Digital Policy Hub — Working Paper

How Can Impact Assessments Improve Protection from TFGBV?

Paula Martins

Fall 2023 cohort

About the Hub

The Digital Policy Hub at CIGI is a collaborative space for emerging scholars and innovative thinkers from the social, natural and applied sciences. It provides opportunities for undergraduate and graduate students and post-doctoral and visiting fellows to share and develop research on the rapid evolution and governance of transformative technologies. The Hub is founded on transdisciplinary approaches that seek to increase understanding of the socio-economic and technological impacts of digitalization and improve the quality and relevance of related research. Core research areas include data, economy and society; artificial intelligence; outer space; digitalization, security and democracy; and the environment and natural resources.

The Digital Policy Hub working papers are the product of research related to the Hub's identified themes prepared by participants during their fellowship.

About CIGI

The Centre for International Governance Innovation (CIGI) is an independent, non-partisan think tank whose peer-reviewed research and trusted analysis influence policy makers to innovate. Our global network of multidisciplinary researchers and strategic partnerships provide policy solutions for the digital era with one goal: to improve people's lives everywhere. Headquartered in Waterloo, Canada, CIGI has received support from the Government of Canada, the Government of Ontario and founder Jim Balsillie.

Partners

Thank you to Mitacs for its partnership and support of Digital Policy Hub fellows through the Accelerate program. We would also like to acknowledge the many universities, governments and private sector partners for their involvement allowing CIGI to offer this holistic research environment.

Key Points

- Technology-facilitated gender-based violence (TFGBV) is a pervasive aspect of digital societies and requires urgent and innovative solutions, which include the adoption of smart regulations. A recent example is the EU Digital Services Act (DSA), which adopts platform accountability provisions that require *ex ante* impact assessments for systemic risks, including TFGBV. The EU DSA approach moves away from piecemeal strategies and toward more holistic responses to online content harms.

- Considering TFGBV as a systemic risk is promising. However, no guidance to companies and very little detail on how exactly to carry out a systemic risk assessment have been released.

- The concept of "systemic risk" adopted by the DSA is foreign to the digital regulatory context, as well as to human rights international law standards and practice. It shifts the standard from rights assessments to risk assessments, providing little explanation about their characteristics.

- Although there is some research and practical experience at the global level on the application of human rights impact assessments (HRIAs), which could be seen as a foundation for the new risk assessments, the use of gender impact assessments or the inclusion of a gender dimension in HRIAs is still underdeveloped. One key concern is that most HRIAs fail to embed a gender lens into their planning, application and evaluation.

- Despite the explicit reference to gender-based violence as a form of systemic risk in the DSA, there are many open questions in relation to its applicability and potential impact in preventing and reducing TFGBV. Intrinsic challenges concerning the DSA language, open definitions and unclear implementation framework may jeopardize its potential for success in addressing TFGBV and must be urgently addressed.

Introduction

TFGBV has been growing with the expansion of the digital world, putting at risk women and gender-diverse persons' psychological and physical health and safety. It jeopardizes not only the well-being of those targeted, but also those around them and society at large. This paper argues that TFGBV requires urgent and innovative solutions.

While acknowledging that any solution to TFGBV requires multi-dimensional responses that must reach beyond the policy realm, this paper will assess one possible response to TFGBV that has been increasingly proposed in new and emerging regulations concerning the operation of internet platforms — the use of impact assessments. In particular, it will assess the potential positive impact on TFGBV of the *ex ante* systemic risk assessment model proposed by the European Union's new DSA.[1]

---

1   Following the adoption of the Digital Services Package in the first reading by the European Parliament in July 2022, both the DSA and Digital Markets Act have been adopted by the Council of the European Union, signed by the presidents of both institutions and published in the *Official Journal of the European Union*, available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32022R2065&qid=1666857835014.

The sections below will look at the definition of TFGBV and its pervasiveness, followed by an explanation of the current mainstream approach to platform accountability and the opportune shift provided by the DSA — from looking at online harms as individual challenges to looking at such harms as a societal problem — and how this shift matters to TFGBV. The section following that will provide a brief overview of the existing efforts to assess human rights and social impacts of business and how they have failed to address gender concerns. Finally, the paper will address the potential risks of a poorly regulated or implemented DSA. It will conclude with recommendations aimed at addressing such risks.

The Problem

Before exploring potential responses to TFGBV, this section will briefly review its concept and extent, recognizing definitional challenges while reaffirming its widespread prevalence and differentiated impacts.

