



DUE DILIGENCE AND ACCOUNTABILITY FOR TECHNOLOGY-FACILITATED VIOLENCE AGAINST WOMEN

Zarizana Aziz
Due Diligence Project

About the author: Zarizana Aziz is a human rights lawyer and is the Director of the Due Diligence Project and co-developed the Due Diligence Framework on State Accountability for Eliminating Violence against Women.

Key words: due diligence, technology-facilitated violence against women, consent, human rights, state, internet intermediaries, international law

Summary:

This paper explores what is technology-facilitated violence against women; what can be done to stem and ultimately eliminate it; and whose responsibility is it to do so. It does this by building upon and furthering the issues identified in two research projects, namely the research on State accountability to eliminate violence against women by the Due Diligence Project (DDP)¹ and the research on corporate and state remedies for dealing with technology-facilitated violence against women by the Association for Progressive Communications (APC).²

The paper further looks at the roles played by both States and private corporations as well as the legislative and non-legislative changes that are needed to ensure that women are able to exercise their right to freedom of expression without the fear of harassment and violence. It recommends that innovations in other fields of online jurisprudence could provide a template for addressing gender-based violence online and ends with the framework on State and Internet Intermediaries Obligation to Eliminate Technology-Facilitated Violence against Women.

Main concepts:

- **Technology-facilitated violence against women:** are acts of gender-based violence 'committed, abetted or aggravated' in part or fully by the use of information and communication technologies, such as cyber stalking; accessing or disseminating a woman's private data (through hacking); identity theft or doxing.

¹ Zarizana Abdul Aziz & Janine Moussa, *Due Diligence Framework: State Accountability Framework for Eliminating Violence against Women*, International Human Rights Initiative (Feb. 2014), available at <http://www.duediligenceproject.org/ewExternalFiles/Due%20Diligence%20Framework%20Report%20Z.pdf>.

² Association for Progressive Communications, *From impunity to justice: Exploring corporate and legal remedies for technology-related violence against women*, <http://genderit.org/onlinevaw/> (last visited June 10, 2017).

- **Due diligence:** International law mandates States to exercise due diligence to promote, protect and fulfill human rights. This includes the obligation to prevent violations, protect victims/survivors of human rights abuses, prosecute violations, punish perpetrators and provide redress and reparation for victims/survivors. This also includes the obligation to remove impunity and preventing human rights abuses by non-state actors. Non-State includes transnational³ and national corporations operating within the jurisdiction of the State.
- **Internet intermediaries** bring together or facilitate transactions between third parties on the internet. They give access to, host, transmit and index content, products and services originated by third parties on the internet or provide internet-based services to third parties.
- **Intermediary liability** in the context of this paper refers to the legal liability of internet intermediaries for content contributed by, or activities carried out by, third parties. The liability approach this paper pursues is “notice and takedown” systems, i.e. systems that require intermediaries to act expeditiously to remove content which is deemed to be unlawful once they have been given notice of the content to ensure that their sites do not serve as vehicles for violating material. Such take down orders should be issued by a judicial authority, be clear and unambiguous, and follow due process.

Key facts:

- Online violence against women presents specific challenges in gauging which data or images constitute violence. What is actionable violence and what is not is gauged by intent to harm, content, imminence of harm (credibility), extent of the harm and context.
- ICT provides a fertile terrain that amplifies reach of transmission. This aggravates the harm to the exercise and enjoyment of human rights and freedoms, particularly the right to privacy or respect for private life caused by the communication of the violating material compared to more traditional media.
- Patriarchy and prevailing interpretations of moral norms, culture and religion places women as the primary bearers of honour and tradition. Women who establish cyber-friendships or relationships may be deemed to have transgressed culturally appropriate behaviour as are women who engage in sexting, exchanging images or who consent to intimate partners taking suggestive images, albeit for private purposes.
- In relation to violence against women, consent is key to differentiating lawful from unlawful and harmful behaviour. Consent in an online context is often complicated by the exact act to which the consent, if any, relates. Because of this, defining consent is crucial in technology-facilitated violence and must be addressed in any relevant mechanisms.
- The enhanced anonymity offered by digital and virtual spaces, through encryption and privacy protocols, provides particular challenges in identifying perpetrators of technology-facilitated violence against women and magnifies impunity.
- It is simplistic to view anonymity as a threat that needs to be removed under all circumstances. As anonymity offers privacy to victims/survivors (whose privacy is often violated by perpetrators) and allows them to re-enter online spaces or to report violence. The anonymity provided by the internet is also beneficial to whistle-blowers, human rights defenders or to those outside current dominant groups, such as LGBTQIA people.

³ Transnational corporations are companies that operate across borders. This raises challenges in terms of the regulating country (where the harm of the crime arose).

Introduction

Increased prevalence of technology-facilitated violence against women, the lack of effective measures to prevent and contain it, and the ensuing impunity must be addressed as part of the struggle to eliminate all forms of gender-based violence. Eliminating technology-facilitated violence against women is made all the more critical given the increasingly central role of online information and communications technology which in many instances has become the main form of communication in commercial dealings as well as personal, political and social interaction.

The internet, once a liberating space, is also, increasingly, a space of violence, particularly violence targeting women. While it is beyond the scope of this paper to explore why women are targeted within online spaces, technology-facilitated violence against women is part of the continuum of violence against women that is committed offline. It reflects and parallels the reality of offline violence against women with the same causes and similar consequences. Like offline violence against women, internet-related violence against women is often in the form of sexual violence such as threats of rape, non-consensual dissemination of intimate data and images, dissemination of rape recordings, cyber stalking, sexual harassment and the exploitation of women and girls.⁴

Another group of persons susceptible to technology-facilitated violence is the LGBTIQ community. In so far as its form, frequency and severity can be compared to approximate the form, frequency and severity of technology-facilitated violence against women, this paper is equally applicable to addressing and eliminating violence against LGBTIQ persons.

Freedom of expression and access to information are key enabling rights to a range of human rights. Online violence prevents women and girls from fully exercising these rights. Thus, removing violence against women from digital and online platforms has the net effect of promoting and strengthening freedom of expression as it creates an environment that allows more individuals, especially sections of society who face discrimination in other public spaces, to participate in these media.⁵

Initiatives by States and internet intermediaries to confront technology-facilitated violence have proven ineffective in stemming technology-facilitated violence, protecting women, bringing the perpetrators to account or in providing satisfactory redress for victims/survivors. In her September 2016 report, the United Nations Special Rapporteur on Violence against Women, its Causes and Consequences, Dubravka Šimonović, identified technology-facilitated violence as a new challenge and one of her priority issues:

"While the use of information and communications technology has contributed to the empowerment of women and girls, its use has also generated technology-facilitated

⁴ Association for Progressive Communications, *Analysis of Incidents of Technology-related Violence Against Women Reported on the "Take Back the Tech!" Ushahidi Platform* (Sept. 9, 2015), available at <http://www.genderit.org/resources/analysis-incidents-reported-take-back-tech-ushahidi-platform>.

⁵ A 2015 report on the status of freedom of expression in Norway cites the Norwegian survey on the status of freedom of speech from 2014 that "shows that hate speech can have harmful effects for those who participate in public debate. In the survey, it emerges that the harm is greater among people with ethnic minority backgrounds than those with majority background," it further notes that "there is no reason to assume that the same harmful effects don't also apply to other groups who are particularly vulnerable to hate speech related to actual or perceived personal characteristics." The report also documents that such speech intimidates people and deters them from speaking publicly. See *The Equality and Anti-Discrimination Ombud's Report: Hate Speech and Hate Crime* (2015), available at http://www.genderit.org/sites/default/upload/hate_speech_and_hate_crime_v3_lr.pdf.

violence. ... [T]here is a need to examine this recent phenomenon, and the applicability of national laws to it, and to make recommendations for States and non-State actors to fight technology-facilitated violence against women and girls while respecting freedom of expression and the prohibition of incitement to violence and hatred, in accordance with article 20 of the International Covenant on Civil and Political Rights.”⁶

This paper explores what is technology-facilitated violence against women; what can be done to stem and ultimately eliminate it; and whose responsibility it is to do so. It does this by building upon and furthering the issues identified in two research projects, namely the research on State accountability to eliminate violence against women by the Due Diligence Project (DDP)⁷ and the research on corporate and state remedies for dealing with technology-facilitated violence against women by the Association for Progressive Communications (APC).⁸

➤ Outline of paper

The paper outlines women experiences in accessing justice; identifies and describes the issues, actors and stakeholders; the role of the State as well as private sector actors; existing mechanisms; application of international human rights law; and good or promising practices in this context. It concludes with recommendations.

Part I will look at violence against women in general and the ability of technology to amplify violence against women. Technology provides platforms capable of masking perpetrators as well as allowing perpetrators to commit violence at increased distance, speed and rate. The capacity of technology to store data and images complicates the provision of remedies.

Part II looks at actors and stakeholders. The primary actor is the perpetrator, namely the originator (author) of the technology-facilitated violence. Layers of encryption allow the perpetrator to remain anonymous. Further, any post can be distributed or accessed online, drawing secondary transmitters who unwittingly or knowingly amplify the harm to the victim/survivor.

Platform providers and intermediaries often deny liability or even responsibility to ensure that their sites do not serve as vehicles for violations. This complicates victims/survivors ability to obtain remedy which requires the cooperation of these intermediaries.

Part III dissects what constitutes infringement. How do we differentiate between legitimate exercise of freedom of expression and violence? The issue of expression in the face of harm has been dealt with in other areas that may prove helpful in defining infringement in violence against women.

Part IV looks at the application of international law and issues of accountability for technology-facilitated violence, exploring international law's contribution toward resolving technology-facilitated violence. This paper will also interrogate whether it is appropriate and feasible to hold internet intermediaries accountable for failure to prevent, respond to and provide remedy for technology-facilitated violence against women committed on their platforms.

⁶ Šimonović Dubravka (Special Rapporteur on violence against women, its causes and consequences), *Report of the Special Rapporteur on violence against women, its causes and consequences*, U.N. Doc. A/HRC/32/42 (Apr. 19, 2016), available at <https://documents-dds-ny.un.org/doc/UNDOC/GEN/G16/080/53/PDF/G1608053.pdf?OpenElement>.

⁷ See *supra* note 1.

⁸ See *supra* note 2.

As internet intermediaries can only be held accountable if they have a positive obligation in this regard, the paper will explore whether it is possible to imbue internet intermediaries with a positive obligation to exercise due diligence in these instances. Issues such as anonymity and extra-territoriality complicates States' prosecuting or holding perpetrators or third party liable. Still, States are not exempt from discharging their obligations on the basis that the wrong is difficult to investigate or prosecute.

Part IV also interrogates the role of the State to exercise due diligence to prevent technology-facilitated violence, protect victims/survivors, investigate and prosecute incidences of technology-facilitated violence, punish perpetrators and provide redress to victims/survivors.

Part V interrogates what measures States have undertaken in addressing technology-facilitated violence and whether these actions, policies, laws and programmes are effective. While many states have attempted to criminalize technology-facilitated violence, its enforcement has proven seriously problematic due to lack of mechanisms, procedures and expertise/skills. As the violating material is posted on a third party platform, often sited beyond the territorial limits and jurisdiction of the state concerned, providing remedies and reparation to the victim/survivor has proven especially difficult. Takedown notices, removal of links and disclosure of identity can only be undertaken by third parties who may or may not be liable for the violating material posted on their platforms.

Finally, Part VI explores ways forward and outlines recommendations and principles to address technology-facilitated violence.

A framework for State and internet intermediaries' obligation to eliminate violence against women is annexed to the end of the paper.

PART I

Definition, gaps and challenges

This part looks at four issues. First, it discusses technology-facilitated violence against women and its manifestations as well as draws parallels to offline violence against women. It then looks at stigmatization of the victim/survivor. Not only are victims/survivors blamed for the violence committed against them, that the violence is not 'physical' tends to mean that State authorities and private sector actors, such as internet intermediaries, minimize its perceived gravity.

The third issue is how online VAW aggravates harm. Online violence is facilitated by instantaneous transmission through vast digital networks. Once uploaded, it may remain online permanently. Finally, this part discusses the issue of consent, which is central to identifying technology-facilitated violence against women as opposed to one's exercise of freedom of expression.

