

“Where you want her, how you want her”

Understanding the violence of deepfake pornography

Amukelani Mnisi¹[0000-0001-5443-8241]

¹ University of the Witwatersrand, Johannesburg, Braamfontein 2000, South Africa
amnisi2000@icloud.com

Abstract. The growing popularity and adoption of generative artificial intelligence (AI) has given rise to the worrying phenomenon of deepfake pornography. A type of synthetic media generated by AI in which a person’s face is superimposed onto existing pornographic material, creating a new, hyper-realistic version in which a new person is depicted in pornographic content they did not actually participate in. Deepfake pornography is often created without the knowledge and consent of those depicted and is a form of technology-facilitated gender-based violence (TFGBV) that overwhelmingly targets women. There is a growing body of literature on the prevalence of deepfake pornography and other forms of TFGBV, however, little effort has been made to understand the purpose and function of this kind of violence against women. This research conceptualises a continuum of TFGBV to illustrate the range of behaviours that constitute TFGBV and identify the common character that underlies these behaviours. An analysis of three cases of deepfake pornography using this continuum reveals that deepfake pornography broadly functions to publicly violate women’s sexual autonomy with the purpose of reinforcing women’s domination and subjugation at the hands of men.

Keywords: Deepfake pornography, technology-facilitated gender-based violence, subjugation, domination.

1 Introduction

In January 2024, the topic “Taylor Swift AI” was trending on X. The global superstar had become the latest famous woman to be victimised by deepfake pornography. A type of synthetic media generated by artificial intelligence (AI) in which a person’s face is superimposed onto an existing pornographic image or video, creating a new, hyper-realistic version in which the new person is depicted performing in pornographic content they did not actually participate in [1, 2]. The fallout was widespread. Her fans flooded the social media platform with positive and supportive posts to overwhelm the pornographic AI-generated images. SAG-AFTRA, the powerful union representing media professionals in the US, condemned the images and vocalised support for legislation to prevent the creation of deepfake pornographic images. Even the Biden administration called for social media companies to play an active role in enforcing their

A. Mnisi

moderation rules to prevent the spread of non-consensual intimate imagery [1]. In response, X (formerly Twitter) wiped all search results of “Taylor Swift nude” and “Taylor Swift AI” from the platform [3].

While those pornographic deepfakes were widely condemned and the social media site quickly moved to erase the images, Swift’s experience greatly differs from the experiences of non-famous women who are victimised by deepfake pornography [4]. When Muslim Indian investigative journalist Rana Ayyub was victimised in 2018, the police attempted to shirk their responsibility and only filed her complaint after she threatened to go to the media with her story. Six months after her complaint, Ayyub had not received any communications or updates from the police on her case [5]. The average woman who discovers that she has become the victim of deepfake pornography can do little else other than submit take down requests to a social media platform that will go ignored [6]. For the many more women who are being victimised on imageboards and websites such as 4chan, they may never know that they are victims of deepfake pornography.

In recent decades, feminist scholars have extensively studied forms of image-based sexual abuse. In this context, deepfake pornography has quickly become seen as an act of technology-facilitated gender-based violence (TFGBV) [7]. Women are overwhelmingly targeted by deepfake pornography, with empirical research suggesting that over 95% of such content features women as the primary subjects [7, 8, 9]. Yet the purpose and function of this violence go largely unquestioned. This paper introduces a continuum of TFGBV as an analytical tool to begin answering the question of why women are being targeted in such a graphic and sexualised way. By placing deepfake pornography and other acts of TFGBV along a continuum, the common character of seemingly disparate acts of TFGBV can be identified and their purpose and function better understood.

Using this continuum to analyse three cases of deepfake pornography, this paper argues that the general function of deepfake pornography is two-fold. Firstly, it publicly violates women’s sexual autonomy. While its secondary function is the creation of an ambient culture of fear in which all women are reminded of their subordinate status and the ever-present threat of victimisation. This paper also argues that the purpose of this violence is to reinforce women’s domination, subjugation and objectification at the hands of men.

The remainder of this paper is organised into five sections. It begins with an explanation of the factors driving the proliferation and consumption of deepfake pornography across the internet and details the process of creating and distributing deepfake pornography on the imageboard 4chan. The three comparative case studies used in this research are also introduced and outlined. The literature review provides an overview of the growing body of research on deepfake pornography, locating this phenomenon in the debates around power, gender, and violence to which it is inextricably linked. It also provides the definitions and conceptual framework used to analyse the three cases

presented. A discussion of the methodology explains why and how a case study methodology was best suited to answer the research question. Analysis of the three case studies using the continuum of TFGBV illuminates how deepfake pornography both emerges from and works in tandem with other acts of TFGBV to target women with sexualised abuse online. The paper concludes with a discussion of the importance of viewing deepfake pornography and other forms of TFGBV as interconnected acts of violence that work to reinforce historical gender inequalities using new and evolving modes of perpetration.

2 Background

The ability to create deepfake pornography using generative AI first gained public attention in 2017 when a deepfake of actress Gal Gadot performing in an incest-themed porn video went viral [10, 11, 12]. It was posted on the online forum Reddit by a user named ‘deepfakes’, who created the video using freely available, open-source machine learning tools [12]. While this early deepfake was derided as unconvincing, it was still viewed as a worrisome harbinger of a growing phenomenon [12]. Since this first high-profile incident of deepfake pornography, the landscape for the creation, distribution and consumption of this content has changed in two significant ways.

