Artikel

Strafbaarstelling van seksuele deepfakes

mr. dr. M.L.R. (Marthe) Goudsmit Samaritter, mr. R.F. (Ruben) Aksay en prof. mr. dr. J.J. (Jan-Jaap) Oerlemans*

1. Inleiding

Apps zoals 'FaceSwap' en 'DeepFakeLive' zijn tegenwoordig mateloos populair. Met een paar klikken is het mogelijk een nepafbeelding of -video te maken. Het leidt vaak tot hilariteit als mensen zich bijvoorbeeld kunnen voordoen als een popster of wanneer het gezicht van een baby op het hoofd van een vader of moeder wordt geplakt. Het is echter ook kinderlijk eenvoudig deze technologie voor schadelijke doeleinden in te zetten.

In 2022 werd bijvoorbeeld Welmoed Sijtsma, een Nederlandse journaliste, slachtoffer van seksuele 'deepfakes'. Zonder Sijtsma's toestemming of medeweten was er een pornografische video gemaakt, waarbij haar gezicht werd geplakt op het gezicht van een porno-actrice in een bestaande pornovideo: een seksuele deepfake, ook wel 'pornografische deepfake' of 'deepnude' genoemd. De documentaire die Sijtsma daarover maakte, vestigde de aandacht op het fenomeen van seksuele deepfakes van en met volwassenen in Nederland.[1]

Welmoed Sijtsma deed aangifte en een verdachte werd opgepakt. In juli 2023 werd bekend dat het Openbaar Ministerie (OM) de man gaat vervolgen. Het is de eerste keer dat het OM een strafzaak begint onder artikel 139h Sr naar aanleiding van dit soort nepporno.[2] Dit roept natuurlijk de vraag op voor welke delicten het OM vervolging kan instellen, maar ook of de huidige delictsomschrijvingen voldoende recht doen aan de schadelijke gedragingen die plaatsvinden bij het in bezit hebben, maken en verspreiden van seksuele deepfakes. In dit artikel beperken wij ons tot beeldmateriaal van volwassenen en het strafrecht.[3] Met de term 'beeldmateriaal' verwijzen wij in deze bijdrage naar zowel bewegende als niet bewegende beelden, dat wil zeggen: zowel foto's als video's. Slachtoffers van seksuele deepfakes kunnen uiteraard een beroep doen op het civiele recht en verwijdering van het materiaal eisen met een beroep op de Algemene verordening gegevensbescherming, de Auteurswet (in het bijzonder portretrechten) of een procedure aanspannen vanwege een onrechtmatige daad.[4] In dit artikel beantwoorden we de vraag in hoeverre de strafbaarstelling voor seksuele deepfakes gerechtvaardigd is, en of deze eventueel dient te worden verbeterd. Daarbij laten we de vraag hoe de wetgeving van seksuele deepfakes zou moeten worden gehandhaafd, buiten beschouwing.[5]

* Marthe Goudsmit Samaritter is Postdoc onderzoeker Max Planck Institute. Ruben Aksay is docent en promovendus straf(proces)recht aan de Radboud Universiteit. Jan-Jaap Oerlemans is bijzonder hoogleraar Inlichtingen en Recht bij de Universiteit Utrecht.

1 Welmoed Sijtsma, 'Welmoed en de Sexfakes', *NPO 3*, 2022.

2 L. van de Ven, 'OM gaat maker deepfakes van Welmoed Sijtsma vervolgen', *NRC*, 28 juli 2023. Volledigheidshalve: het gaat hier dus om nepporno die niet alreeds strafbaar is onder art. 240b Sr. Wij hebben het in deze bijdrage niet over de strafbaarheid en vervolging van (nep)kinderpornografie.

3 De problematiek omtrent 'audiodeepfakes' komt in dit artikel dus niet aan bod. In januari 2023 kwam bijvoorbeeld een audiofake in het nieuws waarin filmster Emma Watson schijnbaar een pagina uit 'Mein Kampf' voorleest. Zie Joseph Cox, 'AI-Generated Voice Firm Clamps Down After 4chan Makes Celebrity Voices for Abuse', *Vice* 30 januari 2023.

4 Zie verder bijvoorbeeld B. van der Sloot, Y. Wagensveld en B.-J. Koops, 'Deepfakes. De juridische uitdagingen van een synthetische samenleving', WODC-rapport 3137, Tilburg, Den Haag: TILT/WODC 2021; S. Kulk, 'Deepnudes aanpakken met IE-recht?', *Intellectuele Eigendom en Reclamerecht* 2021, nr. 1, p. 3-5; A.J. Trouborst, C.J.S. Vrendenbarg en D.J.G. Visser, 'Nep echt onder het naburig recht', *NJB* 2022/93, p. 101-110. Zie ook: D. Yeşilgöz-Zegerius, 'Beleidsreactie op de WODC-onderzoeken naar de regulering van deepfakes en immersieve technologieën', *Kamerstukken II* 2022/23, 26643, nr. 1041.

5 Van der Sloot e.a. 2021; B.W. Schermer en J. van Ham, *Regulering van immersieve technologieën: Considerati*, Den Haag: WODC 2021; M. Galič e.a., *Spioneren met hobbydrones en andere technologieën door burgers: een verkenning van de privacyrisico's en reguleringsmogelijkheden*, Den Haag: WODC 2020.

doi: 10.5553/BSb/266669012023004005007

Het artikel is als volgt opgebouwd. In paragraaf 2 beschrijven we de techniek en ontwikkelingen achter seksuele deepfakes. In paragraaf 3 bespreken wij of en waarom het strafrecht een rol dient te spelen ten aanzien van seksuele deepfakes. In paragraaf 4 behandelen wij de huidige strafbepalingen die van toepassing kunnen zijn op dit digitaal gefingeerde beeldmateriaal. Tot slot evalueren wij in paragraaf 5 of er ruimte is voor verbetering van de strafrechtelijke bepalingen om beter recht te doen aan dit schadelijke beeldmateriaal.

2. De ontwikkeling van deepfakes en seksuele deepfakes

2.1 Deepfakes gedefinieerd

Een deepfake is een beeld, geluid of ander materiaal dat geheel of gedeeltelijk tot stand is gekomen met behulp van kunstmatige intelligentie (geavanceerde technische hulpmiddelen).[6] Het resultaat is vaak niet of nauwelijks van echt te onderscheiden.