➤ [Online violence against women](#)

What constitutes violence against women has been defined in several international instruments including international and regional declarations, treaties, guidelines and recommendations. In line with the 1993 UN Declaration on the Elimination of Violence against Women, this paper defines 'violence against women' as an act of gender-based violence (GBV) that results in, or is likely to result in, physical, sexual, psychological or economic harm or suffering to women,

including threats of such acts, coercion or arbitrary deprivation of liberty, whether occurring in public or in private life.⁹

While the perpetration of technology-facilitated violence against women is somewhat new, which itself poses its own challenges, it shares its basis with other forms of violence against women. Although some forms of technology-facilitated violence against women require and deserve further exploration, at this juncture the paper will not attempt to exhaustively define technology-facilitated violence against women.

Suffice to say that technology-facilitated violence against women are acts 'committed, abetted or aggravated' in part or fully by the use of information and communication technology (ICT) acts of gender-based violence that are committed, abetted or aggravated, in part or fully, by the use of information and communication technologies¹⁰ and include, amongst others, cyber stalking, bullying, threats, blackmail and sexual harassment; assessing or uploading/disseminating intimate photos, videos or audio clips without consent; accessing or disseminating private data without consent; uploading/disseminating altered photos or videos and uploading them to dating, pornography or other kinds of websites; creating fake profiles and other forms of identity theft; mob attacks¹¹, grooming predation (of children in particular), doxing (searching and publicizing personal data of another) and exploitation of women and girls.

Online violence against women presents specific challenges. What is actionable and what is not, is crucial in gauging which data or images constitute violence. Actionable violence (including threats of violence) is gauged by intent to harm, content, credibility or imminence of harm and context.¹² In this paper, data and images that constitute actionable technology-facilitated violence against women are deemed violating material.

Where technology-facilitated violence against women does not involve physical violence, it tends to be trivialized, and thus receive inadequate and inappropriate responses from concerned actors, including the State, the private sector, civil society, and society at large, even women themselves. It is thus crucial to look at the responses of different actors, particularly, the identification and role of first responders (including the police, internet intermediaries and helplines), regulators and the judiciary to map the reality of women's initial experiences when accessing justice/remedies, as this colours the rest of the reporting process.

To some extent, these challenges are shared with other forms of violence against women which does not involve physical harm, such as conventional stalking and sexual harassment. Similar to technology-facilitated violence, harassment and stalking often involve repeated acts. While an individual incident could be lawful expression, repeated unwanted acts constitute unlawful harassment or stalking. It is worth noting that because of the ease with which things can be shared, liked, reposted, stored and downloaded, there is more scope for repetition and dissemination of content constituting technology-facilitated violence.

⁹ Violence against women has been defined and elaborated in many human rights and feminist instruments and discourse including CEDAW. The following forms of violence share similarities to technology-facilitated violence against women: intimate partner violence, domestic violence, sexual harassment, harassment based on gender, stalking and inciting others to commit violence against women.

¹⁰ Women's Legal and Human Rights Bureau, Inc. & Association for Progressive Communications, *From Impunity to Justice: Domestic legal remedies for cases of technology-related violence against women* (March 2015), available at http://www.genderit.org/sites/default/upload/flow_domestic_legal_remedies.pdf.

¹¹ For example, the online attack of Leslie Jones on twitter and the hacking of her iCloud and cell phone. Twitter later suspended one of the attackers. See Katie Rogers, *Leslie Jones, Star of 'Ghostbusters,' Becomes a Target of Online Trolls*, the New York Times (July 19, 2016), available at <http://www.nytimes.com/2016/07/20/movies/leslie-jones-star-of-ghostbusters-becomes-a-target-of-online-trolls.html>; Nicholas Mojica, *Leslie Jones Hacked: A Timeline of the 'Ghostbusters' Star's Twitter Hate and Online Attackers*, International Business Times (August 25, 2016), <http://www.ibtimes.com/leslie-jones-hacked-timeline-ghostbusters-stars-twitter-hate-online-attackers-2407046>.

¹² For discussion on what constitutes actionable acts see *Delfi AS v. Estonia*, App. No. 64569/09, Eur. Ct. H.R. (June 16, 2015).

Perpetrators of technology-facilitated violence against women often employ a continuum of violence against women, both offline and online. Like other forms of violence against women, perpetrators are often known to the survivors and include intimate partners and ex-partners.¹³

As with physical stalking, non-physical stalking can evolve into extreme physical violence. Stalking began receiving recognition after model-actress Rebecca Shaeffer was murdered in 1989 by an obsessed fan who had been stalking her.¹⁴ Since the Shaeffer case, stalking, including cyber stalking, has received somewhat more attention and legal response.¹⁵

Another alarming form of technology-facilitated violence is live streaming on offline acts of violence. Where the cyber stalking and online sexual harassment do not involve physical violence, in this instance, crimes, including gang rape, are committed in a physical offline space and streamed live by perpetrators.¹⁶ With social media, crime involvement and self-promotion are intertwined, resulting in a macabre 'crime performance' where perpetrators share pre-crime plans, live streaming of themselves in the act of committing crime and post-crime bragging. "The social media dynamic that drives offenders to post their crime performances has also influenced the treatment of crime victims, so that 'performance victimization' is also a new reality."¹⁷

Online violence shares similarities with other forms of crimes, quasi-crimes and torts such as defamation, extortion (blackmail) and non-consensual disclosure of private data, communications and images; hate speech; and child pornography. Incitement to harm is yet another possible actionable violation. Incitement comprises of both incitement against a group and incitement against an individual. Harm comprises both physical and psychological harm.

Thus, sending threatening or offensive material or sharing a persons' private data online, and bombarding someone with sexually demeaning emails all constitute violence against women. Furthermore, similar to offline sexual harassment, online harassment or bullying can constitute gross misconduct and grounds for dismissal of an employee, particularly if the employer already has policies on what conduct will be deemed unacceptable irrespective of whether such conduct occurs at the workplace or otherwise. In the Irish case of *Teggart v TeleTech UK Limited*, the Court affirmed the dismissal of an employee, finding, amongst others, that the cumulative impact of the obscene Facebook posts about a co-worker, the intention to create a humiliating work environment and the dissemination of the comments among fellow employees justified the dismissal as having been reasonable.¹⁸

13 APC research indicate that in approximately 40% of the cases of technology-facilitated violence, the perpetrator is known to the victim/survivor.**

14 Associated Press, *The celebrity murder that changed how stalkers are treated*, Page Six (July 14, 2014), <http://pagesix.com/2014/07/14/stars-safer-since-actress-1989-murder/>.

15 Subsequently, California enacted laws criminalizing stalking. Criminal stalking is defined in California as "someone who willfully, maliciously and repeatedly follows or harasses another victim and who makes a credible threat with the intent to place the victim or victim's immediate family in fear of their safety." Continuity of purpose must be established through more than one incident. However, where stalking itself is not a crime, for example in the UK, "offenders get shorter prison sentences that won't make any difference and they go back to stalking". In the UK, a national stalking clinic was opened in London. See Lucy Buckland, *World's first clinic to treat stalkers and prevent violent crimes opens*, DailyMail.com (Dec. 8, 2011), <http://www.dailymail.co.uk/news/article-2071219/Worlds-clinic-treat-STALKERS-prevent-violent-crime-opens.html#ixzz3fSnXU858>.

16 Reuters in Stockholm, *Three men arrested in Sweden after Facebook Live 'gang-rape'*, The Guardian (Jan. 23, 2017), available at <https://www.theguardian.com/world/2017/jan/23/three-men-arrested-sweden-facebook-live-gang-rape-uppsala>; Olivia Solon, *Why a rising numbers of criminals are using Facebook Live to film their acts*, The Guardian (Jan. 27, 2017), available at <https://www.theguardian.com/technology/2017/jan/27/rising-numbers-of-criminals-are-using-facebook-to-document-their-crimes>.

17 Raymond Surette, "Performance Crime and Justice" [2015] *CICrimJust* 21; (2015) 27(2), *Current Issues in Criminal Justice* 195, available at <http://www.austlii.edu.au/au/journals/CICrimJust/2015/21.html> (last visited June 15, 2017).

18 *Teggart v. TeleTech UK Limited*, [2012] NIIT 00704_111T (Mar. 15, 2012).

o *Consent*

Consent is key in differentiating lawful from unlawful and harmful behaviour. Consent in relation to technology-facilitated violence is often complicated by the exact act to which the consent, if any, relates. Because of this, defining consent is crucial in technology-facilitated violence and must be addressed in any mechanism dealing with technology-facilitated violence.

Consent is particularly important in gauging whether there has been violation of privacy with regards to dissemination of private data. Consent that is specific to an individual, like sharing of intimate photos, cannot be expanded to consent for the data to be shared and disseminated more widely.

Focusing on consent also recognizes that women have the right to sexual expression, in other words that there is nothing intrinsically unlawful or immoral about expressing oneself sexually through digital images. It is not the taking, but the spreading of these images, videos or other private data that is unlawful or immoral.

Furthermore, in the digitized world of big data, what is personal and what is public data is blurred. Our personal data is continuously being handled and commoditized by internet corporations.¹⁹ It is stored in servers that are liable to be hacked. Such personal data however, is no less personal even though it may be available in the public domain. This further emphasizes that consent for its dissemination is crucial in determining whether a violation of privacy has been committed.

o *Stigmatization of victims*

Patriarchy and prevailing interpretations of moral norms, culture and religion place women as the primary bearers of honour and tradition. Transgressions or deemed transgressions of culture by women are viewed as more reprehensible and dealt with by society more severely than those committed by men. This renders women more vulnerable and susceptible to 'moral' and 'cultural' attacks, particularly sexually nuanced attacks; and less likely to report gender-based violence.

Victims/ survivors themselves may believe that they transgressed social and cultural norms and are to be blamed for the violence committed against them. Women who establish cyber-friendships or relationships may be deemed to have transgressed culturally appropriate behaviour, as are women who engage in sexting or those who consent to intimate partners taking suggestive images, albeit for private purposes. While the relative anonymity available online allows women to transgress and challenge cultural norms, especially in relation to sexuality, the same anonymity combined with the speed, ease and reach of transmission provides an optimum platform for extortion. If the violence involves the uploading of suggestive or sexually explicit images and conversations either maliciously or without the victim's/survivor's consent, then the victim/survivor herself, more than the perpetrator, tends to bear the brunt of societal condemnation.

As a consequence, victims/survivors may be reluctant to seek assistance, silenced and isolated by shame. State actors' decision to prosecute, may be imbued with biases, and susceptible to negative socio-cultural perceptions that imply that victims/survivors provoked the violence

¹⁹ Personal data means any information relating to an identified or identifiable individual; an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number (e.g. social security number) or one or more factors specific to his or her physical, physiological, mental, economic, cultural or social identity (e.g. name and first name, date of birth, biometrics data, fingerprints and DNA).

through misbehaviour or transgression of socio-cultural norms.²⁰ This ultimately translates to a lack of support for the victims/survivors of technology-facilitated violence against women. As the internet and digital technology become increasingly integrated in our lives, robust policies are required to curb exposure to technology-facilitated violence.²¹

Outreach programmes can end isolation and remove stigma. The availability of a social network also increases women's autonomy and their ability to seek support and assistance. However, women's access to justice lies both within and beyond legal measures and within the interplay of politics, economic and culture, thus both legal and extra-legal (e.g. cultural) remedies are needed.

o *Ease of transmission and persistence*

ICTs provide amplify the transmission of digital material. ICTs allow for the *easy and rapid dissemination* of information and content, provide multiple platforms, and are comprised of vast networks.

Further, violent content, once disclosed or disseminated, is difficult to remove from these networks. It becomes persistent and remains accessible. The nature of the internet facilitates the transmission of the offending messages and images by others. This problem is made worse by the attitude of internet intermediaries. Platform providers have consistently denied requests from victims/survivors to remove harmful content, irrespective of whether the upload and dissemination of the content was done with the victim's/survivor's consent, whether the images were spliced or otherwise altered to appear as that of the victim/survivor or whether sexually explicit or suggestive content was falsely made to appear to originate from the victim/survivor.