In 2023, the sixty-seventh session of the Commission on the Status of Women (CSW67) adopted an annual priority theme related to technology for the first time.[2] In its Agreed Conclusions, the CSW67 stated that "while technology can be used to promote women's and girls' full realization of civil, political, economic, social and cultural rights, it can also be used to perpetuate gender stereotypes and negative social norms and create vicious cycles, in which inequalities are amplified and perpetuated through digital tools" (CSW67 2023).

This text by the CSW67, as well as other language included in the Agreed Conclusions, was an important recent benchmark in the long struggle of women's activists to see TFGBV recognized as a real problem that imposes grave and deep effects on survivors.

Despite this long struggle, even terminology continues to be controversial. Sometimes referred to as online violence against women, digital violence or cyberviolence, there is little agreement about lexicon and the actual contours of this type of violence. This paper will adopt the terminology of TFGBV.

The United Nations Population Fund (UNFPA) defines TFGBV as "an act of violence perpetrated by one or more individuals that is committed, assisted, aggravated and amplified in part or fully by the use of information and communication technologies or digital media, against a person on the basis of their gender" (UNFPA 2023b). TFGBV is a broad concept, encompassing a variety of harms. As pointed out by the UN Special Rapporteur on the right to freedom of expression and opinion, "online gender-based violence includes both harmful speech and behaviour. It is often sexist or misogynistic in nature and contains digital threats or incitement to physical or sexual violence" (UN General Assembly 2021).

Recent research published by the Centre for International Governance Innovation's Supporting a Safer Internet project (in partnership with the International Development Research Centre) demonstrates the pervasiveness of TFGBV. According to the survey

---

2   Innovation and technological change, and education in the digital age for achieving gender equality and the empowerment of all women and girls. See CSW67 (2023).

(conducted by Ipsos) carried out under the project in 18 countries, almost 60 percent (59.7 percent) of all participants had experienced at least one of the 13 forms of online harm[3] surveyed (Dunn, Vaillancourt and Brittain 2023).

Transgender and gender-diverse people reported the highest proportion of incidents experienced, followed by cisgender women. Women were much more likely to rate the various forms of online harm as harmful compared to men (ibid.). Lesbian, gay, bisexual, transgender, queer or questioning (LGBTQ+) people were much more likely to report a serious impact from online harms compared to heterosexual people (ibid.).

This data shows that although TFGBV is widespread, it does not impact everyone the same. Marginalized groups tend to not only be targeted more but also face differentiated effects that arise from the persistent and structural exclusion they face.

As for the effects of TFGBV, research has found wide-ranging impacts of TFGBV that can be as severe as those of offline gender-based violence and that include:

- promotion of cultures of violence, including the normalization of misogyny and violence against women;

- connection to offline acts of physical, sexual and emotional violence;

- contribution to the gender digital divide, and exclusion of women from the digital public square;

- worsening of women's economic exclusion, given the growing role of digital tools and technologies in the modern economy; and

- stifling of women's voices and infringing upon women's rights to political participation (UN Women and World Health Organization 2023).

As this list makes explicit, the impacts of TFGBV extend far beyond the individuals targeted and impose significant threats to society at large, to democracy, to human rights and to gender justice.[4]

---

3   The 13 online harms surveyed by the study are: physically threatened online; blackmailed online; monitored, tracked or spied on online; someone accessing devices or social media accounts belonging to them without permission; called discriminatory names or derogatory cultural terms; personal nude or sexual images of them shared or shown to someone else or posted online without permission; unwanted sexual images sent to them; having personal contact information or address posted online without permission (doxing); lies posted online about them; online impersonation; repeatedly contacted by someone they do not want to be contacted by; networked harassment; and experienced harassment online because of their gender, race, sexual orientation, disability, gender expression or other marginalizing factors.

4   According to Goetz (2007):

Gender justice can be defined as the ending of — and if necessary the provision of redress for — inequalities between women and men that result in women's subordination to men. These inequalities may be in the distribution of resources and opportunities that enable individuals to build human, social, economic, and political capital. Or, they may be in the conceptions of human dignity, personal autonomy and rights that deny women physical integrity and the capacity to make choices about how to live their lives. As an outcome, gender justice implies access to and control over resources, combined with agency. In this sense it does not differ from many definitions of "women's empowerment." But gender justice as a process brings an additional essential element: accountability. Gender justice requires that women are able to ensure that power-holders — whether in the household, the community, the market, or the state — can be held to account so that actions that limit, on the grounds of gender, women's access to resources or capacity to make choices, are prevented or punished. The term "women's empowerment" is often used interchangeably with "gender justice," but gender justice adds an element of redress and restitution that is not always present in discussions of women's empowerment.

