It's Everyone's Problem: Mainstreaming Responses to Technology-Facilitated Gender-Based Violence

By Nina Jankowicz, Isabella Gomez-O'Keefe, Lauren Hoffman, and Andrea Vidal Becker

SIPA | IGP Institute of Global Politics

VITAL VOICES GLOBAL PARTNERSHIP

CONTENTS

It's Everyone's Problem: Mainstreaming Responses to Technology-Facilitated Gender-Based Violence

By Nina Jankowicz, Isabella Gomez-O'Keefe, Lauren Hoffman, and Andrea Vidal Becker

EXECUTIVE SUMMARY

This report assesses the state of research on TFGBV as well as recent global legislative, regulatory, and policy progress made on this issue. Through a case study that explores the data around the online harassment and abuse of Australian eSafety Commissioner Julie Inman Grant, the report documents the real-life effects of TFGBV for women in public life. We argue that TFGBV is not an intractable problem, but one that must be mainstreamed to be mitigated, centering women's experiences in broader policy debates. TFGBV must no longer be the responsibility solely of women's advocacy groups. Technology companies, governments, civic tech organizations, law enforcement, employers, schools, and others must mainstream their work to combat TFGBV to reflect its mainstreamed effects on society. To this end, we recommend a number of practical solutions to the specific and pressing issues that women and girls face online today. Addressing the urgent changes described here will not only make women and girls safer and ensure their voices are heard, but also improve the safety and free expression for everyone who uses the internet, building more robust, representative democracies.

INTRODUCTION

On the morning of April 22, Julie Inman Grant, Australia's eSafety Commissioner, woke up for the seventh day in a row as the target of technology-facilitated gender-based violence (TFGBV). The abuse had begun a week earlier, on April 15, when she issued a takedown notice to the social media platform X (formerly Twitter), requiring that the platform remove a video of a violent stabbing of a bishop that had taken place at a church in Wakeley, Australia. The graphic video was in violation of Australian law and likely in violation of the platform's own terms of service. The irony that the public servant in charge of implementing Australia's online safety laws would be targeted with the very harm she was charged with ameliorating was not lost on Inman Grant.

"She's a goner," read one tweet mentioning her in the early hours of that morning. "Not long before we round up these [World Economic Forum] scumbag traitors."[1]

In the 48 hours that followed, Inman Grant's personal information was released publicly. The identity of her children was exposed. Tens of thousands of instances of abusive content—including rape and death threats—would be directed toward her and her family. In our original analysis that we describe later in this report, our research team found that over 10 percent of the negative content directed at Inman Grant contained gendered narratives and slurs.

Inman Grant is far from alone; the world finds itself in a critical moment in the protection of women's right to free expression and political and civic participation during 2024's "year of elections." Indeed, a 2020 study by the Wilson Center on gendered disinformation looked at hundreds of thousands of pieces of online disinformation and abuse directed at women in public life, including vitriolic speech and sexualized deepfakes during the lead-up to the 2020 U.S. elections, and found that not only did the disinformation and abuse overwhelmingly focus on just 13 female candidates—both Republican and Democrat—but also that 78 percent targeted then Senator and candidate Kamala Harris, demonstrating the intersectional nature of this sort of abuse.[2]

This year, the type of abuse launched at Kamala Harris as she runs as the Democratic nominee for the U.S. presidency[3] and other female candidates around the world is no different; however, the scale and variety of abuse are now much larger as abusers enjoy readily accessible and underregulated deepfake tools, generative artificial intelligence (AI), and 'nudify' apps. Vitriolic, hateful, and threatening language used against women in public life has become normalized by male politicians and media personalities who employ it mostly without consequence and inspire their online followers to replicate it across the internet.

The integrity of elections and democratic processes is therefore undermined by the attacks levied toward female politicians: without access to trustworthy information about women running for office, voters cannot make informed decisions at the ballot box.[4] Online abuse also affects women's decisions to pursue elected office in the first place.[5] For example, Slovakia's former president, Zuzana Čaputová, who received death threats online, said she was not pursuing re-election for personal reasons.[6]

This silencing effect is not only present among adult women but also amongst girls: a 2020 survey by Plan International of 14,000 girls in 22 countries found that of the 98 percent who use social media, more than half reported being "attacked and harassed" online—in many cases, before they were even old enough to vote.[7] As a result of those attacks, "19 percent of girls who were harassed very frequently said they use the social media platform [where they were harassed] less and 12 percent just stopped using it."[8] It is furthermore well-documented that online abuse and harassment on social media has worrying effects on girls' mental health.[9] [10]

On social media platforms, the checks and balances introduced to mitigate and allow the study of TFGBV and other online harms have been rolled back, or in some cases, abandoned entirely. After Elon Musk's purchase of X, the platform's Trust and Safety Council was disbanded,[11] and key teams, including those monitoring human rights and AI ethics, were cut.[12] Meanwhile, several large platforms, including X, Reddit, and Facebook, have ended[13] [14] or monetized[15] access to their application programming interfaces (APIs) for data access, rendering journalists' and researchers' monitoring of TFGBV and other online harms much more difficult.[16] X's API shutdown also closed Block Party, a service that relied on the API to assist users experiencing high levels of vitriol to automatically mute or block people sending them hate online, allowing targets of harassment to continue expressing themselves without needing to subject themselves to further abuse.[17] Overall, in the past two years, the fraying of technological support, trust, safety, and data access problems have proliferated, with broad backlash to content moderation as 'censorship.'[18]

Finally, as generative AI has become widely accessible, so, too, has the proliferation of non-consensual intimate imagery (NCII) in the form of deepfake image-based sexual abuse, also referred to as deepfake pornography.[19] While reliable data on the topic remains difficult to collect, recent studies estimate that 98 percent of all deepfake videos online are deepfake pornography, and that 99 percent of those targeted by deepfake pornography are women.[20] A July 2024 report from Ofcom, the United Kingdom's online safety regulator, found that 43 percent of people aged 16 plus say they have seen at least one deepfake online in the last six months and, of the adults who had seen deepfakes, 14 percent had seen sexualized deepfake content.[21]

This report assesses the state of research on TFGBV as well as recent global legislative, regulatory, and policy progress made on this issue. Through a case study that explores the data around the online harassment and abuse of Australian eSafety Commissioner Julie Inman Grant, the report documents the real-life effects of TFGBV for women in public life. We argue that TFGBV is not an intractable problem, but one that must be mainstreamed to be mitigated, centering women's experiences in broader policy debates. TFGBV must no longer be the responsibility solely of women's advocacy groups. Technology companies, governments, civic tech organizations, law enforcement, employers, schools, and others must mainstream their work to combat TFGBV to reflect its mainstreamed effects on society. To this end, we recommend a number of practical solutions to the specific and pressing issues that women and girls face online today. Addressing the urgent changes described here will not only make women and girls safer and ensure their voices are heard, but also improve the safety and free expression for everyone who uses the internet, building more robust, representative democracies.

BACKGROUND

WHAT IS TECHNOLOGY-FACILITATED GENDER-BASED VIOLENCE (TFGBV)?

Technology-facilitated gender-based violence (TFGBV) is "any act that is committed, assisted, aggravated, or amplified by the use of information communication technologies or other digital tools, that results in or is likely to result in physical, sexual, psychological, social, political, or economic harm, or other infringements of rights and freedoms."[22] This sort of abuse is associated with a number of direct, continuously evolving forms of violence, such as image-based abuse—including distribution of non-consensual intimate imagery (NCII) — "any scenario in which intimate content is being produced, published or reproduced without consent"[23] — AI-generated NCII or deepfake pornography, the sharing of personal information or doxxing, impersonation, catfishing,[24] threats of violence, dogpiling,[25] stalking and monitoring, cyber surveillance including device or app control,[26] and harassment and abuse, gendered hate speech, misogynoir,[27] and gendered disinformation.[28] These forms of violence not only cause distress to victims and survivors but can also jeopardize their physical safety offline and threaten their livelihoods by chilling participation in the workforce and in public life.[29]

TFGBV is perpetuated by a variety of people and groups in women's lives, including strangers,[30] people they know,[31] far-right groups,[32] informal networks (including the online 'manosphere'),[33] political parties,[34] governments,[35] and foreign state actors.[36] Risks and preventive factors of TFGBV vary based on the perpetrator and scale of attack. Individual attacks may stem from having a violent nature or negative attitudes about women, while large-scale, state-backed campaigns are instead connected to broader issues such as the 'health' of democratic processes and national security.[37] Similarly, reports find that marginalized, minority women are the most at-risk for experiencing TFGBV.[38] Despite these divergences, there are a number of societal factors which contribute to the different forms of TFGBV. Broadly speaking, the societal normalization of misogyny, sexism, patriarchy, and gender inequality directly contributes to TFGBV on all levels, applying to most — if not all — cases.

Deepfake Image-Based Sexual Abuse: Deepfake Pornography

In 2017, journalists Samantha Cole and Emanuel Maiberg discovered a Reddit user named "deepfakes" using a machine learning algorithm to swap celebrities' faces onto pornographic performers' bodies, marking the newest emerging form of TFGBV.[39] Since 2017, the technology to create deepfakes has become even easier with the creation of 'nudify' apps that enable users to forge nude versions of pictures of real women and recent advancements in generative AI that enable users to quickly create convincing fake videos and images of their subject with as little as a single photo as a source.[40] In turn, the volume of deepfakes online has increased exponentially: according to a 2023 report by Security Hero, "the total number of deepfake videos online in 2023 [was] 95,820, representing a 550% increase over 2019."[41] Deepfake pornography is part of a broader framework of NCII, which includes revenge porn and child sexual abuse material.

WHAT WE KNOW ABOUT TFGBV

As our lives have become increasingly digital, social media platforms and other online services have become as vital to daily life as any other utility. Likewise, social platforms, tools, and even gaming apps have become a mirror host to the perennial real-world harm of gender-based harassment and abuse, an area in which scholars and civil society are working to keep the rate of our understanding in line with the rapid technological advancements.

This in turn has led to increased attention from governments, multilateral organizations, and think tanks, including from the Global Partnership for Action on Gender-Based Online Harassment and Abuse,[42] the White House and its Task Force to Address Online Harassment and Abuse,[43] United Nations Educational, Scientific, and Cultural Organization (UNESCO)[44] [45], United Nations Population Fund (UNFPA),[46] the Wilson Center,[47] and the Carnegie Endowment for International Peace,[48] among many others.

The findings across the board are clear: TFGBV is widespread, and women — especially women, girls, and LGBTQI+ individuals who face intersecting discrimination on the basis of their race and ethnicity, disability, or sexual orientation — experience higher levels of online harassment and abuse.[49]

A 2021 Economist Intelligence Unit report found that between 2019 and 2020, 85 percent of women globally had witnessed or experienced online violence, with 38 percent having personally been affected.[50] The report found that women face high levels of online abuse across every continent — with minor variation — with the Middle East (98 percent), Latin America and the Caribbean (91 percent), and Africa (90 percent) at the highest end of the range, and Asia Pacific (88 percent), North America (76 percent), and Europe (74 percent) at the lowest.[51] Of the online abuse women face, the forms vary: the most common reported types of online abuse faced by women are misinformation and defamation (67 percent), cyberharassment (66 percent), hate speech (65 percent), and impersonation (63 percent).[52]

TFGBV has detrimental effects, often silencing the women it targets and discouraging them from engaging in online spaces.[53] One 2022 poll reported that 24 percent of women worldwide reduced their internet usage due to harmful online content.[54] The poll found that three of the most common reactions to harmful online contacts recorded were: lower self-esteem, panic attacks, anxiety, or stress, and stopping from saying what they actually think online.[55] The mental health effects of TFGBV should not be underestimated: a 2016 qualitative study focused on the mental health effects of revenge porn survivors found symptoms and mental health issues similar to those of women who had experienced sexual assault: post-traumatic stress disorder, depression, anxiety, and other mental health challenges.[56]

However, statistics may still not fully reflect the scope of the issue: women are often unaware of their options or hesitant to self-report TFGBV they have experienced,[57] face barriers to justice when they do,[58] and often are influenced by societal norms that normalize these forms of violence, causing TFGBV victims and survivors to discount the harms they are facing. Furthermore, deepfakes have proven deeply detrimental to women's livelihoods, job prospects, and community standings.[59]

TFGBV not only affects women's health, but it also affects women's participation in political life, thereby undermining democracy at large.[60] Globally, the Inter-Parliamentary Union conducted a survey of women parliamentarians, where 41.8 percent had had "extremely humiliating or sexually charged images of [themselves] spread on social media."[61] Reporting on the recent June 2024 UK elections revealed that online abuse of women candidates was on the rise.[62] A candidate survey from the UK Electoral Commission even found that 40 percent of female candidates reported that they avoided campaigning on their own to avoid harassment, intimidation, and threats, compared to only 11 percent of male candidates.[63]

This persistent abuse of women in politics has far-reaching consequences: young women are now less likely to run for office because of how women in politics are harassed and abused online.[64] The rise of generative AI has also resulted in growing worries that gendered disinformation targeting politicians will take more pertinent and pervasive forms during upcoming elections, and has resulted in a flurry of laws in many countries targeting mis- and disinformation around elections.[65] In turn, this growing threat to democracy is further exacerbated by attacks against female journalists—a phenomenon extensively documented by Julie Posetti in an International Center for Journalists (ICFJ) report that analyzed millions of social media posts directed at female journalists around the world over three years and found that online and physical violence work in a negative feedback loop, ultimately undermining trust in institutions, ruining women's careers, and amplifying misogyny for political aims.[66]

Globally, TFGBV is also increasingly being used as a political weapon. Around the world, TFGBV has been leveraged to hinder political opposition in countries, including Brazil,[67] Kenya,[68] Ghana,[69] Georgia,[70] India,[71] and Hungary.[72] While each of these cases take place in very different contexts around the world, there is one common thread: female politicians around the world face targeted, gendered disinformation campaigns with attackers ranging from male counterparts, authoritarian regimes, online right-wing groups, and misogynistic detractors on social media. This sort of high-profile abuse dampens political participation for other women, too: research conducted by the Turing Institute found that "three quarters (77 percent) of women are not comfortable expressing political opinions online because of fears they will be targeted by harmful online behaviors such as misogyny, trolling, threats and harassment."[73]

At the geopolitical scale, TFGBV has become yet another weapon for malign actors, such as foreign-backed entities, to carry out broader interference operations abroad. These trends demonstrate that TFGBV is not only an issue of domestic politics, but also one inherently related to national security.[74] Perhaps one of the most notable examples comes from Russian-backed gendered disinformation campaigns, which have been observed around the world. For example, during the 2021 elections in Germany, the Green Party's candidate, Annalena Baerbock, was

Global Partnership for Action on Gender-Based Online Harassment and Abuse

Comprised of fourteen governments, including Australia, Canada, Chile, Denmark, France, Iceland, Kenya, Mexico, New Zealand, the Republic of Korea, Spain, Sweden, the United Kingdom, and the United States, the Global Partnership for Action on Gender-Based Online Harassment and Abuse was announced at the 2021 U.S. Summit for Democracy and launched at the 66th United Nations Commission on the Status of Women in 2022 to convene international organizations, civil society, and the private sector to prevent and address TFGBV.[77] As part of their membership, member states have committed resources to prevent and address TFGBV and are prioritizing expanding research and data collection on TFGBV. Several member states have made individual commitments, such as the U.S.' $15 million in targeted foreign assistance programs since the Global Partnership's launch to respond to TFGBV and the future launch of the Global TFGBV Rapid Response Fund.[78]

specifically targeted and undermined by Russian-backed accounts online, using gendered disinformation that outsized that which her male counterparts received.[75] Though Germany's Network Enforcement Act (NetzDG), one of the world's toughest laws against online hate speech and harassment, was in force during Baerbock's campaign, it was not enough to stop the levels of online abuse she received, in part due to "coordinated sharing, lack of enforcement and oversight, and content deemed legal by social media platforms."[76]

As the research demonstrates, TFGBV is a wide-ranging, ever-evolving, and prevalent issue that will require global collaboration. The work of scholars, civil society, global institutions, and governments has been essential for mapping the issue of TFGBV and has led to wide-scale acceptance that TFGBV is a prominent, multinational issue. However, TFGBV still remains sidelined as a women's issue in research, policy, practice, and in the public eye.