This aggravates the harm to the exercise and enjoyment of human rights and freedoms, particularly the right to privacy or respect for private life caused by the communication of the violating material compared to other forms of more traditional media.²²

PART II Actors and stakeholders

This part discusses three actors and stakeholders involved in technology-facilitated violence. The person initiating the violence, namely the author, or the person who first uploads the offending data or images. This is the primary perpetrator. Secondly the person, who purposefully, recklessly or negligently downloads, forwards, or shares the offending data or images. Lastly, the internet intermediaries on whose platforms technology-facilitated violence is perpetrated.

➤ *Primary perpetrator*

As stated above, ICTs amplify both the anonymity and reach of transmission. The individual who generates the offending data or image is clearly the primary perpetrator. However, legal enforcement officers often lack the training, skill or resources to identify perpetrators who employ protocols to shield their identity, thus offering little or no protection for victims/survivors.

²⁰ See Aziz & Moussa, *supra* note 1, at 59.

²¹ *Id.*

²² Eur. Ct. H.R., Editorial Board of Pravoye Delo and Shtekel v. Ukraine, § 63, App. No. 33014/05, May 5, 2011.

The enhanced anonymity offered by digital and virtual spaces, through encryption and privacy protocols, provides particular challenges in identifying perpetrators of violence against women, including those who engage in harassment, stalking, incitement to harm and defamation. In turn, this magnifies impunity.

The inability of law enforcement and intelligence services to uncloak anonymity or decipher encrypted communications to investigate crimes has raised "legitimate concerns about how bullies and criminals use new technologies to facilitate harassment."²³ Over-regulation on the other hand, can lead to online censorship, mass and targeted surveillance and data collection, digital attacks on civil society and repression. Restrictions to encryption and anonymity tools put the privacy of all internet users at risk.²⁴

The internet thus offers unprecedented capacity for criminals, pranksters, governments and corporations to interfere with the rights to freedom of opinion and expression. To some extent, encryption, anonymity and the concept of security behind them is essential in the face of political censorship as it creates a zone of privacy to protect opinion and belief.²⁵ The internet, having become a "central global public forum", deserves protection. Further, "(S)uch security may be essential for the exercise of other rights, including economic rights, privacy, due process, freedom of peaceful assembly and association, and the right to life and bodily integrity."²⁶

The anonymity provided by these protocols is beneficial to whistle-blowers and human rights defenders, those who oppose current dominant groups or those who are under historical social/cultural/political surveillance because of their identity including black/indigenous/migrant/women, sex workers, queer people, young women and those identifying as LGBTQIA. Anonymity also offers privacy for victims/survivors (whose privacy is often violated by perpetrators and allows them to re-enter online spaces (under pseudonyms, for example). It is simplistic therefore to view anonymity as a threat that needs to be removed under all circumstances.

It is thus critical to formulate principles and guidelines that allow the internet to continue to be the central global public forum that defends the right to privacy and is free from government censorship on the one hand, yet ensure that it is not used as an instrument to commit violations of women's human rights. With warrants and technical skills, the perpetrators can sometimes be identified, especially if the perpetrator is known to the victim/survivor which allows investigators to trace the links to the perpetrator.

➤ Secondary perpetrators

Given the ease and speed of transmission, eliminating technology-facilitated violence against women includes not only addressing and eliminating the primary violation (by the principal perpetrator) but also the dissemination, whether witting or unwitting, by others (secondary perpetrators). Once posted, the offensive material may generally be accessed by others who may download the material, share it by reposting or by creating a link to the material. These others may then take action to discriminate or commit hostile or violent acts against the

23 David Kaye (Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression), U.N. Doc. A/HRC/29/32 (May 22, 2015), available at <https://documents-dds-ny.un.org/doc/UNDOC/GEN/G15/095/85/PDF/G1509585.pdf?OpenElement>.

24 *Id.*

25 *Id.*

26 See Report on encryption, anonymity, and the human rights framework, United Nations Human Rights Office of the High Commissioner, <http://www.ohchr.org/EN/Issues/FreedomOpinion/Pages/CallForSubmission.aspx> (last visited June 16, 2017).

victim/survivor, for example by directly communicating with the victim/survivor or related persons.

Even when primary perpetrators are held liable, little attention and effort is made to hold these secondary perpetrators, who re-transmit offending data and images, liable. Data and images that are tweeted and re-tweeted, downloaded and forwarded, liked and shared may involve a great number of individuals and pose an overwhelming challenge to regulators. Further reflection is needed on how to hold re-transmitters responsible for the transmission of violating materials.

Intent, or more specifically, lack of intent, can be an issue with secondary perpetrators. Still, holding persons accountable despite lack of intent is not without basis under the law. In many jurisdictions, criminal law has developed the concept of reckless indifference where intent cannot be established. For example, a person who drives his vehicle into a restaurant, is liable for the injuries and death caused thereby even though he may not have intended to injure or cause death as he is recklessly indifferent as to whether there are persons in the building who would be injured or killed by his actions.

In the civil (non-criminal) realm, negligence is an established element of some tortious act that does not require intent to be established. Another example is the established liability of persons repeating slanderous or defamatory statements. Generally, a person who repeats slanderous or defamatory information is also liable. Under certain circumstances, this liability is irrespective of whether that person is aware that the statement is defamatory, as dissemination does not render an act less offensive or less harmful.²⁷

If technology-facilitated violence against women follows these paradigms, secondary perpetrators can be made liable for their action in re-transmitting the offending data or images. At the very least, they can be seen as aiders or abettors of a wrongful act although they may not personally know the perpetrator or victim/survivor. After all, ignorance of the identity of the victim/survivor does not make the violence victimless or the harm unforeseeable. It is reasonable to expect that, at the very least, the protection afforded to victims/survivors of offline violence should be made available for technology-facilitated violence.

➤ Internet intermediaries

The internet plays an important role in enhancing access to and facilitating the dissemination of information. It is important that freedom of expression and freedom of information is protected online.²⁸ Internet intermediaries bring together or facilitate transactions between third parties on the internet. They give access to, host, transmit and index content, products and services originated by third parties. This can take place on the internet or by providing internet-based services to third parties.²⁹

Because of the internet's capacity to store and communicate staggering amounts of information, internet intermediaries are placed in a unique position.³⁰ Many policy and lawmakers regard protecting internet intermediaries from liability as a pre-requisite to protecting

²⁷ "A false statement is not less libelous because it is the repetition of rumor or gossip or of statements or allegations that others have made concerning the matter." *Ray v. Citizen-News Co.* (1936) 14 Cal.App.2d 6, 8-9.

²⁸ See Human Rights Council resolutions [20/8](#), [26/13](#) and [32/13](#) on "The promotion, protection and enjoyment of human rights on the Internet", which affirm that the same rights that people have offline must also be protected online.

²⁹ Organisation for Economic Co-operation and Development (OECD), *The Economic and Social Role of Internet Intermediaries* (Apr. 2010), p.9, available at <http://www.oecd.org/internet/ieconomy/44949023.pdf>.

³⁰ European Court of Human Rights, *Ahmet Yildirim v. Turkey*, App. No. 3111/10 (2012), § 48, and *Times Newspapers Ltd*, § 27.

the digital economy to encourage the innovation and creativity that has led to the rapid and successful development of the internet.³¹ However, others try to enforce barriers to expression and innovation through disproportionate or heavy handed liability such as unduly requiring intermediaries to monitor content and data being hosted or transmitted online. This hinders the right to freedom of expression as recognized at the international level.³²

Internet intermediaries are not monolithic. While some merely host or transmit data, like cloud services or small hosting companies, others are increasingly taking on an 'active role' mediating content. This can be by performing different and competing roles simultaneously, providing both hosting services and other categories of services. Shielding internet intermediaries from liability is more straightforward when their roles are limited to merely transmitting, hosting and conveying third party information; their defense being generally referred to as 'hosting defense'. These expanded roles however, challenge the very bases for the 'hosting defense'.³³

Violating materials may not be posted by internet intermediaries nor do these corporations have possession of private data and images which are disclosed and disseminated. Nevertheless, the intermediaries have a responsibility to put in place preventive measures and respond to violating materials, especially when they have the capacity to moderate content and have in place measures to flag and report 'user generated' content.³⁴

Thus, free speech as we understand it and as mediated by these corporations is increasingly becoming nebulous and dependent on the 'protective' measures put in place by the intermediaries themselves. As technology-facilitated violence happens not merely on the first upload by the primary perpetrator, but is repeated every time it is liked and shared, re-tweeted, searched and downloaded or forwarded, internet intermediaries are uniquely situated to stop the recurrence of the violence and provide the necessary relief and remedy needed by victims/survivors.

Freedom of expression requires the free flow and exchange of ideas and knowledge; but for profit-driven internet intermediary corporations, maintaining the free flow and exchange of ideas and knowledge may be more profitable than eliminating violence against women. Profit plays a significant role in deciding where intermediaries lean when tensions arise between the right of women to a safe internet environment and the interest of internet intermediaries to guarantee their users' freedom of expression and access to information. There are precedents where the courts have been "mindful of the risk of harm posed by content and communications on the internet" and demanded greater vigilance from internet intermediaries.³⁵

It is also more cost effective to seek redress from internet intermediaries than all the re-transmitters (which in fact may not even be logistically possible). For these reasons, intermediaries are best placed to bring technology-facilitated violence activities to an end and

31 Farano, referring to see Communications Decency Act (CDA)8 and the Digital Millennium Copyright Act (DMCA)9 were thus passed respectively in 1996 and in 1998, while the Electronic Commerce Directive (e-commerce Directive)10 in Europe was adopted in 2000.

32 *The Manila Principles on Intermediary Liability Background Paper*, Electronic Frontier Foundation (May 30, 2015), available at https://www.eff.org/files/2015/07/08/manila_principles_background_paper.pdf.

33 For an elaboration of the "active role" standard used see Karine Perset, *The Economic and Social Role of Internet Intermediaries*, OECD (Apr. 2010), available at <https://www.oecd.org/internet/ieconomy/44949023.pdf>.

34 Compare this to the more traditional media such as newspapers. Statements carried in newspapers are vetted and edited, as necessary. Thus the level of control over newspapers is much higher than the control exerted by internet and digital platform providers.

35 See *Delfi*, § 157. "While acknowledging the "important role" played by the Internet "in enhancing the public's access to news and facilitating the dissemination of information in general". Although *Delfi* did not involve violence against women, this dicta is persuasive and is applicable to technology-facilitated violence. See also Ahmet Yildirim, cited above, § 48, and *Times Newspapers Ltd*, cited above, § 27. The Court reiterates that it is also mindful of the risk of harm posed by content and communications on the Internet (see *Editorial Board of Pravoye Delo and Shtekel*, cited above, § 63; see also *Mosley*, cited above, § 130)".

their pro-active response and co-operation is necessary to eliminate technology-facilitated violence against women.

PART IV

Application of international law

This part looks at international law and issues of accountability. It explores State responsibility to eliminate technology-facilitated violence which includes States exercising due diligence to prevent technology-facilitated violence, protect victims/survivors, prosecute perpetrators and provide redress and reparation for victims/survivors.

Separately, this part also explores the obligations and duties of internet intermediaries in international law (as opposed to domestic/national laws formulated by States to regulate intermediaries). It looks at the evolution of investing human rights responsibilities and obligations on transnational companies and suggests how these can be complied with.

A. Human Rights and the State

Human rights are universal, inalienable, inter-related, inter-dependent, and indivisible. International human rights law protects the right to dignity and equality, prohibiting gender-based discrimination and gender-based violence.³⁶ International law also protects freedom of expression.

The exercise of these rights under international human rights law is not absolute and may be subject to certain restrictions, but these shall only be such as are provided by law and are necessary for the respect of the rights or reputations of others or for the protection of national security or of public order (*ordre public*), or of public health or morals, and are proportionate to the aim they seek to address.³⁷ The application of these restrictions by States however, “may not put in jeopardy the right itself”.³⁸

Thus an individual's human rights are not absolute in that they cannot be enjoyed at the expense of the human rights of others. Others, in this instance, relates to other persons individually or as members of a community.

