The Weaponisation of Deepfakes

Digital Deception by the Far-Right

Ella Busch and Jacob Ware

International Centre for Counter-Terrorism

The Weaponisation of Deepfakes

Digital Deception by the Far-Right

Ella Busch and Jacob Ware

ICCT Policy Brief

December 2023

International Centre for Counter-Terrorism

About ICCT

The International Centre for Counter-Terrorism (ICCT) is an independent think and do tank providing multidisciplinary policy advice and practical, solution-oriented implementation support on prevention and the rule of law, two vital pillars of effective counter-terrorism.

ICCT's work focuses on themes at the intersection of countering violent extremism and criminal justice sector responses, as well as human rights-related aspects of counter-terrorism. The major project areas concern countering violent extremism, rule of law, foreign fighters, country and regional analysis, rehabilitation, civil society engagement and victims' voices.

Functioning as a nucleus within the international counter-terrorism network, ICCT connects experts, policymakers, civil society actors and practitioners from different fields by providing a platform for productive collaboration, practical analysis, and exchange of experiences and expertise, with the ultimate aim of identifying innovative and comprehensive approaches to preventing and countering terrorism.

Licensing and Distribution

ICCT Policy Brief
December 2023
DOI: 10.19165/2023.2.07
ISSN: 2468-0486

Contents

About ICCT iii

Abstract 1

Introduction 2

Inside the "Mind" of the Deepfake 2

The Infocalypse and its Dividends 3

Deepfakes and the Incitement of Far-Right Violence 4

Detecting the Deepfake - Current Mitigation Strategies, Proposals, & Challenges 5

Next Steps - Policy Recommendations & CVE Strategies 8

Bibliography 12

About the Authors 15

Abstract

In an ever-evolving technological landscape, digital disinformation is on the rise, as are its political consequences. In this policy brief, we explore the creation and distribution of synthetic media by malign actors, specifically a form of artificial intelligence-machine learning (AI/ML) known as the deepfake. Individuals looking to incite political violence are increasingly turning to deepfakes–specifically deepfake video content–in order to create unrest, undermine trust in democratic institutions and authority figures, and elevate polarised political agendas. We present a new subset of individuals who may look to leverage deepfake technologies to pursue such goals: far-right extremist (FRE) groups. Despite their diverse ideologies and worldviews, we expect FREs to similarly leverage deepfake technologies to undermine trust in the American government, its leaders, and various ideological 'out-groups.' We also expect FREs to deploy deepfakes for the purpose of creating compelling radicalising content that serves to recruit new members to their causes. Political leaders should remain wary of the FRE deepfake threat and look to codify federal legislation banning and prosecuting the use of harmful synthetic media. On the local level, we encourage the implementation of "deepfake literacy" programs as part of a wider countering violent extremism (CVE) strategy geared towards at-risk communities. Finally, and more controversially, we explore the prospect of using deepfakes themselves in order to "call off the dogs" and undermine the conditions allowing extremist groups to thrive.

Keywords: deepfakes, far-right, disinformation, extremism, radicalisation

Introduction

On 2 March 2022, shortly following Russia's invasion of Ukraine, Ukraine24 released a video of President Volodymyr Zelensky taking to lectern and asking his fellow Ukrainians to put down their arms and surrender to Russia. Users quickly pointed out discrepancies in the video, from its odd pixelation to the President's varied skin tone between his face and his neck. Indeed, this video was not a true statement from Zelensky but rather a deepfake.[1] Deepfakes are a form of synthetic media, audio, images, and/or videos that are either partially or wholly manipulated through artificial intelligence (AI) technologies and used in a maliciously deceptive or misinformative manner.[2] The Zelensky incident served as the first high-profile use of a deepfake during an armed conflict, and will likely be far from the last.[3] This paper aims to explore the potential usage of deepfake technologies by a new subset of perpetrators–far right extremist (FRE) groups–and the subsequent law enforcement and policy challenges that may arise.[4] It finds that deepfakes will offer considerable new opportunities for FRE actors—and that law enforcement and intelligence agencies must urgently plan countermeasures against this new information tool.

Inside the "Mind" of the Deepfake

As the name suggests, the word "deepfake" is derived from the concept of deep learning, a subset of artificial intelligence-machine learning (AI/ML). Deep learning algorithms are composed of deep neural networks, which simulate the human brain in a way that enables the AI to "learn" from large amounts of data. Deep learning is unique from machine learning for its ability to process unstructured data, such as text and images, which make it ideal for creating deepfake videos, audio, images, or text.[5] Deepfakes are created using a specific deep learning algorithm called a generative adversarial network (GAN). GANs, first developed by researcher Ian Goodfellow in 2014, consist of two neural networks: a generator algorithm and a discriminator algorithm.[6] The generator algorithm creates a fake image (or other form of media) and the discriminator judges the media's authenticity. The action repeats for hours, or even days, until reaching a stable state in which neither the generator nor the discriminator can improve their performance. [7]

The process of creating a well-made deepfake is time-consuming and costly, requiring highly-sophisticated technology. One experiment attempting to make a Tom Cruise deepfake spent two months merely training the GAN, which required a pair of NVIDIA RTX 8000 graphics processing units (GPUs), which, in 2022, cost upward of USD $5,795 each.[8] However, there already exists a

1 James Pearson and Natalia Zinets, "Deepfake Footage Purports to Show Ukrainian President Capitulating," Reuters, March 17, 2022. https://www.reuters.com/world/europe/deepfake-footage-purports-show-ukrainian-president-capitulating-2022-03-16/.
2 Nina Schick, "Introduction." Essay. In Deepfakes: The Coming Infocalypse, 10–11. New York, NY: Grand Central Pub, 2021.
3 Daniel L Byman, Chongyang Gao, Chris Meserole, and V.S. Subrahmanian, "Deepfakes and International Conflict – Brookings," Foreign Policy at Brookings, January 2023. https://www.brookings.edu/wp-content/uploads/2023/01/FP_20230105_deepfakes_international_conflict.pdf?mc_cid=3f7678e334&mc_eid=ee94395166.
4 For more, see Dilrukshi Gamage, Jiayu Chen, Piyush Ghasiya, and Kazutoshi Sasahara, "Deepfakes and Society: What Lies Ahead?" in Mahdi Khosravy, Isao Echizen, and Noboru Babaguchi (eds.), Frontiers in Fake Media Generation and Detection (Singapore: Springer, 2022).
5 "What Is Deep Learning?" IBM, n.d. https://www.ibm.com/topics/deep-learning#:~:text=the%20next%20step-,What%20is%20deep%20learning%3F,from%20large%20amounts%20of%20data.
6 Nina Schick, "Chapter One." Essay. In Deepfakes: The Coming Infocalypse, 45. New York, NY: Grand Central Pub, 2021.
7 Daniel L. Byman, Chongyang Gao, Chris Meserole, and V.S. Subrahmanian. "Deepfakes and International Conflict – Brookings,"Foreign Policy at Brookings, January 2023. https://www.brookings.edu/wp-content/uploads/2023/01/FP_20230105_deepfakes_international_conflict.pdf?mc_cid=3f7678e334&mc_eid=ee94395166.
8 Todd C. Helmus, "Artificial Intelligence, Deepfakes, and Disinformation: A Primer," RAND Corporation, July 6, 2022. https://www.rand.org/pubs/perspectives/PEA1043-1.html.