In de samentrekking van 'deep' en 'fake' verwijst 'deep' naar 'deep learning'- of 'machine learning'-algoritmen die worden gebruikt om het materiaal middels kunstmatige intelligentie te creëren, en 'fake' naar de kunstmatigheid van het materiaal. Deepfakes zijn dus eigenlijk door computers gegenereerde synthetische media.[7]

Seksuele deepfakes zijn deepfakes van seksuele aard. Bij een van de eerste toepassingen van deepfake technologie, in 2017, werden al gezichten van bekende mensen geplaatst op de lichamen van pornoactrices.[8] Deze gebeurtenissen zorgden voor grote bekendheid van deepfaketechnologie. Het gebruik daarvan groeide daarna dan ook enorm. Uiteindelijk voerden platforms als Reddit een verbod op seksuele deepfakes in.[9] Het overgrote merendeel van deepfakes is pornografisch van aard en wordt zonder toestemming van afgebeelde personen gemaakt.[10]

2.2 Typen deepfakes

Doorgaans worden de volgende vier typen deepfakes onderscheiden:

1. *Gezichtsverwisseling* ('face swap'): het verwisselen van het gezicht van een persoon in een video met een ander gezicht.
2. *Attribuut verwerken*: het wijzigen van de kenmerken van de persoon in de video, bijvoorbeeld stijl of de kleur van het haar.
3. *Gezicht naspelen*: het overbrengen van de gezichtsuitdrukkingen van het gezicht van een persoon op de persoon in de doelvideo.
4. *Volledig synthetisch materiaal*: echt materiaal van personen gebruiken om een afbeelding te maken, terwijl het eindresultaat een persoon behelst die in het geheel niet bestaat.[11]

Voor het maken van deepfakes worden technologieën zoals neurale netwerken en generatieve kunstmatige intelligentie gebruikt.[12] Deze begrippen worden hierna toegelicht.

2.3 Neurale netwerken

Neurale netwerken bestaan uit een aantal lagen kunstmatige neuronen die met elkaar informatie uitwisselen. De neurale netwerken die gebruikt worden om deepfakes te maken, bootsen daarmee in wezen de werking van onze hersenen na. Deep learning-algoritmen gebruiken die netwerken om patronen in gegevens te vinden.[13] Zo 'leert' het algoritme zichzelf welke kenmerken belangrijk zijn en hoe deze zich tot elkaar verhouden. Op die manier kan een machine learning-algoritme uiteindelijk overtuigend realistische beelden construeren.

De beschikbaarheid van kwalitatief goede gegevens is essentieel voor een goed deepfake-algoritme. Het algoritme baseert zich op de beschikbare informatie en de kwaliteit van die gegevens heeft grote invloed op het resultaat van een machine learning-algoritme. Afhankelijk van de informatie die als input wordt gebruikt, worden meer of minder realistische deepfakes geproduceerd. Wanneer van iemand bijvoorbeeld maar enkele foto's online te vinden zijn, zal het veel lastiger zijn om van diegene een overtuigende deepfake te maken dan wanneer grote hoeveelheden beeldmateriaal van iemand te vinden zijn.

2.4 Generatieve adversariële netwerken

Een grote sprong in de kwaliteit en toegankelijkheid van deepfaketechnologie werd gemaakt door de ontwikkeling van generatieve adversariële netwerken (GAN's), zoals voorgesteld in 2014 door Ian Goodfellow en collega's.[14] Een GAN werkt met twee concurrerende modellen: een generatief en een discriminerend model. Het generatieve model creëert inhoud op basis van de beschikbare trainingsgegevens. Een discriminerend model test vervolgens de resultaten van het generatieve model door te beoordelen hoe groot de kans is dat het geteste voorbeeld afkomstig is van de dataset en niet van het generatieve model. Met de resultaten van deze tests

6 Van der Sloot e.a. 2021a, p. 2.
7 Kulk 2021, p. 3.
8 R. van den Boom, 'De wenselijkheid van zelfstandige strafbaarstelling betreffende de vervaardiging en verspreiding van non-consensuele deepfake-pornografie. Vanuit strafrechtelijk juridisch perspectief bezien' 2021, p. 21; Van der Sloot e.a. 2021.
9 D. Hawkins, 'Reddit bans "deepfakes", pornography using the faces of celebrities such as Taylor Swift and Gal Gadot', *The Washington Post* 8 februari 2018.
10 Zie H. Ajder e.a., *The State of Deepfakes 2019 Landscape, Threats, and Impact*, Amsterdam: Deeptrace 2019; K. Melville, 'The insidious rise of deepfake porn videos — and one woman who won't be silenced', *ABC News* 29 augustus 2019.
11 Wij houden voor dit onderscheid het Europol-rapport over deepfakes aan. Zie Europol, *Facing reality? Law enforcement and the challenge of deepfakes*, Luxemburg: Europol Innovation Lab, Publications Office of the European Union 2022, p. 9.
12 Zie, o.a. ook Van der Sloot e.a. 2021, p. 28-31.
13 Europol 2022, p. 6.
14 I.J. Goodfellow e.a., 'Generative Adversarial Networks', *arXiv* 2014; Van der Sloot e.a. 2021, p. 28.

doi: 10.5553/BSb/266669012023004005007

worden de modellen voortdurend verbeterd totdat het algoritme niet meer vast kan stellen wat het zelf gemaakt heeft en wat de trainingsgegevens zijn.[15]
Een andere techniek, 'autoencoder', maakt gebruik van een primitiever type kunstmatig neuraal netwerk. Deepfakes kunnen deze architectuur gebruiken door twee encoder-decoderparen te hebben, waarbij elk paar wordt gebruikt om te trainen op een beeldset. Tijdens het trainingsproces vindt en leert de ene encoder bijvoorbeeld overeenkomsten tussen de twee gezichten en reduceert deze tot hun gemeenschappelijke kenmerken, zoals de positie van ogen, neus en mond. Zo kan een 'face swap' gemaakt worden wanneer het beeld van gezicht A gecodeerd wordt met de gemeenschappelijke encoder, en gedecodeerd met de decoder van gezicht B. De decoder reconstrueert vervolgens het gezicht van persoon B met de uitdrukkingen en oriëntatie van gezicht A. Om een overtuigende video te produceren, moet dit op elk frame gebeuren.[16] Op deze manier kan iemands gezicht bijvoorbeeld in een bestaande pornofilm geplaatst worden.

2.5 Toegankelijkheid van deepfaketechnologie

De technische hulpmiddelen om deepfakes te maken zijn breed beschikbaar. Zo kan iedereen open-source tools gebruiken die beschikbaar zijn op GitHub of verspreid worden op online forums. Ook worden tools verkocht op illegale ondergrondse (digitale) markten. Bovendien bieden grote softwarebedrijven, zoals NVIDIA en Google, functies aan die (ook) in deepfakes gebruikt kunnen worden.[17]
Daarnaast zijn er apps ontwikkeld waarmee gebruikers, zonder technologische kennis, deepfakes kunnen maken. Zo was de app 'Deepnude' speciaal gemaakt om alledaagse foto's van vrouwen om te zetten in naaktfoto's.[18] Veel software die algemeen beschikbaar is voor het publiek kan dus worden misbruikt voor het maken van seksuele deepfakes.