“Freedom of speech, especially when it concerns expression on the internet, is the absolute foundation of our societal discourse, nonetheless freedom of speech naturally ends where threats abound. It is not freedom of expression to consciously intimidate people on Facebook and Twitter, especially women, insult them, express the wish to rape them or to threaten physical harm. One has to act on this even across borders ...”³⁹

36 G.A. Res. 217 (III) A, Universal Declaration of Human Rights (Dec. 10, 1948), art. 2; Convention on the Elimination of All Forms of Discrimination Against Women, Dec. 18, 1979, 1249 U.N.T.S. 13.

37 G.A. Res. 2200A (XXI), International Covenant on Civil and Political Rights (Dec. 16, 1966), art. 3.

38 UN Human Rights Committee, *General Comment No. 34: International Covenant on Civil and Political Rights, Article 19: Freedoms of opinion and expression*, Sept. 12, 2011, CCPR/C/GC/34, para. 21.

39 Terry Reintke, *Violence against women online: “Freedom of speech ends where threats abound,”* European Parliament News, (Apr. 27, 2016) available at <http://www.europarl.europa.eu/news/en/news-room/20160425STO24559/Violence-against-women-online-Freedom-of-speech-ends-where-threats-abound>; See also Terry Reintke, (Rapporteur for the Committee on Women's Rights and Gender Equality), *Report on gender equality and empowering women in the digital age, 2015/2007(INI)*, European Parliament, AB-80048/2016, (Apr. 8,

This is different from freedom of opinion. The right to hold opinions without interference is an absolute right and “permits no exception or restriction”.⁴⁰ However, the expression of an opinion, that is the right to freedom of expression bears “special duties and responsibilities”.

The Sustainable Development Goals recognize that “gender equality is not only a fundamental human right, but a necessary foundation for a peaceful, prosperous and sustainable world. Providing women and girls with equal access to education, health care, decent work, and representation in political and economic decision-making processes will fuel sustainable economies and benefit societies and humanity at large”.⁴¹ Violence against women, offline and online, must be acknowledged as a manifestation of systemic marginalization of women throughout society. Enhancing “the use of enabling technology, in particular information and communications technology,⁴² to promote the empowerment of women” requires the elimination of technology-facilitated violence against women.

Article 20 of the *International Covenant on Civil and Political Rights* prohibits any advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence.⁴³ Advocacy of gender-based hatred that constitutes incitement to discrimination, hostility or violence should similarly be regarded as a violation of human rights. Effective measures to limit the dissemination of hate speech and speech inciting discrimination, hostility or violence can by no means be equated to “private censorship”.⁴⁴ Although the *Rabat Plan of Action* prohibits advocacy of national, racial or religious hatred (and not gender-based hate speech), it is still useful at this juncture to refer to the three types of expression mentioned in the Plan as constituting hate speech, namely expression: (i) that constitutes a criminal offence; (ii) that is not criminally punishable but may justify a civil suit or administrative sanctions; (iii) that does not give rise to criminal, civil or administrative sanctions but still raises a concern in terms of tolerance, civility and respect for the rights of others.⁴⁵

The European Union has also entered into agreements with prominent internet intermediaries, such as Facebook, Twitter and YouTube, to prevent the spread of illegal hate speech online, to educate and raise awareness with their users about illegal hate speech, to develop internal “procedures and staff training to guarantee that they review the majority of valid notifications for removal of illegal hate speech in less than 24 hours and remove or disable access to such content, if necessary”.⁴⁶ Internet intermediaries also announced that they would “continue to work with the EU to identify and discredit extremist speech by promoting so-called “counter-

2016), available at <http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//NONSGML+REPORT+A8-2016-0048+0+DOC+PDF+V0//EN>.

40 See *supra* note 38, at para. 9

41 Goal 5: Achieve gender equality and empower all women and girls, Sustainable Development Goals: 17 Goals to Transform Our World, <http://www.un.org/sustainabledevelopment/sustainable-development-goals/> (last visited June 16, 2017).

42 “Goal 5 targets,” Goal 5: Achieve gender equality and empower all women and girls, Sustainable Development Goals: 17 Goals to Transform Our World, <http://www.un.org/sustainabledevelopment/gender-equality/> (last visited 16 June 2017).

43 G.A. Res. 2200A (XXI), *International Covenant on Civil and Political Rights* (Dec. 16, 1966); See also *Rabat Plan of Action on the prohibition of advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence* which similarly only prohibits advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence. Gender-based hatred should be similarly prohibited.

See “Rabat Plan of Action,” available at http://www.ohchr.org/Documents/Issues/Opinion/SeminarRabat/Rabat_draft_outcome.pdf (last visited May 23, 2017).

44 *Id.*

45 Report of the United Nations High Commissioner for Human Rights on the expert workshops on the prohibition of the incitement to national, racial or religious hatred, U.N. Doc. A/HRC/22/17/Add.4, Jan. 11, 2013, available at http://www.ohchr.org/Documents/Issues/Opinion/SeminarRabat/Rabat_draft_outcome.pdf (last visited May 23, 2017).

46 European Commission Press Release: *European Commission and IT Companies announce Code of Conduct on illegal online hate speech* (May 31, 2016), http://europa.eu/rapid/press-release_IP-16-1937_en.htm.

narratives" and supporting educational programs that encourage critical thinking.⁴⁷ The focus of this initiative however is racism, xenophobia and the radicalization of young people and racist use of platforms to spread violence and hatred.⁴⁸ This 'code of conduct' was however heavily criticized for undermining legal speech, circumventing the rule of law and for the absence of independent oversight.⁴⁹

However, several domestic laws similarly prohibit a narrow class of hate crimes, namely on the basis of race, religion, or national origin but not gender, gender identity, sexual orientation or disability.⁵⁰ Further, there are some States that recognize hate speech on the basis of gender or sex, e.g. Canada, Croatia, Netherlands and South Africa. In order for international and regional initiatives on hate speech to apply to gender based technology-facilitated violence against women, gender must be included as a category of hate speech that is illegal.

Hate speech however, must be narrowly defined. For hate speech to be criminalized, it must be of a public nature, at the very minimum present a real and imminent danger, and contain the obvious intention to harm.⁵¹

Lastly, privacy is another protected human right entrenched in, among others, the Universal Declaration of Human Rights.⁵² Invasion of privacy can be established when an individual, in possession of private information, makes a public disclosure of such information without consent.⁵³

B. Human rights and internet intermediaries

➤ *State obligation to ensure compliance by business enterprises*

Eliminating technology-facilitated violence requires the intercession of internet intermediaries, including transnational corporations serving the role of internet intermediaries. In 2005, the United Nations Secretary General appointed John Ruggie as his Special Representative on Human Rights and Transnational and Other Business. In 2011, Ruggie released a set of *Guiding Principles on Business and Human Rights on Implementing the United Nations "Protect, Respect and Remedy" Framework*. The principles provide that "Business enterprises should respect human rights. This means that they should avoid infringing on the human rights of others and should address adverse human rights impacts with which they are involved."

Ruggie called on States to set out clearly the expectation that all business enterprises domiciled in their territory and/or jurisdiction respect human rights throughout their operations:⁵⁴

47 Natalia Drozdiak, *U.S. Tech Firms Agree to EU Code of Conduct on Terror and Hate Content*, Wall Street Journal (May 31, 2016), available at <http://www.wsj.com/articles/u-s-tech-companies-sign-up-to-eu-code-of-conduct-on-terror-1464689959>.

48 *Id.*

49 European Digital Rights (EDRI), *EDRI and Access Now withdraw from the EU Commission IT Forum discussions* (May 31, 2016), <https://edri.org/edri-access-now-withdraw-eu-commission-forum-discussions/>.

50 This was the case in the US until the passing of The Matthew Shepard and James Byrd, Jr Act.

51 *Promotion and protection of the right to freedom of opinion and expression*, Note by the Secretary-General, U.N. Doc. A/67/357 (Sept. 7, 2012).

52 Article 12 states, "No one shall be subjected to arbitrary interference with his privacy, family, home or correspondence, nor to attacks upon his honour and reputation. Everyone has the right to the protection of the law against such interference or attacks."

53 In some cases, for the action to succeed, the public disclosure of the facts in question must be highly offensive to a reasonable person of ordinary sensibilities. In other case, such as test is not applicable for example, where the data consist of a person's phone number, address or bank account details.

54 John Ruggie, *Guiding Principles on Business and Human Rights: Implementing the United Nations "Respect, Protect and Remedy Framework"*, (UN Human Rights Office of the High Commissioner, 2011).

“There are strong policy reasons for home States to set out clearly the expectation that businesses respect human rights abroad, especially where the State itself is involved in or supports those businesses. The reasons include ensuring predictability for business enterprises by providing coherent and consistent messages, and preserving the State’s own reputation.”⁵⁵

To this end, States should “provide effective guidance to business enterprises on how to respect human rights throughout their operations” and encourage or require business enterprises to address their human rights impacts.

➤ *Transnational corporations’ international human rights responsibilities independently of State obligations*

While international human rights law principally focuses on States as subjects of international law, there have been attempts to recognize corporations, especially transnational corporation as having been imbued with international personality and thus recognized as subjects of international law. This has occurred both at the behest of transnational corporations that seek to operate in the international realm and to access international law, as well as at the behest of States that respond by attempting to regulate transnational corporations’ activities and imbue them with responsibilities similar to those vested in states.⁵⁶

When can transnational companies be held to be subjects of international law? Courts and international human rights instruments have traditionally been focused on limiting the power of public (State) and not private actors. Courts adjudicating human rights matters generally preclude cases being brought against non-State defendants/respondents. Likewise constitutional guarantees on fundamental liberties and rights are generally enforceable only against the State.

While States have a vested interest in maintaining their power and monopoly in international law by not acknowledging transnational corporations as subjects of international law, transnational corporations wield tremendous influence. Their burgeoning budgets rival the largest of States and they have access to tremendous resources which directly influence, if not directly participate in, the international law-making process.⁵⁷

Current realities compel more and more scholars and practitioners alike to consider transnational corporations as having acquired a limited personality in international law.⁵⁸ A

The Guiding Principles were proposed to the United Nations Human Rights Council as part of the 2011 report to the Council by then-UN Special Representative on business & human rights, John Ruggie: *Report of the Special Representative of the Secretary-General on the Issue of Human Rights and Transnational Corporations and Other Business Enterprises*, John Ruggie, UN Doc. A/HRC/17/31, Mar. 21, 2011, available at http://www.ohchr.org/Documents/Publications/GuidingPrinciplesBusinessHR_EN.pdf. See also Guidelines for Multinational Enterprises of the Organisation for Economic Co-operation and Development.

⁵⁵ Ruggie, p. 4

⁵⁶ For example, OECD Guidelines, ILO Tripartite Declaration and UN Global Compact.

⁵⁷ The International Court however adopts a more mundane definition based on the capacity to have rights and obligations under international law and the capacity to bring international claims. See ICJ, *Reparation for Injuries Suffered in the Service of the United Nations*, Advisory Opinion, ICJ Reports 1949, 174. This case pertains to the international legal personality of the United Nations.

⁵⁸ Transnational corporations are increasingly parties to internationalized contracts which specifically states that these contracts are to be governed by international law thus conferring transnational corporations has specific international capacities, as well as international treaties, particularly those related to investments. See *Texaco Calasiatic v. Libyan Arab Republic (Merits)*, in *International Legal Materials*, 17, 1978, 1–37, at 17, para. 47. States too attempt to regulate the behavior of transnational corporations. See for example OECD Guidelines for Multinational Enterprises, OECD (2008), available at <http://www.oecd.org/dataoecd/56/36/1922428.pdf>. : UN Doc. E/CN.4/RES/2005/, 2005 (last visited June 16, 2017).

concomitant of international legal personality is the responsibility to respect human rights that exists over and above compliance with national laws and regulations and independently of States' human rights obligations internationally.

