Rikosoikeudellinen vastuu deepfake-teknologian väärinkäytöstä

Lapin yliopisto
Oikeustieteiden tiedekunta
Oikeusinformatiikka
Maisteritutkielma
Maija Kunnas
Syksy 2022

Lapin yliopisto, oikeustieteiden tiedekunta
Työn nimi: Rikosoikeudellinen vastuu deepfake-teknologian väärinkäytöstä
Tekijä: Maija Kunnas
Opetuskokonaisuus ja oppiaine: Oikeustiede, oikeusinformatiikka
Työn laji: Maisteritutkielma
Sivumäärä: XIV+77
Vuosi: 2022

Tiivistelmä

Helposti saatavilla oleva deepfake- eli syväväärennösteknologia mahdollistaa tekoälyn avulla valheellisten tallenteiden luomisen. Aidon oloisen, mutta väärennetyn videon laadinnassa ja etenkin levittämisessä voi syyllistyä lukuisiin eri rikoksiin.

Tutkielmassa havainnollistetaan syväväärentämisen väärinkäytöstä aiheutuvia haasteita niin yksilön kuin yhteiskunnankin tasolla. Päätarkoituksena on tuoda esille väärinkäytön eri muotoja sekä osoittaa, kuinka rikoslain säännökset soveltuvat eri väärinkäyttötilanteisiin. Tutkielmassa pohditaan lisäksi, mitkä teot jäävät tällä hetkellä rangaistavuuden ulkopuolelle.

Suomen rikoslaki soveltuu varsin kattavasti syväväärennösten väärinkäyttöön. Jos video on tehty ilman siinä näkyvän henkilön lupaa, on todennäköistä, että jonkin rikoksen tunnusmerkistö täyttyy riippumatta videon sisällöstä. Teko voi olla identiteettivarkaus, kunnianloukkaus, uuden seksuaalirikoslainsäädännön mukainen seksuaalirikos tai esimerkiksi väärennys. Välillisesti syväväärennöstä voidaan käyttää esimerkiksi petoksessa tai kiristyksessä. Ongelmallisimpia ilmiöitä ovat disinformaation levittäminen syväväärennöksiä hyödyntäen sekä vaalivaikuttamisen eri muodot. Toistaiseksi näihin ei voida puuttua rikoslain puitteissa, ellei tekoon sovellu jokin muu rikoslain säännös, kuten kunnianloukkaus.

Suurimpia haasteita tuovat tekijöiden vaikea jäljitettävyys sekä se, ettei videoiden aitoutta voida todentaa luotettavasti. EU:n tulossa oleva tekoälyasetus pyrkii vastaamaan syväväärennösteknologian tuomiin haasteisiin, mutta käytännön toteutuksessa on omat haasteensa. Näitä ovat esimerkiksi teknologian käyttöön liittyvät, keskenään ristiriitaiset perusoikeudet.

Avainsanat: oikeusinformatiikka, rikosoikeus, deepfake, syväväärennös, tekoäly, disinformaatio

SISÄLLYSLUETTELO

LÄHTEET

Kirjallisuus

Aitamurto, Tanja, Syvät valevideot ovat uusin ase informaatiosodassa – ja niitä on nykyään huolestuttavan vaivaton tehdä. Helsingin Sanomat 13.11.2018. [https://www.hs.fi/mielipide/art-2000005897195.html] (10.10.2022)

Ajder, Henry – Patrini, Giorgio – Cavalli, Francesco – Cullen, Laurence, The State of Deepfakes: Landscape, Threats, and Impact, 2019. (Ajder ym. 2019)

Alén-Savikko, Anette – Vesala, Juha – Havu, Katri – Pihlajarinne, Taina & Koulu, Riikka, Uutisautomaatiota tutkimassa. Päivi Korpisaari (toim.) Data, viestintä ja sääntely. Viestintäoikeuden vuosikirja 2018. (Alén-Savikko ym. 2018)

Bressan, Sarah, Can the EU Prevent Deepfakes From Threatening Peace? Carnegie Europe 19.9.2019. [https://carnegieeurope.eu/strategiceurope/79877] (30.9.2022)

Brown, Nina I., Deepfakes and the Weaponization of Disinformation. Virginia Journal of Law & Technology 23(1) 2020, s. 1–59.

Burt, Tom – Horvitz, Eric, New steps to combat disinformation. Microsoft On the Issues 1.9.2020. [https://blogs.microsoft.com/on-the-issues/2020/09/01/disinformation-deepfakes-newsguard-video-authenticator/] (1.9.2022)

Ciancaglini, Vincenzo – Gibson, Craig – Sancho, David – McCarthy, Odhran – Eira, Maria – Amann, Philipp – Klayn, Aglika, Malicious Uses and Abuses of Artificial Intelligence, Trend Micrro Research, United Nations Interregional Crime and Justice Research Institute & Europol's European Cybercrime Centre, 19.11.2020. (Ciancaglini ym. 2020)

Citron, Danielle K. – Chesney, Robert, Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security. California Law Review 2019, s. 1753–1820.

Coats, Daniel R., Statement for the Record: Worldwide Threat Assessment of the US Intelligence Community. Senate Select Committee on Intelligence 29.1.2019. [https://www.dni.gov/files/ODNI/documents/2019-ATA-SFR---SSCI.pdf] (10.9.2022)

Dan, Viorela – Paris, Britt – Donovan, Joan – Hameleers, Michael – Roozenbeek, Jon – van der Linden, Sander, von Sikorski, Christian, Visual Mis- and Disindormation, Social Media, and Democracy. Journalism & Mass Communication Quarterly 98(3) 2021, s. 641–663. (Dan ym. 2021)

Desowitz, Bill, How Mark Hamill Was De-Aged as Luke Skywalker for 'The Mandalorian' Season 2 Finale – Exclusive. IndieWire 27.8.2021. [https://www.indiewire.com/2021/08/the-mandalorian-season-2-mark-hamill-interviews-de-aging-1234660159/] (15.9.2022)

Dietmar, Julia, GANs And Deepfakes Could Revolutionize The Fashion Industry. Forbes 21.3.2019. [https://www.forbes.com/sites/forbestechcouncil/2019/05/21/gans-and-deepfakes-could-revolutionize-the-fashion-industry/?sh=1451a5123d17] (11.10.2022)

Edwards, Cathy, Google blog: Search On 2022: Search and explore information in new ways. The Keyword 28.9.2022. [https://blog.google/products/search/search-on-2022-announcements/] (9.11.2022)

Forss, Marko – Keinänen, Anssi, Rikoslakia koskeva lainvalmistelu – Miten internet ja erityisesti sosiaalinen media huomioitiin vuosina 2009–2016 annetuissa hallituksen esityksissä. Edilex-sarja 2017/38, 21.9.2017. [https://www.edilex.fi/artikkelit/18068].

Frände, Dan – Matikkala, Jussi – Tapani, Jussi – Tolvanen, Matti – Viljanen, Pekka – Wahlberg, Markus, Keskeiset rikokset. Edita Publishing Oy 2018. (Frände ym. 2018)

Gagliani, Gabriele, Cybersecurity, Technological Neutrality, and International Trade Law. Journal of International Economic Law 23(3) 2020, s. 723–745.

Greene, Viveca S., "Deplorable" Satire: Alt-Right Memes, White Genocide Tweets, and Red-pilling Normies. Studies in American Humor 5(1) 2019, s. 31–69.

Guo, Hui – Wang, Xin – Lyu, Siwei, Detection of Real-time DeepFakes in Video Conferencing with Active Probing and Corneal Reflection. ArXiv preprint arXiv:2210.14153 2022. (Guo ym. 2022)

Huang, Yihao – Juefei-Xu, Felix – Wang, Run – Guo, Qing – Ma, Lei – Xie, Xiaofei – Li, Jianwen – Miao, Weikai – Liu, Yang – Pu, Geguang, Fake Polisher: Making DeepFakes More Detection-Evasive by Shallow Reconstruction, arXiv:2006.07533v3 [cs.CV], 17.8.2020. (Huang ym. 2020)

Kaisto, Janne, Lainoppi ja oikeusteoria: oikeusteorian perusteista aineellisen varallisuusoikeuden näkökulmasta. Edita 2005.

Kelleher, John D., Syväoppiminen – Kuinka tekoäly toimii. Suom. Kimmo Pietiläinen. Helsinki 2020.

Korkman, Julia, Pitävätkö muistosi paikkansa? Duodecim 136(24) 2020, s. 2713–2717.

Korpisaari, Päivi – Pitkänen, Olli – Warma-Lehtinen, Eija, Tietosuoja. 2., uudistettu painos. Alma Talent Oy 2022. (Korpisaari ym. 2022)

Kugler, Matthew B. – Pace, Carly L., Deepfake Privacy: Attitudes and Regulation. Northwestern University Law Review 116(3) 2021, s 1–72.

Lahti, Raimo, Rikosoikeuden ultima ratio -periaattesta ja hallintosanktioiden asemasta. s. 97–115. Rikoksesta rangaistukseen. Juhlajulkaisu Pekka Viljanen 1952–26/8–2012. Turun yliopiston oikeustieteellisen tiedekunnan julkaisuja. Tatu Hyttinen (toim.). Jyväskylä 2012.

Lalla, Vejay – Mitrani, Adine – Harned, Zach, Artificial intelligence: deepfakes in the entertainment industry. WIPO Magazine 6/2022. [https://www.wipo.int/wipo_magazine/en/2022/02/article_0003.html] (17.11.2022) (Lalla ym. 2022)

Langa, Jack, Deepfakes, Real Consequences: Crafting Legislation to Combat Threats Posed by Deepfakes. Boston University Law Review 101(2) 2021, s. 761–802.

Lappi-Seppälä, Tapio – Hakamies, Kaarlo – Helenius, Dan – Melander, Sakari – Nuotio, Kimmo – Ojala, Timo – Rautio, Ilkka, Rikosoikeus. Alma Talent Oy, päivittyvä teos (viimeisin päivitys helmikuussa 2022). (Lappi-Seppälä ym. 2022)

Lidz, Victor, Afterword: A Functional Analysis of the Crisis in American Society, 2020. The American Sociologist 52(1) 2021, s. 214–242.

Mak, Tim – Temple-Raston, Dina, Where are the Deepfakes in This Presidential Election? NPR 1.10.2020. [https://www.npr.org/2020/10/01/918223033/where-are-the-deepfakes-in-this-presidential-election?t=1647328536681] (4.2.2022)

Matikkala, Jussi, Tahallisuudesta rikosoikeudessa. Suomalainen lakimiesyhdistys 2012.

Meaker, Morgan,Clearview Stole My Face and the EU Can't Do Anything About It: One man's battle to reclaim his face shows regulators across the bloc are failing to reprimand the US face search engine. Wired 7.11.2022. [https://www.wired.co.uk/article/clearview-face-search-engine-gdpr] (9.11.2022)

Moreno, Johan, Google Is Evolving Search As Zoomers Use TikTok, Instagram To Find Things Online. Forbes 19.7.2022. [https://www.forbes.com/sites/johanmoreno/2022/07/19/google-is-evolving-search-as-zoomers-are-using-tiktok-instagram-to-find-things-online/] (9.11.2022)

Moshel, Michoel L. – Robinson, Amanda K. – Carlson, Thomas A. – Grootswagers, Tijl, Are you for real? Decoding realistic AI-generated faces from neural activity. Vision Research 199 2022, s. 1–12. (Moshel ym. 2022)

Mustak, Mekhail – Salminen, Joni – Mäntymäki, Matti – Rahman, Arafat – Dwivedi, Yogesh K., Deepfakes: Deceptions, mitigations, and opportunities. Journal of Business Research 2022. (Mustak ym. 2022)

Niemi, Hannele, Miten tekoäly muuttaa oppimista ja koulutusta? teoksessa Älykäs huominen: Miten tekoäly ja digitalisaatio muuttavat maailmaa? Toim. Gaudeamuksen työryhmä, Tallinna 2021.

Niemi, Valtteri: Yksityisyys tekoälyn aikakaudella teoksessa Älykäs huominen: Miten tekoäly ja digitalisaatio muuttavat maailmaa? Toim. Gaudeamuksen työryhmä, Tallinna 2021.

Pöysti, Tuomas, Tehokkuus, informaatio ja eurooppalainen oikeusalue. Helsinki 1999.

Pöysti, Tuomas, Julkisen vallan velvoite edistää sähköisen identiteetin ja verkkoyhteiskunnan infrastruktuurin turvallisuutta. Oikeus 1/2000, s. 91–112.

Pöysti, Tuomas, ICT and Legal Principles: Sources and Paradigm of International Law. Teoksessa Wahlgren Peter (toim.), IT-Law. Tukholma 2004, s. 560–598.

Reynolds, Matt, Courts and lawyers struggle with growing prevalence of deepfakes. Abajournal 9.6.2020. [https://www.abajournal.com/web/article/courts-and-lawyers-struggle-with-growing-prevalence-of-deepfakes]. (4.10.2022)

Rini, Regina, Deepfakes and the Epistemic Backstop. Philosophers' Imprint 20(24) 2020, s. 1–16.

Roth, Andrew, European MPs targeted by deepfake video calls imitating Russian opposition. The Guardian 22.4.2021. [https://www.theguardian.com/world/2021/apr/22/european-mps-targeted-by-deepfake-video-calls-imitating-russian-opposition] (28.2.2022)

Riekkinen, Juhana, Sähköiset todisteet rikosprosessissa. Alma Talent Oy 2019.

Rissanen, Kaisa – Sofy, Laura, Kostoporno – somerikos vai ei? Someturva 9.6.2020. [https://www.someturva.fi/blog/kostoporno-somerikos-vai-ei] (28.2.2022)

Rusanen, Anna-Mari: Algoritmien aakkoset teoksessa Älykäs huominen: Miten tekoäly ja digitalisaatio muuttavat maailmaa? Toim. Gaudeamuksen työryhmä, Tallinna 2021.

Saarenpää, Ahti, Yksityisyydensuoja tietämättömyyden yhteiskunnan uteliaisuusympäristössä. Tietosuoja 1/2004, s. 12–19.

Saarenpää, Ahti, Oikeusinformatiikka, s. 17–203 teoksessa Marja-Leena Niemi (toim.), Oikeus tänään/osa 1. 3. uudistettu painos. Bookwell Oy 2015.

Saaripuu, Tuire, Vahingonkorvausvastuun määräytyminen luonnollisen henkilön sähköisen identiteetin tunnistus- ja allekirjoituspalveluissa. Helsinki 2019.

Siltala, Raimo, Oikeustieteen tieteenteoria. Suomalainen lakimiesyhdistys 2003.

Simonite, Tom, A Zelensky Deepfake Was Quickly Defeated. The Next One Might Not Be. Wired 17.3.2022. [https://www.wired.com/story/zelensky-deepfake-facebook-twitter-playbook/] (11.4.2022)

Simonite, Tom, What Happened to the Deepfake Threat to the Election? Wired 16.11.2020. [https://www.wired.com/story/what-happened-deepfake-threat-election/] (4.2.2022)

Stupp, Catherine, Fraudsters Used AI to Mimic CEO's Voice in Unusual Cybercrime Case. The Wall Street Journal 30.8.2019. [https://www.wsj.com/articles/fraudsters-use-ai-to-mimic-ceos-voice-in-unusual-cybercrime-case-11567157402] (7.10.2022)

Tarkoma, Sasu, Miten tekoäly vaikuttaa kokonaisturvallisuuteen? Teoksessa Älykäs huominen: Miten tekoäly ja digitalisaatio muuttavat maailmaa? Toim. Gaudeamuksen työryhmä, Tallinna 2021.

Tuori, Kaarlo, Oikeusjärjestys ja oikeudelliset käytännöt. Helsingin yliopiston oikeustieteellisen tiedekunnan julkaisut 2013.

Tuori, Kaius, Lakimieslatinan käsikirja (toim. Kalle Kärkkäinen). Edita Publishing Oy 2007.

Vincent, James, 'Deepfake' that supposedly fooled European politicians was just a look-a-like, say pranksters. The Verge 30.4.2021. [https://www.theverge.com/2021/4/30/22407264/deepfake-european-polticians-leonid-volkov-vovan-lexus] (14.9.2022)

Voutilainen, Tomi, ICT-oikeus sähköisessä hallinnossa – ICT oikeudelliset periaatteet ja sähköinen hallintomenettely. Edita Publishing Oy 2009.

Westerlund, Mika, The Emergence of Deepfake Technology: A Review. Technology Innovation Management Review 9(11) 2019, s. 39–49.

Yamaoka-Enkerlin, Anna, Disrupting Disinformation: Deepfakes and the Law. New York University Journal of Legislation and Public Policy 22(3) 2020, s. 725–750.

Yoon, Leonard – Yang, Dongseok – Chung, Choongho – Lee, Sung-Hee, A Mixed Reality Telepresence System for Dissimilar Spaces Using Full-Body Avatar. SIGGRAPH Asia 2020 XR, Association for Computing Machinery 2020, s. 1–2. (Yoon ym. 2020)

Virallislähteet

HaVM 19/2021 – Valtioneuvoston selonteko sisäisestä turvallisuudesta.

HE 66/1988 vp, Hallituksen esitys eduskunnalle rikoslainsäädännön kokonaisuudistuksen ensimmäisen vaiheen käsittäväksi rikoslain ja eräiden muiden lakien muutoksiksi.

HE 94/1993 vp, Hallituksen esitys eduskunnalle rikoslainsäädännön kokonaisuudistuksen toisen vaiheen käsittäviksi rikoslain ja eräiden muiden lakien muutoksiksi.

HE 96/1998 vp, Hallituksen esitys Eduskunnalle henkilötietolaiksi ja eräiksi siihen liittyviksi laeiksi.

HE 169/2005 vp, Hallituksen esitys Eduskunnalle laiksi rikoslain muuttamisesta ja eräiksi siihen liittyviksi laeiksi.

HE 153/2006 vp, Hallituksen esitys eduskunnalle Euroopan neuvoston tietoverkkorikollisuutta koskevan yleissopimuksen hyväksymisestä, laiksi sen lainsäädännön alaan kuuluvien määräysten voimaansaattamisesta sekä laeiksi rikoslain, pakkokeinolain 4 luvun, esitutkintalain 27 ja 28 §:n ja kansainvälisestä oikeusavusta rikosasioissa annetun lain 15 ja 23 §:n muuttamisesta.

HE 36/2009 vp, Hallituksen esitys Eduskunnalle laiksi vahvasta sähköisestä tunnistamisesta ja sähköisistä allekirjoituksista sekä eräiksi siihen liittyviksi laeiksi.

HE 19/2013 vp, Hallituksen esitys eduskunnalle laeiksi rikoslain, pakkokeinolain 10 luvun 7 §:n ja poliisilain 5 luvun 9 §:n muuttamisesta.

HE 232/2014 vp, Hallituksen esitys eduskunnalle laiksi rikoslain eräiden tietoverkkorikoksia koskevien säännösten muuttamisesta ja eräiksi siihen liittyviksi laeiksi.

HE 13/2022 vp, Hallituksen esitys eduskunnalle seksuaalirikoksia koskevaksi lainsäädännöksi.

KOM (2020) 852, lopull. Ehdotus Euroopan parlamentin ja neuvoston asetukseksi digitaalisten palvelujen sisämarkkinoista (digipalvelusäädös) ja direktiivin 2000/31/EY muuttamisesta.

KOM (2021) 206, lopull. Ehdotus Euroopan parlamentin ja neuvoston asetukseksi tekoälyä koskevista yhdenmukaistetuista säännöistä (tekoälysäädös) ja tiettyjen unionin säädösten muuttamisesta.

KOM (2022) 105, lopull. Ehdotus Euroopan parlamentin ja neuvoston direktiiviksi naisiin kohdistuvan väkivallan ja lähisuhdeväkivallan torjumisesta.

LaVM 15/2005 vp, Lakivaliokunnan mietintö hallituksen esityksestä laiksi rikoslain muuttamisesta ja eräiksi siihen liittyviksi laeiksi.

OM 4/41/2013, Oikeusministeriön arviomuistio Identiteettivarkaus; henkilöllisyyden luomista koskevan hankkeen (identiteettiohjelma) työryhmän loppuraportin arviointia ja ehdotuksia jatkotoimiksi. Mikko Monto 15.3.2013.

PeVM 25/1994 vp, Perustuslakivaliokunnan mietintö n:o 25 hallituksen esityksestä perustuslakien perusoikeussäännösten muuttamisesta.

P9_TA(2022)0014, Euroopan parlamentin tarkistukset 20. tammikuuta 2022 ehdotukseen Euroopan parlamentin ja neuvoston asetukseksi digitaalisten palvelujen sisämarkkinoista (digipalvelusäädös) ja direktiivin 2000/31/EY muuttamisesta

Viranomaisten työryhmä- ja hankeaineisto

Euroopan komissio, Disinformation: Commission welcomes the new stronger and more comprehensive Code of Practice on disinformation, lehdistötiedote 16.6.2022.

Euroopan komissio, Code of Practice on Disinformation, 2018.

Euroopan komissio, Strengthened Code of Practice on Disinformation, 16.6.2022.

Henkilöllisyyden luomista koskeva hanke. Identiteettiohjelma. Työryhmän loppuraportti. Sisäinen turvallisuus. Sisäasiainministeriön julkaisu 32/2010. Sisäasiainministeriö 2010.

Kansallinen riskiarvio 2018. Sisäinen turvallisuus. Sisäministeriön julkaisuja 2019:5. Sisäministeriö 2019.

Lausunto 4/2007 henkilötietojen käsitteestä, WP 136. Tietosuojatyöryhmä 2007.

Seksuaalirikoslainsäädännön kokonaisuudistus. Lausuntotiivistelmä. Oikeusministeriön julkaisuja, Mietintöjä ja lausuntoja 2021:2. Oikeusministeriö 2021.

Verkkolähteet

Baranowicz, Tom, ”How to make DeepFake in 10 mins – Tutorial”. Youtube-video 12.8.2020. [https://www.youtube.com/watch?v=eq55Qy4RPiA] (10.3.2022)

DeepBrain AI. [https://deepbrainai.io/about-company] (17.11.2022)

Deepfake Detection Challenge Results: An open initiative to advance AI. Meta AI 12.6.2020. [https://ai.facebook.com/blog/deepfake-detection-challenge-results-an-open-initiative-to-advance-ai/] (17.2.2022)

European Data Protection Board. [https://edpb.europa.eu/news/national-news/2022/french-sa-fines-clearview-ai-eur-20-million_en] (9.11.2022)

FOC issues Joint Statement on Spread of Disinformation Online. Freedom Online Coalition 24.11.2020. [https://freedomonlinecoalition.com/foc-issues-joint-statement-on-spread-of-disinformation-online/] (7.3.2022)

Liikenne ja viestintävirasto Traficom. [https://www.kyberturvallisuuskeskus.fi/fi/toimintamme/saantely-ja-valvonta/sahkoinen-tunnistaminen] (25.10.2022)

Tackling deepfakes in European policy. European Parliamentary Research Service (EPRS) 30.7.2021. [https://www.europarl.europa.eu/at-your-service/en/stay-informed/research-and-analysis] (5.5.2022)

Vaalimainoksessa voi huijata ilman rangaistusta – oikeusoppinut: ”Ei sitä valvo oikein kukaan”. Yle 22.1.2022. [https://yle.fi/uutiset/3-12277841] (14.3.2022)

PimEyes: Face Search Engine Reverse Image Search. [https://pimeyes.com/en] (17.11.2022)

Almost Three-Quarters of UK Public Unaware of Deepfake Threat, New Research Reveals. iProov 1.10.2019. [https://www.iproov.com/press/uk-public-deepfake-threat] (17.11.2022)

Oikeuskäytäntö

C-131/12 Google Spain SL ja Google Inc. v. Agencia de Protección de Datos (AEPD) ja Mario Costeja González (13.5.2014)

Hannover v. Saksa (7.2.2012)

KHO 2018:112

Muut lähteet

Facing reality? Law enforcement and the challenge of deepfakes, an observatory report from the Europol Innovation Lab, Publications Office of the European Union, Luxembourg. Europol 2022. [https://www.europol.europa.eu/publications-events/publications/facing-reality-law-enforcement-and-challenge-of-deepfakes] (25.10.2022)

Internet Organised Crime Threat Assessment (IOCTA). Europol 9.10.2019. [https://www.europol.europa.eu/publications-events/main-reports/internet-organised-crime-threat-assessment-iocta-2021] (25.20.2022)

Väestön tieto- ja viestintätekniikan käyttö -tutkimus. Tilastokeskus 2020. [https://stat.fi/til/sutivi/2020/sutivi_2020_2020-11-10_tie_001_fi.html] (13.10.2022)

LYHENTEET

EIS	Euroopan ihmisoikeussopimus (SopS 63/1999)
EIT	Euroopan ihmisoikeustuomioistuin
EPRS	Euroopan parlamentin tutkimuspalvelu
EU	Euroopan unioni
GAN	generative adversarial network, suom. generatiivinen kilpaileva verkosto
HaVM	hallintovaliokunnan mietintö
HE	hallituksen esitys
HS	Helsingin Sanomat
KHO	korkein hallinto-oikeus
KOM	komission ehdotus
KP-sopimus	kansalaisoikeuksia ja poliittisia oikeuksia koskeva kansainvälinen yleissopimus (SopS 7–8/1976)
LaVM	lakivaliokunnan mietintö
OK	oikeudenkäymiskaari (4/1734)
PeVL	perustuslakivaliokunnan mietintö
PL	Suomen perustuslaki (731/1999)
RL	rikoslaki (39/1889)
TSA	yleinen tietosuoja-asetus (EU) 2016/679
TTS	Text-To-Speech, suom. tekstistä puheeksi -synteesi

TAULUKOT

1. JOHDANTO

1.1. Tutkimuksen taustaa

Siirtyminen informaatioyhteiskunnasta verkkoyhteiskuntaan on ollut matka tietotekniikkaa hyödyntävästä ajasta täysin digitaaliseen aikakauteen sekä toimintaympäristöön. Tietotekniikkaa ei voida pitää enää omana ja erityisenä, muusta maailmasta erillisenä ilmiönä. Tämä kehitys vaikuttaa tavallisten ihmisten arkeen niin työssä kuin vapaa-ajallakin, yrityksiin sekä myös verkkoyhteiskuntaa ympäröivään oikeudelliseen viitekehykseen. Kehittyvä teknologia lisää tarvetta luoda uutta oikeudellista sääntelyä, mikä aiheuttaa oikeudellisten käsitteiden määrän kasvua sekä soveltuvien periaatteiden muutosta.[1] Oikeudellisen verkkoyhteiskunnan tunnusmerkkejä ovat muun muassa verkossa vaikuttavat ja hyödynnettävät perusoikeudet, informaation luonteen ja informaation käsittelyn muuttuminen, verkkoviestintä eri muotoineen, tieto(verkko)rikokset sekä tietoturvallisuuden korostuminen oikeuksien toteutumisessa.[2]

2010-luvun loppupuolella kehitettiin jälleen uusi teknologian muoto, deepfake-älyteknologia, jonka avulla voidaan väärentää ääntä, kuvia sekä videokuvaa. Deepfake-videoille on ominaista, että niissä näkyvät ja kuuluvat ihmiset sanovat tai tekevät jotain, mitä he todellisuudessa eivät ole sanoneet tai tehneet.[3] Suomessa deepfake-termin lisäksi kyseisellä tekniikalla tehdystä materiaalista käytetään myös nimeä syväväärennös.[4] Deepfake-videoiden suosio ja syväväärennösteknologian kuluttajakäyttö alkoivat vuonna 2017 Reddit-nimisellä verkkosivustolla. Tuolloin Deepfake-niminen käyttäjä julkaisi sivustolle pornografisia videoita, joissa esiintyneille ihmisille tämä oli syväväärentämällä vaihtanut julkisuuden henkilöiden kasvoja. Julkaisut saivat paljon huomiota ja ne keräsivät lyhyessä ajassa paljon katselukertoja.[5] Muita jo havaittuja syväväärennöksillä aiheutettuja ongelmia ovat olleet esimerkiksi valtioiden välinen informaatio- ja

[1] Saarenpää 2015 s. 23–29.
[2] Saarenpää 2015 s. 55–56.
[3] Esim. Citron – Chesney 2019 s. 1753.
[4] Syväväärennöksestä on myös käytetty termejä kuten *syvä feikki* ja *syvä valevideo*. Ks. esim. John D. Kelleher ”Syväoppiminen – Kuinka tekoäly toimii” (2019) suomenkielinen versio sekä ”Syvät valevideot ovat uusin ase informaatiosodassa – ja niitä on nykyään huolestuttavan vaivaton tehdä” (HS 13.11.2018). Euroopan unionin oikeudessa viralliseksi käsitteeksi on vakiintunut *syvävääreennös*. Tässä tutkielmassa käytän termejä *deepfake* ja *syvävääreennös* kuvaamaan tekoälyn avulla väärennettyä videota sekä lisäksi sanaa *syvävääreentäminen* kuvaamaan syväväärennösten tekemistä.
[5] Kugler – Pace 2021 s. 10.

vaalivaikuttaminen sekä muu disinformaation levittäminen.[6] Tällaista informaatiovaikuttamista on havaittu hiljattain jopa sodankäynnissä. Venäjän helmikuussa 2022 Ukrainaan aloittamaan hyökkäyssotaan liittyen venäläiset hakkeroivat ukrainalaisen uutissivuston. Sivustolle ladattiin valheellinen video presidentti Volodymyr Zelenskyista, jossa Zelenskyi käski omia joukkojaan antautumaan.[7] Samoin ennen Venäjän hyökkäystä Ukrainaan helmikuussa 2022 Yhdysvallat paljasti venäläisten aikeet käyttää deepfake-teknologialla tuotettua materiaalia hyökkäyksen perusteluna.[8]

Uudenlaiset teknologiamuodot tuovat mukanaan uusia mahdollisuuksia. Syväväärennöksiä voidaan hyödyntää esimerkiksi opetuksessa, kaupallisessa tarkoituksessa, yhteiskunnallisissa kannanotoissa ja yksilön itseilmaisussa. Älyteknologiakehityksen mukanaan tuomien uudenlaisten käyttötarkoitusten lisäksi kehitys lisää yhtä lailla uusia haasteita, joihin yhteiskunnassa on pyrittävä vastaamaan. Syväväärennösten väärinkäytöstä aiheutuneet seuraukset voivat olla vakaviakin. Haasteita oikeudelliseen näkökulmaan tuo syväväärennösten moniulotteisuus. On osattava huomioida kaikki deepfake-videoiden elinkaareen liittyvät toimijat ja niiden oikeudet sekä velvollisuudet. Syväväärennöksiin liittyviä henkilöitä tai oikeushenkilöitä ovat deepfake-videon tekijät, videolla esiintyvät henkilöt (ml. uhrit sekä alkuperäisellä videolla esiintyvät henkilöt), alkuperäisen videon tekijät tai sen immateriaalioikeuksien haltijat, syväväärennösteknologian kehittäjät, alustat, joita käytetään videon levittämiseen sekä alustojen käyttäjät, jotka lataavat, katsovat tai jakavat videota.[9]

Väärinkäytön uhreja ovat luonnolliset henkilöt, joskus yritykset ja kumuloituvien yhteisvaikutusten seurauksena myös kokonaiset yhteiskunnat. Esimerkiksi syväväärennösten avulla helposti toteutettavissa oleva disinformaation levittäminen on teko, jolla on todettu olevan mahdollisesti jopa yhteys ihmisoikeuksien, demokratian ja oikeusvaltion heikkenemiseen.[10] Väärinkäyttöön syyllistyvät voivat olla joko yksittäisiä henkilöitä tai valtiollisia toimijoita.

