

Submission by the Association for Progressive Communications (APC) on the issue of the impact of disinformation on the enjoyment and realisation of human rights

Human Rights Council Advisory Committee (HRC resolution 55/10): Call for inputs

APC welcomes the opportunity to provide input via the call published by the HRC Advisory Committee on the impact of disinformation on the enjoyment and realisation of human rights, which aims to collect contributions that will be used in the preparation of a study to be presented to the HRC during its 61st session.

According to HRC Resolution 55/10, the study will seek to review the methods used to disseminate disinformation and will promote tools and approaches to counter these challenges while protecting and reinforcing human rights standards.

APC is an international network organisation dedicated to empowering and supporting people working for peace, human rights, development and the protection of the environment, through the strategic use of information and communication technologies (ICTs). The APC network has 74 organisational members and 41 associates active in 74 countries, mostly in the global South. We work to build a world in which all people have easy, equal and affordable access to the creative potential of ICTs to improve their lives and create more democratic and egalitarian societies.

This submission follows the questionnaire published by the HRC Advisory Committee as part of the call.

1. What human rights are impacted the most by disinformation? Provide specific examples.

Disinformation has profound and far-reaching impacts on human rights, with certain fundamental rights being particularly affected due to the nature and targets of disinformation campaigns. Disinformation should be “understood as a symptom of much broader information disorders” that have “acquired new dimensions – in terms of reach, speed and volume – with the expansion” of digital technologies and social media.¹

Freedom of expression and access to information

One of the most directly impacted rights is freedom of expression, including the right to seek and receive information. According to APC's analysis, "disinformation causes confusion and has a chilling effect on freedom of expression and information. It directly impacts on the level of trust in the public sphere as a space for democratic deliberation."² People become hesitant to express their views for fear of online harassment or being targeted by disinformation campaigns, while others withdraw from public debate entirely due to the uncertainty created by information pollution.

Disinformation campaigns are characterised by coordination and malign intention, including false or harmful content that exploits gender inequalities or weaponises gender stereotypes. APC's report *Placing “gender” in disinformation* documents how 71% of online harm incidents occur on social media platforms, where algorithmic recommendation systems and attention-based business models actively amplify harmful content.³

¹ APC. (2021). *APC policy explainer: Disinformation*. Association for Progressive Communications. <https://www.apc.org/en/pubs/apc-policy-explainer-disinformation>

² Martins, P., et al. (2024). *Placing “gender” in disinformation*. Association for Progressive Communications. <https://www.apc.org/sites/default/files/genderDisinformation.pdf>

³ Martins, P., et al. (2024). *Op. cit.*

The chilling effect imposed by disinformation is particularly evident in cases of gendered disinformation targeting women in public life. Women politicians, journalists and rights defenders face particularly nasty and broad-reaching coordinated disinformation campaigns designed to "silence and restrict the credibility of their work and the support of public opinion" for their causes and struggles.⁴ These attacks often move beyond mere censorship to pose genuine dangers to targets' safety and lives, as well as threatening their families and mental health. This not only has an impact on the women targeted and their families, but also creates serious disincentives to other women taking on public life roles and work.

Right to political participation

Disinformation severely undermines the right to participate in public affairs and democratic processes. Our research shows how disinformation campaigns are often deployed to deter women and marginalised groups from political participation. Women politicians face "sexualised rumours" and body shaming intended to "weaken their credibility, and over time to erode public trust and delegitimise their leadership."⁵ In some cases, these attacks have been so destructive that women political leaders have been forced to withdraw from politics entirely.

The relationship between disinformation and the right to political participation is particularly striking when examined through an intersectional lens.⁶ Accounts considered low in credibility, including bots and trolls, attack female political candidates at higher rates than their male counterparts. The impact is even more severe for women from minority communities – the report indicates how female political leaders and activists

⁴ Baltazar, F. (2024, 6 June). How disinformation targets women and environmental journalists. *Association for Progressive Communications*. <https://www.apc.org/en/blog/how-disinformation-targets-women-and-environmental-journalists>

⁵ Finlay, A. (2024, 17 October). Integrating policy, research and technical standards in gender approaches to cybersecurity: Key takeaways from a recent APC-hosted roundtable. *Association for Progressive Communications*. <https://www.apc.org/en/news/integrating-policy-research-and-technical-standards-gender-approaches-cybersecurity-key>

⁶ APC. (2021). Op. cit.

from racial, ethnic, religious or other minority groups are targeted far more than their white colleagues.⁷

Right to non-discrimination

Disinformation frequently targets and exacerbates discrimination against already marginalised groups by exploiting existing social divides and tension points. The compounding challenges imposed by intersecting identities heighten both vulnerability to attacks and resulting harms. Studies cited in Martins et. al.'s report show that Black, Asian and minority ethnic women politicians receive almost half of abusive content despite representing a much smaller percentage of political figures.⁸

In Sri Lanka, for example, during the COVID-19 pandemic, state-sponsored disinformation specifically targeted the Muslim minority community, stoking communal hatred and supporting discriminatory policies like the mandatory cremation of COVID-19 victims.⁹

Right to life, liberty and security

In its most extreme manifestations, disinformation can threaten a person's fundamental right to life and security. APC's publications have addressed how online disinformation campaigns can translate into real-world violence, particularly when amplified by state actors. Delegitimation campaigns based on false or distorted information are very commonly used against human rights defenders, "social leaders" and other community activists in Latin America.¹⁰ This is particularly concerning when combined with findings about state-sponsored gendered disinformation campaigns being accompanied by surveillance and spyware attacks against women journalists and activists.¹¹ The real-

⁷ Martins, P., et al. (2024). Op. cit.

⁸ Ibid.

⁹ Corea, H. (2022, 10 August). In Sri Lanka, state-sponsored disinformation and suppression of dissent taint COVID-19 response. *Association for Progressive Communications*. <https://www.apc.org/en/news/sri-lanka-state-sponsored-disinformation-and-suppression-dissent-taint-covid-19-response>

¹⁰ APC. (2021). Op. cit.

¹¹ Martins, P., et al. (2024). Op. cit.

world violence caused by online disinformation campaigns has been demonstrated more recently in the case of Palestine, where "atrocities propaganda" and "gendered disinformation" have been used to justify violence against civilians.¹²

Right to privacy

Privacy rights are often compromised in the context of disinformation campaigns, particularly through doxxing and the non-consensual sharing of personal information. Disinformation attacks frequently involve "the non-consensual access, use, manipulation, and dissemination of private data, information and/or content," creating permanent digital records that can be widely distributed and are extremely difficult to remove.¹³ Platforms' business models, based on the expropriation of personal data, make women and gender-diverse individuals more vulnerable to privacy violations and data breaches – we have documented how deepfake technology is predominantly being used to create non-consensual sexual content targeting women, with 96% of deepfake videos portraying women without their consent.¹⁴

Economic and social rights

Disinformation can also impact economic and social rights. Disinformation thrives where public information regimes are weak and independent investigative journalism is constrained.¹⁵ This impacts people's ability to access reliable information about health, education and economic opportunities. The monetisation of disinformation through advertising revenue further compounds these impacts. In the economic sphere, disinformation can affect livelihoods and economic participation.

¹² APCNews. (2024, 8 March). Upcoming webinar on Palestinian feminist voices: Atrocities propaganda and gendered disinformation. *Association for Progressive Communications*. <https://www.apc.org/en/news/upcoming-webinar-palestinian-feminist-voices-atrocities-propaganda-and-gendered-disinformation>

¹³ APC. (2023). *Feminist Principles of the Internet: Advocacy brief on violence*. <https://genderit.org/FPI-paper-on-violence>

¹⁴ Martins, P., et al. (2024). Op. cit.

¹⁵ APC. (2021). Op. cit.

During the COVID-19 pandemic, health-related disinformation directly affected people's right to health. The COVID-19 pandemic created the ideal conditions for disinformation to thrive due to two key factors. First, it rapidly triggered widespread fear, uncertainty and doubt across the globe. Second, it emerged at a time when information – both accurate and misleading – could be easily created, shared and accessed through the internet, mobile communication, traditional media and social media platforms. As the pandemic spread, social media and instant messaging became flooded with posts questioning treatments, vaccine safety and effectiveness and the value of social distancing. This fuelled social unrest, delayed vaccine adoption and, in some cases, contributed to higher mortality rates.¹⁶

Online misinformation and disinformation contribute to increased discord within and fragmentation of movements, affecting collective action for social and economic justice. Disinformation can introduce conflicting narratives within a movement, causing members to disagree on goals, strategies or even basic facts.¹⁷ Fake news or misleading information can undermine solidarity and weaken the movement from within. This also has external effects – the spreading of false or extreme narratives about a movement can alienate potential supporters. If disinformation paints activists as violent, irrational or dangerous, the broader public may disengage, reducing momentum and legitimacy for the struggles for social and economic justice.

