Bridging the gap:

Addressing technology-facilitated gender-based violence in global AI governance

APC
ASSOCIATION FOR PROGRESSIVE COMMUNICATIONS

Contents

Bridging the gap: Addressing technology-facilitated gender-based violence in global AI governance

Author: Daniela Schnidrig[1]

Coordination and review: Verónica Ferrari (APC)

Proofreading: Laura Pérez Carrara

Design and layout: Jaspreet Singh

Cover illustration: Cathy Chen

Publication production: Cathy Chen (APC) and Lori Nordstrom (APC)

Published by the Association for Progressive Communications (APC), 2026

ISBN 978-92-95113-89-3
APC-202602-APC-R-EN-DIGITAL-379

[1] Daniela is a digital rights and policy consultant with over a decade of experience working at the intersection of technology governance, human rights and gender justice. She previously served as co-head of policy and advocacy at Global Partners Digital, where she led work on technology policy and rights-based advocacy.

Acknowledgements

APC would like to thank the following people for the time and expertise they contributed to provide feedback to support the development of this publication: Hija Kamran, Karla Velasco Ramos, Pavitra Ramanujam and Paula Martins.

We would also like to thank Zuzanna Wojciak and Bruna Martins dos Santos from WITNESS for their insights. Finally, APC is grateful to the experts who generously shared their time and expertise during the validation workshop that informed this report.

Executive summary

Artificial intelligence (AI) technologies are rapidly transforming digital ecosystems, creating new opportunities for innovation while also reinforcing existing structural inequalities – particularly those rooted in gender. Among the most urgent and under-regulated consequences of AI deployment is the rise of technology-facilitated gender-based violence (TFGBV), including automated harassment, deepfakes, gendered disinformation and surveillance-enabled abuse.

This research provides a comprehensive analysis of the intersection between AI governance and TFGBV. It examines how international, regional and national policy frameworks engage with gendered risks introduced or exacerbated by AI and identifies critical gaps in regulation, protection and enforcement.

The study draws on four primary methods: a literature review, a comparative analysis of 24 national and seven international AI governance frameworks, a review of global, regional and national TFGBV legal frameworks and norms and a validation workshop with experts from civil society, academia and policy fields. A taxonomy of AI-TFGBV risks was developed and used to assess both the content and scope of relevant frameworks.

Key findings include:

- Fragmented and siloed governance: Most AI frameworks neglect TFGBV, while TFGBV norms fail to incorporate AI-related risks. This leaves victims without protection or recourse and undermines policy coherence.
- Lack of enforceability: References to fairness or inclusion often lack implementation tools, such as audits, indicators or gender-responsive safeguards.
- Underrepresentation and digital exclusion: Women and gender-diverse people – especially from the Global Majority – remain marginalised in AI development, access and governance.
- Emerging national innovation: Countries such as Chile, Canada, Australia and the UK offer promising models that integrate gender-based analysis, human rights impact assessments or binding regulation of deepfakes and platform accountability.

Recommendations are structured across three pillars:

1. AI-specific actions, such as mandating gender-responsive impact assessments and integrating TFGBV as a risk category in AI governance.
2. TFGBV-specific actions, including updating legal definitions of violence to capture AI-enabled harms and strengthening redress mechanisms.
3. Crosscutting actions, including building regulatory capacity, funding feminist innovation, ensuring intersectional data practices and enforcing corporate responsibility throughout the AI lifecycle.

Introduction

Context and relevance

The rapid integration of artificial intelligence (AI) technologies across digital ecosystems has brought both opportunities and profound risks. While AI has the potential to serve people by optimising decision making and fostering innovation, it also reinforces and amplifies existing structural inequalities, particularly those rooted in gender.

One of the most concerning consequences of the technological shifts that AI has brought about is the proliferation and exacerbation of technology-facilitated gender-based violence (TFGBV), defined by the Association for Progressive Communications as “acts of gender-based violence that are committed, abetted or aggravated, in part or fully, by the use of information and communication technologies (ICTs), such as phones, the internet, social media platforms and email.”[2]

AI systems – especially those involving large-scale data processing, such as machine learning, generative models and large language models (LLMs) – can enable new and more sophisticated forms of TFGBV. These include, among others, gendered disinformation campaigns, automated harassment, non-consensual synthetic intimate imagery (deepfakes) and the algorithmic reinforcement of gender stereotypes.[3] In some contexts, including in conflict-affected settings, AI-enabled surveillance and location tracking technologies have been used to facilitate targeted attacks on women and their families.[4] Such harms can be further exacerbated by biased datasets and programming, a lack of transparency, weak regulatory oversight, inadequate and discriminatory consultation processes and the underrepresentation of women and gender-diverse voices (particularly from the Global Majority and marginalised communities) in AI design, development and governance processes.

As these harms intensify, different stakeholders, such as governments, civil society organisations, the private sector, the tech community, academia and multilateral institutions, are grappling with how to develop AI governance frameworks that are inclusive, rights-based and responsive to these gendered risks introduced by AI. These responses, however, have remained largely fragmented and there is a lack of comprehensive and enforceable approaches centred on gender-specific risks. At the same time, most international norms and policy declarations related to TFGBV have not yet adapted to address AI-specific harms.

This research examines that gap, offering a comparative analysis of AI and TFGBV frameworks and assessing the extent to which they address (or fail to address) the gendered risks

[2] https://www.apc.org/es/node/40128

[3] UN Women. (2024). *Artificial intelligence and gender equality. Explainer*. https://www.unwomen.org/en/articles/explainer/artificial-intelligence-and-gender-equality; Fournier-Tombs, E. et al. (2024). *Artificial intelligence and the women, peace and security agenda in South-East Asia*. UN Women Asia and the Pacific. https://asiapacific.unwomen.org/sites/default/files/2024-05/ap-c871-ai-research-report-2024-full.pdf; WITNESS. (2025, 6 March). Deepfakes and Digital Abuse: Dismantling Technology-Facilitated Gender-Based Violence. https://blog.witness.org/2025/03/technology-facilitated-gender-based-violence/

[4] FIRN team. (n/d). Understanding of technology-facilitated gender-based violence beyond social media-centred analysis. https://firn.genderit.org/blog/understanding-technology-facilitated-gender-based-violence-beyond-social-media-centred

introduced by AI.[5] It advocates for inclusive, multistakeholder governance models, where marginalised and underrepresented communities, civil society, technologists and policy makers collectively shape AI and TFGBV responses.

Research questions and objectives

This study is guided by the following research questions:

- What are the key risks, challenges and emerging themes at the intersection of AI and TFGBV, as identified in the existing literature?
- To what extent do international AI frameworks acknowledge and address TFGBV-related harms?
- To what extent do national AI strategies and governance frameworks in selected countries acknowledge and address TFGBV-related harms?
- How do international norms and policy instruments on TFGBV incorporate the risks introduced by AI technologies? And how can gender responsive AI policies and frameworks be used to the same end?
- Are existing global standards equipped to address the evolving nature of TFGBV in the context of AI?
- What policy recommendations can strengthen the integration of TFGBV concerns into AI governance frameworks and ensure that TFGBV-related norms effectively respond to the challenges posed by AI technologies?

Document structure

This document contributes to the emerging body of research on AI and TFGBV by assessing the degree to which international and national policy frameworks and instruments engage with gendered AI harms. It is structured as follows:

- Section 1 details the methodology, sources of information and limitations.
- Section 2 presents a literature review on key themes at the intersection of AI and TFGBV.
- Section 3 analyses selected national and international AI frameworks and their treatment of TFGBV-related themes.
- Section 4 analyses global and national TFGBV legal frameworks and norms and their responsiveness to AI-specific risks.
- Section 5 synthesises crosscutting findings and insights.
- Section 6 offers action-oriented recommendations for strengthening AI and TFGBV frameworks.

The annex includes a list of frameworks reviewed and literature consulted.

[5] While this study focuses on risks, further research is needed to explore how AI could be used to support prevention, detection and redress of TFGBV. Understanding these underdeveloped opportunities (particularly from a feminist, rights-based lens) will be essential for shaping more holistic responses.

Section 1. Methodology

This research adopted a qualitative, interpretive approach to examine the intersections of AI and TFGBV, drawing on interpretive policy analysis approaches that focus on how policy problems, norms and governance responses are constructed and understood within institutional and socio-political contexts.[6] The methodology was structured around four key components: a literature review; a policy analysis of national and international AI governance frameworks; a review of global and regional TFGBV legal frameworks and norms; and an expert consultation workshop to validate and enrich findings.

1. Literature review

The literature review sought to identify and synthesise key themes, issues, debates and gaps in the existing body of knowledge on AI and its intersection with TFGBV. Sources included peer-reviewed journal articles, policy reports, civil society and NGO publications and relevant legal and technical analyses.

The review prioritised:

- Literature published between 2020 and 2025.
- Works explicitly addressing AI in the context of TFGBV or gendered digital harms.
- Global and region-specific perspectives, with attention to intersectional analyses.

The literature review served as the foundational analytical lens for the study and informed the development of a thematic taxonomy. This taxonomy was refined iteratively over the course of the research and used to guide both the framework analysis and the interpretation of gaps and trends. It comprised three broad thematic clusters:

- Structural gendered risks in AI
  - Gender digital divide
  - Underrepresentation of women in the development of AI technologies
  - Data extraction, consent and gendered surveillance
  - Bias and stereotypes.
- TFGBV harms that are enabled or amplified by AI
  - General TFGBV and automated abuse
  - Deepfakes and synthetic content
  - Cyberstalking and AI surveillance
  - Targeted harassment, doxxing and algorithmic amplification
  - Gendered disinformation
  - Militarisation of AI and gendered warfare.

[6] See: Yanow, D. (2000). *Conducting interpretive policy analysis.* Sage Publications, Inc.

- **Governance gaps and challenges**
  - Proliferation and fragmentation of AI governance instruments
  - Insufficient gender specific safeguards in AI governance frameworks
  - TFGBV norms and national laws lagging behind technological risks
  - Lack of coherence, enforcement and multistakeholder coordination
  - Absence of intersectionality.

2. AI governance frameworks review and analysis

The second component focused on assessing AI frameworks – both national and international – for their treatment of TFGBV themes.

At the international level, the study reviewed leading frameworks and policy instruments, including:[7]

- EU AI Act (2024)
- African Union Continental Artificial Intelligence Strategy (2024)
- Council of Europe AI and Human Rights Framework Convention (2024)
- Global Digital Compact (2024)
- G7 Hiroshima Guiding Principles on AI (2023)
- UNESCO Recommendation on the Ethics of Artificial Intelligence (2021)
- OECD AI Principles (2019).

At the national level, this study analysed existing AI government frameworks in a set of 24 countries across six regions: Canada and the United States (North America); Brazil, Chile and the Dominican Republic (South and Central America); Bulgaria, Croatia, Germany, Greece, Ireland, Italy, the Netherlands, Portugal, Serbia, Slovenia and Spain (Europe); Senegal (Africa); Jordan and Saudi Arabia (Middle East); and Australia, China, Japan, Malaysia and Singapore (Asia Pacific).

Countries were selected based on their inclusion in the gender equality thematic area of the Global Index on Responsible AI (GIRAI),[8] which evaluates how countries promote and protect gender equality in the design, development and use of AI systems. The analysis focused specifically on GIRAI's **government frameworks pillar**, which includes "laws, regulations, policies, strategies and/or guidelines adopted by the national or federal government that address the implications of AI with respect to a particular thematic area."[9] This means that all countries reviewed had at least one AI governance framework that makes reference to gender equality.

[7] At the time of writing, the World Summit on the Information Society 20-year review process (WSIS+20) was ongoing. While emerging discussions include references to AI governance and digital public infrastructure, a detailed assessment of WSIS+20 outcomes falls outside the temporal scope of this report and should be considered in future analyses.

[8] https://www.global-index.ai/thematic-areas-Gender-Equality

[9] Global Center on AI Governance. (2025). *Methodology and Conceptual Framework of the Global Index on Responsible AI*. https://drive.google.com/file/d/1LPE9nEoq5oprQEpJLSaXrQPuGV1cyusd/view

While this approach enabled comparative analyses across jurisdictions, it resulted in limited geographic representation from some regions, notably Sub-Saharan Africa and South Asia. To address this without expanding the core country sample, the study also draws on illustrative examples from GIRAI's other pillars – namely, **government actions** (e.g. awareness campaigns or enforcement efforts) and **non-state actors** (e.g. civil society advocacy or research initiatives). These examples are included in the analysis for contextual insight and serve to highlight promising practices from a broader set of geographies. They are not directly compared to national frameworks.

Each national framework was assessed against TFGBV-related themes derived from the literature review. A comparative matrix was developed to visualise explicit thematic references within frameworks, highlighting patterns and gaps across jurisdictions. Broader governance gaps (such as lack of enforcement or intersectionality) were analysed qualitatively in the narrative sections rather than operationalised as matrix categories. In some cases, the analysis was constrained by language barriers or limited access to up-to-date documents; where necessary, unofficial translations were conducted using digital translation tools.

3. TFGBV legal frameworks and norms review and analysis

This component examined global, regional and national legal frameworks and norms related to gender-based violence, with a focus on whether and how they address specific risks introduced or exacerbated by AI technologies. The review covered:

- International instruments
  - United Nations resolutions and declarations (e.g. A/RES/78/213, A/RES/75/161)
  - Human Rights Council resolutions (e.g. HRC/38/5, HRC/53/29).
- Regional frameworks
  - Istanbul Convention
  - Maputo Protocol
  - Belem do Pará Convention
  - 2024 EU Directive on Combating Violence against Women
  - African Commission Resolutions 522 (LXXII) 2022 and 591 (LXXX) 2024.
- National legal frameworks and policy guidance
  - Australian Human Rights Commission Report on TFGBV and Human Rights Tool
  - UK Online Safety Act.
- Non-state normative frameworks
  - Feminist digital rights principles (e.g. Feminist Principles of the Internet)
  - Feminist Principles for Including Gender in the Global Digital Compact.

The analysis assessed whether these instruments recognise AI and digital technologies as enablers of TFGBV and how they address emerging harms such as deepfakes, algorithmic bias and surveillance.

4. Expert consultation workshop

To complement the desk-based analysis, a participatory consultation workshop was convened with experts working at the intersection of AI policy, gender justice and digital rights.

The workshop aimed to:

- Validate key findings from the literature and policy reviews.
- Identify additional risks, examples and emerging areas of concern not fully captured in formal frameworks.
- Gather feedback on the thematic taxonomy and proposed recommendations.
- Ensure the research is grounded in diverse regional and experiential perspectives.

Participants included civil society actors, researchers and advocates from different regions and sectors. Their inputs were instrumental in refining the findings and strengthening the recommendations.

Limitations

This study acknowledges the following limitations:

- Language constraints and access barriers affected the ability to review some frameworks in their original form.
- The rapidly evolving policy landscape may result in recent updates not being captured at the time of drafting this document (December 2025).
- The qualitative nature of the comparative matrix prioritised analytical depth over exhaustive coverage.
- The analysis focused on the content of frameworks and did not assess the extent of their implementation or enforcement in practice

Section 2. Literature review

The intersection of AI and TFGBV has become a pressing area of concern for gender justice discourse, particularly in the digital rights field.[10] Despite this growing attention, it remains underexplored in AI policy and governance research. A growing body of literature, however, points to multiple ways in which AI systems not only reflect but also reproduce and exacerbate gendered harms.

This section synthesises insights from academic studies, human rights reports and civil society analyses into three thematic clusters: structural gendered risks in AI; TFGBV harms enabled or amplified by AI; and governance gaps and challenges in both AI and TFGBV frameworks.

1. Structural gendered risks in AI

AI systems are shaped by the societal conditions in which they are developed and deployed. As such, structural gendered risks in AI reflect broader inequalities in digital access, data governance, institutional power and economic opportunity. These risks disproportionately affect women and girls, particularly those who also face marginalisation on the basis of race, class, disability, migration status, caste or sexuality. An intersectional lens is therefore essential for understanding how AI systems can reproduce and deepen existing social hierarchies.

1.1. Gender digital divide

Globally, women (particularly women in Global Majority countries) face reduced access to digital tools, connectivity and technical education. This digital divide manifests across multiple dimensions, including physical access to devices and the internet, digital skills and capacity and structural exclusion from digital innovation and governance processes.[11] These barriers limit both their participation in digital economies and their representation in shaping AI development.[12] This divide is both a cause and a consequence of digital exclusion driven by structural inequality and cultural norms, including gendered expectations, safety concerns and restricted mobility.[13] As the International Telecommunications Union notes, women's internet use lags behind that of men globally, with the gap especially stark in low-income countries where access is further shaped by social and cultural barriers.[14] This is exacerbated by socioeconomic inequality, rurality and other intersecting marginalisations, which international normative frameworks have recognised as compounding digital exclusion.[15]

As Kelle Howson notes:

[10] Naim, N. (2023, 28 July). Women in the era of artificial intelligence: Increased targeting and growing challenges. *GenderIT.org*. https://genderit.org/articles/women-era-artificial-intelligence-increased-targeting-and-growing-challenges

[11] Azcona, G. et al. (2023). *Progress on the Sustainable Development Goals: The gender snapshot 2023*. UN Women. https://www.unwomen.org/sites/default/files/2023-09/progress-on-the-sustainable-development-goals-the-gender-snapshot-2023-en.pdf

[12] UNESCO. (2020). *Artificial intelligence and gender equality: Key findings of UNESCO's Global Dialogue*. https://unesdoc.unesco.org/ark:/48223/pf0000374174; and UNESCO. (2023, 7 March). International Women's Day: New factsheet highlights gender disparities in innovation and technology. https://www.unesco.org/en/articles/international-womens-day-new-factsheet-highlights-gender-disparities-innovation-and-technology

[13] GSMA. (2025). *The Mobile Gender Gap Report 2025*. https://www.pta.gov.pk/assets/media/2025-05-20-The-Mobile-Gender-Gap-Report-2025.pdf

[14] International Telecommunication Union. (2024). *Facts and Figures 2024: The gender digital divide*. https://www.itu.int/itu-d/reports/statistics/2024/11/10/ff24-the-gender-digital-divide/

[15] UNESCO. (2022). *Recommendation on the Ethics of Artificial Intelligence, adopted on 23 November 2021*, articles 87-89. https://unesdoc.unesco.org/ark:/48223/pf0000381137

> In the context of the rise of AI-based technologies, victims of the digital divide are still profoundly impacted by the use of AI. They may still be subject to the collection of their data for the purposes of training AI; they are still likely to interact (possibly without their knowledge) with AI-based technologies (for instance facial recognition) in public places; AI-based technologies will still be used to make decisions about them or which affect their lives.[16]

A deeper understanding of these forms of structural exclusion is essential to designing equitable and rights-based governance of AI systems.