[6] Mm. Citron – Chesney 2019 s. 1772–1779.
[7] Simonite 2022.
[8] Europol 2022 s. 5–6.
[9] EPRS 2021 s. 37–38.
[10] Freedom Online Coalition 2020.

Kansainvälinen verkkoympäristö tuo omat haasteensa yhteisten pelisääntöjen luomiselle sekä toiminnan valvonnalle.

Euroopan parlamentin tutkimuspalvelu (EPRS)[11] julkaisi heinäkuussa 2021 *Tackling deepfakes in European policy* -nimisen raportin syväväärennöksistä sekä niitä koskevista teknisistä, yhteiskunnallisista sekä lainsäädännöllisistä näkökohdista. Raportin mukaan syväväärennöksiin liittyvien riskien arviointi osoittaa riskien voivan olla niin psykologisia, taloudellisia kuin yhteiskunnallisiakin. Raportissa arvioidaan, että syväväärennösten käyttö EU-alueella tulee olemaan jopa arkipäivää sosiaalisessa mediassa viiden vuoden kuluessa. Nopeasti tapahtuvan syväväärennösteknologian käytön lisääntymisen oletetaan lisäävän samalla myös sen väärinkäyttöä.[12]

Nopean teknologiakehityksen aikaansaamaa tarvetta uudistuvalle sääntelylle on yleisesti pyritty jarruttamaan lainsäädännön teknologianeutraalisuudella. Lainsäädännön yhteydessä teknologianeutraalisuudella pääsääntöisesti viitataan ajatukseen siitä, että lainsäädännön tulisi kestää aikaa ja olla sovellettavissa sellaiseen teknologiaan, jota vielä lainsäädäntöhetkellä ei välttämättä ole kehitetty.[13] Euroopan unionin tasolla ja kansainvälisesti teknologianeutraali lainsäädäntö ei saisi syrjiä tai suosia mitään tiettyä teknologiaa, vaan säännöksiä tulisi voida soveltaa kaikkiin teknologioihin erotuksetta.[14] Rikosoikeudellisesta näkökulmasta haasteita teknologianeutraalisuudelle tuo laillisuusperiaate, jonka nojalla säännöksen tulisi olla myös täsmällinen ja siten ennakoitavissa. EPRS:n raportissa pohditaan Euroopan unionin vaikutusmahdollisuuksia syväväärennösten haittavaikutusten minimoimiseksi. Tekstissä otetaan myös kantaa Euroopan komission esittelemän uuden tekoälyasetuksen[15] merkitykseen syväväärennösten haittavaikutusten minimoimisessa. Raportti osoittaa, ettei syväväärennöksiä koskevaa

[11] European Parliamentary Research Service. EPRS avustaa Euroopan parlamentin jäseniä tuottamalla riippumattomia, puolueettomia ja luotettavia tutkimuksia sekä analyysejä EU:n toimintaan liittyvistä asioista.
[12] EPRS 2021.
[13] Forss – Keinänen 2017 s. 11.
[14] Gagliani 2020 s. 723.
[15] Ehdotus Euroopan parlamentin ja neuvoston asetukseksi tekoälyä koskevista yhdenmukaistetuista säännöistä ja tiettyjen unionin säädösten muuttamisesta. Komissio on julkaissut ehdotuksensa 21. huhtikuuta 2021. Asetuksen on alun perin ollut tarkoitus tulla voimaan 2022 loppupuolella.

sääntelymahdollisuutta ole hyödynnetty parhaalla mahdollisella tavalla ja että erityisesti rikosoikeudellinen näkökulma on pulmallinen niin väärennöksen tekijän kuin rikoksen uhrin näkökulmasta.[16]

Tutkielmassa käydään läpi syväväärennösteknologiaan soveltuvaa lainsäädäntöä sekä sen merkitystä syväväärennösten väärinkäytön ehkäisyssä. Tällä pyritään osoittamaan EU-oikeuden riittämättömyys syväväärentämisen väärinkäyttöä koskevissa oikeudellisissa kysymyksissä sekä korostamaan kansallisen lainsäädännön merkitystä näiden puutteiden täydentämisessä, nimenomaan rikoslainsäädännön osalta.

1.2. Tutkimuskysymykset sekä aiheen rajaus

Tutkielmassa käsiteltävät kysymykset voidaan tiivistää yhdeksi pääkysymykseksi seuraavalla tavalla: Kuinka olemassa oleva Suomen rikoslainsäädäntö soveltuu syväväärennösteknologian väärinkäyttöön ja onko lainsäädäntö riittävää kyseisen teknologian käytöstä aiheutuvien haittojen minimoimiseksi? Vastaus pääkysymykseen pyritään löytämään pienempien apukysymysten avulla. Apukysymykset ovat seuraavat: 1) Mitä syväväärennökset ovat ja mitkä ovat niiden hyödyt ja haitat? 2) Millaisia vaikutuksia syväväärennöksillä on yksilön ja yhteiskunnan tasolla? 3) Mitä rikoslain säännöksiä voidaan soveltaa syväväärennösten väärinkäyttöön? 4) Onko olemassa oleva rikoslainsäädäntö riittävää syväväärennöksistä aiheutuvien haittojen minimoimiseksi ja miten perusoikeusnäkökulma ja uusiutuva EU-sääntely vaikuttavat aihetta koskeviin sääntelytarpeisiin ja -mahdollisuuksiin? 5) Mitä vaihtoehtoisia näkökulmia on hyvä ottaa huomioon väärinkäytön ongelmien ratkaisussa?

Koska syväväärennösteknologia on varsin uutta, ei sen käytön ja soveltuvan lainsäädännön yhteensovittamiseen ole olemassa pitkäaikaisia tai koeteltuja ratkaisuja. Tavoite on erilaisten skenaarioiden avulla esittää deepfake-teknologian mukanaan tuomia todennäköisiä väärinkäytöstilanteita erityisesti rikosoikeudellisesta näkökulmasta. Tämä tarkoittaa rikosoikeuden systematisointia nykytilanteen valossa.

[16] EPRS 2021 s. 48–50.

Koska tutkimuksen pääpaino on deepfake-teknologian väärinkäytössä ja siihen liittyvässä lainsäädännössä, saattaa herätä kysymys, miksi niin paljon ongelmia aiheuttavaa teknologiaa ja sen käyttöä ei kielletä kokonaan. Yksi tärkeimmistä syistä syväväärentämisen sallimiselle on sanan- ja ilmaisuvapauden turvaaminen. Syväväärennösteknologian käytön kielto rajoittaisi oikeutta ilmaista, julkistaa ja vastaanottaa tietoa. Toisaalta sananvapautta jarruttava toinen perusoikeus on yksityisyyden suoja. Nämä perusoikeudet ja niiden aiheuttamat ristiriidat ovat niin merkittäviä deepfake-teknologiaa koskevassa oikeudellisessa keskustelussa, että niitä pohditaan tutkielman lopussa.

Syväväärennöksiin soveltuu myös paljon muuta lainsäädäntöä, joka jää tämän tutkielman ulkopuolelle aiheen laajuuden rajaamiseksi. Lyhyesti todettakoon, että syväväärennöksissä käytettävien kuvien osalta esimerkiksi esiin voi tulla erilaiset tekijänoikeuskysymykset. Julkaistut syväväärennösvideot voivat myös aiheuttaa julkaisussa väärennetysti esiintyvälle henkilölle esimerkiksi mainehaittaa ja mahdollisia varallisuusvahinkoja. Tällöin tekoälyavusteisen julkaisun erityiskysymysten lisäksi tulee pohdittavaksi yleiset julkaistua viestiä koskevat yksityisoikeudelliset vastuukysymykset. Julkaistuun viestiin kohdistuva vahingonkorvausvastuu pohjautuu Suomessa sananvapauteen liittyvään sääntelyyn sekä vahingonkorvauslakiin (412/1974, VahL).[17] Niin ikään aihetta koskevat erilaiset vastuukysymykset, esimerkiksi syväväärennösteknologian kehittäjän, teknologian tarjoajan sekä verkkosivustojen ja -alustojen näkökulmasta.

Vaikka paljon soveltuvaa lainsäädäntöä jää tutkimuksen ulkopuolelle, on aihe tavanomaista maisteritutkielmaa laajempi. Perustelen aiheen valintaa sillä, että se on tarkoitettu muillekin kuin suppealle, oikeusoppineiden oikeusyhteisölle. Tutkielma on tarkoitettu pikemminkin laajalle oikeusyhteisölle, joka Kaarlo Tuorin mukaan kattaa kaikki, jotka osallistuvat oikeudelliseen keskusteluun ja joilla on performatiivinen asenne yhteiskunnan toimintaa ohjaavaan oikeusjärjestykseen.[18] Laajaan oikeusyhteisöön kuuluvat siten esimerkiksi poliitikot, jotka ovat keskeisessä osassa lakien säätämisessä. Yhtä lailla toivon, että tutkielmasta on hyötyä syväväärennösteknologian käyttäjille, teknologiaa kehittäville tahoille, väärinkäytön uhreille sekä

[17] Havu 2019 s. 24–25.
[18] Tuori 2013 s. 15–17.

muille, jotka tulevat jollakin tavalla linkittymään syvävääreennöksiin, esimerkiksi disinformaation kohteena olemisen kautta.

On tarpeellista tuoda esiin deepfake-teknologian haittavaikutuksia sekä sen käytöstä seuraavia oikeudellisia ongelmia, joita videoiden tekijät eivät välttämättä tule ajatelleeksi teknologiaa käyttäessään. Vanhan oikeudellisen periaatteen mukaisesti kansalaisella on velvollisuus tuntea laki. Lain tunteminen voi olla hankalaa, jos teknologian käyttöön liittyy paljon erilaista, paikoin jopa yllättävää, lainsäädäntöä. Tutkielman yhtenä tarkoituksena on nostaa esille kuluttajan mahdollisesti harmittomana pitämän ”hupiteknologian” käytöstä potentiaalisesti aiheutuvia seurauksia. Yhtä tärkeää on kasvattaa tietoisuutta syvävääreennösten olemassaolosta ylipäätään, jotta ihmiset osaisivat olla mediakriittisiä myös videoiden ja äänitteiden osalta. Syvävääreentämisen ympärillä olevan lainsäädäntökehikon tunteminen on tarpeen, sillä kuten Suomen rikoslain (39/1889, RL) 3:4.2:ssä todetaan, ei lain tuntemuksen puute pääsääntöisesti vapauta tekijää vastuusta; *ignorantia juris non excusat*.[19]

1.3. Tutkimuksen sijoittuminen oikeudenalajaottelussa

Teknologiaan sovellettava oikeudellinen tutkimus kuuluu yleisesti oikeusinformatiikan oikeudenalaan. Siten myös tässä syvävääreennösteknologiaan perehtyvässä tutkielmassa käytetään pääsääntöisesti oikeusinformatiikan alaan kuuluvaa lähdeaineistoa. Rikosoikeudellinen lähestymistapa vaatii rikosoikeudellisten lähteiden käyttöä. Niin ikään tutkielmassa hyödynnetään valtiosääntöoikeuden alle kuuluvaa perusoikeusliitännäistä lähdeaineistoa siltä osin kuin tutkielma käsittelee syvävääreentämiseen liittyviä perusoikeuksia, lähinnä sananvapauden ja yksityisyyden suojan näkökulmista. Suomessa rikosoikeusteoreettisessa keskustelussa on 1990-luvulta lähtien tukeuduttu vahvasti perus- ja ihmisoikeusajatteluun. Tämä on luonnollinen seuraus Suomen liittymisestä Euroopan neuvoston ihmisoikeussopimukseen vuonna 1990 sekä Suomen perusoikeusuudistuksesta vuonna 1995. Perus- ja ihmisoikeuslähtöinen ajattelutapa on rajoittanut siten myös rikosoikeuden käyttöä.[20]

[19] Saarenpää 2015 s. 33.
[20] Lahti 2012 s. 100.

Koen syvävääräntämistä koskevassa oikeudellisessa katsauksessa perusoikeusnäkökulman tärkeänä ensinnäkin siksi, että sillä voidaan perustella syvävääräennösteknologian käytön sallimista – ainakin sananvapauden osalta. Perusoikeuksien rajoittaminen on sallittua vain, jos tavoitetta ei voida saavuttaa kyseiseen perusoikeuteen vähemmän puuttuvilla keinoilla.[21] Tämä on tärkeää muistaa silloin, kun pohditaan, onko Suomessa syvävääräennösteknologian väärinkäytöstä aiheutuva rangaistusuhka tarpeeksi kova, kun otetaan huomioon rangaistusuhan preventiivinen ja retributiivinen tehtävä. Toisaalta perusoikeusnäkökulmaan sisältyy luonnollisesti myös katsaus henkilön yksityisyyden suojaan ja näiden kahden perusoikeuden tasapainoon.

1.4. Metodologia ja lähdeaineisto

Tutkimus on luonteeltaan oikeusdogmaattinen eli lainopillinen, sisältäen lain tulkintaa sekä systematisointia. Metodilla pyritään tyydyttämään käytännön tarpeita tunnistamalla ensin niitä lainsäädännöllisiä heikkouksia ja puutteita, jotka kehittyvä syvävääräennösteknologia aiheuttaa. Lainoppi voidaan jakaa teoreettiseen ja käytännölliseen lainoppiin.[22] Tutkimukseni on tässä jaottelussa enemmän käytännöllinen eli tulkinnallinen, sillä tarkoitus on esittää tulkintakannanottoja syvävääräntämisen ja lainsäädännön välisestä suhteesta. Koska tutkimus on lainopillinen, siinä esitetyn tiedon on teoreettisen päämäärän lisäksi tarkoitus soveltua myös käytäntöön. Systematisoinnilla viittaan soveltamistilannesysteemiin[23], jossa erittelen mahdollisia syvävääräntämisestä aiheutuvia ongelmatilanteita. Tämän jälkeen on selvitettävä, miten olemassa oleva oikeusjärjestyksemme sisältö on sovellettavissa näihin ongelmatilanteisiin.

Vaikka syvävääräennösten haittoihin on jo reagoitu esimerkiksi Euroopan komission ehdotuksessa uudeksi tekoälyasetukseksi, uuden asian äärellä käsitteet eivät ole vakiintuneita eikä kyseinen teknologia ole löytänyt vielä paikkaansa oikeudenalajaotuksessa. Tämän vuoksi tutkimus ei ole puhtaasti lainopillinen vaan siinä on myös pyritty huomioimaan muuttuvaa teknologiaympäristöä sekä hakemaan ratkaisuja maantieteellisesti muualta. Tutkielmassa on siksi myös nähtävissä oikeusvertailevia piirteitä.

---

[21] Lahti 2012 s. 100.
[22] Siltala 2003 s. 109–110; Kaisto 2005 s. 17.
[23] Lainopillisesta systematisoinnista ja soveltamistilannesysteemistä ks. Kaisto 2005 s. 343.

Tutkielman lähdeaineisto pohjautuu voimassa olevaan lainsäädäntöön Suomessa ja Euroopan unionissa, lakien esitöihin, artikkeleihin sekä kirjallisuuteen. Suurin osa käytettävissä olleista syväväärennöksiä koskevista tieteellisistä artikkeleista on ulkomaalaisia, pääasiassa yhdysvaltalaisia. Tämä johtuu siitä, että Yhdysvaltoja lukuun ottamatta aiheesta kirjoitettua oikeustieteellistä tekstiä on toistaiseksi olemassa hyvin vähän. Suomalainen oikeustieteellinen tutkimus syväväärennöksistä on lähes olematonta muutamaa tutkielmaa lukuun ottamatta.

Tutkielma jakautuu viiteen osaan. Ensimmäisessä osassa esitellään syväväärennösteknologiaa sekä lyhyesti sitä, miten kyseinen teknologia toimii. Samassa osassa perehdytään syväväärennöksiin ilmiönä sekä pohditaan, miten syväväärennökset eroavat perinteisistä väärennyksistä, kuten väärennetyistä kuvista tai videoista, joiden luomisessa ei ole käytetty tekoälyä. Ensimmäiseen osaan sisältyy myös lyhyt katsanto syväväärennösten hyötyihin ja riskeihin. Riskien jälkeen käydään läpi syväväärennösten haittavaikutuksiin olennaisesti liittyvää väärennösten havainnointia.

Toisessa osassa käsitellään syväväärennösteknologiaan liittyviä oikeudellisia ongelmia ja niiden väärinkäyttöä niin yksilön kuin yhteiskunnan tasolla.

Kolmas osa käsittelee syväväärennöksiin soveltuvaa rikoslainsäädäntöä. Se käsittää siten ne rikosoikeudelliset säännökset, jotka sellaisenaan ovat sovellettavissa syväväärennösteknologian väärinkäyttöön Suomessa. Näiltä osin tarkastelun kohteena on siis pääsääntöisesti voimassa oleva lainsäädäntö *de lege lata*. Osan lopussa pohditaan erikseen rikoslainsäädännön soveltumista vaalivaikuttamiseen ja disinformaatioon.

Neljäs ja viides osa koostuvat käsiteltyjen aiheiden pohdinnasta sekä johtopäätöksistä. Pohdinnassa käydään läpi myös syväväärennöksiä koskevaa problematiikkaa perusoikeusnäkökulmasta. Syväväärennösten taustalla oleva ristiriitainen perusoikeusverkko tuo haasteita myös uuden lainsäädännön kehittämiselle. Pohdinnassa huomioidaan myös uusi tulossa oleva EU:n tekoälyasetus sekä sen hyödyntämismahdollisuudet syväväärennösten väärinkäytön ehkäisemisessä. Tämä osa kokoaa yhteen tutkimuksessa tehdyt huomiot ja siinä otetaan kantaa

syvävääreennöksiä koskeviin mahdollisiin muutostarpeisiin olemassa olevan lainsäädännön osalta *de lege ferenda*.

2. DEEPFAKE-TEKNOLOGIA

2.1. Deepfake eli syväväärennös

Sana deepfake koostuu termeistä ”deep learning” eli ”syväoppiminen” sekä ”fake” eli ”väärennös”. Deepfaket eli syvävääreennökset ovat realistisia videoita, joita on tekoälyn avulla manipuloitu kuvaamaan ihmisiä sanomassa ja tekemässä asioita, joita ei todellisuudessa ole koskaan tapahtunut. Deepfake voi olla myös pelkkä manipuloitu kuva tai äänite. Syvävääreennösten luomiseen käytetään eräänlaista neuroverkkoarkkitehtuuria, jota kutsutaan nimellä ”Generative adversarial network” tai ”GAN” eli generatiivinen kilpaileva verkosto tai generatiivinen adversiaalinen verkko. Tässä arkkitehtuurissa kaksi neuroverkkoa kilpailee toisiaan vastaan analysoiden samalla suuria määriä tietoa oppiakseen imitoimaan henkilön ilmeitä, tapoja sekä ääntä.[24]

Prosessissa deepfake-algoritmiin syötetään tietoa kahdesta henkilöstä, jotta se oppii vaihtamaan kasvoja näiden kahden henkilön välillä. Syvävääreennetyissä videoissa siis käytetään kasvojen kartoitustekniikkaa ja tekoälyä, joka muuttaa videolla olevan henkilön kasvot toisen henkilön kasvoiksi.[25] Vaihtoehtoisesti syvävääreennösten luominen voi tapahtua niin, että autoenkooderit purkavat kuvia osiin niistä löytämiensä tietojen avulla ja rakentavat näistä osista uusia kuvia ja videoita luoden tällä tavalla erilaisia ilmeitä.[26] Tällöin henkilön kasvot pysyvät samana ja vain ilmeet muuttuvat.

Syvävääreentäminen on osa suurempaa, syväoppimisen (*deep learning*), kokonaisuutta. Jotta syväoppiminen voi toimia, tarvitsee sen ohjelmisto valtavan määrän dataa oppimisensa perustaksi. Kyseinen data voi olla esimerkiksi puhetta, ilmeitä ja eleitä, kuvia, tekstiä tai ihmisestä saatua fysiologista tietoa. Kerätyn data-aineiston lisäksi syväoppiva kone oppii myös sen käyttäjän kanssa käymän vuorovaikutuksen kautta.[27] Tässä tutkielmassa keskitytään vain sellaisiin syvävääreennöksiin, joissa on ihminen. Kyseistä teknologiaa voidaan kuitenkin käyttää myös videoihin, joissa subjekti on jokin muu kuin ihminen, esimerkiksi eläin.[28]

[24] Kelleher 2019 s. 207–208.
[25] Westerlund 2019 s. 38.
[26] EPRS 2021 s. II.
[27] H. Niemi 2021 s. 136.
[28] EPRS 2021 s. 2.

Toistaiseksi useimmat tietoverkosta löytyvät syvävääreennökset keskittyvät julkisuuden henkilöihin, kuten näyttelijöihin, poliitikkoihin tai yritysjohtajiin, joita muutenkin pidetään potentiaalisina deepfake-videoiden uhreina.[29] Tämä johtunee siitä, että yleensä julkisuudessa esiintyvistä henkilöistä on saatavilla paljon digitaalista kuvamateriaalia. Tämä ei kuitenkaan tarkoita sitä, etteikö tavallinen kansalainen voisi joutua syvävääreennöksen kohteeksi.

Tilastokeskuksen tekemän tutkimuksen mukaan vuonna 2020 82 prosenttia 16–89-vuotiaista suomalaisista käytti tietoverkkoa useasti päivässä.[30] Covid 19-pandemian seurauksena myös erilaiset videoneuvottelut yleistyivät huomattavasti. Videopalveluissa on käytössä erilaisia kuvan manipulointityökaluja, kuten virtuaalisia taustoja, ns. filttereitä eli kuvasuodattimia sekä muita keinoja videoiden muokkaukseen. Tällainen kehitys edesauttaa valheellisen maailman sekä todellisuuden sekoittumisen normalisoitumista.[31] Lisäksi erilaisiin sosiaalisen median verkostoihin kuuluminen on tärkeä osa monien ihmisten identiteettiä. Etenkin henkilökohtaisten kuvien lataaminen sosiaalisen median sivustoille on aiheuttanut sen, että tavallisista ihmisistä on saatavilla yhä enemmän visuaalista dataa. Sovellusten tarjoamat mahdollisuudet jaettujen tietojen kontrolloimiseksi vaihtelevat. Joillakin sivustoilla tietojen poistaminen ja jaettavan tiedon vastaanottajien määrittäminen on luotu helpommaksi kuin toisilla.[32]

Henkilö, josta on ladattu paljon kasvokuvia sosiaalisen median sivustoille, on todennäköisesti suuremmassa vaarassa ajautua syvävääreennetylle videolle kuin henkilö, josta kuvia on saatavilla vähemmän tai ei ollenkaan. Mitä enemmän kuvamateriaalia voidaan syöttää deepfake-teknologiaa hyödyntävään sovellukseen, sitä aidompia videoita on mahdollista tehdä. Sanotaankin, että data on tekoälyn ruokaa. Erittäin aidolta vaikuttavaan videoon tarvitaan eri lähteiden mukaan yhdestä moneen sataan kuvaa.[33] Toisaalta lienee todennäköistä, että mitä pidemmälle syvävääreennösteknologiaa kehitetään, sitä vähemmän henkilön kuvia tarvitaan aidolta vaikuttavan videon luomiseksi.

[29] Ks. Esim. Yamaoka-Enkerlin 2020 s. 731–732 ja Brown 2020 s. 15.
[30] Väestön tieto- ja viestintätekniikan käyttö -tutkimus 2020, Tilastokeskus.
[31] Yoon ym. 2020 s. 1–2.
[32] Korpisaari ym. 2022 s. 19.
[33] Eri arvioita ovat esittäneet muun muassa Westerlund 2019 s. 41 sekä Citron – Chesney 2019 s. 1773.

Syvävääreennöksiä voi tehdä myös ilman videokuvaa, sillä tietokoneilla voidaan mallintaa ihmispuhetta keinotekoistesti äänenkloonausteknologian avulla. Myös tällaista puhesynteesin mahdollistavaa teknologiaa on paljolti ja helposti saatavilla. Kirjoitetun tekstin muuttaminen ääneksi (tekstistä puheeksi -synteesi eli Text-To-Speech, TTS) on tekniikka, jota käyttävät myös tutut, arkipäiväiset ääniavustajalaitteet, kuten Applen Siri sekä Amazonin Alexa. Samaa tekniikkaa käyttämällä voidaan väärentää kenen tahansa ihmisen ääntä. Tuolloin hyödynnetään sen henkilön ääninäytettä, jonka puhetta halutaan imitoida. Jo muutaman minuutin pituinen äänite on riittävä laadukkaan ääniväärennöksen tekemiseksi.[34]

Deepfake-teknologiaa voidaan käyttää myös reaaliaikaisesti, eikä syvävääreentäminen siten vaadi siis jälkikäteistä videon muokkausta. Reaaliaikaista videonmuokkausta tekoälyn avulla tarjoavat esimerkiksi Avartarify ja DeepFaceLive. Toistaiseksi reaaliaikaiset syvävääreennökset eivät ole kovin uskottavia, minkä vuoksi niiden käytöllä aikaansaatavat riskit eivät ole suuret. Tilanne saattaa kuitenkin muuttua nopean teknologiakehityksen seurauksena. Videoiden reaaliaikaisuus tekee informaatiovaikuttamisesta ja muusta väärinkäytöstä tehokkaampaa. Syvävääreennösten reaaliaikaisuus on ongelmallista myös väärennösten havainnoinnin näkökulmasta.[35]

2.2. Ero perinteiseen väärentämiseen

Kuvan ja äänen muokkaaminen ei ole uusi ilmiö. Jo pitkään esimerkiksi Adobe Photoshopin kaltaisilla kuvankäsittelyohjelmilla kuka vain on voinut muokata kuvia, joko hienosäätäen tai tehden niihin perinpohjaisia muutoksia. Käsiteltyjen kuvien erottaminen aidoista on ollut haaste jo pidemmän aikaa. Nykyisin niin kutsuttu digitaalinen imitointi on yhä aidompaa ja vakuuttavampaa ja viime vuosina erityisesti tekoälyn hyödyntäminen kuvien ja videoiden muokkauksessa on tehnyt väärennösten silmämääräisestä tunnistamisesta jopa mahdotonta.[36]

Syvävääreennettyjen videoiden sekä äänitteiden aiheuttamat riskit verrattuna tavallisiin kuvamuokkauksiin, joissa ei ole käytetty tekoälyä, ovat merkittävät. Esimerkiksi tekoälyn avulla

[34] EPRS 2021 s. 12.
[35] Guo ym. 2022 s. 1.
[36] Citron – Chesney 2019 s. 1759.

laadituissa syvävääreennöksissä olevilla audiovisuaalisilla elementeillä on huomattavasti vahvempi psykologinen vaikutus kuin tavallisella medialla. Tämä johtuu niiden uskottavuudesta. Uhkana on myös uutisten välityksen osittainen siirtyminen pois luotettavilta mediayrityksiltä yksityishenkilöille. Sosiaalisen median käyttäjien tuottamaa sisältöä on hankalampaa valvoa, eikä yksityishenkilöitä sido samalla tavalla esimerkiksi eettiset ja professionaaliset ohjeet kuin työtään tekeviä toimittajia. Lisäksi erityisesti tietoverkossa kuvista on tulossa hallitseva ilmaisutapa. Näin ollen myös syvävääreennöksistä voi tulla merkittävä informaation lähde tulevaisuudessa.[37]

Yksi ero älyteknologiaa käyttävässä väärennyksessä verrattuna tavalliseen väärennykseen on sen helppous. Syvävääreentämiseen tarkoitettua teknologiaa on tietoverkossa paljon ja helposti saatavilla.[38] Deepfake-videoiden tekemiseen löytyy myös runsaasti opetusvideoita, joiden avulla teknologiaa voi oppia käyttämään peräti kymmenessä minuutissa.[39] Syvävääreennösten tekeminen on siis halpaa ja niin käyttäjäystävällistä, ettei niiden laadintaan tarvita erikoista teknologiaosaamista. Joillakin sosiaalisen median alustoilla syvävääreennösteknologiaa löytyy jopa sovelluksesta sisäänrakennettuna. Näin käyttäjä voi luoda ja julkaista syvävääreennöksen samassa sovelluksessa. Tämä on mahdollista ainakin kahdessa suositussa sosiaalisen median palvelussa, Snapchatissa ja TikTokissa.[40] Tämä mahdollistaa väärennettyjen videoiden massatuotannon sekä niiden vaivattoman levittämisen.[41]

2.3. Hyödyt

Termejä ”väärennös” ja ”fake” käytetään usein laittomuuksien tai muutoin kielteisten tekojen yhteydessä. Ihmiselle syntyy siksi herkästi negatiivinen mielikuva syvävääreennöksistä. Tämä

[37] EPRS 2021 policy s. 1–2.
[38] Lokakuussa 2022 Googlen hakukoneen avulla internetistä löytyi kymmeniä syvävääreentämiseen tarkoitettuja applikaatioita sekä ohjeita videoiden tekemiseen hakusanalla ”deepfake app”.
[39] Opetusvideoita on runsaasti saatavilla YouTube-palvelussa. Ks. esim. Tom Baranowiczin ”How to make DeepFake in 10 mins – Tutorial”.
[40] Kugler – Pace 2021 s. 11.
[41] EPRS 2021 s. 2.

konnotaatio ei kuitenkaan tarkoita sitä, etteikö syvävääreennöksiä voisi käyttää myös laillisesti hyödyksi.[42]

Deepfake-teknologia tuo mukanaan uudenlaisia mahdollisuuksia ja sitä voidaan hyödyntää eri yhteyksissä, kuten opetuksessa, kaupallisessa tarkoituksessa, yhteiskunnallisissa kannanotoissa ja yksilön itseilmaisussa. Opetuksessa syvävääreentämistä voidaan käyttää esimerkiksi luomalla videoita, joissa historialliset henkilöt puhuvat suoraan oppilaille oman aikansa tapahtumista. Vaihtoehtoisen ja siten myös mielenkiintoisen opetustavan kautta oppiminen voi olla tehokkaampaa kuin perinteisempien oppimiskeinojen avulla.[43] Muotiteollisuudessa teknologiaa voidaan käyttää esimerkiksi niin, että kuluttaja sovittaa vaatteita virtuaalisesti verkossa ja voi siten hyödyntää syvävääreennösteknologiaa tehdessään mahdollisen ostopäätöksen.[44]

Erityisen merkittävä segmentti deepfake-teknologialle on elokuvateollisuus. Teknologia on täydellinen ratkaisu, jos esimerkiksi suositun elokuvatrilogian näyttelijä kuolee oikeassa elämässä juuri ennen kuin viimeinen elokuva saadaan kuvattua loppuun. Jos elokuvan juonta ei haluta tai sitä ei ole mahdollista muuttaa, voidaan elokuva tehdä loppuun sijaisnäyttelijän avulla. Älyteknologian ansioista edesmenneen näyttelijän kasvot ja ääni voidaan siirtää sijaisnäyttelijälle. Tuotantoyhtiön kannalta ratkaisun voidaan olettaa olevan toimiva ja halpa.