CASE STUDY

CREATING AN ENEMY IMAGE: THE TFGBV CAMPAIGN AGAINST AUSTRALIAN ESAFETY COMMISSIONER JULIE INMAN GRANT

Exploring the data around the harassment and abuse of Australian eSafety Commissioner Julie Inman Grant in April through June 2024 demonstrates the enormous reach and impact of TFGBV, even against those in positions to fight it. Furthermore, her experience exemplifies the creation and amplification of a gendered enemy image,[79] an academic framework[80] used to understand language deployed on a large-scale via popular media, using stereotypes,[81] dehumanization,[82] and framing the enemy as perpetrating a loss or threat[83] to target individuals or groups, typically from marginalized backgrounds. Enemy images use a number of tactics associated with hate speech, and can be better contextualized through Susan Benesh's dangerous speech framework.[84]

Inman Grant, a technology policy professional who served as an executive at Microsoft, Twitter, and Adobe, was appointed as eSafety Commissioner of Australia in 2017 and has since served three Australian Prime Ministers.[85] The eSafety Office was established in 2015, when the Parliament passed the Enhancing Online Safety for Children Act.[86] Its powers were expanded in 2022 when the Online Safety Act 2021 came into effect, covering Australia's regulatory response to a broad array of online harms, including adult cyber abuse, cyberbullying, image-based abuse, and illegal and restricted content.[87] On its website, the Office of the eSafety Commissioner writes: "eSafety's purpose is to help safeguard Australians at risk of online harms and to promote safer, more positive online experiences."[88]

Among the eSafety's regulatory mechanisms are transparency powers to ensure that online service providers — including social media platforms — are adhering to the country's "Basic Online Safety Expectations."[89] Under this scheme, the eSafety Commissioner can issue notices "requiring online service providers to report on their compliance with the Expectations," publish providers' responses to such requests, and issue fines to those who are found to not be in compliance.[90]

In 2023, Inman Grant issued two such transparency notices relating to child sexual abuse material and online hate to X.[91] In both cases, she found the platform to be in non-compliance, "providing responses that were incorrect, significantly incomplete or irrelevant," and in other cases "[failing] to provide any response to the question, such as by leaving the boxes entirely blank."[92] Inman Grant issued a non-compliance notice to the platform and fined it 610,500 AUD (about 412,000 USD).[93]

According to Inman Grant, the implementation of her office's regulatory powers angered Elon Musk, who had purchased the platform in 2022.[94] "We used our transparency power very effectively to highlight [X's] trust and safety issues," she said, and Musk had a pattern of using vexatious litigation against regulators and advocacy groups to "take on any entity that was critical."[95]

When a video depicting a stabbing of a bishop in Wakeley, Australia was uploaded to X and Meta-owned platforms in April 2024, Inman Grant issued a takedown notice to the platforms.[96] These powers are derived from Australian law, which prohibits content depicting "acts of terrorism,"[97] and allows her to subsequently request certain illegal content be removed from online platforms.[98] Meta complied with the request within the hour, Inman Grant said, but X kept the content up, despite the fact that it likely violated the platform's violent content policy.[99] When the Federal Court of Australia granted an interim injunction compelling X Corp to hide the violent material, Musk began tweeting about Inman Grant on April 22, 2024, calling her an "unelected official” and “eSafety Commissar,”[100] evoking authoritarian sentiments and claiming that Inman Grant “demand[ed] *global* content bans[.]”[101] These dog whistles—“the use of words or symbols with a double (or coded) meaning that is abusive or harmful, sometimes to signal a group of online abusers to attack a specific target”[102]—to his 192 million followers led to increased, targeted harassment against Inman Grant.

On April 23, 2024, there were 73,694 total mentions of Inman Grant or the eSafety Commissioner’s office on X.[103] By comparison, the office and Commissioner’s average daily mentions on X for April through December 2023 were 145.[104]

In order to better understand this harassment and its gendered aspects, the research team used Meltwater, a social listening and sentiment analysis tool, to download a dataset encompassing posts that mentioned “Julie Inman Grant” or Inman Grant’s X handle, “@tweetinjules,” from April 22-23. This subset of posts, which does not encompass all mentions of Inman Grant or the eSafety Commissioner’s office, but is a representative sample, comprised 1,054 posts. The research team then classified these posts using OpenAI’s large language model to analyze sentiment that uses a large language base—including abusive keywords themselves—to understand a variety of aspects of the message content. Additionally, the model uses intent and tone recognition to identify obfuscated harassment and non-keyword-based attacks.

The model initially assessed 70 percent of the posts (736) in the dataset as negative, 20 percent (218) as neutral, and 10 percent (100) as positive. To check the accuracy of the model, as well as add local context on which it likely was not trained, the research team then ran the content assessed as neutral or positive against a list of bespoke keywords encompassing common gendered slurs, abusive keywords, and keywords related to the falsehoods and conspiracies being directed at Inman Grant. This returned an additional 149 negative pieces of content, bringing the overall total of negative content in the dataset to 83 percent. Negative content generated twice as many likes as positive content. Among the negative content, gendered narratives and slurs made up more than 10 percent.

> “
>
> **The overall total of negative content in the dataset [was] 83 percent. Negative content generated twice as many likes as positive content. Among the negative content, gendered narratives and slurs made up more than 10 percent.**

Across the dataset, common gendered stereotypes contributing to the construction of a gendered enemy image of Inman Grant were widely observed, including attacks on Inman Grant’s physical appearance or adherence to beauty standards, claims she had negative characteristics often associated with women, such as aggression or a lack of intelligence. Gendered enemy images intersected with conspiracy theories to falsely claim Inman Grant was a part of a global censorship regime, rather than a domestic regulatory agency focused on online safety. Inman Grant was nicknamed “e-Karen,” a gendered pejorative that typically refers to middle-aged white women who, according to Kansas State University professor Heather Suzanne Woods, embody “entitlement, selfishness, [and] a desire to complain.”[105]

Further, some content claimed that Inman Grant was emotionally driven in her work, and that she was seeking revenge for an unnamed slight when she had worked at Twitter eight years prior to the Wakely stabbing; still others falsely claimed that Elon Musk—who did not own Twitter when Inman Grant worked there from 2014-2016—had fired her and she had an “ax to grind.” All of these narratives—that Inman Grant was allegedly entitled, selfish, power-hungry, emotional, or seeking revenge—are frequently deployed against women in public life.

Dehumanizing[106] language was also widely observed, including content practicing outcasting and the use of political labels and group comparison. Inman Grant was labeled "Big Mother" (a gendered take on Big Brother from George Orwell's *1984*), and "left-wing Barbie," "feminazi," "eNazi," or "Stasi cunt." One tweet read: "captain tampon is a nazi dictator." She was compared to a "terrorist" or "Hitler." Users also practiced dehumanization, labeling her a "humanoid lizard," and "pig." Additionally, they portrayed Inman Grant through a loss/threat framework, including: citing her alleged "harms" to children, families, or gender values.

Outside of the intersection of misogyny and conspiracy, sexist slurs were common, targeting Inman Grant's appearance, intellect, and gender identity. Inman Grant says she was called the "e-Slut of Australia."[107] Gendered enemy images were also achieved through sexualized dehumanization[108] tactics, such as labeling Inman Grant a "dominatrix," "eProstitute," "eSlut," or telling her to "get stuffed." The dataset also includes common slurs such as "bitch," "slag," and "cunt," as well as assertions that women are less intelligent than men or not fit for government roles. One user wrote: "Males to boot these brainwashed females out, and start governing."

The rhetorical tactics observed not only targeted Inman Grant but also worked to normalize enemy images of female public figures[109] and women more generally. Past research has widely documented the tangible impacts of enemy images, including: shifts in policy,[110] public opinion,[111] widespread prejudice (i.e., sexism and misogyny),[112] and motivations for individual and collective violence,[113] including sexual violence.[114] Social media's affordances have emboldened these effects, illustrating the need to re-evaluate the subsequent individual and societal harms to women, both online and offline.

Inman Grant's experience reflects the offline impact of enemy images. Users threatened her, her family, and her employees. OpenAI's sentiment analysis model assessed a full 10 percent of content in this dataset as threatening. For example, one user wrote: "@tweetinjules Vile, white-hating, racist pos. You are one fugly man. Thankfully you marxists will soon be wiped out." In this environment of hate, Inman Grant's family members were doxxed and users directed credible death threats at her, necessitating the involvement of the Australian Federal Police.[115]

"

Users threatened [Inman Grant], her family, and her employees. OpenAI's sentiment analysis model assessed a full 10 percent of content in this dataset as threatening.

In early June, eSafety discontinued its legal action on the Wakeley stabbing against X Corp in Australia's Federal Court to focus on other litigation, including matters involving X Corp. Inman Grant wrote:

> *I have decided to consolidate action concerning my Class 1 removal notice to X Corp in the Administrative Appeals Tribunal. After weighing multiple considerations, including litigation across multiple cases, I have considered this option likely to achieve the most positive outcome for the online safety of all Australians, especially children. As a result, I have decided to discontinue the proceedings in the Federal Court against X Corp in relation to the matter of extreme violent material depicting the real-life graphic stabbing of a religious leader at Wakeley in Sydney on 15 April, 2024.*[116]

Attacks against Inman Grant remained high even after Musk ceased actively mentioning her: her daily mentions averaged 2,585 daily between May and mid-June.[117] Her harassment — and the vexatious litigation and tying up of Australian Government resources — continues today. Currently, eSafety has five further cases against X, while users attempt to continue to — at their own admission — occupy eSafety resources. This has resulted in newly established local organizations starting a campaign to drown eSafety with Freedom of Information requests resulting in more than a 3,000 percent increase in such requests.[118] Sadly, the resources these campaigners are tying up are there to help Australians experiencing online abuse.

> **“**
>
> **[Inman Grant] noted that the lack of regulation, particularly in the United States, is hurting women around the world. “Until the U.S. actually regulates,” she said, “the rest of the world is going to be fighting a losing battle.”**

Inman Grant’s experience at the center of a harassment and abuse campaign instigated by a billionaire tech mogul demonstrates how pervasive TFGBV can be, targeting even those charged with making the internet safer. She told the research team: “There is now a growing awareness that the way online abuse manifests against women is different.”[119] However, she noted that the lack of regulation, particularly in the United States, is hurting women around the world. “Until the U.S. actually regulates,” she said, “the rest of the world is going to be fighting a losing battle.”[120]

THE LEGAL AND REGULATORY LANDSCAPE

Experts have repeatedly indicated that governments must pass laws and regulations in order to properly address the issue of TFGBV,[121] incentivizing platforms to enforce their policies and encouraging civility on and offline. However, legislating against harm and violence in online spaces has been challenging. For many years, the tangible and fundamental impact of social media and digital technologies has been understated, creating a lag for many legal and regulatory systems' understanding of these spaces and the online harms they foster. In the meantime, women and girls have had to make do with existing laws, such as copyright and tort laws, with limited success.[122] Beyond that, digital technologies are developing and transforming at a rapid pace, making it difficult to develop adequate legislation at the pace of technological advancements. Finally, there is the risk that online safety laws will be challenged by critics as undermining freedom of expression and promoting 'censorship.'[123]

Too little has been done around the globe to enact legislation that explicitly accounts for or aims to curb TFGBV—both for women in public life and for the general public alike. However, recent online safety laws that aim to curb general online harms may reduce the presence and impact of TFGBV through oversight and transparency over content moderation or the criminalization of online harms. Further, some of these laws and regulatory regimes have varying degrees of specific provisions for the challenges women face online, including deepfake image-based abuse and cyberflashing,[124] in large part due to the advocacy of civil society. Nonetheless, as we describe below, these laws and regimes have their own shortcomings and can be strengthened in various ways.