Andrew Clapham summarizes the arguments for imbuing non-State actors with human rights obligations to reversing the notion that human rights are the product of the social contract between the State and the individual. This, he argues, allows us to presume that human rights are entitlements enjoyed by everyone to be respected by everyone. The net result being States, corporations and individuals all have human rights obligations. The obligations exist irrespective of whether they are enforceable.⁵⁹

Individuals have been held personally liable for a narrow range of international crimes under humanitarian law that are by no means coextensive with the field of human rights.⁶⁰ Holding corporations, particularly transnational corporations, accountable has, however, been subjected to more intense debates, although some headway has been made to invest corporations with the responsibility to promote, protect and fulfill human rights.⁶¹

Transnational corporations' obligations to respect and protect human rights under international law are being developed, with passionate arguments from advocates on both sides. Thus, it is opportune for us to demonstrate why it is critical to hold internet intermediaries accountable for taking, or failing to take, reasonable steps to eliminate technology-facilitated violence against women on their platforms and to develop a framework and guiding principles for internet intermediaries' obligation to promote, respect and fulfill human rights in relation to eliminating technology-facilitated violence against women. After all, internet intermediaries can better be held accountable if they are vested with a positive duty to promote, protect and fulfill human rights.

This duty however, is not equal to the duty borne by States but merely pertains to the violation of human rights occurring on the respective platforms of the intermediaries. Intermediaries for example do not have the obligation to prevent violence wherever it may occur, but only violence occurring on their platforms.⁶² This may also better accord with the Ruggie principles of not "infringing on the human rights of others" and "addressing adverse human rights impacts with which they are involved".

C. Accountability and the Due Diligence Principle

➤ Due Diligence Principle

The State has an obligation to promote, protect and fulfill human rights. This includes the obligation to prevent violations, protect victims/survivors of human rights abuses, prosecute

⁵⁹ Andrew Clapham, *Human Rights Obligations of Non-State Actors*, (Academy of European Law, European University Institute, Oxford University Press, 2006). Also available at http://graduateinstitute.ch/files/live/sites/iheid/files/sites/international_law/shared/international_law/Prof_Clapham_website/docs/HR%20obligations%20of%20non-State%20actors.pdf (last visited May 6, 2017).

⁶⁰ For example, for war crimes, such as genocide.

⁶¹ The Guiding Principles were proposed to the United Nations Human Rights Council as part of the 2011 report to the Council by then-UN Special Representative on business & human rights, John Ruggie: Report of the Special Representative of the Secretary-General on the Issue of Human Rights and Transnational Corporations and Other Business Enterprises, John Ruggie, UN Doc. A/HRC/17/31, Mar. 21, 2011.

⁶² A comparison can be drawn from imagining a person drowning. Generally, an individual does not owe a duty, even if he is an excellent swimmer, to attempt to save a drowning person. However, the pool owner who obtains economic benefit from the use of the pool by others, owes a duty to ensure that there are sufficient safeguards to prevent death or drowning in his pool.

violations, punish perpetrators and provide redress and reparation for victims/survivors.⁶³ This further includes the obligation to remove impunity and provide for certainty of punishment of perpetrators of technology-facilitated violence against women.⁶⁴ This does not mean that States are per se accountable for acts of non-State actors. All non-State actors are subject to domestic laws and regulations. Non-State includes transnational and national corporations operating within the jurisdiction of the State.

The due diligence principle obligates States to take reasonable measures to prevent violence before it occurs, such as adopting relevant laws and policies, and effectively prosecuting and punishing perpetrators once they occur as well as providing redress and reparation to victims/survivors. Failure to exercise due diligence in taking these measures would render a State accountable.

This principle holds States accountable for violence committed not only by the state or State actors, but also by non-state actors.⁶⁵ Though this principle evolved to focus principally on State obligations, the principle is also useful in guiding internet intermediaries in developing and implementing policies to end violence against women on their platforms.

Such measures should be based on data and meaningful consultation with women's human rights advocates and once developed should be made accessible to women victims/survivors and subjected to continual monitoring and evaluation.

The due diligence principle is further fleshed out by the Due Diligence Project in the areas of prevention, protection, prosecution, punishment and provision of redress (5P's).⁶⁶ These P's are interlinked with overlapping issues.

- *Prevention (P1)*

Prevention includes measures to thwart the occurrence of violence against women. Good prevention programmes provide awareness of technology-facilitated violence against women and of information services and legal protection available following the incident. States have the duty to eliminate discrimination against women in accessing ICTs and promote women's participation and enjoyment of the benefits afforded by ICTs. In this respect, states should develop policies and programmes to educate the public about the issues and develop laws to address technology-facilitated violence against women. They should work to develop a counter-narrative to hate speech based on gender. These counter-narratives should not only address hate crimes but also lawful hate speech based on gender.⁶⁷

States and internet intermediaries should deem technology-facilitated violence not merely as another form of violence but violence that is grounded in discrimination and that prevents women from exercising their freedom of expression and that bars their access to technology and internet spaces. Policies and regulations can be developed for internet intermediary corporations to take preventive measures such as including warnings and reminders against technology-facilitated violence against women and against transmitting content that constitutes technology-facilitated violence.

⁶³ See *supra* note 1.

⁶⁴ *Id.*

⁶⁵ Traditionally, States have only been responsible for their own actions or those of their agents. Gradually, public international law developed to mandate States to exercise due diligence to promote, protect and fulfill human rights.

⁶⁶ See *supra* note 1.

⁶⁷ *The Equality and Anti-Discrimination Ombud's Report: Hate Speech and Hate Crime* (2015), available at http://www.genderit.org/sites/default/upload/hate_speech_and_hate_crime_v3_lr.pdf.

States and internet intermediaries too have the responsibility, independently of states, to develop and publicize policies on technology-facilitated violence and adopt reasonable preventive measures to prevent their platforms from being used to perpetrate technology-facilitated violence.

- o *Protection (P2)*

Protection focuses on avoiding the recurrence of further violence (which should be immediate if the perpetrator can be identified), the provision of accessible services, and adequate training and sensitization of first responders.

States and intermediaries need to implement effective measures to stop the recurrence (and often, escalation) of technology-facilitated violence. For technology-facilitated violence, the violence recurs every time violating materials are accessed, downloaded and shared, so protection of victim/survivors requires the proactive action and cooperation of internet intermediaries.⁶⁸ Thus, the obligation to protect does not only refer to the treatment of the original material, but the uploading and dissemination of that material which constitutes recurrence of the violence. While the protocol to identify, tag and stop specific files has already been developed and employed in some instances of gender-based violence, particularly those involving children, due consideration should be given on how and when this protocol should be used for other forms of violence against women and girls.⁶⁹

Fear of repercussions by perpetrators is the main reason women give for not seeking redress to stop violence.⁷⁰ It is important to note that technology-facilitated violence often accompanies, precedes or escalates into offline violence and protection should therefore include the same protection given to victims/survivors of offline violence, such as the provision of shelters and issuing restraining orders.

- o *Prosecution (P3)*

Prosecution refers to the investigation and institution of proceedings against the perpetrators. Where internet intermediaries are concerned, such proceedings may consist of inquiries. While all violence against women is subject to attitudes of marginalization and victim-blaming, this is more prevalent in cases of technology-facilitated violence, due to the victims/survivors' inability to demonstrate physical harm. Delay is then caused not only by the lower priority accorded to technology-facilitated violence but also by the lack of skills, knowledge and training in investigating technology-facilitated violence. In addition, jurisdictional issues can make it difficult to identify the appropriate law enforcement agency.

All these exacerbate the victims/survivors' often already low confidence in the police. The Due Diligence Project survey found that civil society organization respondents often reserved their worst ratings for the police, particularly in deprioritizing women's safety and security over other concerns. Negative attitudes lead to underreporting, particularly in societies that have a culture of silence surrounding violence against women. The excessive time taken to file charges, delays in the investigations, and the number of years that passed before a case was properly considered were all factors that made women victims/survivors desist from "wasting their time" by filing a complaint.

⁶⁸ Carly Nyst, *End violence: Women's rights and safety online: Technology-related violence against women: Recent legislative trends*, Association for Progressive Communications (APC) (May 2014), p. 4-8, available at http://www.genderit.org/sites/default/upload/flowresearch_cnyst_legtrend_in.pdf.

⁶⁹ Sean Gallagher, *Updated: How Verizon found child pornography in its cloud*, ARS Technica (Mar. 3, 2013), available at <https://arstechnica.com/information-technology/2013/03/how-verizon-found-a-child-pornographer-in-its-cloud/>.

⁷⁰ DDP survey's findings.

The state is obligated to train legal enforcement officers on technology-facilitated violence and establish affirmative duties to investigate and prosecute; to foster confidence in the police and judiciary; establish specialized prosecutors and courts; and develop a multi-sectoral and multi-agency approach.

- *Punishment (P4)*

Punishment refers to the obligation to impose sanctions on perpetrators. The certainty of adequate punishment creates a level of predictability and sends a message that technology-facilitated violence against women will not be tolerated. Punishment should also be capable of preventing recidivism, rehabilitating the perpetrators and deterring others from engaging in violence.

The punishment for technology-facilitated violence is generally lighter than 'physical' offline violence. States should demonstrate a strong political will to eliminate technology-facilitated violence and exercise innovation in formulating appropriate punishment which acknowledges the harm of technology-facilitated violence, not only to the individual victim/survivor but to other women and girls who may be intimidated or influenced by it. This includes the harm of denying women and girls freedom to participate in online spaces as a consequence of technology-facilitated violence against women.

- *Provision of redress and reparation (P5)*

The State is also responsible for providing adequate redress and reparations for victims/survivors. Generally, reparations and restitution to victims of violence include compensation for the costs of quantifiable losses (cost of medical care, loss of wages, and damage to property), injuries and non-quantifiable losses and for the needs of the victims/survivors of violence to re-build their lives in the short, medium and long terms, as they transition from a violent situation to a life free from violence are granted as civil remedies. For technology-facilitated violence, remedies must include the rights of victims/survivors to restitution, where possible.

Victims/survivors of violence against women require that such violence be stopped. Due to the repetitive nature of online gender-based violence (violence is repeated every time a person shares, re-tweets, forwards and downloads the violent content), an injunction against the perpetrator alone will not ensure that the violence stops. Delinking searches⁷¹ from and removal (see EU initiative above) of such content are some of the remedies already provided for other forms of illegal content. Decisions to delink or remove violent content however, must be decided through a transparent process. Such decisions must also be subject to review by relevant independent and impartial judicial tribunals.

PART V

State and intermediaries practices to address technology-facilitated violence

- *The State*

⁷¹ European Court of Justice, *Google Spain SL and Google Inc. v. Agencia Española de Protección de Datos (AEPD) and Mario Costeja González*, Eur. Ct. of Justice (May 13, 2014, available at <http://curia.europa.eu/juris/document/document.jsf?jsessionid=9ea7d0f130d541f8e70d076149b29aa5b05819c20f1e.e34KaxiLc3eQc40LaxqMbN4Pa3eRe0?text=&docid=152065&pageIndex=0&doclang=EN&mode=lst&dir=&occ=first&part=1&cid=47107>).

While many States have attempted to address or even criminalize technology-facilitated violence, its enforcement has proven seriously problematic due to a lack of mechanisms, procedures and expertise/skills. While some countries have specific laws on technology-facilitated violence against women, others rely on a combination of offences in the existing criminal and civil regimes. Offences within the present legal regime include stalking, sexual harassment, defamation, invasion of privacy, hate speech, breach of intellectual property rights, threats, identity and data theft and impersonation.

Without specific legislation, some have sought legal workarounds to have images taken down - most commonly the use of copyright law. However, where a victim/survivor opts to access the intellectual property regime, it is not unknown for victims/survivors to be required to prove that the images that were uploaded pertain (belongs) to her person by transmitting a naked photo of herself to the authorities. Furthermore, copyright seeks to protect the proprietary interest in an intellectual endeavor such as artwork or written work while in technology-facilitated violence against women, the perpetrator should be held accountable for the violation of the victim's/survivor's human rights, dignity and privacy rather than any proprietary interest in the image or conversation as an artwork or written work.