Edesmenneen henkilön kasvojen ja äänen käyttäminen – oli kyse sitten elokuvasta tai opetustilanteesta – on kuitenkin arvioitava erikseen sekä eettisestä että oikeudellisesta näkökulmasta. Edesmenneestä henkilöstä tehdyn digitaalisen kopion käyttäminen kaupallisessa tarkoituksessa on jokseenkin kyseenalaista, vaikka hyödyntämiseen olisikin saatu lupa kuolleen perillisiltä. Kuolleista ylösnousemisen digimaailmassa on mahdollistanut esimerkiksi korealainen DeepBrain AI. DeepBrain AI on yritys, joka on erikoistunut tekoälyteknologiaan ja muun muassa kuolleiden ihmisten tekoälyavatareiksi muuttamiseen. Yrityksen tavoite on luoda ihmiselle keskusteluyhteys edesmenneen perheenjäsenen kanssa, joka luodaan näyttämään ja kuulostamaan oikealta ihmiseltä.[45] New Yorkin osavaltiossa voimassa oleva laki kieltää edesmenneistä

[42] EPRS 2021 s. 3.
[43] Citron – Chesney 2019 s. 1769.
[44] Dietmar 2019.
[45] DeepBrain AI.

näyttelijöistä ja muista esiintyjistä kaupalliseen tarkoitukseen tehdyt syvävääreennökset ilman kuolleen perillisten tai vastaavan lupaa. Tämä henkilöä koskeva suoja on voimassa 40 vuotta kuolemasta.[46]

Elokuvateollisuudessa deepfake-teknologiaa on jo käytetty korvaamaan edesmenneitä näyttelijöitä tai nuorentamalla näyttelijää elokuvaan sopivammaksi. Näin tehtiin esimerkiksi *The Mandalorian* -sarjassa, joka sijoittuu Tähtien sota -universumiin. Sarjassa esiintyi nuori Luke Skywalker, joka viimeksi oli nähty saman ikäisenä Jedin paluu -elokuvassa. *The Mandalorian* kuvattiin noin neljäkymmentä vuotta Jedin paluun jälkeen. Vanhojen kuvien avulla luotu nuori Luke saatiin siirrettyä uuteen sarjaan, vaikkakin vanhoja kuvia käytettäessä niiden huono resoluutio teki teknologian käytöstä hieman haasteellista.[47] Aihetta koskeva lainsäädäntö ympäri maailmaa vaikuttaa olevan pirstaloitunutta tai sitä ei ole ollenkaan. Eettisesti asia ei ole yksiselitteinen. Lähtökohtana voitaneen pitää kuolleen henkilön tietojen käyttäjän velvollisuutta sopia kuvan ja äänen käytöstä henkilön perillisten kanssa. Joskus kuolleella henkilöllä ei ole perillisiä, eikä sellaisessa tapauksessa lupaa voida kysyä. Erikoinen tilanne olisi myös se, jos perillinen antaisi luvan tietojen käyttöön tavalla, joka selvästi olisi ollut vainajan toiveiden vastainen.

Deepfake-teknologiaa voidaan hyödyntää myös lääketieteessä. Teknologian avulla voidaan esimerkiksi luoda sairauden takia puhekyvyn menettäneen henkilön omaa ääntä imitoiva tietokoneääni, jota sairastunut henkilö voi käyttää tietokoneen avulla.[48] Niin ikään on kehitetty terapiamuotoja ahdistuneisuushäiriöihin, joissa deepfake-teknologian avulla potilaalle näytetään videoita, joissa he itse näyttävät voittaneen omat pelkonsa.[49] Syvävääreentämistä voidaan siis mahdollisesti hyödyntää myös terveydenhuollossa.

Edellä mainitun lisäksi syvävääreennökset ovat oiva tapa luoda kantaaottavia satiireja ja parodioita julkisuudessa esiintyvistä henkilöistä. Hyödyntämiskeinoja on varmasti muitakin ja luovuutta käyttämällä niitä keksitään vielä lisää. Mahdollisuuksista huolimatta syvävääreennöksiä

[46] N.Y. Civ. Rights § 50-f(2); Lalla ym. 2022.
[47] Näin Desowitz artikkelissa How Mark Hamill Was De-Aged as Luke Skywalker for 'The Mandalorian' Season 2 Finale – Exclusive (27.8.2021).
[48] Citron – Chesney 2019 s. 1771.
[49] EPRS 2021 s. 29.

voidaan väärinkäyttää monin tavoin ja edellä esitettyihin hyödyllisiinkin tarkoituksiin voi liittyä lainsäädännöllisiä pulmia. Seuraavaksi perehdytään riskeihin, joita deepfake-teknologian käytöstä voi seurata.

2.4. Riskit

EPRS:n syväväärennöksiä koskevassa tutkimuksessa syväväärennöksiin liittyvät riskit on jaettu alla olevan taulukon mukaisesti eri tekoihin, joiden oletetaan johtavan psyykkisiin, taloudellisiin ja sosiaalisiin haittoihin.

Taulukko 1. Syväväärennöksiin liittyviä riskejä.

Psyykkinen haitta	Taloudellinen haitta	Sosiaalinen haitta
• Kiristys • Kunnianloukkaus • Uhkailu • Kiusaaminen • Luottamuksen horjuttaminen	• Kiristys • Identiteettivarkaus • Petos • Osakekurssien manipulointi • Mainehaitta tuotemerkille • Mainehaitta yksilölle	• Uutismedian manipulointi • Haitta taloudelliselle vakaudelle • Haitta oikeusjärjestelmälle • Haitta tieteen uskottavuudelle • Luottamuksen rapautuminen • Haitta demokratialle • Vaalien manipulointi • Haitta kansainvälisille suhteille • Haitta kansalliselle turvallisuudelle

EPRS:n tutkimus syväväärennöksiin liittyvistä riskeistä (2021).[50]

Haittojen aiheuttajina on sekä rikoksia että muita tekoja. Haitallisissa syväväärennöksissä ja niiden levittämisessä voivat yhdistyä julkaistujen tietojen tai kuvien valheellisuus, loukkaukset voivat olla vakavia sekä laajoja ja kuvien tai videoiden uhreille voi aiheutua monia eri yllä esitettyjä seurauksia. Syväväärennökset voivat levitä verkossa ympäri maailmaa ja kuvat tai videot on voitu kuvata ilman lupaa.

[50] EPRS 2021 s. IV.

Psyykkistä haittaa on todettu aiheutuvan kiristyksestä, kunnianloukkauksesta, uhkailusta, kiusaamisesta ja erilaisin keinoin luottamuksen horjuttamisesta. Taloudellinen haitta yksilölle voi syntyä kiristyksestä, identiteettivarkaudesta, petoksesta tai välillisesti mainehaitasta. Laajemmin taloudellinen haitta voi olla seurausta osakekurssien manipuloinnista tai yritykselle kostautuva tuotemerkkiä koskeva mainehaitta. Sosiaaliset haitat tutkimuksessa on nähty laajassa merkityksessä. Näihin lukeutuvat uutismedian manipulointi, haitta taloudelliselle vakaudelle, haitta oikeusjärjestelmälle, haitta tieteen uskottavuudelle, yleinen luottamuksen rapautuminen yhteiskunnassa, haitta demokratialle, vaalimanipulointi, haitta kansainvälisille suhteille tai haitta kansalliselle turvallisuudelle. Syväväärennösten kasvava määrä voi aiheuttaa ihmisissä luottamuspulaa, kun liikkeellä olevan valheellisen tiedon lisäksi sen esittämisestä tehdään yhä vakuuttavampaa.

Suunta, johon verkkoyhteiskunta on tällä hetkellä matkalla, on otollinen syväväärennösten väärinkäytön lisääntymiselle. Syyskuun lopussa 2022 Google ilmoitti siirtävänsä Google hakukoneen uuteen aikakauteen. Aikaisemmin kesällä Googlen hakukoneesta vastaava johtaja Prabhakar Raghavan kertoi, että lähes 40 % 18–24-vuotiaista nuorista käytti videopalvelu TikTokia tai Instagram-sovellusta hakukoneena Googlen sijasta etsiessään ravintolasuosituksia ja muuta informaatiota.[51] Googlen vastaus tähän oli, että tulevaisuudessa yrityksen hakukoneessa sekä Google Mapsissa tullaan panostamaan visuaaliseen hakuun sekä äänen käyttöön. Jatkossa, kun käyttäjä esimerkiksi etsii tietoa tietystä kaupungista, Google korostaa visuaalisia tarinoita ja lyhyitä videoita, joita paikassa vierailleet henkilöt ovat jakaneet.[52] Googlen kannanoton vuoksi on perusteltua olettaa, että tulevaisuudessa Googlen hakukone priorisoi verkosta löytyviä kuvia ja videoita julkaistun tekstin sijaan. Laittomien syväväärennösten osalta voi ilmaantua ongelma, jossa hakukone korostaa disinformaatiota (valheellista tietoa, joka on tarkoituksella luotu ja levitetty vahingoittamistarkoituksessa[53]) tai misinformaatiota (valheellista tietoa, jota ei ole luotu tai levitetty niin, että teon tarkoituksena olisi aiheuttaa vahinkoa[54]) käsittävää visuaalista sisältöä ennemmin kuin faktapohjaista tekstisisältöä.

---

[51] Asiasta uutisoi esimerkiksi Forbes 19.7.2022 otsikolla: ”Google Is Evolving Search As Zoomers Use TikTok, Instagram To Find Things Online”.
[52] Edwards 2022.
[53] Yamaoka-Enkerlin 2020 s. 728.
[54] Yamaoka-Enkerlin 2020 s. 728. Misinformaation levittäjä ei myöskään tiedä levittämänsä tiedon olevan valheellista.

Kuten edellisessä kappaleessa todettiin, syväväärennöksiä voidaan hyödyntää yhteiskunnallisessa vaikuttamisessa, esimerkiksi erilaisten parodioiden ja satiirien avulla. Myös näihin liittyy omat riskinsä. On osoitettu, että erilaiset ääriliikkeet voivat käyttää satiirisia ilmaisuja ikään kuin suojana oman ideologiansa levittämiseen.[55]

2.5. Havainnointi

Syväväärennösten havaitseminen on olennaisessa osassa niiden väärinkäytön ehkäisemisessä. Deepfake-videossa näkyvä teko tai kuuluva puhe saatetaan ottaa vastaan totuutena, kunnes video voidaan luotettavasti todeta väärennetyksi. Syväväärennösten erottaminen aidosta videosta on erityisen tärkeää myös esimerkiksi silloin, kun syväväärennöksiä yritetään käyttää todisteena tuomioistuimessa. Tuolloin vaarana on, että syytön henkilö saa väärennetyn todisteen perusteella tuomion teosta, jota hän ei ole tehnyt.

Väärennettyjä videoita voidaan havaita joko manuaalisesti tai teknologian avulla. Manuaalisella havainnoinnilla tarkoitetaan ihmisen tekemää videon tarkistusta silmämääräisesti etsimällä videosta epäjohdonmukaisuuksia tai vihjeitä, jotka voivat viitata väärennökseen. Manuaalinen havainnointi voi olla mahdollista, kun tarkastettavien videoiden määrä ei ole liian suuri ja kun väärennökset ovat ylipäätään silmämääräisesti havainnoitavissa.[56] Joskus ihmisaivot havainnoivat kuva- ja videoväärennöksiä, vaikka pelkkä silmämääräinen tarkastelu ei erottaisi väärennettyä kuvaa aidosta. Näin todettiin Australiassa tehdyssä tutkimuksessa, jossa aivojen sähkökäyriä mittaamalla huomattiin, että aivot tunnistivat niin sanotut huijauskuvat 54 prosentissa tapauksista, kun taas silmämääräisessä arviossa ihmiset tunnistivat valheelliset kuvat vain 37 prosentissa tapauksista.[57]

Manuaalinen havainnointi on kuitenkin hidasta, työlästä ja epävarmaa. Siksi deepfake-teknologian käyttöön liittyviä erilaisia negatiivisia ilmiöitä on pyritty ratkomaan uutta teknologiaa

[55] Greene 2019 s. 33–34; EPRS 2021 s. 28.
[56] EPRS 2021 s. II.
[57] Moshel ym. 2022 s. 5–7.

kehittämällä. Esimerkiksi Microsoft on kehittänyt syvävääreennettyjä kuvia ja videoita tunnistavan Microsoft Video Authenticator -sovelluksen, joka arvioi niiden valheellisuuden todennäköisyyttä. Teknologia toimii, muttei täydellisesti.[58] Erehtymätöntä väärennettyjä videoita tunnistavaa älyteknologiaa ei tiettävästi ole vielä kehitetty. Sosiaalisen median suuryhtiö Facebook (nykyinen Meta) toteutti vuosina 2019–2020 syvävääreennettyjen videoiden tunnistushaasteen (*Deepfake Detection Challenge*), joka järjestettiin sellaisten uusien teknologioiden luomiseksi, jotka havaitsevat deepfake-videoita ja muuta manipuloitua mediaa. Palkintorahoja haasteessa oli jaossa yhteensä miljoona Yhdysvaltain dollaria.[59] Kovasta yrityksestä huolimatta haasteen yhteydessä ei saatu kehitettyä teknologiaa, jonka avulla olisi täysin mahdollista erottaa syvävääreennetty video alkuperäisestä versiosta. Yhdysvalloissa on laissa säädetty, että mikäli Yhdysvaltain hallitus kehittää teknologian, jolla syvävääreennöksiä voidaan havaita, sillä on velvollisuus jakaa tämä teknologia verkkoalustojen kanssa.[60]

Syvävääreennösten havainnointia vaikeuttaa niiden luomiseen tarkoitetun teknologian nopea kehittyminen. Tavallisesti tekoälyä hyödynnettävässä havainnoinnissa epäiltyä syvävääreennöstä verrataan suureen dataan muita videoita, joiden tiedetään olevan deepfake-videoita. Syvävääreennös voidaan tunnistaa, jos se muistuttaa datajoukosta löytyvää toista syvävääreennöstä. Pienetkin muutokset tunnistettavana olevissa syvävääreennöksissä vaikuttavat havainnoinnissa käytettävän tekoälyteknologian toimimisen todennäköisyyteen.[61]

Toinen ongelma on, että valmis syvävääreennös usein pakataan tai pienennetään sen jakamista varten eri verkkoalustoilla, kuten sosiaalisessa mediassa tai pikaviestintäpalveluissa. Pikselien ja havaittavissa olevien ääni- tai kuvahäiriöiden määrän väheneminen tiedostoa pienennettäessä hankaloittaa syvävääreennösten havaitsemista.[62] Esimerkiksi puhelinkeskustelussa huonosti

---

[58] Mustak 2022 s. 8.
[59] Parhaimmillaan syvävääreennettyjen videoiden tunnistamiseen kehitetty teknologia tunnisti väärennetyn videon 82.56 prosentin tarkkuudella. Tässä oli käytetty ohjattua koneoppimista, jossa tietojärjestelmälle tarjottiin yksityiskohtaista dataa liittyen suureen määrään eri tapauksia, joiden lopputulos oli tiedossa. Ohjaamattoman oppimisen osalta lopputulos oli eri: kun opetusdatasta ei kerrottu etukäteen algoritmille mitään, löysi tietojärjestelmä väärennetyn videon vain 65.18 prosentin todennäköisyydellä. Deepfake Detection Challenge Results: An open initiative to advance AI (12.6.2020).
[60] H.R.3230, 116th Cong. DEEP FAKES Accountability Act § 7; Langa 2021 s. 794.
[61] Huang ym. 2020.
[62] EPRS 2021 s. III.

kuuluva ääni voi tehdä oikean ja valheellisen äänen vertailusta vaikeaa. Tämän lisäksi ongelmaksi on havaittu jo olemassa olevat havainnointilaitteet siltä osin, että ne ovat tutkimuksen mukaan olleet kielen suhteen syrjiviä, sillä teknologiaa on kehitetty tunnistamaan lähinnä englanninkielisiä syväväärennöksiä.[63]

Koska tässä suhteessa teknologiasta saatava hyöty ei poissulje syväväärennösten väärinkäytön mahdollisuutta, on perusteltua tarkastella lainsäädännöllisiä keinoja deepfake-videoiden väärinkäytön ehkäisemiseksi.

[63] EPRS 2021 s. 25.

3. SYVÄVÄÄRENNÖSTEN VÄÄRINKÄYTTÖ

3.1. Aluksi

Algoritmien ja älykkäiden teknologioiden ongelmia on tiedostettu ympäri maailmaa ja näiden väärinkäyttöä on pyritty estämään muun muassa erilaisilla eettisillä suosituksilla, joita ovat laatineet sekalaiset työryhmät. Näitä ohjeita, joiden tarkoitus on ohjata tekoälyn ja algoritmien kehittämistä, käyttämistä ja soveltamista, on laadittu melkein kaksisataa erilaista. Tässä ohjeviidakossa on julkaisuja niin kansallisilta kuin kansainvälisiltä toimijoilta, akateemisilta ja uskonnollisilta yhteisöiltä sekä suuryrityksiltä. Näistä voidaan mainita esimerkiksi Euroopan komissio, OECD, G20, katolinen kirkko, Microsoft, Facebook ja Google.[64] Vaikka kyseisillä eettisillä suosituksilla ei ole juridista pohjaa, eivätkä ne velvoita julkaisijaansa tai käyttäjäänsä, on jo yksin ohjeiden määrästä pääteltävissä, että jonkinlaiset eettiset suuntaviivat älykkään teknologian kehittämisessä ja käytössä on nähty välttämättömänä.

Jotta ohjeet saataisiin velvoittavaan muotoon, on ohjeistukset saatava osaksi lainsäädäntöä. Syvävääreennöksiin liittyviä riskejä ja haittoja on siis pyrittävä ennaltaehkäisemään juridisin keinoin. Tehtävä on monesta syystä haasteellinen: ensinnäkin ongelmaksi tulee kysymys vastuun kohdentamisesta. Kun kehitetään teknologiaa, jota voidaan sekä hyödyntää että väärinkäyttää, on päätettävä, onko epäeettisestä (tai mahdollisesti lainvastaisesta) toimesta vastuussa teknologian kehittäjä, teknologiaa tarjoava yritys vai loppukäyttäjä. Ongelmallista on myös se, missä vaiheessa lainvastaisuuksiin on puututtava. Onko tarkoituksenmukaista kieltää kokonaan jonkin teknologian, esimerkiksi deepfake-applikaatioiden, käyttö, jotta sen käytön negatiiviset vaikutuksen voidaan minimoida? Teknologian hyödyllisyyden vuoksi parhaana lopputuloksena voitaisiin pitää sitä, että teknologian käyttöä rajoittava lainsäädäntö kohdentuisi vain sen väärinkäyttöön.

Eettisen näkökulman lisäksi tekoälyn sallittavuuden arvioinnissa ei pidä unohtaa sen poliittista ja rahallista vaikutusta sekä näiden mukanaan tuomaa valtaa. Tekoälyä kehittämällä ei pyritä ainoastaan arkisten toimintojen helpottamiseen, vaan kehitystyön taustalla on myös kilpailu markkinapaikasta satojen miljardien eurojen arvoisella toimialalla.[65] Teknologian

[64] Rusanen 2021 s. 38–39.
[65] Rusanen 2021 s. 46.

kehittämisellä ja sen mahdollisella rajoittamisella lainsäädännön kautta on siis myös merkittäviä taloudellisia vaikutuksia.

Syväväärentämistä hyödyntävien rikollisten tekoja on hankalaa ennustaa, koska mahdollisia skenaarioita on niin monia erilaisia. Kysymys voi olla harmittomasta pilasta, pyrkimyksestä vaikuttaa valtioiden välisiin suhteisiin tai jotain siltä väliltä. Rangaistavuuteen on suhtauduttava eri tavoin riippuen syväväärentämisen taustalla olevista tosiseikoista. Tuleeko esimerkiksi videota katsoessa selväksi, että kyse on syväväärennöksestä? Onko videon tekemiseen annettu siinä näkyvältä henkilöltä lupa? Onko video selvästi laadittu niin, että sen avulla on tarkoitus synnyttää poliittista keskustelua? Joka tapauksessa, oli kyse sitten tarkoituksella tehdystä rikoksesta tai ei, verkkoympäristö tekee yleisesti rikoksenteosta aiempaa helpompaa[66] ja luo paljon erilaisia mahdollisuuksia myös deepfake-teknologian väärinkäytölle.

Yksi suurimmista haasteista syväväärennösten haitallisessa käytössä lienee videon laatijan vaikea jäljitettävyys. Mikäli laittoman videon tekijää ei saada selville, ei tätä vastaan voida nostaa syytettä, eikä uhrilla ole mahdollisuutta hakea yksityisoikeudellista korvausta kärsitystä haitasta. Usein myös videon tekijä voi olla eri henkilö kuin sen levittäjä. Lisäksi videon levittäjä saattaa toimia anonyymisti esimerkiksi pimeässä verkossa, jossa tällä on mahdollisuus kätkeä oma IP-osoitteensa ja muut tunnistetietonsa.[67] Jos syväväärennöksen alkuperäistä tekijää ei saada selville, esiin voi nousta kysymys vastuun siirtämisestä joko videon levittäjälle, tai sen sivuston ylläpitäjälle, jolle video on ladattu.

Deepfake-teknologian väärinkäyttö voi olla tehokasta myös ilman kuvaa. Hyvänä esimerkkinä toimii vuonna 2019 tapahtunut toimitusjohtajapetos, jossa rikolliset olivat väärentäneet tekoälyn avulla erään yrityksen toimitusjohtajan äänen taloudellisen hyödyn saavuttamiseksi. Tekijät olivat soittaneet yrityksen toimipisteeseen ja puheluun vastannut henkilö, joka oli tunnistanut soittajan saksalaisen aksentin, oli tämän pyydettyä maksanut runsaan 200 000 euron maksun rikoksen tekijän tekaisemalle alihankkijalle.[68] Samalla tekniikalla rikolliset voivat saada haltuunsa

[66] Saarenpää 2015 s. 68.
[67] Citron – Chesney 2019 s. 1792.
[68] Stupp 2019; Europol 2019.

yrityssalaisuuksia, salasanoja tai muuta tärkeää tietoa yritykseltä, mistä seuraa valtavat tietoturvariskit sekä taloudellisia haittoja.[69] Riski tällaisen toiminnan onnistumisesta kasvaa, kun pelkän puhelun sijasta uhriin saadaan reaaliaikainen videoyhteys, jossa yhdistyvät sekä valheellinen ääni että valheellinen kuva.

Syväväärentämistä koskevassa ulkomaisessa akateemisessa tutkimuksessa on perinteisesti jaettu siitä aiheutuvat haitat kahteen kategoriaan: syväväärennöksistä yksilöille aiheutuviin haittoihin sekä laajempiin yhteiskunnallisiin haasteisiin.[70] Tämä tutkielma käsittelee nimenomaan yksilön vastuuta rikosoikeudellisesta näkökulmasta, mutta niin, että syväväärentämisen seuraukset voivat olla sekä yksittäiseen henkilöön että yhteiskuntaan kohdistuvia. Tarkoitus on siten korostaa kyseisen teknologian käytöstä mahdollisesti aiheutuvia seurauksia ja yksilön vastuuta sekä samalla luoda pohjaa keskustelulle siitä, onko nykyinen lainsäädäntömme riittävää syväväärennösteknologian käytöstä seuraavien haittojen minimoimiseksi.

Seuraavaksi esitettävät väärinkäytön vaikutukset on jaettu tässäkin selvyyden vuoksi yksilön sekä yhteiskunnan tasolla syntyviin vaikutuksiin. Näistä vaikutuksista osa liittyy tosielämässä jo havaittuihin haittoihin, kun taas toiset ovat sellaisia, ettei kyseisiä ongelmia ole vielä toistaiseksi havaittu, mutta jotka ovat täysin uskottavia jo nykyisen olemassa olevan teknologian puitteissa. Nämä haitat voivat toteutua tulevaisuudessa, jos erilaisiin toimenpiteisiin ei ryhdytä ongelmien minimoimiseksi.

3.2. Vaikutuksia yksilön tasolla

Muiden kuin julkisuuden henkilöiden osalta yksittäisiin henkilöihin kohdistuvista syväväärentämisen väärinkäytöistä on suomalaisessa mediassa kirjoitettu hyvin vähän, jos lainkaan. Tältä osin myös ulkomainen tutkimus on ollut vähäisempää, pois lukien jo aikaisemmin mainittujen syväväärennettyjen pornovideoiden osalta. Deepfake-teknologian käytön yleistyessä lienee todennäköistä, että yksityishenkilöt ovat yhtä lailla vaarassa joutua väärennettyjen videoiden

[69] Ciancaglini ym. 2020.
[70] Ks. esimerkiksi Kugler – Pace 2021 s. 13.

kohteeksi. Tilanne voi olla myös se, että videoita tehdään jo, mutta toistaiseksi niistä ei ole levinnyt tietoa julkisuuteen.

Syväväärennösteknologia luo erinomaisen alustan mainehaittojen aiheuttamiselle ja siihen syvävääräennösten väärinkäytöllä usein myös pyritään. Ongelmia aiheutuu jo pelkästään videoiden laadinnassa, mutta erityisesti, kun niitä levitetään, joko sosiaalisen median alustoilla tai muutoin. Yksi yleinen ja erittäin haitallinen syvävääräennöksiin liittyvä ilmiö on aikaisemminkin esiin tullut pornograafisten väärennösten tekeminen ja jakaminen.[71] Lisäksi erityisesti väärennetyt pornovideot näyttävät kohdistuvan lähes yksinomaan naisiin, mikä osoittaa sukupuolen merkitystä syvävääräennöksen väärinkäytön kohteeksi joutumisessa.[72]

Nuorten keskuudessa videoiden teossa kyse voi olla yhteisestä hauskanpidosta, mutta syvävääräentäminen avaa myös ovia uudenlaiseen kiusaamiseen. Vaikkakin on mahdollista, että video tai äänite on luotu videossa näkyvän henkilön tai äänitteessä kuuluvan henkilön suostumuksella, on usein myös todennäköistä, että videon tekijä toimii ilman tällaista lupaa.[73] Syvävääräennöksen leviämisestä voi seurata mainehaittaa ja muuta negatiivista julkisuutta videoon väärennetylle henkilölle. Tämä voi johtaa henkiseen ahdistukseen sekä tulevaisuudessa ilmeneviin erilaisiin haasteisiin.

Deepfake-teknologian avulla videon tekijällä on ikään kuin mahdollisuus varastaa toisen henkilön kasvot ja käyttää niitä oman hyödyn tavoittelussa. Ihmisen kasvot sekä ääni ovat molemmat osa ihmisen biometristä identiteettiä, joita voidaan käyttää esimerkiksi henkilötunnistuksessa. EU:n tietosuoja-asetuksen (2016/679, TSA) 4 artiklassa määritellään biometrinen tieto niin, että se kattaa kaikki luonnollisen henkilön fyysisiin ja fysiologisiin ominaisuuksiin tai käyttäytymiseen liittyvät teknisellä käsittelyllä saadut henkilötiedot, kuten kasvokuvat tai sormenjälkitiedot, joiden perusteella kyseinen henkilö on mahdollista tunnistaa tai henkilön tunnistaminen voidaan varmistaa. Suomessa näiden biometristen tunnisteiden käyttö henkilötunnistuksessa on mahdollista myös tunnistus- ja luottamuspalvelulain (617/2009) nojalla.[74] Biometrisiä

[71] Esim. Kugler – Pace 2021 s. 1.
[72] EPRS 2021 s. III.
[73] Citron – Chesney 2019 s. 1758.
[74] Saaripuu 2019 s. 134.

tunnistusominaisuuksia on pidetty ainutkertaisina, pysyvinä sekä muuttumattomina ja siten erinomaisina välineinä henkilötunnistuksessa.[75] Voisi ajatella, että syvävääreennösten myötä yleistyvä ja suhteellisen helppo tapa käyttää toisen henkilön ääntä sekä kasvokuvaa voi tuoda muutoksia myös biometriikan käyttöön tunnistautumisessa etenkin, kun tunnistautuminen tapahtuu verkossa.