Right to cultural life and religious freedom

Disinformation often targets cultural and religious practices, affecting people's right to observe and enjoy their religion and participate in cultural life. Our research provides examples such as the spreading of false narratives about Muslim burial practices during COVID-19 in Sri Lanka, which were used to justify policies that violated religious rights.¹⁸ Disinformation can spread false narratives about religious groups, portraying them as

¹⁶ APC. (2021, 25 February). Op. cit.

¹⁷ APC. (2024, 4 September). The Full Picture: APC launches campaign to help understand and tackle misinformation and disinformation. *Association for Progressive Communications*.
<https://www.apc.org/en/press/full-picture-apc-launches-campaign-help-understand-and-tackle-misinformation-and>

¹⁸ Corea, H. (2022, 10 August). Op. cit.

dangerous, extremist or a threat to national security. This can lead to social stigma, discrimination or even legal restrictions on religious practices. Governments or extremist groups may use disinformation to justify crackdowns on religious minorities.

Intersectional impacts on rights

It's crucial to understand that disinformation's impact on human rights is often intersectional, affecting multiple rights simultaneously and disproportionately impacting those who face multiple forms of marginalisation. Social and legal conditions do not exist in isolation, but instead interact with and reinforce each other. Individuals who belong to multiple marginalised groups face particularly severe impacts. For example, women from ethnic minority backgrounds who are also activists or human rights defenders face layered targeting that combines racial, gender-based and politically-motivated disinformation.¹⁹ As highlighted by Martins, P. et al., "female political leaders and activists from racial, ethnic, religious or other minority groups are targeted far more often than their white colleagues."²⁰

Our work at APC emphasises that disinformation's impact on human rights is systemic rather than isolated. "APC views disinformation as a multifaceted, global and complex issue that should be understood as a symptom of much broader information disorders."²¹ The systemic nature of disinformation's impact means that responses must address not just individual instances of disinformation but the broader structures that enable its spread and impact. Gendered disinformation flourishes in societies where women's freedom of expression is constrained, creating a vicious cycle of rights violations. Religious narratives, cultural beliefs and geopolitical tensions are exploited to target women and gender-diverse people, particularly affecting those from minority communities.²²

¹⁹ Wijesiriwardena, S. (2024). *The machine that fosters shame: The weaponisation of sexuality in anti-gender anti-democracy disinformation*. Association for Progressive Communications. <https://www.apc.org/en/pubs/machine-fosters-shame-weaponisation-sexuality-anti-gender-anti-democracy-disinformation>

²⁰ Martins, P., et al. (2024).

²¹ APC. (2021). Op. cit.

²² Martins, P., et al. (2024). Op. cit.

The impacts are particularly severe when disinformation is state-sponsored or supported by powerful actors. As seen in cases from various regions, when states engage in or support disinformation campaigns, it becomes significantly more difficult for individuals and communities to protect their rights or seek redress for violations. This is exemplified in contexts like Sri Lanka, where state-sponsored disinformation combined with surveillance and the suppression of dissent created a particularly hostile environment for human rights.²³

To effectively protect human rights from disinformation's impacts, our documentation suggests a need for comprehensive approaches that combine regulatory frameworks, platform accountability, digital literacy and support for marginalised communities, recognising the latter's agency and autonomy when creating counternarratives and actions. This must include the recognition of gendered disinformation as a specific phenomenon requiring targeted responses, while simultaneously addressing the broader systemic issues that enable disinformation to flourish and impact human rights.

Looking ahead, the emergence of new technologies, particularly artificial intelligence, presents additional challenges for protecting human rights from disinformation's impacts. The impacts of disinformation have been shown to be exacerbated by platforms' business models and algorithmic amplification: APC's research highlights how artificial intelligence and automated systems can amplify gendered disinformation through recommendation algorithms that prioritise engagement over safety. AI systems can replicate and exacerbate existing inequalities and discriminations, potentially amplifying the harmful effects of disinformation on human rights.²⁴ In another example, "algorithmic news feeds craft automatically generated, highly personalised adversarial content streams that keep users engaged" and monetised while corrupting "the entire global information ecosystem."²⁵ These systems can particularly amplify harmful content targeting vulnerable groups, creating feedback loops that intensify marginalisation.

²³ Corea, H. (2022, 10 August). Op. cit.

²⁴ APC. (2024, 2 April). AI to advance gender equality: Challenges and opportunities. *Association for Progressive Communications*. <https://www.apc.org/en/pubs/ai-advance-gender-equality-challenges-and-opportunities>

²⁵ Martins, P., et al. (2024). Op. cit.

To effectively protect human rights, responses must address both the immediate harms of disinformation and the structural conditions that enable its spread.

2. Which forms of disinformation do you consider to be particularly problematic for human rights and why? Provide specific examples.

Several forms of disinformation pose particularly serious threats to human rights, with some variants proving especially harmful due to their systematic nature and targeted impacts on vulnerable populations. According to APC, disinformation must be understood as a multifaceted, global issue that manifests as a symptom of broader information disorders, making certain forms particularly problematic for human rights.²⁶

Gendered disinformation

Gendered disinformation represents one of the most concerning forms, combining "falsity, malign intent, and coordination" to specifically target women and gender-diverse people.²⁷ According to Martins et al.'s report, this form of disinformation is particularly harmful because it exploits gender inequalities and weaponises gender stereotypes against women and gender-diverse individuals, deterring them from participating in the public sphere. The report documents how 71% of online harm incidents occur on social media platforms, where algorithmic recommendation systems and attention-based business models actively amplify this harmful content.²⁸

The impact of gendered disinformation becomes even more severe when it intersects with existing forms of discrimination. Women politicians face coordinated attacks that

²⁶ APC. (2021). Op. cit.

²⁷ APC. (2023, 19 September). Op. cit.

²⁸ Martins, P., et al. (2024). Op. cit.

combine sexualised rumours and body shaming, aimed at eroding their credibility and legitimacy.²⁹ This aligns with findings that highlight how disinformation campaigns particularly target women engaging politically or culturally.³⁰

Technology-facilitated gender-based violence (TFGBV) that incorporates disinformation elements represents a particularly insidious form of gendered disinformation. This often involves the non-consensual manipulation and distribution of personal information combined with false narratives designed to harm women and gender-diverse people.³¹ Of particular concern is the use of deepfake technology, with 96% of deepfake videos portraying women without their consent.³²

State-sponsored disinformation

State-sponsored disinformation represents another highly problematic form due to its systematic nature and the power imbalance between perpetrators and targets. Government-backed disinformation campaigns have targeted minority communities while simultaneously suppressing counternarratives through surveillance and intimidation.³³ The weaponisation of state resources becomes especially concerning when combined with other forms of repression, as seen in Palestine.³⁴

Environmental disinformation

While digital technology can enable environmental defenders, Indigenous peoples and other marginalised groups to bypass the information barriers that traditional media, the state and private actors may impose, the capabilities that ICTs provide have been

²⁹ Finlay, A. (2024, 17 October). Op. cit.

³⁰ Di Meco, L., & Brechenmacher, S. (2020, 30 November). Tackling Online Abuse and Disinformation Targeting Women in Politics. *Carnegie Endowment for International Peace*. <https://carnegieendowment.org/research/2020/11/tackling-online-abuse-and-disinformation-targeting-women-in-politics?lang=en>

³¹ Raghavan, S., & Hussen, T. S. (2023, 24 August). Global attention to technology-facilitated gender-based violence (TFGBV): Feminist perspectives. *GenderIT*. <https://genderit.org/editorial/global-attention-technology-facilitated-gender-based-violence-tfgbv-feminist-perspectives>

³² Martins, P., et al. (2024). Op. cit.

³³ Corea, H. (2022, 10 August). Op. cit.

³⁴ APC News. (2024, 8 March). Op. cit.

misused by malign actors to spread false or misleading information about the climate crisis and environmental degradation, and have also been abused to coordinate organised online attacks and disinformation campaigns against environmental defenders and affected communities.³⁵ Environmental journalists and activists face targeted disinformation campaigns that not only discredit their work but also pose serious risks to their safety.³⁶ This aligns with findings documenting how online disinformation campaigns can translate into real-world violence.³⁷

Crisis-related disinformation

Crisis-related disinformation is particularly problematic because it can justify or incite violence against vulnerable populations. As seen in the case of Rohingya refugees, false narratives about refugee communities can fuel hatred and incite real-world violence against already marginalised groups.³⁸ Similarly, crisis situations often exacerbate existing gendered disinformation patterns.³⁹

Identity-based disinformation

We consider forms of disinformation that target individuals and groups facing multiple, intersecting forms of marginalisation to be the most problematic. Disinformation campaigns often combine multiple forms of discrimination. Studies show that Black, Asian and minority ethnic women politicians receive almost half of abusive content despite representing a much smaller percentage of political figures.⁴⁰

³⁵ Association for Progressive Communications. (2022). *At the interstice of digital rights and environmental justice: Four issue briefs to inform funding*. Ford Foundation. <https://www.fordfoundation.org/wp-content/uploads/2022/07/apc-01-07-22.pdf>

³⁶ Baltazar, F. (2024, 6 June). Op. cit.