In many contexts, this exclusion is compounded by political and economic barriers. Women and gender diverse individuals are overrepresented in informal labour sectors and underrepresented in formal economies, which limits their ability to benefit from or shape digital transformation.[17] As AI systems become more embedded in digital platforms – such as through algorithmic content curation, searches, biometric verification or service delivery – individuals increasingly interact with AI-mediated infrastructures, often without adequate awareness, safeguards or recourse.[18]

These dynamics have material consequences. Inadequate protections from TFGBV can result in reputational harm, job loss or withdrawal from online and economic spaces altogether.[19] A political economy lens (i.e. an approach that considers how economic and governance structures influence access to and control over AI systems)[20] is essential to understanding how the digital divide not only limits participation but also exposes women and marginalised groups to heightened risks in AI-mediated environments.

1.2. Underrepresentation of women in the development of AI technologies

The AI field remains heavily male dominated, with women and gender diverse people significantly underrepresented in technical, leadership and policy roles. As documented by the World Economic Forum, only 22% of AI professionals globally are women.[21]

This lack of diversity influences how systems are developed and governed, skewing design processes, risk prioritisation and definitions of fairness. It contributes to the reproduction of patriarchal norms in AI systems and reinforces exclusion in decision-making spaces.[22]

1.3. Data extraction, consent and gendered surveillance

[16] Howson, K. (2025). *Human Rights and AI: A Global Index on Responsible AI. Analytical Report Series. Brief 6: Gender Equality*. Global Center on AI Governance. https://genderequalitybrief6.tiiny.site/

[17] Azcona, G. et al. (2023). Op. cit.

[18] Pavel, V. (2022). *Rethinking data and rebalancing digital power*. Ada Lovelace Institute. https://www.adalovelaceinstitute.org/wp-content/uploads/2022/11/Ada-Lovelace-Institute-Rethinking-data-and-rebalancing-digital-power-FINAL.pdf

[19] Amnesty International. (2025). *Human rights implications of technology-facilitated gender-based violence. Submission to the Human Rights Council Advisory Committee.* https://www.amnesty.org/en/documents/ior40/9284/2025/en/

[20] See: Gurumurthy, A., & Barthur, D. (2023). *Reframing AI Governance through a Political Economy Lens*. IT for Change. https://itforchange.net/reframing-ai-governance-through-a-political-economy-lens; and Timcke, S., & Makumbirofa, S. (2024, 22 May). Five things to know about the political economy of AI. *Research ICT Africa*. https://researchictafrica.net/2024/05/22/five-things-to-know-about-the-political-economy-of-ai/

[21] Ramos, G. (2022, 22 August). Why we must act now to close the gender gap in AI. *World Economic Forum*. https://www.weforum.org/stories/2022/08/why-we-must-act-now-to-close-the-gender-gap-in-ai/

[22] UNESCO. (2020). Op. cit.

Beyond issues of representation, AI systems rely on large-scale data extraction, often non-consensual, opaque and disproportionately affecting marginalised and underrepresented communities.[23]

Feminists and decolonial critiques have framed these practices as forms of digital expropriation or data colonialism, where individuals' data is taken without their knowledge and repurposed across systems for commercial or state use.[24] This includes:

- Function creep: where data collected for one purpose (e.g. health care) is reused for surveillance or immigration control.[25]
- Reproductive surveillance: instances where apps have handed over fertility or pregnancy data to authorities.[26]
- Platform opacity: where algorithmic profiling and content moderation are built on training datasets gathered from users without meaningful consent.[27]

These practices disproportionately harm women, LGBTQIA+ people and racialised or poor communities,[28] who are more likely to be subject to intensified surveillance.[29]

1.4. Bias and stereotypes

AI systems are often presented as neutral or objective, but they are built on foundations that reflect societal bias and structural inequality. These biases manifest in different ways:

- **Biased training datasets:** AI models are trained on datasets that are not neutral and often reflect and amplify historical and societal gender norms.[30] Many datasets used to train AI systems lack meaningful demographic diversity, making women, people of colour, LGBTQIA+ individuals and Global Majority users invisible or underrepresented. For instance, facial recognition systems have been shown to misclassify darker-skinned women at significantly higher rates than lighter-skinned men, as demonstrated by the Gender Shades study.[31]
- **Bias introduced during annotation and training:** Even when datasets are more representative, human annotators bring their own societal prejudices, which are often unaccounted for. This can reproduce stereotypes in how content is categorised, moderated or ranked.[32]
- **Design choices that reflect gender norms:** AI assistants are often given female voices and

[23] Amnesty International. (2019). *Surveillance giants: How the business model of Google and Facebook threatens human rights*. https://www.amnesty.org/en/documents/pol30/1404/2019/en/

[24] Das, D. (2025, 20 May). The Binary Is Glitchy: Platform Accountability Through a Decolonial Queer Lens. *GenderIT.org*. https://genderit.org/feminist-talk/binary-glitchy-platform-accountability-through-decolonial-queer-lens

[25] Kornweitz, A. (2021, 4 August). A New AI Lexicon: Function Creep. *AI Now Institute*. https://ainowinstitute.org/publications/collection/a-new-ai-lexicon-function-creep

[26] Felsberger, S. (2025). *The High Stakes of Tracking Menstruation*. MCTD Cambridge. https://www.mctd.ac.uk/wp-content/uploads/2025/06/The-High-Stakes-of-Tracking-Menstruation.pdf

[27] Cannataci, J. A. (2021). *Artificial intelligence and privacy, and children's privacy. Report of the Special Rapporteur on the right to privacy, Joseph A. Cannataci*. UN Human Rights Council. (A/HRC/46/37). https://docs.un.org/en/A/HRC/46/37

[28] Amnesty International. (2019). Op. cit.; Khan, S. (2017, 21 November). Surveillance as a Feminist Issue. *Privacy International*. https://privacyinternational.org/news-analysis/3376/surveillance-feminist-issue

[29] See, for example: Ali Hussen, D. (2023, 26 March). "Dystopian" surveillance "disproportionately targets young, female and minority workers." *The Guardian*. https://www.theguardian.com/global-development/2023/mar/26/dystopian-surveillance-disproportionately-targets-young-female-minority-workers-ippr-report ; Das, S. (2024, 27 January). Facial recognition cameras in supermarkets "targeted at poor areas" in England. *The Guardian*. https://www.theguardian.com/uk-news/2024/jan/27/facial-recognition-cameras-in-supermarkets-targeted-at-poor-areas-in-england

[30] Howson, K. (2025). Op. cit.

[31] Buolamwini, J., & Gebru, T. (2018). Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. *Proceedings of Machine Learning Research, 81*, 1-15. https://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf

[32] Haliburton, L., Leusmann, J., Welsch, R. et al. (2025). Uncovering labeler bias in machine learning annotation tasks. *AI and Ethics, 5*, 2515-2528. https://doi.org/10.1007/s43681-024-00572-w

names and are programmed to respond in submissive, accommodating or flirtatious ways, reinforcing gender stereotypes about care giving and emotional labour.[33]

- **Content moderation failures:** Biases extend to content moderation, where non-English languages and regional dialects are poorly supported. This disproportionately affects women and gender-diverse users in the Global Majority. Whistleblower testimony has shown how a major platform – Meta – consistently underinvested in moderation infrastructure for non-Western contexts, undermining safety and accountability.[34]
- **Compounded discrimination:** As Pollicy highlights,[35] intersecting identity markers – such as race, gender identity, class or disability – directly influence people's experiences with AI. This often results in gendered racial discrimination, where, for example, black women face compounded harms not experienced by white women or black men. AI systems are frequently biased against all women, but especially those at intersections of multiple forms of marginalisation.[36]

UNESCO further underscores that biases can be introduced at any stage of AI development, from design and modelling decisions to data collection processing and deployment.[37] Left unchecked, these biases can entrench inequality at scale.

2. TFGBV harms enabled or amplified by AI

AI has both introduced new forms of TFGBV, as well as intensified the scale, reach and automation of existing ones.[38] The literature identifies several specific threats:

2.1. General TFGBV and automated abuse

Machine learning models used in moderation and platform curation can inadvertently enable abuse by failing to detect harmful content or disproportionately silencing women and marginalised voices.[39] Additionally, automated and AI-enabled tools are increasingly being used to target women in public life with high-volume coordinated harassment, disinformation and defamation.[40] These tools enable harassment at a scale and speed that would be infeasible through manual efforts alone, exponentially increasing the volume and persistence of TFGBV.

[33] West, M., Kraut, R., & Ei Chew, H. (2019). *"I'd blush if I could." Closing gender divides in digital skills through education*. UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000367416

[34] Popli, N. (2021, 26 October). The 5 most important revelations from the "Facebook Papers". *Time*. https://time.com/6110234/facebook-papers-testimony-explained/

[35] Nabulega, S., Achieng, G., & Borokini, F. (2021). *Engendering AI: A Gender and Ethics Perspective on Artificial Intelligence in Africa. Pollicy*. https://pollicy.org/resource/engendering-artificial-intelligence/

[36] Ashwini K. P. (2024). *Racial discrimination and emerging digital technologies: A human rights analysis. Report of the Special Rapporteur on contemporary forms of racism, racial discrimination, xenophobia and related intolerance*. UN Human Rights Council. (A/HRC/44/57). https://docs.un.org/en/A/HRC/44/57

[37] UNESCO-IRCAI. (2024). *Challenging Systematic Prejudices: An Investigation into Gender Bias in Large Language Models*. https://unesdoc.unesco.org/ark:/48223/pf0000388971

[38] UNESCO. (2023). *"Your opinion doesn't matter, anyway." Exposing Technology-Facilitated Gender-based Violence in an era of Generative AI*. https://www.unesco.org/sites/default/files/medias/fichiers/2024/04/Ai-livre-EN-web.pdf

[39] See, for example, studies documenting how algorithmic moderation systems disproportionately flag and remove posts from marginalised users, leading to curtailing their online expression. Lee, C. et al. (2024). People who share encounters with racism are silenced online by humans and machines, but a guideline-reframing intervention holds promise. *Proceedings of the National Academy of Sciences, 121*(38). https://doi.org/10.1073/pnas.2322764121

[40] UN Human Rights Council. (2018). *Resolution 38/5. Accelerating efforts to eliminate violence against women and girls: preventing and responding to violence against women and girls in digital contexts*. https://docs.un.org/en/A/HRC/RES/38/5; UN Commission on the Status of Women. (2023). *CSW67 Agreed Conclusions. Innovation and technological change, and education in the digital age for achieving gender equality*. https://www.unwomen.org/sites/default/files/2023-03/CSW67_Agreed%20Conclusions_Advance%20Unedited%20Version_20%20March%202023.pdf; UN General Assembly. (2020). *Resolution 75/161. Intensification of efforts to prevent and eliminate all forms of violence against women and girls*. https://docs.un.org/en/A/RES/75/161

In many contexts, the consequences of technology-facilitated gender-based violence extend beyond digital spaces and have profound effects on women's and girls' daily lives and economic rights. Targeted harassment, deepfakes and other AI-enabled forms of abuse have led to instances of women being forced to leave their jobs, relocate from their communities or withdraw from public life altogether.[41] These impacts are particularly acute in environments where patriarchal social norms attach high stigma to perceived reputational damage, reinforcing cycles of exclusion and economic marginalisation. As such, TFGBV must be understood not only as a violation of dignity and safety, but also as a driver of material harm and economic displacement.

In addition to private sector deployment, AI technologies are increasingly used by governments, for example, within digital public infrastructure, biometric identification systems and social protection schemes. These uses raise unique TFGBV concerns, particularly in settings where state-led data collection, profiling or surveillance may disproportionately impact women and marginalised communities.[42] The use of AI by authorities in migration control, welfare eligibility or predictive policing can reproduce and exacerbate gendered harms, especially in the absence of oversight mechanisms. At the same time, public sector AI deployment offers an opportunity: if grounded in gender-responsive design and robust accountability standards, government-led systems could set benchmarks for inclusive, rights-based governance.

2.2. Deepfakes and synthetic content

Generative AI has facilitated the creation of hyper-realistic non-consensual intimate images (i.e. deepfake pornography), disproportionately targeting women and girls, including women in public life such as journalists and activists.[43] These images are often deployed for blackmail, humiliation or public discrediting. Legal protections remain uneven and, in some cases, inadequate to provide timely or effective recourse to victims.[44]

The specific needs of local context compound these challenges. As Anmol Irfan highlights, in the context of more conservative countries, a deepfake of a woman in a compromising situation can lead to extreme consequences, including serious threats to safety and, in some contexts, lethal violence.[45]

Furthermore, another newly emerging threat are compositional deepfakes, defined as "the combination of multiple fabricated media sources that seem disparate but corroborate each other, leading to synthetic histories that are very believable."[46]

[41] See: Amnesty International. (n/d). *Online violence*. https://www.amnesty.org/en/what-we-do/technology/online-violence/; and Enock, F. E. et al. (2025). Gendered Inequalities in Online Harms: Fear, Safety Work, and Online Participation. *arXivLabs*. https://arxiv.org/abs/2403.19037

[42] See: Organisation for Economic Co-operation and Development. (2025a). *Governing with Artificial Intelligence. The State of Play and Way Forward in Core Government Functions*. OECD Publishing. https://www.oecd.org/en/publications/governing-with-artificial-intelligence_795de142-en.html; Jansen Reventlow, N. (2021, 29 May). How Artificial Intelligence Impacts Marginalised Groups. https://digitalfreedomfund.org/how-artificial-intelligence-impacts-marginalised-groups/; and Ada Lovelace Institute. (2022). *Countermeasures: The need for new legislation to govern biometric technologies in the UK*. https://www.adalovelaceinstitute.org/wp-content/uploads/2025/06/Ada-Lovelace-Institute-Countermeasures-020625.pdf

[43] See: Security Hero. (2023). *2023 State of Deepfakes. Realities, Threats and Impact*. https://www.securityhero.io/state-of-deepfakes/; Gibson, K. (2025, 9 January). Gender bias, AI, and deepfakes are promoting misogyny online. https://blogs.lse.ac.uk/wps/2025/01/09/gender-bias-ai-and-deepfakes-are-promoting-misogyny-online/; and Kira, B. (2024, 3 June). Deepfakes, the Weaponisation of AI Against Women and Possible Solutions. https://verfassungsblog.de/deepfakes-ncid-ai-regulation/

[44] See: reference to deepfake imagery in paragraph 19 of EU Directive 2024/1385. European Union. (2024). *Directive (EU) 2024/1385 of the European Parliament and Council of 14 May 2024 on combating violence against women and domestic violence*. https://eur-lex.europa.eu/eli/dir/2024/1385/oj/eng; and section 66A of the UK Sexual Offences Act 2003 as introduced by the UK Online Safety Act. UK Government. (2023). *Online Safety Act*. https://www.legislation.gov.uk/ukpga/2023/50; and UK Government. (2003). *Sexual Offences Act*. https://www.legislation.gov.uk/ukpga/2003/42/section/66A

[45] Irfan, A. (2025, 14 March). Generative AI and Deepfakes: is there a way forward? *Gender IT.org*. https://genderit.org/feminist-talk/generative-ai-and-deepfakes-there-way-forward

[46] UNESCO. (2023). Op. cit.

2.3. Cyberstalking and AI surveillance

Tools such as facial recognition, stalkerware[47] and geolocation AI enable more invasive, persistent and often untraceable forms of surveillance, especially in the context of intimate partner violence.[48] These practices are often underpinned by gendered power dynamics and control.[49]

These surveillance tools have a disproportionate impact on marginalised and underrepresented communities, including racialised women and LGBTQIA+ people. For instance, predictive policing systems that rely on historical criminal data and police activity have been shown to perpetuate and reinforce racial biases in policing, disproportionately directing enforcement attention towards racialised communities due to skewed historical data and discriminatory policing practices.[50] Similarly, facial recognition systems routinely misidentify or fail to accurately recognise women of colour and trans individuals, resulting in misgendering, wrongful suspicion or denial of services.[51]

2.4. Targeted harassment, doxxing and algorithmic amplification

Not only can algorithms fail to suppress gendered harassment and hate speech, they can also actively amplify it. Research and whistleblower accounts have shown that platform algorithms often prioritise content that drives engagement, including hateful material, even when it poses real world risks.[52] Cases from Palestine, Myanmar and India illustrate that platforms continue to amplify harmful content despite knowing its effects, due to profit-driven models that reward virality.[53]

Activists (particularly those advocating for gender justice) are often disproportionately targeted through coordinated online attacks that gain reach due to algorithmic amplification.[54] These harms are compounded by automated content moderation tools that are ill equipped to detect abuse in local or non-English languages, especially when such abuse is nuanced, coded or culturally specific.[55]

[47] Defined by the Electronic Frontier Foundation as "commercially-available apps that can be covertly installed on another person's device for the purpose of monitoring their activity without their knowledge or consent." Galperin, E. (2021, 25 December). Stalkerware: 2021 in Review. *Electronic Frontier Foundation*. https://www.eff.org/deeplinks/2021/12/stalkerware-2021-review

[48] Šimonović, D. (2018). *Report of the Special Rapporteur on violence against women, its causes and consequences on online violence against women and girls from a human rights perspective*. UN Human Rights Council. (A/HRC/38/47). https://docs.un.org/en/A/HRC/38/47

[49] UN Commission on the Status of Women. (2023). Op. cit.

[50] O'Donnell, R. M. (2019). Challenging Racist Predictive Policing Algorithms under the Equal Protection Clause. *NYU Law Review*, *94*(3). https://nyulawreview.org/wp-content/uploads/2019/06/NYULawReview-94-3-ODonnell.pdf

[51] Millar, M. (2019, 30 October). Facial recognition technology struggles to see past gender binary. *Reuters*. https://www.reuters.com/article/world/facial-recognition-technology-struggles-to-see-past-gender-binary-idUSKBN1X92OC/; Buolamwini, J., & Gebru, T. (2018). Op. cit.