Esitettyjen esimerkkien perusteella syvävääreennösten kohteena ovat usein yksittäiset henkilöt. Monille yksilöille suunnatut vaikutukset voivat kuitenkin kumuloitua niin, että ne aiheuttavat vahinkoa tietylle ryhmälle tai organisaatiolle. Ensinnäkin yksilöön kohdistuva nöyryytys voi yhtä lailla luoda negatiivisia vaikutuksia uhrin perheeseen. Mainehaitat saattavat koskea koko perhettä ja uhrin kokemukset aiheuttavat usein myös huolta perheen muissa jäsenissä. Ammattitasolla esimerkiksi journalistin tai muun työntekijän päätyminen laittoman syvävääreennöksen uhriksi voi joissakin tilanteissa aiheuttaa mainehaittaa yritykselle, jossa uhri työskentelee. Lopulta hyvin kohdistetut deepfake-videot ja niiden kumulatiivinen yhteisvaikutus voivat johtaa vakaviinkin haittoihin yhteiskunnan tasolla.[76]

3.3. Vaikutuksia yhteiskunnan tasolla

Syvävääreentämisen väärinkäytön haittavaikutuksia voi yhtä lailla ilmetä yhteiskunnan tasolla. Väärinkäytön seurauksia voivat olla esimerkiksi poliittisen keskustelun vääristyminen, vaalien manipulointi, valtiollisiin instituutioihin kohdistuvan luottamuksen heikkeneminen, yhteiskunnan polarisaatio, valtioiden välisten suhteiden häiriintyminen sekä turvallisuusuhkien lisääntyminen.[77] Visuaalisella medialla ylipäätään on tärkeä rooli esimerkiksi disinformaation levittämisessä.[78] Videoilla voidaan myös sabotoida esimerkiksi sotilasoperaatioita ja erilaisia tiedustelutehtäviä.[79] Kaikesta tästä voi seurata valeuutisia, trollien tai robottien aiheuttamaa

[75] Saaripuu 2019 s. 23.
[76] EPRS 2021 s. IV, 49.
[77] Citron – Chesney 2019 s. 1758.
[78] Esim. Dan ym. 2021.
[79] Langa 2021 s. 761.

sosiaalisen median manipulointia tai jopa julkisen epäluottamuksen lietsomista tieteellistä tietoa kohtaan.[80]

Informaatiovaikuttamisen ja muiden hybridiuhkien taustalla olevien toimijoiden tunnistaminen on varsinkin nykypäivänä hyvin hankalaa. Toimijat voivat olla joko yksittäisiä henkilöitä, henkilöryhmittymiä tai esimerkiksi valtiollisia toimijoita. Joskus nämä valtiolliset toimijat voivat olla järjestäytyneen rikollisuuden jäseniä, kuten selvitysten mukaan oli ollut Venäjän toiminnassa vuonna 2014 Krimin valtauksen yhteydessä.[81] Joka tapauksessa rajat ylittävien kyberhyökkäysten tekijöitä ei yleensä löydetä.[82]

Tekoälyllä on siis suuri merkitys valtioiden kokonaisturvallisuuden kannalta, sillä tekoälyä voidaan käyttää jopa valtioiden väliseen hyökkäykseen. Syvävääretämisen osalta hyökkäyksessä käytettävä hyöty saadaan ensin keräämällä dataa siitä, minkälaisella informaatiovaikuttamisella operaatiosta aiheuttaa mahdollisimman paljon haittaa. Tekoäly voidaan ohjelmoida hyödyntämään kaikkia käytettävissä olevia informaatiokanavia yhteiskunnan heikkojen kohtien tunnistamiseen sen eri toiminnoissa.[83] Näin hybridivaikuttamisessa ja disinformaation levittämisessä voidaan vaikuttaa yhteiskunnan heikkoihin kohtiin.

Informaatiovaikuttamisella pyritään usein aiheuttamaan polarisaatiota demokraattisissa yhteiskunnissa. Vaikuttamisen tavoitteena on yleensä joko yhteiskuntarauhan häiritseminen tai ylipäätään vaikuttaminen päätöksentekoon.[84] Informaatiovaikuttaja voi olla myös salaliittoteorioitsija, joka pyrkii levittämään valheellista informaatiota, joko tietoisesti tai vilpittömässä mielessä.[85] Informaatiovaikuttamiselle on joka tapauksessa luonteenomaista päätöksenteon ohjaaminen disinformaation tai vaillinaisen tiedon avulla.[86] Esimerkiksi Covid-19 -tautia ja erityisesti rokotuksia koskevan informaation osalta liikkeellä on ollut paljon myös valheellista tietoa. Informaatiovaikuttaja voi tarkoituksella levittää väärää tietoa esimerkiksi rokotuksista ja aiheuttaa

[80] EPRS 2021 policy s. III.
[81] HaVM 19/2021 s. 14–15.
[82] EPRS 2021 s. 53.
[83] Tarkoma 2021 s. 108–110.
[84] HaVM 19/2021 s 16.
[85] Yamaoka-Enkerlin 2020 s. 733.
[86] HaVM 19/2021 s. 16.

siten valtiollisiin terveysviranomaisiin kohdistuvan luottamuksen heikkenemistä.[87] Tämä voi tapahtua esimerkiksi sellaisen deepfake-videon avulla, johon vaikuttaja ohjelmoi korkeassa asemassa olevan terveysviranomaisen kertomaan rokotteen ”salaisina pidetyistä” haittavaikutuksista. Videon katsoja, joka taas luulee katsovansa aitoa tallennetta, voi levittää videota uskoen sen sisältämään informaatioon. Tällöin on kyse misinformaation levittämisestä.[88]

Tarkasti kohdennettavilla syvävääreennöksillä pyritään vaikuttamaan ihmisiin, jotka uskovat vain sen, mikä sopii heidän maailmankuvaansa – tätä ilmiötä voidaan kutsua ideologiseksi kaikukammioksi (*echo chambers*) tai informaatiokuplaksi (*filter bubbles*). Oikeaan kohderyhmään osuessaan syvävääreennöksellä voidaan aiheuttaa poliittista manipulaatiota tai kohdennettua propagandaa. Luonnollisesti tämä lisää myös salaliittoteorioiden leviämistä.[89]

Euroopan unionissa disinformaation levittämiseen on yritetty puuttua esimerkiksi vuonna 2018, kun komissio julkaisi käytännesäännöt disinformaation torjuntaan[90] ja uudelleen vuonna 2022, jolloin julkaistiin tehostetut disinformaatiota koskevat käytännesäännöt[91]. Monet eri alan toimijat, kuten alustat ja teknologiayritykset, ovat vapaaehtoisesti sitoutuneet toimimaan käytännesääntöjen mukaisesti disinformaation torjumiseksi. Käytännesäännöt ovat joidenkin arvioiden mukaan osoittautuneet tehokkaaksi välineeksi verkossa leviävän disinformaation rajoittamisessa niin vaalivaikuttamisen osalta kuin maailmanlaajuisten kriisienkin, kuten koronaviruspandemian sekä Ukrainan sodan, yhteydessä.[92] Syvävääreennösten osalta näitä toimia on kuitenkin pidetty riittämättöminä.[93]

Eräänä tyypillisenä informaatiovaikuttamisen keinona voidaan pitää vaalivaikuttamista. Kohteena voi olla yksittäinen ehdokas tai puolue, tai esimerkiksi yleisen äänestysinnon vähentäminen tai demokratian uskottavuuden horjuttaminen.[94] Vaaleihin on todistetusti yritetty vaikuttaa

[87] Yamaoka-Enkerlin 2020 s. 748–749.
[88] Yamaoka-Enkerlin 2020 s. 728.
[89] EPRS 2021 s. 33, 52.
[90] Euroopan komissio, *Code of Practice on Disinformation*.
[91] Euroopan komisso, *Strengthened Code of Practice on Disinformation*.
[92] Euroopan komissio, *Disinformation: Commission welcomes the new stronger and more comprehensive Code of Practice on disinformation,* lehdistötiedote 16.6.2022.
[93] Bressan 2019; EPRS 2021 s. 43.
[94] SM:n julkaisu 2019:5 s. 24.

ainakin vuonna 2017, jolloin Venäjä pyrki estämään Ranskan Emmanuel Macronin valinnan maan presidentiksi informaatiomanipuloimalla ranskalaisia äänestäjiä. Pyrkimyksen vaikutukset jäivät kuitenkin vähäisiksi.[95]

Informaatiovaikuttamisen lopputulos olisi saattanut olla toinen, jos manipuloinnissa olisi hyödynnetty deepfake-teknologiaa. Näin totesivat myös tutkijat Bobby Chesney sekä Danielle Citron, joiden mukaan Macronista tehdyllä aidon oloisella deepfake-videolla olisi saattanut olla merkittävä vaikutus kansalaisten äänestyskäyttäytymiseen, etenkin jos Macron olisi tehnyt tai sanonut videolla jotakin shokeeraavaa. Vaalivaikuttamisessa myös syvävääreennöksen julkaisuajankohdalla on suuri merkitys. Syvävääreennetty video, jolla halutaan aiheuttaa haittaa esimerkiksi vaaliehdokkaalle, on julkaistava levitykseen tarpeeksi aikaisin ennen vaaleja, jotta se ehtii levitä mahdollisimman laajalle. Toisaalta on myös tärkeää, ettei aikaa jää julkaisun ja äänestyksen väliin liikaa, jotta videota ei ehditä todeta väärennetyksi (jos sitä voidaan ylipäätään luotettavasti todeta).[96] Deepfake-videon aiheuttama haitta voidaan maksimoida hyödyntämällä pientä aikaikkunaa ennen virallista äänestyspäivää.

Yhdysvaltain kansallisen tiedustelupalvelun vuosittaisessa turvallisuusraportissa varoiteltiin jo vuonna 2019 syvävääreennösten väärinkäytöstä ja vaikuttamisyrityksistä vuoden 2020 presidentinvaaleissa. Tiedustelupalvelussa uskottiin, että videoita väärentäisivät ja levittäisivät useat valtion viholliset sekä eri ehdokkaiden kannattajat vastapuolen ehdokkaista.[97] Odotukset syvävääreennösten ilmaantumiselle olivat suuret. Eri uutislähteiden mukaan vaikuttamisyritykset deepfake-videoiden avulla jäivät kuitenkin vuoden 2020 presidentinvaaleissa asiantuntijoidenkin yllätykseksi olemattomiksi.[98] Joko syvävääreentämisellä aiheutettua vaalivaikuttamista ei siis juurikaan ollut, tai sitä ei huomattu.

---

[95] Citron – Chesney 2019 s. 1778–1779. Venäläisten toiminnassa yhdistyivät kybervakoilu sekä informaatiovaikuttaminen. Lukuisia viestejä ja asiakirjoja varastettiin, niiden sisältöä muutettiin ja dramatisoitiin. Muokattuja asiakirjoja julkaistiin ja levitettiin. Kuitenkin tekijän helpon jäljitettävyyden, Macronin tiimin älykkään toiminnan sekä muiden syiden seurauksena yrityksen vaikutukset jäivät vähäisiksi.
[96] Citron – Chesney 2019 s. 1778.
[97] Näin totesi Yhdysvaltain tiedustelupalvelun johtaja Daniel R. Coats vuoden 2019 maailmanlaajuisessa turvallisuusraportissa ”Statement for the Record: Worldwide Threat Assessment of the US Intelligence Community” s. 7. (29.1.2019).
[98] Ks. esim. Simonite 2020 ja Mak – Temple-Raston 2020.

Poliitikot voivat olla syvävääreennösten kohteena myös muulloin kuin vaalikampanjoinnin aikana. Esimerkiksi vuonna 2018 yhdysvaltalainen koomikko Jordan Peele loi syvävääreennöksen entisestä Yhdysvaltain presidentti Barack Obamasta. Videolla Obama avoimesti kritisoi silloista presidentti Donald Trumpia. Vaikka syvävääreennös ei ollut kovinkaan laadukkaasti tehty, se levisi nopeasti ympäri maailmaa keräten paljon huomiota.[99] Samalla ajatuksella tehtiin myös video Britannian pääministeri Boris Johnsonista. Video antoi olettaa, että hän kannatti silloista poliittista vastustajaansa. Nämä molemmat videot tehtiin kuitenkin varoittamaan katsojia syvävääreennösten aiheuttaman mis- ja disinformaation leviämisestä, eikä niillä ollut tarkoitus aiheuttaa todellista haittaa Obamalle ja Johnsonille.[100]

Vaikka toistaiseksi vaalivaikuttamista syvävääreennösten avulla on tapahtunut yllättävän vähän, on riski kuitenkin olemassa. Kansainvälisissä syvävääreennöksiä koskevissa tiedejulkaisuissa vaalien manipulointi tuntuu nousevan lähes poikkeuksetta yhdeksi merkittävimmistä väärinkäytön ongelmista.[101]

[99] Esim. Yamaoka-Enkerlin 2020 s. 742 sekä Kugler – Pace 2021 s. 16–17.
[100] Kugler – Pace 2021 s. 16.
[101] Aihetta käsittelevät esimerkiksi Citron – Chesney 2019; Kugler – Pace 2021 ja Langa 2021.

4. SOVELTUVA RIKOSLAINSÄÄDÄNTÖ

4.1. Aluksi

Syvävääreennöksiä koskevissa akateemisissa tutkimuksissa ei ole tullut vastaan pohdintaa siitä, mitkä rikoslain pykälät Suomessa soveltuvat näitä koskevaan laittomaan toimintaan. Toistaiseksi on ollut epäselvää, kattaako Suomen rikoslaki kaikki syvävääreentämisellä haittaa aiheuttavat vakavat teot. Tässä kappaleessa syvennytään edellä esitettyjen skenaarioiden sekä muiden mahdollisten laittomien syvävääreentämiseen liittyvien tekojen taustalla vaikuttavaan rikoslainsäädäntöön ja säännöksiin, jotka voisivat olla sovellettavissa väärinkäytöstä rankaisemiseksi.

Maailmalla on ainakin jollakin tasolla koettu, ettei rikoslainsäädännön suomin keinoin ole voitu puuttua syvävääreennösteknologian väärinkäyttöön tarvittavissa määrin. Esimerkiksi Yhdysvalloissa deepfake-videoiden haittoihin on jo lakimuutoksilla reagoitu, sillä olemassa olevaa lainsäädäntöä ei ole pidetty riittävänä. Muutamissa osavaltioissa, kuten Kaliforniassa ja Virginiassa, on pyritty syvävääreentämistä koskevalla lainsäädännöllä estämään valheellisen pornografian luomista ja levittämistä. Lisäksi ainakin Kalifornian ja Texasin osavaltiot ovat kieltäneet poliitikoista tehdyt syvävääreennökset vaaleja edeltäviksi viikoiksi.[102]

Verkossa tapahtuvat identiteettivarkaudet ja muut henkilöllisyyden väärinkäytökset tehdään tyypillisesti kaapatuista verkkoliittymistä, jotta kiinnijäämisen riski olisi pienempi. Tällöin tutkintatoimetkin kohdistuvat ensin sivulliseen tahoon. Usein rikollisin keinoin hyödynnettäviä henkilötietoja hankitaan sieltä, missä niitä on helposti saatavilla. Erilaiset tietojenkäsittelyn prosessit nousevat tässä esiin. Tietokaappaukset voivat tapahtua esimerkiksi sellaisista palvelimista, joissa ihmisten henkilötietoja on käsitelty edellytetyt tietosuojatoimet laiminlyöden tai muutoin sopimusehtojen vastaisesti. Riskialttiita ympäristöjä ovat suojaamaton työasema tai mobiililaite.[103]

[102] Kugler – Pace 2021 s. 4. Säännökset alkuperäisinä: Kalifornian osalta ks. CAL. CIV. CODE § 1708.86 sekä CAL. ELEC. CODE § 20010, Virginian osalta VA. CODE ANN. § 18.2–386.2 ja Texasin osalta TEX. ELEC. CODE ANN. § 255.004(d).
[103] Identiteettiohjelma s. 53–56.

Muutoinkin tietoverkossa tapahtuva rikollisuus on tehnyt rikollisuudesta tuottavampaa: usein saavutettavat hyödyt ovat suuria suhteutettuna toteutuksesta aiheutuneisiin pieniin kustannuksiin sekä vähäiseen kiinnijäämisen riskiin. Verkkoympäristössä identiteettivarkaudet ovat myös automatisoitavissa, jolloin niiden volyymi voi olla moninkertainen verrattuna perinteisiin identiteettivarkauksiin, jotka ovat tavallisesti yksittäistapauksia. Automaattisesti kerättävää identiteettitietoa voidaan helposti myydä eteenpäin tai käyttää itse muussa rikollisessa tarkoituksessa.[104] Nämä asiat houkuttelevat rikollisia siirtymään verkkoon ja hyödyntämään pientä kiinnijäämisen riskiä ja saavutettavissa olevan hyödyn maksimointia.

Rikoslainsäädännön soveltuvuudella on merkitystä myös väärinkäytön kiinnijäämisriskin suhteen. Pakkokeinolaissa (806/2011, PKL) määriteltyjen pakkokeinojen käytön mahdollisuudet riippuvat kunkin epäilyn rikoksen kriminalisoinnista sekä rangaistusasteikosta.

Suomi on kansainvälisten sopimusvelvoitteidensa sekä EU-oikeuden nojalla saattanut voimaan tietoverkkorikoksia koskevaa lainsäädäntöä. Osa velvoitteista juontaa juurensa jo vuoden 2007 Euroopan neuvoston tietoverkkorikollisuutta koskevaan yleissopimukseen (ETS 185).[105] Tuoreempana velvoitteena Suomi on implementoinut EU:n tietoverkkorikosdirektiivin (2013/40/EU) osaksi lainsäädäntöään.[106] Tietoverkkorikosdirektiivi on jossakin määrin samansisältöinen tietoverkkorikollisuutta koskevan yleissopimuksen kanssa. Merkittäviä muutoksia kuitenkin tuli, rikoslainsäädännön osalta 35 lukuun (vahingonteosta) sekä 38 lukuun (tieto- ja viestintärikoksista). Direktiivin seurauksena kriminalisoitiin esimerkiksi syvävääreennösten väärinkäytön osalta merkittävä identiteettivarkaus.[107] Identiteettitiedon väärinkäytön lisäksi uudistuksen kohteena olivat muun muassa datavahingonteko sekä vaaran aiheuttaminen tietojenkäsittelylle.[108] Nämä säännökset astuivat voimaan vuonna 2015. Ensiksi käsitelläänkin syvävääreentämistä hyödyntämällä tehtyä identiteettivarkautta.

---

[104] OM 4/41/2013 s. 10; HE 153/2006 vp s. 59–60; Identiteettiohjelma s. 54.
[105] Euroopan neuvoston tietoverkkorikollisuutta koskeva yleissopimus 60/2007.
[106] Muutokset pohjautuivat Euroopan parlamentin ja neuvoston direktiiviin 2013/40/EU tietojärjestelmiin kohdistuvista hyökkäyksistä ja neuvoston puitepäätöksen 2005/222/YOS korvaamisesta (EU:n tietoverkkorikosdirektiivi).
[107] Lappi-Seppälä ym. 2022 s. 1129.
[108] HE 232/2014 vp. Tietoverkkorikosdirektiivin pohjalta syntyi esimerkiksi laki rikoslain muuttamisesta (368/2015) sekä laki pakkokeinolain 10 luvun 3 ja 6 §:n muuttamisesta (369/2015).

4.2. Identiteettivarkaus

Toisena henkilönä vilpillisesti esiintyminen on yleistynyt verkossa tapahtuvassa toiminnassa, kuten sähköisten asiointipalveluiden yhteydessä. Identiteetin väärinkäytölle on tunnusomaista jonkin identiteettitiedon kerääminen ja hyödyntäminen oikeudettomasti. Identiteettivarkaudesta säädetään rikoslain 38 luvussa, joka koskee eri tieto- ja viestintärikoksia. RL 38:9a:n (368/2015) mukaan ihminen syyllistyy identiteettivarkauteen, jos hän oikeudettomasti käyttää toisen henkilötietoja, tunnistamistietoja tai muuta yksilöivää tietoa erehdyttääkseen kolmatta osapuolta ja tällä tavoin aiheuttaa joko taloudellista vahinkoa tai vähäistä suurempaa haittaa sille henkilölle, jota kyseinen tieto koskee. Rangaistus identiteettivarkaudesta on sakkoa.

Lain esitöiden mukaan henkilö- ja tunnistamistietona sekä muuna yksilöivänä tietona pidetään kaikkea sellaista tietoa, jonka perusteella kolmas osapuoli voi uskoa tiedon käyttäjän olevan se, jota kyseinen tieto koskee. Henkilötietoa on kaikki tiedot, jotka liittyvät jo tunnistettuun tai tunnistettavissa olevaan luonnolliseen henkilöön. Henkilö voidaan suoraan tai epäsuorasti tunnistaa tunnistetietojen perusteella. Tunnistetietoja taas ovat nimi, henkilötunnus, sijaintitieto, verkkotunnistetieto tai henkilölle tunnusomainen fyysinen, fysiologinen, geneettinen, psyykkinen, taloudellinen, kulttuurinen tai sosiaalinen tekijä (TSA 4 artikla).[109] Identiteettivarkauden säännöksessä esitetty yksilöivä tieto on siis olennainen, jos se yksinään tai muihin annettuihin tietoihin yhdistettynä, mahdollistaa henkilön tunnistamisen.[110] Syvävääreennöksissä esitetyistä henkilöistä on yleensä tunnistettavissa vähintäänkin kasvokuva sekä ääni.

Identiteettivarkaus voi kohdistua mihin tahansa olemassa olevaan identiteettitietoon. Verkkoympäristössä henkilötiedon lisäksi identiteettivarkauden kohteena on mikä tahansa muu tunniste, jonka avulla tunnisteen haltija voi erehdyttää kolmatta osapuolta tai muutoin päästä käsiksi tietoon tai palveluun oikean identiteetin haltijan sijasta.[111] Syvävääreennösten avulla lähes kuka tahansa voi audiovisuaalisin keinoin ja jopa reaaliaikaisesti esittää olevansa toinen henkilö.

---

[109] HE 232/2014 vp s. 36. Hallituksen esityksessä henkilötiedoilla viitataan henkilötietolaissa (523/1999) esitettyyn määritelmään henkilötiedosta. Henkilötietolaki on sittemmin kumottu tietosuojalailla (1050/2018), joka taas täydentää yleistä tietosuoja-asetusta. Siksi tässä esitetty henkilötiedon määritelmä on otettu tietosuoja-asetuksesta.
[110] HE 232/2014 vp s. 36.
[111] Identiteettiohjelma s. 48; Saaripuu 2019 s. 18.

Siksi syvävääräntäminen voi rikolliselle olla houkutteleva apuväline identiteettivarkauden toteuttamisessa.

Identiteettivarkaus-termin alle mahtuu monia erilaisia tekokokonaisuuksia. Näille teoille on yhteistä se, että niissä kerätään jotain identiteettitietoa ilman oikeutta. Pelkästään vääränä henkilönä esiintyminen toiselle yksityishenkilölle ei välttämättä ole laitonta, mutta toisen ihmisen henkilötietojen väärinkäyttö taas on rangaistavaa useiden eri rikostunnusmerkistöjen perusteella.[112]

Toisaalta esitöissä todetaan, ettei identiteettivarkauden tunnusmerkistö täyty, jos tietojen käyttö on vähäistä tai muutoin sellaista, ettei kolmas osapuoli todellisuudessa voi erehtyä. Syvävääräntämisessä ei siten syyllistytä identiteettivarkauteen ainakaan silloin, kun on selvää, että kyseessä on syvävääräennös. Näin lienee muun muassa silloin, kun tekijä itse ilmoittaa videolla sen olevan väärennös, tai kun video on niin kelvottomasti laadittu, ettei sen aitoudesta voi erehtyä. Erehtymisen vaaraa ei liene myöskään, kun videosta käy selvästi ilmi sen parodinen tai satiirinomainen luonne.[113]

RL 38:9a:n mukaisesta identiteettivarkaudesta ei rangaista, jos teosta puuttuu tarkoitus erehdyttää toista tai jos syvävääräennöksen avulla on muutoin tarkoitus osallistua anonyymisti keskusteluun.[114] Syvävääräennösten osalta anonyymi toimiminen voisi tulla mahdollisesti kyseeseen ainakin silloin, kun kasvoina käytetään sellaisia tekoälyn itse luomia kasvoja, joita ei todellisuudessa ole olemassa.

Säännöksen sanamuodon mukaisesti teon tulee aiheuttaa myös taloudellista vahinkoa tai jotakin muuta vähäistä suurempaa haittaa. Taloudellisena vahinkona pidetään myös aiheutuneen tilanteen korjaamiseksi syntyneitä selvittelykuluja. Usein säännöksen tarkoittama vähäistä suurempi haitta voi syntyä, vaikka uhrille ei aiheutuisi taloudellista vahinkoa. Lain esitöiden mukaan tällainen haitta voi olla esimerkiksi asian selvittelystä ja oikaisemisesta johtuva vaivannäkö tai

[112] Saaripuu 2019 s. 88–89.
[113] HE 232/2014 vp s 37.
[114] HE 232/2014 vp s. 37.

ylipäätään haitta siitä, että asiaa on mahdotonta oikaista. Selvittely voi liittyä esimerkiksi sellaiseen identiteettivarkauteen, jossa varastetun identiteetin avulla on tehty petoksia. Haittaa voi aiheutua aiheettomien laskujen selvittämisestä.[115]

Edelleen esitöissä mainitaan tilanne, jossa sosiaaliseen mediaan on luotu valeprofiili varastetuilla henkilötiedoilla. Tilanteessa identiteettivarkauden kohteeksi joutuneen henkilön oikeus käyttää omissa nimissään sananvapautta on uhattuna. Tällaisen valeprofiilin poistaminen saattaa olla erittäin hankalaa ja lisäksi valeprofiililla käytyjen keskustelujen selvittelystä voi aiheutua merkittävää haittaa. Kuten esitöissä todetaan, nimenomaan tietoverkon osalta suuria haasteita aiheuttaa se, että tietoverkkoon ladattujen tietojen saaminen poistetuksi saattaa olla erittäin vaikeaa.[116] Nämä esiin tulleet ongelmat koskevat myös syvävääreentämistä. Näidenkin osalta todennäköisesti suuria ongelmia tulee tuottamaan tallenteiden poistaminen verkosta.

Yksityisen henkilön toisena henkilönä esiintyminen on lainsäädännön näkökulmasta ongelmallista, ellei siihen liity vahingoittamis- tai hyötymistarkoitusta. Viranomaiselle toisena henkilönä esiintyminen on kriminalisoitu RL 16:8:ssä. Näihin rikoksiin palataan kappaleessa 4.9. Identiteettivarkautta koskevaa säännöstä laadittaessa on arvioitu, että identiteettivarkaudet usein tapahtuvat osana jotakin muuta rangaistavaa käyttäytymistä.[117]

4.3. Petos

Yleisin identiteettivarkauden kanssa yhdessä kulkeva rikos lienee petos, jonka tunnusmerkistöön kuuluu toisen henkilön erehdyttäminen tai erehdyksen hyväksikäyttö. Petossäännöksillä on pyritty luomaan ja ylläpitämään luottamusta, jota vaihdannan osapuolten välillä tarvitaan. Petoksessa tekijäosapuoli käyttää hyväkseen toisen henkilön informaatioepätasapainoa. Tekoon liittyvät myös erehdyksen vallassa suoritettu määräämistoimi sekä teosta aiheutunut taloudellinen vahinko.[118] Petokseen syyllistyy myös se, joka edellä mainitussa tarkoituksessa syöttää,

[115] HE 232/2014 vp s. 37.
[116] HE 232/2014 vp s. 37.
[117] HE 232/2014 vp s. 5.
[118] Lappi-Seppälä ym. 2022 s. 1195–1197.

muuttaa, tuhoaa tai poistaa dataa tai muutoin puuttuu tietojärjestelmän toimintaan ja saa siten aikaan tietojenkäsittelyn lopputuloksen vääristymisen ja tällä tavoin aiheuttaa toiselle taloudellista vahinkoa (RL 36:1.2). Petoksesta tuomitaan sakkoon tai vankeuteen enintään kahdeksi vuodeksi.

Petoksesta on myös törkeä tekomuoto. Törkeyden kokonaisarvioinnin lisäksi kvalifiointiperusteita ovat huomattavan hyödyn tavoittelu, huomattavan tai erityisen tuntuvan vahingon aiheuttaminen ja rikoksen tekeminen vastuulliseen asemaan perustuvaa erityistä luottamusta tai toisen erityistä heikkoutta tai muuta turvatonta tilaa hyväksi käyttämällä. Törkeästä petoksesta rangaistus on vankeutta vähintään neljä kuukautta ja enintään neljä vuotta (RL 36:2). Jos teko on tavoitellun hyödyn, aiheutetun vahingon määrän tai muihin tekoon liittyvien seikkojen osalta vähäinen, voidaan teko katsoa myös lieväksi petokseksi, josta rangaistus on sakkoa (RL 36:3).

Deepfake-teknologian osalta suurin hyöty petosrikoksen tekijälle syntyy helposta erehdyttämismahdollisuudesta, kun rikos voidaan toteuttaa esiintymällä toisena henkilönä. Deepfake-teknologiaa hyödyntävällä tekijällä voi olla mahdollisuus toimia myös ilman ennakolta tallennettua videota, sillä teknologia mahdollistaa yhtä lailla reaaliaikaisen keskustelun väärennetyillä kasvoilla.

Vuoden 2021 alussa ainakin Britannian, Hollannin ja Ukrainan ulkosuhteista vastaavat poliitikot olivat saaneet valevideosoittoja Venäjältä. Videopuheluissa esiintyi Venäjän oppositiojohtaja Aleksei Navalnyin avustaja Leonid Volkov. Videopuheluiden avulla väitetysti väärennetty Volkov oli pyytänyt rahoitusta Navalnyin kampanjalle sekä haukkunut ukrainalaisia poliitikkoja. Epäilyttävien keskustelujen takia monet ottivat yhteyttä aitoon Volkoviin, joka kielsi soittaneensa videopuheluita kyseisten henkilöiden kanssa.[119] Videopuhelun soittajan tarkoitus oli ilmeisesti hankkia taloudellista hyötyä, mutta myös harjoittaa informaatiovaikuttamista. Journalistit totesivat nopeasti, että kyseessä oli pakko olla deepfake-teknolgian käyttö. Jälkikäteen julkistetussa *The Vergen* haastattelussa Volkovina esiintynyt Alexei Stolyarov kuitenkin kertoi, ettei hän todellisuudessa käyttänyt deepfake-teknologiaa.[120] Vaikka videopuhelussa esiintynyt

[119] Asiasta uutisoi esimerkiksi The Guardian (22.4.2021).
[120] Vincent 2021.

Stolyarov ei ollut syvävääreennetty, se on kuitenkin hyvä esimerkki siitä, kuinka valheellista videokuvaa voidaan hyödyntää petosrikoksissa.

Petosta koskevassa oikeudettomuusarvioinnissa voidaan hyödyntää myös siviilioikeuden säännöksiä, jolloin tarkasteltavaksi saattaa tulla esimerkiksi uhrin selonottovelvollisuuden laajuus. Selonottovelvollisuus korostuu liike-elämässä, jossa yleisesti voidaan asettaa tiukempia vaatimuksia esimerkiksi maksusuorituksista vastaavalle päätöksentekijälle. Tämä selonotto- eli huolellisuusvelvollisuus voi koskea myös kansalaisten välisiä suhteita, kuitenkin vain siltä osin, kuin se ei horjuta yhteiselämän edellyttämää perusluottamusta ihmisten välillä.[121] Syvävääreennösten osalta selonottovelvollisuus voisi olla esimerkiksi tilanteessa, jossa videon katsojalla on selvästi syytä epäillä videon aitoutta. Kuitenkin identiteettivarkauden yhdistäminen petokseen niin, että käytössä on hyvin toimiva deepfake-teknologia, voi luoda tilanteen, jossa uhrilla ei välttämättä ole järkevää syytä epäillä toisen osapuolen aitoutta ja vilpitöntä mieltä.