demand for deepfakes-as-a-service (DaaS), with threat actors willing to pay as much as $16,000 for the service.[9] Additionally, the algorithms underlying GANs are often released with open-source licenses, meaning that anyone, including malicious actors, may download and train a GAN to suit their purposes.[10]

The Infocalypse and its Dividends

Experts fear that advances in deepfake technologies will greatly contribute to the ongoing "infocalypse," the phenomenon in which people are pushed to believe that they cannot trust information outside their social circles or that does not confirm their existing beliefs.[11] Some estimates suggest that, by 2026, as much as 90 percent of online content might be synthetically generated, meaning that the use of deepfakes will likely become a prevalent source of cybercrime. Perhaps more worryingly, the public is relatively uninformed about deepfakes. A 2019 study found that 72 percent of individuals in the UK were unaware of deepfakes and their impact,[12] while a later study found that even those who claimed that they could detect deepfakes were routinely fooled by hyper-realistic content.[13] A 2023 study found that participants could only correctly identify a deepfake 73 percent of the time.[14] Aided by new technologies such as 5G, deepfakes can now alter videoconferences and livestreams in real time, adding to the infocalypse, particularly if such content is circulated in the news.[15]

Deepfakes have been increasingly recognised by international governments and law enforcement as a distinct criminal and national security threat. This was seen in a June 2023 United States Senate Committee on the Judiciary hearing entitled *Artificial Intelligence and Human Rights*, where witness Jennifer DeStefano spoke about her experience as a victim of a deepfake kidnapping and extortion scam. Ms DeStefano revealed that she had received a deepfake phone call that supposedly came from her fifteen-year-old daughter, saying that she had been kidnapped and held ransom for a million dollars.[16] Also in June 2023, news circulated around deepfake "sextortion" scams, where victims, particularly minors, are either led into sharing or are photoshopped onto explicit photos. The perpetrators then threaten to publicly release the content unless they receive payment.[17] Recent scams suggest that deepfake-generated crimes

9 "Facing Reality? Law Enforcement and the Challenge of Deepfakes." Europol Innovation Lab, 2022. https://www.europol.europa.eu/cms/sites/default/files/documents/Europol_Innovation_Lab_Facing_Reality_Law_Enforcement_And_The_Challenge_Of_Deepfakes.pdf.
10 Daniel L. Byman, Chongyang Gao, Chris Meserole, and V.S. Subrahmanian. "Deepfakes and International Conflict – Brookings," Foreign Policy at Brookings, January 2023. https://www.brookings.edu/wp-content/uploads/2023/01/FP_20230105_deepfakes_international_conflict.pdf?mc_cid=3f7678e334&mc_eid=ee94395166.
11 Mika, Westerlund, "The Emergence of Deepfake Technology: A Review | Tim Review." Technology Innovation Management Review, 2019. https://timreview.ca/article/1282.
12 "Facing Reality? Law Enforcement and the Challenge of Deepfakes." Europol Innovation Lab, 2022. https://www.europol.europa.eu/cms/sites/default/files/documents/Europol_Innovation_Lab_Facing_Reality_Law_Enforcement_And_The_Challenge_Of_Deepfakes.pdf.
13 Todd C, Helmus, "Artificial Intelligence, Deepfakes, and Disinformation: A Primer," RAND Corporation, July 6, 2022. https://www.rand.org/pubs/perspectives/PEA1043-1.html.
14 Kimberly T. Mai, Sergi Bray, Toby Davies, and Lewis D. Griffin, "Warning: Humans Cannot Reliably Detect Speech Deepfakes," Public Library of Science, 2023. https://journals.plos.org/plosone/article?id=10.1371%2Fjournal.pone.0285333.
15 "Facing Reality? Law Enforcement and the Challenge of Deepfakes." Europol Innovation Lab, 2022. https://www.europol.europa.eu/cms/sites/default/files/documents/Europol_Innovation_Lab_Facing_Reality_Law_Enforcement_And_The_Challenge_Of_Deepfakes.pdf.
16 "Artificial Intelligence and Human Rights." United States Senate Committee on the Judiciary, June 13, 2023. https://www.judiciary.senate.gov/committee-activity/hearings/artificial-intelligence-and-human-rights.
17 Jared Gans, "FBI Warns of 'deepfakes' in Sextortion Schemes," The Hill, June 7, 2023. https://thehill.com/policy/cybersecurity/4037204-fbi-warns-of-deepfakes-in-sextortion-schemes/. For more, see Audrey de Rancourt-Raymond and Nadia Smaili, "The unethical use of deepfakes," Journal of Financial Crime 30, no. 4 (2023): pp. 1066-1077, https://www.emerald.com/insight/content/doi/10.1108/JFC-04-2022-0090/full/html.

will likely become a cornerstone of law enforcement activity in the future; in 2022, Europol's Innovation Lab brought together 80 law enforcement experts, who collectively cited deepfakes as a major challenge for the field for the years leading up to 2030.[18] A recent memo by the US Department of Homeland Security (DHS) similarly cites the potential for deepfakes to be used as false evidence in criminal trials,[19] as seen in a 2020 UK custody case, where deepfake audio recording was used to create the impression that the father made violent threats against his wife.[20] Byman et al. have considered the potential for deepfakes to legitimise war and uprisings, presenting the example of a deepfake video portraying a powerful world leader burning the Qur'an, which would lead to the incitement of violence.[21] Other academics foresee abuses of power by authority figures, citing a phenomenon that has been dubbed The Liar's Dividend—"the notion that individuals could successfully deny the authenticity of genuine content–particularly if it depicts inappropriate or criminal behavior–by claiming that the content is a deepfake."[22]

Deepfakes and the Incitement of Far-Right Violence

Far-right extremists are likely to implement deepfakes in multiple creative ways. First and foremost, deepfakes will empower far-right extremist influencers with the tools to incite violence, by mass-creating fake content appealing to the violent far right's innate desire for violence and confrontation. In 2018, in a demonstration of the potential ramifications of deepfake videos, Jordan Peele, the director of "Get Out," posted a viral video to Buzzfeed portraying former president Barack Obama calling then-President Donald Trump "a complete and total dipshit" under the title "You Won't Believe What Obama Says in this Video!"[23] The video was widely touted for bringing attention to the dangers of deepfakes and their potential effects on democracy. One year later, the notion came further to fruition through a video of then-Speaker of the House Nancy Pelosi, which was slowed and edited to a degree to make it look as if she were intoxicated (what might be termed a "cheap fake"). Then-president Donald Trump posted the video to his Twitter account, captioned "PELOSI STAMMERS THROUGH NEWS CONFERENCE."[24] The move was significant, as the leader of the United States–wittingly or not–instilled distrust in the legislative branch. Instilling distrust in authorities is one of the many potential repercussions of creating such a deepfake video, and the possibilities for exploitation are endless for malign actors, particularly those who distrust the government altogether.