[52] Popli, N. (2021, 26 October). Op. cit.

[53] Amnesty International. (2022, 29 September). Myanmar: Facebook's systems promoted violence against Rohingya; Meta owes reparations – new report. https://www.amnesty.org/en/latest/news/2022/09/myanmar-facebooks-systems-promoted-violence-against-rohingya-meta-owes-reparations-new-report/; Amnesty International. (2025, 17 February). Meta's new content policies risk fueling more mass violence and genocide. https://www.amnesty.org/en/latest/news/2025/02/meta-new-policy-changes/; Shtaya, M. (2023, 13 November). How Meta's platforms normalize anti-Palestinian racism. *Middle East Institute*. https://www.mei.edu/publications/how-metas-platforms-normalize-anti-palestinian-racism

[54] United Nations. (2024). *Global Digital Compact*. https://www.un.org/global-digital-compact/en; UNESCO. (2022). Op. cit., paragraph 88; Association for Progressive Communications. (2016, 19 August). Feminist Principles of the Internet. https://www.apc.org/en/pubs/feminist-principles-internet-version-20

[55] See: Amnesty International. (2025). Op. cit.; and Posetti, J., & Shabbir, N. (2022). *The Chilling: A global study of online violence against women journalists. International Center for Journalists*. https://www.icfj.org/sites/default/files/2023-02/ICFJ%20Unesco_TheChilling_OnlineViolence.pdf

2.5. Gendered disinformation

AI tools can be weaponised to spread manipulated narratives that discredit women politicians, academics and human rights defenders. These campaigns often intersect with racist, transphobic or xenophobic – and in certain contexts, religious or caste-based – content aimed at undermining credibility, inciting harassment and deterring political participation.[56]

As framed by APC, gendered disinformation is not just false information; it is intentionally structured to target women and gender-diverse people through coordinated attacks that exploit existing power imbalances and reinforce stereotypes.[57] These messages are more effective when they weaponise sexuality, motherhood, morality or religious stigma. APC shows how disinformation campaigns often target marginalised women – including racial minorities, religious minorities and caste-oppressed groups – seeking to silence their voices in public life.[58] These dynamics make disinformation not merely a digital phenomenon, but a tool of violence and exclusion that disproportionately affects those already on society's margins.

2.6. Militarisation of AI and gendered warfare

AI-enabled technologies are increasingly deployed in military and conflict contexts, including for surveillance, predictive threat analysis, autonomous weapons systems and drone targeting. These systems are often presented as improving accuracy or operational efficiency, yet they can undermine human rights because they displace human judgement and may act unpredictably when selecting and engaging targets without meaningful human control.[59] Research and advocacy from feminist peace organisations highlight that the deployment of autonomous weapons and other AI-enabled military systems risks reinforcing existing patterns of gendered and racialised violence by embedding and automating structures of discrimination and exclusion.[60]

In humanitarian and conflict settings, AI-generated and AI-distributed disinformation can fuel escalation, destabilise communities and increase risks of conflict-related sexual violence.[61] Feminist digital rights advocates have documented how gendered disinformation is used to discredit women peacebuilders and human rights defenders, exploit ethnic divisions and legitimise gender-based harassment.[62]

[56] UN Commission on the Status of Women. (2023). Op. cit., paragraph 45.

[57] Martins, P. (2024). *Placing "gender" in disinformation*. Association for Progressive Communications. https://www.apc.org/sites/default/files/genderDisinformation.pdf

[58] Ibid. Examples include coordinated disinformation campaigns targeting women politicians from minority groups and church-based women's networks in Asia and Africa.

[59] See: Stauffer, B. (2025). *A Hazard to Human Rights: Autonomous Weapons Systems and Digital Decision-Making*. Human Rights Watch. https://www.hrw.org/report/2025/04/28/a-hazard-to-human-rights/autonomous-weapons-systems-and-digital-decision-making; and the Stop Killer Robots Campaign at https://www.stopkillerrobots.org/military-and-killer-robots/.

[60] Acheson, R. (2025). *WILPF Inputs to the UN Secretary General's Report on Artificial Intelligence in the Military Domain*. Women's International League for Peace and Freedom. https://docs-library.unoda.org/General_Assembly_First_Committee_-Eightieth_session_(2025)/79-239-WILPF-EN.pdf

[61] Ward, J., Spencer, S., & Kalsi, K. (2023). *Gender-Based Violence and Artificial Intelligence (AI): Opportunities and Risks for Women and Girls in Humanitarian Settings*. Social Development Direct. https://www.sddirect.org.uk/sites/default/files/2023-10/GBV%20AoR%20HD%202023%20GBV%20and%20AI%20final%20.pdf

[62] Martins, P. (2024). Op. cit.

Facial recognition and location-tracking systems have also been used to identify and target residential areas, resulting in strikes that disproportionately affect civilians, including women and children. For example, in Gaza, the Israeli military reportedly used an AI system called Lavender to generate live strike lists with minimal human oversight.[63]

The deployment of autonomous weapons and AI-powered threat detection also raises urgent gender and accountability concerns. Research has shown that algorithmic decision-making in warfare may embed gender bias, overlook culturally specific patterns of harm and exclude women's lived experiences from both threat modelling and peacekeeping frameworks.[64]

Finally, TFGBV and AI-enabled harms can have a profound chilling effect, curbing not only free expression but also broader participation in public life. For women, queer and gender-diverse individuals, the risks of exposure to non-consensual deepfakes, automated harassment or biased decision systems can lead to withdrawal from online platforms, civic spaces or leadership roles.[65] This retreat results not only in psychological or reputational harm, but also in the systemic erasure of marginalised voices from public discourse and decision making. The absence of marginalised individuals from public spaces (driven by fear, distrust or algorithmic exclusion) should be recognised as a harm in itself.[66]

3. Governance gaps and challenges

The literature identifies several critical governance gaps across both AI and TFGBV frameworks, which are outlined below.

3.1. Proliferation and fragmentation of AI governance instruments

The growing number of AI strategies, ethical principles and soft-law initiatives at national, regional and global levels has led to a fragmented governance landscape.[67] This proliferation can result in weak interoperability between standards and confusion around which frameworks take precedence. Furthermore, while many frameworks claim to be rights-based, few offer concrete pathways for implementation, resourcing or gender-responsive accountability.

3.2. Insufficient gender specific safeguards in AI governance frameworks

Most AI governance frameworks refer generically to principles like fairness, inclusivity or non-discrimination, while offering limited guidance on how these principles should be operationalised in practice.[68] Few frameworks articulate specific obligations relating to TFGBV or intersectional gender impacts As a result, very few AI governance instruments meaningfully engage with how AI systems can enable, intensify or institutionalise gender-based violence.[69]

[63] Human Rights Watch. (2024, 10 September). Questions and Answers: Israeli Military's Use of Digital Tools in Gaza. https://www.hrw.org/news/2024/09/10/questions-and-answers-israeli-militarys-use-digital-tools-gaza#_What_is_%E2%80%9CLavender%E2%80%9D

[64] Chandler, K. (2021). *Does Military AI have Gender? Understanding bias and promoting ethical approaches in military applications of AI*. United Nations Institute for Disarmament Research. https://unidir.org/wp-content/uploads/2023/05/UNIDIR_Does_Military_AI_Have_Gender.pdf

[65] See: WITNESS. (2025, 6 March). Op. cit.; European Institute for Gender Equality. (2022). *Combating Cyber Violence against Women and Girls*. https://eige.europa.eu/publications-resources/publications/combating-cyber-violence-against-women-and-girls?language_content_entity=en; and Posetti, J., & Shabbir, N. (2022). Op. cit.

[66] Ibid.

[67] See: United Nations AI Advisory Body. (2024). *Governing AI for Humanity: Final report*. https://www.un.org/sites/un2.un.org/files/governing_ai_for_humanity_final_report_en.pdf

[68] Mittelstadt, B. (2019). Principles alone cannot guarantee ethical AI, *Nature Machine Intelligence, 1*(11), 501-507. https://www.researchgate.net/publication/337015694_Principles_alone_cannot_guarantee_ethical_AI

[69] WITNESS. (2025, 6 March). Op. cit.

3.3. TFGBV norms and national laws lagging behind technological risks

While regional legal instruments like the Istanbul Convention and recent UN resolutions (e.g. UN General Assembly Resolution 75/161) recognise digital violence, most still lack the language and enforcement tools necessary to address AI-generated harms, such as deepfakes, automated abuse and algorithmic amplification. National legislative responses are also struggling to keep pace. As AI technologies evolve rapidly, many governments remain stuck in reactive regulatory modes,[70] addressing harms only as they emerge, rather than establishing proactive, gender-responsive guardrails from a gendered perspective.

3.4. Lack of coherence, enforcement and multistakeholder coordination

AI governance and gender equality policy can often operate in silos, leading to poor coordination, weak enforcement and limited cross-sector accountability. As Maral Niazi argues, despite the proliferation of guiding principles, many AI governance frameworks – particularly at the international level – lack robust enforcement mechanisms for non-compliance.[71]

Multistakeholder coordination is needed to avoid piecemeal responses and reinforce intersectional accountability.

3.5. Absence of intersectionality

Many AI and TFGBV frameworks and studies fail to apply an intersectional lens that captures how gendered harms are shaped by race, disability, sexuality, migration status, religion or caste, resulting in the marginalisation of already vulnerable communities.[72] This lack of intersectionality limits the ability of governance frameworks to anticipate and address the compounded impacts on marginalised communities.[73]

3.6. Political economy and global asymmetries shape AI governance

The distribution of TFGBV risks is deeply shaped by underlying political and economic structures. In many low- and middle-income countries, narratives around digital development and foreign investment have facilitated or legitimised extractive partnerships with large technology firms, including the expansion of data infrastructure without adequate safeguards. These dynamics often reinforce existing inequalities, leaving local communities with limited benefits and disproportionate exposure to harm.[74]

Together, these themes frame the foundation for the analysis that follows in sections 3 and 4. The literature points to an **urgent need for coherent, enforceable and gender-responsive governance approaches, but it is crucial that these are also intersectional, inclusive**

[70] See: Organisation for Economic Co-operation and Development. (2025b). *Steering AI's Future: Strategies for Anticipatory Governance. OECD Artificial Intelligence Papers No. 32*. OECD Publishing. https://www.oecd.org/content/dam/oecd/en/publications/reports/2025/02/steering-ai-s-future_70e4a856/5480ff0a-en.pdf

[71] Niazi, M. (2024). *How do current AI regulations shape the global governance framework? Digital Policy Hub – Working Paper*. Centre for International Governance Innovation. https://www.cigionline.org/static/documents/DPH-paper-Maral_Niazi.pdf

[72] Nabulega, S., Achieng, G., & Borokini, F. (2021). Op. cit.

[73] UNESCO. (2020). Op. cit.

[74] As raised during the November 2025 consultation workshop. See also: Salami, A. O. (2024). Artificial intelligence, digital colonialism, and the implications for Africa's future development. *Data & Policy, 6*. https://doi.org/10.1017/dap.2024.75; Ndemo, B. (2024, 8 August). Addressing digital colonialism: A path to equitable data governance. *UNESCO Inclusive Policy Lab*. https://en.unesco.org/inclusivepolicylab/analytics/addressing-digital-colonialism-path-equitable-data-governance; and Timcke, S., & Makumbirofa, S. (2024, 22 May). Op. cit.

and representative. By inclusive and representative we refer to participatory processes that meaningfully engage a diverse range of stakeholders throughout the policy lifecycle, particularly those from marginalised and underrepresented communities.[75] Governance should not simply protect people – it should empower them. This includes giving marginalised and underrepresented communities the ability to shape how technologies are designed, deployed and regulated, including decisions about how their data is used, by whom and for what purpose. Ensuring **agency, autonomy and meaningful consent** must be central to AI and TFGBV governance reform and not treated as afterthoughts.[76]

[75] This framing draws on resources such as: Global Partners Digital. (2023). *Inclusive Cyber Norms Toolkit*. https://www.gp-digital.org/wp-content/uploads/2023/08/Inclusive-Cyber-Norms-Toolkit_GPD.pdf

[76] See Principle 2 (Consent) of the Feminist Principles of the Internet, which emphasises autonomy and meaningful participation in digital decision-making. Association for Progressive Communications. (2016, 19 August). Op. cit.

Section 3. AI frameworks analysis

This section reviews and analyses how selected national and international AI governance frameworks address issues related to gender inequality and TFGBV. Drawing from a comparative assessment of 24 national frameworks across six global regions, alongside five key international instruments, the analysis evaluates both the depth and specificity of engagement with gender-related risks, particularly those linked to AI-enabled TFGBV.

The findings are structured around two levels of analysis:

1. International frameworks: global and regional instruments that aim to guide or regulate ethical and responsible AI development.
2. National frameworks: country-level strategies, laws and policies that govern AI implementation and oversight.

The frameworks were assessed against the taxonomy presented in the previous section, covering **structural gendered risks** (e.g. gender digital divide, underrepresentation, bias), **TFGBV harms enabled or amplified by AI** (e.g. deepfakes, surveillance, disinformation) and **governance gaps and challenges** (e.g. lack of redress and accountability).

International AI frameworks

This section reviews a selected set of these instruments, chosen for their geopolitical relevance, normative influence and current or proposed role in shaping AI regulation. These frameworks differ in their legal status (ranging from binding treaties to voluntary guidelines), geographic reach and thematic emphasis. Some aim to shape ethical norms (e.g. UNESCO Recommendation, OECD AI Principles), others propose enforceable regulation (e.g. EU AI Act) and some are intended to align policy across multilateral blocs (e.g. G7 Hiroshima Principles, AU Strategy), resulting in differing levels of obligation and accountability.

While many frameworks reference values like fairness or non-discrimination, **their treatment of gender-specific risks and TFGBV varies significantly in scope, specificity and enforceability**. The analysis below explores whether and how these frameworks acknowledge TFGBV-related harms and whether they offer meaningful pathways for addressing them.

Global frameworks

UNESCO Recommendation on the Ethics of Artificial Intelligence (2021)

Status: Non-binding (soft law – recommendation)

The UNESCO Recommendation is a voluntary, rights-based ethical framework adopted by 193 member states. It explicitly links AI to gender equality, calling for transversal integration of gender equality in ethical impact assessments, dedicated public budgets for gender-responsive AI initiatives and redress mechanisms for discriminatory or harmful impacts.[77] While TFGBV is not explicitly named, its principles on inclusion, transparency and redress provide a useful normative foundation for addressing gender-based digital harms.

[77] Articles 87-89.

OECD AI Principles (2019)

Status: Non-binding (soft law – principles)

The OECD framework promotes inclusive growth, sustainability and human rights in AI, noting that systems may have disparate impacts on underrepresented populations, including women.[78] The principles are relevant to AI-related harms more broadly, but lack specificity or enforcement, limiting their application to gendered risks in practice.

Global Digital Compact (2024)

Status: Non-binding (soft law – political framework)

The GDC is a broad digital governance framework that addresses a range of issues including AI, data governance and digital rights. It explicitly acknowledges the need to address and counter TFGBV and includes gender equality and the empowerment of women and girls as a core principle. It also highlights the importance of safe, secure and inclusive digital infrastructure and the ethical use of emerging technologies like AI.

However, the GDC does not explicitly link AI governance to TFGBV risks. Rather, the two themes are addressed in separate sections, without recognition of the intersection. This limits the framework's ability to meaningfully respond to the gendered harms enabled or amplified by AI systems.

G7 Hiroshima Guiding Principles on AI (2023)

Status: Non-binding (soft law – principles)

The G7 Hiroshima Guiding Principles are a non-binding intergovernmental declaration adopted by G7 countries in 2023.[79] They articulate shared values for trustworthy AI but make no mention of gender-specific risks or TFGBV-related concerns.

Freedom Online Coalition (FOC) Joint Statement on AI and Human Rights (2025)

Status: Non-binding (soft law – joint statement)

In 2025, the Freedom Online Coalition (a coalition of 42 governments that work together to advance internet freedom through shaping global norms) issued a Joint Statement on AI and Human Rights, which explicitly highlights TFGBV (including online harassment and non-consensual image-based abuse) as part of broader human rights risks enabled by AI systems.[80] The statement underscores the urgency of integrating gender-responsive safeguards throughout the AI lifecycle and calls for global alignment toward a rights-respecting and ethical AI future

[78] Principle 2.5 on inclusive growth and wellbeing.

[79] The G7 includes Canada, France, Germany, Italy, Japan, the United Kingdom and the United States, along with the European Union.

[80] Freedom Online Coalition. (2025). Joint Statement on AI and Human Rights. https://freedomonlinecoalition.com/joint-statement-on-ai-and-human-rights-2025

Regional frameworks

EU AI Act (2024)

Status: Binding (legislation)

The EU AI Act adopts a risk-based framework, with prohibitions and restrictions on certain AI uses that are deemed high-risk or unacceptable. While the act does not explicitly mention TFGBV, it includes provisions on "high-risk" applications and acknowledges that these can exacerbate discrimination across multiple dimensions – including gender.[81] It also addresses risks associated with reputational damage, disinformation and deceptive AI-generated content, which are often central to TFGBV dynamics, even if not named as such. The act does not require gender-disaggregated impact assessments or explicitly mandate gender-specific audits of AI systems.

African Union (AU) Continental Artificial Intelligence Strategy (2024)

Status: Non-binding (policy roadmap)

This strategy acknowledges structural inequalities, warning that AI risks amplifying gender disparities, digital exclusion and algorithmic bias. It includes commitments to gender equity in AI development, women-led innovation and cultural and linguistic diversity in datasets.[82]

Notably, while the AU strategy identifies deepfakes and disinformation as risks and encourages safeguards for vulnerable groups, it does not link the issues explicitly to gender and it lacks operational guidance for TFGBV prevention.