Petoksella erehdyttäminen voi yksityishenkilön osalta kohdistua esimerkiksi oikeussuhteen olemassaoloon. Näin väärennetyllä videolla esiintyvä henkilö voi saada toisen osapuolen luulemaan, että heidän välillään vallitsee oikeussuhde, jonka perusteella tämä on velvollinen täyttämään jonkin tietyn suoritusvelvoitteen.[122] Petoksen rangaistavuuden edellyttämät erehdyksen hyväksi käyttäminen ja erityisesti erehdyttäminen täyttynevät helposti deepfake-videoiden käytön osalta, jos videolla tai videon esittämisen yhteydessä ei selvästi ilmene sen valheellisuus. Tällöin tekijä välittää tahallisesti virheellistä, vaillinaista tai harhaanjohtavaa tietoa, jonka seurauksena asianomistaja erehtyy.[123]

On kuitenkin huomioitava, että petostapauksessa uhri ei ole se henkilö, jonka kuvaa ja/tai ääntä on käytetty syvävääreennöksessä. Uhri on se, jota syvävääreennöksen avulla erehdytetään. Todennäköisesti syvävääreennöstä hyödyntävässä petoksessa on myös kyse identiteettivarkaudesta, jolloin syvävääreennöksessä käytetty henkilö on identiteettivarkauden uhri ja siltä osin asianomistajan asemassa.

[121] Frände ym. 2018 s. 639–642.
[122] Lappi-Seppälä ym. 2022 s. 1195–1197.
[123] Frände ym. 2018 s. 629.

4.4. Yksityiselämää loukkaava tiedon levittäminen ja kunnianloukkaus

Identiteettivarkauksissa ei aina tavoitella taloudellista hyötyä, vaan rikos voi olla myös henkilön maineeseen kohdistuva, josta ei synny suoraa taloudellista vahinkoa. Erilaiset identiteettirikokset ilman taloudellisen hyödyn tavoittelua tai ylipäätään ilman suoranaista vahingoittamistarkoitusta voivat johtaa kuitenkin henkilön yksityisyyden loukkaamiseen. Käytännössä tällainen rikos olisi joko yksityiselämää loukkaava tiedon levittäminen (RL 24:8) tai kunnianloukkaus (24:9). Molemmista kriminalisoinneista on myös olemassa kvalifioitu tekomuoto. Näitä koskeva yksityiselämä ja kunnia on turvattu myös perustuslain (731/1999, PL) 10.1:n nojalla. Näissä rikoksissa on vaakalaudan toisessa päässä sananvapaus.

Tekijä syyllistyy yksityiselämää loukkaavaan tiedon levittämiseen, jos hän oikeudettomasti joukkotiedotusvälineitä käyttämällä tai muutoin lukuisille ihmisille toimittamalla esittää toisen henkilön yksityiselämästä tiedon, vihjauksen tai kuvan tavalla, joka on omiaan aiheuttamaan vahinkoa tai kärsimystä loukatulle taikka muutoin häneen kohdistuvaa halveksuntaa (RL 24:8.1). Kiellettyä ei ole sellaisen tiedon levittäminen, joka esitetään henkilöstä, joka toimii esimerkiksi politiikassa, elinkeinoelämässä tai julkisessa virassa, jos tämä tieto voi vaikuttaa tämän toiminnan arviointiin kyseisessä tehtävässä ja jos asian esittäminen on tarpeen yhteiskunnallisesti merkittävän asian käsittelemiseksi (RL 24:8.2). Mikäli ilmaisu on esitetty yleiseltä kannalta merkittävän asian käsittelemiseksi ja sen esittäminen, huomioiden sen sisältö, toisten oikeudet ja muut olosuhteet eivät selvästi ylitä sitä, mitä voidaan pitää hyväksyttävänä, ei ilmaisua myöskään pidetä yksityiselämää loukkaavana tiedon levittämisenä (RL 24:8.3). Tiedon toimittaminen lukuisten ihmisten saataville voi tapahtua esimerkiksi tietoverkossa.[124]

Yksityisyyden käsite on pulmallinen, sillä yksityiselämän rajat vaihtelevat suhteessa yhteiskuntaan, ihmiskäsitykseen ja teknologiaan.[125] Lähtökohtana voidaan pitää pyrkimystä tarjota ihmiselle suojaa loukkauksilta, jotka uhkaavat tämän etuja ja koskemattomuutta. Tähän liittyy läheisesti yksilön oma identiteetti, johon hänellä on itsemääräämisoikeus.[126] Lähtökohdaksi voisimme siis asettaa henkilön itsemääräämisoikeuden myös syväväärennösten käytössä: luvatta

[124] HE 19/2013 vp s. 40.
[125] Saarenpää 2004 s. 16–17; Saaripuu 2019 s. 34.
[126] Saaripuu 2019 s. 34–35.

tehty video, jolla esitetään uhri RL 24:8.1:n mukaisella tavalla, olisi yksityiselämää loukkaavaa tiedon levittämistä, jos se oikeudettomasti levitetään lukuisille ihmisille.[127] Asiaa on hieman monimutkaisempi silloin, kun videossa esitetään kuvaa julkisuuden henkilöstä tai jos asian esittäminen on yhteiskunnallisesti tarpeellista. Silloin toki herää kysymys siitä, miksi keskustelu on aloitettava tai käytävä nimenomaan syvävääreennöksen avulla.

Periaatteessa säännös suojaa henkilöä tätä koskevien totuudenmukaisten tietojen levittämiseltä. Yksityiselämää voi loukata toisen yksityiselämää koskevan tiedon, mutta myös vihjauksen tai kuvan, levittäminen. Tieto viittaa totuudenmukaiseen informaatioon, kun taas termi vihjaus ei ole yhtä yksiselitteinen. Lain esitöiden mukaan vihjaus tavallisesti synnyttää jonkinlaisen käsityksen tiedon olemassaolosta, vaikka kyseinen vihjaus ei itsessään tätä tietoa sisältäisikään. Tunnusmerkistön mukainen vihjaus on sellainen, että se ”on omiaan johdattamaan vihjauksen vastaanottajan tiettyyn käsitykseen jostakin yksityiselämää koskevasta seikasta.”[128] RL 24:8 on punnintasäännös, jossa vaakakupissa ovat henkilön yksityisyys ja laajemmin sananvapaus. Tämä on poikkeuksellista rikostunnusmerkistöjä koskevassa sääntelyssä[129] ja luo samalla haasteita niin sananvapautta käyttäville yksilöille omien rajojensa hahmottamisessa kuin myös tuomioistuimille asioiden ratkaisussa.

Esitöissä todetaan, että jos vihjaus synnyttää valheellisia mielikuvia, voisi kyse olla myös kunnianloukkauksesta.[130] Suomessa yksityisiksi tarkoitettujen tavallisten pornograafisten videoiden levittäminen ilman kuvatun ihmisen suostumusta on katsottu joko yksityiselämää loukkaavaksi tiedon levittämiseksi tai kunnianloukkaukseksi.[131] Todennäköisemmin syvävääreennösten yhteydessä, oli kyse pornografiasta tai ei, sovellettavaksi tulee kunnianloukkaus eikä yksityiselämää loukkaava tiedon levittäminen videoiden tai äänitteiden valheellisuuden vuoksi.

---

[127] Viestin jakamisena lukuisille ihmiselle pidetään myös tiedon jakamista sosiaalisen median alustoilla. Melander teoksessa Rikosoikeus, kappaleessa 13. RL 24: Yksityisyyden, rauhan ja kunnian loukkaamisrikokset 2022.
[128] HE 19/2013 vp s. 39.
[129] HE 19/2013 vp s. 42.
[130] HE 19/2013 vp s. 40.
[131] Näin on todettu muun muassa Somerikos vai ei -blogisarjassa, jossa Someturva-nimisen sosiaalisen median ongelmiin keskittyvän yhtiön juristit vertailevat Suomen ja Ruotsin lainsäädäntöä sosiaalisen median näkökulmasta.

Kunnianloukkaukseen syyllistyy se, joka esittää toisesta valheellisen tiedon tai vihjauksen niin, että teko on omiaan aiheuttamaan vahinkoa tai kärsimystä loukatulle taikka häneen kohdistuvaa halveksuntaa, tai muutoin edellä mainitulla tavalla halventaa toista (RL 24:9.1). Tämä valheellinen tieto voi olla kirjallisesti tai suullisesti esitettyä. Valheellinen tieto voidaan niin ikään esittää myös esimerkiksi kuvana[132] ja siten todennäköisesti myös videona. Tuomio kunnianloukkauksesta on sakkoa ja törkeästä kunnianloukkauksesta sakkoa tai vankeutta enintään kahdeksi vuodeksi. Teko on törkeä, jos kunnianloukkauksessa aiheutetaan uhrille suurta kärsimystä tai erityisen suurta vahinkoa ja teko on kokonaisuutena arvostellen törkeä (RL 24:10).

Kyse ei ole kuitenkaan kunnianloukkauksesta, jos esitetty tieto on arvostelua, joka kohdistuu henkilön menettelyyn politiikassa, elinkeinoelämässä, julkisessa virassa tai tehtävässä, tieteessä, taiteessa taikka näihin rinnastettavassa julkisessa toiminnassa. Teko ei saa selvästi ylittää sitä, mitä voidaan pitää hyväksyttävänä (RL 24:9.3). Sama koskee tilanteita, joissa ilmaisu on esitetty yleiseltä kannalta merkittävän asian käsittelemiseksi. Tuolloin kuitenkaan ilmaisun esittäminen ei saa selvästi ylittää hyväksyttävyyden rajaa, kun otetaan huomioon sen sisältö, toisten oikeudet sekä muut olosuhteet (RL 24:9.4). Arvostelu on kuitenkin sallittua vain siltä osin, kuin se kohdistuu henkilön poliittiseen toimintaan tai muuhun työhön, eikä hänen henkilöönsä.

Kunnialla tarkoitetaan kaikille kuuluvaa arvonantoa sekä arvostusta, joka voi määrittyä esimerkiksi yksilön ammatin tai toiminnan kautta. Säännös pyrkii suojaamaan ihmisiä kunnianloukkauksesta aiheutuvilta vahingoilta sekä häpeältä.[133] Yksilöstä tehty väärennetty video, kuuluipa sen sisältöön pornoa, aggressiivista käyttäytymistä tai muuta haitallista sisältöä, voi levitessään aiheuttaa haittaa videossa esitetylle henkilölle. Videon leviämisestä voi äärimmäisessä tilanteessa seurata esimerkiksi työpaikan menettäminen ja työn saamisen vaikeutuminen.[134] Microsoftin hiljattain teettämän tutkimuksen mukaan yli 90 prosenttia työnantajista hakee työnhakuprosessissa verkon hakukoneella tietoa työnhakijoista ja käyttää tätä tietoa valitessaan palkattavan työntekijän. Lisäksi tutkimus osoittaa, että yli 77 prosentissa tapauksissa tämä hakukoneella

[132] Melander teoksessa Rikosoikeus, kappaleessa 13. RL 24: Yksityisyyden, rauhan ja kunnian loukkaamisrikokset 2022.
[133] Melander teoksessa Rikosoikeus, kappaleessa 13. RL 24: Yksityisyyden, rauhan ja kunnian loukkaamisrikokset 2022.
[134] Kugler – Pace 2021 s. 15.

löydetty tieto vaikutti negatiivisesti työnhakijoiden työnsaantimahdollisuuksiin. Erityisesti työnantajien valintaan olivat vaikuttaneet haussa löytyneet työnhakijasta otetut epäasialliset kuvat.[135] Vaikka syvävääreennettyjen tallenteiden leviäminen verkossa voi johtaa edellä esitettyihin seurauksiin, on se todellisuudessa varsin epätodennäköistä. Tämä johtuu jo pelkästään TSA:ssa säädetystä rekisteröidyn oikeudesta tietojen poistamiseen.[136] Vaatimus tietojen poistamiseen voikin juuri liittyä työnhakuun tai mainehaittaan, joka on rekisteröidyn kannalta poistamisperusteeksi kelpaava henkilökohtainen erityinen tilanne. Arviointi tällaisen perusteen todellisesta olemassaolosta kuuluu rekisterinpitäjälle.[137]

Toisin kuin yksityiselämää loukkaava tiedon levittäminen, kunnianloukkaus voi tulla kyseeseen myös, vaikka valheellinen tieto tai vihjaus ei päätyisi lukuisten ihmisten tietoon. Vaikka tällaisen väitteen esittäminen loukattavalle henkilölle esimerkiksi kahden kesken ei todennäköisesti voisi olla omiaan aiheuttamaan kunnianloukkaussäännöksen 1 kohdassa tarkoitettua halveksuntaa tai vahinkoa loukatulle, se voisi kuitenkin aiheuttaa kärsimystä, jolloin 1 kohta tulisi sovellettavaksi.[138] Pornograafisten syvävääreennösten osalta kyse olisi todennäköisesti ennemmin seksuaalirikoksesta (esim. RL 20:3:n mukainen kajoaminen tai RL 20:6:n mukainen seksuaalinen ahdistelu). Toki syvävääreennösten osalta lienee todennäköisempää, että loukkaustarkoituksessa tehty tallenne levitetään ennemmin lukuisten ihmisten tietoon, kuin että se lähetettäisiin vain tallenteessa esitetylle henkilölle.

Kyse on kunnianloukkauksesta myös silloin, kun joku esittää kuolleesta henkilöstä valheellisen tiedon tai vihjauksen ja teko on omiaan aiheuttamaan kärsimystä vainajan erityisen läheiselle ihmiselle (RL 24:9.2). Kuolleen henkilön kunnianloukkauksesta on kyse vain perättömän tiedon osalta ja tältä osin henkilön yksityiselämän suoja näyttäisi muutoin loppuvan kuolemaan. Koska perättömänä tietona pidetään myös valheellista kuvaa ja todennäköisesti myös videota, on oletettavaa, että kunnianloukkauksen tunnusmerkistö voi täyttyä aina, kun syvävääreennöksessä esitetään valheellisesti kuvaa edesmenneestä henkilöstä ja kun teko on omiaan aiheuttamaan kärsimystä vainajan erityisen läheiselle ihmiselle. Kunnianloukkauksen kvalifioitu tekomuoto ei

[135] Burt – Horvitz 2020.
[136] Tietosuoja-asetus 17 artikla, ns. ”oikeus tulla unohdetuksi”.
[137] Korpisaari ym. 2022 s. 246, 255.
[138] HE 19/2013 vp s. 45.

voi tulla kyseeseen vainajasta tehdyn väitteen osalta, oli kyse kuinka törkeästä loukkauksesta tahansa.[139] Säännös yksityiselämää loukkaavasta tiedon levittämisestä suojaa vain elossa olevia henkilöitä.[140]

Näiden ilmaisurikosten tunnusmerkistöjen tulee olla yhteensopivia perus- ja ihmisoikeuksien kanssa, minkä vuoksi on olennaista kiinnittää niiden soveltamistilanteessa huomiota myös EIT:n ratkaisukäytäntöön. Oikeuskäytännössä on erityisesti arvioitu, onko ihmisten sananvapauteen puuttuminen ollut välttämätöntä yhteiskunnassa, onko puuttumiseen ollut painava sosiaalinen tarve, onko puuttumisessa otettu huomioon suhteellisuusperiaate ja onko puuttumiselle esitetyt perustelut olleet riittäviä.[141] Tähän näkökulmaan palataan 5. kappaleessa.

4.5. Seksuaalirikokset

Suomen seksuaalirikoslainsäädäntö on kokonaisuudistuksen kohteena ja muutokset astuvat voimaan vuoden 2023 alusta. Muutos koskee RL 20 luvun seksuaalirikossääntelyä.[142] Tarvetta muutoksille on perusteltu monin eri tavoin, mutta yhtenä merkittävänä syynä voidaan pitää viestintäteknologian kehitystä ja sen aiheuttamaa tarvetta suojella ihmisiä ei-toivotulta seksuaaliselta sisällöltä.[143] Ehdotetussa seksuaalirikosuudistuksessa korostuu seksuaalisen itsemääräämisoikeuden ja henkilökohtaisen koskemattomuuden suojan vahvistaminen määrittelemällä keskeiseksi teon perusteeksi vapaaehtoisen osallistumisen puute.[144] Syvävääränneösten näkökulmasta nykyisen RL 20 luvun seksuaalirikossääntelyn painottuminen seksuaalista itsemääräämisoikeutta fyysisesti loukkaaviin tekoihin on ongelmallista. Tulevalla seksuaalirikoslainsäädännöllä voidaan onneksi paremmin puuttua syvävääränneetyn pornografian väärinkäyttöön. RL 20 luvun säännöksistä syvävääränneöksiin soveltuvat ainakin kajoamisrikokset (3–4 §), seksuaalinen ahdistelu (6 §) sekä seksuaalisen kuvan luvaton levittäminen (7 §).

[139] HE 19/2013 vp s. 47.
[140] HE 19/2013 vp s. 38.
[141] HE 19/2013 vp s. 37–38.
[142] HE 13/2022 vp s. 1.
[143] HE 13/2022 vp s. 40.
[144] HE 13/2022 vp s. 50.

RL 20:3:n mukaan seksuaaliseen kajoamiseen syyllistyy se, joka koskettelemalla tai muulla tavoin tekee muun kuin RL 20:1:ssä tarkoitetun seksuaalisen teon sellaiselle henkilölle, joka ei osallistu siihen vapaaehtoisesti, tai saa tämän ryhtymään sellaiseen tekoon, ja teko olennaisesti loukkaa tämän seksuaalista itsemääräämisoikeutta. Rangaistus seksuaalisesta kajoamisesta on vankeutta vähintään neljä kuukautta ja enintään neljä vuotta. Koskettelua lukuun ottamatta teon rangaistavuuden kannalta ei ole lain esitöiden mukaan merkitystä, tehdäänkö teko fyysisessä läheisyydessä vai viestintäteknologian välityksellä.[145]

Kajoamisrikoksen tunnusmerkistön täyttymistä on arvioitava tapauskohtaisesti, jos teko koskee kuvan valmistamista toisesta salaa tai muuten ilman tämän lupaa. Sellaiset seksuaaliset kuvat tai videot, joissa esimerkiksi kuvataan toisen sukupuolielintä tai sukupuoliyhteyttä ilman kuvattujen lupaa, kuuluvat kuitenkin esitöiden mukaan seksuaalista itsemääräämisoikeutta olennaisesti loukkaavan teon määritelmän piiriin.[146] Siispä myös pornograafisissa deepfake-videoissa ja -kuvissa konkretisoituu sellainen ”muu tapa” tehdä seksuaalinen teko henkilölle, jolta ei ole saatu tekoon lupaa. Sellaisen tallenteen luominen, jossa kuvataan henkilö valheellisesti seksuaalisessa kanssakäymisessä toisen henkilön kanssa, voidaan katsoa olevan RL 20:3:ssa tarkoitettu teko, joka olennaisesti loukkaa uhrin itsemääräämisoikeutta. Näin lienee ainakin silloin, kun video tai kuva vaikuttaa autenttiselta.

Olennaista on se, mitä videolla tehdään sen jälkeen, kun se on luotu. On tulkinnanvaraista voiko tunnusmerkistö täyttyä, jos syvävääreennöksen tekijä ei näytä videota kenellekään. Toisaalta varsinkin tällaisessa tapauksessa syvävääreennöksen tekemisestä jääminen kiinni vaikuttaa erittäin epätodennäköiseltä.

Vapaaehtoisuus katsotaan puuttuvan, jos vapaaehtoisuutta ei ole sanallisesti, käytöksellä tai muulla tavalla ilmaistu, tai jos tekoon on pakotettu käyttämällä henkilöön kohdistuvaa väkivaltaa tai uhkausta tai jos tahdon muodostus tai ilmaisu ei ole ollut mahdollista henkilön tiedottomuuden, sairauden, vammaisuuden, pelkotilan, voimakkaan päihtymistilan, heikentyneen tajunnantilan, tilanteen äkillisyyden, erityisen valta-aseman vakavan väärinkäytön tai muun näihin

[145] HE 13/2022 vp s. 108.
[146] HE 13/2022 vp s. 109.

rinnastettavan syyn vuoksi. Myös yritys on rangaistava (RL 20:3.2–3). Syvävääreennöksiä voidaan tehdä henkilöistä, joista kuvamateriaalia on riittävästi saatavilla. Vapaaehtoisuuden kriteeri puuttuu selvästi ainakin silloin, kun syvävääreennöksessä esitetyltä henkilöltä ei ole ollenkaan kysytty lupaa väärennöksen tekemiseen.

Kajoamisesta on olemassa myös kvalifioitu tekomuoto. RL 20:4:n mukaan seksuaalisessa kajoamisessa on kyse törkeästä tekomuodosta, jos siinä käytetään tai uhataan käyttää vakavaa henkilöön kohdistuvaa väkivaltaa tai jos tekijöitä on useita tai muun syyn vuoksi rikoksella aiheutetaan erityisen tuntuvaa henkistä tai ruumiillista kärsimystä tai jos se tehdään erityisen nöyryyttävällä tavalla. Teko voi olla törkeä myös, jos uhri on alle kahdeksantoistavuotias. Lisäksi teon on oltava kokonaisuutena arvioiden törkeä. Rangaistus törkeästä kajoamisesta on vankeutta vähintään yksi ja enintään kuusi vuotta.

Kajoamisrikos, jossa käytetään syvävääreennösteknologiaa, tuskin tulisi monessakaan tilanteessa täyttämään törkeän tekomuodon tunnusmerkistöä. Koska syvävääreennösten tekemiseen ei lähtökohtaisesti liity väkivallan käyttöä, olisi arvioinnissa keskityttävä tekijöiden lukumäärään, mahdollisen henkisen kärsimyksen tai nöyryyttävyyden tasoon ja uhrin ikään. Tallenteella voi olla monia tekijöitä, mutta se on epätodennäköistä ja lisäksi teon tulisi muutoin kokonaisuudessaan olla törkeä. Periaatteessa suurta henkistä kärsimystä ja erityistä nöyryyttävyyttä voisi aiheuttaa tallenne, joka sisältää lapsia tai eläimeen sekaantumista koskevaa materiaalia. Kyseisessä tapauksessa sekä silloin, jos uhri on alle 16-vuotias, nousevat esiin konkurrenssikysymykset, eikä kyse enää välttämättä olisikaan kajoamisesta, vaan esimerkiksi seksuaalisesta kajoamisesta lapseen (RL 20:14) tai sukupuolisiveellisyyttä loukkaavan kuvan levittämisestä (RL 17:18).

RL 20:6 mukaisesta seksuaalisesta ahdistelusta on kyse silloin, kun henkilö tekee toiselle seksuaalisen teon, joka on omiaan loukkaamaan tämän seksuaalista itsemääräämisoikeutta joko koskettelemalla tai kosketteluun voimakkuutensa tai toistuvuutensa vuoksi vakavuudeltaan rinnastettavalla tavalla sanallisesti, lähettämällä tai esittämällä viestin tai kuvan, ottamalla kuvan tai itseään paljastamalla, taikka muulla vastaavalla tavalla. Teosta voidaan tuomita sakkoa tai vankeutta enintään kuusi kuukautta. Esitöistä ilmenee, että säännös kattaa myös seksuaalisten

video- ja äänitallenteiden lähettämisen tai esittämisen toiselle, mikäli teko on omiaan loukkaamaan tämän seksuaalista itsemääräämisoikeutta.[147] Säännöksessä käytetty ilmaus ”muu vastaava tapa” antaa tilaa tulkinnalle, mutta selvää on, että säännöksen soveltuminen edellyttää teoilta riittävää vakavuutta. Tämä siltikin, vaikka kyseessä on muihin seksuaalirikoksiin nähden toissijainen rangaistussäännös ja vaikka kyseessä on rikoslain 20 luvun lievimmin rangaistava rikos.[148] Pornograafisen syvävääreennöksen luomisen arvioinnissa on tapauskohtaisen tarkastelun seurauksena mahdollista päätyä joko kajoamiseen tai seksuaaliseen ahdisteluun. Vakavammissa teoissa sovellettavaksi tulee kajoamisen säännös, kun taas lievemmissä tapauksissa sovelletaan säännöstä seksuaalisesta ahdistelusta.

Seksuaalisen kajoamisen sekä seksuaalisen ahdistelun ero on tulkinnanvarainen. Kajoamisessa viitataan tekoon, joka loukkaa olennaisesti seksuaalista itsemääräämisoikeutta, kun taas ahdistelu koskee sellaisia tekoja, jotka ovat omiaan loukkaamaan toisen seksuaalista itsemääräämisoikeutta. On selvää, että seksuaalinen kajoaminen koskee vakavampia tekoja. Lakimuutoksen valmisteluvaiheessa Suomen Lakimiesliitto (nykyinen Juristiliitto) korosti lausunnossaan näiden kahden säännöksen välisen eron tarkkarajaisuuden merkitystä.[149] Samoin Etelä-Suomen syyttäjäalue sekä Itä-Suomen hovioikeus toivoivat hallitukselta täsmennyksiä ja perusteluja kajoamisen ja ahdistelun rajanvetoon.[150] Konkreettisesti syvävääreennösten tapauksessa arvioitavaksi tulisi luultavasti videon tai kuvan sisällön loukkaavuus, mahdollisesti tekijän ja uhrin välinen suhde sekä teon toistuvuus. Myös sillä, onko syvävääreennös kuva vai video, saattaa olla merkitystä. Pelkkä syvävääreennetty kuva voisi olla seksuaalista ahdistelua, kun taas erittäin arkaluontoinen ja autenttinen syvävääreennetty pornovideo voisi mahdollisesti täyttää seksuaalisen kajoamisen tunnusmerkistön.

Jos pornograafista syvävääreennöstä levitetään esimerkiksi verkossa, voidaan edellä esitettyjen lisäksi soveltaa myös RL 20:7 säännöstä seksuaalisen kuvan luvattomasta levittämisestä. Ensimmäisen momentin mukaan tällaisen kuvan levittämisestä tuomitaan henkilö, joka oikeudettomasti esittää tai levittää toista seksuaalisesti esittävän todellisuuspohjaisen tai todenmukaisen

[147] HE 13/2022 vp s. 117.
[148] HE 13/2022 vp s. 115–119.
[149] OM: julkaisuja 2021:2 s. 36.
[150] OM: julkaisuja 2021:2 s. 38.

kuvan tai kuvatallenteen siten, että teko loukkaa olennaisesti tämän seksuaalista itsemääräämisoikeutta. Olennaista on henkilön tunnistettavuus, joko kasvojen perusteella tai muutoin.[151] Tässä säännöksessä huomio kiinnittyy termeihin todellisuuspohjainen ja todenmukainen. Termeillä tarkoitetaan 2 momentin mukaan kuvaa tai kuvatallennetta, jos se on valmistettu tilanteesta, jossa henkilö tosiasiallisesti esiintyy (todellisuuspohjainen) tai jos se erehdyttävästi muistuttaa vastaavalla menetelmällä valmistettua kuvaa tai kuvatallennetta tilanteesta, jossa henkilö tosiasiallisesti esiintyy (todenmukainen). Syvävääreennösten osalta kyse on todenmukaisuudesta. Loukkauksen olennaisuutta arvioitaessa tallenteen seksuaalisuuden aste sekä se miten laajasti tai millä tavoin kuva, video tai äänite esitetään tai levitetään.[152] Rangaistus seksuaalisen kuvan luvattomasta levittämisestä on sakkoa tai vankeutta enintään kaksi vuotta (RL 20:7.1).

RL 20:7:n esitöiden mukaan ei-todenmukaiset pilakuvat eivät kuulu säännöksen soveltamisalaan, mutta niihin voidaan joissain tapauksissa soveltaa kunnianloukkaussäännöstä.[153] Tulkinnan mukaan seksuaalisen kuvan luvattomasta levittämisestä ei olisi kyse esimerkiksi silloin, kun pornograafisesta syvävääreennöksestä käy ilmi sen epätodenmukaisuus ja taustalla oleva pilailutarkoitus.

Mikäli tekoon sovelletaan seksuaalisen kuvan luvatonta levittämistä koskevaa säännöstä, kattaa kyseisen säännöksen soveltamisala myös tekoja, joista ennen lakiuudistusta rangaistiin kunnianloukkauksena tai yksityiselämää loukkaavana tiedon levittämisenä.[154] Siksi RL 20:7 soveltuessa tapaukseen, ei tekijää voida rangaista myös kunnianloukkauksesta tai yksityiselämää loukkaavasta tiedon levittämisestä.

Seksuaalinen teko määritellään erikseen RL 20:23:ssä. Kyse on seksuaalisesta teosta, jos se on seksuaalisesti olennainen, kun otetaan huomioon tekijä, kohteena oleva henkilö ja teko-olosuhteet. Suurimmassa osassa tapauksista syvävääreennetyn pornografian tekeminen voidaan lähtökohtaisesti lukea seksuaalisesti olennaiseksi teoksi.

[151] HE 13/2022 vp s. 121.
[152] HE 13/2022 vp s. 121.
[153] HE 13/2022 vp s. 121.
[154] HE 13/2022 vp s. 51.

Komissio on julkaissut 8.3.2022 uuden direktiiviehdotuksen naisiin kohdistuvan väkivallan ja lähisuhdeväkivallan torjumisesta (NKV-direktiiviehdotus)[155]. Direktiiviehdotuksessa ehdotetaan vähimmäisvaatimuksia erilaisten sukupuoleen perustuvien väkivallantekojen kriminalisointien lisäksi myös tietyille verkkoväkivallan muodoille. Näihin kuuluvat verkkoahdistelu ja -häirintä sekä luvattomasti tapahtuvan intiimin materiaalin käsittely ja jakaminen. Kuten ehdotuksessa todetaan, tieto- ja viestintäteknologioiden käyttö lisää tiettyjen verkkoväkivallan muotojen helpon, nopean ja laajan leviämisen riskiä, mikä usein aiheuttaa rikoksen uhrille vakavaa ja pitkäaikaista haittaa. Komissio ehdottaakin jäsenvaltioille manipuloidun materiaalin luvattoman levittämisen kriminalisointia. Kyseisenä materiaalina pidettäisiin syvävääreennöksiä, joissa kuvattu materiaali selvästi muistuttaa olemassa olevaa henkilöä, esineitä, paikkoja tai muita objekteja tai tapahtumia, joissa kuvataan toisen henkilön seksuaalisia tekoja ja jotka voisivat muiden mielestä vaikuttaa virheellisesti aidolta tai totuudenmukaisilta. Rikossäännöksen olisi myös katettava tällaisella toiminnalla uhkaaminen.[156]

Uudessa seksuaalirikoslainsäädännössä ei suoraan ole kriminalisoitu manipuloidun materiaalin luvatonta levittämistä. Tässä kappaleessa tehdyn analyysin perusteella vaikuttaa kuitenkin siltä, että uuden seksuaalirikoslainsäädännön pohjalta voidaan myös syvävääreennetyn pornografian tekijät monissa tilanteissa asettaa rikosoikeudelliseen vastuuseen.