---

18 "Facing Reality? Law Enforcement and the Challenge of Deepfakes." Europol Innovation Lab, 2022. https://www.europol.europa.eu/cms/sites/default/files/documents/Europol_Innovation_Lab_Facing_Reality_Law_Enforcement_And_The_Challenge_Of_Deepfakes.pdf.

19 Tina Brooks, Princess G., Jesse Heatley, Scott Kim, Samantha M., Sara Parks, Maureen Reardon, et al. "Increasing Threat of Deepfake Identities - Homeland Security," Department of Homeland Security, n.d. https://www.dhs.gov/sites/default/files/publications/increasing_threats_of_deepfake_identities_0.pdf.

20 Patrick Ryan, "'deepfake' Audio Evidence Used in UK Court to Discredit Dubai Dad," The National, July 4, 2021. https://www.thenationalnews.com/uae/courts/deepfake-audio-evidence-used-in-uk-court-to-discredit-dubai-dad-1.975764.

21 Daniel L. Byman, Chongyang Gao, Chris Meserole, and V.S. Subrahmanian, "Deepfakes and International Conflict – Brookings," Foreign Policy at Brookings, January 2023. https://www.brookings.edu/wp-content/uploads/2023/01/FP_20230105_deepfakes_international_conflict.pdf?mc_cid=3f7678e334&mc_eid=ee94395166.

22 Kelley M. Sayler, and Laurie A. Harris, "Deep Fakes and National Security," Congressional Research Service , April 17, 2023. https://www.documentcloud.org/documents/23798946-deep-fakes-and-national-security-april-17-2023.

23 Kaylee Fagan, "A Viral Video That Appeared to Show Obama Calling Trump a 'dips---' Shows a Disturbing New Trend Called 'Deepfakes.', Business Insider, April 17, 2018. https://www.businessinsider.com/obama-deepfake-video-insulting-trump-2018-4.

24 Kathryn, Watson, "Trump Tweets Heavily Edited Video of Pelosi Played by Fox Business," CBS News, May 25, 2019. https://www.cbsnews.com/news/trump-tweets-heavily-edited-video-of-pelosi-played-by-fox-news/. See also Thomas Paterson and Lauren Hanley, "Political warfare in the digital age: cyber subversion, information operations and 'deep fakes,'" Australian Journal of International Affairs 74, no. 4 (2020): pp. 439-454, https://www.tandfonline.com/doi/full/10.1080/10357718.2020.1734772.

It is for this reason that anti-government extremists will likely utilise synthetic media to pursue their goals: false, inflammatory information or statements by trusted authority figures, election misinformation, or other distortions of social and political events are all likely to incite violence, presenting the possibility of a second January 6th, or something far worse. The anonymised nature of much modern seditious extremism makes this threat more severe, by broadening the number of identities a deep faker could adopt. The QAnon movement, for instance, has succeeded in "mass radicalising" the American public by sharing alleged "drops" of government malfeasance from a self-proclaimed, privileged perch within government.[25] Imagine if a successful deepfake claiming to depict the Q character from QAnon calling for violence against President Trump's rivals in government was released at a particularly opportune time. Social media companies and the government could attempt to undermine the content by providing proof of its artificial qualities, but the damage would likely be done, particularly given these online spaces' "do your own research" culture would delegitimise the credibility of any counterarguments. These groups will accordingly also particularly benefit from the aforementioned "Liar's Dividend," easily dismissing news that does not fit their narrative, while deepfakes continue to advance a broader climate of distrust and inauthenticity.[26]

Secondly, deepfakes will provide extremists with ideological ammunition, allowing influencers to manufacture "proof" of alleged wrongdoing that justifies extremist views. The use of violence to elevate one religious or racial group over another is a tale as old as time, used to justify wars, genocide, and societal oppression. These extremist movements often share similar foundational conspiracy theories, such as the Great Replacement theory, the belief that minority groups and Jews are orchestrating a plot to overwhelm the white European population, leading to their "replacement," and the New World Order theory, the belief that a [Jewish] cabal of elites is conspiring to create a unified international government that will grant them complete control over the world.[27] Given this conspiratorial wasteland, is not difficult to see the possible exploitation of deepfakes by such groups to incite the racial holy war ("RaHoWa") that they seek to create.

Content serving as "proof" of wrongdoing or conspiracy by these outgroups could be used to validate their respective causes. These groups, many of which see themselves as protectors against government overreach, would benefit greatly from deepfake content that portrays the federal government as corrupt, incompetent, fraudulent, or tyrannical. A far-right influencer recently tweeted a video of President Biden declaring a military draft to counter Russian mobilisation in Ukraine, a blatant fake designed to inspire anti-government sentiment.[28] In another example, after the latest outbreak of hostilities between Israel and Hamas, fake audio purported to show London mayor Sadiq Khan announcing that the city would forgo the annual Armistice Day commemoration in favour of a pro-Palestine march, the mayor's office declaring that the audio was being "circulated and amplified by a far-right group." Moreover, in this case, the Metropolitan Police specifically stated that the faked audio was not a crime.[29] Meanwhile, in the United States, an AI-generated voice mimicking the director of the Anti-Defamation League was used to spread anti-Semitic rhetoric.[30] Deepfakes provide the opportunity to ignite political

25 Zack Stanton, "The Problem Isn't Just One Insurrection. It's Mass Radicalization," Politico, February 11, 2021, https://www.politico.com/news/magazine/2021/02/11/mass-radicalization-trump-insurrection-468746.
26 Matteo Wong, "We Haven't Seen the Worst of Fake News," The Atlantic, December 20, 2022. https://www.theatlantic.com/technology/archive/2022/12/deepfake-synthetic-media-technology-rise-disinformation/672519/.
27 Flores Myles, "The New World Order: The Historical Origins of a Dangerous Modern Conspiracy Theory," Middlebury Institute of International Studies at Monterey, May 31, 2022. https://www.middlebury.edu/institute/academics/centers-initiatives/ctec/ctec-publications/new-world-order-historical-origins-dangerous.
28 Isaac Stanley-Becker and Naomi Nix, "Fake images of Trump arrest show 'giant step' for AI's disruptive power," Washington Post, March 22, 2023, https://www.washingtonpost.com/politics/2023/03/22/trump-arrest-deepfakes/.
29 Jess Warren, "Fake Audio of Sadiq Khan Is Not a Crime, Says Met," BBC News, November 11, 2023. https://www.bbc.com/news/uk-england-london-67389609.
30 Donie O'Sullivan, Curt Devine, and Allison Gordon, "How Antisemitic Hate Groups Are Using Artificial Intelligence

violence, whether an insurrection to remove a disliked political leader, protests against a made-up transgression by the government, false evidence of a government cover-up, fabricated scandals, or proof of a government conspiracy, among other things. As the 2024 elections come to the forefront, polarising misinformation campaigns may be the means by which such groups can pursue their desired governmental changes.[31]