³⁷ Wine, B. (2022, 12 October). Researchers to Study Connection Between Online Misinformation and Real-World Violence. *Georgia Tech*. <https://www.gatech.edu/news/2022/10/12/researchers-study-connection-between-online-misinformation-and-real-world-violence>

³⁸ APC News. (2024, 4 December). Countering misinformation, disinformation and hate speech targeting Rohingya refugees. *Association for Progressive Communications*. <https://www.apc.org/en/news/countering-misinformation-disinformation-and-hate-speech-targeting-rohingya-refugees>

³⁹ Kapantai, E., Christopoulou, A., Berberidis, C., & Peristeras, V. (2021). A systematic literature review on disinformation: Toward a unified taxonomical framework. *New Media & Society*, 23(5), 1301-1326. <https://doi.org/10.1177/1461444820959296>

⁴⁰ Wijesiriwardena, S. (2024). Op. cit.

It is also important to consider that different forms of disinformation usually work in concert, creating complex webs of false narratives that are difficult to untangle and counter. The emergence of artificial intelligence presents new and potentially more problematic forms of disinformation. AI systems can not only amplify existing biases but also create new forms of manipulated content that are increasingly difficult to detect and counter.⁴¹ These technological developments, combined with the expansion of social media, have given disinformation new dimensions “in terms of reach, speed and volume.”⁴² Furthermore, the role of algorithmic systems in spreading and amplifying disinformation poses unique challenges. Platforms’ “content curation builds on the profiling and micro-targeting of individuals, with the ultimate goal of serving platforms’ advertising purposes.” This automated amplification is particularly problematic because “algorithmic news feeds craft automatically generated, highly personalised adversarial content streams that keep users engaged” and monetised while corrupting “the entire global information ecosystem” as clarified by Khoo and already mentioned above.⁴³

To effectively address these particularly problematic forms of disinformation, responses must recognise both their immediate impacts and their broader implications for human rights and understand how different forms of disinformation interact with and reinforce one another, while also developing targeted responses that address the specific challenges posed by each form.

⁴¹ APC. (2024, 2 April). Op. cit.

⁴² APC. (2021). Op. cit.

⁴³ Martins, P., et al. (2024). Op. cit.; Khoo, C. (2018, 1 July). The revolution will not be automated. *Canadian Centre for Policy Alternatives*. <https://www.policyalternatives.ca/news-research/the-revolution-will-not-be-automated/>

3. How can some measures to counter disinformation negatively impact human rights' enjoyment? Provide specific examples.

While measures to counter disinformation are necessary, some approaches can inadvertently or deliberately undermine human rights, often affecting marginalised communities disproportionately. Responses must address both immediate manifestations and underlying systemic issues, yet many current approaches fail to do so and can actually harm those they claim to protect.

Broad anti-disinformation laws

One of the most concerning trends is the implementation of broad anti-disinformation laws that can be misused to suppress legitimate expression. Many countries have passed legislation ostensibly aimed at combating disinformation, legislation which instead becomes a tool for silencing critics and marginalised voices. Existing research identifies “cases in Saudi Arabia, Cuba, Egypt, Jordan, Libya, Nicaragua, Russia, Uganda and Venezuela” where cybercrime laws have been weaponised against women and LGBTQIA+ people.⁴⁴ Such laws often disproportionately impact women and gender-diverse individuals, and activists and outspoken individuals on gender-related themes are repeatedly accused of promoting “gender ideology”, making them vulnerable to prosecution under broad anti-disinformation laws.⁴⁵

⁴⁴ Derechos Digitales. (2023, 30 June). When protection becomes threat: Cybercrime regulation as a tool for silencing women and LGBTQIA+ people around the world. *Association for Progressive Communications*. <https://www.apc.org/en/news/when-protection-becomes-threat-cybercrime-regulation-tool-silencing-women-and-lgbtqia-people>

⁴⁵ Martins, P., et al. (2024). Op. cit.

Content moderation policies

Content moderation policies implemented by social media platforms, while necessary, can have unintended negative consequences. Automated systems often fail to understand context and nuance, leading to the removal of legitimate content, particularly content created by marginalised communities.⁴⁶ Language limitations are also a concern, since policies are applied with double standards and overall, poorly in non-English speaking environments. These problems are also compounded by platforms' lack of transparency and accountability regarding their content moderation practices.⁴⁷

Surveillance and data collection

The issue of surveillance and data collection as counter-disinformation measures raises particular concerns. Counter-disinformation efforts during the pandemic led to expanded surveillance powers that were then used to target critics and minority communities.⁴⁸ An analysis of the gendered impact of such surveillance shows how these measures create additional risks for women and gender-diverse people.⁴⁹

Internet shutdowns and digital blackouts

Internet shutdowns and digital blackouts, sometimes implemented under the guise of preventing the spread of disinformation, represent another problematic countermeasure. Governments increasingly use "digital darkness" as a control mechanism, particularly during periods of social unrest or political transition.⁵⁰ These shutdowns prevent the

⁴⁶ Malhotra, N. A. (2019, 23 September). Are we any better at judging right from wrong? Automation in content moderation. *GenderIT*. <https://genderit.org/articles/are-we-any-better-judging-right-wrong-automation-content-moderation>

⁴⁷ APC. (2021). Op. cit.

⁴⁸ Corea, H. (2022, 10 August). Op. cit.

⁴⁹ Raghavan, S., & Hussen, T. S. (2023, 24 August). Op. cit.

⁵⁰ Sirnate, V., & Jain, G. (2023, 3 April). Democracy is dying in digital darkness. *Association for Progressive Communications*. <https://www.apc.org/en/blog/democracy-dying-digital-darkness>

documentation and reporting of human rights violations, effectively silencing victims while allowing other narratives to dominate.⁵¹

Criminalisation

The criminalisation of certain forms of speech (for example, through cyber laws and other related legislation) in the name of fighting disinformation often leads to self-censorship and the suppression of legitimate discourse. Anti-disinformation measures can particularly impact feminists and human rights defenders who are already struggling against systemic oppression.⁵²

Negative impacts for human rights defenders

Counter-disinformation measures can also negatively impact those working to defend digital rights. Digital rights defenders face increased risks when anti-disinformation laws are used to target their work.⁵³ Such measures can have a chilling effect on civil society organisations and human rights defenders, particularly those working on sensitive issues.⁵⁴

Economic impacts

The economic impacts of some counter-disinformation measures can indirectly affect human rights while failing to address root causes. Platforms' business models and automated advertising systems continue to amplify problematic content, yet many

⁵¹ Sype, E. (2024). The Meta-Israel nexus: Silencing Palestinian voices in the digital landscape. *Association for Progressive Communications*. <https://dev-d10.apc.org/en/event/meta-israel-nexus-silencing-palestinian-voices-digital-landscape>

⁵² Sívori, H., & Mochel, L. (2022, 22 November). Brazilian feminist responses to online hate speech: Seeing online violence through an intersectional lens. *Association for Progressive Communications*. <https://www.apc.org/en/pubs/brazilian-feminist-responses-online-hate-speech-seeing-online-violence-through-intersectional>

⁵³ Abrougui, A. (2023, 28 July). Is Civil Society In MENA Ready To Tackle AI's Human Rights Challenges? *GenderIT*. <https://www.genderit.org/index.php/feminist-talk/civil-society-mena-ready-tackle-ais-human-rights-challenges>

⁵⁴ APC. (2021). Op. cit.

regulatory approaches fail to address these fundamental issues.⁵⁵ Economic penalties can disproportionately affect smaller, independent media outlets and community-based platforms that serve marginalised populations.

Authoritarianism

Perhaps most concerningly, some counter-disinformation measures can legitimise authoritarian control over digital spaces. Without proper safeguards, efforts to combat disinformation can become tools for expanding state control over online discourse. This is particularly problematic for marginalised communities who rely on digital platforms for advocacy and community building.⁵⁶

The complex nature of these impacts requires careful consideration in relation to any proposed countermeasures. Through its work, APC advocates for approaches that strengthen democratic institutions and support community-led responses rather than top-down, punitive measures. This includes addressing the structural conditions that enable disinformation to flourish while ensuring that countermeasures do not further marginalise vulnerable communities.⁵⁷

To mitigate these negative impacts, counter-disinformation measures should be designed with human rights at their core, ensuring transparency, accountability and meaningful participation from affected communities, which requires recognising gendered disinformation as a specific phenomenon requiring targeted responses while addressing the broader systemic issues that enable disinformation to flourish and impact human rights.