Although TFGBV is part of a continuum that characterizes the structural misogyny, discrimination and inequality faced by women and gender non-conforming people offline, the advent of new technology, in particular the widespread use of social media, has led to mechanically amplified messages of hatred and facilitated broad coordination between malignant actors (Association for Progressive Communications [APC] 2018).

In the era of generative artificial intelligence (AI), it is possible to easily create realistic fake media of different types — including images, audio and "synthetic histories" — allowing for the easier coordination of harsh campaigns of disinformation and harassment or, in the most extreme cases, hate speech and incitement (UN Education, Scientific and Cultural Organization 2023). As AI takes on language skills, artificially created narratives create confusion and hallucinations. Adding to the violence perpetrated on specific individuals, these new types of attacks aggravate information disorders[5] that impact whole information ecosystems.

Big tech business models, based on the extraction of personal data and attention economics, have been found to promote controversial and aggressive content, amplifying harmful messaging and exploiting social fissures (Maréchal and Roberts Biddle 2020; O'Reilly n.d.). Social media platforms have been accused of prioritizing profit over people's well-being, in particular, those groups most marginalized and that are, at the same time, also the most targeted or impacted by technology-facilitated violence (PEN America 2021).

A Potential Solution? The DSA Approach to Platform Regulation

UN Women has affirmed that existing law and policy frameworks concerning gender-based violence have not kept pace with progress in technology (UN Women 2022). Public outcry has called for increased platform accountability in relation to online harms, and many have defended the need for improved legal frameworks that could deliver such accountability.

In Europe, a new piece of legislation seeks to propose a novel model of platform regulation. Different from most proposals currently under debate, the DSA model seeks to address online harms as a societal problem, not as an individual one. This approach is evident in the concept of "systemic risk."

Theoretically, the systemic risk approach could be better suited to respond to the structural discrimination and misogyny that are behind TFGBV. However, intrinsic challenges concerning the DSA language, open definitions and unclear implementation framework may jeopardize its potential for success.

---

5   According to Wardle and Derakhshan (2017), information disorders include dis-, mis- and malinformation, which intersect around the concepts of falseness and harm.

The DSA Platform Accountability Provisions

The DSA was adopted in 2022 and brings a number of provisions that establish rules defining responsibilities and obligations for intermediaries providing services in the European Union.[6] Among such obligations, the DSA requires the largest platforms and search engines to regularly assess systemic risks to various social interests arising from "the design, including algorithmic systems, functioning and use made of their services"[7] and to reasonably and proportionately mitigate these risks.

Very large online platforms (VLOPs) and very large search engines (VLOSEs)[8] need to carry out annual risk assessments, which should cover:

- dissemination of illegal content; and

- actual or foreseeable negative effects:

  - on any fundamental rights;

  - on civic discourse and electoral processes, and public security; and

  - *in relation to gender-based violence,* the protection of public health and minors, and serious negative consequences to the person's physical and mental well-being.

The DSA also requires that assessments be carried out before new products are launched.

Following such an assessment, VLOPs and VLOSEs must adopt risk mitigation measures,[9] which should be reasonable, proportionate and effective. They are also subject to an audit regime (as set out in article 37 of the DSA[10] and in the European Commission's forthcoming delegated legislation)[11] and alternative mitigation measures can be required, as well as fines applied. The European Commission and the Board for Digital Services are mandated to review the assessments and develop best practices and guidelines on specific risks.[12] This guidance should then be taken into consideration by companies in their subsequent assessments.

These DSA provisions follow the approach adopted by the United Nations Guiding Principles on Business and Human Rights (UNGPs) (with the important difference that the DSA is a statutory norm and the UNGPs are voluntary principles), which establish that companies are required to conduct human rights due diligence across its business activities and relationships in order to comply with their obligation to respect international human rights standards. Different from the more traditional rights assessment, however, the DSA adopts a risk-assessment approach.

The DSA, however, neither provides a definition of systemic risk nor clarifies how to assess it. What it does is to vaguely indicate some kinds of risks in article 34(1): the

---

6   On August 25, 2023, the DSA came into effect for VLOPs and VLOSEs. It became fully applicable to other entities on February 17, 2024.