ONLINE SAFETY LAWS

A growing number of nations have enacted laws and established regulatory regimes aimed at mitigating online harms. Some, including the U.S., do not have an overarching national online safety law, even though the U.S. Senate recently passed two pieces of legislation aimed at making the internet safer for children and many states have existing or proposed online safety laws.[125] While this report cannot cover all of these laws and regimes in detail, several prominent examples that have varying degrees of specific provisions for the challenges women face online include:

Australia: As mentioned earlier in the report, the Online Safety Act 2021 came into effect in 2022,[126] building on the existing Expanding Online Safety for Children Act that established the Office of the eSafety Commissioner in 2015.[127] The Office of the eSafety Commissioner is empowered to hold platforms accountable, help users report claims of online harms, work with the private sector to embed safety by design, and generally set the regulatory standards for content moderation and online safety in Australia.[128]

Accordingly, the Office of the eSafety Commissioner has a range of 'schemes' detailing the specifications for illegal content that companies are obligated to closely monitor and take down, including cyberbullying,[129] image-based abuse,[130] and adult cyber abuse,[131] online content,[132] and abhorrent violent conduct.[133] The eSafety Commissioner can order a wide range of platforms and messaging services to take down content that is image-based abuse—or remove it directly—within 24 hours, and eSafetyWomen delivers direct and indirect support to women most at risk of online abuse, including evidence-based resources and professional development[134] and social media self-defense resources.[135]

The European Union: The European Union's Digital Services Act (DSA) entered into force in November 2022[136] and, as of February 2024, the DSA rules apply to all platforms.[137] Designed to enable a safer online environment throughout the EU,[138] the DSA sets out to combat disinformation, reduce and prohibit illegal and harmful content on platforms and digital marketplaces, protect rights online, and hold companies, especially "Very Large Online Platforms" (VLOPs) and "Very Large Online Search Engines" (VLOSEs), accountable for the content they host.[139] By mandating corporate accountability and transparency, reporting and oversight mechanisms, and penalties for non-compliance—as well as better support for users' rights by requiring flagging mechanisms and clearly defining illegal content—the DSA is a serious attempt to reduce online harms.[140]

Under the DSA, VLOPS and VLOSEs are required to conduct a risk assessment of their technologies and specifically consider the ways they enable "systemic risks that are linked to their services," including "gender-based violence" and "illegal content," and put measures in place to mitigate such risks.[141]

In May 2024, the EU adopted the "Directive on combating violence against women and domestic violence," the "first ever [EU] law to effectively fight violence against women and domestic violence" that "criminalizes at the EU level certain forms of violence against women offline (female genital mutilation and forced marriage) and online (non-consensual sharing of intimate images, cyberstalking, cyber harassment and incitement to hatred and violence on the ground of gender)."[142] Member states will have until June 14, 2027, to transpose the directive into national law.[143]

In its initial drafting, the Directive faced criticism for vaguely-defined terms such as "cyber harassment" and "cyber incitement to violence or hatred" that observers like the Center for Democracy and Technology (CDT) viewed as major drawbacks, arguing that terms did not "meet the principles of legality, proportionality, and necessity [and] risk being weaponized against the very individuals this Directive aims to protect."[144] The final version of the Directive attempts to address some of these shortcomings, including by setting minimum standards for the criminalization of TFGBV and criminalizing cyberflashing as part of cyber harassment; however, member states will ultimately need to build upon and clarify the Directive when they transpose it into their national laws.[145]

In addition to the DSA, in March 2024, the EU adopted the Artificial Intelligence Act (AI Act), "considered to be the world's first comprehensive horizontal legal framework for AI."[146] The AI Act requires "[P]roviders of AI systems generating synthetic audio, image, video or text content to ensure that the outputs of the AI system are marked in a machine-readable format and detectable as artificially generated or manipulated."[147] Through a tiered system categorizing AI systems by degree of risk, the AI Act imposes different regulatory requirements to ensure "that AI systems used in the EU are safe, transparent, traceable, non-discriminatory and environmentally friendly"[148] Content that is generated or manipulated with AI, such as deepfakes, must be labeled as content that has been generated or manipulated with AI "so that users are aware when they come across such content."[149]

The United Kingdom: In 2017, the UK government announced its plans to make the UK "the safest place in the world to be online."[150] The resulting bill, the Online Safety Act 2023 (OSA), came into force in late 2023 and creates the obligation to protect both children and adults from harmful content and embed better transparency and safety mechanisms into online platforms, in addition to criminalizing users who create harmful content or use online platforms in harmful ways.[151] The OSA creates three main pillars of legal duties: illegal harms, the protection of children, and obligations for larger platforms to report on illegal content and provide transparency around the use of users' data.[152] Ofcom is charged with implementing this legislation as the OSA's regulatory body.[153] Ofcom offers a wide range of support to victims of cybercrime, including making claims, and assistance in taking down illegal content, in addition to providing oversight to ensure companies are in compliance with the appropriate regulations.[154]

The OSA lists a large range of harmful content that companies must take down and report, with the most recent criminal offenses that came into effect on January 31, 2024, and include "encouraging or assisting serious self-harm, cyberflashing, sending false information intended to cause non-trivial harm, threatening communications, intimate image abuse, [and] epilepsy trolling."[155] Much of these provisions were included after many rounds of consultations,[156] advocacy on behalf of women who have been targeted by online abuse, and impactful work by experts and scholars,[157] ensuring that the OSA protected not just children, but that women over 18 and girls had tailored and responsive policies that supported them against the unique online harms they face.[158]

Global Online Safety Regulators Network
In response to the cross-border nature of online harms, a group of national regulators with legislated online safety powers established the Global Online Safety Regulators Network (GOSRN), "the first dedicated forum for independent online safety regulators around the world, in November 2022."[159] Comprised of 9 members and several observers,[160] including the United Kingdom's Ofcom, Australia's eSafety Commissioner, France's Arcom, Ireland's Coimisiún na Meán, and South Korea's Communications Standards Commission, GOSRN's purpose is to share "information, best practice, expertise and experience, to support coherent and coordinated approaches to online safety issues."[161] In its April 2024 Position Paper, where GOSRN mapped the similarities and differences between its regulatory regimes, GOSRN found several areas for collaboration and coherence between regimes.[162] These areas include regulatory tools like risk assessment and transparency reporting, user complaint functions, researcher access to data mechanisms, and information requests to industry.[163]

LEGISLATING RESPONSES TO DEEPFAKE IMAGE-BASED SEXUAL ABUSE

With the advancement of generative AI and, in turn, the rise of deepfake pornography, a growing number of countries have enacted laws or proposed legislation to provide victims and survivors of deepfake image-based sexual abuse with either criminal or civil remedies — or both. While much of the world has yet to enact such laws or even propose such legislation, several countries are leading the way in addressing deepfake pornography and should serve as models to other nations grappling with this growing problem:

Australia: In February 2024, the Office of the eSafety Commissioner released regulatory guidance for the Image-Based Abuse Scheme.[164] The Office of the eSafety Commissioner has the regulatory powers "to take action against a person (end-user) who shares online (or threatens to share) an intimate image [including a deepfake] without the consent of the person shown."[165] The guidance specifies that an individual can make a complaint to eSafety about the image-based abuse, and eSafety can then require online service providers and end-users to remove the intimate image through a notice and — if they don't comply — through a formal warning or civil penalties in court.[166]

Advocates have pointed out that although this civil framework is helpful, individuals who create non-consensual, intimate deepfakes are still within their rights under current law in Australia.[167] This loophole is precisely the target of new legislation introduced to Parliament in June 2024, which would criminalize the creation and non-consensual sharing of deepfake intimate images.[168] In addition to the current fines, violators under this law would receive up to six years of imprisonment for sharing images and face an additional year of imprisonment if they also created it.[169] The legislation also proposes additional funding to the Office of the eSafety Commissioner to pilot age verification tools online, increased resourcing and training for local law enforcement, and better information sharing across jurisdictions about offenders.[170]

The European Union: Under the DSA, VLOPs and VLOSEs are required to conduct a risk assessment of their technologies and specifically consider the ways they enable "systemic risks that are linked to their services," including "gender-based violence" and "illegal content," and put measures in place to mitigate such risks.[171] The DSA provides that the "unlawful non-consensual sharing of private images" and the "sharing [of] images depicting child sexual abuse" is illegal content.[172] In December 2023, the European Commission added three pornography websites — Pornhub, Stripchat, and XVideos — to the list of VLOPs.[173] [174]

In May 2024, the EU adopted the "Directive on combating violence against women and domestic violence," which explicitly criminalizes non-consensual sharing of intimate images, including deepfakes, cyberstalking, harassment, and incitement to hatred or violence, all measures that member states will have to adopt and enforce.[175] Offenders of non-consensual sharing of intimate images are "punishable by a maximum penalty of at least 1 year of imprisonment" and victims will have the right to claim full compensation from offenders for costs and damages.[176] While the Directive importantly does not require proof of specific motives, such as causing distress, it does have some shortcomings, including specific language requiring "engaging in sexual activity," which will likely leave out images produced through 'nudify apps' and shared non-consensually.[177]

South Korea: South Korea was one of the first countries to directly deal with deepfake pornography when K-Pop stars started becoming targets.[178] This trend emerged and gained notoriety when 74 women and girls were targeted by an incident of widespread sexual assault and deepfake pornography created in male messaging groups on Telegram.[179] The groups were dedicated to editing and requesting edits of violent and explicit deepfakes of women and girls in their lives.[180] This situation, called the "Nth Room Case,"[181] led to a major overhaul of the penal code and the conviction of 220 offenders for sex crimes.[182] [183]

Recognizing the deeply troubling issues at play, in 2020, the South Korean government enacted a revision to the "Act on Special Cases Concerning the Punishment of Sex Crimes"[184] to prohibit "the creation and distribution of 'false video products' which 'may cause sexual desire or shame against the will of the person who is subject to video.'"[185] Under this framework, the creation and distribution of non-consensual deepfake pornography is outlawed, regardless of intent.[186] Offenders may be convicted up to five years in prison and — if they sold access to the non-consensual pornographic deepfake for a profit — they could face up to twelve years in prison.[187]

Like the other laws and regimes, the South Korea law does have some shortcomings. Advocates call for punishments for those who purchase or download this content as well, citing the dramatic rise in the proliferation of non-consensual deepfake pornography — especially in Telegram chatrooms where people distribute, create, and collaborate over the creation of explicit edits of acquaintances, even disclosing sensitive personal information.[188] Furthermore, South Korea, like other countries, is finding that taking down content is a very slow process (it only has a 2.3 percent success rate) and a lack of capacity and training within policing is still a barrier.[189] Nonetheless, the Korean Police Agency has developed and deployed new AI detection tools to help combat deepfake pornography's proliferation.[190]

The United Kingdom: In January 2024, an amendment to the OSA, which criminalizes the act of sharing or threatening to share intimate images, including deepfakes, came into effect.[191] The amendment was the result multiple calls for evidence and campaigning from women's rights groups, including Glitch, the End Violence Against Women Coalition, Refuge, Carnegie, the National Society for the Prevention of Cruelty to Children, 5Rights and experts Clare McGlynn and Lorna Woods, to more specifically target intimate image abuse in the OSA.[192] [193]

Offenders who share non-consensual intimate deepfakes can face up to two years of prison time and unlimited fines.[194] Importantly, this amendment does not require proof of intent to humiliate, cause distress, or alarm, or that the content be shared for personal sexual gratification: the basic act of sharing intimate images without consent is sufficient to count as an offense, relieving victims and survivors of the burden to present evidence of intent.[195] Additionally, as this is a sexual offense, victims are automatically protected with anonymity in any court cases or media reporting.[196]

As progressive as the amendment is, the OSA still technically allows for the creation of NCII.[197] Experts like Professor Clare McGlynn note that without criminalization of the entire ecosystem, sexualized deepfakes will continue to grow, "propped up" and enabled by payment processors, search engines, and users.[198]

In an effort to address some of the shortcomings raised in public comments, the UK Ministry of Justice announced in April 2024 that an amendment to the Criminal Justice Bill will be introduced and will criminalize the non-consensual creation of sexually explicit deepfakes.[199] If passed, the act of creating deepfake pornography nonconsensually would result in the offender having a criminal record and an unlimited fine and facing jail time if they share the image widely.[200] As drafted, however, the proposed amendment has been met with criticism because it focuses on the intent of the perpetrator, requiring proof of intent to cause "alarm, distress, or humiliation."[201] Nonetheless, since the announcement of this law, some of the largest deepfake pornography sites have been blocked in the UK, perhaps in reaction or anticipation of this legal development, demonstrating the powerful signal these sorts of laws can have in society.[202]

Beyond the language of the law itself, experts and advocates also raise the observation that the police do not have the proper training or knowledge to assist with this sort of crime, and call for increased education and resourcing in communities.[203] McGlynn also points out that this legislation does not sufficiently hold companies accountable, arguing that search engines need to be modified so that deepfakes are not so highly ranked in search results and that payment processing companies must be held responsible for enabling the proliferation of this type of content.[204] These hurdles remain major challenges to the enforcement of the current legal framework.

The United States: While forty-nine states have laws against NCII or “revenge porn,”[205] and the Violence Against Women Act Reauthorization Act of 2022 (VAWA), amongst other provisions, established a federal civil cause of action for individuals whose “intimate visual depictions” are disclosed without their consent,[206] the U.S. does not yet have a federal law that explicitly regulates or criminalizes deepfake pornography.

However, in July 2024, the U.S. Senate unanimously passed the bipartisan Disrupt Explicit Forged Images and Non-Consensual Edits Act of 2024 (DEFIANCE ACT),[207] introduced by Representative Alexandria Ocasio-Cortez (D-NY) in the House and Senator Dick Durbin (D-IL) and Senator Lindsey Graham (R-SC) in the Senate.[208] If passed by the House of Representatives and signed into law by the President, this legislation would amend VAWA to create a federal civil remedy for “an identifiable individual who is the subject of a digital forgery,” or an “intimate visual depiction of an identifiable individual created through the use of software, machine learning, artificial intelligence, or any other computer-generated or technological means.”[209] Additionally, “[the] identifiable individual may recover the actual damages sustained by the individual or liquidated damages [in the amount of $150,000] and the cost of the action, including reasonable attorney’s fees and other litigation costs reasonably incurred.”[210]

In addition, in June 2024, Senator Ted Cruz (R-TX) introduced the Tools to Address Known Exploitation by Immobilizing Technological Deepfakes on Websites and Networks (TAKE IT DOWN) Act, alongside a bipartisan group of cosponsors, including Senator Amy Klobuchar (D-MN), Senator Richard Blumenthal (D-CT), Senator Bill Cassidy (R-LA), and others.[211] The TAKE IT DOWN Act would make it unlawful for a person to knowingly publish or threaten to publish NCII, including deepfakes, on social media and other online platforms.[212] It would provide criminal penalties for such offenses and require social media and other websites to remove the NCII, pursuant to a request from the identifiable individual, within 48 hours.[213]

In the absence of a federal law regulating deepfake pornography, “29 states have enacted laws addressing sexual deepfakes, and of those laws, 21 restrict the distribution of nonconsenual sexual deepfakes.”[214] Notable examples include Virginia, which prohibits the distribution of deepfake pornography and makes this offense punishable with jail time,[215] and California, which prohibits the distribution and creation of deepfake pornography.[216] [217] Both the Virginia and California laws, however, face drawbacks: Virginia requires victims to prove malice and does not criminalize the creation of non-consensual intimate deepfakes, and the California law does not punish these crimes with jail time.[218]

RECOMMENDATIONS

While the challenge that TFGBV poses to women's and girls' equality and free expression is immense, there are a number of policy solutions that, if implemented, could help prevent and address TFGBV. Technology companies, governments, civic tech organizations, law enforcement, employers, schools, and others must work together to combat TFGBV. To this end, we recommend a number of practical solutions to the specific and pressing issues that women and girls face online today.