⁵⁵ Martins, P., et al. (2024). Op. cit.

⁵⁶ Souter, D. (2023, 2 March). Inside the Digital Society: Making platforms accountable. *Association for Progressive Communications*. <https://www.apc.org/en/blog/inside-digital-society-making-platforms-accountable>

⁵⁷ APC. (2021). Op. cit.

4. What do you consider to be the social and legal conditions that may contribute to disinformation and to the impact of disinformation?

Various social and legal conditions interconnect to create environments where disinformation can flourish and maximise its harmful impacts. Disinformation must “be understood as a symptom of much broader information disorders” wherein certain social and legal conditions amplify its effects, particularly on marginalised communities.⁵⁸

Socioeconomic disparities

Pre-existing social inequalities create fertile ground for disinformation to take root and spread. As stated above, disinformation exploits existing social divides and tension points, targeting groups already in situations of marginalisation.⁵⁹ These campaigns thrive where women's freedom of expression is constrained, creating a conducive environment for discrimination, violence and the spread of disinformation. The intersection of multiple forms of marginalisation can amplify vulnerability to disinformation.⁶⁰

Social polarisation and the formation of echo chambers significantly contribute to disinformation's impact. Online misinformation and disinformation contribute significantly to increased discord and fragmentation within movements, creating self-reinforcing cycles of mistrust and division.⁶¹ These conditions can be particularly damaging for marginalised communities who may already face social isolation or exclusion.

The influence of cultural and religious dynamics deserves attention too, as these can also play a significant role in defining the scope of disinformation's impacts. Power

⁵⁸ APC. (2021). Op. cit.

⁵⁹ Martins, P., et al. (2024). Op. cit.

⁶⁰ Raghavan, S., & Hussen, T. S. (2023, 24 August). Op. cit.

⁶¹ APC. (2024, 4 September). Op. cit.

relations between majority and minority religious groups are exploited in disinformation campaigns, extending from religious identity to broader identity politics. This exploitation can be especially effective when combined with existing social prejudices and a lack of cross-cultural understanding.⁶²

Algorithms and models designed without safety-by-design principles or human rights considerations

Platforms' business models and algorithmic amplification represent another critical condition contributing to disinformation and its impact. As already highlighted above, platforms' "content curation builds on the profiling and micro-targeting of individuals, with the ultimate goal of serving platforms' advertising purposes."⁶³ This is especially problematic because it monetises engagement with harmful content. Algorithms designed to maximise engagement favour more extreme content, including disinformation, because it generates more user interaction.⁶⁴ Platform algorithms can amplify gendered disinformation and harassment, creating cycles of harm that are difficult to break.⁶⁵

Platform accountability, or rather its absence, represents another crucial condition. Platforms' lack of transparency regarding algorithms, content moderation practices and advertising systems contributes to the spread of disinformation. This lack of accountability is particularly problematic in regions where platforms have limited local presence or understanding, often failing to account for local contexts and language variations.⁶⁶

⁶² Martins, P., et al. (2024). Op. cit.

⁶³ Ibid.

⁶⁴ Malhotra, N. A. (2019, 23 September). Op. cit.

⁶⁵ Baltazar, F. (2024, 6 June). Op. cit.

⁶⁶ Martins, P., et al. (2024). Op. cit.

Weak legal frameworks

The absence of robust legal frameworks to protect digital rights – for example, privacy and data protection laws – or the existence of laws that can be manipulated to suppress legitimate expression creates conditions where disinformation can thrive. Inadequate or poorly designed legal frameworks can actually enable rather than prevent harm.⁶⁷ This is particularly evident in cases where cybercrime laws are used to silence women and LGBTQIA+ activists who are actually fighting against disinformation.

Digital gaps

Uneven access to digital resources and varying levels of digital literacy create conditions where disinformation can spread more easily. A lack of education and digital literacy leaves women more vulnerable to disinformation because they are less equipped to check sources due to systemic barriers to literacy.⁶⁸ This digital divide is often gendered, with women and marginalised communities having less access to technology and fewer opportunities to develop digital literacy skills.

Authoritarian contexts

Political conditions that favor authoritarian control create environments where disinformation can be weaponised by those in power. Governments increasingly use digital control and disinformation as tools of state power.⁶⁹ The problem is compounded when state-sponsored disinformation combines with other forms of control, including government control over information, and is then used to target minority communities while suppressing counternarratives.⁷⁰

⁶⁷ Derechos Digitales. (2023, 30 June). Op. cit.

⁶⁸ Irfan, A. (2023, 10 January). Societal Barriers To Credible Information Makes Women Unsafe In South Asia. *GenderIT*. <https://genderit.org/feminist-talk/societal-barriers-credible-information-makes-women-unsafe-south-asia>

⁶⁹ Sirnate, V., & Jain, G. (2023, 3 April). Op. cit.

⁷⁰ Corea, H. (2022, 10 August). Op. cit.

The weakening of independent journalism and lack of diverse media voices creates conditions where disinformation can flourish. Healthy information systems require robust access to public information, plural and diverse media contexts and independent journalism, and disinformation thrives where public information regimes are weak, independent investigative journalism is constrained and fundamental rights are disregarded.⁷¹

Addressing these conditions requires comprehensive approaches that combine regulatory frameworks, platform accountability, digital literacy initiatives and support for marginalised communities. We call for an investment in understanding how different communities are affected by these conditions, particularly those who face multiple forms of discrimination and marginalisation, while developing targeted responses that address both immediate manifestations and underlying systemic issues.

5. Which social groups in vulnerable situations are particularly affected by disinformation?

The impact of disinformation is not evenly distributed across society. Certain groups face disproportionate harm due to the intersection of existing vulnerabilities and systemic inequalities, creating layered or intersecting vulnerabilities specific to particular social groups.

Gender

Women and gender-diverse individuals face particularly severe impacts from disinformation. Disinformation campaigns target not only individual women but feminist struggles and gendered discourse more broadly. As referred to above, 71% of online harm incidents occur on social media platforms, where algorithmic recommendation

⁷¹ APC. (2021). Op. cit.

systems and attention-based business models actively amplify harmful content.⁷² Women in public roles face especially intense targeting. Women journalists and politicians face coordinated disinformation campaigns that combine personal attacks with professional discrediting.⁷³

Sexuality

LGBTQIA+ individuals and communities face targeted disinformation that often intersects with other forms of discrimination. Disinformation campaigns exploit and amplify existing prejudices. This targeting becomes particularly severe when combined with other vulnerabilities, such as a person's gender identity or ethnic background.⁷⁴

Religion and ethnicity

Religious and ethnic minorities are particularly vulnerable to disinformation campaigns that exploit existing prejudices. An example is how the Muslim minority community in Sri Lanka faced intensified discrimination through state-sponsored disinformation during the COVID-19 pandemic.⁷⁵ "Different dynamics between majority religious groups and minority ones are exploited in disinformation campaigns, evidencing power relations that move from religious identity to broader identity politics."⁷⁶

Indigenous peoples face specific vulnerabilities to disinformation, particularly when it intersects with land rights and environmental issues. Disinformation exploits existing social divides and tension points, with the compounded challenges resulting from intersecting identities heightening both vulnerability to attacks and resulting harms. This is particularly evident in cases where indigenous women defenders face multiple layers of discrimination and targeting.⁷⁷

⁷² Martins, P., et al. (2024). Op. cit.

⁷³ Baltazar, F. (2024, 6 June). Op. cit.

⁷⁴ Derechos Digitales. (2023, 30 June). Op. cit.

⁷⁵ Corea, H. (2022, 10 August). Op. cit.

⁷⁶ Martins, P., et al. (2024). Op. cit.

⁷⁷ Ibid.

Refugee and migrant communities are particularly vulnerable to disinformation campaigns that exploit xenophobia and nationalism. False narratives have been used to fuel hatred and incite violence against already vulnerable displaced populations.⁷⁸ Women and girls within these communities face additional risks, particularly when disinformation campaigns exploit gender-based stereotypes and prejudices.⁷⁹

Activism

Human rights defenders, particularly those from marginalised communities, face specific targeting through disinformation. Women human rights defenders and activists face compounded threats.⁸⁰ Environmental defenders face particular risks, where disinformation campaigns are used to discredit their work and sometimes put their physical safety at risk.⁸¹

Age

Young people, particularly girls and gender-diverse youth, face specific vulnerabilities to disinformation. Young people's extensive use of digital platforms can make them particularly vulnerable to targeted disinformation.⁸² This vulnerability is compounded for young people who belong to other marginalised groups, as platforms' content curation and micro-targeting can expose them to particularly harmful content.

Ableism

People with disabilities face particular challenges related to disinformation, especially when it intersects with other forms of marginalisation. Disinformation campaigns particularly target those facing multiple forms of discrimination, making people with disabilities who are also members of other marginalised groups especially vulnerable.⁸³

⁷⁸ APC News. (2024, 4 December). Op. cit.