The first change is that the technical quality of deepfake pornography has significantly improved. Most AI-generated images and videos are produced through a Generative Adversarial Network (GAN) machine learning model, which can generate new content based on existing data [13]. Inspired by the neural architecture and function of the human brain, GANs have two competing neural networks: a generator and a discriminator [13]. The generator creates realistic data samples while the discriminator distinguishes between real and generated samples, pushing the generator to improve its data generation capabilities until the discriminator can no longer differentiate between the two samples [13]. Initially, GANs were limited to small-sized images to maintain the model’s stability [14]. Since 2017, GAN models have improved with the introduction of the progressive learning training strategy, which expanded the capacity of the generator and discriminator, giving GANs the ability to generate large, high-quality images instead [14]. As a result, GANs can now generate high-resolution, photo-realistic images that, in turn, have made pornographic deepfakes look significantly more convincing [13].

The second change is that there are now far more people creating deepfake pornography across various platforms, as a subset of a broader proliferation of AI tools [15]. A McKinsey & Company survey on the adoption of AI in business, its impact, and investment in AI [16] found that from 2017 to 2022, the general use of AI had more than doubled. After the wildly successful launches of OpenAI’s AI tools in 2022 – one of which was DALL-E, an open source, publicly available AI image generator – big tech companies raced to create their own generative AI tools and integrate generative

A. Mnisi

models into their mainstream products. While these AI tools have safety guidelines and employ various misuse prevention techniques to detect and block users' requests to generate harmful content, they are easily circumvented, especially when the AI model is open-sourced [17]. A greater variety of generative AI tools to choose from and the accessibility and popular adoption of these tools have made it easier for users to generate deepfake pornography and as such, the volume of it has grown. Websites such as Mr DeepFakes, which was shut down in May 2025, hosted forty-three thousand deepfake pornographic videos as of November 2023 [18, 15]. On the website, there is a community of hundreds of deepfake content creators and sellers, with 657 users being responsible for over 95% of the site's video content, while other creators frequently respond to users' requests for custom deepfake content [15]. The content on Mr DeepFakes primarily featured celebrity women, with 95.3% of the individuals depicted being women musicians and actors [15].

However, it is not only famous women who are targets of deepfake pornography. On the infamous imageboard 4chan, the "Adult Requests" board has become a hotbed of deepfake pornography production, distribution, and consumption. On the board, anonymous users upload images of women and request a pornographic deepfake to be made of her. The request often includes some variation of the specific sex acts, positions, and number of sexual partners they want depicted in the deepfake. In some cases, requesters will provide an existing pornographic video they want the targeted woman to appear in. If other users are interested in the request, they will comment on the original request with a "Bump" to express interest and hopefully gain the attention of a user who knows how to create the desired deepfake [19]. Once the deepfake is created, the creator responds to the original post with a publicly accessible link through which the deepfake can be accessed by anyone, even those who are not 4chan users. It is not always clear how the user requesting the deepfake knows the woman whose photo they uploaded, as many of the threads are both anonymous, and the requests do not contain explanations. In many cases, these images were innocuous and unremarkable, the kind that anyone might post on their personal social media. They include photographs of women in everyday scenes, such as at their graduations, hanging out with their friends and families, or at the beach.

Whereas the "Adult Requests" board operates as a casual marketplace for the creation, distribution, and consumption of deepfake pornography ostensibly for personal use, other parts of 4chan have created and distributed deepfake pornography to serve as a weapon to be used against women. For several years beginning in 2014, feminist media critic Anita Sarkeesian became the target of a massive coordinated and vitriolic harassment campaign that originated on 4chan [20]. The harassment campaign was in response to a series of YouTube videos in which Sarkeesian critiqued outdated depictions of women in some video games [20]. To the gaming community and 4chan users – both of which are overwhelmingly male-dominated spaces due in part to the harassment women face when they attempt to enter these spaces – these videos were an attack on them and symbolised a larger feminist encroachment on an internet and gaming culture made for men and by men [21, 20]. One of the core features of the coordinated

Understanding the violence of deepfake pornography

harassment campaign was the creation and distribution of pornographic deepfakes of Sarkeesian that were shared across multiple platforms in a bid to harm her professional reputation [20]. Participants in the harassment campaign shared Sarkeesian's contact information online with the hopes of her being confronted with the video in her personal email accounts; it was their goal to make these pornographic deepfakes of her inescapable [22].

The online harassment of Anita Sarkeesian represents just one context in which deepfake pornography is used, but the three cases analysed in this paper show that deepfake pornography is weaponised against women for a myriad of reasons and to serve many different and context-specific purposes.

2.1 Case studies

Rana Ayyub

In April 2018, investigative journalist Rana Ayyub became the victim of what she called "a deepfake porn plot" intended to silence her [5]. Ayyub is one of India's most prominent and globally recognised journalists and the author of the international best-seller *Gujarat Files: Anatomy of a Cover-Up* that exposed prime minister Narendra Modi's role in the Gujarat riots of 2002 [23]. As a Muslim woman in India and vocal critic of the Modi government, Ayyub is frequently targeted with misogynistic and Islamophobic online abuse [5, 23]. After appearing on an Al Jazeera TV panel where she criticised India's protection of child abusers after the rape of an eight-year-old girl, Ayyub received a pornographic deepfake video of herself from a source within India's ruling Bharatiya Janata Party (BJP) [5]. The deepfake spread across online chat groups maintained by BJP supporters, went viral, and was even shared on the Facebook page of the chief minister of India's largest state [23]. Screenshots from the video were posted in the comments section of Ayyub's social media posts and shared with her family members [5].

Ayyub faced a deluge of sexual harassment across various social media platforms, with people messaging her to enquire about her "rates for sex" and some threatening her with gang rape if she continued to criticise the Modi government [5]. The next day, Ayyub was doxxed with her home address and personal phone number being posted publicly on social media [5, 23].