7   EC, *Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC,* [2023] OJ, L 277 [*DSA*], art 34.1.

8   VLOPs and VLOSEs are those that reach an average of 45 million monthly active users within the European Union. The first set of VLOPs and VLOSEs was published in April 2023; see European Commission (2023).

9   *DSA, supra* note 7, art 27.

10   *DSA, supra* note 7.

11   *Ibid,* art 37.7.

12   *Ibid,* art 35 2(b).

dissemination of illegal content through their services; any negative effects for the exercise of the fundamental rights to respect for private and family life, freedom of expression and information, the prohibition of discrimination and the rights of the child; and intentional manipulation of their service, including by means of inauthentic use or automated exploitation of the service, with an actual or foreseeable negative effect on the protection of public health, minors, civic discourse, or actual or foreseeable effects related to electoral processes and public security. It also points out to their potential sources in article 34(2), when it states that "when conducting risk assessments, very large online platforms shall take into account, in particular, how their content moderation systems, recommender systems and systems for selecting and displaying advertisement influence any of the systemic risks" referred to before, including the potentially rapid and wide dissemination of illegal content and of information that is incompatible with their terms and conditions.

Different from the criteria used in the EU AI Act currently under debate — which refers to "high risk"[13] — the adoption of systemic risk by the DSA seems deliberately open in nature. According to Martin Husovec, the DSA regulator opted for a regulatory technique that he refers to as "risk management dialogue" for handling online harms (Stanford Cyber Policy Center 2023). The technique aims at addressing significant information asymmetry and the opacity of the ecosystem. In practice, that means that the regulator has no clear idea of the risks or contributing factors and is in the dark about solutions. According to Husovec, the answer presented by the DSA is to force the digital service providers themselves to think about the risks, let their conclusions be evaluated by other stakeholders (auditors, researchers, field non-governmental organizations) and then the regulator can form its opinion and provide further guidance for future assessments (ibid.).

What Is a Systemic Risk and Why Does It Matter for TFGBV?

The DSA systemic risk approach could be considered a promising model for addressing TFGBV from a regulatory standpoint: first, it provides an alternative to overly strict regulations and the disproportional application of criminal law, which too often can be abused to target gender activists and gender non-conforming individuals; second, it focuses on preventing harms, instead of simply seeking responsibility for attacks already consummated; and third, it understands that piecemeal (individual) responses to online harms are insufficient to counteract structural and prevalent challenges such as gender-based violence.

---

13   EC, *Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts*, [2021] COM(2021) 206 final, 2021/0106(COD), online: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52021PC0206>.

Addressing Systemic Risks While Protecting Freedom of Expression

Most models of platform accountability promoted by states tend to endorse strict regulatory frameworks or recourse to criminalization. Regulating content can be risky since it may endanger fundamental rights, such as freedom of expression and association. Civil society organizations have collected cases that demonstrate how cybercrime laws, fake news bills and other legislation addressing digital issues have been used to criminalize dissenting and oppositional voices, including gender activists (APC and Derechos Digitales 2023).

International freedom of expression standards allow for the establishment of restrictions to this right, as long as such limitations respect principles of legality, necessity and proportionality. Criminal laws and strict regulations should be reserved for the most extreme situations such as hate speech and incitement, because of the chilling effect they can have while at the same time lacking proof of effectiveness. For harmful speech and online content that is not considered illegal, other measures should be explored.

Betting on Prevention

Arguably, a more effective solution to deal with harmful content could be the adoption of preventive measures — the goal, more than simply to punish violators, is to avoid damage in the first place. This is the approach adopted by the DSA, with the establishment of *ex ante* obligations on platforms, in addition to *post facto* responsibility in case of illegality.

Individual versus Societal Harms

Scholars such as Evelyn Douek have criticized the most common approaches to content moderation that seek to respond to online freedom of speech violations on a case-by-case basis, focusing too much on due process and *ex post* individual review — an approach she argues is highly ineffective in view of the speed and scale of online speech. She proposes that content moderation should instead be understood as "a project of mass speech administration" that needs a more "proactive and continuous form of governance than the vehicle of individual error correction allows" (Douek 2022).

The DSA seems to build on similar premises, for its proposal is to move away from piecemeal strategies and toward more holistic responses to online content harms.