⁷⁹ APC. (2021). Op. cit.

⁸⁰ Raghavan, S., & Hussen, T. S. (2023, 24 August). Op. cit.

⁸¹ Baltazar, F. (2024, 6 June). Op. cit.

⁸² Raghavan, S., & Hussen, T. S. (2023, 24 August). Op. cit.

⁸³ Martins, P., et al. (2024). Op. cit.

This includes women with disabilities, who often face compounded discrimination and targeting. As highlighted by Choudhury, “Threats of violence, gendered shaming and intimidation remain major ways to attack disabled women and women human rights defenders, and gendered disinformation that also uses ableism may attack their appearance, speech, and ways of being to devalue their credibility and humanity. False and misleading information about disabled women, which use ableist stereotypes and prejudices, may perpetuate harmful stereotypes about disabled women and women human rights defenders — portraying them as helpless, dependent, or burdensome — and thus, reduce their credibility in ways that do not apply to non-disabled women.”⁸⁴

Socioeconomic standing

Communities with limited economic resources often face greater vulnerability to disinformation due to reduced access to fact-checking resources and digital literacy training. Economic barriers can limit access to reliable information, particularly affecting women and marginalised communities who may already have limited access to digital resources.⁸⁵

6. What is the role of digitalisation in relation to disinformation? Are there particular technologies you consider relevant and/or problematic in (countering) the dissemination of disinformation (e.g. generative AI)? Explain and provide examples.

Digitalisation has fundamentally transformed how disinformation is created, spread and consumed, while also offering potential tools for countering it. Disinformation “has

⁸⁴ Choudhury, A. (2024). *When Disinformation Weaponises Disability: #DefeatDeceit*. <https://takebackthetech.net/blog/when-disinformation-weaponises-disability-defeatdeceit>

⁸⁵ Irfan, A. (2023, 10 January). Op. cit.

acquired new dimensions – in terms of reach, speed and volume – with the expansion” of digital technologies and social media, making certain technologies particularly significant in both amplifying and potentially combating false narratives.⁸⁶

Digitalisation’s role in/impact on disinformation highlights the need for countermeasures that address both the technological and social aspects of disinformation, including developing better content moderation systems, improving digital literacy and ensuring that technological solutions don’t inadvertently create new vulnerabilities for already marginalised communities. The technological dimensions of disinformation require, as already stated above, the recognition of gendered disinformation as a specific phenomenon requiring targeted responses when addressing the broader systemic issues that enable disinformation to flourish and impact human rights.⁸⁷

Social media platforms and algorithmic amplification

Social media platforms play a central role in the spread of disinformation through their algorithmic content recommendation systems. As already affirmed, platforms’ content curation builds on the “profiling and micro-targeting of individuals,” with “algorithmic news feeds crafting automatically generated, highly personalised adversarial content streams that keep users engaged” and monetised while corrupting “the entire global information ecosystem.”⁸⁸ 71% of online harm incidents occur on social media platforms, where business models prioritise engagement over safety.⁸⁹ These systems particularly affect marginalised communities, with platform algorithms amplifying gendered disinformation and harassment while suppressing legitimate speech from vulnerable groups.⁹⁰

Messaging applications, particularly those with encryption, present unique challenges with regard to the spread of disinformation. For instance, WhatsApp and similar

⁸⁶ APC. (2021). Op. cit.

⁸⁷ Martins, P., et al. (2024). Op. cit.

⁸⁸ Ibid.

⁸⁹ Ibid.

⁹⁰ Souter, D. (2023, 2 March). Op. cit.

platforms can become vectors for the rapid spread of false information, particularly in communities where digital literacy levels are low.⁹¹

Platform governance technologies present both opportunities and challenges. There is a need for increased transparency regarding algorithms, content moderation practices and advertising systems.⁹² When these technologies are developed and implemented particular attention must be paid to their potential impacts on marginalised groups. Technological solutions must be developed with input from affected communities.⁹³

Artificial intelligence and generative technologies

The emergence of artificial intelligence, particularly generative AI, presents new challenges in the disinformation landscape. AI systems can replicate and exacerbate existing inequalities and biases.⁹⁴ This is particularly evident in the creation and spread of deepfake content. Deepfake technology is predominantly being used to create non-consensual sexual content, with 96% of deepfake videos portraying women without their consent.⁹⁵

Surveillance technologies and data collection

Digital surveillance technologies play a complex role in the disinformation ecosystem. Surveillance technologies can be used to suppress legitimate information while enabling state-sponsored disinformation.⁹⁶ Surveillance measures can create additional risks for women and gender-diverse people, especially when personal information collected for "security purposes" can be weaponised for harassment or abuse.

⁹¹ Irfan, A. (2023, 10 January). Op. cit.

⁹² APC. (2021). Op. cit.

⁹³ Baltazar, F. (2024, 6 June). Op. cit.

⁹⁴ APC. (2024, 2 April). Op. cit.

⁹⁵ Martins, P., et al. (2024). Op. cit.

⁹⁶ Syype, E. (2024). Op. cit.

Automated content moderation

Content moderation technologies, while intended to combat disinformation, can sometimes contribute to the problem. Content moderators lack gender-sensitive training and are often alien to local contexts, cultures and languages.⁹⁷ These limitations are particularly evident in non-English language contexts. Content moderation systems often perform poorly when dealing with local languages and cultural contexts, leaving many communities vulnerable to disinformation in their native languages.⁹⁸

Digital infrastructure and access

Broader digital infrastructure, including internet access and connectivity, plays a crucial role in both spreading and potentially countering disinformation. Economic and social barriers can limit access both to reliable information and the tools needed to combat disinformation, particularly affecting women and marginalised communities.⁹⁹ This digital divide creates additional vulnerabilities for those already facing discrimination and marginalisation.

A comprehensive approach that combines technological tools with digital literacy programmes, robust journalism and community-based responses will be needed to address the business models of social media platforms, promoting transparency in content moderation and supporting community-led responses. Emerging counter-technologies offer some promise in combating disinformation, but technological solutions alone cannot address the structural conditions that enable disinformation to flourish.

⁹⁷ Martins, P., et al. (2024). Op. cit.

⁹⁸ Machirori, F. (2022, 5 October). Tackling online hate speech in Africa and beyond: “We can’t trust Big Tech to abide by its own rules”. *Association for Progressive Communications*.
<https://www.apc.org/en/news/tackling-online-hate-speech-africa-and-beyond-we-cant-trust-big-tech-abide-its-own-rules>

⁹⁹ Irfan, A. (2023, 10 January). Op. cit.

7. How can the private sector address disinformation?

The private sector, particularly technology companies and social media platforms, bears significant responsibilities in addressing disinformation while upholding human rights principles. Our research and advocacy work have posited that companies must take comprehensive action that recognises the gendered and intersectional dimensions of disinformation.

Platform accountability and algorithmic reform

A fundamental concern we have identified is that current business models and algorithmic systems actively contribute to the spread of disinformation.¹⁰⁰ Platform economics often prioritise engagement over accuracy and safety, incentivising the spread of sensational and harmful content. These algorithms disproportionately impact women and marginalised communities, who are frequently targeted by coordinated disinformation campaigns.¹⁰¹

The algorithmic amplification of harmful content creates particular challenges for women in public life. Content recommendation systems can rapidly spread gendered disinformation targeting women journalists and politicians, combining personal attacks with professional discrediting.¹⁰² This automated amplification means disinformation campaigns targeting women can reach massive audiences before platforms identify and address them.

We emphasise that companies must fundamentally reconsider these systems, calling for "increased access to data and information held by tech companies to allow us to better understand the phenomenon of disinformation."¹⁰³ This transparency would enable researchers, civil society and users to better comprehend how disinformation spreads

¹⁰⁰ APC. (2021). Op. cit.

¹⁰¹ Martins, P., et al. (2024). Op. cit.

¹⁰² Baltazar, F. (2024, 6 June). Op. cit.

¹⁰³ APC. (2021). Op. cit.

and impacts different communities. Without this transparency, it becomes nearly impossible to hold platforms accountable for their role in amplifying harmful content.

Companies should revise recommendation algorithms that currently amplify extreme content, paying special attention “to long-term issue-based disinformation campaigns, especially those targeted against specific groups and themes, including human rights, women's rights and environmental issues.”¹⁰⁴ This requires shifting from engagement-maximising models toward systems that prioritise information integrity and user safety. Such reforms would particularly benefit women and marginalised communities who currently bear the brunt of the algorithmic amplification of harmful content.

Intersectional approaches to content moderation

In our research and publications, we have addressed how current content moderation systems often fail to understand context and nuance.¹⁰⁵ This deficiency particularly affects marginalised communities, whose experiences and communication patterns may not align with dominant norms embedded in automated systems.