Council of Europe (CoE) AI and Human Rights Framework Convention (2024)

Status: Binding once ratified

This convention integrates equality and non-discrimination as central principles, requiring states to adopt measures to ensure that AI systems respect equality and non-discrimination and provide protection against human rights harms in digital and automated contexts.[83] While the framework mandates legal protections, its enforcement depends on national transposition. While it does not explicitly reference TFGBV, its human rights-based approach – particularly in relation to automated decision making – could be applied to address some gendered harms, depending on how member states implement and interpret the provisions.

[81] Recitals 27 and 48 and title III on high-risk systems. European Parliament and Council of the European Union. (2024). *Regulation (EU) 2024/1689*. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32024R1689.

[82] Ibid., section 2.4.4.1.

[83] Ibid., articles 5 and 7.

Table 1: International frameworks[84]

Instrument / theme	Non-discrimination / equality (general)	Gender digital divide	Gender representation in AI	Bias and stereotypes (incl. gender bias)	TFGBV (explicit)	Synthetic media / deepfakes	Online harassment / abuse	Disinformation	Algorithmic amplification of harmful content
UNESCO	X	X	X	X	X		X		
EU AI Act	X			X		X[85]		X[86]	
OECD principles	X			X					
G7 Hiroshima Guiding Principles	X[87]			X[88]					
AU Continental AI strategy	X	X	X	X		X[89]		X[90]	
CoE AI and Human Rights Framework Convention	X								
Global Digital Compact	X	X			X		X[91]	X[92]	

National AI frameworks

This section presents the findings of the analysis of 24 national AI governance frameworks, assessing how they address gendered risks, particularly TFGBV. These countries were selected based on the methodology outlined in section 1 of this report, which focuses on the Global Index on Responsible AI's (GIRAI) gender equality thematic area. All countries included have at least one formal AI governance framework – such as a strategy, policy, regulation or guideline – that references gender or gender equality in the context of AI.

To complement this sample, a small number of illustrative practices drawn from GIRAI's other pillars (government actions and non-state actors) are also referenced in the narrative analysis to address the geographically uneven sample and highlight practices from underrepresented regions.

The analysis is guided by a thematic taxonomy developed through the literature review and visualised through a comparative matrix that maps how national frameworks address key TFGBV-related risks and gendered harms.

[84] Note: The categories reflected in this table represent a *subset* of the broader analytical taxonomy developed through the literature review. The matrix captures only those themes that are explicitly articulated within individual AI governance frameworks themselves (e.g. references to non-discrimination, bias, deepfakes or harassment). Other dimensions of the taxonomy – such as governance gaps, lack of enforcement, fragmentation or absence of intersectionality – are analytical findings derived from cross-framework comparison and interpretation and are therefore examined in the narrative analysis rather than represented as discrete table categories.

[85] Addresses deepfakes or synthetic media without explicitly framing them as a gendered harm.

[86] Addresses disinformation without explicitly framing it as a gendered harm.

[87] Refers to non-discrimination but not gender discrimination specifically.

[88] Refers to bias but not gender bias specifically.

[89] Addresses deepfakes or synthetic media without explicitly framing them as a gendered harm.

[90] Addresses deepfakes or synthetic media without explicitly framing them as a gendered harm.

[91] Broadly referenced.

[92] Addresses disinformation without explicitly framing it as a gendered harm.

Findings

The analysis revealed widely uneven engagement with TFGBV and gender-specific risks. **While references to principles like fairness and inclusion are common, concrete measures to address gendered digital harms** – especially TFGBV – **remain limited**.

Of the 24 national frameworks reviewed, only a small number constitute legally enforceable instruments. The majority are formally adopted national strategies, plans or policy frameworks, complemented by a smaller set of voluntary ethical guidelines and principles.

Table 2 presents a matrix that provides a visual overview of which gender-specific themes are addressed in each national AI framework. These patterns point to a significant gap between high-level gender equality commitments and the operational treatment of AI-enabled gendered harms at the national level.

Key trends include:

- Broad and abstract references:
  Although two thirds of the reviewed countries (16 out of 24) refer to non-discrimination, they generally do so in abstract terms and do not reference AI's role in enabling gendered digital harms. Furthermore, most frameworks also lack requirements for **gender-disaggregated impact assessments**.

- TFGBV remains largely absent:
  Few national frameworks directly engage with the relationship between AI and TFGBV. Most references to gender are limited to the themes identified in the first pillar (gender digital divide, underrepresentation of women in AI development, bias and stereotypes). With a few exceptions (such as the Netherlands and the United States, which reference targeted harassment), TFGBV and its specific manifestations (e.g. deepfakes, algorithmic amplification or surveillance) are almost entirely missing from AI frameworks.

- Broad principles are not matched by implementation tools:
  While many frameworks mention fairness, diversity and inclusion, few are backed by enforceable mechanisms, such as audit protocols, accountability mechanisms or gender-responsive safeguards. This implementation gap limits the practical impact of normative commitments.

- Intersectional harms remain invisible:
  Where gender is addressed, it is often framed through a narrow binary lens. Few frameworks acknowledge intersecting vectors of harm – such as those based on race, disability, sexual orientation, caste or migration status – despite growing evidence of their significance in shaping how individuals experience AI systems.[93]

[93] For example, Chile's National AI Policy explicitly includes intersectionality as a key factor in its gender roadmap and annex III. Ministerio de Ciencia, Tecnología, Conocimiento e Innovación, Gobierno de Chile. (2021). *Política Nacional de Inteligencia Artificial*. https://www.minciencia.gob.cl/uploads/filer_public/bc/38/bc389daf-4514-4306-867c-760ae7686e2c/documento_politica_ia_digital_.pdf

Specific country examples

- Chile's AI policy is one of the few to include an explicit gender roadmap, aiming to:
  - Increase women's participation in science, technology, engineering and math (STEM) and AI fields.
  - Address bias in datasets.
  - Promote the development of equitable audit and evaluation mechanisms for AI applications.
  - Prevent discrimination in automated decision-making.[94]
- Canada mandates Gender-Based Analysis Plus (GBA+) during the development of automated systems, ensuring that algorithms are assessed for disparate impacts across gender and intersecting identities.[95]
- China's guidelines include detailed definitions of discrimination, positive action and fairness in AI systems, alongside calls for diversity in data and teams, though these are not framed explicitly in gendered terms.[96]
- Ireland promotes gender diversity in AI careers and has launched initiatives to attract women into AI through reskilling and education, though this focus is more on the workforce than on harm prevention.[97]
- Although Germany[98] and Brazil[99] recognise bias in AI and gendered data challenges, they do not explicitly mention TFGBV.
- Singapore addresses fairness through legal standards and algorithmic de-biasing, with indirect implications for reducing discrimination.[100]
- Australia's framework emphasises fairness, accessibility and non-discrimination.[101]

[94] Ibid.

[95] Treasury Board of Canada Secretariat. (2019). *Directive on Automated Decision-Making*, Appendix C (Gender-based Analysis Plus). https://www.tbs-sct.canada.ca/pol/doc-eng.aspx?id=32592

[96] Sections 3.5.1.3-4 of China's Governance Principles for the New Generation Artificial Intelligence. Digital Policy Office. (2025). *Ethical Artificial Intelligence Framework*. https://www.digitalpolicy.gov.hk/en/our_work/data_governance/policies_standards/ethical_ai_framework/

[97] Pillar 3 and Skills Development Plan of Ireland's AI Strategy. Government of Ireland. (2021). *AI – Here for Good. A National Artificial Intelligence Strategy for Ireland*. https://enterprise.gov.ie/en/publications/publication-files/national-ai-strategy.pdf

[98] Federal Foreign Office. (2023). *Shaping Feminist Foreign Policy. Federal Foreign Office Guidelines*. https://www.shapingfeministforeignpolicy.org/papers/Guidelines_Feminist_Foreign_Policy.pdf

[99] Ministério da Ciência, Tecnologia e Inovações Secretaria de Empreendedorismo e Inovação. (n/d). *Estratégia Brasileira de Inteligência Artificial* –EBIA. https://www.gov.br/mcti/pt-br/acompanhe-o-mcti/transformacaodigital/ebia.pdf

[100] Personal Data Protection Commission – Singapore. (2020). *Model Artificial Intelligence Governance Framework. Second edition*. https://www.pdpc.gov.sg/-/media/Files/PDPC/PDF-Files/Resource-for-Organisation/AI/SGModelAIGovFramework2.pdf

[101] Department of Industry, Science and Resources. (2019). *Australias's AI Ethics Principles*. https://www.industry.gov.au/publications/australias-artificial-intelligence-ethics-principles/australias-ai-ethics-principles

Morocco: Government-led consultations on AI and human rights

- ***Type:*** Governance action
- ***What:*** In 2023, Morocco's National Human Rights Council released a report[102] summarising the findings of **national consultations and monitoring activities** on the protection of human rights in digital and AI systems. The initiative included analysis of the human rights implications of AI and digital technologies, with a **focus on gender equality and social inclusion**.
- ***Why it matters***: While Morocco does not yet have a comprehensive national AI strategy, this report reflects a proactive governance approach that integrates gender equality considerations in the early stages of AI policymaking. It also demonstrates how participatory, multistakeholder research can shape future regulatory directions with a rights-based lens.

The Philippines: National workshop on gender-based AI policy[103]

- ***Type:*** Governance action
- ***What:*** In partnership with the International Telecommunication Union (ITU), the Philippines' Department of Information and Communications Technology (DICT) hosted the **"AI Dialogue: Gender-Based AI Policy in the Philippines"** workshop in 2023. The event convened public officials, academics and civil society actors to explore **gender-responsive approaches to AI policy**, with a focus on inclusion, safety and bias mitigation.
- ***Why it matters:*** This initiative reflects growing government awareness of gendered risks in AI, even in the absence of a dedicated national AI law. It highlights how dialogue-based policy development – particularly with international support – can build local capacity and advance gender considerations within the national AI governance agenda.

India: Civil society input on business and human rights in AI

- ***Type:*** Non-state action
- ***What:*** The feminist digital rights organisation **IT for Change** submitted a detailed input to the **UN Human Rights B-Tech Project**, advocating for a human rights-based framework for regulating technology companies.[104] Their submission highlights the **gendered and intersectional harms of AI systems**, including the role of platforms in facilitating TFGBV, algorithmic bias and surveillance. It also proposes **ethical and legal standards** grounded in Global Majority realities.
- ***Why it matters:*** This initiative exemplifies how civil society can shape global AI norms from a feminist and decolonial perspective, especially in contexts where national regulation is still emerging. IT for Change's engagement bridges local experience with international standard setting, making gendered digital harms visible in global governance spaces.

[102] Bensalah, M., & National Human Rights Council. (2022). *Rapport de synthèse des résultats du monitoring et des rencontres nationales de concertation sur la protection des droits humains dans le digital et les systèmes de l'intelligence artificielle.* https://www.researchgate.net/publication/359982730_Rapport_de_synthese_des_resultats_du_monitoring_et_des_rencontres_nationales_de_concertation_sur_la_protection_des_droits_humains_dans_le_digital_et_les_systemes_de_l%27intelligence_artificielle

[103] https://ictstatistics.dict.gov.ph/dict-partners-with-itu-to-deliver-ai-dialogue-gender-based-ai-policy-in-the-philippines-workshop/

[104] IT for Change. (2023). *Input to the UN Human Rights B-Tech Project.* https://itforchange.net/sites/default/files/2440/Input%20to%20the%20UN%20Human%20Rights%20B-Tech%20Project%20.pdf

Ethiopia: Academic research on AI ethics and African gendered identities

- ***Type:*** Non-state action
- ***What:*** The **GIGA Institute for African Affairs** published a study titled "Machine Ethics and African Identities: Perspectives of Artificial Intelligence in Africa," which critically examines the cultural, gendered and ethical implications of AI systems for the continent.[105] The research explores how dominant AI models may **erase African epistemologies and reinforce gendered inequalities**, particularly through imported data regimes and top-down ethical standards.
- ***Why it matters:*** This work contributes a gender-aware, decolonial perspective to AI ethics – challenging Eurocentric framings and contributing to African feminist scholarship on AI ethics. It underscores the importance of contextualising AI harms within local identities and of grounding ethical standards in intersectional African realities.

Table 2: National frameworks[106]

Instrument / theme	Type[107]	Non- discrimination/ equality (general)	Gender digital divide	Gender representation in AI	Bias and stereotypes (incl. gender bias)	TFGBV (explicit)	Synthetic media / deepfakes	Online harassment / abuse	Disinformation	Algorithmic amplification of harmful content
Australia	Voluntary / advisory framework	X								
Brazil	Formally adopted policy / strategy (not legally binding)	X	X[108]		X					
Bulgaria	Formally adopted policy / strategy (not legally binding)	X								
Canada	Legally enforceable framework (within federal public administration)	X[109]								

[105] Kohnert, D. (2022). Machine Ethics and African Identities: Perspectives of Artificial Intelligence in Africa. *SSRN Electronic Journal*. http://dx.doi.org/10.2139/ssrn.4163096

[106] Note: The categories reflected in this table represent a *subset* of the broader analytical taxonomy developed through the literature review. The matrix captures only those themes that are explicitly articulated within individual AI governance frameworks themselves (e.g. references to non-discrimination, bias, deepfakes or harassment). Other dimensions of the taxonomy – such as governance gaps, lack of enforcement, fragmentation or absence of intersectionality – are analytical findings derived from cross-framework comparison and interpretation and are therefore examined in the narrative analysis rather than represented as discrete table categories.

[107] Note on legal status: The Global Index on Responsible AI (GIRAI) classifies many national AI strategies, plans and policy frameworks as "binding" on the basis of their **formal adoption by government authorities**. In this report, however, a narrower legal definition is applied. Frameworks are classified as **legally enforceable** only where they create binding legal obligations, compliance requirements or sanctions. National strategies, action plans, roadmaps and policy documents adopted through executive or ministerial processes are therefore treated as **non-legally binding policy frameworks**, even where they are formally adopted by government.

[108] Discusses inclusion and access, but not framed as a gender digital divide.

Instrument / theme	Type[107]	Non- discrimination/ equality (general)	Gender digital divide	Gender representation in AI	Bias and stereotypes (incl. gender bias)	TFGBV (explicit)	Synthetic media / deepfakes	Online harassment / abuse	Disinformation	Algorithmic amplification of harmful content
Chile	Formally adopted policy / strategy (not legally binding)	X[110]	X	X	X					
China[111]	Formally adopted policy guidance (non-legally binding)			X[112]	X					
Croatia	Formally adopted policy / action plan (non-legally binding)	X								
Dominican Republic	Formally adopted policy / strategy (not legally binding)			X						
Germany	Formally adopted policy framework (non-legally binding)	X								
Greece	Legally enforceable framework	X								
Ireland	Formally adopted policy / strategy (not legally binding)	X	X	X	X[113]					
Italy	Formally adopted policy / strategy (not legally binding)			X						
Japan	Voluntary / advisory framework				X[114]					
Jordan	Voluntary / advisory framework	X			X[115]					
Malaysia	Formally adopted policy / roadmap (non-legally binding)			X						

[109] It mandates conducting a gender-based analysis.

[110] Furthermore, it contemplates intersectionality.

[111] GIRAI lists "China, Hong Kong Special Administrative Region" as a combined entry. However, for the purposes of this analysis, only Mainland China's national AI governance framework was reviewed, as Hong Kong and China maintain distinct legal and regulatory systems.

[112] Calls for diversity in data and teams, though these are not framed explicitly in gendered terms.

[113] It contemplates bias but not gender bias explicitly.

Instrument / theme	Type[107]	Non- discrimination/ equality (general)	Gender digital divide	Gender representation in AI	Bias and stereotypes (incl. gender bias)	TFGBV (explicit)	Synthetic media / deepfakes	Online harassment / abuse	Disinformation	Algorithmic amplification of harmful content
The Netherlands	Formally adopted policy framework (non-legally binding)	X						X		
Portugal	Voluntary / advisory framework	X	X[116]		X[117]					
Saudi Arabia	Voluntary / advisory framework	X			X					
Senegal[118]	Formally adopted policy / strategy (not legally binding)	X	X	X	X					
Serbia	Voluntary / advisory framework	X		X	X					
Singapore	Voluntary / advisory framework				X					
Slovenia	Formally adopted policy / programme (non-legally binding)	X		X						
Spain	Formally adopted policy / strategy (not legally binding)	X		X						
United States	Formally adopted policy framework (non-legally binding)	X	X	X		X		X		

[114] It contemplates bias but not gender bias explicitly.

[115] It contemplates bias but not gender bias explicitly.

[116] Indirect. It contemplates inclusion and accessibility although it does not address gender specifically.

[117] It contemplates bias but not gender bias explicitly.

[118] Senegal's assessment is based on the publicly available summary of its national AI strategy and UNESCO's AI Readiness Assessment (RAM) report, as the full strategy text was not accessible at the time of analysis. The RAM was used to contextualise ethical and inclusion-related priorities but does not replace formal policy analysis.

Section 4. TFGBV norms analysis

This section analyses a selection of international and regional norms, resolutions and legal instruments that address gender-based violence, assessing the extent to which these frameworks integrate AI-specific risks. **While there has been increasing recognition of TFGBV, references to AI remain limited, often generalised and lacking enforceable or operational guidance**.

The analysis draws on instruments from the United Nations (UN), regional human rights systems (Europe, Africa and the Americas), national-level legal responses and initiatives from feminist civil society groups, focusing on how they respond to the TFGBV harms previously identified in this study.

UN resolutions and declarations

Human Rights Council Resolution 38/5 (2018), Accelerating efforts to eliminate violence against women and girls: preventing and responding to violence against women and girls in digital contexts

Status: Non-binding

One of the earliest UN texts to specifically address violence against women in digital contexts, this Human Rights Council resolution recognises forms of TFGBV such as online harassment, surveillance and hacking. It calls for states to develop effective responses to systemic gender discrimination online but does not mention AI or automated technologies explicitly.[119]

UN General Assembly Resolution 75/161 (2020), Advancing women's rights in the digital age

Status: Non-binding

Focuses on violence against women in the digital age, recognising that online sexual harassment and cyberviolence have social and economic impacts. It calls for improved protections but does not incorporate AI-driven risks such as automated surveillance or deepfake generation.[120]

UN General Assembly Resolution 78/213 (2023), Promotion and protection of human rights in the context of digital technologies

Status: Non-binding

This landmark resolution highlights the risks AI poses when deployed without human rights, technical and ethical safeguards, noting that such systems can reinforce gender-based discrimination and TFGBV. It also affirms the need to apply human rights law in the design and regulation of these technologies.[121] However, it remains high-level and does not provide detailed guidance on specific AI-enabled harms such as deepfakes.