3. De rol van het strafrecht ten aanzien van seksuele deepfakes

3.1 Rechtvaardiging van strafbaarstelling

Seksuele deepfakes kunnen worden beschouwd als een vorm van seksueel misbruik met beeldmateriaal, ook wel 'wraakporno' genoemd.[19] Het gaat dan specifiek over seksueel beeldmateriaal dat zonder toestemming van de afgebeelde persoon of personen wordt gemaakt en/of verspreid. Uit de beleidsreactie van minister Yeşilgöz-Zegerius op de motie uit 2022 blijkt dat de regering aanpassing van het Wetboek van Strafrecht niet nodig vindt voor de aanpak van seksuele deepfakes.[20] Wij zullen deze positie in dit artikel onderzoeken.
Grofweg kan worden gesteld dat strafbaarstelling gerechtvaardigd kan zijn wanneer handelingen schadelijk en onrechtmatig zijn.[21] Dit valt te herkennen als de eerste criteria van De Roos.[22] In dit deel bekijken wij of er op basis van deze criteria ten aanzien van seksuele deepfakes een rol weggelegd kan zijn voor het strafrecht.[23] Deze vragen zijn, ten aanzien van seksueel misbruik met beeldmateriaal, door Goudsmit Samaritter ook behandeld in haar proefschrift over seksuele deepfakes.[24]

3.2 Schadelijkheid

De schadelijkheid van seksuele deepfakes blijkt bijvoorbeeld uit onderzoek van McGlynn e.a.[25] Dat onderzoek toont duidelijk aan dat seksueel misbruik middels beeldmateriaal, waaronder seksuele deepfakes, substantiële schade veroorzaakt. De schade bestaat gewoonlijk uit een breuk met iemands sociale gemeenschap, een voortdurende angst, het risico op isolatie, en een inperking van vrijheid. Dit kan ervoor zorgen dat het leven van slachtoffers ronduit ondraaglijk wordt.[26] In sommige gevallen gaat dat zelfs zo ver dat slachtoffers zelfmoord plegen.[27] De schadelijkheid van seksuele deepfakes behoeft hiermee volgens ons geen nadere toelichting.

3.3 Onrechtmatigheid

De onrechtmatigheid van seksuele deepfakes is gelegen in de schending van een grote hoeveelheid mensenrechten- en grondrechten.[28] Seksuele deepfakes zijn dus *mala in se*. Hieronder wordt toegelicht hoe deze *mala* zich in seksuele deepfakes kunnen uiten, namelijk als een (1) schending van de autonomie, (2) schending van seksuele integriteit, (3) schending van privacy, (4) schen-

15 Zie ook V. Ciancaglini e.a., *Malicious Uses and Abuses of Artificial Intelligence*: Trend Micro Research 2020, p. 54.
16 T.T. Nguyen e.a., 'Deep Learning for Deepfakes Creation and Detection: A Survey', *arXiv* 2022, p. 3.
17 Ciancaglini e.a. 2020, p. 56.
18 S. Cole, 'This Horrifying App Undresses a Photo of Any Woman With a Single Click', *Vice* 26 juni 2019.
19 De term 'wraakpornografie' is misleidend, daar die de indruk kan wekken dat het slechts om materiaal gaat dat de bedoeling heeft 'kwetsend' te zijn, zie C. McGlynn en E. Rackley, 'More than 'Revenge Porn': Image-Based Sexual Abuse and the Reform of Irish Law', *Irish Probation Journal* 2017, 38-52.
20 D. Yeşilgöz-Zegerius, 'Beleidsreactie op de WODC-onderzoeken naar de regulering van deepfakes en immersieve technologieën', *Kamerstukken II* 2022/23, 26643, nr. 1041.
21 Voor 'public harmful wrongs', strafbaarstellingscriteria in democratische rechtsstaten, zie R.A. Duff e.a., *The Boundaries of the Criminal Law*, Oxford: Oxford University Press 2010, p. 7-8.
22 Th.A. de Roos, Strafbaarstelling van economische delicten: een crimineel-politieke studie (diss. Utrecht), Arnhem: Gouda Quint 1987, p. 54.
23 Evenwel zijn er meer criteria of benaderingen uit wetenschappelijke literatuur te destilleren. Voor deze bijdrage is gekozen om de nadruk te leggen op de schadelijkheid en onrechtmatigheid.
24 M.L.R Goudsmit, *The Wrongness of Image-based Sexual Abuse* (diss. Oxford), Oxford 2022.
25 C. McGlynn e.a., '"It's Torture for the Soul": The Harms of Image-Based Sexual Abuse', *Social & Legal Studies* 2020, 1-22, p. 17.
26 McGlynn e.a. 2020, 10-6.
27 Zie bijvoorbeeld L. Baard, 'Onur (14) pleegt zelfmoord na ontdekken naaktfoto op Instagram', *AD* 21 februari 2017; K. Forster, 'Tiziana Cantone: Woman's Suicide After Sex Tape Went Viral Prompts Calls for Stronger Online Privacy Laws', *The Independent* 16 september 2016.
28 Goudsmit 2022, p. 292-3.

doi: 10.5553/BSb/266669012023004005007

ding van uitingsvrijheid, en (5) schending van het gelijkheidsbeginsel.[29]

3.3.1 Schending van de autonomie

Volgens onderzoek van Deeptrace is ruim 96% van alle deepfakes pornografisch.[30] Doorgaans worden die deepfakes zonder toestemming van de afgebeelde persoon gemaakt, hetgeen inbreuk maakt op de autonomie van de afgebeelde persoon. Deze persoon kan er namelijk niet meer voor kiezen om *niet* pornografisch te worden afgebeeld.[31]

Door het creëren van seksuele deepfakes zonder toestemming van de afgebeelde persoon wordt dus diens recht op autonomie geschonden. Dat recht omvat het zelf beslissingen kunnen nemen over de eigen persoon, waaronder zijn of haar seksualiteit.[32]

3.3.2 Schending van de seksuele integriteit

Seksuele deepfakes die zonder toestemming worden gemaakt, maken daarnaast inbreuk op de seksuele integriteit. Dat gebeurt wanneer iemand aan een handeling wordt onderworpen die betrekking heeft op diens seksualiteit, maar waarover diegene geen zeggenschap kon hebben.[33]

Seksuele gedragingen, inclusief jezelf op seksuele wijze weergeven of uiten, zijn keuzes die onder de seksuele integriteit vallen. Seksuele weergaven middels een deepfake van iemand die daarvoor geen toestemming heeft gegeven, maken daarmee een inbreuk op de seksuele integriteit en dergelijke inbreuken worden doorgaans als seksueel misdrijf gekwalificeerd.[34]

3.3.3 Schending van de privacy

Ook de privacy wordt geschonden door seksuele deepfakes die zonder toestemming gemaakt worden. Deze privacyschending bestaat eruit dat de publicatie van seksuele deepfakes iemands relaties en sociale positie aan kan tasten. Volgens het Europees Hof voor de Rechten van de Mens (EHRM) hebben verdragsstaten een positieve verplichting om ook de privacy tussen burgers onderling te waarborgen.[35] Waar het gaat om bijzonder belangrijke aspecten van iemands wezen of identiteit, hebben verdragsstaten een nauwe *margin of appreciation* ten aanzien van de bescherming van privacy.[36] Seksualiteit wordt beschouwd als een dergelijk belangrijk privacyaspect.[37] Waar het recht op privacy door burgers jegens elkaar op seksuele wijze geschonden wordt, hebben staten dus algauw een verplichting tot ingrijpen om het recht op privacy in artikel 8 van het Europees Verdrag voor de Rechten van de Mens (EVRM) afdoende te beschermen. Uit de jurisprudentie van het EHRM volgt logischerwijs dat seksuele deepfakes, zeker wanneer deze gepubliceerd en verspreid zijn, een horizontale privacyschending opleveren waartegen staten verplicht moeten optreden.[38]