The limitations of these systems become even more pronounced when addressing gendered disinformation. Culturally specific forms of harassment and disinformation often evade platform detection systems.¹⁰⁶ This failure creates environments where harmful narratives targeting women, particularly those with intersecting marginalised identities, can flourish unchecked.

Content moderators often lack gender-sensitive training and are frequently unfamiliar with local contexts, cultures and languages. Martins et al.’s report documents how “reporting mechanisms, when available, are cumbersome ... and often force users to

¹⁰⁴ APC. (2021). *Disinformation and Freedom of Expression: Submission in response to the call by the UN Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression for contributions to the annual thematic report to be presented to the Human Rights Council at its 47th session in June 2021*. Association for Progressive Communications.
<https://www.apc.org/sites/default/files/APCSubmissionDisinformationFebruary2021.pdf>

¹⁰⁵ Malhotra, N. A. (2019, 23 September). Op. cit.

¹⁰⁶ Baltazar, F. (2024, 6 June). Op. cit.

attribute their experiences to predetermined categories that fail to capture the multifaceted nature of the abuse faced.”¹⁰⁷ Women from religious and ethnic minorities face particular challenges in having their experiences recognised within these limited frameworks.

We advocate for companies to adopt a human rights-based approach to content moderation guided by principles of accountability, equality, non-discrimination, participation, inclusion, transparency and empowerment. This approach must be sensitive to how disinformation tactics vary across cultural contexts and target different communities in specific ways. Content moderation systems should be designed with input from diverse communities, particularly those most affected by gendered disinformation.

Language and cultural context

We emphasise the critical importance of addressing language and cultural gaps in disinformation responses. Content moderation often fails in non-English contexts, leaving many communities vulnerable to disinformation in their native languages.¹⁰⁸

This gap creates a significant disparity in protection. While English-language users may receive some safeguards against disinformation, speakers of other languages – particularly in the global South – experience minimal protection. Women and gender-diverse people in these communities face compounded vulnerabilities when gendered disinformation spreads unchecked in their languages.

Language barriers in content moderation particularly affect feminist and human rights defenders who are already struggling against systemic oppression.¹⁰⁹ When platforms fail to understand cultural contexts and nuances of language, legitimate feminist

¹⁰⁷ Martins, P., et al. (2024). Op. cit.

¹⁰⁸ Machirori, F. (2022, 5 October). Op. cit.

¹⁰⁹ Sívori, H., & Mochel, L. (2022, 22 November). Op. cit.

discourse may be wrongly removed while harmful gendered disinformation remains online.

Companies must invest substantially in improving their capacity to moderate content across different languages and cultural contexts. This includes hiring moderators fluent in local languages and familiar with cultural nuances, as well as developing context-sensitive detection tools that can identify disinformation in various linguistic and cultural environments.

Protecting women human rights defenders (WHRDs) and journalists

Our work has highlighted the particular risks faced by WHRDs and journalists targeted by coordinated disinformation campaigns.¹¹⁰ These attacks often combine false information with gendered harassment to discredit and silence women's voices in public discourse.

Private sector actors must develop specific mechanisms to protect these individuals and groups, including expedited review processes for reports from verified human rights defenders and journalists facing targeted attacks. Rapid responses are particularly important when disinformation is state-sponsored and targets minority communities.¹¹¹

Companies should also create specialised teams familiar with the patterns of gendered disinformation used against public-facing women, particularly those from marginalised communities. These mechanisms must be transparent, with clear information about how companies identify and address disinformation targeting vulnerable groups. Response times should be minimised, and affected individuals should receive comprehensive support throughout the process.¹¹²

¹¹⁰ Raghavan, S., & Hussen, T. S. (2023, 24 August). Op. cit.

¹¹¹ Corea, H. (2022, 10 August). Op. cit.

¹¹² Souter, D. (2023, 2 March). Op. cit.

Safety-by-design and preventive approaches

Rather than merely reacting to disinformation after it spreads, we advocate for preventive approaches embedded in platform design. Platform architecture can either enable or prevent the spread of disinformation and related harms.¹¹³

Safety-by-design principles would require companies to conduct thorough gender impact assessments before launching new features or products, considering how they might be exploited to spread gendered disinformation. These assessments should particularly examine impacts on women and gender-diverse people with intersecting marginalised identities, who often face the most severe targeting.

Practical implementations could include developing features that help users identify potential disinformation, creating friction in sharing processes for content flagged as potentially false and implementing systems to detect and address coordinated disinformation campaigns before they cause significant harm. Such preventive approaches are particularly important given how 71% of online harm incidents occur on social media platforms where algorithmic recommendation systems and attention-based business models actively amplify harmful content.¹¹⁴

Investing in digital literacy and civil society collaboration

Our “The Full Picture” campaign underscores the importance of helping users understand how to identify and respond to disinformation.¹¹⁵ Companies should allocate resources to support educational programmes, particularly those targeting vulnerable communities who may have less access to digital literacy education.

¹¹³ APC. (2023, 1 September). Overview of the manifestations and impacts of technology-facilitated gender-based violence and the need for safety by design. *Association for Progressive Communications*. <https://www.apc.org/en/pubs/overview-manifestations-and-impacts-technology-facilitated-gender-based-violence-and-need>

¹¹⁴ Martins, P., et al. (2024). Op. cit.

¹¹⁵ <https://www.apc.org/en/press/full-picture-apc-launches-campaign-help-understand-and-tackle-misinformation-and>

As noted above, lack of education and digital literacy leaves women more vulnerable to misinformation because they are less equipped to check sources due to systemic barriers to literacy.¹¹⁶ Companies have a responsibility to address such disparities by supporting programmes specifically designed to reach women and marginalised communities.

These initiatives are most effective when developed in collaboration with civil society organisations representing affected communities. We stress the importance of including diverse voices when developing policies and solutions for addressing disinformation. This collaboration should include regular consultation with affected communities, support for civil society research and mechanisms for incorporating feedback from marginalised groups into platform policies.

Data protection and resource allocation

The collection and excessive use of personal data can make individuals vulnerable to targeted disinformation attacks.¹¹⁷ Companies must strengthen their data protection practices to prevent such misuse, adopting strong privacy protections while ensuring transparency about data practices.

The risks are particularly acute for women and gender-diverse individuals. Platforms' business models, based on the expropriation of personal data, make women and gender-diverse individuals more vulnerable to privacy violations and data breaches.¹¹⁸

Additionally, platforms need to invest adequately in addressing disinformation in all global contexts, not just in major markets. This includes ensuring appropriate staffing for content moderation in different languages and cultural contexts, and providing comprehensive training and support for moderators, particularly regarding the gendered dimensions of disinformation.

¹¹⁶ Irfan, A. (2023, 10 January). Op. cit.

¹¹⁷ APC. (2024, 2 April). Op. cit.

¹¹⁸ Martins, P., et al. (2024). Op. cit.

Addressing disinformation requires a comprehensive approach from the private sector that goes beyond reactive content removal. Companies must take responsibility for how their platforms and technologies can enable or combat disinformation, while ensuring their actions respect human rights and apply them equally, regardless of identity.

The private sector's role must be part of a broader effort that includes collaboration with civil society, governments and affected communities. Most importantly, companies must be willing to make fundamental changes to their business practices and platform designs to effectively address gendered disinformation while protecting the rights of women and gender-diverse people, particularly those with intersecting marginalised identities.

8. What are good practices at the national, regional, or international level to address the negative effects of disinformation on human rights? How can the integrity of the information systems be strengthened? Please provide examples.

Addressing the harmful effects of disinformation on human rights requires multi-layered approaches that operate at national, regional and international levels. This is particularly true when considering disinformation's disproportionate impact on women and marginalised communities. Drawing from APC's extensive documentation, several effective practices have emerged that strengthen information integrity while protecting fundamental rights, with particular attention paid to gender and intersectionality.

Regulatory frameworks with human rights safeguards

Developing balanced regulatory frameworks is one of the most important national-level interventions. Throughout our research and publications, we argue that any legal frameworks must be built on truly participatory processes and must avoid broad criminalisation provisions that could be weaponised against legitimate expression. We support the OAS' caution against "general prohibitions on the dissemination of information based on vague and ambiguous ideas, including 'false news' or 'non-objective information'," which are incompatible with international standards for freedom of expression.¹¹⁹

Effective regulatory approaches focus on platform accountability rather than content criminalisation. We have documented how cybercrime laws ostensibly designed to protect people have been weaponised against women and LGBTQIA+ communities. In response, we recommend regulatory approaches that incorporate a "gender perspective throughout the conceptualisation, implementation and monitoring of cybercrime laws."¹²⁰

Kenya's National Cohesion and Integration Commission (NCIC) offers an instructive example of institutional monitoring: before Kenya's elections in 2022, the NCIC gave Facebook a seven-day ultimatum to address hate speech and incitement on its platform. Though the commission lacks prosecution power, this approach demonstrates how national institutions can create accountability mechanisms for platforms.¹²¹

The European Union's Digital Services Act provides another model for comprehensive regulation. This regulation requires very large online platforms to conduct systemic risk assessments, including risks related to gender-based violence, and implement mitigation measures.¹²² Such mandatory due diligence obligations could help address forms of gendered disinformation that might not rise to the level of illegal content, but still cause significant harm.