4.6. Väärennys

Rikoslain 33 luvussa säädetään todistelurikoksina pidettävistä väärennysrikoksista, jotka heikentävät oikeudenkäyntien luotettavuutta.[157] Henkilö, joka valmistaa väärän asiakirjan tai muun todistuskappaleen tai väärentää sellaisen käytettäväksi harhauttavana todisteena taikka käyttää väärää tai väärennettyä todistuskappaletta tällaisena todisteena, tuomitaan väärennyksestä sakkoon tai enintään kahdeksi vuodeksi vankeuteen (RL 33:1). Säännöksen mukainen väärennyksen kielto viittaa ensisijaisesti todistuskappaleiden ulkoisen luotettavuuden varmistamiseen,

[155] KOM (2022) 105, lopull. Ehdotus Euroopan parlamentin ja neuvoston direktiivi naisiin kohdistuvan väkivallan ja lähisuhdeväkivallan torjumisesta.
[156] KOM (2022) 105, lopull. Ehdotus Euroopan parlamentin ja neuvoston direktiivi naisiin kohdistuvan väkivallan ja lähisuhdeväkivallan torjumisesta.
[157] Saaripuu 2019 s. 93.

todisteiden oikeaperäisyyteen sekä koskemattomuuteen.[158] Esitöiden mukaan sanamuodolla ”käytettäväksi harhauttavana todisteena” pyrittiin siihen, että harmittomat väärennykset jäisivät rangaistavuuden ulkopuolelle.[159]

RL 33:1:n mukaisella todistuskappaleella tarkoitetaan saman luvun 6 §:n määritelmän mukaan muun muassa asiakirjaa ja sen näköisjäljennöstä, ääni- ja kuvatallennetta, laskimen tai muun vastaavan teknisen laitteen tuottamaa tallennetta sekä automaattiseen tietojenkäsittelyyn soveltuvaa tallennetta, jos sitä käytetään tai voidaan käyttää oikeudellisesti merkityksellisenä todisteena oikeuksista, velvoitteista tai tosiasioista (RL 33:6.1). Todistuskappale katsotaan vääräksi, jos se todisteena käytettäessä on omiaan antamaan erehdyttävän kuvan alkuperästään tai sen antajan henkilöllisyydestä (RL 33:6.2). Todistuskappaletta pidetään väärennettynä, jos sen sisältöä on oikeudettomalla tavalla muutettu niin, että muutettu tieto on merkityksellistä jonkin todistelun kannalta (RL 33:6.3).

Väärennyssäännöksessä käytetyllä todistelu-termillä viitataan tuomioistuimessa ja viranomaisessa tapahtuvan todisteiden esittämisen lisäksi myös muihin tilanteisiin, joissa on kysymys tietyn tosiasian vahvistamisesta todisteen avulla.[160] Siten syvävärennöksen valheellinen käyttö todisteena, joko tuomioistuimessa tai muutoin, näyttäisi lähtökohtaisesti täyttävän väärennyksen tunnusmerkistön.

Väärennettyä äänitallennetta on käytetty esimerkiksi eräässä ulkomaalaisessa lapsen huoltajuustapauksessa, jossa lapsen äiti oli väitetysti esittänyt todisteena äänitteen, jolla lapsen isä oli uhkaillut äitiä. Äänite oli kuitenkin niin sanottu *cheap fake*, eli väärennös ei ollut kovinkaan laadukas. Huonosti toteutetun väärennyksen vuoksi äänitettä ei voitu käyttää isää vastaan.[161] Kuitenkin tällaisten yritysten seurauksena laadukkaasti tehdyt deepfake-videot tai -äänitteet voivat mennä tuomioistuimessa läpi, mikä voi johtaa asian käsittelyssä epäoikeudenmukaiseen lopputulokseen ja heikentää yleisellä tasolla luottamusta oikeusjärjestelmään.

[158] HE 66/1988 vp s. 26.
[159] HE 66/1988 vp s. 112.
[160] HE 66/1988 vp s. 109.
[161] Reynolds 2020.

Rikolliset voivat hyödyntää sitä, että liikkeellä on paljon väärennettyä ja vääristettyä tietoa. Vastuu tehdystä rikoksesta tai muusta lainvastaisesta toiminnasta voi olla vältettävissä sillä, että aitoa todistusaineistoa väitetään väärennetyksi. Tätä ilmiötä on kutsuttu ainakin valehtelijan osuudeksi (*liar's dividend*).[162] Yleinen tietoisuus siitä, että kuvien, videoiden ja äänitteiden manipulointi on helppoa ja voi olla yleistä, lisää syytä uskoa väärennösväitettä todeksi. Kyseinen ilmiö voi jatkossa tulla vastaan myös tuomioistuimissa ja erityisesti todistelussa.[163] Todisteena esitetty video voi olla syvävääreennetty tai vaihtoehtoisesti aitoa videotodistetta voidaan väittää syvävääreennökseksi. Haasteita syntyy, jos kumpaakaan väitettä ei voida näyttää toteen. Esitöistä ilmenee, että todistuskappale katsotaan vääräksi, jos se "todisteena käytettäessä on omiaan antamaan erehdyttävän kuvan alkuperästään tai antajansa henkilöllisyydestä, mikä viime kädessä saa olennaisen merkityksensä asiakirjan käyttöyhteydestä."[164]

Todisteen luotettavuutta arvioitaessa pyritään usein etsimään näyttöä tukevaa muuta todistusaineistoa. Jos sähköisenä todisteena esitettävää videota epäillään tai väitetään väärennetyksi, on luonnollista, että tätä väitettä tukevaa muuta todistusaineistoa yritetään löytää muista, toisistaan riippumattomista lähteistä. Usein tilanne voi olla se, että muuta aiheeseen liittyvää informaatiota ei ole saatavilla ja arvio todisteen aitoudesta täytyy tehdä tallenteen sisältämän informaation uskottavuuden perusteella. Tässä arviossa täytyy käyttää perustana yleistä elämänkokemusta sekä huomioida kaikki ne seikat, jotka tallenteesta ovat tiedossa.[165]

RL 33:1:n säännöksen sanamuoto on sellainen, että siihen sisältyy epäaitouden lisäksi erehdyttämistarkoitus. Olennaista siis on, että todistuskappaletta käytetään harhauttavana todisteena. Lisäksi vaatimuksena on, että väärennöstä käytetään hyötymis- tai vahingoittamistarkoituksessa. Siten harmittomat ja vahinkoa aiheuttamattomat väärennykset jäävät rangaistavuuden ulkopuolelle.[166] Samoin ulkopuolelle jäisivät sellaiset todisteena käytettävät syvävääreennökset, joiden esittäjä tuomioistuimessa ei tiedä esittävänsä epäaitoa tallennetta ja mahdollisesti edellä

[162] Citron – Chesney 2019 s. 1758.
[163] Rini 2020 s. 1–2.
[164] HE 66/1988 vp s. 112–113.
[165] Riekkinen 2019 s. 419.
[166] Saaripuu 2019 s. 94.

mainittu tilanne, jossa todistaja kuvailee syvävääreennöksessä näkemäänsä tapahtumaa omana muistonaan.

4.7. Kiristys

Syvävääreentämisen taustalla ei aina välttämättä ole taloudellisen hyödyn tavoittelu. Syvävääreennöksiä voidaan kuitenkin käyttää kiristämiseen niin, että videon tai äänitteen laatija uhkailee uhria syvävääreennetyn tallenteen levittämisellä. Tällöin tekijä voi syyllistyä rikoslain 31 luvun 3 §:ssä säädettyyn kiristykseen, jonka mukaan muulla kuin ryöstön tunnusmerkistössä tarkoitetulla uhkauksella rikoksentekijä tai se, joka hänen puolestaan toimii, pakottaa toisen luopumaan sellaisesta taloudellisesta edusta, johon hänellä ei ole laillista oikeutta. Kiristysuhkaus eroaa ryöstöuhkauksesta siinä, että kiristysuhkaukselle on tunnusomaista muu kuin uhkauksen vakavuus ja välittömyys.[167] Kiristyksestä tuomitaan sakkorangaistus tai vankeutta enintään kaksi vuotta (RL 31:3).

Säännöksessä esitetty sanamuoto ei uhkauksen osalta sisällä vaatimusta laittomuudesta, eli että uhkauksen toteutuminen johtaisi rikokseen tai vaikkapa väärien ja harhaanjohtavien tietojen levittämiseen uhrista. Merkitystä on vain sillä, kokeeko uhri uhkauksen itselleen negatiivisena asiana.[168] Näin ollen, jos kiristäjä uhkaa julkaisevansa uhrista tehdyn syvävääreennöksen saadakseen taloudellista etua, on kyseessä todennäköisesti RL 31:3:n mukainen kiristys, riippumatta siitä, onko itse syvävääreennös laiton vai ei.

Kiristyksestä on olemassa myös törkeä tekomuoto, joka voi täyttyä esimerkiksi silloin, jos rikoksentekijä käyttää häikäilemättömästi hyväksi toisen erityistä heikkoutta tai muuta turvatonta tilaa tai jos taloudellinen etu, josta toinen pakotetaan luopumaan, on erittäin arvokas. Lisäksi teon on oltava kokonaisuudessaan törkeä. Rangaistus törkeästä kiristyksestä on vankeutta vähintään neljä kuukautta ja enintään neljä vuotta (RL 31:4). Molemmissa tekomuodoissa yritys on rangaistavaa.

[167] Frände ym. 2018 s. 598–599.
[168] Frände ym. 2018 s. 599.

4.8. Yleistä järjestystä vastaan tehdyt rikokset

Syvävääreennöksen tekijällä voi myös olla tavoitteena yhteiskuntarauhan horjuttaminen. Videossa voidaan esittää esimerkiksi tunnettu poliitikko kannustamassa kansalaisia nousemaan vastarintaan valtion edustajia vastaan. Karismaattisen poliitikon kannustus saattaa johtaa siihen, että tyytymättömät ihmiset seuraavat aidolta vaikuttavaa auktoriteettia aina väkivaltaisuuksiin saakka. Näin tapahtui tammikuussa vuonna 2021, kun presidentti Donald Trumpin kannattajat valtasivat Yhdysvaltain Capitolin kongressirakennuksen.[169] Trumpin kannustuspuheet eivät olleet väärennettyjä, mutta tapaus kuvastaa hyvin yhden henkilön, oli kyse aidosta tai syvävääreennetystä puhujasta, mahdollisuutta saada huomattava määrä ihmisiä liikkeelle. Yllytystarkoituksessa tehdyn syvävääreennetyn videon seurauksena voi tilanteesta riippuen rangaistuksena olla esimerkiksi julkinen kehottaminen rikokseen, josta tuomitaan sakkoon tai vankeuteen enintään kahdeksi vuodeksi (RL 17:1) tai väkivaltaisen mellakan johtaminen, josta rangaistus on vankeutta enintään neljä vuotta (RL 17:4).

4.9. Oikeudenkäyttöä ja viranomaista vastaan tehdyt rikokset

Syvävääreennösten käytön yleistyessä kasvaa myös vaara niiden väärinkäytön kohdistumisesta valtiollisiin toimijoihin. Rikoslain 15 ja 16 luvuissa säädetään rikoksista oikeudenkäyttöä ja viranomaisia vastaan.

Periaatteessa on mahdollista, että syvävääreennöksen katsomisen seurauksena joku syyllistyisi tuottamukselliseen perättömään lausumaan. RL 15:4:n mukaan tuomioistuimessa esiintyvä todistaja tai asiantuntija tai henkilö, joka valan tai vakuutuksen nojalla tuomioistuimessa tai oikeudenkäyntiin rinnastettavassa viranomaismenettelyssä antaa huolimattomuuttaan väärän tiedon asiassa tai salaa siihen kuuluvan seikan tuomitaan sakkoon tai vankeuteen enintään kuudeksi kuukaudeksi. Tällainen tilanne voisi syntyä esimerkiksi, jos todistaja on nähnyt syvävääreennetyn videon, jonka todenperäisyyttä ei ole itse tarkistanut, mutta jonka pohjalta esittää virheellisen todistajalausunnon tuomioistuimessa.

[169] Lidz 2021 s. 233.

Rikollinen voi luoda syvävääreennöksen, jossa valheellisesti näkyy, kun henkilö tekee rikoksen. Syvävääreennöksessä henkilö voi myös ilmiantaa itsensä tai toisen henkilön. Videon viranomaiselle tai tuomioistuimelle lähettänyt henkilö voidaan tuomita RL 15:6:n mukaisesti väärästä ilmiannosta sakkoon tai vankeuteen enintään kolmeksi vuodeksi, jos väärä tieto aiheuttaa vaaran ilmiannetun pidättämisestä, vangitsemisesta tai muun pakkokeinon kohteeksi joutumisesta tai jos tieto johtaa ilmiannetun syyttämiseen tai rangaistukseen tuomitsemiseen tai muuhun rikosoikeudelliseen seuraamukseen virheellisin perustein.

RL 15:7:ssä säädetään todistusaineiston vääristelemisestä. Todistusaineiston vääristelemisestä tuomitaan se, joka yrittää saada syyttömän ihmisen tuomituksi rangaistukseen tai muulla tavalla vahingoittaakseen ja siksi kätkee, hävittää, turmelee, muuntaa tai muuten vääristää tuomioistuimessa tai rikosjutun esitutkinnassa tarpeellisen esineen, asiakirjan tai muun todisteen. Todisteella on oltava tekijän ymmärryksen mukaan merkitystä asian kannalta. Todistusaineiston vääristelemisestä tuomitaan myös henkilö, joka tietää todisteen olevan perätön tai vääristelty ja antaa sen 1 momentissa esitetyssä tarkoituksessa todisteeksi tai itse käyttää sitä harhauttavalla tavalla tuomioistuimessa tai rikosasian esitutkinnassa (RL 15:7.2). Tuomio kyseisestä rikoksesta on sakkoa tai vankeutta enintään kaksi vuotta. Teko on törkeä, jos siinä aiheutetaan vakava vaara, että syytön tuomitaan ankaraan seuraamukseen, jos rikoksen kohteena on erityisen merkityksellinen todiste tai jos rikos tehdään erityisen suunnitelmallisesti ja kokonaisarvioinnissa teko katsotaan törkeäksi. Kvalifioidun tekomuodon rangaistusasteikko on vankeutta neljästä kuukaudesta kuuteen vuoteen (RL 15:8). Todistusaineiston vääristely saattaa toteuttaa myös RL 15:11:n mukaisen rikoksentekijän suojelemisen tai RL 33:1:n mukaisen väärennyksen tunnusmerkistön.[170]

RL 15:11:n mukaiseen rikoksentekijän suojelemiseen voi syyllistyä esimerkiksi henkilö, jolla on hallussaan alkuperäinen, todisteeksi sopiva video, jonka hän syvävääreentää ja tällä tavalla estää tai yrittää estää rikoksentekijän saattamista vastuuseen rikoksesta. Säännöksen mukaan

[170] Lappi-Seppälä ym. 2022 s. 386.

rikoksentekijää suojelevan henkilön on täytynyt olla tietoinen tapahtuneesta rikoksesta. Rikoksentekijän suojelemisesta tuomitaan sakkoon tai vankeuteen enintään yhdeksi vuodeksi.

RL 16 luvun säännöksistä syvävääreennöksiin soveltuvat väärän henkilötiedon antaminen (5 §) sekä väärän todistuksen antaminen viranomaiselle (8 §). Henkilö, joka erehdyttääkseen viranomaista ilmoittaa nimensä väärin tai antaa muutoin henkilöllisyydestään väärän tai harhaanjohtavan tiedon taikka sanotussa tarkoituksessa käyttää toisen henkilötodistusta, passia, ajokorttia tai muuta sen kaltaista todistusta, tuomitaan väärän henkilötiedon antamisesta. Teosta tuomitaan sakkoon tai vankeuteen enintään kuudeksi kuukaudeksi (RL 16:5). Toisaalta jos henkilötodistusta tai muuta todistusta on jollakin keinolla muutettu, sovelletaan tapaukseen todennäköisesti RL 33 luvun säännöksiä väärennyksestä RL 16:5:n soveltamisen sijasta.[171]

Väärän henkilötiedon antaminen yksityishenkilölle ei ole kuitenkaan edellä esitetyn säännöksen perusteella rangaistavaa. Yksityishenkilöitä koskeva vastaava säännös on ollut aikaisemmin olemassa, mutta se on kumottu vuonna 1999. Aikanaan rikoslain mukaan oli rangaistava sakolla sitä, joka erehdyttääkseen yksityistä henkilöä käytti toisen passia, työtodistusta tai sen kaltaista todistusta.[172] Hallituksen vuonna 2005 tekemän esityksen perusteella kyseinen säännös otsikolla ”Toiselle kuuluvan henkilötodistuksen tai vastaavan asiakirjan käyttäminen yksityisen henkilön erehdyttämiseksi” yritettiin kriminalisoida uudestaan.[173] Eduskunnan lakivaliokunta kuitenkin hylkäsi kriminalisointiehdotuksen. Kriminalisointi olisi lakivaliokunnan mukaan edellyttänyt painavaa yhteiskunnallista tarvetta, sen olisi pitänyt olla ennaltaehkäisevä ja sen olisi tullut täyttää kriminalisointiperiaatteiden vaatimukset.[174] Tämän perusteella esimerkiksi reaaliaikaisen syvävääreennöksen mahdollistavaa teknologiaa hyödyntävä henkilö, joka esiintyy toisena jollekin muulle kuin viranomaiselle, ei syyllisty väärän henkilötiedon antamisen osalta rikokseen.

Väärän todistuksen antamisesta viranomaiselle on kyse silloin, kun viranomaiselle annetaan oikeudellisesti merkityksellinen, mutta totuudenvastainen, kirjallinen todistus tai siihen rinnastettava tekninen tallenne. Samoin rangaistaan sitä, joka laadittuaan sellaisen todistuksen tai

[171] Lappi-Seppälä ym. 2022 s. 406.
[172] HE 169/2005 vp s. 5.
[173] HE 169/2005 vp.
[174] LaVM 15/2005 vp s. 3–4, 6.

tallenteen antaa sen toiselle samaan tarkoitukseen käytettäväksi. Jos muualla laissa ei teosta säädetä ankarampaa rangaistusta, teosta tuomitaan sakkoa tai vankeutta enintään kuusi kuukautta (RL 16:8). Jos teonkuvauksessa täyttyvät esimerkiksi sekä väärän todistuksen antaminen viranomaiselle että petoksen tunnusmerkistöt, tuomitaan teko petoksena eikä väärän todistuksen antamisena.[175] Vaihtoehtoisesti kyse voi ehkä olla myös väärennyksestä.

Vilppiä tuomioistuimissa sekä viranomaisasioinnissa on tapahtunut jo ennen syväväärennösteknologian kehittämistä. Deepfake-videot tuovat kuitenkin uudenlaisia haasteita oikeudenkäyttöön erityisesti niiden vaikean havainnoinnin vuoksi. Kuten tutkielmassa on aiemmin tullut esille, toistaiseksi ei ole kehitetty teknologiaa, joka varmuudella erottaisi syväväärennöksen aidosta videosta. Siten väärennetyt videot aiheuttavat vakavaa huolta audiovisuaalisen materiaalin uskottavuudesta sekä hyväksyttävyydestä sähköisenä todisteena tuomioistuimissa.[176] Ongelma ei kokonaan poistu, vaikka syväväärennösten havainnointiin kehitettäisiin täysin toimiva teknologia. Suullisessa todistelussa todistaja voi kuvailla tuomioistuimen ulkopuolella näkemäänsä tai kuulemaansa syväväärennöstä, eikä videon aitoutta välttämättä ole enää jälkikäteen mahdollista tutkia.

Lisäksi ihmisen muisti on jokseenkin epäluotettava, varsinkin jos käsiteltävästä asiasta on kulunut enemmän aikaa. On mahdollista, että todistaja on nähnyt syväväärennöksen, mutta muistaa nähneensä tapahtuman itse. Usein myös esimerkiksi medialta vastaanotettu informaatio vaikuttaa ihmisen muistiin. Jos syväväärennös saa paljon julkista huomiota, on varsin todennäköistä, että mediaa seuranneille henkilöille (kuten todistajille) syntyy annetun tiedon pohjalta virheellisiä muistikuvia tapahtumista.[177] Sama koskee mitä tahansa internetissä laajalle levinnyttä materiaalia, esimerkiksi sosiaalisessa mediassa tavallisten käyttäjien tekemiä julkaisuja ja niiden sisältöä.

Tuomioistuimen on oikeudenkäymiskaaren (4/1734, OK) 17 luvun 8 §:n mukaan evättävä näyttö muun muassa silloin, kun näyttö voidaan korvata olennaisesti luotettavammalla

[175] Lappi-Seppälä ym. 2022 s. 411.
[176] Ciancaglini ym. 2020.
[177] Esim. Korkman 2020 s. 2713.

todisteella (parhaan todistusaineiston periaate).[178] Jos todisteena on esimerkiksi tallenne, jonka sisällöstä saatu informaatio on ristiriidassa muun olemassa olevan näytön kanssa ja tämä muu näyttö on luotettavampaa, on tuomioistuimen evättävä tallenteen käyttö todisteena. Sama koskee henkilötodistelua, joka tukeutuu nähtyyn syväväärennökseen ja syväväärennöksen pohjalta esitettyyn informaatioon tai vääristyneeseen muistikuvaan, joka on seurausta nähdystä syväväärennöksestä. Joskus kuitenkin tällaisen näytön lisäksi ei ole olemassa muuta näyttöä ja voi olla hankalaa olla huomioimatta todistelua, varsinkin, jos tallenne tai kertomus on uskottava.

RL 16:5:n ja 16:8:n soveltuvuutta rajoittavat lainkonkurrenssikysymykset, minkä seurauksena ne eivät välttämättä todellisuudessa tulisi sovellettavaksi syväväärennösten väärinkäyttötilanteissa. On epävarmaa, voisiko RL 16:5:ssä säädetyn väärän henkilötiedon antamisesta tuomita syväväärennöksen väärinkäyttäjää ylipäätään, koska syväväärennösten osalta tällaisissa tilanteissa lienee aina kyse väärennetyn (henkilö)todistuksen antamisesta, jolloin tekoon soveltuu RL 33:1:n väärennys. Vastaavasti RL 16:5:ssä esitetty väärän todistuksen antaminen viranomaiselle varmaankin useissa tilanteissa korvattaisiin syytteellä väärennyksestä (RL 33:1) tai petoksesta (RL 36:1).

4.10. Vaalivaikuttamisesta ja disinformaatiosta

Yhteiskunnallisen vaikuttavuutensa osalta yksi merkittävimpiä syväväärennösten aiheuttamia seurauksia ovat toistaiseksi olleet erilaiset informaatiovaikuttamiset. Näihin lukeutuvat vaalivaikuttaminen sekä valeuutisten ja muun disinformaation levittäminen eri tarkoituksissa. Mutta voiko syväväärennöksen avulla väärää tietoa levittävä ylipäätään syyllistyä rikokseen? Entä muuttuuko tilanne, jos pyrkimyksenä on vaikuttaa vaalitulokseen vilpillisin keinoin? Lähtökohta on se, ettei valehtelu sellaisenaan ole rangaistavaa, eikä sitä ole myöskään järkevää säätää rangaistavaksi.[179] Toisaalta valehtelu ja salaaminen tietyissä tilanteissa voi olla rangaistavaa, kuten silloin, kun kyse on RL 15 luvun mukaisesta perättömästä lausumasta tai väärästä ilmiannosta, petoksesta (RL 36:1) tai väärän henkilötiedon antamisesta (RL 16:5).

[178] Riekkinen 2019 s. 384.
[179] HE 66/1988 vp s. 109.

Rikoslain 14 luvussa säädetään rikoksista poliittisia oikeuksia vastaan. Näihin kuuluvat vaalirikos (1 §), vaalilahjonta (2 §), vilpillinen äänestäminen (3 §), vaalituloksen vääristäminen (4 §), poliittisten toimintavapauksien loukkaaminen (5 §) sekä kokouksen estäminen (6 §). Näistä vaalirikos, poliittisten toimintavapauksien loukkaaminen sekä kokouksen estäminen vaativat soveltuakseen väkivaltaista tai väkivallalla uhkaavaa vaalivaikuttamista, eivätkä ne siksi lähtökohtaisesti sovellu deepfake-teknologiaa hyväksi käyttävään ja tässä tarkoitettuun vaalivaikuttamiseen. Periaatteessa kyllä syvävääreennöksiä voidaan hyödyntää väkivallalla uhkaamiseen. Tuolloin kuitenkin syvävääreennöksen merkitys rikoksen näkökulmasta jäisi varsin vähäiseksi eikä deepfake-teknologian käyttö sinänsä vaikuttaisi kyseisen tunnusmerkistön täyttymiseen.

Väärän tiedon levittämisessä ei ole myöskään kyse lahjonnasta eikä konkreettisesta äänestämisestä, minkä vuoksi myös vaalilahjonta sekä vilpillinen äänestäminen jäävät tässä soveltuvuuden ulkopuolelle. Edellä mainituista poliittisia oikeuksia koskevista kriminalisoinneista ainoa mahdollisesti soveltuva olisi 4 §:n mukainen vaalituloksen vääristäminen.

Vaalituloksen vääristämisestä on kyse silloin, kun jonkin teon seurauksena yleisen vaalin tai äänestyksen tulos ei ole oikea tai jos sitä ei saada selville. Tämä tapahtuu laskemalla ääniä väärin, turmelemalla, muokkaamalla, hävittämällä tai lisäämällä äänestyslippuja taikka muulla näihin rinnastettavalla tavalla puuttumalla vaalin tai äänestyksen asianmukaiseen toimittamiseen. Rangaistavuuden edellytyksenä on, että tekijän tarkoitus on ollut aikaansaada väärä vaalitulos tai se, että tulosta ei saada selville.[180]

Lain esitöiden mukaan vaalituloksen vääristämisen rangaistavuus ei edellytä sitä, että teko tosiasiallisesti vaikuttaisi äänestyksen tulokseen tehden siitä virheellisen.[181] On epävarmaa, olisiko 4 § sovellettavissa syvävääreennöksellä aiheutettuun vaalituloksen vääristymiseen. Tätä tulee arvioida sen pohjalta, voiko kyseessä olla muihin tunnusmerkistön mukaisiin tekomuotoihin rinnastettava tapa puuttua vaalin tai äänestyksen asianmukaiseen toimittamiseen.

[180] HE 94/1993 vp s. 76.
[181] HE 94/1993 vp s. 80.

RL 14:4:ssa esitetyn ”muun rinnastettavan teon” on liityttävä nimenomaan vaalin tai äänestyksen toimittamiseen eli joko äänestystilanteeseen tai äänten laskentaan.[182] Vaikka vaalituloksen vääristymiseen tähtäävällä syvävääreennöksellä saataisiin aikaan väärä vaalitulos, ei syvävääreennöksiä hyödyntämällä todennäköisesti voida konkreettisesti vaikuttaa itse äänestystilanteeseen tai äänten laskentaan. Uhka poliittisten syvävääreennösten taustalla liittyy nimenomaan äänestyskäyttäytymiseen vaikuttamiseen ja vaikuttaa siltä, ettei rikoslaissa ole keinoja puuttua tällaisiin tekoihin.

Poliittisia oikeuksia koskevan rikoslain 14 luvun esitöissä mainitaan, että silloin, kun teko kohdistuu vaalimainontaan, eikä konkreettiseen äänestämiseen, voivat vahingontekorikossäännökset tulla kyseeseen poliittisia oikeuksia suojaavien säännösten sijasta.[183] Vuonna 2022 ei kuitenkaan vahingontekorikoksia koskevasta luvusta löydy vilpilliseen vaalimainontaan soveltuvia säännöksiä. Asiaa kommentoi myös rikos- ja prosessioikeuden professori Matti Tolvanen Ylen tekemässä haastattelussa 22. tammikuuta 2022.[184] Haastattelussa Tolvanen totesi, ettei ehdokkaita koskevaa mainontaa pahemmin rajoita rikoslaki kuin kuluttajansuojakaan. Tolvanen korosti äänestäjien omaa vastuuta ottaa selvää, onko vaalimainonnassa esitetyt väitteet totuudenmukaisia vai mahdollisesti harhaanjohtavaa tai väärää tietoa. Syvävääreennösten osalta voimmekin siis olettaa, että niiden avulla toteutettu vaalivaikuttaminen ja -manipulointi ei ole kiellettyä, jos tekoon ei voida soveltaa jotakin toista rikossäännöstä (kuten kunnianloukkausta).

Syvävääretämisen avulla disinformaation levittämiseen ja vaalivaikuttamiseen rikoslaki ei siten ota kantaa, ellei videon sisältö täytä jonkin muun rikossäännöksen tunnusmerkistöä. Tämä tarkoittanee sitä, että väärän tiedon, valeuutisten ja informaatiovaikuttamisen osalta vastuu on siirretty tiedon vastaanottajalle. Videon katsojan odotetaan itse ottavan selvää siitä, ovatko videolla esitetyt asiat totta vai ei. Jo tavallisenkin tietovirran osalta lähdekriittisyyden opettaminen olisi ensiarvoisen tärkeää. Videoiden osalta haasteellisuus tulee esiin siinä, että pelkän tietosisällön lisäksi ihmisten olisi pyrittävä arvioimaan myös näkemänsä ja kuulemansa todenmukaisuutta. Varsinkin silloin, kun syvävääreennös on taidokkaasti tehty eikä sitä ole silmämääräisesti

[182] Lappi-Seppälä ym. 2022 s. 365.
[183] HE 94/1993 vp s. 81.
[184] Ylen internetartikkeli: Vaalimainoksessa voi huijata ilman rangaistusta – oikeusoppinut: ”Ei sitä valvo oikein kukaan” (22.1.2022).

(tai edes teknologian avulla) mahdollista todeta väärennetyksi, on kriittinen arviointikykymme koetuksella.