Research findings underline the urgent need for States to address the remedies available to victims/survivors.⁷² Like many repetitive forms of violence against women, victims/survivors require cessation of the violence and immediate protection from repercussions either in the form of retribution from the perpetrator or his family/friends or in the form of the victim/survivors being blamed and stigmatized. Yet, technology-facilitated violence poses new challenges in this regard.

o *Extra-territoriality*

States attempting to hold perpetrators, re-transmitters and internet intermediaries accountable are faced with a major complication, namely that some of these individuals and entities may be beyond the reach of a State's jurisdiction. Only in rare cases do States assert territorial jurisdiction over matters occurring outside their physical boundaries. Yet, the global nature of the internet has added an urgent need to re-examine the meaning of extra-territoriality.

In comparative law, a principle exists that even if the act in question originated from outside the physical jurisdiction of the State, the State may assert jurisdiction if the harm arose within the State. For example, if you discharge a gun from one side of a national border, and the bullet crosses the border and kills a person on the other side of the border, which State has jurisdiction? Arguably, the State where the harm occurred has jurisdiction to prosecute the perpetrator, if and when the perpetrator enters the State; or where both States have reciprocal arrangements.

As a result of the global nature of the internet, many courts have commenced asserting jurisdiction even when the intermediary is located not within their jurisdictions. Using this principle, the French Courts have, for example, asserted jurisdiction over a California based company because disputed goods were accessible to French public, namely, the website "targeted" the relevant public in their jurisdiction.⁷³

⁷² See Nyst, *supra* note 68.

⁷³ See Daniel Seng, referencing *Yahoo! v. Association Amicale des déportés d'Auschwitz et des camps de Haute Silesie, le MRAP* (jurisdiction of a Paris Court); see generally Joel R. Reidenberg, *Yahoo and Democracy on the Internet*, 42 *Jurimetrics J.* 261, 262; Ali-baba Course of study materials, *Internet Distribution, e-commerce and other computer related issues: current developments in liability on-line, business methods patents and software distribution, licensing and copyright protection question* (June 2010); WIPO Report, *Comparative analysis of the national approaches to the liability of Internet Intermediaries* (2011).

As the violating material is posted on a third party platform, often sited beyond the territorial limits of the State concerned, providing remedies and reparation to the victim/survivor has proven to be especially difficult. Takedown notices, removal of links and disclosure of identity can only be undertaken by third parties who may or may not be liable for the violating material having been posted on their platforms. Like any profit-driven entities, intermediaries would prefer to take the path that generates the most traffic and income.

The European Court of Justice bridged the extra-territorial arguments by finding that search engines and (by implication, other corporations) with sales and marketing subsidiaries in the European Union, are subject to European law relating to European Union citizens irrespective of where that data is processed. The Court further ordered Google to delink certain websites in its search engines based in Europe as well as in the US on the grounds that although the English language search engine is based in the US, the search engine can be accessed by individuals in Europe and therefore continually causes harm in Europe.⁷⁴

Assertion of extra-territorial jurisdiction is not without problems. States contest the assertion of extra-territorial jurisdiction by other States in areas as diverse as drugs, taxation, trade sanctions and national security trade controls. Extra-territorial jurisdiction may be deemed as challenging other States' sovereignty and violating international law.⁷⁵

Even where laws are enacted to address technology-facilitated violence against women, weak political infrastructure and the inaction of enforcement officers results in these laws being poorly implemented. Existing domestic laws can be gauged by their ability to address the culture of impunity, and the participation and power of women as active agents in this process. It is imperative that States articulate what constitutes technology-facilitated violence against women (when does an author's ill-will or animus toward another become actionable, when does hostility constitute intimidation or threats) and establish training and sensitization programmes for legal and judicial officers to handle cases of technology-facilitated violence against women competently and effectively.

In other instances, victims/survivors have sought to obtain justice through claims of sexual harassment, invasion of privacy, defamation and misappropriation of name and likeness. Where the criminal or quasi-criminal processes fail to meet women's needs, victims'/survivors' are normally expected to commence expensive civil actions.

- *Specific laws and policies*

Laws on technology-facilitated violence have been passed in several countries including Canada, England, Germany, Israel, New Zealand, South Africa, Wales and several US states. The contents of these laws will briefly be reviewed in this section.

The criminal justice system appears, for the most part, ill-equipped and unable to meet the challenges presented by technology-facilitated violence against women. This includes challenges in investigation, prosecution and adjudication of cases involving technology-facilitated violence against women. Even where laws are enacted to address technology-facilitated violence against women, weak political infrastructure, inaction of enforcement officers results in non-efficacy of these laws; examples of these are the cases in DRC, Bosnia and Herzegovina, and Colombia.⁷⁶

⁷⁴ See Mario Costeja Gonzalez v. Google, decision of the Court of Justice of the European Union in 2014. See also Dave Lee, *What is the 'right to be forgotten'?*, BBC News (May 13, 2014), <http://www.bbc.com/news/technology-27394751>.

⁷⁵ For example, unilateral prohibition of exports to unauthorized foreign destinations and US investigation into the North Atlantic Aviation.

⁷⁶ See Nyst, *supra* note 68.

Still, APC's research indicates that the first responders most approached by women who encounter technology-facilitated violence are the police. Victims/survivors are however referred from one agency to another because it is unclear who is responsible or how the complaint should be handled.⁷⁷ Furthermore, misogyny and gender insensitivity still exist among those charged to enforce the law as a result of inadequate training. This results in loss of confidence in the justice systems and discourages women from asserting their rights. It also serves to silence women.

Notable reforms were implemented in California after Shaeffer's death. These include laws that make stalking a crime (felony stalking), availability of long term protection orders (up to ten years) for stalking, restrictions on public access to information from driving records in California, and a specialized Los Angeles police unit that works with prosecutors, attorneys and security details to keep stalkers a safe distance away from their target.⁷⁸

Los Angeles, with its high population of celebrities appears to have undergone a mindset change, with institutional transformation and policy reform. Changes in police and judicial attitudes to stalking, pro-active preventive intervention which includes searching social media and online sites for evidence of stalking, vigilance over the unauthorized release of personal information, including home addresses, investigations that include tracking of digital fingerprints and collaboration with non-state actors all provide a safer environment. Although the impetus for these changes was to protect celebrities, these laws and policies should be made equally applicable to address and eliminate technology-facilitated violence against the general population.

Over the past decade, there have been several prominent incidents of harassment and stalking in South Africa, including the tragic killing of a television journalist, Shadi Rapitso, in 2009. The *Protection from Harassment Act* came into force on 27 April 2013; enabling individuals subject to online or offline harassment to apply to a competent court for a protection order lasting up to five years. The Act also contains provisions requiring electronic communications service providers to assist courts in identifying perpetrators responsible for harassment; and creates the offence of contravention of protection orders and failure of an electronic communications service provider to furnish required information.⁷⁹

The Cyber-safety Act of Nova Scotia (Canada) came into force in August 2013; enabling individuals subjected to cyber bullying (or, in the case of minors, their parents) to apply to a judicial officer for a protection order against an individual. The legislation came about as a direct result of the death of 17-year-old Nova Scotia student Rehtaeh Parsons, who took her own life after having been subjected to months of harassment and humiliation stemming from the dissemination online of a photo of her being allegedly sexually assaulted. The Act also contains provisions requiring electronic communications service providers to assist courts in identifying individuals responsible for cyber bullying, and creates the tort of cyber bullying, which enables individuals to sue for damages arising out of cyber bullying.⁸⁰

In New Zealand, the *Harmful Digital Communications Bill* was introduced in the aftermath of the October 2013 "Roast Busters" sex scandal in which a group of Auckland men allegedly lured young girls into group sex and then posted the video of the incidents online. The Act provides victims with a quick and efficient means of redress for harm (defined broadly) caused to individuals by digital communications (including any text message, writing, photograph, picture

⁷⁷ Rima Athar, *End violence: Women's rights and safety online: From Impunity to justice: Improving corporate policies to end technology-related violence against women*, Association for Progressive Communications (APC) (March 2015), p. 43, available at http://www.genderit.org/sites/default/upload/flow_corporate_policies_formatted_final.pdf.

⁷⁸ See *supra* note 14.

⁷⁹ See Nyst, *supra* note 68.

⁸⁰ *Id.* at 9-13.

or recording). The Act also creates an agency to which victims can turn when they face online abuse; a set of court orders that can be served against Internet hosts and authors upon referral by the aforementioned agency; new civil and criminal offences; and a 48-hour content takedown process whereby individuals can demand that online hosting providers remove content they allege is harmful.⁸¹

The English and Welsh law defines 'revenge porn as ""photographs or films which show people engaged in sexual activity or depicted in a sexual way or with their genitals exposed, where what is shown would not usually be seen in public". It covers images shared on and offline without the subject's permission and with the intent to cause harm.⁸²

Data protection regulation which exists in some countries, may similarly be applied to cases of technology-facilitated violence. Data protection law was held to have applied primarily to outdated and irrelevant data in search results, unless there is a public interest in the data remaining available and even where the search results link to lawfully published content.⁸³ The European Court of Justice ordered Google search engine to delink the result of searches from a specific outdated data.⁸⁴ Google in that case was deemed a data controller of personal data.

A case for the right of victims/survivors "to be forgotten" online can be made out by applying data protection regulation on data and images (fake or otherwise) constituting violence against women that were uploaded either maliciously or without consent. Still, there is no absolute right to be forgotten and the "right to be forgotten" is difficult in practice and may be in conflict with the right to freedom of expression and access to information if abused.⁸⁵ Regulators are divided on whether the Google judgment signals the beginning of a changed approach.⁸⁶

➤ *Internet intermediaries and platform providers*

Whether, when and to what extent platform or service providers should be held liable for third-party content remains unsettled. Mainly the imposition of liability on service providers for third-party content depends on the intermediaries' role. First, did the intermediary provide, for economic purposes, a platform for user-generated comments? Second, did users – whether identified or anonymous – engage in speech which infringes the personal rights of others or amount to either direct threats of violence or hate speech and incitement to violence against them?

Judicial solutions in civil and common law jurisdictions gradually started allowing claims in authorizing infringement; vicarious and contributory liability; inducing infringement; joint wrongdoing (tortfeasorship); aiding and abetting; and negligence. All these developments portend recognition by judges and policy makers that intermediaries should be made "more" responsible.⁸⁷

81 *Id.* at 18-22.

82 'Revenge porn' illegal under new law in England and Wales, BBC News (Feb. 12, 2015), <http://www.bbc.com/news/uk-31429026>.

83 See the Spanish Data Protection Directive.

84 See *supra* note 74.

85 Derechos Digitales, *What are the implications of the right to be forgotten in the Americas?*, IFEX (Sept. 22, 2015), https://www.ifex.org/americas/2015/09/22/derecho_olvido/.

86 Taylor Wessing, *Google Spain and the 'right to be forgotten'*, Global Data Hub (Nov. 2014), http://united-kingdom.taylorwessing.com/globaldatahub/article_2014_google_spain.html.

The European Union has however initiated steps to put in place a policy to protect the right of individuals to have their data fully removed when it is no longer needed for the purposes for which it was collected.

87 Daniel Seng, *Comparative Analysis of the National Approach to the Liability of Internet Intermediaries*, p. 5, available at http://www.wipo.int/export/sites/www/copyright/en/doc/liability_of_internet_intermediaries.pdf (last visited June 16, 2017).

Intermediaries can serve as the informational and access gateways for infringing activities and are able to prevent or stem the flood of violating materials which are facilitated through the intermediaries' facilities and services. Furthermore, intermediaries are profiting from these activities. Still, intermediaries' responsibility is not one of strict liability and provision must be made for when and how intermediaries' responsibility should be engaged. One example may be that intermediaries should be made responsible after the violating material has been brought to the intermediaries' attention and opportunity given to the intermediary to take the requisite action such after an inquiry.⁸⁸

The courts have held, where warranted, shifting the risk of the victim/survivor obtaining redress to the internet company, which was usually in a better financial position than the perpetrator, was not as such a disproportionate interference with the media company's right to freedom of expression.⁸⁹

Presently, liability of internet intermediaries largely pertain to copyright infringements. Indeed potential liability of internet intermediaries for content posted on their platforms have raised one of the "most spirited and fascinating debates in the legal arena, putting right holders, service providers and Internet users at loggerheads".⁹⁰

Copyright interests are represented by huge concerns within a multi-billion dollar industry. Internet intermediaries similarly can have resources and income to rival those of many States. These disputes concern billions of dollars in potential revenue, expenditure and loss. With nearly bottomless financial resources, stakeholders in these disputes are able to engage the best of minds and exert influence over the highest-ranking lawmakers.