Yale Information Society scholar Niklas Eder (2023) points to other reasons for adopting such a complex system of accountability for online harms in the DSA. He stresses that the DSA, in part, continues to promote the individual remedy solution, and even strengthens it through the introduction of the "demotion of content" mechanism and by establishing individual remedies as a legal requirement in certain cases. However, he considers that individual remedies are inadequate and built on neo-liberal constructions of individual liberty and responsibility. Safety online should be guaranteed by the "welfare state," taking the burden off of the individual's shoulders and addressing this challenge as a societal problem. This is what the DSA seeks to promote with the introduction of a risk-assessment regulatory model, which complements its own system of individual remedies.

In that sense, systemic risk could be an interesting approach to tackle TFGBV, which is a structural type of violence that both results from and causes gender inequality and discrimination. Online gender-based violence cannot be seen as a separate phenomenon from the gender-based violence that occurs offline — TFGBV is part of the online-offline continuum of structural misogyny faced historically by women (APC 2023): it "not only reflects and further entrenches inequalities in the online world, it also increasingly interacts with and influences the offline world" (UNFPA 2023a). That is why academics and activists researching and working to stop TFGBV have called for technical, social and legal strategies that promote comprehensive, multi-faceted and multi-stakeholder responses to this grave societal problem (Bailey and Burkell 2021).

Jane Bailey, from the University of Ottawa, and other researchers argue that framing TFGBV as merely an issue of interpersonal behaviour is simplistic and risky, for it may render invisible broader issues that affect, in particular, women from marginalized communities and groups. These women, communities and groups face intersectional challenges, which create a complex system of discrimination and exclusion composed of the many compounding layers of challenges they face in view of their multiple identities (Bailey et al. 2022). Structural, systemic and design factors contribute to TFGBV and, therefore, solutions to TFGBV must be framed through intersectional and structural lenses.

A trend among current responses to TFGBV is the establishment of better procedures for reporting harms and the building of better ways for women to curate their safety online (World Wide Web Foundation, Feminist Internet and Walker 2021). These solutions, however, put the burden on women to protect themselves from harms and vulnerabilities they have not created. In addition, many of the existing reporting systems are not only burdensome but also subject women to revictimization. As highlighted by PEN America (2023), "the fact that people who are harassed online experience trauma and other forms of psychological harm can make the troublesome reporting process all the more frustrating."

Considering TFGBV as a systemic risk, therefore, is promising. However, no guidance to companies and very little detail on how exactly to carry out a systemic risk assessment have been released. Any value judgment on the potential of success of this model remains open to regulatory arrangements still to be adopted and uncertain implementation efforts.

The DSA Intrinsic Limitations

Despite the explicit reference to gender-based violence as a form of systemic risk in the DSA, there are many open questions in relation to its applicability and potential impact in preventing and reducing TFGBV.

The concept of systemic risk adopted by the DSA is foreign to the digital regulatory context, as well as to human rights and gender international law standards and practice. As mentioned before, it shifts the standard from rights assessments to risk assessments, providing little explanation about their characteristics.

Additionally, there are some broader grounds for concern in relation to the model proposed by the European regulation. First, by setting the category of VLOPs and

VLOSEs, the regulator used size as a proxy for high risk; however, that may exclude from the systemic-risk assessment obligation smaller service providers that excel as spaces for the cultivation and radicalization of gender hate and misogyny. Second, by including four main categories of risks in article 34(1), the regulator may incentivize the fragmentation of tools and analysis at the expense of a more holistic view of potential risks. Third, the list of factors to be considered in the risk assessments as per article 34(2) are all endogenous to the platforms' systems (content moderation systems; applicable terms and conditions and their enforcement; systems for selecting and presenting advertisements; data-related practices; intentional manipulation of services; amplification and dissemination of content that is illegal or violates terms and conditions), leaving aside exogenous aspects, in particular, gender-responsive context analysis and the broader effects of the "platformization" of social interaction and their consequences (Mantelero 2022) to gender rights. Fourth, the adoption of self-assessment as a model and the later auditing by third parties requires further safeguards in relation to independence and transparency of auditors that should be addressed in upcoming complementary regulations. And finally, a fifth point of concern relates to the fact that the regime proposed by the DSA assumes and accepts a residual risk when opting for the adoption of mitigation measures; this approach is based on a degree of techno-solutionism that minimizes the negative effects of new and emerging technology, in contrast with the potential benefits for society. This approach fails to address a tendency by technological researchers and large technological companies to explore and seek superfluous technological applications, irrespective of need considerations, building on the premise that innovation in itself is good and desirable. This is especially problematic when we consider that the negative impacts of emerging technology, as well as its benefits, are not evenly distributed in society, as already addressed above.

