingen te doen. En die krijgen daartoe nog meer tools in handen. Dan gaat uiteindelijk de samenleving ten onder als je die hier niet tegen beschermt.'

Van der Sloot: 'Bedrijven als Google, Meta en Bytedance – eigenaar van TikTok – gaan een enorme zeggenschap krijgen over wat wij denken wat waar is en hoe wij de werkelijkheid ervaren. Zij hebben invloed op welke instituties we vertrouwen en hoe we met elkaar omgaan. Daar denken ze op een hele strategische manier over na. Als zij de waarheid in handen hebben en die kunnen vormgeven, dan kunnen ze het hele leven commercialiseren.'

Het doet denken aan een rapport van de veiligheidsdienst AIVD van afgelopen mei waarin de dienst schrijft dat onze democratische rechtsorde wordt bedreigd door een beweging die uitdraagt dat een 'kwaadaardige elite' bij de overheid, de rechtspraak, kranten en tv-zenders, wetenschap, grote bedrijven en de politie uit is



AI in de jaren tachtig werd er fake nieuws verspreid over de herkomst van het aidsvirus.

op totale wereldcontrole. Dit wereldbeeld is feitelijk onjuist, zegt de AIVD. Maar, stelt de dienst: als een groeiende groep extremisten dat toch zegt, kunnen mensen onterecht hun vertrouwen verliezen in instituties waar onze democratische rechtsorde op gebaseerd is. Van der Sloot: 'De AI-technologie zit verknoopt met dit soort sociale ontwikkelingen. Met polarisatie en post-truth. Er zijn al nu al traditionele media die een aparte variant van de waarheid aanhangen. In de Verenigde Staten is onderzocht dat meer dan 40 procent van de Amerikanen meent dat Biden niet legitiem de president

'DE ONTWIKKELINGEN GAAN DOOR. DE KRACHTEN DIE DAAR ACHTER ZITTEN ZIJN IMMENS. HET GAAT OVER MACHT, BEÏNVLOEDING EN MANIPULATIE'

is. Gooi daar hele realistische deepfakes bij en je krijgt dat nog veel meer mensen in de verkeerde dingen gaan geloven. Of andersom, dat je denkt dat een verhaal niet klopt als ware informatie wordt getoond. We gaan naar een maatschappij van *trust nothing, question everything*. Onze waarheid, de basis van al onze sociale instituties en persoonlijke communicatie, staat onder druk. Conclusie: onze democratie en onze rechtsstaat staan op het spel.'

Den Dunnen: 'De KGB-agent die ik aanhaalde, gaf nog aan: wij zijn van de lange termijn. We hebben daarvoor één generatie gereserveerd. Aanvankelijk liepen hun desinformatiecampagnes via bijvoorbeeld mediabedrijven in Afrika en India. Op die manier duurde het zes jaar voordat het verhaal viraal ging dat aids uit een Amerikaans laboratorium kwam. De theorie van een paar jaar geleden dat corona het gevolg was van 5G-masten, ging veel sneller viraal. De ontwikkelingen gaan straks nog harder. Daarom is het belangrijk gezamenlijk te kijken naar alternatieve strategieën. Laatst luisterde ik de podcast *The AI-Dilemma*. Daarin werd AI vergeleken met kernwapens en dus ook met de noodzaak tot verdergaande regulering en internationale afspraken. Dat is iets dat nu moet gebeuren. Met opsporing alleen gaan we de wapenwedloop met malafide deepfakes verliezen.' ✖