ENSURING PLATFORM ACCOUNTABILITY AND ACTION

Most very large social media platforms and other technology companies have a spotty record, at best, of enforcing their own terms of service in this area. As such, the current will to make important changes in favor of women's online safety — in effect, to self-regulate — is lacking. It is critical, therefore, to enact broad-based legislative frameworks to protect the safety of marginalized communities, including women. As more governments explore regulation of social media and technology firms — and some firms seek to implement progressive, inclusive trust and safety policies — the following reforms would greatly improve the online landscape for women;

Governments should explore oversight that encourages platforms to exercise their duty of care to ensure women are able to safely express themselves online. Given the widespread rollback of investments in trust and safety by platforms and the demonization of content moderation policies as 'censorship,' governments should explore such oversight. These might, as in Australia's eSafety mechanisms, include transparency powers that allow regulators to understand platforms' content moderation processes and decision-making on issues related to protected groups' speech. Platforms should consider radically transparent self-reporting on these issues as well, inviting the scrutiny and oversight from the public, legislatures, and affected communities that leads to better policy.

Legislators should consider creating a transparency and oversight mechanism to ensure researchers and journalists have access to social media data — including data related to online abuse — for public service and research purposes, while protecting the privacy of platform users. The DSA mandates VLOPs and VLOSEs to provide data access to vetted researchers for the purpose of scrutinizing systemic risks and mitigating harmful content. Other countries should consider a similar provision in their own regulation, which would provide important accountability and oversight over platforms. Technology platforms could also consider providing this data access ahead of regulatory action.

Online safety laws and regulations must include explicit provisions that address the harms women face online. Rather than having to use copyright or tort laws to seek remedies for online harms, women and other victims of TFGBV should be able to seek remedies via legal infrastructure that specifically addresses challenges like cyberflashing, cyberstalking, deepfake pornography and others. To achieve this, it is critical that legislators and regulators regularly consult with civil society in the development of such laws. GOSRN could also consider holding public meetings on TFGBV where civil society groups and survivors could share experiences and recommendations for the benefit of the network and other bodies watching the proceedings.

Technology companies must address gender imbalances in tech. At the largest tech companies — Google, Apple, Meta, Amazon, and Microsoft — "on average only 31 percent of…[U.S.] employees are female."[219] This gender imbalance, combined with a lack of understanding among male employees of the harms women face on their platforms, can lead to issues in addressing and regulating TFGBV occurring on these platforms. Platforms should consider adopting scholar Sasha Costanza-Chock's "Design Justice" framework, a "growing social movement that aims to ensure a more equitable distribution of design's benefits and burdens; fair and meaningful participation in design decisions; and recognition of community-based design traditions, knowledge, and practices."[220]

URGENTLY ADDRESSING DEEPFAKE IMAGE-BASED SEXUAL ABUSE

As documented throughout this report and by the drumbeat of stories about deepfake pornography in international media, the challenge of deepfake pornography is affecting more women and girls by the week. Governments and technology companies should make the following reforms:

Legislatures should pass federal or nationwide bills that may include both civil and criminal penalties for both the creation and distribution of non-consensual deepfake pornography. They should also ensure that such laws do not require victims to prove intent to distress, humiliate, or otherwise harm the victim or malice to the victim, and might also consider the introduction of excess penalties for the creation or distribution of deepfake pornography against political candidates during elections. While the introduction of such legal frameworks will not end the phenomenon entirely, they do have a cascading effect that allow technology platforms to take more muscular action against accounts that distribute or amplify such content and app developers who create applications that create it. As more similar laws come into force around the world, those who profit and generate enjoyment from this most base violation of women's privacy and bodily autonomy will be driven out of business.

Search engines and app stores should delist and demonetize websites and applications that are dedicated to the creation and distribution of deepfake pornography and ensure underage users are not able to access apps that 'nudify' individuals. Additionally, payment processors should refuse to process transactions for organizations that provide such services. This is a relatively trivial act for very large online platforms to implement, and any technology company that claims to care about women's free and equal participation in public life should move to do so immediately. Google recently made changes prohibiting users from advertising services that enable the creation of deepfake or nudified images and videos, demonstrating the ease of implementation for companies.[221] Technology companies must invest in the necessary resources to ensure that these changes are systematically enforced.

Technology companies should prioritize making the technologies that challenge deepfakes, such as immunizing or poisoning an image,[222] more accessible and make it easier to detect deepfakes, with methods such as watermarking. The practice of immunizing images uses pixels invisible to the human eye to ensure that AI models classify them incorrectly, whereas poisoning can teach an AI model that an image is something it is not, disrupting models in the long term. Watermarking can help people and systems differentiate what content has been AI generated or altered.

Governments should support public awareness campaigns and educational resources that highlight the pervasiveness of deepfake pornography, the adverse mental health and employment effects of deepfake pornography on victims and survivors, and how victims and survivors can get justice and support.

SUPPORTING VICTIMS AND SURVIVORS OF TFGBV

Around the world, there is currently little support for victims and survivors of TFGBV, leading to a drop-off in reporting and self-silencing among this population. Governments, employers, and schools should make the following reforms:

Governments should invest in training for and community building with law enforcement entities. Victims and survivors of TFGBV often do not feel the law enforcement bodies and representatives with whom they interface to be particularly helpful or understanding. Governments at the national and local level must improve awareness and training on the use of existing legal infrastructure to aid victims of online abuse, and consider holding community events to improve trust between TFGBV survivors and the law enforcement community. Funding should be made available to entice more investigators to units that specialize in these types of crimes.

Governments should implement and properly resource victim hotlines and helplines to ensure that victims of TFGBV receive assistance and counseling in a robust and timely fashion. These 24/7 resources should be well coordinated with the relevant law enforcement and social service authorities.[223]

Employers and schools should have policies and support for employees and students who are victims and survivors of TFGBV. In today's highly polarized political environment, in which the internet acts as an accelerant to abuse, employers and schools must have proactive strategies in place to support employees who may become targets of TFGBV. These may include doxxing support, mental healthcare, assistance navigating law enforcement, and other emotional, physical, and psychosocial provisions to ensure women's online freedom of expression is maintained.[224] Policies should delineate a clear escalation framework so employees and students know how and when in the cycle of online abuse to request support. This framework allows employees and students to feel supported as they express themselves online, and underscores that their online self-expression should not cause them or their career harm in the future.

Governments should enact laws and protections for victims and survivors of TFGBV that would enable them to take the time they need off of work to meet with their attorney, attend legal proceedings, or obtain psychological counseling.

DEEPENING RESEARCH AND MAINSTREAMING ADVOCACY

Finally, conducting deeper research and using it to inform policies along with mainstreaming advocacy is essential to preventing and addressing TFGBV. Researchers and governments should make the following reforms:

Deepening Research Inquiries. Studies have laid out the existence and scope of TFGBV, providing essential overviews of the phenomena. However, many reports focus on top-line insights, failing to examine diverse cases in depth or the many underlying contributing factors to TFGBV. Researchers should prioritize research that explores a richer analysis of how and why TFGBV occurs on a societal level, including factors such as communication structures, historical context, socio-political dynamics, intersectional dynamics, and distinctive empirical contexts.

Researchers should systematize and streamline research. As recommended in a Global Partnership report, studies should attempt to produce unified and centrally agreed upon frameworks which can be used to guide future work in the field and provide a consensus around key indicators that may be helpful in preventing and proactively intervening in cases of TFGBV.[225] Research designs—specifically methodology and data collection approaches—also vary greatly, resulting in difficulties striking a balance between quantitative and qualitative methods and not frequently coordinated across sectors. Future research would greatly benefit from multi-stakeholder, cross-disciplinary efforts.[226]

Mainstream advocacy, cultural communication, and awareness building. Governments should aim to empower civil society to raise public awareness of TFGBV through strategic communications campaigns and cultural engagement activities that educate citizens about the realities of being a woman online. These programs can reverse the normalization of TFGBV. Campaigns should also aim to reach younger audiences by engaging where those audiences are: on social media. Crucially, these campaigns should not only target women, but focus on building male allies across societies.

CONCLUSION

For too long, the abuse and harassment that women and girls have faced on the internet has been a feature of their online experiences. In nearly every new technology that is developed, users identify a way to use it to demean, degrade, or demonize women and girls. We envision a world that actively addresses this inherent inequality in our society, rather than enabling it. This cannot happen if women's advocacy groups are working in isolation. Technology companies, governments, civic tech organizations, law enforcement, employers, schools, and others must also work together to prevent and address TFGBV. In a year in which half of the world's population has an opportunity to vote, and in an era in which key information informing citizens' everyday lives is increasingly distributed and consumed on the internet, TFGBV affects women's full and equal participation in society. Until the online harms that disproportionately target women and girls are addressed, modern democracy has much work left to do—and that work is everyone's problem.

ENDNOTES

[1] The research team is deliberately excluding citations of direct links to abusive content throughout this report so as not to amplify the users who posted it.

[2] Nina Jankowicz et al., "Malign Creativity: How Gender, Sex, and Lies are Weaponized Against Women Online," The Wilson Center, January 2021, https://www.wilsoncenter.org/publication/malign-creativity-how-gender-sex-and-lies-are-weaponized-against-women-online.

[3] Just three months before the 2024 U.S. Presidential election, Elon Musk, tech billionaire and endorser of Donald Trump, shared a convincing political deepfake on X depicting Kamala Harris saying falsehoods about the election and claiming that Harris is a "diversity hire" for being a woman and person of color. See: Ali Swenson, "A parody ad shared by Elon Musk clones Kamala Harris' voice, raising concerns about AI in politics," AP News, July 29, 2024, https://apnews.com/article/parody-ad-ai-harris-musk-x-misleading-3a5df582f911a808d34f68b766aa3b8e.

[4] Julia Thalin, "Gender equality: driver and objective to electoral integrity," International IDEA, February 29, 2024, https://www.idea.int/news/gender-equality-driver-and-objective-electoral-integrity.

[5] Experiences of intimidation are common among those in office: 43 percent of state legislators experienced threats during their last term and most recent campaign, and 18 percent of local officeholders faced threats within the prior three months. This abuse impedes the normal functioning of representative government, with more than 40 percent of local officeholders reporting that the hostility had chilled their desire to work on controversial topics or run for re-election or higher office. See: Grady Youthok Short and Maya Kornberg, "Hostility and Abuse Threaten to Undermine Gains in Representative Democracy," Brennan Center for Justice, February 27, 2024, https://www.brennancenter.org/our-work/analysis-opinion/hostility-and-abuse-threaten-undermine-gains-representative-democracy.

[6] Agence France-Presse, "Slovakian president Čaputová says she will not run for re-election," The Guardian, June 20, 2023, https://www.theguardian.com/world/2023/jun/20/slovakia-president-zuzana-caputova-says-she-will-not-run-for-re-election.

[7] Sharon Goulds, Miriam Gauer, Aisling Corr and Jacqui Gallinetti, "State of the World's Girls 2020: Free to Be Online?" Plan International, 2020, https://plan-international.org/publications/free-to-be-online/.

[8] Sharon Goulds, Miriam Gauer, Aisling Corr and Jacqui Gallinetti, "State of the World's Girls 2020: Free to Be Online?" Plan International, 2020.

[9] Jennifer Gerson, "The complicated ties between teenage girls and social media – and what parents should know," The 19th News, September 13, 2023, https://19thnews.org/2023/09/social-media-teenage-girls-mental-health-body-image/.

[10] Anjali Popat and Carolyn Tarrant, "Exploring Adolescents' Perspectives on Social Media and Mental Health and Well-Being: A Qualitative Literature Review," Clinical Child Psychology and Psychiatry 28, no. 1 (2023): 323–37. https://doi.org/10.1177/13591045221092884.

[11] "Twitter's decision to disband safety council threatens wellbeing of users," Amnesty International, December 13, 2022, https://www.amnesty.org/en/latest/news/2022/12/global-twitters-decision-to-disband-safety-council-threatens-wellbeing-of-users.

[12] Taylor Hatmaker, "Elon Musk just axed key Twitter teams like human rights, accessibility, AI ethics and curation," TechCrunch, November 4, 2022, https://techcrunch.com/2022/11/04/elon-musk-twitter-layoffs.

[13] Sara Scire, "A window into Facebook closes as Meta sets a date to shut down CrowdTangle," Nieman Lab, March 14, 2024, https://www.niemanlab.org/2024/03/a-window-into-facebook-closes-as-meta-sets-a-date-to-shut-down-crowdtangle/.

[14] Brandi Geurkink and Sarah Gilbert, "Why Reddit's decision to cut off researchers is bad for its business — and humanity," Fast Company, January 22, 2024, https://www.fastcompany.com/91014116/reddit-researchers-bad-for-business.

[15] Justine Calma, "Twitter just closed the book on academic research," The Verge, May 31, 2023, https://www.theverge.com/2023/5/31/23739084/twitter-elon-musk-api-policy-chilling-academic-research.

16 Sheila Dang, “Exclusive: Elon Musk’s X restructuring curtails disinformation research, spurs legal fears,” Reuters, November 6, 2023, https://www.reuters.com/technology/elon-musks-x-restructuring-curtails-disinformation-research-spurs-legal-fears-2023-11-06/ .

17 “Block Party’s Twitter product is on indefinite hiatus as of May 31,” Block Party, May 30, 2023, https://www.blockpartyapp.com/blog/twitter-hiatus/. See also: Jess Weatherbed, “Anti-harassment service block party leaves Twitter amid API changes,” The Verge, May 31, 2023, https://www.theverge.com/2023/5/31/23743538/block-party-hiatus-twitter-app-anti-harassment-service-api. However, Block Party and its counterpart “PrivacyParty” continue to exist on different platforms.

18 Jim Rutenberg and Steven Lee Meyers, “How Trump’s Allies Are Winning the War Over Disinformation,” The New York Times, March 17, 2024, https://www.nytimes.com/2024/03/17/us/politics/trump-disinformation-2024-social-media.html.