The stage set between internet intermediaries and violence against women victims/survivors cannot be further removed from the stage set between copyright concerns and internet intermediaries. Unlike intellectual property protection, which involves big corporations with limitless funds pursuing violators, internet intermediaries and influencing governments, victims/survivors of technology-facilitated violence are everyday women. The high cost of litigation and such formidable opponents in the form of internet intermediaries with resources that rival States can combine to defeat victims/survivors at the outset. These obstacles are especially acute for women who already face greater challenges in accessing justice, such as poor women, female teenagers, younger women and sexual minorities. It also has the effect of bringing more unwanted attention to and can prompt recurring instances of the violation, since courts are not always willing to shield the victims/survivors by giving them anonymity.

Internet intermediaries must further establish comprehensive policies on technology-facilitated violence against women. The posting of disclaimers stating that the writers of the comments – and not the applicant company - are accountable for them does not necessarily result in zero liability when violence occurs. Even if legal obligations cannot be proven, advocates are increasingly insisting that social media platforms have an ethical duty to ensure that technology remains accessible to all. This means that online discrimination and violence must end. Firms that refuse to take substantive measures to curb technology-facilitated violence will increasingly become the centre of controversy.

Internet companies are increasingly finding themselves facing a conflict between profits and social justice; and between freedom of expression and the freedom from discrimination. The

88 Dia Kayyali & Danny O'Brien, *Facing the Challenge of Online Harassment*, Electronic Frontier Foundation (Jan. 8, 2015), available at <https://www.eff.org/deeplinks/2015/01/facing-challenge-online-harassment>.

89 Krone Verlags GmbH & Co. KG v. Austria (no. 4), no. 72331/01, § 32, 9 (November 2006).

90 Béatrice Martinet Farano, *Internet Intermediaries' Liability for Copyright and Trademark Infringement: Reconciling the EU and U.S. Approaches*, (Transatlantic Technology Law Forum Working Paper No. 14, Sept. 1, 2012), <https://law.stanford.edu/publications/internet-intermediaries-liability-for-copyright-and-trademark-infringement-reconciling-the-eu-and-u-s-approaches/>.

recent controversy involving Nextdoor, a “private social network for your neighborhood” is an example. In response to accusations of racial profiling by users, Nextdoor initiated simple anti-profiling measures. The site warns user of racial profiling, “Ask yourself – is what I saw actually suspicious, especially if I take race or ethnicity out of the equation?”⁹¹ The concept behind these warnings is what activists have advocated. This is similar to copyright warnings employed by other technology companies before allowing members to upload material.

These companies also release community guidelines emphasizing that the posting of comments that are contrary to good practice or contain threats, insults, obscene expressions or vulgarities, or incite hostility, violence or illegal activities, are prohibited. Many portals have an automatic system to delete comments based on stems of certain vulgar words with a notice-and-take-down system in place, whereby anyone could notify the administrator of inappropriate comments by simply clicking on a button designated for that purpose. In addition, on some occasions, administrators have removed inappropriate comments on their own initiative. Both Twitter and Facebook have taken the positive step to opening a dialogue with women's rights groups to receive input into the design of policies and processes.

Still, there has only been one known recent incident of a user having been permanently banned for “participating in or inciting targeted abuse of individuals”.⁹²

PART VI Way forward

A. The State

Although access to the internet and other digital spaces is most often facilitated by private entities, it is crucial to regard this space not as private but public, albeit controlled by private entities. After all, some of these spaces are accessed by millions of users.

Specific laws on technology-facilitated violence as well as specialized mechanisms with trained and skilled personnel are required to confront and eliminate technology-facilitated violence. However, merely criminalizing technology-facilitated violence does not provide the remedy required by technology-facilitated violence victims/survivors. Experience has shown that women's access to justice should be a mix of criminal, civil and administrative processes and include the areas of all the 5Ps, namely in prevention of technology-facilitated violence; protection of victims/survivors; prosecution and punishment of perpetrators and provision of redress and reparation for the victims/survivors.

The State is responsible for establishing regulating mechanisms consisting of an independent authorizing entity; though the independent entity should not serve to authorize itself. The regulatory framework must include provision for the possibility of ordering internet and digital intermediaries to divulge information required to identify the perpetrators where circumstances warrant it, through injunctions or injunction like orders. It must also respect and provide for the

⁹¹ Sam Levin, *What happens when tech firms end up at the center of racism scandals?*, The Guardian (Aug. 30, 2016), <https://www.theguardian.com/technology/2016/aug/30/tech-companies-racial-discrimination-nextdoor-airbnb>.

⁹² Milo Yiannopoulos was banned in relation to the online abuse of Leslie Jones. See Laura Bates, *Leslie Jones's Twitter abuse proves relying on users to report bullies isn't enough*, The Guardian (July 21, 2016), <https://www.theguardian.com/lifeandstyle/womens-blog/2016/jul/21/leslie-jones-twitter-abuse-proves-relying-on-users-to-report-bullies-isnt-enough>.

right of victims to restitution. This redress for example should be specific and proportional to the harm, as well as necessary under the circumstances (see the Manila Principles).⁹³

Where voluntary self-regulation by intermediaries fails to deliver the remedies needed, States need to establish independent judicial or quasi-judicial mechanisms to assist victims/survivors in obtaining these remedies.

Admittedly, intermediaries are not responsible nor can they be made liable for the initial act of violence, namely that of posting the violating material online. However, the continued accessibility or dissemination of these materials means that the victim/survivor is continually subjected to violence. Under these circumstances, the State must, in compliance with its international obligation of exercising due diligence to eliminate violence against women, hold intermediaries accountable for failure to remedy the harm or allow their platforms to be the instrument of continued violence after notice of the violence is drawn to their attention.

State regulation must be conscious of not violating freedom of expression yet at the same time, prioritizing women's access to online technology in a safe environment where perpetrators of technology-facilitated violence do not enjoy impunity. The State has a positive role in creating an enabling environment for freedom of expression and equality, while recognizing that this opens up avenues for potential violence. Strong democratic structures — including free and fair elections, an independent judiciary and a vibrant civil society — are needed to prevent abuse and to realize more fully the goals of pluralism and equitable access.⁹⁴ States must also include women's rights organizations in the development of regulations, and adopt a human rights' approach.

B. Internet intermediaries

Self-regulation by internet intermediaries and platform providers remains the most viable method of imbuing corporations with responsibility. As with off-line violence, consent must be the pillar around which both preventive and post-incident policies are formulated. Content that speaks of rape or sexual and physical violence toward an individual or identified individuals should not be treated as freedom of expression.

Victims/survivors of violence, whether they live on college campuses or in remote villages require that violence cease; yet postings on the internet have a level of permanence and can repeatedly be searched, accessed and disseminated. Cessation of technology-facilitated violence and the restoration of privacy can only be provided by internet intermediaries and platform providers.

Intermediary corporations must recognize violence against women as unlawful behaviour, and demonstrate increased and expedited cooperation in providing relief to victims/survivors within the corporations' capacities. This could be through systems for cooperating with law enforcement, takedown procedures for abusive and harmful content, and/ or the possibility of account termination for misconduct. The intermediaries' reporting procedure and mechanisms, as well as remedies, must be accessible and transparent. Exercising due diligence includes setting out when and how intermediaries are deemed to have had notice of such violence.

Corporations should also create appropriate record keeping systems specific to violence against women, and classify and share the ways in which they have responded to reports of

⁹³ See *supra* note 32.

⁹⁴ See for example, Article XIX, *Camden principles on Freedom of Expression and Equality* (April 2009), available at <https://www.article19.org/data/files/pdfs/standards/the-camden-principles-on-freedom-of-expression-and-equality.pdf>. □

such violence.⁹⁵ Internet intermediaries must also commit to and implement comprehensive human rights standards as well as committing to, and operationalising, the UN Guiding Principles on Business and Human Rights.⁹⁶

The mechanisms set up to respond to violence against women must be available and accessible to victims/survivors regardless of their geographic location. When developing liability rules for intermediaries, it is important that legal requirements are appropriate and proportional to the function and size of the intermediary. Policies must be responsive to all women including those outside Europe and North America. Given that the reach of the internet and digital media is neither limited by nor respectful of geo-political boundaries, complaint mechanisms should be equally global.

National regulators and regional courts alike have recognized victims/survivors' rights to restitution, namely their right to have violating materials taken down or de-linked from the result of searches. This right, sometimes referred to, rightly or erroneously, as "the right to be forgotten" compels intermediaries to exercise due diligence under certain circumstances.⁹⁷ These circumstances should include materials that constitute violence against women. However, it may be impossible to ensure a complete take down of the violating material. In such circumstances, certain actions, such as delinking the result of searches to the violating material, may be deemed reasonable and sufficient to stop the harm.⁹⁸ These actions must be proportionate and capable of remedying the harm caused.

Intermediaries should also seek to empower users through hotlines, awareness raising and education. More pro-active measures such as formulating and publicizing anti-violence against women policies and posting reminders and warnings that the content of materials about to be uploaded should not constitute violence against women may go some way toward corporations' meeting their due diligence responsibilities to protect and respect human rights and to provide remedy in case of violations.

However, the ensuing jurisprudence from multiple jurisdictions has resulted in confusing or conflicting court decisions.⁹⁹ What is required is an international multi-stakeholder framework that harmonizes and prescribes the factors to be considered for indirect internet intermediary liability and the defenses available against such liability.¹⁰⁰

95 This includes the use of multi-stakeholder policy platforms, such as Global Network Initiative, as opportunities.

96 See also Athar, *supra* note 77.

97 The "right to be forgotten" is still a debatable concept. The right to be forgotten, outside the gender-based violence context, is sometimes used to compel intermediaries to take down criticisms and political dissent. Alternatively, it is also sought by expunge criminal past. Most recently a Japanese court dismissed the claims of a man convicted of violating child prostitution and pornography laws for his criminal past to be removed from Google search results. "The deletion (of references to the charges from search engines) can be demanded only when value of privacy protection clearly exceeds freedom of expression of search sites", said the Court. See Justin McCurry, *Japanese court rules against paedophile in 'right to be forgotten' online case*, The Guardian (Feb. 1, 2017), <https://www.theguardian.com/world/2017/feb/02/right-to-be-forgotten-online-suffers-setback-after-japan-court-ruling>.

98 See *supra* note 71. See also Google policy on delinking the content: <https://support.google.com/websearch/troubleshooter/3111061#ts=2889054%2C2889099> (last visited 16 June 16, 2017).

99 Compare the court decisions of *A&M Records Inc v. Napster* (9th Cir. 2001) and *UMG Recordings Inc. et al. v. Veoh Networks Inc et al.* (9th Cir. 2011). Napster was held liable for third party infringing content and Youtube not liable despite a high amount of infringing content existing on both platforms.

100 See also Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, Note by the Secretariat, U.N. Doc. A/HRC/32/38 (May 11, 2016), available at <https://documents-dds-ny.un.org/doc/UNDOC/GEN/G16/095/12/PDF/G1609512.pdf?OpenElement>.

PRINCIPLES AND POLICIES ON STATE AND INTERNET INTERMEDIARIES OBLIGATION TO ELIMINATE TECHNOLOGY-FACILITATED VIOLENCE AGAINST WOMEN

1. *Human Rights Approach*

Principle: The State has an obligation under international human rights law to eliminate violence against women, both online and offline and to promote, protect and fulfill human rights. Business enterprises, including internet intermediaries also have the responsibility to respect and protect human rights, and remedy adverse human rights impacts in which they are involved.

Eliminating technology-facilitated violence requires States fulfilling their obligation to promote, protect and fulfill human rights, in collaboration with and with the intercession of internet intermediaries, particularly transnational corporations serving the role of internet intermediaries.