More specifically, the manufacturing of proof will be particularly beneficial for extremist movements rallying against perceived overlords, such as the male supremacist and incel movements that claim victimhood against a new feminist order that suppresses men and masculinity.[32] In 2017, a Reddit user calling himself "Deepfakes" began the discussion thread "r/deepfakes," which he dedicated to fake, AI-generated videos depicting famous actresses' faces onto porn videos. These videos, among the first deepfakes, were convincing, inspiring other users to make their own versions and post them online. Although Reddit shut down the forum within weeks for "involuntary pornography," the damage had already been done; "Deepfakes" had shared his algorithm with the rest of the internet. Today, nonconsensual pornography may comprise up to 96 percent of deepfake videos.[33] This trend of "image-based sexual abuse" may be exploited by various misogynist groups, sometimes called "male supremacists."[34] While these groups' communications take place largely online,[35] instances of offline violence have led law enforcement, such as the Texas Department of Public Safety, to issue warnings that "this particular threat could soon match, or potentially eclipse, the level of lethalness demonstrated by other domestic terrorism types."[36]

Male supremacists' anti-female actions and rhetoric make them likely candidates for the creation of nonconsensual deepfake pornography or other forms of online harassment against women. An example of this threat may be seen in the 2014 case of Zoe Quinn, a video game developer who became the main victim in an event known as "Gamergate." Gamergate began when Quinn's ex-boyfriend posted unfounded content on his blog, stating that she had slept with men in exchange for positive reviews of the games she created. This precipitated a harassment campaign against women in the gaming industry, who received rape and death threats from male supremacists. Their addresses and contact information were posted online–a process known as doxxing–forcing the victims to leave their homes out of fear of retaliation.[37] One could imagine that deepfakes could reproduce an event like Gamergate, where manufactured pornographic videos may be used to create false evidence against certain women, leading to similar threats

---

in the Wake of Hamas Attacks," CNN, November 15, 2023. https://www.cnn.com/2023/11/14/us/hamas-israel-artificial-intelligence-hate-groups-invs/index.html.

31 Daniel I Weiner and Lawrence Norden, "Regulating AI Deepfakes and Synthetic Media in the Political Arena," Brennan Center for Justice, December 12, 2023. https://www.brennancenter.org/our-work/research-reports/regulating-ai-deepfakes-and-synthetic-media-political-arena.

32 For more on AI and gender, see Nick Zuroski, "Weapon or Tool?: How the Tech Community Can Shape Robust Standards and Norms for AI, Gender, and Peacebuilding," Global Network on Extremism & Technology, December 6, 2023, https://gnet-research.org/2023/12/06/weapon-or-tool-how-the-tech-community-can-shape-robust-standards-and-norms-for-ai-gender-and-peacebuilding/.

33 Nina Schick, "Chapter One." Essay. In Deepfakes: The Coming Infocalypse, 37-42. New York, NY: Grand Central Pub, 2021. See also Nina Jankowicz, "The threat from deepfakes isn't hypothetical. Women feel it every day." Washington Post, March 25, 2021, https://www.washingtonpost.com/opinions/2021/03/25/threat-deepfakes-isnt-hypothetical-women-feel-it-every-day/.

34 Chandell Gosse and Jacquelyn Burkell, "Politics and Porn: How News Media Characterizes Problems Presented By ..." Taylor & Francis Online, September 30, 2020. https://www.tandfonline.com/doi/full/10.1080/15295036.2020.1832697.

35 "Male Supremacy." Southern Poverty Law Center. Accessed August 14, 2023. https://www.splcenter.org/fighting-hate/extremist-files/ideology/male-supremacy.

36 Bruce Hoffman Jacob Ware, and Ezra Shapiro, "Assessing the Threat of Incel Violence - Taylor & Francis Online." Taylor & Francis Online, April 19, 2020. https://www.tandfonline.com/doi/full/10.1080/1057610X.2020.1751459.

37 "Male Supremacy." Southern Poverty Law Center. Accessed August 14, 2023. https://www.splcenter.org/fighting-hate/extremist-files/ideology/male-supremacy.

and harassment. Deepfake content may also be used to elevate misogynists' beliefs that women are inherently evil or discriminatory against men, leading to the incitement of violence or even the validation of the male supremacist cause, leading to the radicalisation of new members.

Deepfakes also present the potential to serve as a recruitment tool for extremist groups, particularly as online recruitment becomes ever-more important to violent extermism (VE) radicalisation; in 2016 alone, 90 percent of extremists were recruited at least in part via social media, especially "lone wolf" actors who do not rely on a formal organisational structure to carry out extremist activities. Far-right extremists were also found to participate in online dialogues and create content at a greater rate than their far-left or jihadist counterparts.[38]

The importance of online content to FRE radicalisation presents important implications for the potential use of deepfake technology as a recruitment tool. Deepfakes may be employed to create compelling content that glamorises the so-called "pull" factors of radicalisation[39] – such as group identity and camaraderie, the most prominent pull factor[40] – by presenting a false visual of life in the group and the positive experiences of members. Such propaganda tactics have already been employed by jihadist organisations like ISIS to appeal to younger generations and may be greatly improved using deepfakes.[41] The Islamic State has also pioneered the use of artificial intelligence and deepfakes to advance divine justifications for violence. Researchers Daniel Siegel and Bilva Chandra have for instance explored one case of deepfakes amplifying the story of a Pakistani Muslim man claiming to directly communicate with the Prophet Muhammad and God, and calling for the "deepest depth of hellfire" for those questioning his prophecies.[42] Deepfakes could also be created for humorous purposes, or to advance messages subliminally, both of which can further strengthen in-group bonds.[43]

Deepfakes may also have tactical benefits, especially as the technology strengthens. The ability to insert deepfake technology into livestreams as they occur may particularly motivate lone actors, who often post social media livestreams of their attacks to inspire fear and motivate others to follow in their footsteps.[44] Livestreams may be manipulated in order to make an attack appear justified or more successful than it truly is, which in turn may motivate further violence, or at least garner more attention. Deepfakes may also appeal to recruits' emotions by presenting emotionally-charged content that confirms their existing biases and calls them to arms, such as

38 Michael Jensen, Patrick James, Gary LaFree, Aaron Safer-Lichtenstein, and Elizabeth Yates, "Use of Social Media by US Extremists - UMD." START.UMD.edu, July 2018. https://www.start.umd.edu/pubs/START_PIRUS_UseOfSocialMediaByUSExtremists_ResearchBrief_July2018.pdf. See also Jacob Ware, "The Third Generation of Online Radicalization," George Washington University Program on Extremism, June 16, 2023, https://extremism.gwu.edu/third-generation-online-radicalization.