Gendered Harms and Impact Assessments — Previous Experiences and Their Limitations

As discussed above, the UNGPs and other international human rights "soft" standards already point to the obligation of companies to respect human rights and engage in due diligence practices, including HRIAs. Along with HRIAs, social impact assessments (SIAs) and environmental impact assessments (EIAs) have been in use for some time, derived from environmental regulations or development projects best practices. The experience accumulated thus far in carrying out HRIAs (and, to a lesser extent, SIAs and EIAs) will be an important starting point for the development of the new risk assessments required by the DSA.

However, although there is some research and practical experience at the global level on the application of HRIAs, SIAs and EIAs, the use of gender impact assessments or the inclusion of a gender dimension in HRIAs, SIAs and EIAs is still underdeveloped. One key concern in this regard is that most HRIAs fail to embed a gender lens in their planning, application and evaluation. The results are gender-blind assessments that risk

"perpetuating and exacerbating systemic gender discrimination in societies, or even creating new forms of discrimination" (Götzmann and Bainton 2021).

These are some of the concerns observed by scholars and practitioners (Götzmann and Bainton 2021; Peletz and Hanna 2019; Hill, Madden and Collins 2017; Levac et al. 2021):

- HRIAs fail to recognize that women and gender non-conforming individuals face a disproportionate burden of the negative impacts of business activities and benefit less from them at the same time. This recognition not only needs to be present, but it must inform impact assessment and management.

- HRIAs too often ignore gendered relations, roles and power dynamics and fail to integrate that analysis throughout the assessment process, questioning assumptions, interrogating premises and prioritizing nuances.

- HRIAs tend to apply gender stereotypes that consider women as a homogeneous category and apply male-female binary views. As a result, actions such as women-only consultations or gender-disaggregated data are seen as sufficient for checking the "gender box" in risk assessment and management.

- Related to the above, HRIAs rarely embed an intersectional approach and methodology. That means that the complexities of life in society for communities and individuals who face discrimination and structural inequalities as a result of intersecting systems of oppression (sexism, racism, classism, casteism, ableism, compulsory heterosexuality and so on) are simplified and distorted. In some cases, a broad approach of "addressing the needs of vulnerable communities" is adopted in a manner that leads to a cumulative perspective on the multiple impacts of identities on their lived experiences and exercise of rights. This type of approach contrasts with an intersectional one, which rejects the simple layering of impacts for "intersectionality demands a multi-dimensional and multi-scalar view. This demands an impact assessment process that is able to account for the spatial and temporal dimensions of specific impacts — how specific impacts and forms of discrimination are historically produced and reproduced," clarified by Götzmann and Bainton (2021, 173).

- Another limitation of conventional HRIAs is the idea that the baseline information for assessing negative impacts of projects is the status quo. In that sense, maintaining the original conditions is seen as ideal, while a gender lens may require a more critical assessment that leads to incremental change to improve pre-existing discriminatory patterns.

- Related to the above, conventional HRIAs in general promote limited gender mainstreaming or gender integration views — in the rare cases where gender is considered — not really embedding a gender justice lens that could see the process of risk assessment and management in its transformative potential for substantive gender equality.

- Specific accounts of HRIA processes have also pointed to the limited participation of women and LGBTQ+ voices and the acceptance of cultural views that disenfranchise women, justifying their exclusion.

- Lack of understanding of private versus public dynamics often leads to incomplete analysis in conventional HRIAs.

- Gender analysis of specific contexts is hindered by the lack of disaggregated and whole sets of gender-relevant data that could provide an understanding of social and economic settings.

- When gender and other identity factors are considered — with the limitations above — it is normally too late in the assessment cycle. Gender analysis will be most effective and beneficial if it is integrated into all stages of the impact assessment process.

- Tools, methods and skills for more gender-responsive impact assessment and management are limited.

Conclusion and Recommendations

It is too soon to judge whether the model proposed by the DSA can have any real influence on the ongoing plague of TFGBV. From a fundamental conceptual level, a systemic risk regulatory model that imposes *ex ante* obligations on digital service providers does hold the promise of a better response to the problem of gender-based violence in digital spaces, which is structural and entrenched in so many institutional and social practices, relations, norms, laws and policies. However, many questions regarding the operationalization of this proposal remain unanswered and should be addressed promptly. Responding to these questions requires:

- Learning from the past decades of experience in the implementation of HRIAs, EIAs and SIAs, taking into account their poor regard to gender responsive and intersectional methodologies.

- The establishment of a truly multi-stakeholder approach to setting the upcoming delegated norms, guidelines, best practices and codes of conduct that will complement the provisions of the DSA.