5. TULEVAISUUDEN NÄKYMIÄ

5.1. Sananvapaus vs. oikeus yksityisyyteen

Avoimessa tietoverkossa toimiminen mahdollistaa laajan, reaaliaikaisen ja tehokkaan ilmaisun. Syvävääreennösten luomisessa on perusoikeuksien näkökulmasta kyse ilmaisuvapauden käytöstä, mutta toisaalta väärinkäytettynä väärennetyt videot voivat rikkoa yksilön yksityisyyden suojaa. Kunniaa ja yksityisyyttä loukkaavissa rikoksissa ovat keskinäisessä jännitteessä sananvapaus, yksityisyys ja kunnia.[185] Sananvapautta ja yksityisyyden suojaa ei voida saman aikaisesti toteuttaa kuin tiettyyn rajaan saakka ilman, että toinen perusoikeuksista tulee loukatuksi. Jotta hyväksyttävyyden raja perusoikeuden käyttämiselle voidaan löytää, voidaan ratkaisua pyrkiä löytämään perusoikeuspunninnan kautta. Perusoikeuspunninnassa tavoitellaan perus- ja ihmisoikeuksien optimaalista toteuttamista eli perus- ja ihmisoikeusmyönteisellä tavalla.[186]

Syvävääreennösten avulla tapahtuva ilmaisu saa tukea eurooppalaisista hallinto-oikeuden keskeisistä periaatteista, kuten julkisuusperiaatteesta, ilmaisuvapaudesta, informaation vapaasta liikkuvuudesta sekä informaation muoto- sekä käyttövapaudesta.[187] Sananvapaus kuuluu perus- ja ihmisoikeuksiin ja se on suojattu monissa eri laeissa ja sopimuksissa. Sananvapauden suojasta säädetään perustuslaissa (12 §), YK:n kansalais- ja poliittisia oikeuksia koskevassa sopimuksessa (19 artikla), Euroopan unionin perusoikeuskirjassa (11 artikla) sekä Euroopan ihmisoikeussopimuksessa (10 artikla). Sananvapautta voidaan pitää yhteiskunnan kannalta toimivan demokratian kulmakivenä sekä yksilön näkökulmasta keskeisenä osana itseilmaisua. Sananvapaus suojaa sanallisen ilmaisun lisäksi kuvaa, eleitä, videokuvaa, ääniä sekä jopa fyysistä kosketusta.[188]

Euroopan ihmisoikeustuomioistuimen antamien tuomioiden perusteella Suomessa sananvapautta ja kunnianloukkausta tai yksityiselämää loukkaavaa tiedon levittämistä koskevassa rajanvedossa suomalaiset tuomioistuimet ovat antaneet ihmisoikeuksien näkökulmasta turhan paljon sananvapautta rajoittavia tuomioita. Suomi onkin saanut toista kymmentä EIS 10 artiklan rikkomista koskevaa tuomiota EIT:lta. Tuomioissaan EIT on alleviivannut kysymystä siitä,

[185] Lappi-Seppälä ym. 2022 s. 809.
[186] Matikkala 2012 s. 148.
[187] Pöysti 1999 s. 405; Saaripuu 2019 s. 9.
[188] Matikkala 2012 s. 148.

onko sananvapauden rajoittaminen ollut EIS 10 artiklan sanamuodon mukaisesti ”välttämätöntä demokraattisessa yhteiskunnassa”.[189]

Syvävääreennösten osalta kysymys siis kuuluu: voidaanko ilmaisukeinona toimivan syvävääreentämisen rajoittamista perustella niin, että se on *välttämätöntä* demokraattisessa yhteiskunnassa? Välttämättömyyden tulee EIS 10 artiklan mukaan perustua lisäksi kansalliseen turvallisuuteen, alueelliseen koskemattomuuteen tai yleiseen turvallisuuteen tai rajoitusten tulee olla välttämättömiä epäjärjestyksen tai rikollisuuden estämiseksi, terveyden tai moraalin suojaamiseksi, muiden henkilöiden maineen tai oikeuksien turvaamiseksi, luottamuksellisten tietojen paljastumisen estämiseksi tai tuomioistuinten arvovallan ja puolueettomuuden varmistamiseksi. EIT on ilmaisuvapauden puuttumista koskevassa ratkaisukäytännössään tehnyt myös suhteellisuusarviointia, eli pohtinut, onko puuttuminen ollut suhteellista tavoiteltuihin ja hyväksyttäviin päämääriin nähden. Punninnassa tärkeää on intressien välinen tasapano sekä puuttumisen oikeasuhtaisuus legitiimin päämäärän näkökulmasta.[190]

Suomen perustuslaissa esitetyistä perusoikeuksista jotkut ovat ehdottomia. Suurinta osaa perusoikeuksista, sananvapaus mukaan lukien, voidaan kuitenkin rajoittaa perusoikeuksien yleisten rajoitusedellytysten mukaisesti. Jos esimerkiksi sananvapautta halutaan rajoittaa, on rajoituksessa seuraavien edellytysten täytyttävä: 1) lailla säätämisen vaatimus, 2) täsmällisyys ja tarkkarajaisuus, 3) hyväksyttävyys, 4) ydinalueen koskemattomuus, 5) suhteellisuusvaatimus, 6) oikeusturvavaatimus sekä 7) ihmisoikeusvelvoitteiden noudattamisen vaatimus.[191] Mikäli syvävääreennösten käyttöä halutaan lainsäädännöllä rajoittaa, on edellä mainittujen rajoitusedellytysten täytyttävä. Tämä tuo oman haasteensa lainsäätäjälle.

EIT on ottanut kantaa sananvapauden ja yksityiselämän suojan väliseen tasapainodilemmaan esimerkiksi tapauksessa Hannover v. Saksa (7.2.2012). Ratkaisussa pohdittiin mikä merkitys on annettava sille, liittyykö julkaisu julkisuudessa tapahtuneeseen toimintaan vaiko jonkin yksityishenkilön yksityiselämään. Ratkaisussaan EIT korosti yksityishenkilön saamaa vahvaa

[189] Matikkala 2012 s. 153–162.
[190] HE 19/2013 vp s. 14.
[191] PeVM 25/1994 vp s. 4–5.

yksityiselämän suojaa, vaikka kyse oli saksalaisesta kuninkaallisesta ja siten julkisuuden henkilöstä. Merkitystä annettiin myös sille, kuinka kuvaa tai tekstiä oli esitetty, olivatko julkaistut tiedot todenmukaisia, kuinka vakavasti ja laajasti yksityiselämää oli loukattu ja oliko julkaisusta aiheutunut julkaisussa esitetylle henkilölle seurauksia. Lisäksi huomioitiin tietojen levittämisen laaja-alaisuus sekä erityisesti se, oliko julkaistut kuvat otettu kuvatun henkilön luvalla vai salaa, tai jopa lainvastaisesti.[192]

Syvääärennösten osalta tapauksella on merkitystä ensinnäkin yksityisyyden suojan rikkomisesta julkisuuden henkilön tapauksessa. Lähtökohtaisesti julkisuuden henkilöiden osalta esimerkiksi kunnianloukkauksissa rikoksen tunnusmerkistö ei täyty yhtä helposti kuin tavallisilla henkilöillä. Silti myös julkisuuden henkilöitä koskee vahva yksityisyyden suoja. Huomio on merkityksellinen, sillä toistaiseksi suurin osa verkossa olevista syvääärennöksistä – etenkin syvääärennetyistä pornovideoista – on tehty julkisuudessa esiintyvistä henkilöistä.[193]

Ihmisellä on jonkinlainen luonnollinen tarve yksityisyyteen. Kehittyvä tieto- ja viestintäteknologia lisää informaation käsittelyä, minkä seurauksena tarve yksityisyyteen erityisesti tiedon osalta korostuu.[194] Myös yksityiselämän suoja on vahvasti suojattu perus- ja ihmisoikeus. Suoja kattaa henkilön omaan identiteettiin sisältyviä seikkoja, kuten henkilön nimen ja kuvia. Tähän perusoikeuteen sisältyy myös oikeus elää rauhassa ilman joutumista sellaisen huomion kohteeksi, joka ei ole toivottua.[195] Niin ikään yksityisyyden suojan alle kuuluu oikeus turvalliseen sähköiseen identiteettiin.[196] Sähköinen identiteetti voi kuulua luonnolliselle henkilölle, oikeushenkilölle tai viranomaiselle ja sillä tarkoitetaan informaatiota, joka voidaan teknisesti tai oikeudellisesti luotettavalla tavalla varmistaa, ja jonka perusteella henkilö voidaan tunnistaa sähköisessä toimintaympäristössä.[197]

[192] Hannover v. Saksa (7.2.2012) kohdat 76–80.
[193] Esim. Kugler – Pace 2021 s. 3 ja Yamaoka-Enkerlin 2020 s. 731.
[194] Korpisaari ym. 2022 s. 1.
[195] Alén-Savikko 2018 s. 43.
[196] Saaripuu 2019 s. 67.
[197] Esim. Voutilainen 2009 s. 244.

Oikeus yksityisyyteen on turvattu muun muassa kansalaisoikeuksia ja poliittisia oikeuksia koskevassa kansainvälisessä yleissopimuksessa[198] (17 artikla), ihmisoikeuksien yleismaailmallisessa julistuksessa[199] (12 artikla), EIS:ssa (8 artikla) sekä EU:n perusoikeuskirjassa (7 artikla). Lisäksi TSA:n 85 artiklassa säädetään jäsenvaltioiden velvollisuudesta sovittaa yhteen sananvapaus ja henkilötietojen suoja. Tarve suojella yksityisyyttä sekä identiteettiä korostuu informaatioyhteiskunnassa, jossa turvallisuus ja yksityisyys painottuu fyysisen koskemattomuuden sijasta enemmän informaatioyksityisyyteen, ensin mainittua tietenkään täysin korvaamatta.[200]

Rajanvetoa sananvapauden ja yksityisyyden välillä on pohdittu esimerkiksi yksityiselämää loukkaavaa tiedon levittämistä koskevissa lain esitöissä.[201] Rajoittamisen arvioinnissa on huomioitava levitetyn tiedon merkityksellisyys sekä kuinka syvälle yksityiselämään levitetyllä tiedolla on kajottu. Vaikka itse säännöksessä ei nimenomaisesti mainita kuvan saamaa vahvaa suojaa, esitöiden mukaan kuva nauttii usein laajempaa yksityisyyden suojaa kuin pelkkä teksti. Merkitystä annetaan myös kuvassa (tai videossa) esitettävän henkilön tunnistettavuudelle.[202]

Keskeisessä osassa yksityisyyden ja henkilötietojen suojan kannalta on myös tietosuoja-asetus, jonka voimaantulon myötä realisoitui syvävääränennösten näkökulmasta oleellinen henkilön oikeus tulla unohdetuksi. TSA 17 artiklan mukaan rekisteröidyllä henkilöllä on oikeus saada tietyin edellytyksin tietonsa poistetuksi. Sillä, pyydetäänkö tiedon poistoa hakukoneen julkaisemista hakutuloksista vai internetsivulta, jossa kuva tai video on alun perin julkaistu, on merkitystä.[203] Oikeus tietojen poistamiseen ei ole kuitenkaan absoluuttinen. Samaisessa 17 artiklassa todetaan, ettei rekisterinpitäjän tarvitse poistaa henkilötietoja, jos ne ovat perustellusti tarpeellisia esimerkiksi sananvapauden ja tiedonvälittämisen turvaamiseksi. Olennainen ongelma on, että vaikka alkuperäinen deepfake-video tai muu tallenne poistettaisiin ensijulkaisijan

[198] Kansalaisoikeuksia ja poliittisia oikeuksia koskeva kansainvälinen yleissopimus 8/1976 (KP-sopimus).
[199] Ulkoasiainministeriön ilmoitus Euroopan ihmisoikeussopimuksesta (Yleissopimus ihmisoikeuksien ja perusvapauksien suojaamiseksi) sellaisena kuin se on muutettuna yhdennellätoista pöytäkirjalla 63/1999.
[200] Pöysti Oikeus 1/2000 s. 97; Saaripuu 2019 s. 11.
[201] HE 19/2013 vp s. 37–45.
[202] HE 19/2013 vp s. 43.
[203] Esim. Euroopan Unionin tuomioistuimen tuomio C-131/12 Google Spain SL ja Google Inc. v. Agencia de Protección de Datos (AEPD) ja Mario Costeja González (13.5.2014) ja KHO 2018:112.

internetsivulta, on ulkopuolinen henkilö saattanut ladata tallenteen omalle tietokoneelleen. Saman tallenteen loputonta uudelleen lataamista internetiin on lähes mahdotonta estää.

Tietosuoja-asetuksen tuoma henkilötietojen suoja ei toimi moitteettomasti. Haasteita aiheuttaa esimerkiksi EU:n ulkopuolella sijaitsevat yhtiöt ja niissä tapahtuva tietojenkäsittely. Yhdysvaltalainen Clearview AI on kerännyt yli 30 miljardia kasvokuvaa ihmisitä ympäri maailmaa. Tekoälyyn turvautuvan teknologian avulla Clearview AI on koonnut kuvat avoimista internetlähteistä ja yhtiö myy informaatiota lähinnä liittovaltion sekä osavaltioiden lainvalvontaviranomaisille.[204] Lokakuussa 2022 Ranskan tietosuojavalvontaviranomainen asetti Clearview AI:lle suurimman mahdollisen seuraamusmaksun eli 20 miljoonaa euroa eri rikkomuksista tietosuoja-asetusta koskien.[205] Sekä Clearviewin että muiden vastaavanlaisten tapausten osalta pelkona on, että EU:n tietosuojarikkomukset kiistetään ja seuraamusmaksut jätetään maksamatta. Ongelmallista esimerkiksi yhdysvaltalaisten yritysten osalta on seuraamusmaksujen täytäntöönpanon ontuminen.[206] Clearview AI:n tapaisia yhtiöitä on muitakin. PimEyes on verkkoalusta, joka toisin kuin Clearview AI ei tallenna kaikkia löydettyjä kasvokuvia tietokantaan, vaan etsii kuvan internetistä sen jälkeen, kun palvelun käyttäjä on ladannut etsimänsä henkilön kuvan alustalle.[207] PimEyes-palvelua voi käyttää kuka vain ja käyttö on tiettyyn rajaan asti ilmaista.[208] Palvelun avulla syvävääreennöksen tekijä voi löytää helposti uusia kuvia kenestä tahansa henkilöstä, josta haluaa tehdä syvävääreennöksen. Kuvanetsintäpalvelut siten voivat helpottaa syvävääreennösten tekoa.

Kuvaa henkilöstä pidetään henkilötietona, vaikka kuva tai muu tallenne ei olisi aito tai varmennettu.[209] Jos henkilötieto liittyy tunnistettavissa olevaan henkilöön, on kyse henkilötiedosta. Video, jossa on yhdistetty esimerkiksi yhden henkilön kasvot ja toisen vartalo, sisältää kahden henkilön henkilötietoja, jos molemmat ovat siitä tunnistettavissa. Jos henkilöä ei voida enää

---

[204] Clearview AI.
[205] European Data Protection Board.
[206] Näin arvioitiin Wired-verkkosivustolla julkaistussa artikkelissa ”Clearview Stole My Face and the EU Can’t Do Anything About It: One man’s battle to reclaim his face shows regulators across the bloc are failing to reprimand the US face search engine”.
[207] Meaker 2022.
[208] PimEyes: Face Search Engine Reverse Image Search.
[209] Esim. tietosuojatyöryhmä 2007 s. 6. Tietosuojatyöryhmän lausunto 4/2007 henkilötietojen käsitteestä on laadittu henkilötietodirektiivin aikana, mutta lausunnossa avatut käsitteet ovat pitkälti yhtenevät TSA:n käsitteiden kanssa.

varmuudella tunnistaa syvävääreennöksestä, ei TSA välttämättä sovellu, jos muokkauksen jälkeen kuvassa tai videossa näkyvää henkilöä ei voida enää liittää johonkin todelliseen ihmiseen. Joka tapauksessa tietosuoja-asetusta sovelletaan vain tunnistettavissa olevaan ja elävään ihmiseen. Asetus ei suojaa kuolleen henkilön henkilötietoja.[210]

Jos deepfake-teknologiaa käytetään yhdistämään kahden henkilön kasvot, riippuu TSA:n soveltuminen videoon siitä, ovatko kyseiset henkilöt tunnistettavissa. Jos kumpaakaan henkilöä ei tunnisteta älyteknologialla tehdyn kasvojen yhdistämisen seurauksena, ei tietosuoja-asetus todennäköisesti sovellu. Tilanne on tulkinnanvaraisempi, jos esimerkiksi toisella henkilöllä on helposti tunnistettava syntymämerkki kasvoissa, josta hänen ”osuutensa” videolla voisi tunnistaa. Jos taas syvävääreentämällä luodaan täysin fiktiivinen hahmo, jonka tekemisessä ei ole käytetty kenenkään tiedossa olevan henkilön tunnistetietoja, ei TSA luonnollisestikaan ole sovellettavissa. Asetusta voitaneen soveltaa silloin, jos täysin fiktiivinen hahmo (jonka tekemiseen ei ole käytetty kenenkään todellisen henkilön kuvia) on luotu muistuttamaan todellista henkilöä ja tämä henkilö on tallenteesta tunnistettavissa.[211]

Vaikka sananvapaus on yksilön oikeus, on sen olemassaolo välttämätöntä toimivan demokratian takaamiseksi. Näin ollen kysymys on myös yhteiskunnan oikeushyvästä. Yksityisyys taas on vahvasti yksilöön kohdistuva oikeus. PL 10 §:ssä säädetty yksityiselämän suoja tarkoittaa pääsääntöisesti (eli ellei erikseen muuta ole säädetty) sitä, että henkilöllä on oikeus hallita omia henkilötietoja ja päättää mitä niillä tehdään. Koska henkilöllisyyden ja yksityisyyden suoja on perustavaa laatua oleva oikeus, on myös valtiolla oma roolinsa tämän oikeuden turvaamisessa. Näiden oikeuksien turvaamisen keskiössä on vahva tietoturvallisuuden toteuttaminen.

[210] Tietosuoja-asetus 27 artikla; Korpisaari ym. 2022 s. 58.
[211] Tietosuojatyöryhmä 2007 s. 10–11. Kun selvitetään liittyvätkö jotkin tiedot henkilöön niin, että kyse on henkilötiedoista, on yksi peruste tarkoitus-tekijän täyttyminen. Tarkoitus-tekijä on olemassa, jos tietoja käytetään niin, että käytön tarkoituksena on ”arvioida tai kohdella tietyllä tavalla kyseistä henkilöä tai vaikuttaa hänen asemaansa tai käyttäytymiseensä”.

5.2. Apua uudesta tekoälyasetuksesta ja digipalvelusäädöksestä

Koska syväväärennösteknologian toiminta perustuu tekoälyyn, on unionin tekoälysääntelyllä tärkeä rooli kyseisen teknologian käytön rajojen määrittämisessä. Komissio onkin vuonna 2021 julkistanut ehdotuksensa asetukseksi, jolla säädettäisiin tekoälyn käyttöä.[212] Tekoälyasetuksen tavoitteena on luoda ratkaisuja, jotka mahdollistavat tekoälyn turvallisen ja luotettavan käytön ottaen samalla huomioon unionin kansalaisten arvot ja perusoikeudet. Komission ehdotuksessa esitetään riskinarviointimenetelmä sellaisten ”suuririskisten tekoälyjärjestelmien” määrittelemiseksi, jotka aiheuttavat merkittäviä riskejä ihmisten terveydelle ja turvallisuudelle tai perusoikeuksille.[213] Ehdotuksessa syväväärentämistä ei pidetä niin sanottuna suuririskisenä tekoälyjärjestelmänä, mutta sille esitetään joitakin läpinäkyvyyttä koskevia vähimmäisvaatimuksia.

Asetusehdotuksessa syväväärennöksiä pidetään rajoitetun riskin tekoälyjärjestelminä. Ehdotuksen 52(3) artiklassa säädetään syväväärennöksiä koskevasta läpinäkyvyydestä. Ehdotuksen mukaan läpinäkyvyysvelvoitteita sovellettaisiin järjestelmiin, i) jotka ovat vuorovaikutuksessa ihmisten kanssa, ii) joilla havaitaan tunteita tai määritetään yhteys (sosiaalisiin) luokkiin biometristen tietojen perusteella tai iii) joilla tuotetaan tai manipuloidaan sisältöä (’syväväärennökset’). Jos tekoälyjärjestelmää käytetään tuottamaan tai manipuloimaan sellaista kuva-, ääni- tai videosisältöä, joka selvästi muistuttaa aitoa sisältöä, olisi velvollisuus ilmoittaa, että sisältö on tuotettu automatisoitua teknologiaa käyttäen. Ilmoitusvelvollisuus syntyy, kun ihmiset ovat vuorovaikutuksessa tällaisen tekoälyjärjestelmän kanssa tai, jos heidän tunteitaan tai ominaisuuksiaan tunnistetaan automatisoidun teknologian avulla. Tätä ilmoitusvelvollisuutta sisällön keinotekoisuudesta tai sen manipuloinnista ei kuitenkaan sovelleta, jos ”käyttö on lain mukaan sallittua rikosten paljastamiseksi, estämiseksi ja tutkimiseksi ja rikoksia koskevissa syytetoimissa tai jos se on tarpeen EU:n perusoikeuskirjassa taattua sananvapautta ja taiteen ja tieteen vapautta koskevan oikeuden käyttämiseksi ja edellyttäen, että toteutetaan kolmansien osapuolten oikeuksia ja vapauksia koskevat asianmukaiset suojatoimet.”[214]

[212] KOM (2021) 206, lopull. Ehdotus Euroopan parlamentin ja neuvoston asetus tekoälyä koskevista yhdenmukaistetuista säännöistä (tekoälysäädös) ja tiettyjen unionin säädösten muuttamisesta.
[213] KOM (2021) 206, lopull. s. 14.
[214] KOM (2021) 206, lopull. s. 15–16.

Ehdotuksessa esitetään, että suuririskiseksi tekoälyjärjestelmäksi olisi kuitenkin luokiteltava lainvalvontaviranomaisten käyttämät tekoälyjärjestelmät esimerkiksi silloin, kun niitä käytetään syvävääreennösten havaitsemiseen.[215] Syvävääreennösten havaitsemiseen käytettävän teknologian määrittely suuririskiseksi johtuu siitä, että sen käyttöön liittyy uhka yksilön oikeuksille ja vapauksille.[216] Syvävääreennösten väärinkäytön ehkäisemisen kannalta tarpeellisen ja oleellisen havainnointiteknologian käyttö tulee siis olemaan tiukasti rajoitettua.

Tekoälyasetuksen myötä syvävääreennösteknologian käyttöä ei siis kielletä. Asetuksen mukaan syvävääreennökset eivät ole suuririskisiä järjestelmiä, mutta suuririskisyys on tulkinnanvaraista ja ajatus on varsin helposti haastettavissa. Edellä esitettyjen esimerkkien perusteella syvävääreennöksillä voidaan saada aikaan suuriakin ongelmia, mikä perustelee niiden väärinkäytön suuria riskejä. Asetusehdotuksessa ei kerrota tarkemmin vaadittavasta merkinnästä syvävääreennöksissä eikä myöskään siitä, kenen vastuulla ylipäätään tällainen merkitseminen on. Ehdotuksessa ei määrätä seuraamuksia syvävääreennöksen merkitsemisen laiminlyönnistä, vaan rangaistuskäytännöt on jätetty jäsenvaltioiden päätettäväksi (artikla 71). Tällaisen merkintävelvollisuuden käyttöönotto EU-alueella on erinomainen asia teknologian väärinkäytön näkökulmasta, jos se pystytään toteuttamaan toivotulla tavalla.

Kuten EPRS:n syvävääreennöksiä koskevassa tutkimuksessakin todetaan, syvävääreennösteknologiaa käyttävät rikoksentekijät toimivat usein anonyymisti eikä heidän tunnistamisensa ole helppoa. Voitaneen olettaa, etteivät kyseiset tekijät ole halukkaita noudattamaan tekoälyasetuksessa säädettyä merkintävaatimusta.[217] Jos merkitsemisvelvollisuudesta seuraa olettamus, että kaikki verkossa vastaan tulevat syvävääreennökset on merkitty, saattaa ihmisten kriittisyys videoita kohtaan vähentyä. Silloin merkintävaatimuksen laiminlyövät rikolliset ovat vahvemmassa asemassa näitä videoita hyödyntäessään eri rikoksissa. Jos syvävääreennöksiä merkitsevät vain ne, jotka käyttävät niitä lailliseen tarkoitukseen, hankaloituu teknologian sallittu käyttö ikään kuin tarpeettomien lainsäädäntörajoitusten vuoksi.

[215] KOM (2021) 206, lopull. s. 29.
[216] EPRS 2021 s. 38.
[217] EPRS 2021 s. VII.

Euroopan unionissa tuli marraskuussa 2022 voimaan digipalvelusäädös[218], joka myös muuttaa direktiiviä sähköisestä kaupankäynnistä[219]. Koska digipalvelusäädös koskee sosiaalisen median alustojen sisältöä, sillä on merkitystä myös syväväärennösten levittämisen kannalta. Euroopan parlamentti oli tehnyt ensimmäisessä käsittelyssä tarkistuksen komission säädösehdotukseen[220], jossa se muun muassa ehdotti omaa artiklaa syväväärennöksille (artikla 30a). Artiklassa esitetään, että erittäin suurilla verkkoalustoilla[221] olisi velvollisuus merkitä alustallaan näkyvissä oleva kuva-, ääni- tai videosisältö, joka on tuotettu tai manipuloitu muistuttamaan olemassa olevia henkilöitä, esineitä, paikkoja tai muita kokonaisuuksia tai tapahtumia ja joka lisäksi antaisi henkilölle virheellisen vaikutelman tallenteen aitoudesta ja todenmukaisuudesta. [222] Lopulliseen digisäädöksen versioon tätä artiklaa ei kuitenkaan otettu.

On odotettavissa, että uusi säädös tulee silti tehostamaan laittoman sisällön – kuten joissakin tapauksissa syväväärennösten – tunnistamista, ilmiantamista, poistamista sekä uudelleenlevittämisen estämistä. Toisaalta säädöksellä tavoitellaan myös läpinäkyvyyttä sisällön ilmianto- ja poistomenettelyssä. Syväväärennösten kannalta on ongelmallista, ettei uuden säädöspaketin perusteella ole pääteltävissä minkälainen sisältö luokitellaan lainvastaiseksi.[223] Myös sananvapauden ja tiedonvälityksen turvaaminen laittoman materiaalin poistomenettelyssä tulee olemaan haasteellista. Säädöksen johdanto-osan 22 kohdan mukaan palvelun tarjoaja voi tulla tietoiseksi laittomasta sisällöstä joko oma-aloitteisesta tutkimuksesta tai joissakin tilanteissa yksityishenkilöiden tai yhteisöjen sille säädöksen mukaisesti tekemistä ilmoituksista.

---

[218] Euroopan parlamentin ja neuvoston asetus (EU) 2022/2065, annettu 19 päivänä lokakuuta 2022, digitaalisten palvelujen sisämarkkinoista ja direktiivin 2000/31/EY muuttamisesta (digipalvelusäädös).
[219] Euroopan parlamentin ja neuvoston direktiivi 2000/31/EY, annettu 8 päivänä kesäkuuta 2000, tietoyhteiskunnan palveluja, erityisesti sähköistä kaupankäyntiä, sisämarkkinoilla koskevista tietyistä oikeudellisista näkökohdista (direktiivi sähköisestä kaupankäynnistä).
[220] KOM (2020) 852, lopull. Ehdotus Euroopan parlamentin ja neuvoston asetukseksi digitaalisten palvelujen sisämarkkinoista (digipalvelusäädös) ja direktiivin 2000/31/EY muuttamisesta.
[221] Erittäin suuret verkkoalustat on määritelty digipalvelusäädöksen 33 artiklassa. Verkkoalusta on erittäin suuri, jos sen unionissa olevien palvelun aktiivisten vastaanottajien kuukausittainen keskimäärä on vähintään 45 miljoonaa.
[222] P9_TA(2022)0014, Euroopan parlamentin tarkistukset 20. tammikuuta 2022 ehdotukseen Euroopan parlamentin ja neuvoston asetukseksi digitaalisten palvelujen sisämarkkinoista (digipalvelusäädös) ja direktiivin 2000/31/EY muuttamisesta, tarkistus 399.
[223] EPRS 2021 s. 42.

Digipalvelusäädöksessä säädetään myös riskien vähentämisestä. Säädöksen 35 artiklan mukaan erittäin suurten verkkoalustojen sekä erittäin suurten verkossa toimivien hakukoneiden tarjoajien on otettava käyttöön riskinvähentämistoimenpiteitä, joilla vastataan tunnistettuihin riskeihin niiden järjestelmissä. Artikla velvoittaa huomion kiinnittämistä erityisesti tällaisten toimenpiteiden perusoikeuksiin kohdistuviin vaikutuksiin. Syvävääreennöksiä koskeva toimenpide on 35 artiklan 1 kohdan k alakohdan mukaisesti varmistaa, että ”tietty tieto on erotettavissa selvästi erottuvien merkintöjen avulla, kun se esitetään niiden verkkorajapinnoilla, jos se on tuotettua tai manipuloitua kuva-, ääni- tai videosisältöä, joka selvästi muistuttaa olemassa olevia henkilöitä, esineitä, paikkoja tai muita tahoja tai tapahtumia ja antaa henkilölle valheellisen vaikutelman siitä, että se on aito tai totuudenmukainen; lisäksi sellaisen helppokäyttöisen toiminnon tarjoaminen, jonka avulla palvelun vastaanottajien on mahdollista ilmoittaa tällaisista tiedoista.”

Syvävääreennösten näkökulmasta digipalvelusäädöksessä esitetyt manipuloitua materiaalia koskevat ratkaisut ovat siis pitkälti samat kuin tekoälyasetuksessa. Ne jakavat myös samat ongelmat: manipuloitua materiaalia merkitsevät todennäköisesti ne henkilöt, jotka eivät muutenkaan väärinkäytä deepfake-teknologiaa. Samoin materiaalin toteaminen laittomaksi on haasteellista ja palvelun tarjoajat joutuvat hankaliin tilanteisiin joutuessaan arvioimaan materiaalin poistamisen ja sananvapauden turvaamisen välistä rajanvetoa.

Unionin tasoisista säädöksistä emme löydä täysin moitteettomia ratkaisuja laittomien syvävääreennösten teon ja jakamisen estämiseksi.