De bovenstaande drie schendingen kunnen allemaal leiden tot een schending van artikel 8 van het EVRM. Als zodanig zijn ze ook in jurisprudentie van het EHRM terug te vinden.[39]

3.3.4 Schending van de uitingsvrijheid

Naast de bovengenoemde schendingen van artikel 8 EVRM, zijn ongewenste seksuele deepfakes ook een schending van het recht op uitingsvrijheid van de afgebeelde persoon in artikel 10 EVRM. Ten aanzien van dit argument heeft een van ons beargumenteerd dat het dan gaat om een schending van het *negatieve* recht op vrije meningsuiting: het recht om *niet* te uiten wanneer dat gewenst is.[40] Seksuele deepfakes brengen slachtoffers in een positie waarin gesuggereerd wordt dat ze een pornografische uiting doen, zonder dat de afgebeelde personen dat hadden gewild of zelf hebben gedaan. Het EHRM erkent het recht op negatieve uitingsvrijheid.[41] Goudsmit Samaritter betoogt dat het EHRM niet-consensuele seksuele deepfakes als een schending van het negatieve recht op uitingsvrijheid zou kunnen beschouwen.[42] Omdat er echter nog geen EHRM-jurisprudentie over negatieve uitingsvrijheid bij seksuele deepfakes is, laten wij hier in het midden in hoeverre dit recht inderdaad door seksuele deepfakes wordt geschonden.

3.3.5 Schending van het gelijkheidsbeginsel

Ten slotte kan worden betoogd dat niet-consensuele seksuele deepfakes een inbreuk maken op het gelijkheidsbeginsel.[43]

Het gelijkheidsbeginsel in het EVRM gaat specifiek over de toepassing en bescherming van andere rechten in het EVRM, en moet dus worden gelezen in combinatie met andere rechtenschendingen. In het geval van seksuele

29 Zie voor uitgebreidere behandeling van deze punten Goudsmit 2022, p. 26 e.v. en p. 152 e.v.

30 Ajder e.a. 2019, p. 1.

31 J. Raz, *The Morality of Freedom*, Oxford: Oxford University Press 1986, p. 374-375.

32 Zie bijvoorbeeld EHRM 31 juli 2000, ECLI:CE:ECHR:2000:0731JUD003576597 (*ADT t. United Kingdom*).

33 Zie voor het onderscheid tussen autonomie en integriteit bijvoorbeeld J.W. Herring en J. Wall, 'The Nature and Significance of the Right to Bodily Integrity', *The Cambridge Law Journal* 2017, 566-588, p. 576. Zie ook EHRM 27 juli 2021, ECLI:CE:ECHR:2021:0727JUD007263117 (*X/Y t. The Netherlands*).

34 S.P. Green, *Criminalizing Sex: A Unified Liberal Theory*, New York: Oxford University Press 2020, p. 4.

35 EHRM 10 april 2007, ECLI:CE:ECHR:2007:0410JUD000633905 (*Case of Evans t. United Kingdom*), par. 75.

36 Council of Europe en European Court of Human Rights, Guide on Article 8 of the European Convention on Human Rights: Right to Respect for Private Life, Home and Correspondence, 31 august 2020, par. 7.

37 EHRM 29 april 2002, ECLI:CE:ECHR:2002:0429JUD000234602 (*Pretty t. United Kingdom*), par. 71.

38 Zie bijvoorbeeld EHRM 25 december 2001, ECLI:CE:ECHR:2001:0925JUD004478798, par. 56-58 (*P.G. & J.H. t. United Kingdom*). Bij de publicatie van seksuele deepfakes zonder toestemming is overigens ook sprake van een inbreuk op het recht op bescherming van persoonsgegevens.

39 Zie bijvoorbeeld EHRM 31 juli 2000, ECLI:CE:ECHR:2000:0731JUD003576597 (*ADT t United Kingdom*); EHRM 12 juli 1977, ECLI:CE:ECHR:1976:0519DEC000695975 (*Brüggemann and Scheuten t. Federal Republic of Germany*), par. 55; EHRM 26 maart 1985, ECLI:CE:ECHR:2021:0727JUD007263117, par. 27 (*X/Y t. The Netherlands*).

40 Goudsmit 2022, p. 218 e.v.

41 EHRM 3 april 2012, ECLI:CE:ECHR:2010:1102JUD004172306 (Case of Gillberg t. Sweden), par. 86. Zie ook Council of Europe en European Court of Human Rights, Guide on Article 10 of the European Convention on Human Rights: Freedom of expression, 31 december 2020, par. 25.

42 Goudsmit 2022, p. 227.

43 Goudsmit 2022, p. 272 e.v..

doi: 10.5553/BSb/266669012023004005007

deepfakes gaat het dan over de hiervoor behandelde rechten. Vrijwel alle seksuele deepfakes zijn afbeeldingen van vrouwen.[44] De rechtenschendingen die hierboven zijn besproken, komen in die context dan ook bijna alleen maar bij vrouwen voor. Wanneer onvoldoende tegen seksuele deepfakes wordt opgetreden, levert dat dus voor vrouwen meer kans op rechtenschendingen op dan voor mannen. Een schending van artikel 14 EVRM (non-discriminatie) moet gezocht worden in het verschil in de mate waarin (in casu) mannen versus vrouwen ongestoord van hun rechten gebruik kunnen maken. Bij seksuele deepfakes geldt dat de kans op slachtofferschap voor vrouwen aanzienlijk hoger is dan voor mannen: de meest toegankelijke software – een telefoonapplicatie – voor het maken van seksuele deepfakes is zelfs zo geprogrammeerd dat deze alleen ten aanzien van vrouwen werkt.[45] Het gaat ons hier niet over de schade die slachtoffers ondervinden, maar enkel over de algemene bescherming van rechten. Die bescherming is, door de scheve verhoudingen waarmee seksuele deepfakes voorkomen bij mannen versus bij vrouwen, ongelijk. Daardoor kan dit als een schending van artikel 14 EHRM worden aangemerkt.