States should also set out clearly the expectation that all business enterprises domiciled in their territory and/or jurisdiction and to protect, respect and remedy human rights throughout their operations.¹⁰¹ Internet intermediaries also have the responsibility to ensure that their platforms are not abused to perpetrate and perpetuate violence against women, and if they are, take immediate action to remedy it.¹⁰²

It is thus crucial to look at responses of different actors, particularly, the identification and role of first responders (including the police, internet intermediaries and helplines), regulators and the judiciary to map the reality of women's experiences and facilitate women's access justice/remedies.

Even where perpetrators are held liable, further reflection is needed on how to hold re-transmitters responsible for the re-transmission of violating materials. This is because sheer volume of persons mob attacking victims/survivors or re-transmitting violating materials result in aggravated harm. In many jurisdictions, the law has developed the concept of reckless indifference and the concept of negligence where intent cannot be established.

2. *Definition: Violence against women*

Principle: Violence against women as a legal concept has been extensively interpreted and its definition can be found in various international human rights instruments. Actionable online gender-based violence (including threats of violence) is gauged by intent to harm, content, credibility or imminence of harm and context.

101 See *supra* note 54.

102 *Id.*

- (a) Violence against women has been rigorously defined in several international instruments. The 1993 UN Declaration on the Elimination of Violence against Women defines 'violence against women' as an act of gender-related violence (GBV) that results in, or is likely to result in, physical, sexual, psychological or economic harm or suffering to women, including threats of such acts, coercion or arbitrary deprivation of liberty, whether occurring in public or in private life.¹⁰³
- (b) Online violence against women are acts 'committed, abetted or aggravated' in part or fully by the use of information and communication technology (ICT) acts of gender-based violence that are committed, abetted or aggravated, in part or fully, by the use of information and communication technologies¹⁰⁴
- (c) While states do not have the obligation to protect individuals from offense, it does have the obligation to protect individuals from harm. Online violence against women is part of the continuum of violence against women that is committed offline. Even if technology-facilitated violence takes the form of non-physical gender-based violence, such violence can approximate and sometimes exceed the harm of physical violence. Furthermore, human rights protected offline must also be protected online.
- (d) Gender-based hate speech (advocacy of gender-based hatred that constitutes incitement to discrimination, harm, hostility or violence) should be prohibited.¹⁰⁵ Incitement to harm comprises of both incitement against a group and incitement against an individual.

3. Freedom of Expression and Technology-Facilitated Violence against women

Principle: Eliminating technology-facilitated violence ensures that the digital space remains a platform for everyone to exercise their rights to freedom of opinion and expression. Measures to eliminate online gender-based violence should respect freedom of expression and be limited to what is necessary and proportionate to address technology-facilitated violence against women.

Human rights are universal, inalienable, inter-related, inter-dependent and indivisible. Freedom from gender-based violence against women, freedom of expression and rights to privacy are protected by international human rights law. An individual's human rights are not absolute in that it cannot be enjoyed at the expense of the human rights of others.

The exercise of these rights under international human rights law is not absolute and may be subject to certain restrictions. This is different from freedom of opinion. The right to hold opinions without interference is an absolute right and "permits no exception or restriction".¹⁰⁶ However, the expression of an opinion that is the right to freedom of expression bears

¹⁰³ Violence against women has been defined and elaborated in many human rights and feminist instruments and discourse including CEDAW. The following forms of violence share similarities to technology-facilitated violence against women: intimate partner violence, domestic violence, sexual harassment, harassment based on gender, stalking and inciting others to commit violence against women.

¹⁰⁴ See *supra* note 10.

¹⁰⁵ Currently Rabat Plan of Action prohibits advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence. Gender-based hatred should be similarly prohibited.

See http://www.ohchr.org/Documents/Issues/Opinion/SeminarRabat/Rabat_draft_outcome.pdf (last visited May 23, 2017).

¹⁰⁶ *Id.* at 9.

“special duties and responsibilities”. The free exchange of and access to information does not equate to unregulated violence. Freedom of expression or access to information cannot be bought at the expense of women’s security and safety.¹⁰⁷

Restrictions shall only be such as are provided by law and are necessary for respect of the rights of others; and for the protection of national security or of public order. The application of these restrictions by States however, “may not put in jeopardy the right itself”.¹⁰⁸

4. Consent and privacy

Principle: Consent is critical in online gender-based violence and key in differentiating lawful behaviour from unlawful and harmful behaviour. It must be addressed in any mechanism dealing with technology-facilitated violence.

Consent is important in gauging whether there has been violation of privacy with regards to dissemination of private data. Consent that is specific to an individual cannot be expanded to consent for the data to be shared and disseminated. Protection of women’s agency is intrinsically linked to the rights to give informed consent on what and with whom material can be shared.

In the digitalized world of big data, what is personal and public data is blurred. Our personal data is continuously being handled and commoditized.¹⁰⁹ Such personal data however, is no less personal for the fact that it may be available in the public domain. Consent for its dissemination is crucial in determining whether a violation of privacy has been committed.

5. Independent regulating mechanism

Principle: The State is responsible to set up an independent entity that is authorized to hear and decide on cases involving technology-facilitated violence against women and issue effective remedies for the victim/survivor.

The regulatory framework must address the role and responsibility/liability of the primary perpetrators, secondary perpetrators (re-transmitters) and internet intermediaries. It must also include provision for the possibility of ordering the internet service provider to divulge the information required to identify the perpetrator where circumstances warrant it, injunction or injunction-like orders, take down or de-linked from the result of searches orders. An independent monitoring entity should not serve to authorize itself. Neither should it be used as a form of surveillance in violation of human rights. Ensuring timeliness, affordability and responsiveness would facilitate women’s access to these mechanisms.

107 “It is not an exercise of freedom of expression to consciously intimidate women online, express the wish to rape them, threaten to harm them or incite others to do so”, declared Terry Reintke, (Rapporteur for the Committee on Women’s Rights and Gender Equality), Report on *gender equality and empowering women in the digital age, 2015/2007(INI)*, European Parliament, AB-80048/2016, (Apr. 8, 2016), available at <http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//NONSGML+REPORT+A8-2016-0048+0+DOC+PDF+V0//EN>.

108 UN Human Rights Committee, *General Comment No. 34: International Covenant on Civil and Political Rights, Article 19: Freedoms of opinion and expression*, Sept. 12, 2011, CCPR/C/GC/34, para. 21.

109 Personal data means any information relating to an identified or identifiable individual; an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number (e.g. social security number) or one or more factors specific to his or her physical, physiological, mental, economic, cultural or social identity (e.g. name and first name, date of birth, biometrics data, fingerprints, DNA...)

The State has a positive role in creating an enabling environment for freedom of expression and equality, while recognising that this brings potential for abuse. State regulation must be conscious of not violating freedom of expression yet at the same time, prioritize women's access to online technology in a safe environment where perpetrators of technology-facilitated violence do not enjoy impunity. Strong democratic structures — including free and fair elections, an independent judiciary and a vibrant civil society — are needed to prevent abuse and to realise more fully the goals of pluralism and equitable access.¹¹⁰ States must also include women's rights organizations in development of the regulations, and adopt a human rights' approach.

6. Duties and responsibilities of information communication technology intermediaries

Principle: Self-regulation, where effective, remains the most appropriate way to address professional issues relating to internet and digital intermediaries. Reporting procedure and mechanisms, as well as remedies must be accessible and transparent.

Victims/survivors of violence require that the violence cease; yet postings on the internet not only maintains a certain level of permanence, it can continually be searched, accessed and disseminated. Cessation of technology-facilitated violence and restoration of privacy can only be provided by internet intermediaries and platform providers.

Reversing the notion that human rights are the product of the social contract between the State and the individual allows us to presume that human rights are entitlements enjoyed by everyone to be respected by everyone.¹¹¹ The net result being States, corporations and individuals all have human rights obligations. The obligation exists irrespective of whether they are enforceable.

Internet intermediaries' duty however, is not on all fours with the duty borne by the States. Intermediaries for example do not owe the obligation to prevent violence wherever it may occur but only violence occurring on their platforms.¹¹² This may also better accord with the Ruggie principles of not "infringing on the human rights of others" and "addressing adverse human rights impacts with which they are involved".

Intermediaries should seek to empower users to work through hotlines, awareness raising and education. Complaints and remedies should be archived. Intermediary corporations must recognize violence against women as unlawful behaviour, and demonstrate increased and expedited cooperation in providing relief to victim/survivors within the corporations' capacities, for example, as systems for cooperating with law enforcement, takedown (including system-wide removal of content, where possible), or de-linking from the result of searches procedures for abusive and harmful content and the possibility of account termination for misconduct. In this regard, the intermediaries' reporting procedure and

¹¹⁰ See for example, Article XIX, *Camden principles on Freedom of Expression and Equality* (April 2009), available at <https://www.article19.org/data/files/pdfs/standards/the-camden-principles-on-freedom-of-expression-and-equality.pdf>.

¹¹¹ Andrew Clapham, *Human Rights Obligations of Non-State Actors*, Academy of European Law, European University Institute, Oxford University Press (2006), available at http://graduateinstitute.ch/files/live/sites/iheid/files/sites/international_law/shared/international_law/Prof_Clapham_website/docs/HR%20obligations%20of%20non-State%20actors.pdf.

¹¹² Comparison can be drawn from imagining a person drowning. Generally, an individual does not owe a duty, even if he is an excellent swimmer, to attempt to save a drowning person. However, the pool owner who obtains economic benefit from the use of the pool by others, owes a duty to ensure that there are sufficient safeguards to prevent death or drowning in his pool.

mechanisms, as well as remedies must be accessible and transparent. Exercising due diligence includes setting out circumstances under which intermediaries are deemed to have had notice of such violence.

Where voluntary self-regulation by intermediaries fails to deliver the remedies needed, States need to establish independent judicial or quasi-judicial mechanisms to assist victims/survivors in obtaining these remedies.

7. Due Diligence

Principle: States and internet intermediaries should exercise due diligence to eliminate online gender-based violence. This extends to reasonable measures to prevent violence before they occur, and effectively investigate and take action against perpetrators once they occur as well as provide redress and reparation to victims/survivors.

Though this principle evolved to focus principally on State obligations, the principle is also useful in guiding internet intermediaries in developing and implementing policies to end violence against women on their platforms. The five areas where states and internet intermediaries need to exercise due diligence are –

(a) Prevention

Prevention includes measures to thwart the occurrence of violence against women. Good prevention programmes provide awareness of technology-facilitated violence against women and of information services and legal protection available post the incident. States and internet intermediaries should deem technology-facilitated violence not merely as another form of violence but violence that is grounded in discrimination and prevents women from exercising their freedom of expression and access to technology and internet spaces and develop preventive policies accordingly.

(b) Protection

Protection focuses on avoiding the recurrence of further violence (which should be immediate if the perpetrator can be identified) the provision of accessible services, and adequate training and sensitization of first responders.

As information and communication technology allows for the easy and rapid dissemination of information and content, the harm of technology-facilitated violence can be quickly amplified. The obligation to protect does not merely refer to the treatment of the original material, but the uploading and dissemination of that material which constitutes recurrence of the violence. Protection from further violence requires not only immediate action on the part of States and internet intermediaries; it sometimes necessitates extra-territorial reach.

(c) Prosecution

Prosecution refers to investigation and instituting proceedings against the perpetrators. Where internet intermediaries are concerned, such proceedings may consist of inquiries. Given the speed and breadth of dissemination of data and material online, States and

internet intermediaries must take prompt effective action upon having notice of the perpetration of the act of violence.

(d) Punishment

Punishment refers to the obligation to impose sanctions/negative consequences on perpetrators. The certainty of adequate punishment creates a level of predictability and sends a message that technology-facilitated violence against women will not be tolerated.

(e) Provision of redress and reparation

Generally, reparations and restitution to victims of violence include compensation for the costs of quantifiable losses (cost of medical care, loss of wages, and damage to property), injuries and non-quantifiable losses and for the needs of the victims/survivors of violence to re-build their lives in the short, medium and long terms, as they transition from a violent situation to a life free from violence are granted as civil remedies. For technology-facilitated violence, remedies must include the ability of victims/survivors to have content removed (where possible) or delinked from the results of searches.