5.3. Rikosoikeudellisen sääntelyn riittävyydestä

Edellä läpikäytyjen rikosnimikkeiden osalta huomataan, että mahdollisia syvävääreentämistä koskevia rikoksia on monia ja että olemassa oleva rikoslakimme soveltuu varsin kattavasti teknologian väärinkäytön ehkäisyyn ja väärinkäytöstä rankaisemiseen. Kaikkiin tässä tutkielmassa esitettyihin skenaarioihin ei ole kuitenkaan mahdollista puuttua rikoslain puitteissa. Haasteita luovat erityisesti videot, joiden tekijöiden tarkoituksena on levittää valeuutisia tai muuta väärennettyä tietoa tai vaikuttaa vaalikäyttäytymiseen. Vaikuttaisi myös siltä, että kuolleista henkilöistä voi tehdä monenlaisia syvävääreennöksiä ilman, että niistä joutuisi rikosoikeudelliseen

vastuuseen. Ainoa kyseeseen tuleva säännös on edesmenneiden ihmisten osalta kunnianloukkaussäännös. Edelleen deepfake-teknologian uhriksi voivat joutua lapset ja nuoret erilaisissa kiusaamistapauksissa, jolloin tekojen kriminalisoinneista ei pääsääntöisesti ole apua, jos tekijä on alle 15-vuotias.

Soveltuvista rikoslain säännöksistä ilmenee, että monesti rikostunnusmerkistö täyttyy vasta, jos syväväärentämistä koskevaan väärinkäyttöön liittyy oman hyödyn tavoittelu tai tarkoitus vahingoittaa toista. Pelkkää kiusantekoa, pilailua tai trollausta ei vielä välttämättä pidetä rikoksena. Tarkasteltaessa syväväärennöksiin soveltuvia rikoslain pykäliä huomataan myös, että tunnusmerkistökuvauksia yhdistää usein väärennöstä hyödyntävä aktiivinen tekeminen. Rikoslain näkökulmasta syväväärennöksen luominen ei itsessään ole rikos, jos tallenne tehdään omaan käyttöön eikä sitä laajalti levitetä tai muuten näytetä muille. Rikoksen tunnusmerkistö voi siis täyttyä vasta, kun luotua väärennöstä käytetään väärin. Lähtökohtaisesti kriminalisoinnit eivät ole sovellettavissa, jos tallenteesta selviää sen epäaitous ja valheellisuus. Esimerkiksi silloin kyse ei voi olla identiteettivarkaudesta, petoksesta tai väärennyksestä. Selvästi epäaidon syväväärennöksen väärinkäyttö ei todennäköisesti voi johtaa myöskään tuomioon julkisesta kehottamisesta rikokseen tai väkivaltaisen mellakan johtamisesta.

Selvästi epäaidon tallenteen käyttö eri tilanteissa voi johtaa rikokseen esimerkiksi, jos kyseessä olevalla videolla kiristetään (kiristys) toista tai jos video halventaa siinä näkyvää henkilöä (kunnianloukkaus). Seksuaalirikosten näkökulmasta selvästi epäaidon syväväärennöksen teko tai levittäminen voi johtaa uuden seksuaalirikoslainsäädännön perusteella rangaistukseen seksuaalisena kajoamisena tai seksuaalisena ahdisteluna. Teoilta vaaditaan kuitenkin riittävää vakavuutta, jotta säännöksiä voidaan soveltaa. Seksuaalisen kuvan luvaton levittäminen ei voi tulla kyseeseen, jos tallenne ei ole todenmukainen.

Suurin osa verkossa olevista deepfake-videoista sisältää seksuaalista materiaalia. Vuonna 2019 on arvioitu, että kaikista olemassa olevista deepfake-videoista 90–95 prosenttia sisältää suostumuksen vastaista pornoa. Syväväärennettyä pornoa varten luoduilla nettisivuilla olevista

syvävääreennetyistä videoista 100 prosenttia oli tehty naisista.[224] Nämä sekä muut[225] tutkimustulokset osoittavat, että deepfake-videot luovat uudenlaista seksuaalisen hyväksikäytön uhkaa.

Yhdysvaltalaisessa Northwestern University Pritzker School of Law'ssa järjestettiin tutkimus, jossa selvitettiin tavallisten ihmisten asenteita syvävääreennöksiä kohtaan sekä mielipiteitä siitä, kuinka syvävääreentäminen tulisi ottaa huomioon lainsäädännössä. Tutkimuksen mukaan ihmiset eivät lähtökohtaisesti kokeneet syvävääreennöksiä vahingollisiksi silloin, kun niissä oli selvästi ilmoitettu videon olevan vääreennös. Ilmoitus vääreennyksestä ei kuitenkaan vähentänyt videon vahingollisuutta silloin, kun kyseessä oli pornograafinen video. Toisaalta myöskään sillä, oliko video laadittu siinä näkyvän henkilön luvalla, ei näyttänyt olevan suurta merkitystä.[226]

Tämän tutkielman tekohetkellä Suomessa (valheellisten) pornovideoiden levittäminen voitaisiin tuomita ainakin kunnianloukkauksena ja mahdollisesti yksityiselämää loukkaavana tiedon levittämisenä. Näistä perusmuotoisten tekojen rangaistuksena tuomitaan sakkoa. Ajatus tällaisen videon leviämisestä ja tieto siitä, ettei videota saa todennäköisesti lopullisesti poistettua tietoverkosta, on epämiellyttävä. Seksuaalirikoslainsäädännön uudistus onkin sopiva lisä syvävääreennösten väärinkäyttöön reagoimiseksi juuri pornografisten vääreennösten osalta. RL 20 luvun uusista säännöksistä syvävääreennöksiin soveltuvat ainakin kajoamisrikokset, seksuaalinen ahdistelu sekä seksuaalisen kuvan luvaton levittäminen. Näitä seksuaalirikosten kriminalisointeja sovellettaessa rangaistusasteikko on korkeimmillaan jopa kuusi vuotta vankeutta. Toisaalta edelleen ainakin seksuaalisen kuvan luvattomassa levittämisessä soveltamisalan ulkopuolelle jäävät pilailutarkoituksessa tehdyt pornograafiset videot, jos on selvää, että video ei ole aito. Tällaisessakin tilanteessa uhrille voi aiheutua psyykkistä haittaa. Tapaukseen voisi kuitenkin todennäköisesti soveltua kunnianloukkaussäännös. Joka tapauksessa muutokset parantavat syvävääreennöksellä aiheutetun seksuaalirikoksen uhriksi joutuneen asemaa. Uusi seksuaalirikoslainsäädäntö kattaa varsin hyvin kostopornoon liittyvät teot.

[224] Ajder ym. 2019 s. 1–2.
[225] Myös esim. EPRS:n teettämässä tutkimuksessa Tackling deepfakes in European policy korostetaan nimenomaan (naisiin kohdistuvien) seksuaalisten syvävääreennösten aiheuttamia ongelmia, kuten kostoporno, seksillä kiristäminen ja muu seksuaalinen hyväksikäyttö.
[226] Kugler – Pace 2021 s. 29–51.

Disinformaatio on ollut yhteiskunnallinen ongelma jo aikana ennen syvävääreentämistä. Väärän tiedon (tahallisen) levittämisen kasvu ja kehitys on aiheuttanut huolta yhteiskunnassa. On tiedossa, että disinformaatioon syyllistyvät niin valtiolliset kuin ei-valtiollisetkin toimijat, ja toimien taustalla piilevät usein poliittiset, ideologiset, kaupalliset tai muut vastaavat syyt.[227] Hybridi- ja informaatiovaikuttamisen osalta kyse on usein eräänlaisista häirintätoimista, eivätkä toimet välttämättä pelkästään valheellisen tiedon levittämisen osalta täytä minkään rikoksen tunnusmerkistöä.[228] On kuitenkin huomioitava, että erilaisessa hybridivaikuttamisessa myös hyödynnetään sellaisia mahdollisuuksia häiritä ja vaikuttaa, joita ei ole osattu huomioida vielä viranomaistoiminnassa tai lainsäädännössä. Tämän vuoksi olisikin tärkeää, että lainsäädännössä otettaisiin huomioon ennakollisesti erilaiset hybridivaikuttamistilanteet.[229] On perusteltua ajatella, että lainsäädännön avulla voimme ennakollisesti varautua laajavaikutteisiin sekä väärinkäytettynä varsin haitallisiin deepfake-teknologian mukanaan tuomiin ongelmiin.

Tässä tutkielmassa hyödynnetyssä lähdeaineistossa on keskustelun ulkopuolelle pääosin jäänyt nuorten syvävääreennösten käyttö etenkin kiusaamistilanteissa. Kuten aiemmin todettiin, syvävääreennöksiä voi tehdä jo nyt monella suositulla älypuhelimeen ladattavalla sovelluksella. Myös syvävääreennösten tekemistä varten luotuja sovelluksia voi ladata helposti puhelimeen. Nuorilla on mahdollisuus tehdä kiusaamismielessä ilkeitä ja suurta haittaa aiheuttavia syvävääreennöksiä, jotka rikosoikeudellisessa vastuussa olevan ihmisen kohdalla voisivat johtaa moniin eri rikossyytteisiin. Ennaltaehkäiseviä ratkaisuja on löydyttävä muualta kuin rikoslainsäädännöstä, jotta alle 15-vuotiaiden osalta vakavat kiusaamiset syvävääreennöksillä voidaan estää.

Yleisellä tasolla rikosoikeudellisen sääntelyn riittävyyttä arvioitaessa on muistettava ultima ratio -periaate, joka kuvastaa kriminaalipoliittisen puuttumisen viimekätisyyttä.[230] Syvävääreennösteknologian väärinkäytön erilaisin keinoin estämisessä sen kriminalisoinnilla on ongelmana se, että niin deepfake-videoiden tekijöitä kuin videoiden levittäjiäkin on varsin ongelmallista

[227] Avoimen ja turvallisen internetin puolesta toimiva, hallitustenvälinen *Freedom Online Coalition* (FOC) on vuoden 2020 lopulla julkaissut kannanoton disinformaation levittämisestä ja ihmisoikeuksista. Tämän kannanoton tavoitteena on FOC:n 32 jäsenmaan, ml. Suomen, yhteistyössä torjua disinformaation käyttöä ihmisoikeuksien, demokratian ja oikeusvaltion heikentämiseen.
[228] HaVM 19/2021 s. 14–15.
[229] HaVM 19/2021 s. 15.
[230] Lahti 2012 s. 103.

saada kiinni. Jotta kriminalisointi toimisi syvävääreennösteknologian väärinkäyttöä ennaltaehkäisevästi, pitäisi videoiden laadinnan sekä levittämisen valvonta ja kontrollointi olla varmalla tasolla. Soveltuipa siis rikoslainsäädäntö syvävääreennöksiin kuinka hyvin tahansa, jää keskeiseksi ongelmaksi niiden havainnointiin tarkoitetun teknologian puutteellinen toimivuus ja syvävääreennöksen tekijän tai levittäjän vaikea jäljitettävyys.

5.4. Muita ratkaisuehdotuksia

Koko ajan kehittyvä älyteknologia on nykypäivää ja tulevaisuutta. Tekoälyä yleisesti voidaan hyödyntää laajasti monissa eri yhteyksissä, kuten teollisuudessa, markkinoinnissa, lääketieteessä, suunnittelussa ja peleissä. Syvävääreennöksiä taas voidaan hyödyntää esimerkiksi elokuvateollisuudessa, opetuksessa, yhteiskunnallisessa vaikuttamisessa sekä muussa itseilmaisussa. Siispä kannattaa tukeutua vanhaan lakimieslatinan ilmaisuun *abusus non tollit usum* eli väärinkäyttö ei sulje pois käyttöä. Eihän ole järkeä kieltää esimerkiksi puukkojen ja muiden teräaseidenkaan käyttöä kokonaan, vaikka niitä käytetäänkin usein tekovälineinä vakavissa väkivaltarikoksissa.[231] Siksi on olennaista keskittyä siihen, kuinka voimme elää tällaisen teknologian kanssa, kuitenkin riskit minimoiden.

Syventyminen syvävääreennösten maailmaan ja kasvava tietoisuus niiden väärinkäytön mahdollisuuksista voivat aiheuttaa epätoivoa. Ongelmiin on silti löydettävissä monenlaisia ratkaisuja. Teknologian näkökulmasta yksi komissionkin ehdotuksista on vesileimojen käyttö deepfake-videoissa.[232] Tämä takaisi ainakin sen, että vilpittömässä mielessä ja laillisesti tehdyt syvävääreennökset eivät aiheuttaisi hämmennystä ja väärän informaation leviämistä. Tässäkin on kuitenkin omat ongelmansa, kuten aikaisemmin tuli esille.

Yksi keino vähentää syvävääreennösten väärinkäyttöä olisi estää niiden luomiseen ja levitykseen liittyvä anonyymisyys. Teknologialähtöinen ratkaisu voisi löytyä vahvasta tunnistautumisesta tai esimerkiksi lohkoketjuista. Näiden avulla laittomien syvävääreennösten tekijät olisi helpompaa saada kiinni. Vahva sähköinen tunnistus edellyttää, että sen käyttäjällä on kaksi hallussa

[231] Tuori 2007 s. 15.
[232] KOM (2021) 206, lopull. s. 15–16.

olevaa tietoa tai ominaisuutta.[233] Näitä voivat olla esimerkiksi tunnusluku, fyysinen ominaisuus (esim. iiris) ja salasana.[234] Lohkoketjuteknologiaa hyödyntämällä taas voidaan varmistua tiedon alkuperästä ja suojata sitä ei-toivotulta muokkaamiselta.[235]

Toisaalta anonyymiin verkkoviestintään puuttuminen on mutkikas ratkaisu. Kiinassa verkkoalustojen käyttäjien tulee rekisteröityä sivustoille omilla henkilötiedoilla, ennen kuin he voivat käyttää kyseisiä palveluita. Näin ihmisten käyttäytymistä verkossa voidaan valvoa ja laittomuuksiin on helpompaa puuttua. Tällaiset rajoitukset voivat vaikuttaa yksilön sananvapauden käyttöön, vaikka se tapahtuisikin sallituissa rajoissa. Anonyymisyys suojaa myös esimerkiksi aktivisteja ja henkilöitä, jotka tuovat laitonta toimintaa päivänvaloon (*whistleblowing*).[236]

Yksi ratkaisu on siirtää vastuuta osittain myös verkkoalustoille. Alustoilta voidaan vaatia esimerkiksi syväväärennösten ja disinformaation tunnistusmenetelmien käyttöönottoa. Samoin vastuu voi syntyä vaatimuksesta merkata jo tunnistetut syväväärennökset tai velvoitetta poistaa alustalta sellaiset tunnistetut syväväärennökset, joita ei ole merkattu väärennöksiksi.[237] Näin on yleisesti jonkin verran jo tehtykin, sillä lähtökohtaisesti verkkoalustoilla on velvollisuus poistaa sivuiltaan selvästi laiton julkaisu. Edelleen syntyy tilanne, jossa vastakkain ovat ilmaisuvapaus sekä sananvapaus. Sisällön laittomuusarvioinnin siirtäminen alustoille on ongelmallista, sillä se saattaa johtaa oikeudettomaan sananvapauden rajoittamiseen. Ylipäätään näin tulkinnanvaraisen harkintavallan ja ”tuomarin viitan” siirtäminen verkkoalustojen ylläpitäjille on mutkikas asia.

Globaalissa verkkoympäristössä jatkuvan muutoksen kohteena ovat niin ihmisten oikeudet ja vapaudet kuin siellä toimimiseen liittyvät velvollisuudet sekä riskit. Lainsäädännön uudistushankkeissa on huomioitava myös teknologia-alan nopeasti muuttuva toimintaympäristö. Teknologian kehittyessä uusia ongelmia saattaa syntyä nopeasti. Tämän vuoksi on ensiarvoisen tärkeää kiinnittää huomiota lainsäädännön teknologianeutraalisuuteen. Nopeita muutoksia

[233] HE 36/2009 vp s. 41.
[234] Liikenne ja viestintävirasto Traficom.
[235] Niemi 2021 s. 126.
[236] EPRS 2021 s. 62.
[237] EPRS 2021 s. 62.

tasapainottamassa ovat kuitenkin yleiset oikeusperiaatteet, jotka kestävät eri tavalla aikaa.[238] Syvävääreennösten käyttö väärennysrikosten tekovälineenä on yksi isoimmista huolenaiheista deepfake-teknologiaa koskien. EPRS:n tekemässä tutkimuksessa esitetään yhtenä vaihtoehtona sitä, että tuomioistuimissa luovuttaisiin täysin audiovisuaalisen todistelun vastaanottamisesta. Toki kuten raportissa todetaan, tällä voisi olla merkittäviä seurauksia oikeuden toteutumiseen sekä ylipäätään oikeusjärjestelmän toimivuuteen.[239]

On myös muistettava, ettei uusi lainsäädäntö ole aina välttämättä kaikista järkevin ratkaisu, kun pohditaan teknologian haittavaikutusten minimointia. Liian laaja ja vaikeaselkoinen sääntely voi hidastaa innovaatioiden syntyä. On selvää, että tekoälyn kehityksestä seuraavat uhkakuvat kaipaavat kokonaisvaltaisia ratkaisuja. Jotta uhkia voidaan ehkäistä ja torjua, olisi asiantuntijoiden mukaan tärkeää kehittää niin osaamispohjaa kuin peruskansalaistaitojakin.[240] Tällä viitataan mediakriittisyyttä koskevaan valistukseen ja erityisesti syvävääreennöksiä koskevaan tiedon levitykseen. Hyvä idea voisi olla myös kehittää kansalaisille ohjeistuksen siitä, miten kannattaa toimia, jos verkossa tulee vastaan varma syvävääreennös tai syvävääreennökseltä vaikuttava video, jota levitetään kuin video olisi alkuperäinen.

Tietoisuuden lisääminen on tärkeää kuluttajien mediakriittisyyden näkökulmasta sekä niiden henkilöiden kannalta, jotka haluavat tehdä syvävääreennöksiä. Vuonna 2019 teetetyssä tutkimuksessa 72 prosenttia Britannian kansalaisista eivät olleet kuulleet syvävääreennöksistä eivätkä siten olleet tietoisia niiden mahdollisista haitoista.[241] Tilanne lienee parantunut muutamassa vuodessa, kun syvävääreennökset ovat yleistyneet. Silti aiheesta informointi on tärkeää sekä videoiden yleisön että niiden tekijöiden kannalta. Epäilyttävien videoiden suhteen on helpompaa olla kriittinen, jos tietää, että tällaisia videoita voidaan ylipäätään tehdä helposti deepfake-teknologian avulla. Toisaalta näitä laillisia sovelluksia käyttävien olisi hyvä tiedostaa mahdollisen väärinkäytön rajat ja se, että väärinkäytöstä voi joutua rikosoikeudelliseen vastuuseen.[242]

---

[238] Saaripuu 2019 s. 52.
[239] EPRS 2021 s. 34.
[240] Ks. esim. Tarkoma 2021 s. 111.
[241] iProov 1.10.2019.
[242] Europol 2022.

6. JOHTOPÄÄTÖKSET

Verkkoyhteiskunnassa valheellisuudesta on tullut normi. Sosiaalisessa mediassa autenttisuus on vähissä muun muassa siksi, että kuvan- ja videonkäsittely älypuhelimen avulla on tehty vaivattomaksi.[243] Erityisesti verkossa tapahtuvasta kommunikoinnista on tullut visuaalista. Visuaalisuus näkyy esimerkiksi siinä, että Google on ilmoittanut hakukoneensa siirrosta uuteen aikakauteen, jossa hakutuloksina käyttäjät saavat yhä enemmän kuvia ja videoita.

Tekoälyyn pohjautuva syvävääreentäminen mahdollistaa vaivattoman kuva-, ääni- ja videomanipulaation. Syvävääreennöksiä voidaan hyödyntää opetuksessa, elokuvateollisuudessa ja esimerkiksi yksilön itseilmaisussa. Syvävääreennösten hyötykäyttöön liittyy kuitenkin pulmia, kuten kysymykset kuolleiden henkilöiden tietojen käytöstä. Syvävääreennösten väärinkäyttöön taas liittyy monia eri riskejä ja vääreennösten havainnointi on haasteellista. Väärinkäyttöön liittyvät riskit voivat kohdistua sekä yksilöön että yhteiskuntaan. Yksilölle aiheutuvia seurauksia voivat olla esimerkiksi mainehaitta, häpeä ja pelko. Yhteiskunnassa syvävääreennösten aiheuttamat vaikutukset voivat olla seurausta informaatiovaikuttamisesta ja lopputuloksena voi olla esimerkiksi vääristynyt vaalitulos, valtiollisiin instituutioihin kohdistuva luottamuspula, polarisaatio tai muu turvallisuusuhkien lisääntyminen.

Syvävääreennöksiin löytyy sovellettavissa olevia säädöksiä kansallisella ja Euroopan unionin tasolla sekä myös kansainvälisessä oikeudessa. Säädöksiä voidaan soveltaa teknologianeutraalisuudesta johtuen ilman erityistä mainintaa syvävääreennöksistä.

Kaikkiaan soveltuvien säädösten määrä on suuri. Kansallisesta lainsäädännöstä esimerkkinä voidaan esittää Suomen perustuslaki, rikoslaki sekä vahingonkorvauslaki. Unionin oikeudessa syvävääreennöksiin soveltuvat jo nyt yleinen tietosuoja-asetus, EU:n tekijänoikeusjärjestelmä, audiovisuaalisia mediapalveluja koskeva direktiivi[244], käytännesäännöt disinformaation torjuntaan, toimintasuunnitelma disinformaation torjumiseksi, demokratian toimintasuunnitelma sekä syksyllä 2022 voimaan tullut digipalvelusäädös. EU:n tasolla syvävääreennösten käyttöä on

[243] EPRS 2021 s. 22.
[244] Euroopan parlamentin ja neuvoston direktiivi 2010/13/EU, annettu 10 päivänä maaliskuuta 2010, audiovisuaalisten mediapalvelujen tarjoamista koskevien jäsenvaltioiden tiettyjen lakien, asetusten ja hallinnollisten määräysten yhteensovittamisesta.

tulevaisuudessa tarkoitus rajoittaa myös uuden tekoälyä koskevan lainsäädäntökehikon ja erityisesti siihen sisältyvän uuden tekoälyasetuksen myötä. Kansainvälisessä oikeudessa syvävääreennöksiin voidaan liittää esimerkiksi Istanbulin sopimus[245]. Kaikkia näitä säädöksiä ei ole käsitelty tässä tutkielmassa, eikä se olisi ollut tarkoituksenmukaistakaan.

Kaikkiaan syvävääreennöksiin soveltuvien säädösten määrä on suuri, mutta oikeuden toteutuminen erityisesti syvävääreennösten väärinkäytön uhrien näkökulmasta on silti haasteellista.[246] Kansainvälisessä oikeudessa ja EU-oikeudessa voidaan luoda jonkinlaiset raamit syvävääreennösten ympärillä olevalle ilmiölle, mutta niitä koskevien rikosoikeudellisten ongelmien ratkaiseminen jää kansallisen sääntelyn varaan.

Todennäköisimmin deepfake-teknologian väärinkäyttöön soveltuvia rikoslain säännöksiä ovat identiteettivarkaus, petos, kunnianloukkaus, muutamat uuden seksuaalirikoslainsäädännön mukaiset seksuaalirikokset, väärennys, kiristys sekä erilaiset yleistä järjestystä vastaan tehdyt rikokset. Joissakin tapauksissa sovellettavaksi voivat tulla oikeudenkäyttöä ja viranomaista vastaan tehdyt rikokset ja niitä koskevat säännökset. Vaalivaikuttamisen sekä disinformaation osalta rikoslain soveltuvuus jää epäselväksi. Todennäköisenä voidaan kuitenkin pitää sitä, etteivät rikoslain mukaiset rikokset poliittisia oikeuksia vastaan ole sovellettavissa deepfake-teknologian väärinkäyttöön niiden tavoitteista riippumatta. Disinformaation levittämisen osalta tärkeä lähtökohta on se, ettei valehtelu itsessään ole rikos. Muiden mahdollisten rikosnimikkeiden osalta voidaan sanoa, että rikoslain soveltuvuus on mahdollista monissakin eri tilanteissa, jotka riippuvat lähinnä rikoksentekijän mielikuvituksen laajuudesta.

Vaikka Suomen rikoslaki on sovellettavissa erilaisissa väärinkäyttötilanteissa, jää tutkimuksen perusteella myös pieni epävarmuus tämänhetkisen rikoslaillisen sääntelyn riittävyydestä. Oikeudellisen sääntelyn suunnittelussa on tärkeää pyrkiä seuraamaan teknologisen toimintaympäristön kehitystä. Mikäli kehitetään teknologiaa, joka pystyy täydellä varmuudella tunnistamaan syvävääreennöksen, ei rikos- tai muu oikeudellinen tarve reagoida syvävääreennöksiin välttämättä

[245] Euroopan neuvoston yleissopimus naisiin kohdistuvan väkivallan ja perheväkivallan ehkäisemisestä ja torjumisesta (SopS 53/2015).
[246] EPRS 2021 s. VI.

ole yhtä merkittävä, kuin mitä se on nyt. Tilanne voi olla sama, jos syväväärennösteknologian valmistajat päättävät kehittää – tai he muuttuvan lainsäädännön vuoksi joutuvat kehittämään – itse ratkaisuja deepfake-teknologian mukanaan tuomiin ongelmiin. Ennen näitä oikeudellisia muutoksia tai teknologiasta löytyviä ratkaisuja, olemme olemassa olevan lainsäädännön sekä oikeusperiaatteiden varassa. Joka tapauksessa rikosoikeuden osalta voidaan todeta, että jo nyt olemassa olevat säännökset kattavat suhteellisen monipuolisesti syväväärentämisen mahdollistamia rikoksia.

Uuden lainsäädännön kehittämistä hankaloittavat eri perusoikeudet, jotka toisaalta puoltavat syväväärennösten käyttöä ja toisaalta vastustavat sitä. Perusoikeuksille onkin ominaista niiden päällekkäisyys. Myös syväväärennöksiin soveltuvat oikeudet ovat sillä tavalla toisiinsa kietoutuneita, ettei niitä kaikkia voida täysimääräisesti soveltaa loukkaamatta jotakin toista oikeutta. Esimerkiksi henkilötietojen käsittelyyn liittyvä yksityisyyden suoja ei koostu pelkästään henkilötietojen ja yksityisyyden suojasta, vaan siihen voidaan liittää yhtä lailla muitakin oikeuksia, kuten oikeudet kunniaan, henkilökohtaiseen koskemattomuuteen ja ihmisarvoiseen kohteluun, oikeus turvallisuuteen sekä oikeus vaikuttaa itseään koskeviin asioihin.[247]

Luottamuksen heikkeneminen uutisiin ja tietoon sekä faktojen ja mielipiteiden sekoittuminen on huolestuttavaa. Tätä disinformaatio ja syväväärennökset ruokkivat. Tutkimuksen mukaan jo pelkkä syväväärennösten olemassaolo aiheuttaa ihmisissä epäluottamusta saatavilla olevaa tietoa kohtaan, olipa kyse sitten faktatiedosta tai valheesta.[248] Verkkoympäristössä tieto – sekä oikea että valheellinen – voi myös levitä nopeasti ympäri maailmaa. Kun haittaa aiheuttava syväväärennös leviää, on todennäköistä, ettei sitä voida poistaa verkosta kokonaan. Vaikka alustoille olisi asetettu vastuu laittomien videoiden poistamisesta, ja tätä vastuuta noudatettaisiin, ovat alustojen käyttäjät silti saattaneet ladata videon omalle laitteelle, jolloin ulkopuoliset eivät pääse materiaaliin käsiksi. Tämä on yksi syy, miksi laittomien syväväärennösten osalta niiden levittämisen ja luomisen ennaltaehkäisy on erityisen tärkeää. Toinen syy on se, että laittomien syväväärennösten tekijät todennäköisesti jakavat videoita anonyymisti. Huolellisesti toteutettuna syväväärennökseen liittyvä rikos on tehtävissä hyvin pienellä kiinnijäämisen riskillä.

[247] HE 96/1998 vp s. 5, Korpisaari ym. 2022 s. 16.
[248] EPRS 2021 s. III.

On merkityksetöntä, mikä rikoslain pykälä soveltuu missäkin tilanteessa, jos laittoman videon tekijää ja/tai levittäjää ei saada kiinni. Digipalvelusäädöksessä sekä uudessa tekoälyasetuksessa säädetyt vaatimukset joidenkin syväväärennösten merkitsemisestä ja ilmoitusvelvollisuudesta voivat toimiessaan olla hyviä. Kuitenkin on odotettavissa, että vilpillisessä mielessä *mala fide* toimiva henkilö yksinkertaisesti jättää tämän merkitsemis- ja ilmoitusvelvollisuuden täyttämättä. Silloin ei kattavasti soveltuvasta rikoslainsäädännöstäkään ole välttämättä hyötyä väärinkäytön seurauksia setvittäessä. Digipalvelusäädöksessä säädetään myös laittoman materiaalin poistomenettelyistä. Laittomuuden arviointi on oma ongelmansa ja syväväärennösten poistoa alustoilta varjostaa mahdollisuus tarpeettomasta sananvapauden rajoittamisesta.

Väärinkäyttöön voidaan puuttua myös muilla keinoin. Tärkeää olisi saada syväväärennösten väärinkäyttäjät kiinni. Erilaiset vahvat tunnistautumiskeinot deepfake-teknologian käytössä tai lohkoketjuteknologian hyödyntäminen voisivat poistaa ongelman koskien tekijöiden vaikeaa jäljitettävyyttä. Näissäkin ratkaisuissa tulisi vastaan muun muassa sananvapauden rajoittamista koskevat ongelmat. Tärkeämpää olisikin keskittyä esimerkiksi syväväärennöksiä ja disinformaatiota koskevaan valistukseen.

On selvää, että tulevaisuudessa syväväärennösteknologian väärinkäyttöön on voitava puuttua tavalla tai toisella. Nopealla tahdilla kehittyvä teknologia tuo mukanaan paljon uusia mahdollisuuksia ja hyötyä, mutta luo herkästi myös pohjaa pahantahtoisten tarkoitusperien toteuttamiselle. Tämän tutkielman pohjalta voidaan todeta, että syväväärennösten osalta suurimmaksi ongelmaksi jää disinformaation levittäminen sekä vaalivaikuttaminen. Näiden tekojen vaikutukset voivat olla merkittäviä ja niihin on vaikeaa puuttua rikosoikeudellisin keinoin. Kyseiseen ongelmaan ei toistaiseksi löytynyt vastausta ja aiheen moniulotteisen ongelmallisuuden vuoksi siitä riittäisi pohdittavaa myös tulevaisuuden tutkimustyöhön.