[119] UN Human Rights Council. (2018). *Resolution 38/5 adopted by the Human Rights Council on 5 July 2018. Accelerating efforts to eliminate violence against women and girls: preventing and responding to violence in digital contexts*. (A/HRC/RES/38/5). https://documents.un.org/doc/undoc/gen/g18/214/82/pdf/g1821482.pdf

[120] UN General Assembly. (2020). *Resolution 75/161 adopted by the General Assembly on 16 December 2020. Intensification of efforts to prevent and eliminate all forms of violence against women and girls*. (A/RES/75/161). https://documents.un.org/doc/undoc/gen/n20/372/46/pdf/n2037246.pdf

[121] UN General Assembly. (2023). *Resolution 78/213 adopted by the General Assembly on 19 December 2023. Promotion and protection of human rights in the context of digital technologies*. (A/RES/78/213). https://docs.un.org/en/A/RES/78/213

CSW67 Agreed Conclusions (2023). Innovation and technological change, and education in the digital age for achieving gender equality

Status: Non-binding

A significant normative step, the Agreed Conclusions of the 67th Session of the Commission on the Status of Women (CSW67) acknowledge that AI and digital technologies can entrench bias, amplify TFGBV and disproportionately harm women and girls, including through surveillance, algorithmic discrimination and exclusion from digital labour markets. It urges governments to regulate AI and integrate gender-responsive safeguards, addressing:

- The need for data audits and algorithmic transparency.
- The underrepresentation of women in AI development.
- The role of algorithmic discrimination in TFGBV.[122]

Regional instruments and frameworks

EU Directive on Combating Violence against Women (2024)

Status: Binding (once transposed into national law)

One of the most detailed regional efforts to explicitly regulate TFGBV, including harms that may be enabled or amplified by AI. It defines and criminalises:

- Cyberstalking, cyberflashing and deepfake pornography
- Non-consensual dissemination of synthetic sexual content
- Use of ICTs for psychological manipulation, harassment and surveillance.[123]

It also stresses intersectional discrimination and mandates digital platforms and states to address amplification risks. This directive presents a model of binding, enforceable protections against AI-enabled TFGBV.

Istanbul Convention (Council of Europe) (2011)

Status: Binding (for ratifying states)

While pioneering in addressing gender-based violence, the convention is limited in its digital specificity. Article 17 encourages the private sector to develop policies against online violence but does not directly address AI or its distinct harms.[124]

Maputo Protocol (African Union) (2003)

Status: Binding (for ratifying states)

Mentions the need for women's participation in the development and use of technologies but offers no concrete measures on digital or AI-facilitated violence (article XVIII).[125]

[122] UN Commission on the Status of Women. (2023). Op. cit.

[123] European Union. (2024). Op. cit.

[124] Council of Europe. (2011). *Council of Europe Convention on preventing and combating violence against women and domestic violence (Istanbul Convention)*. https://rm.coe.int/168008482e

[125] African Union. (2003). *Protocol to the African Charter on Human and Peoples' Rights on the Rights of Women in Africa (Maputo Protocol)*. https://au.int/sites/default/files/treaties/37077-treaty-charter_on_rights_of_women_in_africa.pdf

African Commission on Human and Peoples’ Rights Resolution 522 (LXXII) 2022 – Protection against digital violence

Status: Non-binding regional resolution

Recognises digital violence as a growing threat and expands the understanding of gender-based violence to include cybersecurity abuse (cyberstalking, sexist hate speech, non-consensual image sharing). Does not explicitly reference AI, but provides an important normative foundation for addressing gender-based harms in online spaces.

African Commission on Human and Peoples’ Rights Resolution 591 (LXXX) 2024 – Call for study on digital violence

Status: Non-binding regional resolution

Mandates a one-year continental study on digital violence against women, intends to inform new standards and assist states in tackling digital gender harms. The resolution does not mention AI, but may lay the groundwork for future frameworks that respond to AI-specific risks.

Inter-American Convention of Belém do Pará (OAS) (1994)

Status: Binding (for ratifying states)

The convention itself does not explicitly address digital or AI-enabled forms of violence.[126] However, its scope has been progressively expanded through interpretative and normative developments led by its Follow-up Mechanism (MESECVI).[127]

Inter-American Model Law to Prevent, Punish and Eradicate Gender-based Digital Violence Against Women (2025)

Status: Non-binding (model law / regional normative instrument)

In December 2025, the Follow-up Mechanism to the Belém do Pará Convention (MESECVI) formally adopted the Inter-American Model Law to Prevent, Punish and Eradicate Gender-based Digital Violence against Women, marking a major regional milestone in the regulation of TFGBV.

The Model Law provides a comprehensive, rights-based framework to guide domestic legislative reforms. It explicitly recognises a wide range of technology-facilitated harms, including online harassment, surveillance, identity theft, hate speech, deepfakes, non-consensual sharing of intimate content, digital extortion and gender-biased disinformation campaigns, many of which may be enabled or amplified by AI systems

Grounded in a gender-sensitive and intersectional approach, the Model Law establishes state obligations to prevent, investigate, punish and provide reparations for digital violence, alongside responsibilities for digital platforms and internet intermediaries, including cooperation with authorities, transparency and evidence preservation.

Feminist civil society analysis has highlighted the significance of the Model Law in explicitly recognising digital violence as a form of gender-based violence, expanding state responsibility

[126] Organization of American States. (1994). *Inter-American Convention on the Prevention, Punishment and Eradication of Violence Against Women (Convention of Belém do Pará).* https://www.oas.org/juridico/english/treaties/a-61.html

[127] Equality Now. (2024, 21 June). The Belém do Pará Convention at 30: Commitments to Accelerate its Implementation at the XI Conference of States Parties. https://equalitynow.org/news_and_insights/the-belem-do-para-convention-at-30-commitments-to-accelerate-its-implementation-at-the-xi-conference-of-states-parties/

into online and algorithmically mediated spaces and affirming survivors' rights to access justice, remedies and support services. As Equality Now notes, the Model Law is particularly notable for its intersectional framing, its recognition of structural discrimination and its potential to serve as a legislative blueprint for states seeking to address emerging harms such as AI-generated abuse and coordinated digital attacks against women in public life.[128]

While non-binding, the Model Law constitutes a highly influential normative and technical instrument and represents one of the most advanced regional efforts to address TFGBV in the context of digital and AI-enabled harms.

National legal frameworks and policy guidance addressing TFGBV and AI

Australia: Australian Human Rights Commission Report on TFGBV and Human Rights Tools (2023)

Status: Non-binding

The Australian Human Rights Commission has identified growing risks of TFGBV in digital environments, particularly with the integration of AI-enabled technologies. Its 2023 report outlines:

- Legal reforms targeting deepfake pornography, the need for stronger regulatory oversight of AI-enabled technologies and strengthening the Online Safety Act.
- A critique of industry self-regulation, noting a rollback of protections by platforms like Meta and emphasising that voluntary compliance is no longer sufficient to protect women and girls online.
- Supporting the adoption of human rights impact assessments (HRIAs) in the development of AI tools, to ensure they address gendered harms before deployment.[129]

This approach blends human rights frameworks with technical oversight, signalling a proactive model for regulating AI-TFGBV intersections.

UK Online Safety Act (2023)

Status: Binding (legislation)

The UK's Online Safety Act introduces binding obligations on platforms to reduce online harms and explicitly recognises the use of AI and machine learning as part of "proactive technologies" for content detection and moderation.[130]

The act also introduces criminal penalties for the creation and distribution of sexually explicit deepfakes, non-consensual intimate imagery, including content digitally altered via AI to appear real, and behaviour motivated by intent to humiliate or cause distress.

[128] Equality Now. (2025, 11 December). Six key points to understand the new Inter-American Model Law on Digital Violence against Women. https://equalitynow.org/news/news-and-insights/six-key-points-to-understand-the-new-inter-american-model-law-on-digital-violence-against-women/

[129] Australian Human Rights Commission. (2025). *Technology-facilitated gender-based violence. Submission to Human Rights Council Advisory Commission*, paragraphs 23-25. https://www.ohchr.org/sites/default/files/documents/hrbodies/hrcouncil/advisorycommittee/cfi-neurotechnology/subm-technology-facilitated-gender-nhri-australian-hr-commission.pdf

[130] UK Government. (2023). Op. cit., part 12, section 231, definition of proactive technologies (https://www.legislation.gov.uk/ukpga/2023/50/section/231/enacted), and, section 66A of the Sexual Offences Act, as amended by the Online Safety Act (https://www.legislation.gov.uk/ukpga/2003/42/section/66A).

Civil society and non-state normative frameworks

Feminist Principles of the Internet

These civil society principles emphasise the right to safety and bodily autonomy online, calling for actions against online harassment, doxxing and surveillance. They recognise the gendered power imbalances in tech and advocate for feminist governance of digital spaces, but stop short of naming AI as a distinct risk domain.[131]

Feminist Digital Justice and Global Digital Compact proposals

Emerging feminist digital justice frameworks have made critical contributions to the Global Digital Compact (GDC) process, arguing that algorithmic amplification, AI-generated abuse and predictive profiling should be understood not as peripheral concerns but as core manifestations of TFGBV. Notably, the Feminist Principles for Including Gender in the Global Digital Compact – developed by a coalition of feminist and digital rights organisations – outline a rights-based, intersectional agenda for the GDC.[132] These principles call for recognition of gendered power asymmetries in AI systems, the need for structural safeguards and the agency of marginalised communities in shaping global digital norms.

While these proposals have begun to influence multistakeholder discourse and are reflected in civil society submissions to the UN process, they have not yet been codified into binding instruments. Their inclusion in the final GDC text is partial and key feminist concerns – such as TFGBV and intersectional discrimination in AI – remain underdeveloped or absent in the official language.

Key gaps identified

Despite progress in acknowledging TFGBV in digital spaces, several gaps persist in how international norms engage with AI-specific risks:

- Limited specificity: Most frameworks refer generically to “digital technologies” without naming AI-enabled abuses such as deepfakes, algorithmic profiling or automated targeting.
- Non-enforceability: Many UN resolutions and regional declarations promote state action but lack enforcement or monitoring mechanisms.
- Disconnect from AI and data policy domains: TFGBV norms remain siloed from broader AI governance and are rarely cross-referenced in policy spaces like data protection or platform accountability. Yet as noted in earlier sections, data extraction and algorithmic profiling function as central enablers of TFGBV, raising urgent questions about consent, redress and power in data-driven systems. While a detailed review of data frameworks is beyond the scope of this study, better integration of TFGBV concerns into data and AI regulation is essential.
- Lack of intersectionality: Few instruments account for compounded vulnerabilities based on disability, race, migration status, sexuality, caste or other aspects of identity in AI and TFGBV contexts.

[131] Association for Progressive Communications. (2016). Op. cit.

[132] Association for Progressive Communications. (2023, 6 October). The Feminist Principles for Including Gender in the Global Digital Compact. https://www.apc.org/en/pubs/feminist-principles-including-gender-global-digital-compact-0

While awareness of TFGBV in digital environments is growing, **most global and regional norms lag behind the technological reality of AI-driven harms**. The 2024 EU Directive and the CSW67 Agreed Conclusions mark significant progress, but the **systemic integration of AI-specific TFGBV risks across legal, regulatory and normative instruments remains urgently needed**.

Section 5. Findings and analysis

This section distils key crosscutting insights from the comparative analysis of AI and TFGBV frameworks, literature review and the thematic taxonomy. It identifies patterns, gaps and tensions across governance approaches and highlights opportunities for convergence.

Policy fragmentation and siloed governance

One of the most significant findings is the disconnect between AI governance and TFGBV frameworks. While many AI frameworks reference ethical principles such as fairness and non-discrimination, only a handful meaningfully engage with gendered digital harms or TFGBV-specific risks. Conversely, most TFGBV norms (even those acknowledging digital technologies) do not account for the risks posed by AI tools such as deepfakes, profiling or algorithmic targeting.

This gap is particularly concerning in the context of state-led AI deployment, where digital public infrastructure, welfare automation[133] and surveillance systems can introduce or exacerbate TFGBV risks – especially for already marginalised communities.[134]

This siloed approach can result in:

- **Gaps in protection:** Women and marginalised communities face harms not covered by AI audits or by traditional TFGBV protections.
- **Ineffective enforcement:** Without legal recognition of AI-specific TFGBV, victims lack clear avenues for redress.
- **Duplicative or inconsistent efforts:** TFGBV and AI regulators operate on parallel tracks, often without intersecting mandates or data-sharing frameworks.
- **Lack of private sector accountability:** Many AI systems that contribute to TFGBV are developed, deployed and maintained by private companies – often operating across jurisdictions with minimal regulation or oversight. Without enforceable obligations on these actors, gendered harms are unlikely to be prevented or addressed.

Broad principles are not matched by implementation tools

While many frameworks mention fairness, diversity and inclusion, few establish accountability mechanisms, gender-disaggregated data requirements, audit protocols or gender-responsive safeguards. In several cases, gender equality is referenced in preambles or guiding principles but not reflected in concrete implementation measures (such as indicators, timelines, budgetary allocations or enforcement procedures). This mirrors trends observed in adjacent digital policy fields, such as cybersecurity[135] or data protection. Without concrete mechanisms, gender integration risks remaining rhetorical rather than operational.

[133] See: Alston, P. (2019). *Report of the Special Rapporteur on extreme poverty and human rights (Digital welfare states and human rights)*. UN General Assembly. (A/74/493). https://docs.un.org/en/A/74/493

[134] See: Amnesty International. (2024). *Coded injustice: Surveillance and discrimination in Denmark's automated welfare state*. https://www.amnesty.org/en/documents/eur18/8709/2024/en/

[135] See: Finlay, A. (2023). *A framework for developing gender-responsive cybersecurity policy: Assessment tool*. Association for Progressive Communications. https://www.apc.org/sites/default/files/apcgendercyber-assessmenttool.pdf

Disproportionate attention to certain risks

The thematic assessment reveals a skewed distribution of attention across AI-TFGBV themes. Well-represented areas include bias and stereotypes, especially in datasets and algorithmic decision-making (noted in national frameworks in Brazil and Canada and in international frameworks such as the UNESCO Recommendation and the CSW67 Agreed Conclusions, among others), and the gender digital divide and underrepresentation of women in STEM (covered in the CSW67 Agreed Conclusions, the AU AI strategy and several national frameworks).

Underrepresented or under-addressed areas include deepfakes and non-consensual synthetic content, addressed most robustly only in the UK, EU and Australia; and algorithmic amplification of harmful content, which receives little to no references in AI frameworks.

Few frameworks integrate intersectionality, despite well-documented disparities in how AI harms are distributed along the lines of race, class and disability. The absence of gender- and other equalities-related disaggregated data requirements in most frameworks compounds this issue.

Emerging national and regional innovation

While international norms lag in AI-TFGBV integration, several national and regional frameworks are pioneering new approaches:

- Australia combines content regulation, criminal law and human rights impact assessments (HRIAs) to address deepfakes, platform accountability and AI risks.
- The UK defines deepfake pornography as a criminal offence and requires platforms to use proactive technologies to mitigate harm.
- Chile and Canada provide policy models for embedding gender-based analysis into algorithmic systems.
- At the regional level, the EU AI Act introduces binding obligations on high-risk profiling, algorithmic targeting and biometric assessment. While it does not name TFGBV directly, these provisions are directly relevant to predictive harms and surveillance abuses that disproportionately affect women and marginalised communities.

These frameworks show the value of combining technical oversight, legal reform and gender analysis – a model that could be scaled up or used to inform future regional and multilateral efforts.

Absence as harm: The chilling effect of TFGBV in the age of AI

An emerging insight from the literature review and consultations is the chilling effect that AI-enabled TFGBV has on the participation of women and gender-diverse individuals in public life. Deepfakes, targeted harassment, algorithmic amplification of abuse and non-consensual surveillance can produce significant reputational and psychological harm. Beyond individual impacts, these forms of abuse often lead to withdrawal from digital spaces, reduced visibility and, in some cases, forced exits from professional or community roles.

This absence is a form of structural harm. It restricts the exercise of core human rights such as freedom of expression, political participation and access to essential services. It also reduces the presence of marginalised voices in civic, academic and policy discourse, reinforcing existing power imbalances. This exclusion is both a consequence of TFGBV and a systemic form of gender-based harm.

Conclusions

AI-related TFGBV risks remain under-regulated, inconsistently addressed and poorly integrated across global, regional and national frameworks. Few instruments adequately capture the specific harms introduced or amplified by AI systems, such as deepfakes, automated harassment or algorithmic amplification.

The findings point to an urgent need for coordinated, inclusive and intersectional responses from policy makers, regulators, multilateral institutions and other actors involved in AI governance and gender-based violence prevention and response. Key areas for action include:

- Stronger policy coherence between AI governance and TFGBV frameworks: Breaking down silos by embedding TFGBV risks into AI standards, strategies and regulatory frameworks and ensuring that AI governance explicitly incorporates violence prevention and response as core objectives.

- Sharing and adaptation of promising national and regional practices: Building on emerging models – such as the EU AI Act, Australia's HRIA tools or Chile's gender-responsive algorithmic audits – to inform international guidance and regional frameworks, while remaining sensitive to diverse legal, political and cultural contexts.

- Greater investment in enforcement mechanisms and transparency tools: Strengthening oversight through requirements for gender audits, gender- and equalities-disaggregated data collection and impact assessments that centre the experience of women and marginalised groups. This should include sustained support for civil society, researchers and regulators to monitor harms, track compliance and support accountability over time.

Section 6. Policy recommendations

To close the gaps identified at the intersection of AI and TFGBV, this section proposes concrete, actionable recommendations to support the development of inclusive, enforceable and human rights-based frameworks that recognise and prevent AI-related gender harms, while also strengthening protections within TFGBV mechanisms.

The following recommendations are directed at policy makers, multilateral institutions, national governments, regulatory bodies and standard-setting organisations working at the intersection of AI governance and TFGBV prevention. They are also relevant to civil society organisations, researchers and feminist advocates engaged in digital technology processes.