¹¹⁹ Derechos Digitales. (2023, 30 June). Op. cit.

¹²⁰ Ibid.

¹²¹ Machirori, F. (2022, 5 October). Op. cit.

¹²² Martins, P., et al. (2024). Op. cit.

Digital literacy and education initiatives

Comprehensive digital literacy programmes represent another effective intervention at the national level. We specifically advocate for "digital and media literacy programmes to counter information disorders," recommending that such programmes be both independently implemented and "embedded into the regular educational system curricula by states."¹²³

These programmes are particularly effective when they target vulnerable communities. Digital literacy gaps leave women particularly vulnerable to health misinformation.¹²⁴ Effective digital literacy initiatives must address these gaps while considering how social and cultural barriers affect different communities' access to information and education.

Successful examples include Pollicy's work in Uganda promoting women leaders' digital resilience. Their initiatives specifically build resilience against gendered disinformation targeting women in politics, demonstrating how literacy programmes can be tailored to address specific vulnerabilities.¹²⁵

HER Internet in Uganda implemented an exemplary project to create awareness about gendered disinformation, particularly focusing on sexual minorities and sex workers. The organisation convened dialogues to share experiences and strategies to counter disinformation effects, showing how educational initiatives can be tailored to specific community needs.¹²⁶

¹²³ APC. (2021). Op. cit.

¹²⁴ Irfan, A. (2023, 10 January). Op. cit.

¹²⁵ Uiras, M. J. (2024, 18 May). Why gendered disinformation is dangerous for African democracy. *Association for Progressive Communications*. <https://www.apc.org/en/blog/why-gendered-disinformation-dangerous-african-democracy>

¹²⁶ Kyogabirwe, L. (2023, 6 February). Pushing Back Against Gendered Disinformation in Uganda. *CIPESA*. <https://cipesa.org/2023/02/pushing-back-against-gendered-disinformation-in-uganda/>

Independent media support

Supporting independent journalism and fact-checking organisations represents another crucial intervention. Addressing disinformation requires "healthy information systems that include robust access to public information; plural, accessible and diverse media contexts; independent and qualified journalism".¹²⁷ Supporting journalists requires not just general media strengthening but also specific protection mechanisms and resources – in this sense, the effectiveness of independent media support is enhanced when it includes attention to groups targeted specifically. Environmental journalists, for example, face targeted disinformation campaigns that not only discredit their work, but also threaten their safety.¹²⁸

Fact-checking infrastructure

Projects like Tattle and FactCheck.org provide models for systematic disinformation monitoring and debunking.¹²⁹ These initiatives are particularly effective when they consider how disinformation affects different communities and tailor their approaches accordingly. In Nepal, Panos South Asia's monitoring of gendered online violence against women in politics exemplifies another effective approach: by systematically documenting gendered disinformation in electoral contexts, this initiative helps build understanding of disinformation patterns and provides evidence for advocacy.¹³⁰

Multi-stakeholder collaborative approaches

Regional and international collaborative approaches have shown promise in addressing cross-border disinformation. We have emphasised the importance of regional

¹²⁷ APC. (2021). Op. cit.

¹²⁸ Baltazar, F. (2024, 6 June). Op. cit.

¹²⁹ Dash, S. (2022, 26 August). Gendered disinformation on monkeypox: "If we keep pushing it as a gay disease, people will die". *Association for Progressive Communications*. <https://www.apc.org/en/news/gendered-disinformation-monkeypox-if-we-keep-pushing-it-gay-disease-people-will-die>

¹³⁰ Panos South Asia. (2022). *Analysis of Gendered Violence in Social Media against Women in Politics in Nepal*. Panos. https://southasiacheck.org/wp-content/uploads/2022/07/Analysis-of-Gendered_web.pdf

cooperation and knowledge-sharing, particularly in contexts where similar disinformation patterns affect multiple countries.¹³¹

The Freedom Online Coalition, a key space for engagement, offers one model for international collaboration. This coalition brings together governments committed to advancing internet freedom and addressing online threats, including disinformation.¹³²

The Global Partnership for Action on Gender-Based Online Harassment and Abuse is another important initiative in this regard.¹³³

The UN Strategy and Plan of Action on Hate Speech represents another important international framework.¹³⁴ This initiative includes commitments to engage with social media companies to get them to both support UN principles against hate speech and victims of hate speech, demonstrating how international organisations can create frameworks for coordinated responses.¹³⁵

Community-led responses and civil society engagement

Community-led responses have proven particularly effective in addressing context-specific disinformation. Feminist communities have developed tailored approaches to countering gendered disinformation, building on local knowledge and networks.¹³⁶

The #YourSlipIsShowing campaign provides another example of effective community response – Black feminist activists used this hashtag to identify and flag fake accounts

¹³¹ Abrougui, A. (2023, 28 July). Op. cit.

¹³² APC. (2021). Op. cit.

¹³³ See, for example: Foreign, Commonwealth & Development Office. (2024, 24 September). Global Partnership for Action on Gender-Based Online Harassment and Abuse calls for urgent action on countering gendered disinformation. GOV.UK. <https://www.gov.uk/government/news/global-partnership-for-action-on-gender-based-online-harassment-and-abuse-calls-for-urgent-action-on-countering-gendered-disinformation>

¹³⁴ <https://www.un.org/en/genocideprevention/documents/UN%20Strategy%20and%20Plan%20of%20Action%20on%20Hate%20Speech%2018%20June%20SYNOPSIS.pdf>

¹³⁵ Machirori, F. (2022, 5 October). Op. cit.

¹³⁶ Sívori, H., & Mochel, L. (2022, 22 November). Op. cit.

masquerading as Black women to spread disinformation, demonstrating how community knowledge and coordination can effectively counter targeted campaigns.¹³⁷

TrollBusters in the United States offers another instructive model. This network supports women journalists facing online harassment by monitoring continued attacks, sending positive counter-messaging and helping report content to platforms. Such community-based support systems can provide critical assistance to those targeted by coordinated disinformation campaigns.

Platform accountability and design standards

Promoting platform accountability and design standards represents another critical intervention. We argue for the need for platforms to be accountable "to users and society in general," while avoiding government control, particularly by authoritarian regimes.¹³⁸

The Santa Clara Principles on Transparency and Accountability in Content Moderation provide a framework for platform accountability that emphasises transparency about content removal, provision of notice to affected users and meaningful appeal rights.¹³⁹

Safety-by-design approaches represent another important standard. This approach emphasises that "technological development must ... embed safety-by-design standards" to mitigate "inherent gendered and other harmful biases and threats" before technologies are released.¹⁴⁰

Strengthening information system integrity

Strengthening information system integrity requires comprehensive approaches that address multiple vulnerabilities simultaneously. This includes promoting diversity and

¹³⁷ Malhotra, N. A. (2019, 23 September). Op. cit.

¹³⁸ Souter, D. (2023, 2 March). Op. cit.

¹³⁹ Malhotra, N. A. (2019, 23 September). Op. cit.

¹⁴⁰ APC. (2023, 1 September). Op. cit.

plurality in media ecosystems, ensuring access to reliable information, strengthening institutional responses to disinformation and building resilience among vulnerable communities.

We have also addressed the importance of transparency and access to information as fundamental to combating disinformation.¹⁴¹ When governments control access to information, they can more easily spread false narratives while suppressing counternarratives. Ensuring open access to information represents a crucial foundation for information system integrity.

Building robust cross-sector collaborations strengthens information integrity by bringing together diverse expertise and resources. Addressing disinformation “requires a coordinated global response, involving governments, international organisations, civil society and the private sector”.¹⁴²

Early warning systems represent another promising approach to strengthening information integrity. DEWARD (Disinformation Early Warning Data Tool) monitors social conversations to provide warnings about emerging misinformative content.¹⁴³ Such systems can potentially help identify disinformation campaigns before they cause significant harm.

There is no single solution to addressing disinformation's impact on human rights. Instead, effective approaches combine multiple interventions across different levels, ensuring that responses are comprehensive, context-appropriate and attentive to the specific needs of different communities, particularly women and gender-diverse people with intersecting marginalised identities.

By combining regulatory frameworks with human rights safeguards, digital literacy initiatives, support for independent media, collaborative approaches, community-led

¹⁴¹ Sirnate, V., & Jain, G. (2023, 3 April). Op. cit.

¹⁴² Baltazar, F. (2024, 6 June). Op. cit.