19 This report uses the terms “deepfake pornography” and “non-consensual intimate deepfakes” interchangeably to refer to this phenomenon.

20 “2023 State Of Deepfakes: Realities, Threats, And Impact,” Security Hero, accessed July 22, 2024, https://www.homesecurityheroes.com/state-of-deepfakes/.

21 A deep dive into deepfakes that demean, defraud and disinform. See: “A deep dive into deepfakes that demean, defraud and disinform,” Ofcom, July 23, 2024, https://www.ofcom.org.uk/online-safety/illegal-and-harmful-content/deepfakes-demean-defraud-disinform/.

22 “Technology-facilitated Violence against Women: Towards a common definition. Report of the meeting of the Expert Group” UN Women and World Health Organization, accessed July 22, 2024, https://www.unwomen.org/sites/default/files/2023-03/Expert-Group-Meeting-report-Technology-facilitated-violence-against-women-en.pdf.

23 “What is NCII?” InHope, February 17, 2023, accessed August 5, 2024, https://inhope.org/EN/articles/what-is-ncii.

24 The British charity Anti-bullying Pro, a non-profit dedicated to stopping bullying and harassment, defines catfishing as “creating fake profiles or providing fake pictures or fake information on social media networking sites, apps and online… often used by online predators to lure or trick people into starting online relationships.” See: “Cyberbullying Behavior and Catfishing: How can we recognize if someone is fake online?” Anti-Bullying Pro, accessed July 2024, https://www.antibullyingpro.com/support-and-advice-articles/cyberbullying-behaviour-and-catfishing-how-can-we-recognise-if-someone-is-fake-online.

25 In its “Online Harassment Field Manual,” PEN America defines dogpiling or cyber-mob attacks as “When a large group of abusers collectively attacks a target through a barrage of threats, slurs, insults, and other abusive tactics.” See: Online Harassment Field Manual, “Defining “Online Abuse”: A Glossary of Terms,” PEN America, accessed July 2024, https://onlineharassmentfieldmanual.pen.org/defining-online-harassment-a-glossary-of-terms/.

26 Cyber surveillance is the “misuse [of] connected devices to monitor, harass, isolate and otherwise harm” victims and can include actions such as controlling personal devices, monitoring logs of activity such as through the use of GPS or home surveillance devices, hacking into personal accounts, eavesdropping, and controlling home devices connected to the internet like stereo systems. See WomensLaw.org’s explainer on this form of abuse: “About Abuse: Abuse Using Technology,” WomensLaw.org, accessed July 2024, https://www.womenslaw.org/about-abuse/abuse-using-technology/all.

27 In its “Misogynoir Report: Ending the dehumanising of Black women on social media,” the British non-profit Glitch defines digital misogynoir as “ the continued, unchecked, and often violent dehumanisation of Black women on social media, as well as through other forms such as algorithmic discrimination.” See: Glitch, Gabriela de Oliveira, and Dr. Julia Slupska, “The Digital Misogynoir Report: Ending the dehumanizing of Black women on social media,” Glitch, 2023, accessed July 18, 2024, https://glitchcharity.co.uk/wp-content/uploads/2023/07/Glitch-Misogynoir-Report_Final_18Jul_v5_Single-Pages.pdf.

28 Veronica Ahlenback, Erika Fraser, Kavita Kalsi and Maria Vlahakis, “Technology-Facilitated Gender-Based Violence: Preliminary Landscape Analysis,” The Global Partnership, July 10, 2023, accessed July 22, 2024, https://assets.publishing.service.gov.uk/media/64abe2b21121040013ee6576/Technology_facilitated_gender_based_violence_preliminary_landscape_analysis.pdf.

29 20 percent of women journalists in a UNESCO survey said they had been attacked or abused offline in connection with online violence they had experienced. Julie Posetti, Nermine Aboulez, Kalina Bontcheva, Jackie Harrison, and Silvio Waisbord, “Online violence against women journalists: a global snapshot of incidence and impacts,” UNESCO Digital Library, 2020, accessed July 30, 2024, https://unesdoc.unesco.org/ark:/48223/pf0000375136. See also: “Decoding technology-facilitated gender-based violence: a reality check from seven countries,” Rutgers International, 2024, accessed July 30, 2024, https://rutgers.international/resources/decoding-technology-facilitated-gender-based-violence-a-reality-check-from-seven-countries/. See also: “Hot Yoga Tallahassee,” Secret Service, accessed July 30, 2024, https://www.secretservice.gov/sites/default/files/reports/2022-03/NTAC%20Case%20Study%20-%20Hot%20Yoga%20Tallahassee_0.pdf. A case study report written by the Secret Service noted the growing problem of extremist misogyny and documented the worrying ties between violent attackers, such as the Hot Yoga Tallahassee attacker who killed two women and injured four more at a yoga studio, and misogynistic communications on online forums, often but not exclusively within the ‘manosphere.’

30 “Technology-Facilitated Gender-Based Violence as an Attack on Women’s Public Participation: Review of Global Evidence and Implications,” IREX, August 2023, https://www.irex.org/sites/default/files/Technology-Facilitated%20Gender-Based%20Violence%20as%20an%20Attack%20on%20Women%E2%80%99s%20Public%20Participation_Review%20of%20Global%20Evidence%20and%20Implications%20.pdf.

31 Ahlenback et al, “Technology.”

32 Sara Bundtzen, “Misogynistic pathways to radicalisation: Recommended Measures for Platforms to Assess and Mitigate Online Gender-Based Violence,” Institute for Strategic Dialogue, 2023, https://www.isdglobal.org/wp-content/uploads/2023/09/Misogynistic-Pathways-to-Radicalisation-Recommended-Measures-for-Platforms-to-Assess-and-Mitigate-Online-Gender-Based-Violence.pdf.

33 Ahlenback et al, “Technology.”

34 Ahlenback et al, “Technology.”

35 Ahlenback et al, “Technology.”

36 “DRG Learning Digest - Combatting Technology-Facilitated Gender-Based Violence in Politics,” U.S. Agency for International Development, March 7, 2023, https://content.govdelivery.com/accounts/USAIDHQ/bulletins/34c7e57.

37 Ahlenback et al, “Technology.”

38 “Preventing Technology-Facilitated Gender-Based Violence (TF GBV),” United Nations Population Fund, accessed July 2024, https://www.un.org/techenvoy/sites/www.un.org.techenvoy/files/GDC-Submission_UNFPA.pdf.

39 Katherine Noel, “Journalist Emanuel Maiberg Addresses AI and the Rise of Deepfake Pornography,” Institute of Global Politics, April 22, 2024, accessed July 22, 2024, https://igp.sipa.columbia.edu/news/rise-deepfake-pornography. Samantha Cole, “We are truly fucked: Everyone is making ai-generated fake porn now,” VICE, January 24, 2018, https://www.vice.com/en/article/bjye8a/reddit-fake-porn-app-daisy-ridley.

40 Noel, “Journalist,” and Cole, “We are truly fucked.”

41 “2023 State Of Deepfakes: Realities, Threats, And Impact,” Security Hero.

42 Ahlenback et al, “Technology.”

43 “White House Task Force To Address Online Harassment and Abuse: Final Report and Blueprint,” The White House, May 2024, accessed July 15, 2024, https://www.whitehouse.gov/wp-content/uploads/2024/05/White-House-Task-Force-to-Address-Online-Harassment-and-Abuse_FINAL.pdf.

44 “Safety of Women Journalists” UNESCO, accessed July 30, 2024, https://www.unesco.org/en/safety-journalists/safety-women-journalists.

45 Rumman Chowdhury and Dhanya Lakshmi, “Your opinion doesn’t matter, anyway: exposing technology-facilitated gender-based violence in an era of generative AI,” UNESCO, 2023, https://unesdoc.unesco.org/ark:/48223/pf0000387483.

46 “Technology-Facilitated Gender-Based Violence: A Growing Threat,” United Nations Population Fund, 2023, https://www.unfpa.org/TFGBV.

47 Nina Jankowicz et al., “Malign Creativity: How Gender, Sex, and Lies are Weaponized Against Women Online,” The Wilson Center, January 2021, https://www.wilsoncenter.org/publication/malign-creativity-how-gender-sex-and-lies-are-weaponized-against-women-online.

48 “Launch of Women LEAD: Women Leading Effective and Accountable Democracy in the Digital Age,” Carnegie Endowment for International Peace, July 8, 2024, accessed July 29, 2024, https://carnegieendowment.org/events/2024/07/launch-of-women-lead-women-leading-effective-and-accountable-democracy-in-the-digital-age.

49 Jacqueline Hicks, “Global evidence on the prevalence and impact of OGBV,” Institute of Development Studies, December 7, 2021, https://www.ids.ac.uk/publications/global-evidence-on-the-prevalence-and-impact-of-online-gender-based-violence-ogbv/. See also: “Crowdsourced Twitter study reveals shocking scale of online abuse against women,” Amnesty International, December 18, 2018, accessed July 29, 2024, https://www.amnesty.org/en/latest/press-release/2018/12/crowdsourced-twitter-study-reveals-shocking-scale-of-online-abuse-against-women/. A 2018 Twitter study by Amnesty International and Element AI revealed that women of color were 34 percent more likely to be mentioned in abusive or problematic tweets with Black women, in particular, 84 percent more likely to face abuse of this nature. See also: “Protecting LGBTIQ+ voices online: resource development research,” eSafety Commissioner, https://www.esafety.gov.au/sites/default/files/2021-08/LGBTQI%2B%20cyber%20abuse%20resource%20development%20-%20Report.pdf. In 2019, the Australian eSafety Commissioner found that 30 percent of LGBTQI+ individuals experienced online hate speech, more than double the national average of 14 percent.

50 “Measuring the prevalence of online violence against women,” The Economist, March 1, 2021, accessed July 22, 2024, https://onlineviolencewomen.eiu.com/.

51 “Measuring,” The Economist.

52 “Measuring,” The Economist.

53 National Democratic Institute “Tweets That Chill: Analyzing Online Violence Against Women in Politics,” National Democratic Institute, July 14, 2019, https://www.ndi.org/tweets-that-chill.

54 Stacy Dixon, “Effects of harmful online contact worldwide 2022, by gender,” Statista, April 12, 2022, https://www.statista.com/statistics/1301606/effects-of-harmful-online-contact/.

55 Dixon, “Effects.”

56 Samantha Bates, “Revenge Porn and Mental Health: A Qualitative Analysis of the Mental Health Effects of Revenge Porn on Female Survivors,” Feminist Criminology, 12(1), 22-42, (2017), https://doi-org.ezproxy.cul.columbia.edu/10.1177/1557085116654565.

57 “Measuring,” The Economist.

58 Suzie Dunn, “Addressing Gaps and Limitations in Legal Frameworks and Law Enforcement on Technology-Facilitated Gender-based Violence,” UN Women, October 2022, accessed July 29, 2024, https://www.unwomen.org/sites/default/files/2022-12/EP.15_Suzie%20Dunn.pdf.

59 Tatum Hunter, “AI porn is easy to make now. For women, that’s a nightmare,” Washington Post, February 13, 2023, accessed July 29, 2024, https://www.washingtonpost.com/technology/2023/02/13/ai-porn-deepfakes-women-consent/.

60 Madeleine Songy, Madeleine Pelton, Olivia Cretella, Areala Mendoza, Kathryn Hopp, Olivia Jackson, and Ailya Banks, “Technological Threats: How Online Harassment of Female Political Figures Undermines Democracy,” Texas A&M University, June 2022, https://bush.tamu.edu/wp-content/uploads/2022/06/WPS-Report-FINAL-draft.pdf.

61 “Sexism, harassment and violence against women parliamentarians,” Inter-Parliamentary Union, 2016, accessed July 29, 2024, 4, https://www.ipu.org/resources/publications/issue-briefs/2016-10/sexism-harassment-and-violence-against-women-parliamentarians.

62 Hannah Al-Othman, “Female UK election candidates report increased abuse,” The Guardian, June 14, 2024, https://www.theguardian.com/politics/article/2024/jun/14/abuse-female-election-candidates-getting-worse-uk-targeted.

63 “Candidate Survey May 2024 Data Tables,” The Electoral Commission, July 3, 2024, https://www.electoralcommission.org.uk/news-and-views/media-centre/candidate-survey-may-2024-data-tables?utm_source=blognotification&utm_medium=email&utm_campaign=Blog%20Post%20Notification%20Women%20Around%20the%20World&utm_term=WomenAroundWorld.

64 Kirsten Zeiter, Sandra Pepera, and Molly Middlehurst, “Tweets That Chill: Analyzing Online Violence Against Women in Politics,” National Democratic Institute, May 2019, accessed July 29, 2024, https://www.ndi.org/sites/default/files/NDI%20Tweets%20That%20Chill%20Report.pdf.

65 Daniel Funke and Daniela Flamini, “A guide to anti-misinformation actions around the world,” Poynter, Last Modified August 13, 2019, accessed July 29, 2024, https://www.poynter.org/ifcn/anti-misinformation-actions/#uk. Also see: Kristen Kern, “Artificial Intelligence's Impact on Election Mis- and Disinformation,” League of Women Voters, December 21, 2023, https://www.lwv.org/blog/artificial-intelligences-impact-election-mis-and-disinformation.

66 Julie Posetti and Nabeelah Shabbir, “The Chilling: A Global Study On Online Violence Against Women Journalists,” International Center for Journalists and UNESCO, 2022, accessed July 29, 2024, https://www.icfj.org/our-work/chilling-global-study-online-violence-against-women-journalists.

67 A 2024 study on gendered disinformation in Brazil found a close relationship between Bolsonaro's authoritarian regime and a pointed rise in gendered disinformation operations. See: “She Persisted: Brazil Report,” ShePersisted.Org, 2024, https://she-persisted.org/wp-content/uploads/2024/04/ShePersisted-Brazil-Report-ENG.pdf.

68 A UN Women report from 2023 documented rising online harassment and gendered disinformation intended to, and often successful in, discrediting and deplatforming female candidates. See: “Online Harassment Risks Pushing Kenyan Women Out of Politics,” UN Women, September 13, 2023, https://www.unwomen.org/en/news-stories/feature-story/2023/09/online-harassment-risks-pushing-kenyan-women-out-of-politics.