- This multi-stakeholder process requires an intersectional gender lens as a crucial component of participation and engagement.

- Finally, something that has not been addressed in this paper due to scope and format limitations, is the development of comparative studies to understand how systemic risk has been defined and assessed in other regulatory areas, such as finance, where the term supposedly originates (Broughton Micova and Calef 2023), and science.

Acknowledgement

About the Author

Paula Martins is a policy advocacy lead at the Association for Progressive Communications where she follows digital and human rights policy issues at the global level and helps shape the organization's responses to emerging policy trends. As a Digital Policy Hub doctoral fellow, her research delves into freedom of expression and digital rights. She is also pursuing her doctorate in law at McGill University.

Acronyms and Abbreviations

AI   artificial intelligence

APC   Association for Progressive Communications

CSW67   sixty-seventh session of the Commission on the Status of Women

DSA   Digital Services Act

EIAs   environmental impact assessments

HRIAs   human rights impact assessments

LGBTQ+   lesbian, gay, bisexual, transgender, queer or questioning

SIAs   social impact assessments

TFGBV   technology-facilitated gender-based violence

UNFPA   United Nations Population Fund

UNGPs   United Nations Guiding Principles on Business and Human Rights

VLOPs   very large online platforms

VLOSEs   very large search engines

Works Cited

APC. 2018. "Providing a gender lens in the digital age: APC Submission to the Office of the High Commissioner for Human Rights' Working Group on Business and Human Rights." November. www.ohchr.org/sites/default/files/Documents/Issues/Business/Gender/APC.pdf.

— — — . 2023. "The Feminist Principles for including gender in the Global Digital Compact." October 6. www.apc.org/en/pubs/feminist-principles-including-gender-global-digital-compact-0.

APC and Derechos Digitales. 2023. "When Protection Becomes an Excuse for Criminalisation: Gender considerations on cybercrime frameworks." August. www.apc.org/sites/default/files/gender_considerations_on_cybercrime_0.pdf.

Bailey, Jane and Jacquelyn Burkell. 2021. "Tech-Facilitated Violence: Thinking Structurally and Intersectionally." *Journal of Gender-Based Violence* 5 (3): 531–42. https://doi.org/10.1332/239868021X16286662118554.

Bailey, Jane S., Carys Craig, Suzie Dunn and Sonia Lawrence. 2022. "Reframing Technology-Facilitated Gender-Based Violence at the Intersections of Law & Society." *Canadian Journal of Law & Technology* 19: 209–16. https://digitalcommons.osgoode.yorku.ca/scholarly_works/2984.

Broughton Micova, Sally and Andrea Calef. 2023. "Elements for Effective Systemic Risk Assessment Under the DSA." July. https://doi.org/10.2139/ssrn.4512640.

CSW67. 2023. "Agreed Conclusions, 67th Session." www.unwomen.org/en/csw/csw67-2023.

Douek, Evelyn. 2022. "Content Moderation as Systems Thinking." *Harvard Law Review* 136 (2): 528–607. https://harvardlawreview.org/print/vol-136/content-moderation-as-systems-thinking/.

Dunn, Suzie, Tracy Vaillancourt and Heather Brittain. 2023. *Supporting Safer Digital Spaces*. CIGI Special Report. Waterloo, ON: CIGI. www.cigionline.org/publications/supporting-safer-digital-spaces/.

Eder, Niklas. 2023. "Making Systemic Risk Assessments Work: How the DSA Creates a Virtuous Loop to Address the Societal Harms of Content Moderation." June 26. https://doi.org/10.2139/ssrn.4491365.

European Commission. 2023. "Digital Services Act: Commission designates first set of Very Large Online Platforms and Search Engines." Press Release, April 25. https://digital-strategy.ec.europa.eu/en/news/digital-services-act-commission-designates-first-set-very-large-online-platforms-and-search-engines.

Goetz, Anne Marie. 2007. "Gender Justice, Citizenship and Entitlements: Core Concepts, Central Debates and New Directions for Research." In *Gender Justice, Citizenship and Development*, edited by Maitrayee Mukhopadhyay and Navsharan Singh, 15–57. Ottawa, ON: International Development Research Centre.

Götzmann, Nora and Nicholas Bainton. 2021. "Embedding gender-responsive approaches in impact assessment and management." *Impact Assessment and Project Appraisal* 39 (3): 171–82. https://doi.org/10.1080/14615517.2021.1904721.