39 Peter Simi, "Recruitment and Radicalization among U.S. Far-Right Terrorists." Recruitment and Radicalization among U.S. Far-Right Terrorists | START.umd.edu, November 2016. https://www.start.umd.edu/publication/recruitment-and-radicalization-among-us-far-right-terrorists.

40 Peter Simi, "Recruitment and Radicalization among U.S. Far-Right Terrorists," Recruitment and Radicalization among U.S. Far-Right Terrorists | START.umd.edu, November 2016. https://www.start.umd.edu/publication/recruitment-and-radicalization-among-us-far-right-terrorists

41 Jordan Bach-Lombardo and Charlie Winter, "Why Isis Propaganda Works," The Atlantic, February 13, 2016. https://www.theatlantic.com/international/archive/2016/02/isis-propaganda-war/462702/.

42 Daniel Siegel and Bilva Chandra, "'Deepfake Doomsday': The Role of Artificial Intelligence in Amplifying Apocalyptic Islamist Propaganda," Global Network on Extremism & Technology, August 29, 2023, https://gnet-research.org/2023/08/29/deepfake-doomsday-the-role-of-artificial-intelligence-in-amplifying-apocalyptic-islamist-propaganda/.

43 Olivier Cauberghs, "For the Lulz?: AI-Generated Subliminal Hate is a New Challenge in the Fight Against Online Harm," Global Network on Extremism and Technology, November 13, 2023. https://gnet-research.org/2023/11/13/for-the-lulz-ai-generated-subliminal-hate-is-a-new-challenge-in-the-fight-against-online-harm/.

44 Stephane J. Baele, Lewys Brace, and Travis G. Coan, "Uncovering the Far-Right Online Ecosystem: An Analytical Framework and ...S," Taylor & Francis Online, December 30, 2020. https://www.tandfonline.com/doi/full/10.1080/1057610X.2020.1862895.

content portraying the wrongdoings of out-groups. They could also be created to downplay or mock law enforcement responses to extremist groups, perhaps encouraging members to be more blatant and overt in their violent activism.

It should be noted that these possibilities are not exhaustive—the applications of this technology are "only limited by the creativity of those designing the deepfakes."[45]

Detecting the Deepfake - Current Mitigation Strategies, Proposals, & Challenges

Identifying, addressing, and prosecuting the deepfake threat, especially when propagated by domestic extremists, will be no small feat for law enforcement. The hyperrealism of deepfakes makes them very difficult to identify and discern from real media for humans and machines alike. Even when identified, it is an additional challenge to locate and prosecute the individual responsible, particularly given the lack of a federal law regarding deepfakes. As of June 2023, only five US States have passed deepfake-related legislation. In 2019, Texas passed SB751 and California passed AB730, both of which banned the use of deepfakes in influencing upcoming elections. That same year, California also passed AB602, Georgia passed SB337, and Virginia passed SB1736, all of which prohibit the creation and dissemination of nonconsensual deepfake pornography. In 2020, New York passed S6829A, which provides legal recourse for the unlawful publication of deepfakes.[46]

At the federal level, recent actions have illustrated growing concern with deepfakes. The Deepfake Report Act of 2019 requires the Secretary of Homeland Security to annually report on "the extent of digital content forgery technologies, also known as deepfake technologies, [which] are being used to weaken national security, undermine our nation's elections, and manipulate media." A provision in the 2020 National Defence Authorization Act (NDAA) similarly also stipulates that the Director of National Intelligence issue a comprehensive report on the weaponisation of deepfakes, warn Congress of foreign deepfakes being used to target US elections, and create a competition that would award the creation of deepfake detection technologies. The 2020 Identifying Outputs of Generative Adversarial Networks Act directed the National Science Foundation (NSF) and National Institute of Standards and Technology (NIST) to support research on GANs, and the Defense Advanced Research Projects Agency (DARPA) created two programmes, MediFor (concluded in FY2021) and SemaFor (ongoing) devoted to the detection of deepfakes.[47]

A handful of legislative initiatives have been introduced, but not passed, in Congress since 2018.[48] In 2019, New York Representative Yvette Clark introduced the 2019 DEEP FAKES Accountability Act, requiring that all deepfake audio, visual, or moving-picture content be clearly labelled as deepfakes.[49] The bill was reintroduced in September 2023; as noted by ABC, the

45 Daniel L Byman, Chongyang Gao, Chris Meserole, and V.S. Subrahmanian, "Deepfakes and International Conflict – Brookings," Foreign Policy at Brookings, January 2023. https://www.brookings.edu/wp-content/uploads/2023/01/FP_20230105_deepfakes_international_conflict.pdf?mc_cid=3f7678e334&mc_eid=ee94395166.
46 Caroline Quirk, "The High Stakes of Deepfakes: The Growing Necessity of Federal Legislation to Regulate This Rapidly Evolving Technology," Princeton University, June 19, 2023. https://legaljournal.princeton.edu/the-high-stakes-of-deepfakes-the-growing-necessity-of-federal-legislation-to-regulate-this-rapidly-evolving-technology/.
47 Kelley M. Sayler and Laurie A. Harris, "Deep Fakes and National Security," Congressional Research Service , April 17, 2023. https://www.documentcloud.org/documents/23798946-deep-fakes-and-national-security-april-17-2023.
48 "S.2559 - Deepfake Task Force Act 117th Congress (2021-2022)." Congress.Gov, August 4, 2021. https://www.congress.gov/bill/117th-congress/senate-bill/2559/text.
49 Todd C. Helmus, "Artificial Intelligence, Deepfakes, and Disinformation: A Primer," RAND Corporation, July 6, 2022. https://www.rand.org/pubs/perspectives/PEA1043-1.html.

bill emerges ahead of the 2024 elections, "in which accessibility to generative-AI tools is giving candidates or their supporters the ability to produce real-looking fakes in order to advance partisan messages."[50] The following month, Wisconsin Senator Andre Jacque began circulating a deepfake-pornography related bill for co-sponsorship, hoping that the bill would, among other things, "give investigators more tools to prosecute [deepfake pornography]."[51] Notably, Congressional efforts have not been limited to either party, raising the potential that legislative measures to combat deepfakes might prove successful.