After an international outcry, which included the United Nations writing to the Indian government, the abuse slowed [24]. However, Ayyub continues to feel its impact. She lives in fear that she could be targeted like this again and, as such, uses social media more cautiously than before [5]. Despite the caution, Ayyub still experiences waves of online abuse, particularly after the publication of any of her reporting or commentary

A. Mnisi

on communal violence, human rights violations, and politics in India [23]. The abuse is misogynistic, often sexually explicit, and seeks to undermine her credibility as a journalist [23].

Taylor Swift

In January 2024, pornographic deepfake images of Taylor Swift circulated across social media, accumulating tens of millions of views [1, 4]. The images depicted Swift in several sexualised positions and partially clothed in Kansas City Chiefs merchandise, the National Football League (NFL) team her boyfriend Travis Kelce plays for.

The origin of the deepfake images was traced back to a community on the imageboard 4chan, which encouraged users to evade the safety policies of AI image generator tools by attempting to create pornographic and sometimes violent images of famous women [25]. It became something of a game on the imageboard with users being encouraged to “be creative,” share tips for generating the pornographic images, and commending each other for good work [25]. Having first appeared in this 4chan community on January 6, 2024, the images appeared eleven days later on the encrypted messaging app Telegram – in a group dedicated to abusive images of women – before appearing on mainstream social media a day later [25].

As an incredibly successful woman in the public eye, Swift has experienced misogyny throughout her career [26]. Critics view her as undeserving of her success, are dismissive of her talents, and express a deep disdain for her public persona [26]. Swift’s public persona is notable for how clearly it is designed to appeal to women and girls, rather than men, and the intense devotion this persona has inspired amongst her global fanbase. [27]. This persona, and the fierce support it engenders amongst Swift’s fanbase of predominantly young women, has been a frequent target of misogynistic derision [26].

However, once Swift’s relationship with Kelce became public knowledge, there was a noticeable increase in the hostile sexism she experienced online [27]. When Swift began attending Kelce’s football games, she brought with her the intense media interest she generates and her large fanbase, which precipitated a 24% increase in women aged 18-24 who watched the 2024 Super Bowl [28]. Where misogynistic rhetoric against Swift and her fans was once largely confined to dedicated online anti-fan communities, it was now also coming from the NFL’s large, and largely male, fanbase as well [27].

While Swift never publicly commented on the pornographic deepfakes made of her, they drew widespread ire and were viewed by some as an attempt by men to put a successful and powerful woman “back in her box” and remind her that no matter how

Understanding the violence of deepfake pornography

powerful she may be, men can still “reduce you to a sex object and there’s nothing you can do about it” [4].

4chan

Since its creation in 2003, 4chan has become infamous in popular culture for its incredibly lenient content moderation policies, controversial subcultures, and alleged hosting of illegal and unethical activity, including deepfake pornography [9, 29]. As previously described, the “Adult Requests” board on 4chan is a hotbed of deepfake pornography creation, distribution, and consumption. While the board is often also used by users attempting to find the names of porn performers and OnlyFans content creators, most requests on the board are requests for the creation of pornographic deepfakes.

The board operates as a kind of impersonal online marketplace. Both the users who upload images of women requesting pornographic deepfakes and the users who create them are anonymous, and no money or exchange for the ‘service’ of creating the deepfake takes place. Compared to other parts of 4chan known for their offensive language and hostile communication, the “Adult Requests” board is relatively tame, and discussion rarely goes beyond users specifying how they want the women depicted in the deepfake (i.e.: topless, legs spread, etc.) and later thanking the user who creates the desired deepfake. On the “Adult Requests” board, the objectification of women is simply routine.

On one thread on the board, dated April 14, 2025, an anonymous user offers their services to others, posting on the board, “I’m making deepfakes for the next 2 hours. Post your brown or black-haired crushes” [30]. All the user asks for is to be told “how you know her, where you want her, how you want her.” The post received numerous responses with one user uploading an image of a woman and expressing a desire to “see this broad’s titties” [30]. Another user posted an image of a woman identified as a former coworker, and a request to see her depicted in a deepfake with her legs spread; the user provided a reference image for the desired deepfake [30]. With no fanfare, the original poster provides the other users with their requested deepfakes, posting the links on the publicly available thread.

Whether women are highly visible celebrities, posting innocuous images of themselves online, or simply doing their jobs, the violence of being targeted with deepfake pornography remains ever-present for all women. As AI tools grow more popular and sophisticated with each year and the volume of deepfake pornography online grows, greater awareness and research of this phenomenon becomes increasingly necessary and urgent.

3 Literature review

The problem of deepfake pornography is inextricably linked to issues of gender, violence, and power. Therefore, furthering the existing research on deepfake pornography and understanding the function and purpose of this violence requires this research to grapple with violence against women and its evolving modes of perpetration, pornography, and the pervasiveness of patriarchal violence in all these issues. The literature reviewed reflects the layers and often interconnected nature of these issues in trying to make sense of deepfake pornography.

The first body of literature this research engages with is violence against women and its changing modes of perpetration. In the growing body of research on deepfake pornography, it is classified as a form of technology-facilitated gender-based violence (TFGBV) generally and an act of image-based sexual abuse specifically [7]. TFGBV is a form of gender-based violence that, like other forms of gender-based violence, is rooted in the systemic reinforcement of gender norms and inequality and uses technology to broaden the scope of violence it inflicts on its victims, primarily women and girls [7]. The use of technology has introduced new forms of abuse, such as ‘revenge porn’ and the creation of deepfake pornography [31]. Acts of TFGBV are considered acts of technology-facilitated violence because, like other technology-facilitated attacks, they have the ability to dominate, intimidate, dehumanise, and create reasonable feelings of fear in those targeted by them [7, 32, 11].