These recommendations are structured in three categories:

1. AI specific actions
2. TFGBV-specific actions
3. Crosscutting and structural actions.

AI-specific actions

1. Mandate gender-responsive risk and impact assessments.

Require public and private sector developers to conduct gender-based analysis and human rights impact assessments (HRIAs) at every stage of the AI lifecycle, especially for high-risk applications (e.g. biometric analysis, content moderation).[136]

2. Establish enforceable standards for AI bias mitigation.

Introduce binding obligations for the detection and correction of gender bias in AI datasets and models. These standards should include requirements for representative training data and auditable fairness metrics.[137]

3. Explicitly integrate TFGBV risks in AI governance frameworks.

Amend AI ethics guidelines, strategies and legislation to recognise TFGBV as a distinct risk category, particularly with respect to deepfakes and synthetic content, algorithmic amplification of abuse and automated profiling for harassment.[138]

4. Ensure meaningful and resourced participation of women and gender diverse people across the AI lifecycle

Governments, international institutions and private sector actors should ensure **inclusive, equitable and sustained participation** of women, gender diverse communities and civil society actors across all stages of the AI lifecycle – from problem definition and design, to deployment, evaluation and oversight.

[136] Treasury Board of Canada Secretariat. (2019). Op. cit.; Australian Human Rights Commission. (2025). Op. cit., paragraphs 23-25.

[137] UNESCO. (2022). Op. cit., articles 88-89.

[138] Ibid. articles 88-89; European Union. (2024). Op. cit., articles 5-9; UK Government. (2023). Op. cit., section 66A, as inserted into the Sexual Offences Act 2003.

This should include:

- **Establishing mechanisms for gender-diverse and underrepresented groups** – particularly from the Global Majority – to co-develop, monitor and evaluate AI policies, standards and enforcement frameworks.
- **Providing sustained resources to community-led expertise**, especially in contexts where state or corporate-led efforts may lack legitimacy or gendered insight.
- **Recognising and addressing power asymmetries in AI governance**, including the concentration of influence among large technology firms and Global North institutions and committing to redistributive, participatory models of decision-making.

TFGBV specific recommendations

1. Integrate AI-related risks into TFGBV norms and frameworks.

Update domestic and international TFGBV laws and policies to explicitly include AI-generated and AI-enabled harms, including synthetic sexual imagery, AI-facilitated stalking, automated doxxing and algorithmic amplification of abuse. Existing instruments such as the EU Directive (2024) and the UK Online Safety Act offer useful models for this integration.[139]

2. Update definitions of digital violence.

Ensure definitions of TFGBV explicitly encompass AI-enabled harms, including non-consensual synthetic content, algorithmic harassment, profiling and coordinated disinformation campaigns targeting women and gender-diverse people.[140]

3. Strengthen accountability and redress mechanisms.

Create clear pathways for reporting and redress in cases of AI-enabled TFGBV, including transparent reporting and appeals mechanisms for victims.[141]

Crosscutting and structural recommendations

1. Foster regulatory coordination and multistakeholder governance by:

- Building institutionalised links between AI ethics bodies, data protection authorities and gender equality mechanisms.
- Ensuring meaningful participation of civil society, feminist networks, marginalised and underrepresented communities, technologists, legal experts, academics and private sector actors, in both agenda setting and implementation.
- Promoting global alignment across emerging AI and data governance processes, including coordinated engagement between the Global Dialogue on AI Governance[142] and the Working Group on Data Governance of the UN Commission on Science and

[139] European Union. (2024). Op. cit., articles 5-9; UK Online Safety Act (2023), section 66A, as inserted into the Sexual Offences Act 2003.

[140] UN Commission on the Status of Women. (2023). Op. cit., paragraphs 38-46.

[141] Ibid.

[142] The Global Dialogue on AI Governance is a multistakeholder process initiated by the UN Tech Envoy to build consensus around a potential global governance framework for AI. https://www.un.org/global-dialogue-ai-governance/en

Technology for Development (CSTD),[143] grounded in shared commitments to human rights, gender equality and intersectional feminist principles.

2. Fund feminist and community-led digital innovation.

Support grassroots and feminist technology initiatives focused on safety, accountability and ethical AI design, including harm documentation and capacity building for local and regional actors. These initiatives are often best positioned to identify emerging threats and model responsive tools.[144]

This support should be accompanied by policy environments that enable a more diverse and plural AI ecosystem, rather than reinforcing the current concentration of AI development in the hands of a small number of actors, predominantly based in the Global North.

3. Strengthen corporate accountability for AI-driven TFGBV.

Private sector actors (particularly large technology platforms, infrastructure providers and AI developers) play a central role in shaping the AI ecosystem[145] and can contribute to the amplification of TFGBV.[146] However, corporate initiatives to address these harms often rely on voluntary commitments and self-regulation and may lack transparency, enforceability or meaningful community accountability.[147]

Under international human rights law,[148] states bear the primary obligation to protect individuals and communities from human right abuses by business enterprises, including those arising from AI systems. In the context of AI-driven TFGBV, this requires governments to adopt and enforce effective legal, policy and regulatory measures to ensure companies identify, prevent, mitigate and account for TFGBV risks throughout the AI lifecycle. Multilateral bodies can support this through norm setting and coordination, while civil society can play a key role in monitoring impacts.

[143] The UN CSTD Working Group on Data Governance is a dedicated expert body tasked with developing policy recommendations to implement the Global Digital Compact's vision for equitable and inclusive data governance. See: UN Trade and Development. (n/d). Working group on data governance at all levels. https://unctad.org/topic/commission-on-science-and-technology-for-development/working-group-on-data-governance

[144] See: IT for Change & Development Alternatives with Women for a New Era (Dawn). (2023). *The Declaration of Feminist Digital Justice*. https://feministdigitaljustice.net/wp-content/uploads/2023/03/JNC-WG-Declaration-of-Feminist-Digital-Justice_2023.pdf; and Association for Progressive Communications. (2023). *Joint Submission to the Global Digital Compact on Gender*. https://www.apc.org/sites/default/files/gdc_joint_submission_on_gender_final.pdf

[145] See: UN General Assembly. (2025). *Artificial intelligence procurement and deployment: ensuring alignment with the Guiding Principles on Business and Human Rights. Report of the Working Group on the issue of human rights and transnational corporations and other business enterprises*. (A/HRC/59/53). https://docs.un.org/en/A/HRC/59/53; WITNESS. (2025, 6 March). Op. cit.; UNESCO. (2023). Op. cit.; Jankowicz, N., Gómez-O'Keefe, I., Hoffman, L., & Vidal Becker, A. (2024). *It's Everyone's Problem: Mainstreaming Responses to Technology-Facilitated Gender-Based Violence*. Institute of Global Politics & Vital Voices Global Partnership. https://igp.sipa.columbia.edu/sites/igp/files/2024-09/IGP_TFGBV_Its_Everyones_Problem_090524.pdf

[146] Wilkinson, I., Hofstetter, J-S., Shires, J., & Siba Yahaya, M. (2024). *The role of the private sector in combatting gendered cyber harms: How private sector technologies can both endanger and advance gender-transformative cybersecurity*. Royal Institute of International Affairs. https://www.chathamhouse.org/sites/default/files/2024-06/2024-06-03-private-sector-gendered-cyber-harms-wilkinson-hofstetter-shires-yahaya_0.pdf; Amnesty International. (2018). *Toxic Twitter: A Toxic Place for Women*. https://www.amnesty.org/en/latest/research/2018/03/online-violence-against-women-chapter-1-1/

[147] UN Human Rights Council. (2022). *The practical application of the Guiding Principles on Business and Human Rights to the activities of technology companies. Report of the Office of the United Nations High Commissioner for Human Rights*. (A/HRC/50/56). https://docs.un.org/en/A/HRC/50/56; Amnesty International. (2019). Surveillance Giants: How the business model of Google and Facebook threatens human rights. https://www.amnesty.org/en/tech/surveillance-giants/; and Pírková, E., Sampieri, A., & Zaghdoud, A. (2024). *Platform accountability: A rule-of-law checklist for policymakers*. Access Now. https://www.accessnow.org/wp-content/uploads/2024/12/Platform-accountability-a-rule-of-law-checklist-for-policymakers-report-2024.pdf

[148] UN Human Rights Council. (2022). Op. cit.; *UN Office of the High Commissioner for Human Rights. (2025). UNGPs compass for tech regulation: A policymaker's guide on how to apply the UNGPs to require rights-respecting corporate responsibility*. https://www.ohchr.org/sites/default/files/documents/issues/business/b-tech/b-tech-ungps-compass-vf.pdf

This includes, for example, state-led measures to:

- Mandate transparency, safety and redress obligations for companies, including gender audits and risk assessments for AI systems with high societal impact.
- Require gender-responsive and human rights-based due diligence in private sector AI development, aligned with frameworks such as the UN Guiding Principles on Business and Human Rights.
- Establish clear public-interest safeguards for corporate involvement in digital infrastructure and data governance, particularly in Global Majority contexts where AI-related investments may carry risks of gendered harms.

4. Build capacity among regulators and oversight bodies to address AI-related TFGBV

Ensure that institutions tasked with enforcing AI and TFGBV frameworks (including national regulators, data protection authorities or judicial bodies) receive adequate training on structural gendered harms and intersectional impacts.

5. Ensure intersectional data collection and infrastructure that supports the development of rights-based AI and TFGBV governance. This includes:

- Disaggregated data by gender, race, disability, age, migration status and other relevant identities.
- Use of anonymisation and strong privacy protections, in compliance with data protection frameworks and ethical data standards.
- Community-informed consent protocols that prioritise agency, transparency and safety for affected individuals and groups.
- Collaboration with civil society organisations, researchers and feminist networks who often lead the production of gender-responsive data, especially where state data systems are incomplete or exclusionary.
- Governments and international organisations should fund, recognise and partner with non-state actors in the development of safe, inclusive and context-aware data ecosystems that support TFGBV prevention and AI accountability.

Section 7. Conclusions

Addressing TFGBV in the age of AI requires a shift from reactive, siloed regulation toward integrated, rights-based governance. By rights-based governance, we mean an approach that places human rights – including privacy, equality, freedom from violence and non-discrimination – at the centre of AI design, development, deployment and oversight.

Taken together, the recommendations outlined in this report aim to strengthen both the technical and normative foundations of AI and TFGBV governance, advancing feminist digital justice and intersectional accountability.

Annexes

Glossary

The definitions below are provided for clarity and consistency in this report. They reflect a combination of commonly accepted usage, feminist digital justice frameworks and relevant policy sources. Where definitions are directly drawn from external sources, attribution is noted.

Artificial intelligence (AI)

A field of computer science focused on creating systems that can perform tasks typically requiring human intelligence, such as pattern recognition, language processing or decision making.

Algorithmic amplification

The process by which algorithms – especially those used in recommendation systems – intensify the visibility and spread of certain types of content, including gendered disinformation, hate speech or abuse. Often driven by engagement metrics, this amplification can disproportionately target or harm women and marginalised groups.

Algorithmic bias

Systematic and repeatable errors in AI outputs that reflect and reinforce existing societal prejudices, often due to biased training data or assumptions in model design.

Data extractivism

As defined by APC, it is the massive extraction of user data by private companies such as Meta, Google, Amazon and Microsoft, often without appropriate privacy checks and protocols.[149] This collection of data – a form of surveillance capitalism similar to the extraction of raw materials and exploitation of labour in colonised territories in the Global South that enriched states and companies based in the North – is then used to sell content and advertising back to users, crippling the openness of the internet and its usefulness for building democracy and social justice. APC's concern with extractivism extends to all forms of the extractive economy that operates in the production and use of ICTs, including the extraction of raw materials and the impact of this on local communities.

Digital governance

As defined by APC, this is a term currently used more frequently instead of "internet governance", taken to involve the governance of a broader spectrum of digital services and infrastructures in a time of rapid digitalisation and datafication.[150]

Deepfakes and non-consensual synthetic content

Digitally manipulated images, videos or audio generated using AI (often, deep learning), typically to depict people in fabricated situations. When used non-consensually, they represent a form of TFGBV.

[149] https://www.apc.org/en/glossary/data-extractivism

[150] https://www.apc.org/en/glossary/digital-governance

Feminist Principles of the Internet

As defined by APC, it is a series of statements that offer a gender and sexual rights lens on critical internet-related rights.[151] They were drafted at the first Imagine a Feminist Internet meeting that took place in Malaysia in April 2014, organised by APC, with 50 activists and advocates working in sexual rights, women's rights, violence against women and internet rights.

Gender digital divide

As per APC, access is not just about being online.[152] Access is about the inclusion and meaningful participation of all women and gender-diverse people in the digital space and decision making at all levels of use, design, management and governance of digital technologies. A feminist perspective on access allows us to consider the gendered impact of exclusion and access, thinking about "access" intersectionally and holistically and as a domain of power in any society. According to the ITU, there are four types of gender digital divides or four main categories of the global digital gender divide: a gap in access and use of the internet; a gap in digital skills and use of digital tools; a gap in participation in STEM fields; and a gap in tech sector leadership and entrepreneurship.[153]

Gender

Gender refers to the set of ideas, representations, practices and social prescriptions based on the anatomical difference between the sexes. These ideas and practices create social, economic and legal hierarchies in society that result in discrimination and inequality.[154]

Gender norms are changeable over time; they inform individual identities, social relations and the distribution of resources and power in society. Although gender is often understood as expressing expectations regarding appropriate behaviour for men and women, gender is non-binary and diverse. It refers to people of all gender identities and expressions. Gender equality therefore refers to equal rights, opportunities and outcomes for men, women, girls, boys and people of diverse gender identities and expressions. Equal treatment on the basis of gender is a human right enshrined in international law.[155]

Gendered disinformation

As per APC, there is no one definition of gendered disinformation that is broadly agreed on or commonly accepted.[156] The term is often used interchangeably with TFGBV or is referred to as the gender dimensions of disinformation. Disinformation itself remains an expression that lacks a single agreed-upon definition and is too often conflated with other concepts, such as propaganda and advocacy to incite discrimination, violence and hostility. The UN Special Rapporteur on freedom of expression and opinion has expressed concern with the growing use of manipulation, deception and the distortion of information, aimed at creating confusion. These conceptual challenges testify to the fact that disinformation is a multifaceted phenomenon.

[151] https://www.apc.org/en/glossary/feminist-principles-internet

[152] https://www.apc.org/en/glossary/gender-digital-divide

[153] International Telecommunication Union. (2023). Bridging the gender divide. https://www.itu.int/en/mediacentre/backgrounders/Pages/bridging-the-gender-divide.aspx

[154] Finlay, A. (2023). Op. cit.

[155] Global Partners Digital. (2023). Op. cit.

[156] Martins, P. (2024). Op. cit.

Gender-responsive

Gender-responsive policymaking extends beyond merely addressing inequalities within specific policies; it actively designs policies that contribute to advancing gender equality. Gender-responsive policymaking is increasingly considered international good practice.[157] Because gender inequality is intersectional, rather than a single issue, this report comments on gender alongside other patterns of historical or contextual intersectional marginalisation, where relevant, including potential implications for minority communities.[158]

Gendered digital harms

Harms that disproportionately affect individuals based on their gender in digital or technology-mediated environments. These include online harassment, deepfakes, non-consensual data collection, biased algorithms and exclusion from digital access or decision-making processes.

Gender analysis

A structured process for assessing the differentiated impacts of policies, technologies or programmes on people based on gender and often other intersecting identities. In AI governance, gender analysis can inform more equitable data use, design processes and risk assessments.

Generative AI

A subset of AI technologies capable of creating new content – such as text, images or audio – based on training data. Includes tools like large language models (LLMs) and image generators.

Global South

As per APC, the region traditionally known as the developing world that encompasses the majority of humankind. APC advocates on behalf of marginalised peoples in the Global South.[159] However, its use of the term "Global South" is relatively fluid. It refers to issues of social justice and the marginalisation of people and communities in countries typically identified as in the Global South, but includes developed countries in the "Global North" where similar and relevant social justice issues might emerge or groups and communities have allied experiences. For example, this includes the marginalisation of Indigenous peoples in Canada or the social exclusions faced by Black people in New York or working-class people in London. APC's membership is predominantly from the Global South, but a number of allies in the Global North are also members.

Government frameworks

As per the Global Index on Responsible AI, these are laws, regulations, policies, strategies and/or guidelines adopted by the national or federal government that address the implications of AI with respect to a particular thematic area.[160]

[157] See: UN Women. (2021). *UN Women Strategic Plan 2022-2025. Building a Gender-Equal World*. https://www.unwomen.org/en/digital-library/publications/2021/09/un-women-strategic-plan-2022-2025

[158] Ferrari, V., & Millar, K. (2025). *A Novel Approach to the 11 UN Norms for Responsible State Behaviour in Cyberspace: Guidelines for Gendered Implementation*. Organization of American States – Inter-American Committee against Terrorism (CICTE) & UN Institute for Disarmament Research. https://unidir.org/wp-content/uploads/2025/05/UNIDIR_Novel_Approach_11_UN_Norms_Responsible_State_Behaviour_Cyberspace_Guidelines_Gendered_Implementation.pdf

[159] https://www.apc.org/en/glossary/global-south

[160] https://www.global-index.ai/methodology

Intersectionality

The intersectional perspective identifies a system of diverse and interconnected oppressions – among them gender, but which include issues such as race, religion and class – that creates sometimes complex social, economic and other hierarchies among people in a society. Individuals are rarely subject to one form of oppression on its own.[161]

Technology-facilitated gender-based violence (TFGBV)

As per APC, technology-facilitated gender-based violence (TFGBV) – such as cyberstalking, online harassment and doxxing – encompasses acts of gender-based violence that are committed, abetted or aggravated, in part or fully, by the use of information and communication technologies (ICTs), such as phones, the internet, social media platforms and email.[162]

[161] Finlay, A. (2023). Op. cit.