Tech giants have begun to enact policies and deepfake detection mechanisms on their platforms. Meta, TikTok, Reddit, and YouTube all have content moderation policies that call for a ban and remove synthetic media that produces misleading and deceptive content (although the 2024 presidential campaign to date has revealed flaws in the defences).[52] Meta has developed an AI tool that reverse-engineers AI-generated images to track their origin, while Google has released a large dataset of visual deepfakes (now included in the FaceForensics benchmark, a deepfake detection dataset) and Microsoft has launched the Microsoft Video Authenticator, which provides a percentage chance that a photo or video has been artificially manipulated.[53] In September 2023, Google also announced that verified elections advertisers would have to label synthetically-manipulated content.[54] Some companies have launched efforts to incentivise the development of deepfake detection technologies, such as Meta's Deep Fake Detection Challenge, while others have banned the training of AI systems to produce deepfakes, such as Google.[55]

The law enforcement experts at Europol's Innovation Lab noted that criminal actors tend to stay one step ahead of law enforcement in the implementation, use, and adaptation of technologies, making it difficult for deepfake detection algorithms to keep up with progressively-more-advanced deepfakes. While deepfakes can be assessed manually to find inconsistencies in pixelation or human attributes, it is a labour-intensive process and is complicated by human biases. Automated detection processes have been developed as well, including those that track biological signals (such as inconsistencies in skin tone), look for phoneme-viseme mismatches (mouth movements that are inconsistent with speech), assess patterns in facial movements (looking for discrepancies), and utilise recurrent convolutional models to find inconsistencies in video frames. Even automated detection systems are flawed, however, and can falsely flag a deepfake as real content if, for example, a video was compressed in a way that made it more difficult to detect inconsistencies in pixelation.[56] The issue is rendered yet more complicated by

50 Emmanuelle Saliba and Jessie DiMartino, "Bill Would Criminalize 'Extremely Harmful' Online 'Deepfakes,'" ABC News, September 25, 2023. https://abcnews.go.com/Politics/bill-criminalize-extremely-harmful-online-deepfakes/story?id=103286802.
51 Brittany Schmidt, "Proposed Wisconsin Bill to Address Artificially Made 'Deep Fake' Pornography," WBay, October 3, 2023. https://www.wbay.com/2023/10/02/proposed-wisconsin-bill-address-artificially-made-deep-fake-pornography/.
52 "Meta, X Questioned by Lawmakers Over Lack of Rules Against AI-Generated Political Deepfakes," CBS News, October 5, 2023. https://www.cbsnews.com/sanfrancisco/news/meta-x-political-deepfakes-lawyers-question-platforms-over-lack-of-rules/.
53 "Facing Reality? Law Enforcement and the Challenge of Deepfakes." Europol Innovation Lab, 2022. https://www.europol.europa.eu/cms/sites/default/files/documents/Europol_Innovation_Lab_Facing_Reality_Law_Enforcement_And_The_Challenge_Of_Deepfakes.pdf.
54 Emmanuelle Saliba and Jessie DiMartino. "Bill Would Criminalize 'Extremely Harmful' Online 'Deepfakes,'" ABC News, September 25, 2023. https://abcnews.go.com/Politics/bill-criminalize-extremely-harmful-online-deepfakes/story?id=103286802.
55 Jack Cook, "Deepfake Technology: Assessing Security Risk," American University, July 27, 2022. https://www.american.edu/sis/centers/security-technology/deepfake_technology_assessing_security_risk.cfm.
56 "Facing Reality? Law Enforcement and the Challenge of Deepfakes." Europol Innovation Lab, 2022. https://www.europol.europa.eu/cms/sites/default/files/documents/Europol_Innovation_Lab_Facing_Reality_Law_Enforcement_And_The_Challenge_Of_Deepfakes.pdf.

the fact that, like any other new technology, AI will primarily be used for positive purposes—such as education or ensuring online privacy—and that countermeasures will therefore be met with an outcry against blocking this technology solely because of manipulations by a fringe of bad actors.

Next Steps - Policy Recommendations & CVE Strategies

Although ongoing steps to combat deepfakes are promising and relatively robust, a more aggressive effort is nevertheless needed to prevent the more nefarious uses and impacts of this new technology.[57] Firstly, as prior measures have displayed, the most important antidote to deepfakes is education. This can be performed both prior to and in the aftermath of video releases. Governments and countering violent extremism (CVE) nonprofits should explore digital literacy efforts to educate the public on the tell-tale signs of deepfakes, as well as on the practice of seeking secondary and tertiary sources to confirm information or content. Such "pre-bunking" measures have proven effective in other environments and may help equip the public with the tools required to dismiss manipulated media once it breaks onto their screens.[58] In addition to "pre-bunking" measures, a community-level response to combatting deepfakes could come in the form of strengthened digital literacy programs. We encourage the implementation of "deepfake literacy" as part of a broader CVE program geared towards communities that are at-risk of radicalisation and recruitment by far-right extremist groups. Such a program would inform individuals of all ages of the dangers of deepfakes while teaching them how to identify deepfake content. While current CVE programs aim to counter the attractive online messaging promoted by FRE groups, this added layer would protect against the use of deepfakes to promote a false positive image of far-right recruitment. In cases where prevention fails, social media companies should flag content that is produced using artificial intelligence, perhaps by watermarking artificial content. Although such measures are in place, they will need to be strengthened as deepfakes grow more sophisticated and realistic. As Amy Zegart and Michael Morell warn, "Deception has always been part of espionage and warfare, but not with this level of precision, reach, and speed."[59]

Moving further upstream, lawmakers should intensify measures to ban certain kinds of deepfakes, particularly those stealing an individual's likeness and voice or intentionally intended to deceive the public for nefarious purposes.[60] These efforts should give energy to bipartisan measures that do not seek to target one politician or party more than any other—a focus on the technology, as opposed to the content it produces, might protect any laws from First Amendment challenges. Such measures should also be pushed at an international level, under the auspices of groups such as the United Nations, European Union, and NATO, as well as tech-focused consortiums such as the Christchurch Call. Collaboratives, such as the Content Authenticity Initiative, can also be employed to produce industry best practices and promote cross-platform partnerships.[61] Notably,

57 For a counter-argument, see Bryan C. Taylor, "Defending the state from digital Deceit: the reflexive securitization of deepfake," Critical Studies in Media Communication 38, no. 1 (2021), https://www.tandfonline.com/doi/abs/10.1080/15295036.2020.1833058.

58 See, for example, Jigsaw, "Prebunking Anti-Vaccine Narratives: An Effective Alternative to Debunking Individual False Claims," Medium, March 2, 2022, https://medium.com/jigsaw/prebunking-anti-vaccine-narratives-an-effective-alternative-to-debunking-individual-false-claims-78f0047a8b47.

59 Amy Zegart and Michael Morell, "Spies, Lies, and Algorithms: Why U.S. Intelligence Agencies Must Adapt or Fail," Foreign Affairs 98, no. 3 (May/June 2019), pp. 85-97, https://www.jstor.org/stable/26798154.

60 See, for example, Jack Langa, "Deepfakes, Real Consequences: Crafting Legislation to Combat Threats Posed by Deepfakes Notes," Boston University Law Review 101 (2021): pp. 761-802, https://heinonline.org/HOL/Page?handle=hein.journals/bulr101&id=767&collection=journals&index=.