Naast het feit dat seksuele deepfakes vrijwel exclusief vrouwen treffen, ervaren vrouwen gemiddeld ook meer negatieve gevolgen van seksueel misbruik middels beeldmateriaal in algemene zin: zij worden bijvoorbeeld sneller onderworpen aan 'slutshaming' en negatieve seksualisering.[46] Goudsmit Samaritter betoogt daarom dat niet-consensuele seksuele deepfakes een schending van het gelijkheidsbeginsel met zich meebrengen.[47]

3.4 De rol voor het strafrecht

Op basis van de bovenstaande criteria van schadelijkheid en onrechtmatigheid van seksuele deepfakes kan strafbaarstelling gelegitimeerd worden. Daar het gaat om potentieel ernstige mensenrechtenschendingen zou het zelfs zo kunnen zijn dat strafbaarstelling een internationaalrechtelijke verplichting is.[48] Het EHRM heeft eerder bepaald dat zeer ernstige schendingen van het EVRM verdragsstaten verplichten tot inzet van het strafrecht. In de zaak *Volodina t. Russia* ging het over ernstig huiselijk en seksueel geweld, waarvan de onrechtmatigheid en schadelijkheid enige gelijkenissen vertonen met seksuele deepfakes.[49] Zowel huiselijk geweld als seksuele deepfakes maken bijvoorbeeld (op vergelijkbare wijze) inbreuk op het recht op autonomie en privéleven, en leveren een vergelijkbaar soort schade op.[50]

Het is dus mogelijk dat het Hof een vergelijkbare conclusie zou trekken ten overstaan van seksuele deepfakes. Wat echter duidelijk is, is dat de staat – met of zonder strafrecht – verplicht is tot optreden tegen niet-consensuele seksuele deepfakes omdat die mensen- en grondrechten schenden. Uit bovenstaande analyse blijkt dat het strafrecht daarvoor een gelegitimeerd middel is, omdat ongewenste seksuele deepfakes schadelijk en onrechtmatig zijn. Daarnaast kunnen particulieren elkaar middels civiele procedures aanspreken en bijvoorbeeld de verwijdering van materiaal en een schadevergoeding afdwingen.

4. Huidige strafbaarstelling van seksuele deepfakes

4.1 Strafrechtelijk onderscheid tussen deepfakes en seksuele deepfakes

De ontwikkeling van deepfakes (sinds 2017) is relatief jong, zeker in juridische termen. Er is op dit moment niet één '*one size fits all*' strafrechtsartikel voor de aanpak van ongewenste seksuele deepfakes. Desondanks biedt het Nederlandse strafrecht al verschillende mogelijkheden voor de strafrechtelijke aanpak van niet-consensuele seksuele deepfakes.

Seksuele deepfakes zijn een verbijzondering van deepfakes in het algemeen. Het voornaamste deel van de bestaande bepalingen ziet echter niet specifiek op seksuele deepfakes. In het onderstaande staan wij kort stil bij huidige strafrechtbepalingen waar seksuele deepfakes onder zouden kunnen vallen, ook als het seksuele element geen bestanddeel van de strafrechtbepaling is. Dit betreffen achtereenvolgens smaad (art. 261 Sr) en laster (art. 262 Sr), het misbruik van biometrische gegevens (art. 231b Sr), en de strafbaarstelling van seksueel misbruik middels beeldmateriaal (ook wel 'wraakporno' genoemd) (art. 139h Sr). In het onderstaande behandelen we hoe geschikt deze bepalingen zijn voor de aanpak van seksuele deepfakes. Daarbij kijken wij ook kort naar het principe van *fair labelling*, dat inhoudt dat de naam en categorie van delicten de aard van de onrechtmatigheid van de strafbaar gestelde handeling dienen weer te geven.[51] Een wet is voor alle betrokkenen eerlijker wanneer aan het principe van fair labelling wordt voldaan.[52]

44 Zie Ajder e.a. 2019, p. 2.
45 Cole 2019.
46 N. Henry e.a., Image-Based Sexual Abuse: A Study on the Causes and Consequences of Non-Consensual Nude or Sexual Imagery, London: Routledge 2020, p. 39.
47 Goudsmit 2022, p. 277.
48 Een van de auteurs heeft beargumenteerd dat de ernst van de mensenrechtenschendingen die tot stand komen door seksuele deepfakes en andere vormen van seksueel misbruik middels beeldmateriaal zodanig groot kan zijn dat die in strijd zijn met het folterverbod in art. 3 EHRM. Zie Goudsmit 2022, p. 243 e.v.
49 EHRM 4 november 2019, ECLI:CE:ECHR:2019:0709JUD004126117 (*Case of Volodina t Russia*), par. 128, en de *separate opinion* van Judge Pinto de Albuquerque par. 20.
50 Zie voor de schadevergelijking McGlynn e.a. 2020, p. 17; en J. Herring, 'The Meaning of Domestic Violence: Yemshaw v London Borough of Hounslow [2011] UKSC 3', *Journal of Social Welfare and Family Law* 2011, 297-304, p. 300 e.v.
51 A. Ashworth, 'The Elasticity of Mens Rea', in: C.F.H. Tapper (red.), *Crime, Proof & Punishment: Essays in Memory of Sir Rupert Cross*, London: Butterworths, 1981, p. 53.
52 J. Chalmers & F. Leverick, 'Fair Labelling in Criminal Law', *The Modern Law Review* 2008, 217-246, p. 29.

doi: 10.5553/BSb/266669012023004005007

4.2 Smaad en laster

Voor smaad en laster is vereist dat het slachtoffer in diens eer of goede naam wordt aangerand. Voor een van ons was dat reden om eerder te betogen dat deze bepalingen niet geschikt zijn om iemand te vervolgen voor het maken van niet-consensueel seksueel beeldmateriaal.[53] In 2014 wees toenmalig minister van Justitie Opstelten naar smaad en laster als strafbepalingen waar seksueel misbruik met beeldmateriaal onder zou kunnen vallen.[54]

Seksuele deepfakes zullen gepubliceerd of gedeeld moeten zijn om onder smaad en laster te vallen. Seksuele deepfakes die puur voor 'eigen gebruik' gemaakt worden, vallen hier niet onder.[55] Het moet gaan om 'een bepaald feit' dat geschikt is om 'iemands integriteit aan te tasten'.[56] Aan dat feit moet ruchtbaarheid worden gegeven, wat impliceert dat het feit aan een groter publiek wordt tentoongesteld. De Hoge Raad heeft eerder bepaald dat seksuele gedragingen die zichtbaar zijn op een film, 'bepaalde feiten' kunnen zijn in de zin van artikel 261 Sr.[57]

Het zou kunnen dat het louter weergeven van een seksuele deepfake – zoals het uploaden van de deepfake op een voor het publiek toegankelijke website – onvoldoende is om te spreken van smaad of laster. Het is onduidelijk of daaruit voldoende blijkt dat de verdachte opzettelijk de eer of goede naam van het slachtoffer aanrandde door hem een bepaald feit ten laste te leggen. Wanneer echter een naam gekoppeld wordt aan een (seksueel) filmpje, ligt dat anders. Daarmee worden de seksuele gedragingen in de beelden namelijk toegedicht aan de genoemde persoon (ook al is dat niet daadwerkelijk de afgebeelde persoon).[58]

Het verwijzen naar bijvoorbeeld een seksuele deepfake alsmede een bijgevoegde titel en/of tekst moeten dus tezamen worden bezien om te kunnen beoordelen of er sprake is van smaad.[59] Iets dergelijks was aan de orde in 2019, toen een man in hoger beroep werd veroordeeld voor smaad nadat hij seksuele deepfakes van zijn ex-partner via internet had verspreid.[60] De verdachte had naaktfoto's verspreid waarop hij het hoofd van zijn ex-partner gemonteerd had op afbeeldingen van lichamen die deelnamen aan seksuele handelingen. Omdat het bij smaad (en laster) erom gaat dat het slachtoffer door 'een bepaald feit' in een ongunstig daglicht is komen te staan, is het voor de veroordeling onder smaad irrelevant of het om een daadwerkelijke naaktfoto gaat of om een gemanipuleerde naaktfoto (in casu seksuele deepfakes). Omdat smaad en laster echter specifiek gericht zijn op het 'in een ongunstig daglicht' zetten van de afgebeelde persoon, geven deze delicten de aard van de schending niet correct weer.

