

IT for Change's Submission to the Public Consultation on the OECD Draft Recommendation on Information Integrity

IT for Change

November 2024



IT for Change's Submission to the Public Consultation on the OECD Draft Recommendation on Information Integrity

IT For Change*

IT for Change (ITfC) is an NGO based in Bengaluru, India, and is in Special Consultative Status with the Economic and Social Council of the United Nations. For over two decades, ITfC has worked to create a society where digital technologies contribute to human rights, social justice, and equity. Through research, model building, and policy advocacy, ITfC has worked extensively on issues of information integrity, social media platform accountability, and the intersection of gender and the digital public sphere.

ITfC welcomes the OECD's Draft Recommendations on Information Integrity as a timely and holistic intervention to safeguard the information space and hold technology platforms and governments to account. We believe the focus of the Recommendations on enhancing transparency, accountability, and plurality of information sources is in the right direction, and recognize its potential to serve as an important international standard to guide multi-stakeholder actions to strengthen information integrity, rooted in democratic principles of international human rights law. That being said, the Recommendations fall short of addressing the core threats in the algorithmified public sphere to communicative rationality, moving beyond symptomatic remedies to preserving information integrity. The algorithmic virality of social media business models has skewed the values of pluralism and diversity in media into a post-truth regime that threatens the democratic principle of verity¹ — the pursuit of truth as public reason.

The liberal regulatory playbook of supplier-focused and consumer-focused remedies to restoring the health and vibrancy of the platformized public sphere does not address the root problem — the techno-design architectures that engender a hostile environment inimical to equality, liberty, and solidarity. Challenging this paradigm necessitates structural changes to the social media platform architectures and not just

¹ Truth as a constitutional value, LSE Event, 13 Nov 2024, <https://www.youtube.com/watch?v=EvirDw5fDv8&t=1s>.

*Submission by Merrin Muhammed Ashraf with input and review from Anita Gurumurthy and Nandini Chami. To know more about IT for Change visit <https://itforchange.net>

procedural rules to mitigate specific harms. It requires us to fundamentally reimagine the terms on which public discourse is mediated and organized on the internet currently such that values of truth and democratic integrity take center stage, enabling public discourse and democratic deliberations in a critical and agentic manner.

In light of this, ITfC urges the OECD to take into account the following suggestions to ensure that this Recommendation sets the highest standards for information integrity in accordance with the democratic values of equality, liberty, solidarity, and also, verity/public reason.

- 1. Definition of Information Integrity (II.para 6. p. 3):** The definition of information integrity needs to be broadened to account for all stages of the communication cycle. This includes not only individuals' ability to access information, but also their ability to generate ideas and opinions, express and speak, and most importantly, have their voices heard and understood.² The definition should also be concerned with the imperative to ensure the safety of citizens, particularly, marginalized communities, journalists, and human rights defenders from violence within the information ecosystem.³

Rationale: Defining information integrity solely from the perspective of access to information assumes that merely providing citizens with accurate, reliable, and plural information, would automatically equip them to fully exercise their rights and citizenship.⁴ This ignores the wider economic, social, cultural, and technological factors that shape the current communication and media ecosystem and impact the ability of citizens to engage meaningfully with information, have their views heard, and participate in respectful and reasoned public deliberations and debates. Further, the violence in the information ecosystem in the form of disinformation, hate speech, trolling, etc amplified by platform algorithms adversely affects journalists, human rights defenders, and marginalized communities by silencing them and denying their right to be heard.⁵

² Siochrú, Seán O and Gulumurthy, Anita. (Feb 2024). Digital platforms versus democratic political discourse: Challenges and the way forward. WACC Global.

<https://waccglobal.org/digital-platforms-versus-democratic-political-discourse-challenges-and-the-way-forward/>

³ Working towards healthier information ecosystems – Collective visions from civil society in Latin America and the Caribbean. (2024). The Engine Room.

<https://www.theengineroom.org/library/new-report-working-towards-healthier-information-ecosystems-collective-visions-from-civil-society-in-latin-america-and-the-caribbean/>

⁴ Nina Santos. (Mar 2024). Tech Policy Press. <https://www.techpolicy.press/why-do-we-need-to-discuss-so-called-information-integrity/>

⁵Posetti, Julia and Nabeelah Shabbir. 2022. "The Chilling: A Global Study of Online Violence

- 2. Strengthen recommendation on data access (III. 2.b. p.5):** The recommendation should clearly state that the criteria and process of vetting researchers to give access to data of platforms should be transparent and an independent state body do the vetting. There should be a provision for response and appeal in case a data access request is opposed by the platform or denied by the regulator. Further, access to public data should not be just limited to independent vetted researchers, but also available to public interest actors such as civil society organizations, research institutions, and journalists.