61 Emmanuelle Saliba and Julia Cherner, "Amid Spread of AI Tools, Advocates for New Digital Standard Say It Would Help Sort Fact from Fiction," ABC News, August 20, 2023. https://abcnews.go.com/Technology/amid-spread-ai-tools-new-digital-standard-users/story?id=102397146.

laws banning deepfakes would have benefits far beyond counter-terrorism, including protecting persons from pornographic blackmail and Western publics from authoritarian disinformation campaigns. A successful law might also provide a framework for efforts to regulate other online harms. If measures to ban nonconsensual deepfakes fail, lawmakers could explore criminal or civil penalties for their production.[62]

Finally, and far more controversially, deepfakes could also be used against extremist groups—by mocking leaders or ideology, "calling off the dogs," and otherwise undermining the conditions allowing extremist groups to thrive. Video taken on January 6th showed rioters reading tweets published by President Trump, in real time, as they determined to push further into the US Capitol complex. Imagine if they had instead seen a deepfake in which Trump denigrated his supporters and their decision to riot, and called for them to leave and accept the election results, thus ending the violence. An endeavour clearly best left for nonstate actors specialising in online wars (democratic governments should steer well clear of such psychological operations on their own populations), the utility of deepfakes themselves in counter-terrorism and countering violent extremism cannot be overlooked.

62 Williams Kaylee, "Tightening Restrictions on Deepfake Porn: What US Lawmakers Could Learn from the U,." Tech Policy Press, October 24, 2023. https://techpolicy.press/tightening-restrictions-on-deepfake-porn-what-us-lawmakers-could-learn-from-the-uk/.

Bibliography

"Artificial Intelligence and Human Rights." United States Senate Committee on the Judiciary, June 13, 2023. https://www.judiciary.senate.gov/committee-activity/hearings/artificial-intelligence-and-human-rights.

Bach-Lombardo, Jordan, and Charlie Winter. "Why Isis Propaganda Works." The Atlantic, February 13, 2016. https://www.theatlantic.com/international/archive/2016/02/isis-propaganda-war/462702/.

Baele, Stephane J., Lewys Brace, and Travis G. Coan. "Uncovering the Far-Right Online Ecosystem: An Analytical Framework and ...Research Agenda." Taylor & Francis Online, December 30, 2020. https://www.tandfonline.com/doi/full/10.1080/1057610X.2020.1862895.

Brooks, Tina, Princess G., Jesse Heatley, Scott Kim, Samantha M., Sara Parks, Maureen Reardon, et al. "Increasing Threat of Deepfake Identities - Homeland Security." Department of Homeland Security, n.d. https://www.dhs.gov/sites/default/files/publications/increasing_threats_of_deepfake_identities_0.pdf.

Byman, Daniel L, Chongyang Gao, Chris Meserole, and V.S. Subrahmanian. "Deepfakes and International Conflict - Brookings." Foreign Policy at Brookings, January 2023. https://www.brookings.edu/wp-content/uploads/2023/01/FP_20230105_deepfakes_international_conflict.pdf?mc_cid=3f7678e334&mc_eid=ee94395166.

Cauberghs, Olivier. "For the Lulz?: AI-Generated Subliminal Hate is a New Challenge in the Fight Against Online Harm." Global Network on Extremism and Technology, November 13, 2023. https://gnet-research.org/2023/11/13/for-the-lulz-ai-generated-subliminal-hate-is-a-new-challenge-in-the-fight-against-online-harm/.

Cook, Jack. "Deepfake Technology: Assessing Security Risk." American University, July 27, 2022. https://www.american.edu/sis/centers/security-technology/deepfake_technology_assessing_security_risk.cfm.

Daniel Siegel and Bilva Chandra, "'Deepfake Doomsday': The Role of Artificial Intelligence in Amplifying Apocalyptic Islamist Propaganda," Global Network on Extremism & Technology, August 29, 2023, https://gnet-research.org/2023/08/29/deepfake-doomsday-the-role-of-artificial-intelligence-in-amplifying-apocalyptic-islamist-propaganda/.

"Facing Reality? Law Enforcement and the Challenge of Deepfakes." Europol Innovation Lab, 2022. https://www.europol.europa.eu/cms/sites/default/files/documents/Europol_Innovation_Lab_Facing_Reality_Law_Enforcement_And_The_Challenge_Of_Deepfakes.pdf.

Fagan, Kaylee. "A Viral Video That Appeared to Show Obama Calling Trump a 'dips---' Shows a Disturbing New Trend Called 'Deepfakes.'" Business Insider, April 17, 2018. https://www.businessinsider.com/obama-deepfake-video-insulting-trump-2018-4.

Flores, Myles. "The New World Order: The Historical Origins of a Dangerous Modern Conspiracy Theory." Middlebury Institute of International Studies at Monterey, May 31, 2022. https://www.middlebury.edu/institute/academics/centers-initiatives/ctec/ctec-publications/new-world-order-historical-origins-dangerous.

Gans, Jared. "FBI Warns of 'deepfakes' in Sextortion Schemes." The Hill, June 7, 2023. https://thehill.com/policy/cybersecurity/4037204-fbi-warns-of-deepfakes-in-sextortion-schemes/.

Gosse, Chandell, and Jacquelyn Burkell. "Politics and Porn: How News Media Characterizes

Problems Presented By ...” Taylor & Francis Online, September 30, 2020. https://www.tandfonline.com/doi/full/10.1080/15295036.2020.1832697.

Helmus, Todd C. “Artificial Intelligence, Deepfakes, and Disinformation: A Primer.” RAND Corporation, July 6, 2022. https://www.rand.org/pubs/perspectives/PEA1043-1.html.

Hoffman, Bruce, Jacob Ware, and Ezra Shapiro. “Assessing the Threat of Incel Violence - Taylor & Francis Online.” Taylor & Francis Online, April 19, 2020. https://www.tandfonline.com/doi/full/10.1080/1057610X.2020.1751459.

Isaac Stanley-Becker and Naomi Nix, “Fake images of Trump arrest show ‘giant step’ for AI’s disruptive power,” Washington Post, March 22, 2023, https://www.washingtonpost.com/politics/2023/03/22/trump-arrest-deepfakes/.

Jacob Ware, “The Third Generation of Online Radicalization,” George Washington University Program on Extremism, June 16, 2023, https://extremism.gwu.edu/third-generation-online-radicalization.

Jensen, Michael, Patrick James, Gary LaFree, Aaron Safer-Lichtenstein, and Elizabeth Yates. “Use of Social Media by US Extremists - UMD.” START.UMD.edu, July 2018. https://www.start.umd.edu/pubs/START_PIRUS_UseOfSocialMediaByUSExtremists_ResearchBrief_July2018.pdf.