While pornographic deepfakes are the product of a technologically advanced epoch, they are also a continuation of historical patterns of sexual violence that overwhelmingly victimise women. Therefore, it can be placed on the continuum of sexual violence against women theorised by Liz Kelly [33]. Kelly uses the concept of a continuum in two ways. The first is to allow for a general discussion of sexual violence and identify the common character that underlies the different forms of violence as the “abuse, intimidation, coercion, intrusion, threat, and force men use to control women” [33]. The second is to use the continuum to document and name the range of abuse, intimidation, coercion, intrusion, threat, and force men use to control women while acknowledging that there are no defined and distinct analytic categories into which men’s behaviour can be categorised [33]. In this case, the range of the continuum does not denote severity but rather the prevalence of certain acts of sexual violence [33]. Therefore, acts of sexual violence, such as sexual harassment, that most women experience in their lives and are likely to experience on multiple occasions due to these acts being viewed as acceptable male behaviour, are placed at the beginning of the continuum [33]. Other acts of sexual violence, such as rape, which are defined as criminal, appear at the end of the continuum. The continuum does not suggest that there is a linear progression connecting different acts of sexual violence, but does illuminate how common and accepted forms of sexual violence lay the foundation for the perpetration of other, sometimes criminal, forms of sexual violence [33].

Understanding the violence of deepfake pornography

McGlynn et al. [34] build on Kelly's concept of the continuum in the context of ubiquitous technology. They argue that while image-based sexual abuse (IBSA) can and should be placed on Kelly's originally theorised continuum, our shared concept of IBSA itself is unique to a time in history in which technology is an integral part of our everyday lives and offers new modes of perpetration of violence against women [33]. Thus, on the continuum formed by practices and behaviours of image-based sexual abuse, the common character of the various forms of IBSA are: the sexual nature of the imagery; the gendered nature of the abuse; the sexualised nature of the harassment and abuse; the harms as breaches of fundamental rights to dignity, sexual autonomy and sexual expression; and the minimisation of these forms of abuse [34]. Though the two continuums were conceptualised in vastly different socio-technological landscapes, they are united in analysing the harms created by behaviours on their respective continuums, including 'abuse, intimidation, intrusion, threat and force' [33].

From the existing literature on image-based sexual abuse comes an expanded definition of violence that focuses not only on physical violence and its impact but also foregrounds the cumulative impact of psychological attacks [7]. This violence, whether physical and psychological, is rooted in a reality significantly shaped by existing gendered and racialised systems of power. The continuums of sexual violence and IBSA theorised by Kelly [33] and McGlynn et al. [34], respectively, are helpful conceptual tools that ground and locate deepfake pornography in a long history of violence against women and its different modes of perpetration.

In the context of deepfake pornography and the violence of creating, distributing, and consuming it, this research must contend with a second body of literature on the politics of pornography. For Dworkin [35], pornography reinforces and reflects the widespread and socially sanctioned objectification and debasement of women. Moreover, the legitimisation of pornography as "depictions of the erotic" communicates to the viewer that the real pleasure of sex is the debasement of women [35]. While Dworkin's arguments represent only one side of the feminist debate about pornography, in the case of deepfake pornography they are invigorated with renewed relevance because of how this kind of pornography acts as a violation of women's autonomy. Since Dworkin made these arguments, pornography has only grown more violent and more widely accessible. This primes generations of young men who watch pornography to find sexual satisfaction and masculine affirmation in seeing women be sexually humiliated [36].

In many ways, Dworkin already observed and understood that the pornography industry would only continue to grow and with it, the need for more bodies and the sexual brutalisation of those bodies. Dworkin knew that "The numbers [of pictures needed to meet the demand for pornography] grow as the technology and its accessibility grow" [35]. Deepfake pornography is the result of this unrelenting demand, and this research question is an effort to understand the violence that ensues when the demand never ceases, and the technology and accessibility only continue to grow.

A. Mnisi

In the existing research on deepfake pornography, McGlynn et. al [34] have come the closest to making sense of the function and purpose of its violence, albeit solely from the perspective and experiences of victim-survivors. This paper and the methodology employed to answer this question attempt to contextualise incidents and processes of producing, distributing, and consuming deepfake pornography and, in doing so, come closer to understanding the motivations of the perpetrators and the effects of these acts of TFGBV.

3.1 Conceptual framework

A continuum is specifically designed to illustrate range, and when applied to sexual violence against women and the behaviours that constitute image-based sexual abuse, it allows for a general discussion about these types of violence and the discovery of the common character that underlies a range of violent behaviours. Thus, to understand how we arrive at the violence of creating, distributing, and consuming deepfake pornography, a continuum is perfectly suited to present the range of TFGBV and the common character that underlies different acts of TFGBV. Moreover, a continuum demonstrates the relationship between technology, gender, and power – the concepts at the core of this research – and how they operate in tandem to victimise women online.

Unlike the originally theorised continuum of sexual violence, the continuum of TFGBV conceptualised in this research is organised according to severity, as some of the acts listed have the potential to cause serious, offline harms for victims, whether they be physical or psychological.

The acts of TFGBV that make up the first portion of the continuum can be broadly categorised as ‘e-bile.’ E-bile is hostile misogynistic rhetoric laced with profanity and graphic threats of sexual violence [37]. While it is often deeply disturbing to read, e-bile has become a standard discursive tool when internet users want to express their disagreement with or disapproval of women [37]. It is rooted in hostile sexism – negative attitudes and contempt for women, specifically those who are perceived as transgressing traditional gender roles – and is most often directed at women who are highly visible and outspoken in the public sphere, such as woman politicians, although all women can and are targeted as well [37, 27]. Closely related to e-bile, although it should be considered as a separate act of TFGBV, is ‘lascivious contempt’. Lascivious contempt targets women who are perceived as hypersexualised and then derogated as ‘sluts’ who are unsatisfactory because they are too fat, too ugly, etc. [37]. What makes lascivious contempt different from e-bile is that it makes a competition of misogynist rhetoric, usually playing out in social media comment sections where users compete to make the most offensive comment, often escalating to increasingly sexualised and graphic insults [37].