69 Gendered disinformation is routinely used by male political candidates and officeholders Kenya to discredit and undermine their female opponents. See: Kwaku Krobea Asante, “How misinformation will be gender-based in Ghana's upcoming elections,” Poynter, March 29, 2024, accessed July 19, 2024, https://www.poynter.org/commentary/2024/how-misinformation-will-be-gender-based-in-ghanas-upcoming-elections/.

70 A study in Georgia looked at gendered disinformation, harassment, and hate speech on social media over the course of one year and found that “the most prevalent manifestations” rose during times of political instability. See: “MDF Report on Sexist Language and Gender Disinformation in Georgia,” Civil Georgia, April 5, 2024, accessed July 29, 2024, https://civil.ge/archives/590094.

71 Women politicians in India face severe abuse on Twitter, with Muslim women politicians facing 94.1 percent more abuse. See: “New Study Shows Shocking Scale of Abuse on Twitter Against Women Politicians in India,” Amnesty International USA, January 23, 2020, https://www.amnestyusa.org/press-releases/shocking-scale-of-abuse-on-twitter-against-women-politicians-in-india/.

72 A study on gendered disinformation in Hungary found that women who speak out against Prime Minister Victor Orban's autocratic regime face severe misogynistic attacks online. See: “She Persisted: Hungary Report” ShePersisted.Org, 2023, https://she-persisted.org/wp-content/uploads/2023/03/ShePersisted_Hungary.pdf.

73 “Three in four women not comfortable expressing political opinions online,” The Alan Turing Institute, March 20, 2024, https://www.turing.ac.uk/news/three-four-women-not-comfortable-expressing-political-opinions-online.

74 Lucina Di Meco and Kristina Wilfore, “Gendered disinformation is a national security problem,” Brookings Institution, March 8, 2021, https://www.brookings.edu/articles/gendered-disinformation-is-a-national-security-problem/.

75 Adam Satariano, “An Experiment to Stop Online Abuse Falls Short in Germany,” The New York Times, September 23, 2021, https://www.nytimes.com/2021/09/23/technology/online-hate-speech-germany.html; See also: Julia Smirnova, Hannah Winter, Nora Mathelemuse, Mauritius Dorn, and Helena Schwertheim, “Digitale Gewalt und Desinformation gegen Spitzenkandidat:innen vor der Bundestagswahl 2021,” ISD, (2021), https://www.isdglobal.org/wp-content/uploads/2021/09/Digitale-Gewalt-und-Desinformation_v5.pdf for original report in German studying gendered disinformation during the 2021 German elections.

76 “The gendered disinformation playbook in Germany is a warning for Europe,” Brookings Institution, October 29, 2021, accessed July 30, 2024, https://www.brookings.edu/articles/the-gendered-disinformation-playbook-in-germany-is-a-warning-for-europe/.

77 “White House Task Force To Address Online Harassment and Abuse: Final Report and Blueprint,” The White House, May 2024, https://www.whitehouse.gov/wp-content/uploads/2024/05/White-House-Task-Force-to-Address-Online-Harassment-and-Abuse_FINAL.pdf. Accessed July 15, 2024.

78 “White House Task Force To Address Online Harassment and Abuse: Final Report and Blueprint,” The White House, May 2024.

79 The concept ‘enemy images’ has been used for a variety of empirical case studies in the last 30+ years. While the concept is not *exclusive* to TFGBV, researchers (including co-author Isabella Gomez-O’Keefe) have approached the framework with a ‘gendered’ lens.

80 Louis Oppenheimer, “The Development of Enemy Images: A Theoretical Contribution,” Peace and Conflict, 12(3), 269-292; Ragnhild Fiebig-von Hase, Ursula Lehmkuhl, eds, Enemy Images in American History (Berghahn Books, 1998).

81 Ragnhild Fiebig-Von Hase and Ursula Lehmkuhl, eds., Enemy Images in American History (Herndon, VA: Berghahn Books, 1998.); Bo Petersson, “Hot Conflict and Everyday Banality: Enemy images, scapegoats and stereotypes,” Development 52(4), (2009), 460-465; Byron B. Renz, Our own worst enemy as protector of ourselves: Stereotypes, schemas, and typifications as integral elements in the persuasive process (University Press of America, 2010); Walter G. Stephan, Oscar Ybarra, and Kimberly Rios Morrison, “Intergroup threat theory,” In Handbook of prejudice, stereotyping, and discrimination (Psychology Press, 2015), 255-278.

82 Daniel Bar-Tal, “Causes and Consequences of Delegitimization: Models of conflict and ethnocentrism,” Journal of Social Issues, 46(1), (1990),65–81, https://doi.org/10.1111/j.1540-4560.1990.tb00272.x, Nick Haslam, “Dehumanization: An Integrative Review,” Personality and Social Psychology Review, 10(3), (2006), 252–264. https://doi.org/10.1207/s15327957pspr1003_4; Brock Bastian, Simon M Laham, Sam Wilson, Nick Haslam, and Peter Koval, “Blaming, praising, and protecting our humanity: The implications of everyday dehumanization for judgments of moral status,” British Journal of Social Psychology, 50(3), (2011), 469–483. https://doi.org/10.1348/014466610x521383.

83 Robert Jervis, “Political implications of loss aversion,” Political psychology, (1992), 187-204; Walter G. Stephan, Oscar Ybarra, and Kimberly Rios Morrison, “Intergroup threat theory,” In Handbook of prejudice, stereotyping, and discrimination (Psychology Press, 2015), 255-278; Alberto Voci, “The link between identification and in-group favouritism: Effects of threat to social identity and trust-related emotions,” British Journal of Social Psychology, 45(2), (2006), 265–284. https://doi.org/10.1348/014466605x52245; Marco Brambilla, Simona Sacchi, Stefano Pagliaro, and Naomi Ellemers, “Morality and intergroup relations: Threats to safety and group image predict the desire to interact with outgroup and ingroup members,” Journal of Experimental Social Psychology, 49(5), (2013), 811–821, https://doi.org/10.1016/j.jesp.2013.04.005.

84 “Dangerous Speech: A Practical Guide,” The Dangerous Speech Project, March 2021, https://dangerousspeech.org/wp-content/uploads/2020/08/Dangerous-Speech-A-Practical-Guide.pdf.

85 “About the Commissioner,” eSafety Commissioner, accessed July 2024, https://www.esafety.gov.au/about-us/who-we-are/about-the-commissioner.

86 “Submission to the Australian Competition and Consumer Commission Digital Platforms Inquiry,” Office of the eSafety Commissioner, February 2019, https://www.accc.gov.au/system/files/Office%20of%20the%20eSafety%20Commissioner%20%28February%202019%29.PDF.

87 “Regulatory Information,” eSafety Commissioner, accessed July 31, 2024, https://www.esafety.gov.au/industry/regulatory-information.

88 “Regulatory Schemes,” eSafety Commissioner, accessed July 18, 2024, https://www.esafety.gov.au/about-us/who-we-are/regulatory-schemes.

89 “Basic Online Safety Expectations,” eSafety Commissioner, accessed July 30, 2024, https://www.esafety.gov.au/industry/basic-online-safety-expectations.

90 “Responses to Transparency Notices,” eSafety Commissioner, accessed July 30, 2024, https://www.esafety.gov.au/industry/basic-online-safety-expectations/responses-to-transparency-notices.

91 “Full Report: Basic Online Safety Expectations - Summary of Response from X Corp/Twitter to eSafety’s Transparency Notice on Online Hate,” eSafety Commissioner, January 2024, https://www.esafety.gov.au/sites/default/files/2024-01/Full-Report-Basic-Online-Safety-Expectations-Summary-of-response-from-X-CorpTwitter-to-eSafetys-transparency-notice-on-online%20hate.pdf?v=1722370491682.

92 "Full Report: Basic Online Safety Expectations - Summary of Response from X Corp/Twitter to eSafety's Transparency Notice on Online Hate," eSafety Commissioner, January 2024, https://www.esafety.gov.au/sites/default/files/2024-01/Full-Report-Basic-Online-Safety-Expectations-Summary-of-response-from-X-CorpTwitter-to-eSafetys-transparency-notice-on-online%20hate.pdf?v=1722370491682.

93 "Responses to transparency notices," eSafety Commissioner, accessed July 18, 2024, https://www.esafety.gov.au/industry/basic-online-safety-expectations/responses-to-transparency-notices.

94 Interview with Julie Inman Grant, conducted via Zoom, July 8, 2024.

95 Interview with Inman Grant, 2024.

96 Michael Gordon, "eSafety Commissioner Orders X and Meta to Remove Violent Videos Following Sydney Church Stabbing," The Guardian, April 16, 2024, https://www.theguardian.com/australia-news/2024/apr/16/esafety-commissioner-orders-x-and-meta-to-remove-violent-videos-following-sydney-church-stabbing.

97 "Learn about the Online Safety Act," eSafety Commissioner, March 18, 2024. https://www.esafety.gov.au/newsroom/whats-on/online-safety-act.

98 "Statement on Federal Court Order," eSafety Commissioner, April 24, 2024, https://www.esafety.gov.au/newsroom/media-releases/statement-on-federal-court-order. Accessed July 18, 2024.

99 "Violent Content," X, (May 2024), https://help.x.com/en/rules-and-policies/violent-content.

100 Elon Musk, "That is exactly the issue. Should the eSafety Commissar (an unelected official) in Australia have authority over all countries on Earth?" X, April 22, 2024, accessed July 18, 2024, https://x.com/elonmusk/status/1782584820549202024.

101 Elon Musk "The Australian censorship commissar is demanding *global* content bans!" X, April 19, 2024, accessed July 18, 2024, X.Com https://x.com/elonmusk/status/1781394185951563973.

102 "Defining 'Online Abuse': A Glossary of Terms," PEN America, accessed August 5, 2024, https://onlineharassmentfieldmanual.pen.org/defining-online-harassment-a-glossary-of-terms/.

103 Data obtained via Meltwater and confirmed with the eSafety Commissioner's office.

104 Data obtained via Meltwater and confirmed with the eSafety Commissioner's office.

105 Kaitlyn Tiffany, "How 'Karen' Became a Coronavirus Villain," The Atlantic, May 6, 2020, https://www.theatlantic.com/technology/archive/2020/05/coronavirus-karen-memes-reddit-twitter-carolyn-goodman/611104/.

106 Daniel Bar-Tal, "Causes and Consequences of Delegitimization: Models of conflict and ethnocentrism," Journal of Social Issues, 46(1), (1990), 65–81, https://doi.org/10.1111/j.1540-4560.1990.tb00272.x.

107 Interview with Inman Grant, 2024.

108 Nick Haslam, "Dehumanization: An Integrative Review," Personality and Social Psychology Review, 10(3), (2006), 252–264, https://doi.org/10.1207/s15327957pspr1003_4.

109 For political figures see: Maria Giuseppina Pacilli, Michele Roccato, and Silvia Russo, "From political opponents to enemies? The role of perceived moral distance in the animalistic dehumanization of the political outgroup," Group Processes & Intergroup Relations/Group Processes and Intergroup Relations, 19(3), (2015), 360–373. https://doi.org/10.1177/1368430215590490.

110 Silverstein, B., & Holt, R. R. (1989). Research on enemy images: Present status and future prospects. Journal of Social Issues, 45(2), 159-175; Whitford, A. B., & Yates, J. (2003). Policy signals and executive governance: Presidential rhetoric in the war on drugs. The Journal of Politics, 65(4), 995-1012; Hall, J. (2021). In search of enemies: Donald Trump's populist foreign policy rhetoric. Politics, 41(1), 48-63, Badie, D. (2010). Groupthink, Iraq, and the war on terror: Explaining US policy shift toward Iraq. Foreign Policy Analysis, 6(4), 277-296.

111 Debra Merskin, (2004). "The construction of Arabs as enemies: Post-September 11 discourse of George W. Bush," (2004); Brett Silverstein, "Enemy images: The psychology of US attitudes and cognitions regarding the Soviet Union," (1989); Lennat Soberon, "Reaganite America and Its Mnemonic Menaces: A Thematic Analysis of Collective Trauma and Enemy Image Construction in the 1980s American Action Film," (2020).

112 Bo Petersson, “Hot Conflict and Everyday Banality: Enemy images, scapegoats and stereotypes,” Development 52(4), (2009), 460-465; Nira Yuval-Davis, “Theorizing identity: Beyond the ‘us’ and ‘them’ dichotomy,” Patterns of Prejudice, 44(3), (2010), 261-280.

113 Muzaffer Ercan Yilmaz, “Enemy images and conflict,” I.Ü. Siyasal Bilgiler Fakültesi Dergisi No: 32 (March 2005); R.K. White, “Enemy Images: A Resource Manual on Reducing Enmity,” Psychologists for Social Responsibility; William Eckhardt, “Making and breaking enemy images,” Bulletin of Peace Proposals 22(1) (1991), 87-95; Ulrich Beck, “The Sociological Anatomy of Enemy Images: The Military and Democracy After the End of the Cold War,” in Enemy Images in American History, edited by Ragnhild Fiebig-von Hase and Ursula Lehmkuhl (New York, Oxford: Berghahn Books, 1998), pp. 65-88; Mina Cikara, Emile G. Bruneau, Rebecca R. Saxe, “Us and Them: Intergroup Failures of Empathy,” Current Directions in Psychological Science 20(3) (2011), 149-153.

114 Dima Salih, Through the Lens of ISIS: The Portrayal of the Female Enemy and Sexual Violence in ISIS Online Magazines Dabiq and Rumiyah, University of Helsinki Master’s Thesis (2019).

115 John Bourne, “Julie Inman Grant Says Elon Musk’s ‘Dog Whistle’ Remarks Have Contributed to Abuse,” ABC News, June 5, 2024, https://www.abc.net.au/news/2024-06-05/julie-inman-grant-says-elon-musks-dog-whistle/103938880.

116 “Statement from the eSafety Commissioner re: Federal Court proceedings,” eSafety Commissioner, June 5, 2024, https://www.esafety.gov.au/newsroom/media-releases/statement-from-the-esafety-commissioner-re-federal-court-proceedings.

117 Data obtained via Meltwater and confirmed with the eSafety Commissioner’s Office.

118 Interview with Inman Grant, 2024.

119 Interview with Inman Grant, 2024.

120 Interview with Inman Grant, 2024.

121 “No Trade-off between Women’s Right to Safety and Their Right to Speak, Says UN Expert,” United Nations Human Rights Office of the High Commissioner, 2023, Last Modified October 13, 2023, https://www.ohchr.org/en/press-releases/2023/10/no-trade-between-womens-right-safety-and-their-right-speak-says-un-expert.