4.3 Misbruik van (biometrische) gegevens

Misbruik van (biometrische) gegevens of kenmerken is strafbaar gesteld in artikel 231a en artikel 231b Sr. Dit wettelijk kader gaat niet specifiek over misbruik met seksuele beelden, maar over het misbruik van iemands identiteit *in abstracto*. Zowel artikel 231a als artikel 231b Sr gaan identiteitsmisbruik tegen. Wij bespreken hier de toepasbaarheid van die strafbepalingen op seksuele deepfakes.

In artikel 231a Sr gaat het specifiek over biometrische kenmerken of persoonsgegevens.[61] Het kan daarbij bijvoorbeeld gaan over het gezicht, de ogen of de irissen. Het kan echter ook over bepaalde gedragskenmerken gaan, bijvoorbeeld een manier van spreken, schrijven of lopen. Hoewel ook bij seksuele deepfakes gebruik wordt gemaakt van iemands biometrische kenmerken, is artikel 231a Sr niet zonder meer geschikt voor het aanpakken van misbruik van seksuele deepfakes. Er moet namelijk ook sprake zijn van een oogmerk om iemands identiteit te verhelen of te misbruiken. Wanneer seksuele deepfakes gebruikt worden om live (online) seksdiensten aan te bieden met het voorkomen van een ander, zou artikel 231a Sr wellicht wél van toepassing kunnen zijn. Andere omstandigheden waar het artikel op ziet, zijn bijvoorbeeld het aanbrengen van iemands vingerafdrukken op een gebruikt wapen, of middels een deepfakevideo de indruk wekken dat iemand oproept tot geweld.[62]

Volgens Van der Sloot e.a. biedt artikel 231b Sr in beginsel een geschikt kader voor het vervolgen van seksuele deepfakes.[63] Het wetsartikel ziet echter op niet-biometrische kenmerken, terwijl die kenmerken met name relevant zijn bij seksuele deepfakes. Om seksuele deepfakes onder de reikwijdte van artikel 231b Sr te laten vallen, zullen deze vergezeld moeten worden door andere, niet-biometrische kenmerken van de afgebeelde persoon. Dat kan bijvoorbeeld een naam, telefoonnummer of adres zijn. Er is dan sprake van 'misbruik van de aanduiding van de persoon'.[64] Dit is al snel het geval als seksuele deepfakes in de titel of het bijschrift bijvoorbeeld de naam van de afgebeelde persoon bevatten.

Het is echter de vraag of artikel 231a en artikel 231b Sr wel bedoeld zijn om het soort misbruik dat met seksuele deepfakes tot stand komt, aan te pakken. Bij seksuele deepfakes gaat het vaak om daders die seksuele prikke-

53 Het argument dat seksualiteit niet de eer en goede naam aantast, is verder uitgewerkt in M.L.R. Goudsmit, 'De wijzende vinger bekeken: over de strafbaarstelling van wraakpornografie', *NJB* 2018/24, 1721-1729, p. 1722-4.

54 Vragen van het lid Rebel over het strafbaar stellen van «wraakporno» (ingezonden 24 november 2014) en antwoord van Minister Opstelten (Veiligheid en Justitie) (ontvangen 23 december 2014). *Handelingen II*, 2014/15, nr. 933.

55 Zie bovendien het klachtvereiste, art. 269 lid 1 Sr.

56 Janssens in *T&C Strafrecht*, art. 261 Sr, aant. 9 (online, laatst bijgewerkt op 1 februari 2023).

57 HR 3 december 2013, ECLI:NL:HR:2013:1556, r.o. 2.5.

58 HR 3 december 2013, ECLI:NL:HR:2013:1556, r.o. 2.5. Zie ook HR 4 december 2018, ECLI:NL:HR:2018:2240, r.o. 2.5.

59 Onduidelijke tekst die niet voldoende toegespitst is op het vermoedelijke slachtoffer is geen smaad. Zie HR 4 december 2018, ECLI:NL:HR:2018:2240.

60 Gerechtshof 's-Hertogenbosch 28 januari 2019, ECLI:NL:GHSHE:2019:252.

61 Van der Sloot e.a. 2021, p. 55-57.

62 *Kamerstukken II* 2011/12, 33352, nr. 3, p. 24.

63 Van der Sloot e.a. 2021, p. 66.

64 Van der Sloot e.a. 2021, p. 57.

doi: 10.5553/BSb/266669012023004005007

ling op willen wekken, macht uit willen oefenen over het slachtoffer, of bij een bepaalde groep willen horen.[65] Bij misbruik van biometrische kenmerken en gegevens gaat het daarentegen over een veelal frauduleus doel.[66] Bij misbruik van biometrische gegevens staan vooral het gemanipuleerde beeld en de gevolgen die daaruit voortvloeien centraal en niet de *seksuele* connotatie van de beelden.

Wetssystematisch is het dan wellicht niet geheel passend om seksuele deepfakes onder deze artikelen te scharen. Immers: de indeling van bepaalde strafbepalingen is meestal gebaseerd op het rechtsgoed in kwestie. Vanuit dat licht bezien is het misbruik van seksueel beeldmateriaal mogelijk niet geschikt om onder te brengen in Titel XII (Valsheid met geschriften, gegevens en biometrische kenmerken).[67]

4.4 Seksueel misbruik met beeldmateriaal ('wraakporno')

Het laatste wetsartikel dat wij hier bespreken is artikel 139h Sr, dat het vervaardigen en/of verspreiden van afbeeldingen van seksuele aard strafbaar stelt. Hierbij gaat het veelal om 'wraakporno'.[68] Uit de memorie van toelichting bij artikel 139h Sr blijkt niet dat de wetgever ook seksuele deepfakes heeft bedoeld met 'seksueel beeldmateriaal'. Toch wordt inmiddels aangenomen, bijvoorbeeld door het OM (in de vervolging van de zaak met Welmoed Sijtsma) en door minister Yeşilgöz-Zegerius, dat dergelijk materiaal onder de reikwijdte van artikel 139h Sr valt, ook wanneer dit materiaal niet gepubliceerd wordt.[69] Het vervaardigen van seksueel beeldmateriaal zonder toestemming van de afgebeelde persoon is een schending van de persoonlijke levenssfeer. Omdat dit materiaal intiem en gevoelig is, zouden mensen zelf moeten kunnen beslissen of zij willen dat dit wordt vervaardigd.[70]

De eerste vraag ten aanzien van seksuele deepfakes is of het maken van een deepfake als 'vervaardiging' geldt op dezelfde manier als bijvoorbeeld 'filmen' dat doet. Indien dat het geval is, zouden seksuele deepfakes onder artikel 139h lid 1 Sr vallen, waarvoor publicatie niet nodig is. Volgens Van der Sloot e.a. zou die uitleg van 'vervaardiging' moeten worden afgewezen. Volgens hen gaat het bij 'voorbeelden in de wetsgeschiedenis' steeds om 'een fysieke situatie waarbij iemand met een foto- of videocamera beelden maakt'.[71] Aangezien deepfaketechnologie zeer nieuw is, spreekt de wetsgeschiedenis daar inderdaad van. Dat sluit onzes inziens echter niet uit dat deepfaketechnologie een nieuwe manier van vervaardigen van afbeeldingen mogelijk maakt, die ook onder artikel 139h lid 1 Sr kan vallen.