[162] https://www.apc.org/en/glossary/technology-facilitated-gender-based-violence

List of sources consulted[163]

- Acheson, R. (2025). *WILPF Inputs to the UN Secretary General's Report on Artificial Intelligence in the Military Domain*. Women's International League for Peace and Freedom. https://docs-library.unoda.org/General_Assembly_First_Committee_-Eightieth_session_(2025)/79-239-WILPF-EN.pdf

- Ada Lovelace Institute. (2022). *Countermeasures: The need for new legislation to govern biometric technologies in the UK*. https://www.adalovelaceinstitute.org/wp-content/uploads/2025/06/Ada-Lovelace-Institute-Countermeasures-020625.pdf

- Ali Hussen, D. (2023, 26 March). "Dystopian" surveillance "disproportionately targets young, female and minority workers." *The Guardian*. https://www.theguardian.com/global-development/2023/mar/26/dystopian-surveillance-disproportionately-targets-young-female-minority-workers-ippr-report

- Amnesty International. (n/d). Online Violence. https://www.amnesty.org/en/what-we-do/technology/online-violence/

- Amnesty International. (2018). *Toxic Twitter: A Toxic Place for Women*. https://www.amnesty.org/en/latest/research/2018/03/online-violence-against-women-chapter-1-1/

- Amnesty International. (2019). Surveillance Giants: How the business model of Google and Facebook threatens human rights. https://www.amnesty.org/en/tech/surveillance-giants/

- Amnesty International. (2022, 29 September). Myanmar: Facebook's systems promoted violence against Rohingya; Meta owes reparations – new report. https://www.amnesty.org/en/latest/news/2022/09/myanmar-facebooks-systems-promoted-violence-against-rohingya-meta-owes-reparations-new-report/

- Amnesty International. (2024). *Coded injustice: Surveillance and discrimination in Denmark's automated welfare state*. https://www.amnesty.org/en/documents/eur18/8709/2024/en/

- Amnesty International. (2025). *Human rights implications of technology-facilitated gender-based violence. Submission to the Human Rights Council Advisory Committee*. https://www.amnesty.org/en/documents/ior40/9284/2025/en/

- Amnesty International. (2025, 17 February). Meta's new content policies risk fueling more mass violence and genocide. https://www.amnesty.org/en/latest/news/2025/02/meta-new-policy-changes/

- Association for Progressive Communications. (2023). *The Feminist Principles for including gender in the GDC*. https://www.apc.org/sites/default/files/femprinciples-gdc.pdf

- Association for Progressive Communications. (2023). *Joint Submission to the Global Digital Compact on Gender*. https://www.apc.org/sites/default/files/gdc_joint_submission_on_gender_final.pdf

- Azcona, G. et al. (2023). *Progress on the Sustainable Development Goals: The gender snapshot 2023*. UN Women. https://www.unwomen.org/sites/default/files/2023-09/progress-on-the-sustainable-development-goals-the-gender-snapshot-2023-en.pdf

[163] This list includes all sources consulted during the research, including those explicitly cited in the text as well as additional materials that informed the analytical framing and contextual understanding of the study.

- Buolamwini, J., & Gebru, T. (2018). Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. *Proceedings of Machine Learning Research, 81*, 1-15. https://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf
- Center for Democracy & Technology. (2019). *AI & Machine Learning*. https://cdt.org/ai-machine-learning/?ref=internet.exchangepoint.tech
- Chandler, K. (2021). *Does Military AI have Gender? Understanding bias and promoting ethical approaches in military applications of AI*. United Nations Institute for Disarmament Research. https://unidir.org/publication/does-military-ai-have-gender-understanding-bias-and-promoting-ethical-approaches-in-military-applications-of-ai/
- Das, S. (2024, 27 January). Facial recognition cameras in supermarkets "targeted at poor areas" in England. *The Guardian*. https://www.theguardian.com/uk-news/2024/jan/27/facial-recognition-cameras-in-supermarkets-targeted-at-poor-areas-in-england
- Das, D. (2025, 20 May). The Binary Is Glitchy: Platform Accountability Through a Decolonial Queer Lens. *GenderIT.org*. https://genderit.org/feminist-talk/binary-glitchy-platform-accountability-through-decolonial-queer-lens
- Derechos Digitales. (2023, 29 June). *Reflexiones feministas para el desarrollo de inteligencia Artificial*. https://www.derechosdigitales.org/noticias/reflexiones-feministas-para-el-desarrollo-de-inteligencia-artificial/
- Derechos Digitales et al. (2025). *Applying a gender lens to the implementation of the UNGPs in the digital age. Joint comments to the draft paper*. https://www.derechosdigitales.org/wp-content/uploads/JOINT-COMMENTS-TO-THE-DRAFT-PAPER-OHCHR-B-Tech-Gender-Lens_May2025_Finalversion.pdf
- Dunn, S. (2020). *Technology-Facilitated Gender-Based Violence: An Overview*. Centre for International Governance Innovation. https://www.cigionline.org/static/documents/SaferInternet_Paper_no_1_coverupdate.pdf
- Dunn, S., Vaillancourt, T., & Brittain, H. (2023). *Supporting Safer Digital Spaces*. Centre for International Governance Innovation. https://www.cigionline.org/static/documents/SaferInternet_Special_Report.pdf
- Enock, F. E. et al. (2025). Gendered Inequalities in Online Harms: Fear, Safety Work, and Online Participation. *arXivLabs*. https://arxiv.org/abs/2403.19037
- Equality Now. (2025, 11 December). Six key points to understand the new Inter-American Model Law on Digital Violence against Women. https://equalitynow.org/news/news-and-insights/six-key-points-to-understand-the-new-inter-american-model-law-on-digital-violence-against-women/
- European Institute for Gender Equality. (2022). *Combating Cyber Violence against Women and Girls*. https://eige.europa.eu/publications-resources/publications/combating-cyber-violence-against-women-and-girls?language_content_entity=en
- Felsberger, S. (2025). *The High Stakes of Tracking Menstruation*. MCTD Cambridge. https://www.mctd.ac.uk/femtech-high-stakes-tracking-menstruation/
- Feminist Principles of the Internet. (n/d). https://feministinternet.org/principles
- Ferrari, V., & Millar, K. (2025). *A Novel Approach to the 11 UN Norms for Responsible State Behaviour in Cyberspace: Guidelines for Gendered Implementation*. Organization of American States – Inter-American Committee against Terrorism (CICTE) & UN Institute for

Disarmament Research. https://unidir.org/wp-content/uploads/2025/05/UNIDIR_Novel_Approach_11_UN_Norms_Responsible_State_Behaviour_Cyberspace_Guidelines_Gendered_Implementation.pdf

- Finlay, A. (2023). *A framework for developing gender-responsive cybersecurity policy: Assessment tool*. Association for Progressive Communications. https://www.apc.org/sites/default/files/apcgendercyber-assessmenttool.pdf
- FIRN team. (n/d). Understanding of technology-facilitated gender-based violence beyond social media-centred analysis. https://firn.genderit.org/blog/understanding-technology-facilitated-gender-based-violence-beyond-social-media-centred
- Fournier-Tombs, E. et al. (2024). *Artificial Intelligence and the Women, Peace and Security Agenda in South-East Asia*. UN Women Asia and the Pacific. https://asiapacific.unwomen.org/en/digital-library/publications/2024/05/artificial-intelligence-and-the-women-peace-and-security-agenda
- Freedom Online Coalition. (2025). Joint Statement on Artificial Intelligence and Human Rights. https://freedomonlinecoalition.com/joint-statement-on-ai-and-human-rights-2025
- GenderIT.org. (2023). *Global Attention to TFGBV*: Feminist Perspectives. https://www.genderit.org/edition/global-attention-technology-facilitated-gender-based-violence-tfgbv-feminist-perspectives
- GenderIT.org. (2023, 28 July). Algorithmic anxieties and Feminist futures in MENA. https://www.genderit.org/edition/algorithmic-anxieties-feminist-futures-mena
- Gibson, K. (2025, 9 January). Gender bias, AI, and deepfakes are promoting misogyny online. https://blogs.lse.ac.uk/wps/2025/01/09/gender-bias-ai-and-deepfakes-are-promoting-misogyny-online/
- Global Center on AI Governance. (2025). *Global Index on Responsible AI. Gender Equality Thematic Area*. https://www.global-index.ai/thematic-areas-Gender-Equality
- Global Partners Digital. (2023). *Inclusive Cyber Norms Toolkit*. https://www.gp-digital.org/wp-content/uploads/2023/08/Inclusive-Cyber-Norms-Toolkit_GPD.pdf
- Global Partnership for Action on Gender Based Online Harassment and Abuse. (2023). *Technology-Facilitated Gender-Based Violence: Preliminary Landscape Analysis*. https://assets.publishing.service.gov.uk/media/64abe2b21121040013ee6576/Technology_facilitated_gender_based_violence_preliminary_landscape_analysis.pdf
- Global Partnership on Artificial Intelligence. (2025). *Policy Guide for Implementing Transformative AI Policy Recommendations*. https://wp.oecd.ai/app/uploads/2025/05/policy-guide-for-implementing-transformative-AI-policy-recommendations-2.pdf
- GSMA. (2025). *The Mobile Gender Gap Report 2025*. https://www.pta.gov.pk/assets/media/2025-05-20-The-Mobile-Gender-Gap-Report-2025.pdf
- Gurumurthy, A., & Barthur, D. (2023). *Reframing AI Governance through a Political Economy Lens*. IT for Change. https://itforchange.net/reframing-ai-governance-through-a-political-economy-lens
- Haliburton, L., Leusmann, J., Welsch, R. et al. (2025). Uncovering labeler bias in machine learning annotation tasks. *AI and Ethics 5*, 2515-2528. https://doi.org/10.1007/s43681-024-00572-w

- Hoegg, E. (2025). Technology-Facilitated Gender-Based Violence (TFGBV) and Generative Artificial Intelligence (AI). *Resistance, 7*(1). https://ojs-o.library.ubc.ca/index.php/thatswhatwesaid/article/view/622
- Howson, K. (2025). *Human Rights and AI: A Global Index on Responsible AI Analytical Report Series. Brief 6: Gender Equality*. Global Center on AI Governance. https://genderequalitybrief6.tiiny.site/
- Human Rights Watch. (2024, 10 September). Questions and Answers: Israeli Military's Use of Digital Tools in Gaza. https://www.hrw.org/news/2024/09/10/questions-and-answers-israeli-militarys-use-digital-tools-gaza#_What_is_%E2%80%9CLavender%E2%80%9D
- International Telecommunication Union. (2024). *Facts and Figures 2024: The gender digital divide*. https://www.itu.int/itu-d/reports/statistics/2024/11/10/ff24-the-gender-digital-divide/
- Irfan, A. (2025, 14 March). Generative AI and Deepfakes: is there a way forward? *GenderIT.org*. https://genderit.org/feminist-talk/generative-ai-and-deepfakes-there-way-forward
- IT for Change. (2023). *Input to the UN Human Rights B-Tech Project*. https://itforchange.net/sites/default/files/2440/Input%20to%20the%20UN%20Human%20Rights%20B-Tech%20Project%20.pdf
- IT for Change & Development Alternatives with Women for a New Era (Dawn). (2023). *The Declaration of Feminist Digital Justice*. https://feministdigitaljustice.net/wp-content/uploads/2023/03/JNC-WG-Declaration-of-Feminist-Digital-Justice_2023.pdf
- Jankowicz N., Gómez-O'Keefe, I., Hoffman, L., & Vidal Becker, A. (2024). *It's Everyone's Problem: Mainstreaming Responses to Technology-Facilitated Gender-Based Violence*. Institute of Global Politics & Vital Voices Global Partnership. https://igp.sipa.columbia.edu/sites/igp/files/2024-09/IGP_TFGBV_Its_Everyones_Problem_090524.pdf
- Jansen Reventlow, N. (2021, 29 May). How Artificial Intelligence Impacts Marginalised Groups. https://digitalfreedomfund.org/how-artificial-intelligence-impacts-marginalised-groups/
- Kamran, H. (2025, 18 June). Resisting extraction and centring justice in feminist futures for AI. *GenderIT.org*. https://genderit.org/feminist-talk/resisting-extraction-and-centring-justice-feminist-futures-ai
- Khan, S. (2017, 21 November). Surveillance as a Feminist Issue. *Privacy International*. https://privacyinternational.org/news-analysis/3376/surveillance-feminist-issue
- Kira, B. (2024, 3 June). Deepfakes, the Weaponisation of AI Against Women and Possible Solutions. https://verfassungsblog.de/deepfakes-ncid-ai-regulation/
- Kohnert, D. (2022). Machine Ethics and African Identities: Perspectives of Artificial Intelligence in Africa. *SSRN Electronic Journal*. http://dx.doi.org/10.2139/ssrn.4163096
- Kornweitz, A. (2021, 4 August). A New AI Lexicon: Function Creep. *AI Now Institute*. https://ainowinstitute.org/publications/collection/a-new-ai-lexicon-function-creep
- Lee, C. et al. (2024). People who share encounters with racism are silenced online by humans and machines, but a guideline-reframing intervention holds promise. *Proceedings of the National Academy of Sciences, 121*(38). https://doi.org/10.1073/pnas.2322764121
- Martins, P. (2024). *Placing "gender" in disinformation*. Association for Progressive Communications. https://www.apc.org/sites/default/files/genderDisinformation.pdf

- Millar, M. (2019, 30 October). Facial recognition technology struggles to see past gender binary. *Reuters*. https://www.reuters.com/article/world/facial-recognition-technology-struggles-to-see-past-gender-binary-idUSKBN1X92OC/

- Mittelstadt, B. (2019). Principles alone cannot guarantee ethical AI, *Nature Machine Intelligence, 1*(11), 501-507. https://www.researchgate.net/publication/337015694_Principles_alone_cannot_guarantee_ethical_AI

- Nabulega, S., Achieng, G., & Borokini, F. (2021). *Engendering AI: A Gender and Ethics Perspective on Artificial Intelligence in Africa. Pollicy*. https://pollicy.org/resource/engendering-artificial-intelligence/

- Naim, N. (2023, 28 July). Women in the era of artificial intelligence: Increased targeting and growing challenges. *GenderIT.org*. https://genderit.org/articles/women-era-artificial-intelligence-increased-targeting-and-growing-challenges

- Ndemo, B. (2024, 8 August). Addressing digital colonialism: A path to equitable data governance. *UNESCO Inclusive Policy Lab*.

- Niazi, M. (2024). *How do current AI regulations shape the global governance framework? Digital Policy Hub – Working Paper*. Center for International Governance Innovation. https://www.cigionline.org/static/documents/DPH-paper-Maral_Niazi.pdf

- O'Donnell, R. M. (2019). Challenging Racist Predictive Policing Algorithms under the Equal Protection Clause. *NYU Law Review, 94*(3). https://nyulawreview.org/wp-content/uploads/2019/06/NYULawReview-94-3-ODonnell.pdf

- Organisation for Economic Co-operation and Development. (2025a). *Governing with Artificial Intelligence. The State of Play and Way Forward in Core Government Functions*. OECD Publishing. https://www.oecd.org/en/publications/governing-with-artificial-intelligence_795de142-en.html

- Organisation for Economic Co-operation and Development. (2025b). *Steering AI's Future: Strategies for Anticipatory Governance. OECD Artificial Intelligence Papers No. 32*. OECD Publishing. https://www.oecd.org/content/dam/oecd/en/publications/reports/2025/02/steering-ai-s-future_70e4a856/5480ff0a-en.pdf

- Pauwels, E. (2020). *Artificial Intelligence and Data Capture Technologies in Violence and Conflict Prevention: Opportunities and Challenges for the International Community*. Global Center on Cooperative Security. https://www.jstor.org/stable/resrep27551?seq=2

- Pavel, V. (2022). *Rethinking data and rebalancing digital power* Ada Lovelace Institute. https://www.adalovelaceinstitute.org/wp-content/uploads/2022/11/Ada-Lovelace-Institute-Rethinking-data-and-rebalancing-digital-power-FINAL.pdf

- Perera, S. (2022). *White paper on feminist internet research*. Association for Progressive Communications. https://firn.genderit.org/sites/default/files/2022-11/FIRN-whitepaper-2022.pdf

- Pírková, E., Sampieri, A., & Zaghdoud, A. (2024). *Platform accountability: A rule-of-law checklist for policymakers*. Access Now. https://www.accessnow.org/wp-content/uploads/2024/12/Platform-accountability-a-rule-of-law-checklist-for-policymakers-report-2024.pdf

- Popli, N. (2021, 26 October). The 5 most important revelations from the "Facebook Papers". *Time*. https://time.com/6110234/facebook-papers-testimony-explained/

- Posetti, J., & Shabbir, N. (2022). *The Chilling: A global study of online violence against women journalists.* International Center for Journalists. https://www.icfj.org/sites/default/files/2023-02/ICFJ%20Unesco_TheChilling_OnlineViolence.pdf
- Ramos, G. (2022, 22 August). Why we must act now to close the gender gap in AI. World Economic Forum. https://www.weforum.org/stories/2022/08/why-we-must-act-now-to-close-the-gender-gap-in-ai/
- Rutgers. (2025). *Decoding Technology-Facilitated Gender-Based Violence: A Reality Check from Seven Countries*. https://rutgers.international/wp-content/uploads/2024/06/Decoding-TFGBV-Report-2024.pdf
- Salami, A. O. (2024). Artificial intelligence, digital colonialism, and the implications for Africa's future development. *Data & Policy, 6*. https://doi.org/10.1017/dap.2024.75
- Security Hero. (2023). *2023 State of Deepfakes. Realities, Threats and Impact*. https://www.securityhero.io/state-of-deepfakes/
- Shtaya, M. (2023, 13 November). How Meta's platforms normalize anti-Palestinian racism. *Middle East Institute*. https://www.mei.edu/publications/how-metas-platforms-normalize-anti-palestinian-racism
- Stauffer, B. (2025). *A Hazard to Human Rights: Autonomous Weapons Systems and Digital Decision-Making*. Human Rights Watch. https://www.hrw.org/report/2025/04/28/a-hazard-to-human-rights/autonomous-weapons-systems-and-digital-decision-making
- Stop Killer Robots Campaign. https://www.stopkillerrobots.org/military-and-killer-robots/
- Timcke, S., & Makumbirofa, S. (2024, 22 May). Five things to know about the political economy of AI. *Research ICT Africa*. https://researchictafrica.net/2024/05/22/five-things-to-know-about-the-political-economy-of-ai/
- UN Women. (2021). *Strategic Plan 2022-2025. Building a Gender Equal World*. https://www.unwomen.org/en/digital-library/publications/2021/09/un-women-strategic-plan-2022-2025
- UN Women. (2024). *Artificial intelligence and gender equality. Explainer*. https://www.unwomen.org/en/articles/explainer/artificial-intelligence-and-gender-equality
- UN Women. (2025). *Partnering for Gender-Responsive AI*. https://www.unwomen.org/en/digital-library/publications/2025/01/advancing-gender-equality-through-partnerships-for-gender-responsive-artificial-intelligence
- UN Women. (2025, 5 February). How AI reinforces gender bias – and what we can do about it. Interview with Zinnya del Villar on AI gender bias and creating inclusive technology. https://www.unwomen.org/en/news-stories/interview/2025/02/how-ai-reinforces-gender-bias-and-what-we-can-do-about-it
- UNESCO. (2020). Artificial intelligence and gender equality: Key findings of UNESCO's Global Dialogue, https://unesdoc.unesco.org/ark:/48223/pf0000374174
- UNESCO. (2023). *"Your opinion doesn't matter, anyway." Exposing Technology-Facilitated Gender-based Violence in an era of Generative AI*. https://www.unesco.org/sites/default/files/medias/fichiers/2024/04/Ai-livre-EN-web.pdf