122 Anna Pechenik Gieseke, “The New Weapon of Choice”: Law’s Current Inability to Properly Address Deepfake Pornography,” Vanderbilt Law Review, October 2020, https://scholarship.law.vanderbilt.edu/cgi/viewcontent.cgi?article=4409&context=vlr.

123 Vjosa Isai, “Canada Wants to Regulate Online Content. Critics Say It Goes Too Far,” The New York Times, April 9, 2024, https://www.nytimes.com/2024/04/09/world/canada/canada-hate-speech.html.

124 This issue has become so commonplace that the Cambridge Dictionary defines it as “the act of someone using the internet to send an image of their naked body, especially the genitals (= sexual organs), to someone that they do not know and who has not asked them to do this.” See also: “Cyberflashing: Supporting victims of cyberflashing and giving preventative advice,” UK Safer Internet Centre, accessed July 2024, https://saferinternet.org.uk/online-issue/cyberflashing.

125 Nicholas Fandos. “Senate Passes Child Online Safety Bill, Setting Up House Showdown,” The New York Times, July 30, 2024. https://www.nytimes.com/2024/07/30/us/politics/senate-child-online-safety-bill-house.html; See also: Tech Policy Press’ policy tracker for the 100+ US state bills on online safety, spanning roughly 40 states, in lieu of federal regulation: “Policy Tracker,” Tech Policy Press, accessed July 30, 2024, https://www.techpolicy.press/tracker/. See also: Tech Policy Press’ 2023 snapshot of US state bills for child online safety: Bernard, Tim. “144 State Bills Aim to Secure Child Online Safety As Congress Flounders,” Tech Policy Press, May 22, 2023, https://www.techpolicy.press/144-state-bills-aim-to-secure-child-online-safety-as-congress-flounders/.

126 “Regulatory information,” eSafety Commissioner, July 1, 2024, https://www.esafety.gov.au/industry/regulatory-information.

127 “Submission to the Australian Competition and Consumer Commission Digital Platforms Inquiry,” eSafety Commissioner, February 2019, https://www.accc.gov.au/system/files/Office%20of%20the%20eSafety%20Commissioner%20%28February%202019%29.PDF.

128 “Basic Online Safety Expectations: Regulatory Guidance,” eSafety Commissioner, July 2024, https://www.esafety.gov.au/sites/default/files/2024-07/Basic-Online-Safety-Expectations-regulatory-guidance-July-2024.pdf.

129 “Cyberbullying Scheme,” eSafety Commissioner, December 2023, accessed July 22, 2024, https://www.esafety.gov.au/sites/default/files/2023-12/Cyberbullying-Scheme-Regulatory-Guidance-Updated-December2023.pdf.

130 “Image-Based Abuse Scheme - Regulatory Guidance,” eSafety Commissioner, November 2021, accessed July 22, 2024, https://www.esafety.gov.au/sites/default/files/2022-03/Image-Based%20Abuse%20Scheme%20Regulatory%20Guidance.pdf.

131 “Adult Cyber Abuse Scheme Regulatory Guidance,” eSafety Commissioner, December 2023, accessed July 22, 2024, https://www.esafety.gov.au/sites/default/files/2023-12/Adult-Cyber-Abuse-Scheme-Regulatory-Guidance-Updated-December2023.pdf.

132 “Online Content Scheme Regulatory Guidance,” eSafety Commissioner, December 2023, accessed July 22, 2024, https://www.esafety.gov.au/sites/default/files/2023-12/Online-Content-Scheme-Regulatory-Guidance-Updated-December-2023.pdf.

133 “Abhorrent Violent Conduct Powers - Regulatory Guidance,” eSafety Commissioner, accessed July 22, 2024, https://www.esafety.gov.au/sites/default/files/2022-03/Abhorrent%20Violent%20Conduct%20Powers%20Regulatory%20Guidance.pdf.

134 “How the new Online Safety Act supports women,” eSafety Commissioner, accessed July 22, 2024, https://www.esafety.gov.au/sites/default/files/2022-03/OSA%20Fact%20sheet%20Women_0.pdf.

135 “Women In The Spotlight,” eSafety Commissioner, accessed July 2024, https://www.esafety.gov.au/women/women-in-the-spotlight.

136 “Questions and Answers on the Digital Services Act,” The European Commission, February 2024, https://ec.europa.eu/commission/presscorner/detail/en/QANDA_20_2348.

137 “The Digital Services Act,” The European Commission, accessed July 2024, https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/digital-services-act_en.

138 “The Digital Services Act,” European Commission, 2024, https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/digital-services-act_en.

139 “Digital Services Act - Consilium,” European Council. Accessed July 22, 2024, from https://www.consilium.europa.eu/en/policies/digital-services-act/

140 “Digital Services Act - Consilium,” European Council.

141 “DSA: Very large online platforms and search engines,” European Commission, February 21, 2024, accessed July 22, 2024, https://digital-strategy.ec.europa.eu/en/policies/dsa-vlops.

142 “Ending gender-based violence,” European Commission, accessed July 22, 2024, https://commission.europa.eu/strategy-and-policy/policies/justice-and-fundamental-rights/gender-equality/gender-based-violence/ending-gender-based-violence_en.

143 “Ending gender-based violence,” European Commission.

144 “International Women’s Day: The Need for a Harmonised, Human Rights-Centred EU Legal Framework on Online Gender-Based Violence,” Center for Democracy and Technology, Last Modified March 7, 2023, https://cdt.org/insights/international-womens-day-the-need-for-a-harmonised-human-rights-centred-eu-legal-framework-on-online-gender-based-violence/.

145 “CDT Europe Reacts to EU Directive on Gender-Based Violence (GBV) – New Rules to Tackle Online GBV Create Free Expression Concerns,” Center for Democracy and Technology, Last modified March 11, 2024, https://cdt.org/insights/cdt-europe-reacts-to-eu-directive-on-gender-based-violence-gbv-new-rules-to-tackle-online-gbv-create-free-expression-concerns/.

146 Kirk J. Nahra, “The European Parliament Adopts the AI Act, ” Wilmer Hale, 2024, https://www.wilmerhale.com/en/insights/blogs/wilmerhale-privacy-and-cybersecurity-law/20240314-the-european-parliament-adopts-the-ai-act.

147 Gernot Fritz, Teresa Ehlen, and Tina F. Cuvan, “EU AI Act unpacked #8: New rules on deepfakes,” Lexology, 2024, https://www.lexology.com/library/detail.aspx?g=c25237ee-a37f-4959-837c-f32a624f54ab.

148 “EU AI Act: First Regulation on Artificial Intelligence,” European Parliament, Last Modified June 8, 2023, https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence. Last modified June 8, 2023.

149 “EU AI Act,” European Parliament.

150 “Making Britain the safest place in the world to be online,” UK Government, October 11, 2017, https://www.gov.uk/government/news/making-britain-the-safest-place-in-the-world-to-be-online.

151 “Online Safety Act: explainer,” UK Government, May 8, 2024, accessed July 22, 2024, https://www.gov.uk/government/publications/online-safety-act-explainer/online-safety-act-explainer.

152 “Online,” UK Government.

153 “Ofcom’s approach to implementing the Online Safety Act,” Ofcom, October 26, 2023, https://www.ofcom.org.uk/online-safety/information-for-industry/roadmap-to-regulation.

154 “Online safety,” Ofcom, accessed July 23, 2024, https://www.ofcom.org.uk/online-safety/.

155 “Online safety,” Ofcom.

156 “Consultation: Protecting people from illegal harms online,” Ofcom, November 9, 2023, accessed July 23, 2024, https://www.ofcom.org.uk/online-safety/illegal-and-harmful-content/protecting-people-from-illegal-content-online.

157 P. Collings, “EFF’s submission to Ofcom’s consultation on illegal harms,” Electronic Frontier Foundation, March 11, 2024, https://www.endviolenceagainstwomen.org.uk/wp-content/uploads/2024/02/Ofcoms-Illegal-Harms-Consultation-VAWG-Roundtable-public.pdf.

158 Clare McGlynn, “Intimate image abuse,” Clare McGlynn, accessed July 2024, https://www.claremcglynn.com/intimate-image-abuse.

159 “Position Statement Regulatory coherence and coordination: the role of the Global Online Safety Regulators Network,” Global Online Safety Regulators Network, May 2024, https://www.esafety.gov.au/sites/default/files/2024-05/GOSRN-Position-Statement-on-Regulatory-Coherence.pdf?v=1719249816530.

160 “The Global Online Safety Regulators Network”, eSafety Commissioner, July 4, 2024, https://www.esafety.gov.au/about-us/who-we-are/international-engagement/the-global-online-safety-regulators-network#membership.

161 “The Global Online Safety Regulators Network”, eSafety Commissioner, 2024.

162 “Position Statement,” Global Online Safety Regulators Network.

163 “Position Statement,” Global Online Safety Regulators Network.

164 “Image-Based Abuse Scheme - Regulatory Guidance,” eSafety Commissioner, February 2024, accessed July 22, 2024, https://www.esafety.gov.au/sites/default/files/2024-02/Image-Based-Abuse-Scheme-Regulatory-Guidance-Feb2024.pdf?v=1717440210883.

165 “Image-Based,” eSafety Commissioner.

166 “Image-Based,” eSafety Commissioner.

167 Asher Flynn, Anastasia Powell, Adrian J. Scott, and Elena Cama, “Deepfakes and digitally altered imagery abuse: A cross-country exploration of an emerging form of image-based sexual abuse,” The British Journal of Criminology, 62(6), (2021), 1341–1358. https://doi.org/10.1093/bjc/azab111.

168 Karen Middleton, “Jail time for those caught distributing deepfake porn under new Australian laws,” The Guardian, June 2, 2024, https://www.theguardian.com/australia-news/article/2024/jun/01/creating-or-sharing-deepfake-porn-without-consent-to-be-under-proposed-new-australian-laws.

169 Karen Middleton, “Jail time for those caught distributing deepfake porn under new Australian laws,” The Guardian.

170 Paul Karp, Josh Butler, Jordyn Beazley, “Australian government pledges almost $1bn to help women leave violent relationships,” The Guardian, April 30, 2024, https://www.theguardian.com/australia-news/2024/may/01/leaving-violence-payment-australia-women-violent-relationships.

171 “DSA: Very large online platforms and search engines,” European Commission, February 21, 2024, https://digital-strategy.ec.europa.eu/en/policies/dsa-vlops.

[172] “The final text of the Digital Services Act (DSA),” EU Digital Services Act, 2022, https://www.eu-digital-services-act.com/Digital_Services_Act_Preamble_11_to_20.html.

[173] “DSA: Three pornographic websites added to list of ‘very large platforms,’” INCYBER News, December 26, 2023, https://incyber.org/en/article/dsa-three-pornographic-websites-added-to-list-of-very-large-platforms.

[174] Morgan Meaker, “Europe Has Traded Away Its Online Porn Law,” WIRED, April 27, 2022, https://www.wired.com/story/digital-services-act-deepfake-porn/.

[175] “Ending gender-based violence,” European Commission, accessed July 22, 2024, https://commission.europa.eu/strategy-and-policy/policies/justice-and-fundamental-rights/gender-equality/gender-based-violence/ending-gender-based-violence_en.

[176] “Proposal for a Directive of the European Parliament and of the Council on Combating Violence against Women and Domestic Violence,” European Parliament, 2019, https://www.europarl.europa.eu/meetdocs/2014_2019/plmrep/COMMITTEES/LIBE/DV/2024/02-15/VAW-provisionalagreement_EN.pdf.

[177] Carlotta Rigotti and Clare McGlynn, “Towards an EU criminal law on violence against women: The ambitions and limitations of the Commission’s proposal to criminalise image-based sexual abuse,” New Journal of European Criminal Law 13, no. 4, (November 2022): 452-477, https://durham-repository.worktribe.com/js/pdfjs/web/viewer.html?file=https://durham-repository.worktribe.com/previewfile/1184722/37656.pdf.

[178] E. J. Dickson, “Deepfake Porn Is Still a Threat, Particularly for K-Pop Stars,” Rolling Stone, October 7, 2019, https://www.rollingstone.com/culture/culture-news/deepfakes-nonconsensual-porn-study-kpop-895605/.

[179] Nicole de Souza and R. Shanahan, “The Nth Room case and modern slavery in the digital space,” Lowy Institute, April 20, 2020, accessed July 23, 2024, https://www.lowyinstitute.org/the-interpreter/nth-room-case-modern-slavery-digital-space.

[180] Nicole de Souza and R. Shanahan, “The Nth Room case and modern slavery in the digital space,” Lowy Institute.

[181] Nicole de Souza and R. Shanahan, “The Nth Room case and modern slavery in the digital space,” Lowy Institute.

[182] Lina Park, “South Korea’s Deepfake Sex Crime Epidemic Exposes Lack of Legal Protections,” Korea Pro, June 14, 2024, accessed July 3, 2024, https://koreapro.org/2024/06/south-koreas-deepfake-sex-crime-epidemic-exposes-lack-of-legal-protections/.

[183] Nicole de Souza and R. Shanahan, “The Nth Room case and modern slavery in the digital space,” Lowy Institute.

[184] “Statutes of the Republic of Korea,” Statutes of the Republic of Korea, accessed July 22, 2024, https://elaw.klri.re.kr/eng_service/lawView.do?hseq=40947&lang=ENG.

[185] Kaylee Williams, “Exploring Legal Approaches to Regulating Nonconsensual Deepfake Pornography,” Tech Policy Press, May 15, 2023, accessed July 22, 2024, https://www.techpolicy.press/exploring-legal-approaches-to-regulating-nonconsensual-deepfake-pornography/.

[186] Kaylee Williams “Exploring Legal Approaches to Regulating Nonconsensual Deepfake Pornography,” Tech Policy Press.

[187] Kaylee Williams “Exploring Legal Approaches to Regulating Nonconsensual Deepfake Pornography,” Tech Policy Press.

[188] Park, “South Korea’s Deepfake.”

[189] Park, “South Korea’s Deepfake.”

[190] Park, “South Korea’s Deepfake.”

[191] “Online Safety Act: new criminal offences circular,” UK Government, accessed July 22, 2024, https://www.gov.uk/government/publications/online-safety-act-new-criminal-offences-circular/online-safety-act-new-criminal-offences-circular.