Lasarte, Diego. “As Fake Photos of Trump’s ‘Arrest’ Went Viral, Trump Shared an AI-Generated Photo Too.” Quartz, March 23, 2023. https://qz.com/trump-ai-photo-arrest-truthsocial-twitter-1850259197.

Mai, Kimberly T., Sergi Bray, Toby Davies, and Lewis D. Griffin. “Warning: Humans Cannot Reliably Detect Speech Deepfakes.” Public Library of Science, 2023. https://journals.plos.org/plosone/article?id=10.1371%2Fjournal.pone.0285333.

“Male Supremacy.” Southern Poverty Law Center. Accessed August 14, 2023. https://www.splcenter.org/fighting-hate/extremist-files/ideology/male-supremacy.

Mariëtte van Huijstee et al., “Tackling deepfakes in European policy,” European Parliament, July 2021, https://www.europarl.europa.eu/RegData/etudes/STUD/2021/690039/EPRS_STU(2021)690039_EN.pdf.

Matteo Wong, “We Haven’t Seen the Worst of Fake News,” The Atlantic, December 20, 2022, https://www.theatlantic.com/technology/archive/2022/12/deepfake-synthetic-media-technology-rise-disinformation/672519/.

“Meta, X Questioned by Lawmakers Over Lack of Rules Against AI-Generated Political Deepfakes.” CBS News, October 5, 2023. https://www.cbsnews.com/sanfrancisco/news/meta-x-political-deepfakes-lawyers-question-platforms-over-lack-of-rules/.

O’Sullivan, Donie, Curt Devine, and Allison Gordon. “How Antisemitic Hate Groups Are Using Artificial Intelligence in the Wake of Hamas Attacks.” CNN, November 15, 2023. https://www.cnn.com/2023/11/14/us/hamas-israel-artificial-intelligence-hate-groups-invs/index.html.

Pearson, James, and Natalia Zinets. “Deepfake Footage Purports to Show Ukrainian President Capitulating.” Reuters, March 17, 2022. https://www.reuters.com/world/europe/deepfake-footage-purports-show-ukrainian-president-capitulating-2022-03-16/.

Quirk, Caroline. “The High Stakes of Deepfakes: The Growing Necessity of Federal Legislation to Regulate This Rapidly Evolving Technology.” Princeton University, June 19, 2023. https://legaljournal.princeton.edu/the-high-stakes-of-deepfakes-the-growing-necessity-of-federal-legislation-to-regulate-this-rapidly-evolving-technology/.

Ryan, Patrick. "'deepfake' Audio Evidence Used in UK Court to Discredit Dubai Dad." The National, July 4, 2021. https://www.thenationalnews.com/uae/courts/deepfake-audio-evidence-used-in-uk-court-to-discredit-dubai-dad-1.975764.

"S.2559 - Deepfake Task Force Act 117th Congress (2021-2022)." Congress.Gov, August 4, 2021. https://www.congress.gov/bill/117th-congress/senate-bill/2559/text.

Saliba, Emmanuelle, and Jessie DiMartino. "Bill Would Criminalize 'Extremely Harmful' Online 'Deepfakes.'" ABC News, September 25, 2023. https://abcnews.go.com/Politics/bill-criminalize-extremely-harmful-online-deepfakes/story?id=103286802.

Saliba, Emmanuelle, and Julia Cherner. "Amid Spread of AI Tools, Advocates for New Digital Standard Say It Would Help Sort Fact from Fiction." ABC News, August 20, 2023. https://abcnews.go.com/Technology/amid-spread-ai-tools-new-digital-standard-users/story?id=102397146.

Sayler, Kelley M., and Laurie A. Harris. "Deep Fakes and National Security." Congressional Research Service , April 17, 2023. https://www.documentcloud.org/documents/23798946-deep-fakes-and-national-security-april-17-2023.

Schick, Nina. "Introduction." Essay. In Deepfakes: The Coming Infocalypse, 10–11. New York, NY: Grand Central Pub, 2021.

Schmidt, Brittany. "Proposed Wisconsin Bill to Address Artificially Made 'Deep Fake' Pornography." WBay, October 3, 2023. https://www.wbay.com/2023/10/02/proposed-wisconsin-bill-address-artificially-made-deep-fake-pornography/.

Simi, Peter. "Recruitment and Radicalization among U.S. Far-Right Terrorists." Recruitment and Radicalization among U.S. Far-Right Terrorists | START.umd.edu, November 2016. https://www.start.umd.edu/publication/recruitment-and-radicalization-among-us-far-right-terrorists.

Warren, Jess. "Fake Audio of Sadiq Khan Is Not a Crime, Says Met." BBC News, November 11, 2023. https://www.bbc.com/news/uk-england-london-67389609.

Watson, Kathryn. "Trump Tweets Heavily Edited Video of Pelosi Played by Fox Business." CBS News, May 25, 2019. https://www.cbsnews.com/news/trump-tweets-heavily-edited-video-of-pelosi-played-by-fox-news/.

Weiner, Daniel I, and Lawrence Norden. "Regulating AI Deepfakes and Synthetic Media in the Political Arena." Brennan Center for Justice, December 12, 2023. https://www.brennancenter.org/our-work/research-reports/regulating-ai-deepfakes-and-synthetic-media-political-arena.

Westerlund, Mika. "The Emergence of Deepfake Technology: A Review | Tim Review." Technology Innovation Management Review, 2019. https://timreview.ca/article/1282.

"What Is Deep Learning?" IBM, n.d. https://www.ibm.com/topics/deep-learning#:~:text=the%20next%20step-,What%20is%20deep%20learning%3F,from%20large%20amounts%20of%20data.

Williams, Kaylee. "Tightening Restrictions on Deepfake Porn: What US Lawmakers Could Learn from the UK." Tech Policy Press, October 24, 2023. https://techpolicy.press/tightening-restrictions-on-deepfake-porn-what-us-lawmakers-could-learn-from-the-uk/.

Zack Stanton, "The Problem Isn't Just One Insurrection. It's Mass Radicalization." Politico, February 11, 2021, https://www.politico.com/news/magazine/2021/02/11/mass-radicalization-trump-insurrection-468746.

About the Author

Ella Busch

Ella Busch is a researcher at Georgetown University studying Government and Psychology. She has a particular interest in domestic terrorism and hopes to specialise in security in the future.

Jacob Ware

Jacob Ware is a research fellow at the Council on Foreign Relations and an adjunct professor at Georgetown University's Walsh School of Foreign Service and at DeSales University. With Bruce Hoffman, he is the co-author of the forthcoming God, Guns, and Sedition: Far-Right Terrorism in America. He serves on the editorial boards for the academic journal Studies in Conflict & Terrorism and the Irregular Warfare Initiative at the Modern War Institute at West Point.

International Centre for Counter-Terrorism

International Centre for Counter-Terrorism (ICCT)
T: +31 (0)70 763 0050
E: info@icct.nl
www.icct.nl