In antwoord op Kamervragen over de strafbaarstelling van seksuele deepfakes verwijst minister Yeşilgöz-Zegerius naar een vonnis van de rechtbank Den Haag waarbij iemand is veroordeeld voor seksuele deepfakes, kennelijk onder artikel 139h lid 1 Sr.[72] Op de telefoon van de verdachte werden pornografische beelden gevonden, waarin door deepfakemanipulatie het gezicht van een collega en het gezicht van een kennis van de verdachte waren verwerkt. De beelden waren niet gepubliceerd en de slachtoffers waren niet van op de hoogte van het bestaan van deze beelden. Evenwel resulteerde dit in een veroordeling voor het in artikel 139h lid 1 Sr strafbaar gestelde feit.

Daarnaast valt het openbaar maken van seksuele deepfakes, zijnde afbeeldingen van seksuele aard, onder de reikwijdte van artikel 139h lid 2 Sr wanneer de verdachte wist dat die openbaarmaking nadelig kon zijn voor de afgebeelde persoon. In het licht van sterke maatschappelijke normen over 'waar' en 'wanneer' naaktheid en seksualiteit passend zijn, zou eenvoudig vastgesteld moeten kunnen worden dat een verdachte wist dat openbaar maken van seksuele afbeeldingen nadelig kan zijn.

In een nota naar aanleiding van het verslag ter voorbereiding van de Wet seksuele misdrijven stelt de minister derhalve dat 'niet alleen het verspreiden, maar ook het zonder toestemming van de afgebeelde persoon vervaardigen en voorhanden hebben van seksueel deepfake beeldmateriaal van diegene' strafbaar is op grond van artikel 139h Sr.[73] De uitleg van Van der Sloot e.a. dat enkel het verspreiden van seksuele deepfakes door de strafbaarstelling van artikel 139h Sr wordt gedekt, strookt dus niet met de opvatting van de minister.[74] Wij achten op basis van het bovenstaande dan ook de kans groot dat naast het verspreiden, ook het maken van niet-consensuele seksuele deepfakes strafbaar is op grond van artikel 139h Sr.

Ten slotte kan naast de strafbaarheid van individuele daders, ook worden gedacht aan strafbaarheid van bedrijven die het creëren en verspreiden van seksuele deepfakes mogelijk maken. Wij gaan hier niet nader in op de vraag in hoeverre dat haalbaar, wenselijk, en/of rechtvaardigbaar is. Over zaken als 'platform responsibility' is reeds veel informatie beschikbaar.[75] Deson-

65 C. Mortreux e.a., Understanding the Attitudes and Motivations of Adults who engage in Image-Based Abuse: Australian Government eSafety Commissioner 2019. Zie ter vergelijking ook over de niet-seksuele motieven van verkrachters: C.T. Palmer, 'Twelve reasons why rape is not sexually motivated: A skeptical examination', *The Journal of Sex Research* 1988, p. 512-530.

66 Het lijkt erop dat seksuele deepfakes ook voor financieel gewin gemaakt worden. Daarover zijn echter geen gegevens beschikbaar. Hier is dus ruimte voor vervolgonderzoek.

67 J. de Hullu, Materieel Strafrecht: Over algemene leerstukken van strafrechtelijke aansprakelijkheid naar Nederlands recht, Deventer: Wolters Kluwer 2021, p. 69-70.

68 Een van ons heeft beargumenteerd dat dit geen geschikte of gepaste term is, zie M.L.R. Goudsmit, 'What Makes a Sex Crime? A Fair Label for Image-based Sexual Abuse', *Bsb* 2021. afl. 2, p. 67-75. Zie ook McGlynn en Rackley 2017.

69 Zie D. Yeşilgöz-Zegerius en F.M. Weerwind, 'Antwoorden Kamervragen over het verbieden van deepfakes', 14 november 2022.

70 Zie paragraaf 3. Zie ook *Kamerstukken II* 2018/19, 35080, nr. 3, p. 4.

71 Van der Sloot e.a. 2021, p. 62.

72 Zie D. Yeşilgöz-Zegerius en F.M. Weerwind, 'Antwoorden Kamervragen over het verbieden van deepfakes', 14 november 2022; en rechtbank Den Haag 4 maart 2021, ECLI:NL:RBDHA:2021:1885, r.o. 6.3.

73 *Kamerstukken II* 2022/23, 36222, nr. 7, p. 13.

74 Van der Sloot e.a. 2021.

75 Zie bijvoorbeeld A. Kuczerawy, 'Intermediary liability & freedom of expression: Recent developments in the EU notice & action initiative', *CL&SR* 2015/31, p. 46-56.

doi: 10.5553/BSb/266669012023004005007

danks blijft het een lastig vraagstuk of en in welke gevallen bijvoorbeeld het 'hosten' van seksueel beeldmateriaal dat zonder toestemming openbaar gemaakt is, onder de strafrechtelijke verantwoordelijkheid van de internetprovider of website-eigenaar kan vallen. Daarnaast kan men zich afvragen of het maken of faciliteren van software die het creëren van deepfakes mogelijk maakt, strafbaar zou kunnen of moeten zijn. Wij stellen dat nader onderzoek noodzakelijk is om vast te stellen of bijvoorbeeld een medeplichtigheidsconstructie tot strafrechtelijke aansprakelijkheid kan leiden in die gevallen waarin een bedrijf bijvoorbeeld opzettelijk gelegenheid (of middelen) verschaft tot het plegen van het misdrijf.[76]

4.5 Mogelijke uitzondering op strafbaarstelling seksuele deepfakes

Een punt dat wij hier kort aan willen kaarten, is dat er bij het maken en verspreiden van seksuele deepfakes een spanningsveld bestaat tussen enerzijds de in paragraaf 3.3 besproken rechten van de afgebeelde persoon en anderzijds het recht op uitingsvrijheid van degene die deze afbeeldingen maakt en/of verspreidt. Het creëren van beelden – al dan niet met behulp van AI – valt immers in beginsel onder de artistieke vrijheid. Voor sommige potentieel strafwaardige beelden kan een beroep worden gedaan op de kunstexceptie.[77]

Het is de vraag of dat ook bij seksuele deepfakes het geval zal zijn.[78] In de zaak *Vereinigung Bildender Künstler t. Austria* werd de kunstexceptie erkend als onderdeel van de vrijheid van meningsuiting (art. 10 EVRM). Het ging om een serie van pornografische schilderijen waar de hoofden van onder meer bekende Oostenrijkse personen op waren geplakt.[79] Omdat de beelden satirisch waren en niet suggereerden een weergave van de werkelijkheid te zijn, honoreerde het EHRM het beroep op de kunstexceptie.[80] Bij seksuele deepfakes is dat echter doorgaans niet aan de orde, omdat die steeds realistischer worden en ook die schijn misbruiken. Wanneer een seksuele deepfake overduidelijk satirisch is bedoeld, kan eventueel een beroep worden gedaan op de kunstexceptie. Bij realistisch gemanipuleerde pornografische beelden is het niet zeer waarschijnlijk dat een beroep op de kunstexceptie succes zal hebben.