Platforms should also proactively disclose to the public, in accessible formats and country-appropriate languages, information on, but not limited to: (a) content moderation tools, including procedures for quality assurance or evaluation, and measures taken to mitigate any harm from incorrect decisions; (b) the logic driving the platform recommender algorithms and the nature of personal data collected and used by the algorithms; (c) types of complaints received concerning content hosted and the action taken by platforms and timeline for resolution; and (d) number of human moderators, their expertise, and employment status.

Rationale: It is important to have clear and transparent criteria to ensure that vetting of researchers does not become prohibitive for researchers and civil society actors to conduct research. Further, the vetting process should be done by an independent state regulator, with provisions for response and appeal if the request is met with opposition from the platform or denied by the regulator.

It is also important to ensure that access to public data is just limited to independent vetted researchers and is also available to public interest actors such as civil society organizations, research institutions, and journalists. Such access is possible under the European Digital Services Act (DSA) and the OECD Recommendation should not provide for less data access than the DSA.

Apart from data access to researchers and public interest actors, the Recommendations should also require platforms to proactively disclose to the public vital information about their functioning and practices. This recommendation is grounded in the principle of public interest in information and

Against Women Journalists.” International Centre for Journalists.

https://www.icfj.org/sites/default/files/2022-11/ICFJ_UNESCO_The%20Chilling_2022_1.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/>.

data sharing. A rule-of-law approach is critical to ensure transparency, accountability, and trust in the digital ecosystem. Proactive disclosure empowers individuals and institutions to understand how platforms function and is essential for fostering informed public discourse, combating disinformation, misinformation, and hate speech, and reducing opacity around platform governance. It also helps to hold platforms accountable by enabling public scrutiny of their practices.

- 3. Greater obligations for advertisers and ad tech platforms (III. 2c. p.5):** Recommendations should seek to establish human rights-responsible advertising practices by requiring advertisers to ensure that advertising does not risk the human rights of individuals and manipulate the information system; to not engage in targeted advertisements based on sensitive data and perceived user traits; to carry out thorough audits of advertising campaigns; compel transparency from ad tech platforms; and vet ad exchange supply partners for their adherence to human rights standards and commitment to uphold information integrity.⁶

Rationale: Advertising in digital spaces, wields significant power to influence opinions, behaviors, and access to information. When unchecked, it can exacerbate societal harm, violate individual rights, and undermine the integrity of the information environment.⁷ Taking this into account, the UN Global Principles on Information Integrity recommended the measures mentioned in the previous paragraph.⁸

Advertisers should ensure their campaigns do not infringe upon human rights or manipulate public opinion. Human rights frameworks like the UN Guiding Principles on Business and Human Rights emphasize corporate responsibility to prevent activities that negatively impact rights such as privacy, equality, and freedom of expression. Targeting users based on sensitive traits amplifies the risk of discrimination and exploitation. Transparency in how advertisements are targeted, displayed, and monetized is critical for accountability.⁹ Ad tech platforms often operate as opaque systems, making it difficult to understand how decisions are made. Greater transparency would

⁶ UN Global Principles for Information Integrity. (2024). <https://www.un.org/en/information-integrity/global-principles>

⁷ Maréchal, Nathalie et al. (May 2020) Ranking Digital Rights.

https://rankingdigitalrights.org/wp-content/uploads/2023/07/Getting_to_the_Source_of_Infodemics_Its_the_Business_Model_2020-05.pdf

⁸ UN Global Principles for Information Integrity. (2024). <https://www.un.org/en/information-integrity/global-principles>

⁹ UN Global Principles for Information Integrity. (2024). <https://www.un.org/en/information-integrity/global-principles>

empower advertisers, regulators, and the public to scrutinize these processes for adherence to human rights principles.¹⁰

4. Expressly recognize online hate speech and trolling as threats to information integrity (III. 3.d.

p.9): The Recommendations should expressly recognize hate speech and online trolling as acts of violence within the information ecosystem and risks to information integrity and call for actions by the governments and platforms. It is also important to have these terms defined in the definitional clause by referring to international and national legal and policy developments.

Rationale: Hate speech and online trolling, often conducted in a coordinated and systematic manner, are a potent form of patriarchal social censorship – a systematic intimidation, that silences women, especially from marginal social locations, and dissuades them from entering public-political life.¹¹ This undermines efforts to achieve the sustainable development goal of gender equality and constitutes a veritable threat to democracy, overall.¹²

While many jurisdictions are now enacting laws to penalize online hate speech and trolling,¹³ social media platforms have done very little to address its amplification. Opaque platform operations vis-à-vis grievance redressal, actions taken on problematic content, and the logic of algorithms, prevent regulatory efforts to address gendered disinformation and hate speech online.¹⁴

- 5. Call for fair revenue sharing and collective bargaining for news publishers (IV. 1f. p.7):** The Recommendation on encouraging remuneration models for journalistic content should be strengthened by expressly requiring governments to devise fair and equitable revenue sharing arrangements between news publishers and platforms and institute enforcement mechanisms for the same. Further, where direct negotiations between the news industry and platform companies