¹⁴³ mediafutures.eu/projects/disinformation-early-warning-data-tool-deward/

responses and platform accountability, stakeholders can work together to mitigate disinformation's harmful effects while strengthening the integrity of information systems. These approaches must consistently centre an intersectional gender perspective to address how disinformation disproportionately impacts women and gender-diverse people, particularly those with multiple marginalised identities.

9. What international organization, bodies, or agencies would be in your opinion best placed to tackle disinformation in line with international human rights law?

Several international organisations are particularly well-positioned to address disinformation while upholding human rights standards. Their effectiveness stems from their mandates, expertise and established frameworks for promoting rights-based approaches to information governance. As disinformation disproportionately impacts women and marginalised communities, these organisations must centre intersectional approaches in their work.

Regional human rights mechanisms

Regional human rights bodies are well suited to playing important roles in addressing disinformation while protecting human rights. Regional mechanisms are particularly valuable because they can develop context-specific approaches while maintaining alignment with international human rights standards. This regional specificity becomes crucial since region-specific patterns of gendered disinformation require tailored responses.¹⁴⁴

¹⁴⁴ Uiras, M. J. (2024, 18 May). Op. cit.

One interesting example is how regional human rights bodies can develop contextually appropriate approaches while maintaining alignment with international standards.¹⁴⁵ In another example, regional special rapporteurs on freedom of expression from the UN, OSCE, African Commission, and Inter-American Commission jointly addressed issues of freedom of expression and gender justice, demonstrating the potential for coordinated regional responses.¹⁴⁶

Freedom Online Coalition

The Freedom Online Coalition (FOC), while not a traditional international organisation, represents an important multi-stakeholder initiative specifically focused on internet freedom issues. We identify the FOC as a key space for engagement on disinformation.¹⁴⁷ The FOC's value lies in its ability to bring together governments committed to promoting human rights online, civil society organisations and other stakeholders. This multi-stakeholder approach aligns with our emphasis on inclusive governance processes that centre the experiences of marginalised communities.

The FOC's NotTheCost campaign against online violence targeting women in politics demonstrates how this coalition can address specific dimensions of gendered disinformation. By bringing together governmental and civil society perspectives, the FOC can develop nuanced approaches that balance addressing harmful content with protecting freedom of expression.

International Telecommunication Union

Technical standards and infrastructure governance significantly influence how information circulates online. The International Telecommunication Union (ITU) has an established role in governing the technical aspects of telecommunications and increasingly engages with broader internet governance issues. In a recent article titled

¹⁴⁵ Machirori, F. (2022, 5 October). Op. cit.

¹⁴⁶ Baltazar, F. (2024, 6 June). Op. cit.

¹⁴⁷ APC. (2021). Op. cit.

Integrating policy, research and technical standards in gender approaches to cybersecurity, Alan Finlay describes the ITU's Women in Standardization Expert Group as an example of efforts to incorporate gendered considerations into technical standards development.¹⁴⁸ Effective responses to disinformation must address both content-level issues and underlying technical systems, making the ITU's work on gender-sensitive technical standards increasingly relevant.¹⁴⁹

Multi-stakeholder initiatives and hybrid approaches

Beyond traditional international organisations, hybrid governance arrangements that bring together diverse stakeholders show particular promise. We emphasise the importance of governance approaches that are “independent from both governments and business interests”.¹⁵⁰ The Internet Governance Forum (IGF) is an example of a space that engages with disinformation and exemplifies this multi-stakeholder approach.¹⁵¹ While the IGF lacks decision-making authority, it provides a valuable platform for dialogue across sectors and regions, and demonstrates how multi-stakeholder processes can address specific dimensions of disinformation while ensuring diverse participation.

Within the UN system

United Nations Human Rights Council and Special Procedures

The UN Human Rights Council (HRC) and its Special Procedures are central actors in addressing disinformation from a rights-based perspective. We identify the HRC and its Special Procedures as key spaces for engagement on disinformation issues.¹⁵² Their established human rights framework provides a solid foundation for developing

¹⁴⁸ Finlay, A. (2024, 17 October). Op. cit.

¹⁴⁹ Martins, P., et al. (2024). Op. cit.

¹⁵⁰ Souter, D. (2023, 2 March). Op. cit.

¹⁵¹ APC. (2021). Op. cit.

¹⁵² Ibid.

approaches that balance combating disinformation with protecting freedom of expression and other fundamental rights.

The UN Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression has been particularly active in this area. In Martins et al.'s *Placing "gender" in disinformation* we described how we have collaborated with the UN Special Rapporteur to organise regional consultations in Africa, Asia, Latin America and the Caribbean, and the MENA region, as well as a consultation in Geneva alongside the Human Rights Council session, which offers an example of how the Special Procedures can facilitate global dialogue while ensuring representation from diverse regions and perspectives.¹⁵³ The Special Rapporteur has engaged with complex questions around disinformation, emphasising that responses must adhere to international human rights standards. What is more, this Special Rapporteur has explicitly argued that gendered disinformation targets women and gender non-conforming individuals "because of the gendered nature of the attacks and their gendered impact, and, very importantly, because it reinforces prejudices, bias and structural and systemic barriers that stand in the way of gender equality and gender justice."¹⁵⁴

Office of the United Nations High Commissioner for Human Rights

The Office of the United Nations High Commissioner for Human Rights (OHCHR) plays a crucial role in monitoring and promoting compliance with international human rights standards in digital contexts. Its established expertise in human rights monitoring and reporting makes it well-positioned to document how disinformation impacts human rights, particularly for marginalised communities. The OHCHR's global mandate allows it to address disinformation in diverse contexts while applying consistent human rights standards. This is particularly important given the disproportionate impact of disinformation on women and gender-diverse individuals from marginalised

¹⁵³ Martins, P., et al. (2024). Op. cit.

¹⁵⁴ Goldsman, F. (2021, 24 September). A mine-ridden internet and six rules for understanding anti-rights narratives. *Association for Progressive Communications*. <https://www.apc.org/en/blog/mine-ridden-internet-and-six-rules-understanding-anti-rights-narratives>; APC. (2021, 25 February). Op. cit.

communities.¹⁵⁵ In APC condemns the weaponisation of communication technology against civilians in Palestine and Lebanon, we have emphasised the importance of the OHCHR in holding both states and businesses accountable within the human rights framework.¹⁵⁶

However, there are concerns about inadequate support for the OHCHR, and about other UN initiatives potentially weakening the office's work by making it dependent on a voluntary funding mechanism available upon request.¹⁵⁷ This highlights the need for strengthening the OHCHR's resources and mandate to effectively address disinformation, particularly its gendered dimensions.

Addressing disinformation while upholding human rights requires a coordinated international response that combines the strengths of different organisations and governance approaches. The most effective international bodies will be those that ensure meaningful participation from diverse stakeholders, particularly those most affected by disinformation. This includes creating mechanisms for civil society participation, ensuring regional diversity and addressing barriers to participation such as language, resources and technical expertise.

It is important to note that no single organisation can effectively address disinformation alone. Instead, a coordinated approach that leverages different organisations' strengths while ensuring coherence is necessary. We have called for "better coordination among UN and other global digital processes and spaces that have been multiplying and at times replicating and even contradicting each other."¹⁵⁸ This coordination is essential for developing consistent approaches to disinformation that uphold human rights standards. We also have emphasised the importance of linking disinformation discussions to other

¹⁵⁵ Martins, P., et al. (2024). Op. cit.

¹⁵⁶ APC. (2024, 3 October). APC condemns the weaponisation of communication technology against civilians in Palestine and Lebanon. *Association for Progressive Communications*. <https://www.apc.org/en/pubs/apc-condemns-weaponisation-communication-technology-against-civilians-palestine-and-lebanon>

¹⁵⁷ Souter, D. (2023, 2 March). Op. cit.

¹⁵⁸ Finlay, A., & Betancourt, V. (2024, 28 October). How the Global Digital Compact is implemented will determine our digital future. *Association for Progressive Communications*. <https://www.apc.org/en/news/how-global-digital-compact-implemented-will-determine-our-digital-future>

relevant processes, including "WSIS+20, the IGF and Beijing+30," as well as "cybersecurity, where APC has been advocating for a holistic gender approach to standards settings and design."¹⁵⁹ In conclusion, we once again stress the importance of ensuring that responses to disinformation centre the experiences and needs of marginalised communities. International processes must address the specific ways in which disinformation affects women and gender-diverse individuals.¹⁶⁰

¹⁵⁹ Ibid.

¹⁶⁰ Baltazar, F. (2024, 9 April). Civil society and governments agree on the urgent need for a feminist Global Digital Compact (GDC). *Association for Progressive Communications*.
<https://www.apc.org/en/news/civil-society-and-governments-agree-urgent-need-feminist-global-digital-compact-gdc>