Een situatie waar nog meer onderzoek naar kan worden gedaan, is wat de juridische situatie is en/of zou moeten zijn ten aanzien van beeldmateriaal dat op synthetische wijze tot stand is gekomen, waarbij het eindresultaat een persoon weergeeft die niet echt bestaat, maar waarbij voor het maken wel degelijk beeldmateriaal van echte personen is gebruikt (bijvoorbeeld voor het trainen van algoritmes). In de besproken strafbaarstellingen in deze paragraaf gaan wij ervan uit dat het basismateriaal en de uiteindelijke beelden tot het slachtoffer te herleiden zijn. Er kunnen echter grijze gebieden ontstaan wanneer afbeeldingen bijvoorbeeld slechts zeer geringe gelijkenissen vertonen met een bestaand persoon. Het geleden nadeel voor die persoon is dan lastiger aantoonbaar.[81] Ook kan de vraag worden gesteld of de betrokkene een natuurlijk persoon moet zijn of dat expliciet in de delictsomschrijving moet staan dat het een virtuele creatie betreft, zoals bij ontucht (art. 248a Sr) en grooming (art. 248e Sr). Wij beantwoorden deze vraag in dit artikel verder niet. Duidelijk is dat volledig synthetische seksuele deepfakes reeds werkelijkheid zijn en daarmee nader onderzoek waard zijn.

5. Slotbeschouwing

Hierboven hebben wij uitgelegd wat seksuele deepfakes zijn, dat die gerechtvaardigd strafbaar gesteld kunnen worden, en zijn de huidige strafbepalingen toegelicht.

Wij concluderen dat de strafwet al enige aanknopingspunten biedt voor het aanpakken van deepfakes in het algemeen, en in mindere mate van seksuele deepfakes in het bijzonder. De huidige strafbaarstelling van seksueel misbruik middels beeldmateriaal richt zich onvoldoende op de aard en ernst van de schade en onrechtmatigheid van de strafbare handeling.[82] Het plan het artikel te verplaatsen naar een nieuwe titel over seksuele misdrijven juichen wij daarom toe.[83] Op deze wijze bieden de naam en categorisering van het strafbare feit meer inzicht in de onrechtmatigheid van de strafbare handeling. Dit doet onzes inziens meer recht aan het principe van *fair labelling*.

Ten aanzien van de wetgeving voor de strafbaarstelling van seksuele deepfakes doen wij de volgende aanbevelingen. De tekst van artikel 139h Sr gaat voorbij aan maatschappelijke en sociale normen over waar en wanneer gepast naaktheid en seksualiteit getoond kan worden. Daardoor kan een verdachte zich beroepen op 'niet weten' dat het maken en/of verspreiden nadelig zou zijn voor de afgebeelde persoon, terwijl de nadelige gevolgen van het 'naakt' of 'seksueel' tonen van personen in het openbaar juist als algemeen bekend mogen worden verondersteld.

Daarnaast is op dit moment niet duidelijk of de nadelige gevolgen door het slachtoffer bevestigd moeten worden, of dat die ook vastgesteld kunnen worden wanneer de

76 Zie art. 48 onder 2 Sr. Zie ook de strafbaarstelling in art. 240c Sr ten aanzien van personen die de leeftijd van zestien jaren nog niet hebben bereikt.

77 S.R. Bakker, Uitzonderlijke excepties in het strafrecht. Een zoektocht naar systematiek bij de beslissingen omtrent uitsluiting van strafrechtelijke aansprakelijkheid in bijzondere contexten (diss. Rotterdam), Boom juridisch 2021, p. 89-100.

78 Van den Boom 2021, p. 25.

79 EHRM 25 januari 2007, ECLI:NL:XX:2007:BA2629 (*Vereinigung Bildender Künstler t. Austria*).

80 EHRM 25 januari 2007, ECLI:NL:XX:2007:BA2629 (*Vereinigung Bildender Künstler t. Austria*), par. 32-4.

81 S. Royer en C. Conings, 'Catfishing, Cyberbullying, Deepfakes, Dickpics, Doxing, Grooming, Sextortion… Cyberfenomenen en hun strafrechtelijke kwalificaties', *IP- en ICT-recht* 2023/125, p. 81-153, p. 144, 147 en 150.

82 M.L.R. Goudsmit, 'Criminalising Image-Based Sexual Abuse: An Analysis of the Dutch Bill against "Revenge Pornography"', *AA* 2019, p. 442-447.

83 Op 4 juli 2023 stemde de Tweede Kamer in met een voorstel om art. 139h Sr te verplaatsen van Titel V 'Misdrijven tegen de openbare orde' van Boek II van het Wetboek van Strafrecht naar een nieuwe titel voor seksuele misdrijven naar aanleiding van de nieuwe Wet seksuele misdrijven (*Kamerstukken II* 2022/23, 36222, nr. 20).

doi: 10.5553/BSb/266669012023004005007

afgebeelde persoon niet weet dat er een seksuele deepfake van hem of haar bestaat, of wanneer de afgebeelde persoon niet opgespoord kan worden. Het is niet ondenkbaar dat seksuele deepfakes met onbekende slachtoffers niet onder de strafbaarstelling van artikel 139h Sr vallen, zelfs wanneer kan worden vastgesteld dat de verdachte geen toestemming had voor het maken en/of verspreiden van het seksuele beeldmateriaal.

Op basis van de in dit artikel gepresenteerde overwegingen, menen wij ons aan te kunnen sluiten bij de conclusie die Goudsmit Samaritter in haar proefschrift trekt: seksuele deepfakes zouden strafbaar gesteld kunnen worden als het 'opzettelijk maken en/of verspreiden van seksueel beeldmateriaal, met nalatigheid ten aanzien van de vraag of de afgebeelde persoon toestemming heeft gegeven voor het maken en/of verspreiden van de beelden'.[84]

Van den Boom stelt dat iedereen wier 'beeld digitaal is vastgelegd, slachtoffer kan worden'.[85]In een digitaal tijdperk, waarin vrijwel iedereen een online aanwezigheid heeft, valt het dan ook niet meer aan slachtoffers over te laten om zichzelf tegen seksuele deepfakes te beschermen. Daarin ligt een taak voor de staat, die onzes inziens nog vollediger – ook door middel van strafrecht – kan worden uitgevoerd.

84 Goudsmit 2022, p. 298.
85 Van den Boom 2021, p. 25.

doi: 10.5553/BSb/266669012023004005007