Understanding the violence of deepfake pornography

The midpoint of the continuum of TFGBV is when online behaviour begins to create offline harms such as negatively impacting the targeted person's professional reputation and threatening their current and future employment. The 'Gamergate' coordinated harassment campaign against Anita Sarkeesian is an example of this. The impact of the campaign was felt not only online in the deluge of death and rape threats Sarkeesian received, but also offline. A speech that Sarkeesian was scheduled to deliver at Utah State University was cancelled after the university received an email threatening "a massacre-style attack" if the talk proceeded and a bomb threat was called in at the Game Developer's Choice Awards where Sarkeesian was set to receive an award [38, 39]. These hate campaigns often emerge and escalate in the face of feminist activism and perceived feminist gains in spaces traditionally geared towards men (Jane, 2014). This was the case for Sarkeesian, with a member of the gamer community celebrating her being targeted with deepfake pornography because "she attacked us first" by attempting to critique portrayals of women in some video games [22].

With its overlapping presence in coordinated harassment campaigns, it could be argued that deepfake pornography and other forms of image-based sexual abuse (IBSA) are the next addition to the continuum of TFGBV. In the case of Rana Ayyub, the distribution of deepfake pornographic images of her operated as a coordinated harassment campaign with the impacts being felt both on and offline [5, 23]. These acts of IBSA, which include revenge porn and the non-consensual distribution of intimate images, may be directed at one specific woman at any given moment, but they serve as a reminder to all women that it could happen to them. For victims of IBSA, it creates a near-constant fear that these images will be discovered by family members, future employers, or partners, and thereby negatively affect their futures offline [11, 4].

The furthest end of the continuum of TFGBV would be 'Doxxing,' because of its potential to cause serious physical harm to its targets. Doxxing is the non-consensual publication of another person's private or personal identifying information – such as their phone number and home address – on the internet, often to enable their online harassment [40, 20]. Doxxing is often a feature of coordinated harassment campaigns and is compounded by the other features of harassment campaigns like e-bile, however, it is placed at the furthest end of the continuum for three reasons. Firstly, it gives those who are making violent threats against others on the internet the information needed to actualise those threats and cause real physical harm if they choose [40]. Secondly, once that private information is published, it can be diffused across the internet which may invite further harm against the person doxxed long after its initial publication [40]. Finally, for victims of doxxing, they are aware that because of the spread of their information, there is no foreseeable endpoint to the potential danger being doxxed has put them in [40]. For victims, the harm of being doxxed is not experienced as a discrete event but as a repetitive one, an experience which can become a great source of psychological anguish for them [40].

A. Mnisi

The most cursory examination of the three cases used in this research already reveals the presence of other forms of TFGBV, such as e-bile and coordinated harassment campaigns, in instances where women were targeted with deepfake pornography. That women are being targeted with the same acts of TFGBV suggests that these acts share a common character and are being deployed to serve similar, if not the same, function and purpose. While the continuum of TFGBV comprises seemingly distinct acts of violence, in practice, these acts of violence are combined or escalate to create enormous harm for victims and targets. Without the continuum of TFGBV, deepfake pornography is decontextualized and treated as separate from the behaviours that created and enabled it.

Therefore, a continuum of TFGBV illuminates the range of violence women experience online, as well as how these acts of violence seamlessly shade into each other as they progress along the continuum. Moreover, by placing deepfake pornography on the continuum of TFGBV alongside other acts of violence, we can identify the common character of these acts of violence and ultimately understand the function and purpose of the violence of deepfake pornography.

4 Methodology

This research seeks to understand the function and purpose of the of violence deepfake pornography. This requires understanding the motivations of creators, distributors, and consumers of deepfake pornography and its effects on them, as well as the women targeted. This research uses a case study methodology because it provides the opportunity for multi-perspectival analysis [41]. A case study model also offers the best opportunity to analyse and understand these differing perspectives and to examine them from a macro, societal level.

This research makes use of three case studies to maximise how much can be learnt about the function and purpose of the violence of deepfake pornography. The case studies selected for this research were chosen because they illustrate different contexts in which deepfake pornography was created and allow for reasonable inferences to be made about the different functions and purposes of the creation, distribution, and consumption of deepfake pornography in different contexts.

Each case required different units of observation and offered different sources of information. For Rana Ayyub, this research was able to draw from Ayyub's first-hand account of the events surrounding the deepfake pornography campaign against her and understand its direct effect on her as a victim of this kind of violence [5]. For Taylor Swift, Swift did not make any public statements about the viral deepfakes of herself and thus this research relied on the societal reaction to the deepfakes as it played out on social media. In the case of deepfake pornography creation, consumption, and distribu-

Understanding the violence of deepfake pornography

tion on 4chan, I monitored the “Adult Requests” board for a week to witness and understand the process and sought out older posts on the board to see if there were any changes in the process or in the behaviour and language of users.

While each unit of observation and source illuminated the details and complexities of each case, they weren’t without their own limitations. The use of Rana Ayyub’s first-hand account, while valuable as a primary source, by its nature focuses on one woman’s experience of being targeted with deepfake pornography and cannot and should not be used as representative of all victims’ experiences. For the Taylor Swift case, it is impossible to capture each reaction to the pornographic deepfakes and to keep track of the discussions it sparked, especially in such a fragmented online space like X (formerly Twitter). Similarly, a week of following posts on 4chan could not capture every user interaction on the board, nor could it keep pace with the sheer number of posts being created throughout that week.