[192] “UK Online Safety Bill: A long journey towards the Act,” Glitch, accessed July 18, 2024, https://glitchcharity.co.uk/uk-online-safety-bill/.

193 Clare McGlynn, “Can Deepfake Sites Still Be Accessed in the UK?” Glamour UK, (April 23, 2024), https://www.glamourmagazine.co.uk/article/new-deepfake-laws-whats-next-opinion. Written evidence submitted by Professor Clare McGlynn, Durham Law School, Durham University (OSB0014) Summary Key Omissions from (n.d.), UK Parliament Committees, accessed July 22, 2024, https://committees.parliament.uk/writtenevidence/39012/pdf/. “End Violence Against Women,” February 23, 2024, https://www.endviolenceagainstwomen.org.uk/wp-content/uploads/2024/02/VAWG-letter-to-OFCOM.pdf.

194 “Online Safety Act,” UK Government.

195 “The UK’s Online Safety Act Is Not Enough To Address Non-consensual Deepfake Pornography,” TechPolicy, March 13, 2024, https://www.techpolicy.press/the-uks-online-safety-act-is-not-enough-to-address-nonconsensual-deepfake-pornography/.

196 Elouise Spensley, “Law Column: Update to anonymity provisions for victims of image abuse,” Hold the Front Page, accessed July 23, 2024, https://www.holdthefrontpage.co.uk/2024/news/law-column-update-to-anonymity-provisions-for-victims-of-image-abuse/.

197 “The UK’s Online Safety Act Is Not Enough To Address Non-consensual Deepfake Pornography,” Tech Policy Press, March 13, 2024, accessed July 23, 2024, https://www.techpolicy.press/the-uks-online-safety-act-is-not-enough-to-address-nonconsensual-deepfake-pornography/.

198 “Deepfake porn: why we need to make it a crime to create it, not just share it,” Durham University, April 9, 2024, accessed July 22, 2024, https://www.durham.ac.uk/research/current/thought-leadership/2024/04/deepfake-porn-why-we-need-to-make-it-a-crime-to-create-it-not-just-share-it/.

199 “Government cracks down on ‘deepfakes’ creation,” UK Government, April 16, 2024, accessed July 22, 2024, https://www.gov.uk/government/news/government-cracks-down-on-deepfakes-creation.

200 “Government cracks down,” UK Government.

201 “Government criminalises creation of deepfakes, but with a major loophole,” End Violence Against Women, April 16, 2024, accessed July 22, 2024, https://www.endviolenceagainstwomen.org.uk/government-criminalises-creation-of-deepfakes-but-with-a-major-loophole/.

202 Matt Burgess, “The Biggest Deepfake Porn Website Is Now Blocked in the UK,” WIRED, April 19, 2024, https://www.wired.com/story/the-biggest-deepfake-porn-website-is-now-blocked-in-the-uk/.

203 “The UK’s Online Safety Act,” Tech Policy Press.

204 Clare McGlynn, “Deepfake porn: why we need to make it a crime to create it, not just share it,” The Conversation, April 9, 2024, http://theconversation.com/deepfake-porn-why-we-need-to-make-it-a-crime-to-create-it-not-just-share-it-227177.

205 Steve LeBlanc, “Massachusetts on verge of becoming second-to-last state to outlaw ‘revenge porn’,” AP News, June 13, 2024, accessed July 22, 2024, https://apnews.com/article/revenge-porn-ban-massachusetts-bill-passed-ace13a5bf9e8b5131066dda8ac8b8b65. “Governor Healey Signs Bill Banning Revenge Porn, Expanding Protections Against Abuse and Exploitation,” Mass.gov, June 20, 2024, accessed July 22, 2024, https://www.mass.gov/news/governor-healey-signs-bill-banning-revenge-porn-expanding-protections-against-abuse-and-exploitation.

206 “Text - S.3623 - Violence Against Women Act Reauthorization Act of 2022,” U.S. Congress, accessed July 22, 2024, https://www.congress.gov/bill/117th-congress/senate-bill/3623/text.

207 “Text - S.3696 / H.R. 7569 - 118th Congress (2023-2024): DEFIANCE Act of 2024,” U.S. Congress, January 30, 2024, accessed July 22, 2024, https://www.congress.gov/bill/118th-congress/senate-bill/3696/text.

208 Kat Tenbarge, “The Defiance Act passes in the Senate, potentially allowing deepfake victims to sue over nonconsensual images,” NBC News, July 24, 2024, https://www.nbcnews.com/tech/tech-news/defiance-act-passes-senate-allow-deepfake-victims-sue-rcna163464.

209 “Text - S.3696 / H.R. 7569,” U.S. Congress.

210 “Text - S.3696 / H.R. 7569,” U.S. Congress.

211 “Text - S. 4569 - 118th Congress (2023-2024): TAKE IT DOWN Act,” U.S. Congress, accessed July 22, 2024, https://www.congress.gov/bill/118th-congress/senate-bill/4569/text.

212 “Text - S. 4569,” U.S. Congress.

213 “Text - S. 4569,” U.S. Congress.

214 “Combating Sexual Deepfakes,” multistate.ai, accessed July 22, 2024, https://www.multistate.ai/deepfakes-sexual.

215 “§ 18.2-386.2. Unlawful dissemination or sale of images of another; penalty,” Virginia Law, accessed July 23, 2024, https://law.lis.virginia.gov/vacode/title18.2/chapter8/section18.2-386.2/.

216 Alan Riquelmy, “California legislative committee approves deepfake bill,” Courthouse News Service, April 23, 2024, accessed July 23, 2024, https://www.courthousenews.com/california-legislative-committee-approves-deepfake-bill/.

217 K.C. Halm, Ambika Kumar, Jonathan Segal, and Caesar Kalinowski IV, “Two New California Laws Tackle Deepfake Videos in Politics and Porn,” Davis Wright Tremaine, accessed July 25, 2024, https://www.dwt.com/blogs/media-law-monitor/2020/02/two-new-california-laws-tackle-deepfake-videos-in.

218 Kaylee Williams, “Exploring Legal Approaches to Regulating Nonconsensual Deepfake Pornography,” Tech Policy Press.

219 “Women in Tech Stats 2024: Uncovering Trends and Unseen Data,” WomenTech Network, 2024, accessed August 1, 2024, https://www.womentech.net/en-us/women-in-tech-stats.

220 Sasha Costanza-Chock, “Design Justice: towards an intersectional feminist framework for design theory and practice,” in Storni, C., Leahy, K., McMahon, M., Lloyd, P. and Bohemia, E. (eds.), Design as a catalyst for change - DRS International Conference 2018, 25-28 June, Limerick, Ireland, p. 529, https://doi.org/10.21606/drs.2018.679.

221 “Update to Inappropriate Content Policy (May 2024),” Google Support, May 1, 2024, accessed August 5, 2024, https://support.google.com/adspolicy/answer/14720423?hl=en&ref_topic=29265&sjid=17682351251798728111-NC.

222 Lukas Struppek et al, “Leveraging Diffusion-Based Image Variations for Robust Training on Poisoned Data,” HuggingFace, October 10, 2023, accessed August 5, 2024, https://huggingface.co/papers/2310.06372.

223 See: the UK’s Revenge Porn Helpline where victims of intimate image abuse can go for help in taking down non-consensual intimate images and receive counseling on resources available: https://revengepornhelpline.org.uk/; See also: the White House Task Force to Address Online Harassment and Abuse’s newly launched national 24/7 Image Abuse Helpline and Safety Center https://www.whitehouse.gov/wp-content/uploads/2024/05/White-House-Task-Force-to-Address-Online-Harassment-and-Abuse_FINAL.pdf; See also: South Korea’s Cyber Sexual Violence Response Center where victims can receive counseling, investigation assistance, legal advice, mental health support, and content takedown services in addition to educational resources and literacy campaigns https://www.cyber-lion.com/helpline and https://www.kocsc.or.kr/mainPage.do; See also: Australia’s online portal to report intimate image abuse and access guides and resources https://www.esafety.gov.au/report/forms.

224 Nina Jankowicz, “Being online should not come with worry,” Centre for Information Resilience, December 21, 2023, accessed July 22, 2024, https://www.info-res.org/being-online-should-not-come-with-worry.

225 Ahlenback et al, “Technology.”

226 Ahlenback et al, “Technology.”

ABOUT THE AUTHORS

NINA JANKOWICZ

Nina Jankowicz, the co-founder and CEO of The American Sunlight Project, is an internationally recognized expert on disinformation and democratization, one of TIME magazine's 100 Most Influential People in AI, and the author of two books: *How to Lose the Information War* (2020), which the New Yorker called "a persuasive new book on disinformation as a geopolitical strategy," and *How to Be A Woman Online* (2022), an examination of online abuse and disinformation and tips for fighting back, which Publishers Weekly named "essential." Jankowicz has advised governments, international organizations, and tech companies, and testified before the U.S. Congress, UK Parliament, and European Parliament.

In 2022, Jankowicz was appointed to lead the Disinformation Governance Board, an intra-agency best practices and coordination entity at the Department of Homeland Security; she resigned the position after a sustained disinformation campaign caused the Biden Administration to abandon the project. From 2017-2022, Jankowicz has held fellowships at the Wilson Center, where she led accessible, actionable research about the effects of disinformation on women and freedom of expression around the world. She advised the Ukrainian Foreign Ministry on strategic communications under the auspices of a Fulbright-Clinton Public Policy Fellowship in 2016-17. Early in her career, she managed democracy assistance programs to Russia and Belarus at the National Democratic Institute. She completed her MA at Georgetown University's School of Foreign Service and her BA, magna cum laude, at Bryn Mawr College.

ISABELLA GOMEZ-O'KEEFE

Isabella Gomez-O'Keefe is a PhD Student in Political Sociology at the University of Cambridge, and a Stamps Scholar at Queens' College Cambridge. Isabella also works as a research associate at the Centre for Information Resilience, where she has been part of a number of research projects examining identity-based disinformation and hate speech online. She previously earned an MPhil in Political & Economic Sociology from the University of Cambridge, BA in International Affairs & Security Studies from the University of Georgia, and completed two visiting terms at the University of Oxford studying International Conflict.

Isabella conducts interdisciplinary research in the fields of politics, sociology, and social psychology, with a focus on American politics. Her current work examines mainstream American political discourse during US elections, protests, and moments of political contention in the last five years, looking specifically at the use of socio-cognitive 'enemy images' on YouTube and TikTok. The research builds on her previous projects exploring the links between the U.S. Capitol insurrection and 'enemy images' in mainstream right-wing rhetoric on YouTube and X. Isabella also has a background in Middle Eastern politics, including working on U.S.-Iran policy at the State Department and researching discursive radicalization and sexual violence in the Islamic State.

LAUREN HOFFMAN

Lauren Hoffman is the special assistant to Secretary Hillary Rodham Clinton and the associate director of the Institute of Global Politics at Columbia University's School of International and Public Affairs (SIPA).

Prior to joining Columbia SIPA, Hoffman was the associate director of the Women's Initiative at the Center for American Progress (CAP), focusing on women's economic security. Hoffman's work has been cited by the U.S. House of Representatives Committee on Oversight and Reform, Deloitte, the Urban Institute, amicus briefs, and law review articles, and her commentary has appeared in CNBC, Business Insider, Ms. Magazine, The Daily Beast, and MarketWatch.

Hoffman has also worked at Cahill Gordon & Reindel LLP, TIME'S UP, the National Women's Law Center, the Sheryl Sandberg & Dave Goldberg Family Foundation, and with the Women and Foreign Policy program at the Council on Foreign Relations. She started her career as a legal analyst at Goldman Sachs & Co. Hoffman graduated with a JD from American University Washington College of Law and a BA from Yale University in Political Science and French.

ANDREA VIDAL BECKER

Andrea Vidal Becker is currently a consultant for the Institute of Global Politics where she was an inaugural Student Scholar for the class of 2023-2024. Previously, Andrea spent four years serving newly arrived refugees as a Program Supervisor with the International Rescue Committee, where her innovative programming for asylum-seeking youth was awarded by global leadership. She has also worked as a development associate with Human Rights Watch in Silicon Valley and served on the board of Urban Refuge, a non-profit aid-mapping tool that she co-founded as an undergraduate studying International Relations at Boston University. She also holds a Masters in International Affairs from Columbia University's School of International and Public Affairs (SIPA) with focuses in human rights, technology, media and policy. Andrea's research interests range from the ways in which emerging technology impacts human rights to tech policy and humanitarian aid.

THANK YOU

The research team would like to thank:

- Australian eSafety Commissioner Julie Inman Grant for allowing us to analyze her story of online harassment and abuse.
- Tom Jarvis, an open source analyst who focuses on large-scale data analysis and human rights abuses, for his expertise leading the quantitative analysis conducted for this report.
- IGP program assistants, Chelsea Tabachnik and Harrison Gerson, for their indispensable help with research and fact-checking.

The research team would like to also thank the following experts, who provided invaluable advise and guidance. We are grateful to them for sharing their time, expertise, and experience:

- Seyi Akiwowo, *Founder and CEO, Glitch*
- Tracy Chou, *Founder & CEO, Block Party*
- Gabriela de Oliveira, *Head of Policy, Research and Campaigns, Glitch*
- Shauna Dillavou, *CEO, Brightlines*
- Camille François, *Associate Professor of Practice of International and Public Affairs; IGP Affiliated Faculty Member, Columbia School of International and Public Affairs*
- Susanna Gibson, *Nurse Practitioner, Virginia State House Candidate, Founder and Executive Director, MyOwn Image*
- Carrie Goldberg, Esq., *Founding Attorney, C.A. Goldberg PLLC*
- Nadine Farid Johnson, *Policy Director, Knight First Amendment Institute at Columbia University*
- Clare McGlynn, *Professor of Law, Durham University*
- Ella McPherson, *Associate Professor of the Sociology of New Media and Digital Technology and Co-Director of the Centre of Governance and Human Rights*
- Ella Serry, *Manager of International Engagement, eSafety Commissioner*
- Lorna Woods, *Professor of Internet Law, Essex University*
- Celia Zolynski, *Professor of Private Law, Université Paris 1 Panthéon-Sorbonne*

SIPA | IGP Institute of Global Politics

International Affairs Building, 15th Floor
420 West 118th Street
New York, NY 10027

212-853-4720
igp.sipa.columbia.edu
igp@sipa.columbia.edu

VITAL VOICES
GLOBAL PARTNERSHIP

1509 16th Street NW
Washington, D.C. 20036

(202) 861-2625
vitalvoices.org