- UNESCO. (2023, 7 March). International Women's Day: New factsheet highlights gender disparities in innovation and technology. https://www.unesco.org/en/articles/international-womens-day-new-factsheet-highlights-gender-disparities-innovation-and-technology
- UNESCO-International Research Centre on Artificial Intelligence. (2024). *Challenging Systematic Prejudices: An Investigation into Gender Bias in Large Language Models*. https://unesdoc.unesco.org/ark:/48223/pf0000388971
- United Nations AI Advisory Body. (2024). *Governing AI for Humanity: Final report*. https://www.un.org/sites/un2.un.org/files/governing_ai_for_humanity_final_report_en.pdf
- Venema, A. E. (n/d). Deepfakes as a Security Issue: Why Gender Matters. *Women in International Security*. https://wiisglobal.org/deepfakes-as-a-security-issue-why-gender-matters/
- Verma, P. (2023, 5 November). "AI fake nudes are booming. It's ruining real teens' lives." *The Washington Post*. https://www.washingtonpost.com/technology/2023/11/05/ai-deepfake-porn-teens-women-impact/
- Viteri Almeida, D. (2024, 25 July). Artificial Intelligence and Gender-Based Public Policies. *The Organization for World Peace*. https://theowp.org/artificial-intelligence-and-gender-based-public-policies/
- Ward, J., Spencer, S., & Kalsi, K. (2023). *Gender-Based Violence and Artificial Intelligence (AI): Opportunities and Risks for Women and Girls in Humanitarian Settings*. Social Development Direct. https://www.sddirect.org.uk/resource/gender-based-violence-and-artificial-intelligence-ai-opportunities-and-risks-women-and
- West, M., Kraut, R., & Ei Chew, H. (2019). *"I'd Blush If I Could." Closing gender divides in digital skills through education*. UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000367416
- Wilkinson, I. , Hofstetter, J-S., Shires, J., & Siba Yahaya, M. (2024). *The role of the private sector in combatting gendered cyber harms*. Royal Institute of International Affairs. https://www.chathamhouse.org/sites/default/files/2024-06/2024-06-03-private-sector-gendered-cyber-harms-wilkinson-hofstetter-shires-yahaya_0.pdf
- WITNESS. (2025, 6 March). Deepfakes and Digital Abuse: Dismantling Technology-Facilitated Gender-Based Violence. https://blog.witness.org/2025/03/technology-facilitated-gender-based-violence/
- Wu, C. (2024, 24 September). The Impact of Generative AI on Technology-Facilitated Gender-Based Violence: Part I. *Human Rights Research Center*. https://www.humanrightsresearch.org/post/the-impact-of-generative-ai-on-technology-facilitated-gender-based-violence-part-i
- Yavuz, C. (2025). Adverse human rights impacts of dissemination of nonconsensual sexual deepfakes in the framework of European Convention on Human Rights: A victim-centered perspective. *Computer Law & Security Review, 56*. https://doi.org/10.1016/j.clsr.2025.106108

Instruments

International

- African Commission on Human and Peoples' Rights. (2022). *Resolution on the Protection of Women Against Digital Violence in Africa – ACHPR/Res. 522 (LXXII) 2022*. https://achpr.au.int/en/adopted-resolutions/522-resolution-protection-women-against-digital-violence-africa-achpr
- African Commission on Human and Peoples' Rights. (2024). *Resolution on the Need to Undertake a Study on Digital Violence against women's Rights In Africa – ACHPR/Res.591 (LXXX) 2024*. https://achpr.au.int/en/adopted-resolutions/achprres591-lxxx-2024-resolution-need-undertake-study-digital-violence-against
- African Union. (2003). *Protocol to the African Charter on Human and Peoples' Rights on the Rights of Women in Africa (Maputo Protocol)*. https://au.int/sites/default/files/treaties/37077-treaty-charter_on_rights_of_women_in_africa.pdf
- African Union. (2024). *Continental Artificial Intelligence Strategy. Harnessing AI for Africa's* "Development and Prosperity." https://au.int/sites/default/files/documents/44004-doc-EN-_Continental_AI_Strategy_July_2024.pdf
- Alston, P. (2019). *Report of the Special Rapporteur on extreme poverty and human rights (Digital welfare states and human rights).* UN General Assembly. (A/74/493). https://docs.un.org/en/A/74/493
- Ashwini K.P.A (2024). *Contemporary forms of racism, racial discrimination, xenophobia and related intolerance. Report of the Special Rapporteur on contemporary forms of racism, racial discrimination, xenophobia and related intolerance, Ashwini K.P.A.* UN Human Rights Council. (A/HRC/56/68). https://docs.un.org/en/A/HRC/56/68
- Cannataci, J. A. (2021). *Artificial intelligence and privacy, and children's privacy. Report of the Special Rapporteur on the right to privacy, Joseph A. Cannataci.* UN Human Rights Council. (A/HRC/46/37). https://docs.un.org/en/A/HRC/46/37
- Council of Europe. (2011). *Council of Europe Convention on preventing and combating violence against women and domestic violence (Istanbul Convention)*. https://rm.coe.int/168008482e
- Council of Europe. (2024). *Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law*. https://rm.coe.int/1680afae3c
- European Union. (2023). *Hiroshima Process International Code of Conduct for Organizations Developing Advanced AI Systems.* https://digital-strategy.ec.europa.eu/en/library/hiroshima-process-international-code-conduct-advanced-ai-systems
- European Union. (2023). *Hiroshima Process International Guiding Principles for Organizations Developing Advanced AI Systems*. https://digital-strategy.ec.europa.eu/en/library/hiroshima-process-international-guiding-principles-advanced-ai-system
- European Union. (2024). *Directive (EU) 2024/1385 of the European Parliament and of the Council of 14 May 2024 on combating violence against women and domestic violence*. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:L_202401385
- European Union. (2024). *EU Artificial Intelligence Act*. https://artificialintelligenceact.eu/
- Organization for Economic Cooperation and Development. (2019). *OECD AI Principles*. https://oecd.ai/en/ai-principles

- Organization of American States. (1994). *Inter-American Convention on the Prevention, Punishment and Eradication of Violence Against Women (Convention of Belém do Pará).* https://www.oas.org/juridico/english/treaties/a-61.html

- Organization of American States. (2025). *Inter-American Model Law to Prevent, Punish and Eradicate Gender-based Digital Violence Against Women (MESECVI, 2025).* https://belemdopara.org/cim_mesecvi/inter-american-model-law-to-prevent-punish-and-eradicate-gender-based-digital-violence-against-women-mesecvi-2025/

- Šimonović, D. *Report of the Special Rapporteur on violence against women, its causes and consequences on online violence against women and girls from a human rights perspective.* UN Human Rights Council. (A/HRC/38/47). https://docs.un.org/en/A/HRC/38/47

- UN Commission on the Status of Women. (2023). *CSW67 Agreed Conclusions. Innovation and technological change, and education in the digital age for achieving gender equality*. https://www.unwomen.org/sites/default/files/2023-03/CSW67_Agreed%20Conclusions_Advance%20Unedited%20Version_20%20March%202023.pdf

- UN General Assembly. (2018). *Resolution 73/173. Intensification of efforts to eliminate all forms of violence against women and girls*. (A/RES/78/173) https://docs.un.org/en/A/RES/73/148

- UN General Assembly. (2020). *Resolution 75/161. Advancing women's rights in the digital age*. (A/RES/75/161). https://docs.un.org/en/A/RES/75/161

- UN General Assembly. (2023). *Resolution 78/213. Promotion and protection of human rights in the context of digital technologies*. (A/RES/78/213). https://docs.un.org/en/A/RES/78/213

- UN General Assembly. (2025). *Artificial intelligence procurement and deployment: ensuring alignment with the Guiding Principles on Business and Human Rights. Report of the Working Group on the issue of human rights and transnational corporations and other business enterprises.* (A/HRC/59/53). https://docs.un.org/en/A/HRC/59/53

- UN Human Rights & B-Tech Project. (2025). *UNGPs Compass for Tech Regulation: A Policymaker's Guide on how to apply the UNGPs to require rights-respecting corporate responsibility.* https://www.ohchr.org/en/documents/tools-and-resources/ungps-compass-tech-regulation-policymakers-guide-how-apply-ungps

- UN Human Rights Council. (2018). *Resolution adopted by the Human Rights Council on 5 July 2018. 38/5. Accelerating efforts to eliminate violence against women and girls: preventing and responding to violence against women and girls in digital contexts.* (A/HRC/RES/38/5). https://docs.un.org/en/A/HRC/RES/38/5

- UN Human Rights Council. (2021). *Resolution adopted by the Human Rights Council on 13 July 202147/23. New and emerging digital technologies and human rights*. https://docs.un.org/en/A/HRC/RES/47/23

- UN Human Rights Council. (2022). *Practical application of the Guiding Principles on Business and Human Rights to the activities of technology companies*. (A/HRC/50/56). https://docs.un.org/en/A/HRC/50/56

- UN Human Rights Council. (2023). *Resolution adopted by the Human Rights Council on 14 July 2023. 53/29. New and emerging digital technologies and human rights*. https://docs.un.org/en/A/HRC/RES/53/29

- UNESCO. (2022). *Recommendation on the Ethics of Artificial Intelligence, adopted on 23 November 2021*. https://www.unesco.org/en/articles/recommendation-ethics-artificial-intelligence
- United Nations. (2024). *Global Digital Compact*. https://www.un.org/digital-emerging-technologies/global-digital-compact

National

- Australia – Department of Industry, Science and Resources. (2019). *Australia's AI Ethics Principles*. https://www.industry.gov.au/publications/australias-artificial-intelligence-ethics-principles/australias-ai-ethics-principles
- Australia – Australian Human Rights Commission. (2025). *Technology-facilitated gender-based violence Australian Human Rights Commission Submission to Human Rights Council Advisory Committee*. https://www.ohchr.org/sites/default/files/documents/hrbodies/hrcouncil/advisorycommittee/cfi-neurotechnology/subm-technology-facilitated-gender-nhri-australian-hr-commission.pdf
- Brazil – Brazilian AI Strategy. Ministério da Ciência, Tecnologia e Inovações Secretaria de Empreendedorismo e Inovação. (n/d). *Estratégia Brasileira de Inteligência Artificial –EBIA*. https://www.gov.br/mcti/pt-br/acompanhe-o-mcti/transformacaodigital/ebia.pdf
- Bulgaria – Concept for the Development of Artificial Intelligence in Bulgaria until 2030. https://www.mtc.government.bg/sites/default/files/20pr0724pr.pdf
- Canada – Treasury Board of Canada Secretariat. (2019). *Directive on Automated Decision-Making*. https://www.tbs-sct.canada.ca/pol/doc-eng.aspx?id=32592
- Chile – National Artificial Intelligence Policy. Ministerio de Ciencia, Tecnología, Conocimiento e Innovación, Gobierno de Chile. (2021). *Política Nacional de Inteligencia Artificial*. https://www.minciencia.gob.cl/uploads/filer_public/bc/38/bc389daf-4514-4306-867c-760ae7686e2c/documento_politica_ia_digital_.pdf
- China – Digital Policy Office. (2025). *Ethical Artificial Intelligence Framework*. https://www.digitalpolicy.gov.hk/en/our_work/data_governance/policies_standards/ethical_ai_framework/
- Croatia – National Plan for Gender Equality, for Period until 2027. Government of the Republic of Croatia. (2017). *Nacionalni plan za ravnopravnost spolova, za razdoblje do 2027*. https://ravnopravnost.gov.hr/UserDocsImages//dokumenti/NPRS%202027%20APRS%202024//Nacionalni%20plan%20za%20ravnopravnost%20spolova,%20za%20razdoblje%20do%202027..pdf
- Dominican Republic – National Artificial Intelligence Strategy. Gobierno de la República Dominicana. (2023). *ENIA – Estrategia Nacional de Inteligencia*. https://agendadigital.gob.do/wp-content/uploads/2023/10/Final_ENIA-Estrategia-Nacional-de-Inteligencia-Artificial-de-la-Republica-Dominicana-.pdf
- Germany – Federal Foreign Office. (2023). Shaping *Feminist Foreign Policy. Federal Foreign Office Guidelines*. https://www.shapingfeministforeignpolicy.org/papers/Guidelines_Feminist_Foreign_Policy.pdf
- Greece – Law 4961/2022. Emerging information and communication technologies, strengthening digital governance and other provisions. Available in Greek at: https://search.et.gr/el/fek/?fekId=618783

- Ireland – Government of Ireland. (2021). *AI – Here for Good. A National Artificial Intelligence Strategy for Ireland*. https://enterprise.gov.ie/en/publications/publication-files/national-ai-strategy.pdf
- Italy – National Strategy for Artificial Intelligence. Ministerio dello viluppo economico. (2020). *Strategia Nazionale per l'Intelligenza Artificiale*. https://www.mimit.gov.it/images/stories/documenti/Strategia_Nazionale_AI_2020.pdf
- Japan – The Conference toward AI Network Society. (2019). *AI Utilization Guidelines*. https://www.soumu.go.jp/main_content/000658284.pdf
- Jordan – Ministry of Digital Economy and Entrepreneurship. (2022). *The National Artificial Intelligence Code of Ethics*. https://dig.watch/resource/the-national-artificial-intelligence-code-of-ethics-of-jordan
- Malaysia – Malaysia Science and Technology Information Centre. (2023). Malaysia *National Artificial Intelligence Roadmap 2021-2025 (AI-RMAP)*. https://dig.watch/resource/malaysia-national-artificial-intelligence-roadmap-2021-2025
- Morocco – Summary report on the results of monitoring and national consultation meetings on the protection of human rights in digital and artificial intelligence systems. Bensalah, M., & National Human Rights Council. (2022). *Rapport de synthèse des résultats du monitoring et des rencontres nationales de concertation sur la protection des droits humains dans le digital et les systèmes de l'intelligence artificielle*. https://www.researchgate.net/publication/359982730_Rapport_de_synthese_des_resultats_du_monitoring_et_des_rencontres_nationales_de_concertation_sur_la_protection_des_droits_humains_dans_le_digital_et_les_systemes_de_l%27intelligence_artificielle
- The Netherlands – Emancipation: a Task for All of Us. Ministerie van Onderwijs, Cultuur en Wetenschap. (2022). *Emancipatie: een opdracht voor ons allen*. https://open.overheid.nl/documenten/ronl-9442234d31a1e83aaed7b1a7dece2205bb92e2fe/pdf
- Portugal – Guide to an Ethical, Transparent and Responsible Artificial Intelligence in Public Administration. Agência para a Modernização Administrativa. (2025). *GuIA – Guia para a Inteligência Artificial*. https://digital.gov.pt/en/documentos/guia-para-a-inteligencia-artificial
- Saudi Arabia – Saudi Data and Artificial Intelligence Authority. (2024). *AI Ethics Principles*. https://sdaia.gov.sa/en/SDAIA/about/Documents/ai-principles.pdf
- Senegal – National Strategy and Roadmap for Senegal on Artificial Intelligence. Ministère de la Communication, des Télécommunications et de l'Économie numérique. (2023). *Stratégie nationale et feuille de route du Sénégal sur l'intelligence artificielle*. https://africadataprotection.org/sources/Synth%C3%A8se%20de%20la%20strat%C3%A9gie%20IA%20du%20S%C3%A9n%C3%A9gal.pdf
- Serbia – Government of Serbia. (2023). *Ethical Guidelines for Development, Implementation and Use of Robust and Accountable Artificial Intelligence*. https://www.ai.gov.rs/extfile/en/471/Ethical%20guidelines%20for%20development%20implementation%20and%20use%20of%20robust%20and%20accountable%20AI.pdf
- Singapore – Personal Data Protection Commission. (2020). *Model Artificial Intelligence Governance Framework. Second edition*. https://www.pdpc.gov.sg/-/media/Files/PDPC/PDF-Files/Resource-for-Organisation/AI/SGModelAIGovFramework2.pdf

- Slovenia – Government of the Republic of Slovenia. (2021). *National Programme to Promote the Development and Use of Artificial Intelligence in the Republic of Slovenia by 2025 (NpAI).* https://www.gov.si/assets/ministrstva/MDP/National_Programme_for_AI_2025.pdf
- Spain – Government of Spain. (2020). *ENIA –National Artificial Intelligence Strategy*. https://portal.mineco.gob.es/RecursosArticulo/mineco/ministerio/ficheros/National-Strategy-on-AI.pdf
- United Kingdom – UK Government. (2023). *Online Safety Act*. https://www.legislation.gov.uk/ukpga/2023/50
- United States – The White House. (2021). *National Strategy on Gender Equity and Equality.* https://bidenwhitehouse.archives.gov/wp-content/uploads/2021/10/National-Strategy-on-Gender-Equity-and-Equality.pdf

Bridging the gap:

Addressing technology-facilitated gender-based violence in global AI governance

APC
ASSOCIATION FOR
PROGRESSIVE
COMMUNICATIONS