Despite these challenges and limitations, a multiple case study methodology is the best approach to answering this research question because it highlights the similarities in seemingly disparate cases. The different sources that came with each case provided the multi-perspectival analysis needed to understand the phenomenon of deepfake pornography not just from the perspective of those involved in it (i.e.: the creators, consumers and victims) but from a broader society who is now aware, perhaps for the first time, of the existence of this form of violence and its implications for their own online safety.

5 Analysis and Discussion

In each of the cases studied in this research, deepfake pornography served a specific and unique function and purpose. An analysis and understanding of these specificities and differences helps us understand the broader function and purpose of the violence of this phenomenon.

For Rana Ayyub, the function of the deepfake pornography campaign against her was to silence and undermine her as a critic of the Modi government. When Ayyub’s case is placed on the continuum of technology-facilitated violence (TFGBV), the distinction between different acts of TFGBV disappears almost entirely. Ayyub became the target of increased misogynistic and Islamophobic e-bile, a harassment campaign with deepfake pornography at its centre, and was later doxxed by social media users. These acts of TFGBV were simultaneously deployed against Ayyub to form a multi-pronged assault that is increasingly used against women journalists around the world to shame, silence, undermine, and discredit them professionally [42]. While Ayyub had grown used to e-bile as an outspoken critic, the experience of being targeted by deepfake pornography impacted her so deeply that she remains cautious of what she posts online and self-censors as a means of protecting herself from being targeted again [5].

A. Mnisi

When Taylor Swift was targeted with deepfake pornography, those deepfakes functioned to reassert masculine dominance in the face of perceived feminist encroachment on male spaces. Taylor Swift's public persona, success, and deeply devoted fanbase had long attracted misogynistic derision and hostile sexism from dedicated anti-fan communities online [26]. However, that hostile sexism had never before prompted an act of violence such as the creation and distribution of pornographic deepfakes until Swift began dating NFL player Travis Kelce. When Kelce and Swift's relationship became public knowledge and Swift began attending his games, she brought her fanbase along with her and precipitated a 24% increase in women aged 18-24 who watched the 2024 Super Bowl [28]. The increased presence and participation of young women in the male-dominated football fandom was also compounded by the subversion of traditional gender norms that Swift's presence represented [27]. Despite playing the role of devoted girlfriend to a decorated athlete, Swift's status as a billionaire cultural juggernaut positions her as the proverbial breadwinner in her relationship and in violation of the masculinist ideals found amongst the NFL's fanbase [27]. Thus, when Swift was targeted with pornographic deepfakes, it read to women online as an attempt to place a powerful woman "back in her box" and reduce her to little more than a sex object [4]. Moreover, to have Swift wearing the merchandise of her boyfriend's football team in those deepfakes could be read as an attempt to define Swift and her sexuality in relation to the man in her life, asserting his masculine dominance over her and challenging the independence she enjoys in reality. While the use of pornographic deepfakes was specifically designed and targeted to negatively affect Rana Ayyub, the pornographic deepfakes of Taylor Swift affected a much larger audience. It is impossible to know how the deepfakes impacted Swift – she never made a public statement – but for the women online who saw them, it was a frightening demonstration of the unique violence women face online and emphasised the ever-present threat of victimisation for all women.

For both Rana Ayyub and Taylor Swift, the weaponisation of pornographic deepfakes against them was spurred by external events. Both women were frequent targets of hostile sexism and had experienced other acts of TFGBV such as e-bile, but the creation and distribution of pornographic deepfakes were an escalation in response to other events. For Ayyub, it was speaking out against sexual violence against women and children in India and her public perception amongst Hindu nationalists as a hired anti-Hindu and anti-Modi mouthpiece [23]. For Swift, it was her highly publicised and influential entrance into the male-dominated world of the NFL [27]. In both cases, the women's behaviour and presence were deemed as transgressing traditional gender norms and threats to patriarchy that had to be neutralised. Where other acts of TFGBV had failed to curb the women's success and outspokenness, deepfake pornography was deployed as a new and deeply intimate act of violence that would not only directly impact these women but would send a clear message to all women.

Unlike the other two cases, the creation, distribution, and consumption of deepfake pornography on 4chan does not represent an escalation of violence but rather an evolution of a prevailing culture of misogyny. On 4chan, a standard response to users finding

Understanding the violence of deepfake pornography

out that a woman is participating in a discussion or another user claims to be a woman is to demand that she show ‘tits or get the fuck out’ [20]. On 4chan, these acts of TFGBV are normalised and are merely the cost of entry. As previously described, the “Adult Requests” board is largely free from the hostile communication that characterises other parts of the infamous website. When there is brief discussion on the board, as the process of creating and distributing deepfakes takes place, users often employ the violent and degrading language that is frequently used in pornography. Women are described as “sluts” and requests are made for the women to be depicted in deepfakes enduring violent sex acts like “face fucking” [19]. This language and the creation of deepfake pornography are not out of place on a website that was the epicentre of the harassment campaign against Anita Sarkeesian and was the origin point of the pornographic deepfakes of Taylor Swift [25]. What is different about the process taking place on the “Adult Requests” board is that it appears almost purposeless. It has become mundane for users on the board to strip women of their autonomy and treat them as publicly available objects whose identities and likeness can be hijacked to fulfill any manner of sexual fantasies. In the case of 4chan, deepfake pornography functions to normalise women’s sexual objectification.