Hill, Christina, Chris Madden and Nina Collins. 2017. *A Guide to Gender Impact Assessment for the Extractive Industries*. Melbourne, Australia: Oxfam. https://oxfamilibrary.openrepository.com/bitstream/handle/10546/620782/gt-gender-impact-assessement-extractives-010117-en.pdf?sequence=1.

Levac, Leah, Jane Stinson, Susan M. Manning and Deborah Stienstra. 2021. "Expanding evidence and expertise in impact assessment: informing Canadian public policy with the knowledges of invisible communities." *Impact Assessment and Project Appraisal* 39 (3): 218–28. https://doi.org/10.1080/14615517.2021.1906152.

Mantelero, Alessandro. 2022. "Fundamental rights impact assessments in the DSA." Putting the DSA into Practice: Enforcement, Access to Justice, and Global Implications debate. *Verfassungsblog* (blog), January 11. https://verfassungsblog.de/dsa-impact-assessment/.

Maréchal, Nathalie and Ellery Roberts Biddle. 2020. *It's Not Just the Content, It's the Business Model: Democracy's Online Speech Challenge*. New America, March 17. www.newamerica.org/oti/reports/its-not-just-content-its-business-model/.

O'Reilly, Tim. n.d. "Tim O'Reilly on AI's Role in the Attention Economy." *Bellagio Conversations in AI* (blog). The Rockefeller Foundation. www.rockefellerfoundation.org/bellagio-conversations-in-ai/tim-oreilly-on-the-attention-economy/.

Peletz, Nicole and Kevin Hanna. 2019. "Gender Analysis and Impact Assessment: Canadian and International Experiences." University of British Columbia and Canadian International Resources and Development Institute. https://internationalwim.org/wp-content/uploads/2020/11/WEB_Gender_Analysis_Impact_Assessment.pdf.

PEN America. 2021. *No Excuse for Abuse*. March 31. https://pen.org/report/no-excuse-for-abuse/.

— — —. 2023. "New Report: To Reduce Online Abuse, Social Media Platforms Must Overhaul Flawed System for Users to Report Harassment and Threats." Press Release, June 29. https://pen.org/press-release/new-report-to-reduce-online-abuse-social-media-platforms-must-overhaul-flawed-system-for-users-to-report-harassment-and-threats/.

Stanford Cyber Policy Center. 2023. "How Does the EU's Digital Services Act Regulate Content Moderation? And Will It Work?" April 11. YouTube video, 1:00:53. www.youtube.com/watch?v=np05wM3h2mc.

UN Education, Scientific and Cultural Organization. 2023. *Technology-Facilitated Gender-Based Violence in an Era of Generative AI*. Paris, France: UN Education, Scientific and Cultural Organization. https://unesdoc.unesco.org/ark:/48223/pf0000387483.

UN General Assembly. 2021. "Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, Irene Khan." A/HRC/47/25. https://digitallibrary.un.org/record/3925306?ln=en.

UNFPA. 2023a. "Preventing Technology-Facilitated Gender-Based Violence (TF GBV)." www.un.org/techenvoy/sites/www.un.org.techenvoy/files/GDC-Submission_UNFPA.pdf.

— — —. 2023b. "What is technology-facilitated gender-based violence?" March. www.unfpa.org/resources/brochure-what-technology-facilitated-gender-based-violence.

UN Women. 2022. "Accelerating Efforts to Tackle Online and Technology-Facilitated Violence against Women and Girls." Policy Brief. www.unwomen.org/en/digital-library/publications/2022/10/accelerating-efforts-to-tackle-online-and-technology-facilitated-violence-against-women-and-girls.

UN Women and World Health Organization. 2023. "The State of Evidence and Data Collection on Technology-facilitated Violence against Women." March. www.unwomen.org/sites/default/files/2023-04/Brief-The-state-of-evidence-and-data-collection-on-technology-facilitated-violence-against-women-en.pdf.

Wardle, Claire and Hossein Derakhshan. 2017. *Information Disorder: Toward an interdisciplinary framework for research and policy making*. Council of Europe report. DGI(2017)09. https://rm.coe.int/information-disorder-toward-an-interdisciplinary-framework-for-researc/168076277c.

World Wide Web Foundation, Feminist Internet and Craig Walker. 2021. "Tech Policy Design Lab: Online Gender-Based Violence and Abuse." June. https://assets.website-files.com/617a5f094309b93ce9ab25b9/618c03493db5a4df42017dc8_OGBV_Report_June2021.pdf.