Despite the differences in each case, when placed on the continuum of TFGBV, it becomes clear that deepfake pornography is an act of violence that often works in tandem with other acts of TFGBV. Moreover, in each instance in which it is used, common characteristics emerge: the sexual nature of the imagery used to target women; the gendered nature of the abuse; the sexualised nature of the harassment and abuse; the harms as breaches of fundamental rights to dignity, sexual autonomy, and sexual expression [34]. Whether deepfake pornography is being used to silence and undermine political critics, reassert masculine dominance, or reinforce the mundanity and inevitability of misogyny, it does so by relying on sexualised imagery of women designed to reflect their objectification and subjugation at the hands of men and uses a mode of perpetration that leaves all women vulnerable to victimisation. In these cases, we see that the broad function of deepfake pornography is to sexually humiliate women and publicly violate their sexual autonomy with the purpose of reinforcing women’s domination and subjugation at the hands of men.

6 Conclusion

The range of available generative AI tools and their growing popularity have led to a significant increase in the volume of deepfake pornography online. While women in the public eye are often the most at risk of being targeted, all women are at risk of victimisation. The purpose of deepfake pornography is to create an ambient fear of victimisation in all women and thus reinforce their subordination and objectification at the hands of men. Deepfake pornography serves this purpose by functioning as an act of sexual humiliation and violation by depicting women – particularly women whose success or outspokenness challenge traditional gender roles – in pornographic content

A. Mnisi

that reinforces their role as sexual objects and vessels of male sexual desires and aggression.

The cases of Rana Ayyub, Taylor Swift, and the “Adult Requests” board on 4chan represent the different contexts in which deepfake pornography is used and weaponised. Analysis of these cases using the continuum of TFGBV reveals how deepfake pornography and other acts of TFGBV are often simultaneously deployed against women online because of their common character and goal of breaching women’s fundamental rights to dignity, sexual autonomy, and sexual expression. The continuum also highlights the importance of viewing deepfake pornography and other forms of TFGBV as interconnected and evolving acts of violence that are used to terrorise women online and reinforce the inevitability of their victimisation as a means of fortifying entrenched gender inequalities online.

References

1. Hibberd, J., Phillips, Z.G. (2024) Taylor Swift AI Nudes Provoke Fandom Uproar on X: “Disgusting as Hell”. The Hollywood Reporter. <https://www.hollywoodreporter.com/news/general-news/taylor-swift-ai-nudes-twitter-x-1235806979>, last accessed 2025/05/29.
2. Öhman, C: Introducing the pervert’s dilemma: a contribution to the critique of Deepfake Pornography. *Ethics and Information Technology*, 133-140 (2019).
3. Piña, C. (2024) Taylor Swift Search Re-enabled on X Following AI Nudes. The Hollywood Reporter. <https://www.hollywoodreporter.com/news/general-news/taylor-swift-search-re-enabled-x-ai-nudes-1235810070>, last accessed 2025/03/17.
4. Saner, E. (2024) Inside the Taylor Swift deepfake scandal: ‘It’s men telling a powerful woman to get back in her box.’ The Guardian. <https://www.theguardian.com/technology/2024/jan/31/inside-the-taylor-swift-deepfake-scandal-its-men-telling-a-powerful-woman-to-get-back-in-her-box>, last accessed 2025/03/21.
5. Ayyub, R (2018) I Was The Victim Of A Deepfake Porn Plot Intended To Silence Me: Rana Ayyub. HuffPost. https://www.huffpost.com/archive/in/entry/deepfake-porn_in_5c1201cfe4b0508b213746bd, last accessed 2025/02/15.
6. Human Rights Watch (2021) “My Life is Not Your Porn”: Digital Sex Crimes in South Korea. Human Rights Watch, New York.
7. Dunn, S.: Technology-Facilitated Gender-Based Violence: An Overview. Centre for International Governance Innovation, Waterloo (2020)
8. Security Hero (2024) State of Deepfakes: Realities, Threats, and Impacts. <https://www.securityhero.io/state-of-deepfakes/>, last accessed 2025/03/20.
9. Deeptrace (2019) The State of Deepfakes: Landscape, Threats, and Impact. https://regmedia.co.uk/2019/10/08/deepfake_report.pdf, last accessed 2025/04/20.
10. Wagner, T.L., Blewer, A.: “The Word Real Is No Longer Real”: Deepfakes, Gender, and the Challenges of AI-Altered Video. *Open Information Science* 2019(3), 32 – 46 (2019) <https://doi.org/10.1515/opis-2019-0003>.
11. Brieger, A. R.: Empowerment or exploitation: A qualitative analysis of online feminist communities’ discussions of deepfake pornography. Two-year Master's Thesis. Uppsala Universitet (2024).
12. Cole, S. (2017) AI-Assisted Fake Porn Is Here and We’re All Fucked. Vice. <https://www.vice.com/en/article/gal-gadot-fake-ai-porn>, last accessed 2025/02/27
13. Banh, L., Stobel, G: Generative artificial intelligence. *Electronic Markets*, 62-79 (2023)
14. Li, C., Zhuang, B., Wang, G., Liang, X., Chang, X., Yang, Y. (2024) Automated Progressive Learning for Efficient Training of Vision Transformers. Paper presented at Conference on Computer Vision and Pattern Recognition, Seattle, United States, 2024, <https://doi.org/10.48550/arXiv.2410.00350>.
15. Han, C., Li, A., Kumar, D., Durumeric, Z.: Characterizing the MrDeepFakes sexual deepfake marketplace. In: Proceedings of the 34th USENIX Conference on Security Symposium, pp. 5169–5188. USENIX Association, Seattle (2025).
16. McKinsey & Company. (2024) What is generative AI? <https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-generative-ai>, last accessed 2025/04/15.
17. Clifford, B. (2023) Preventing AI Misuse: Current Techniques. Centre for the Governance of AI. <https://www.governance.ai/analysis/preventing-ai-misuse-current-techniques>, last accessed 2025/04/23.

A. Mnisi

18. Wise, A. (2025) Major deepfake porn site shuts down. NPR. <https://www.npr.org/2025/05/06/nx-s1-5388422/mr-deepfakes-porn-site-ai-shut-down>, last accessed 2025/10/3.
19. 4chan. n.d.
20. Nagle, A: *Kill All Normies: Online Culture Wars from 4chan and Tumblr to Trump and the Alt-Right*. Zero Books: United Kingdom (2017).
21. Andersen, I. V.: *Hostility online: Flaming, trolling, and the public debate*. *First Monday* 26(3). <http://dx.doi.org/10.5210/fm.v26i3.11547> (2021)
22. Harwell, D. (2018) Fake-porn videos are being weaponized to harass and humiliate women: 'Everybody is a potential target'. *The Washington Post*. <https://www.washingtonpost.com/technology/2018/12/30/fake-porn-videos-are-being-weaponized-harass-humiliate-women-everybody-is-potential-target/>, last accessed 2025/03/21
23. Posetti, J., Bontcheva, K., Zaffar, H., Shabbir, N., Maynard, D., Pandya, M.: *Rana Ayyub: Targeted online violence at the intersection of misogyny and Islamophobia*. International Center for Journalists: Washington DC (2023).
24. United Nations (2018) UN experts call on India to protect journalist Rana Ayyub from online hate campaign. United Nations Office of the High Commissioner for Human Rights. <https://www.ohchr.org/en/press-releases/2018/05/un-experts-call-india-protect-journalist-rana-ayyub-online-hate-campaign>, last accessed 2025/03/10.
25. Hsu, T. (2024) Fake and Explicit Images of Taylor Swift Started on 4chan, Study Says. *The New York Times*. <https://www.nytimes.com/2024/02/05/business/media/taylor-swift-ai-fake-images.html>, last accessed 2025/03/25.
26. Kinnunen, L. (2025) In: Aker, P., Kaun, A. (eds) *Gendering Media: Framing of AI, Interacting with ChatGPT and Anti-fandom*. Södertörn University: Stockholm.
27. Elder, L., Greene, S., Gulatti, J., Lizotte, M. (2025) *The Politics of Taylor Swift: Partisanship, Gen Z, and Sexism*. Paper presented at Southern Political Science Association Conference 2025, San Juan, Puerto Rico.
28. Feldman, J. (2024) One Year Later, Taylor Swift Effect Stretches Far Beyond NFL. *Sportico*. <https://www.sportico.com/personalities/people/2024/taylor-swift-effect-nfl-attendance-girls-social-marketing-1234798352/>, last accessed 2025/05/1.
29. Satter, R. (2025) Notorious messageboard 4chan has been hacked, posts claim." *Reuters*. <https://www.reuters.com/technology/cybersecurity/notorious-internet-messageboard-4chan-has-been-hacked-posts-claim-2025-04-15/#:~:text=It%20has%20long%20been%20notorious,nonconsensual%20pornography%20%2C%20opens%20new%20tab..>, last accessed 2025/04/16.
30. Anonymous. (2025) "Making Deepfakes of your brunettes" [4chan post]. 4chan.
31. Dunn, S. (2020) *Is it Actually Violence? Framing Technology-Facilitated Abuse as Violence*. In: Bailey, J.F. (ed) *The Emerald International Handbook of Technology-Facilitated Violence and Abuse*. Emerald Publishing Limited: Leeds.
32. McGlynn, C., Johnson, K., Rackley, E., Henry, N., Gavey, N., Flynn, A., Powell, A. 'It's Torture for the Soul': The Harms of Image-Based Sexual Abuse. *Social & Legal Studies*, 541-562 (2021).
33. Kelly, L.: *Surviving Sexual Violence*. University of Minnesota Press, Minneapolis (1988).
34. McGlynn, C., Rackley, E., Houghton, R.: *Beyond 'Revenge Porn': The Continuum of Image-Based Sexual Abuse*. *Feminist Legal Studies*, 25-46 (2017).
35. Dworkin, A.: *Pornography: Men Possessing Women*. Penguin, New York (1981).
36. DeKeseredy, W. S., Schwartz, M. D.: *Thinking Sociologically About Image-Based Sexual Abuse: The Contribution of Male Peer Support Theory*. *Sexualization Media & Society*, 1-8 (2016).

Understanding the violence of deepfake pornography

37. Jane, E.A.: 'Back to the kitchen, cunt': speaking the unspeakable about online misogyny. *Continuum: Journal of Media and Cultural Studies* 28(4), 558-570 (2014) <http://dx.doi.org/10.1080/10304312.2014.924479>.
38. Kelion, L. (2014) Feminist video-games talk cancelled after massacre threat. BBC. <https://www.bbc.com/news/technology-29626809>, last accessed 2025/05/25
39. Seppala, T.J. (2014) GDC award for culture critic Anita Sarkeesian led to bomb threat. *Engadget*. <https://www.engadget.com/2014-09-17-anita-sarkeesian-gdc14-bomb-threat.html>, last accessed 2025/05/30
40. Anderson, B., Wood, M.A.: Harm Imbrication and Virtualised Violence: Reconceptualising the Harms of Doxxing. *International Journal for Crime, Justice and Social Democracy*, 11(1), 196-209 (2022).
41. Tellis, W.M.: Application of a Case Study Methodology. *The Qualitative Report* 3(3) (1997).
42. International Center for Journalists. (2022). *The Chilling: A global study of online violence against women journalists*. International Center for Journalists: Washington DC.