¹⁰ Maréchal, Nathalie et al. (May 2020) Ranking Digital Rights.

https://rankingdigitalrights.org/wp-content/uploads/2023/07/Getting_to_the_Source_of_Infodemics_Its_the_Business_Model_2020-05.pdf

¹¹ Ashraf, Merrin Muhammed et al. (June 2024). Gendering the G20 Information Integrity Agenda. T20

Brasil. <https://itforchange.net/sites/default/files/add/Gendering%20the%20G20%20Information%20Integrity%20Agenda.pdf>

¹² Sobieraj, Sarah. 2020. Credible Threat: Attacks against Women Online and the Future of

Democracy. New York: Oxford University Press. <https://doi.org/10.1093/oso/9780190089283.001.0001>

¹³ Law N° 14.19 (Brazil) https://www.europarl.europa.eu/doceo/document/A-9-2023-0234-AM-298-298_EN.pdf

¹⁴ Ashraf, Merrin Muhammed et al. (June 2024). Gendering the G20 Information Integrity Agenda. T20

Brasil. <https://itforchange.net/sites/default/files/add/Gendering%20the%20G20%20Information%20Integrity%20Agenda.pdf>

are provided for, collective bargaining arrangements ought to be conducted on genuinely collective, inclusive, and transparent principles.¹⁵

Rationale: The growing dependence of news publishers on online platforms like social media for traffic and revenue, has created an imbalance in bargaining power. This dynamic often results in news organizations struggling to negotiate fair compensation for the content they produce, while digital platforms continue to profit from advertising.¹⁶ To ensure the sustainability of news media in this platform era, especially small and local media, it is essential to allow media organizations to benefit from the earnings that platforms receive from their use of news content in a manner that does not unduly advantage larger media organizations at the expense of smaller ones.

Further, recognition of collective bargaining for news publishers to engage with platforms should be conducted on transparent and inclusive principles to ensure that details and terms of any bilateral deals are available to all and the interests of small media houses are also protected.¹⁷

6. Mandate interoperability, data portability, and unbundling of content hosting and content curation to enable the sustainability of alternative platforms and recommender algorithms. Also consider public funding to support the development of public alternatives to for-profit recommender systems and diverse, alternative media and communication platforms (IV. 1j.

p.7): The Recommendations should require States to mandate interoperability and data portability to address the market concentration of a few big platforms and facilitate the emergence of newer platforms and platform services like recommender algorithms.

Rationale: To promote the plurality of information sources, it is essential to ensure the emergence of new and alternative models of platforms and recommender algorithms with diverse value propositions. Existing structures and network effects have made large platforms larger and more successful while making it increasingly difficult for new/smaller platforms to participate. To address this market concentration, a number of measures have been proposed.

¹⁵ Working Group on the Sustainability of Journalism. (June 2021). A New Deal for Journalism. Forum on Information and Democracy. https://informationdemocracy.org/wp-content/uploads/2021/06/ForumID_New-Deal-for-Journalism_16Jun21.pdf

¹⁶ Singh, Madhavi. (Dec 2023). Big Tech and News Media in India: The Frenemies Who Control What We Read. BotPopuli. <https://botpopuli.net/big-tech-and-news-media-in-india-the-frenemies-who-control-what-we-read/>

¹⁷ Working Group on the Sustainability of Journalism. (June 2021). A New Deal for Journalism. Forum on Information and Democracy. https://informationdemocracy.org/wp-content/uploads/2021/06/ForumID_New-Deal-for-Journalism_16Jun21.pdf

Firstly, unbundling of content hosting from content curation on large platforms will enable a marketplace of alternative recommender systems to emerge which will present platform users with viable alternatives to recommender systems driven solely by maximizing engagement and profits, without needing to migrate to other platforms, where they may have no existing connections.¹⁸

Secondly, interoperability and data portability requirements could counteract network effects and create more fair competition by lowering entry barriers for new players.¹⁹

Thirdly, States should be encouraged to consider providing public funding to support meaningful, public alternatives to for-profit recommender systems so that timely, accurate, local knowledge is always available.²⁰ States should also support the emergence of diverse, alternative media and communication platforms that do not follow the surveillance capitalist business model.

7. Require technology platforms to shift away from business models and practices that

incentivize threats to information integrity (IV. 2.a. p.8): The Recommendations should expressly require online platforms to move away from business models that incentivize disinformation, misinformation, and other forms of information manipulation.²¹ This should not be left to the voluntary initiative of platforms, but should be legislatively enforced by the state through legal and policy measures. These measures could include imposing a statutory duty of care on platform owners for addressing the individual and societal harms stemming from their business model and techno-design choices, comprehensive ex-post and ex-ante accountability measures including periodic human rights risk assessment, and developing a liability framework to hold platforms and those directly in charge of for the conduct of business accountable for enabling or facilitating harms including disinformation, hate speech, incitement to violence, etc.

Rationale: There is ample evidence to show that the business model of online platforms and their techno-design choices based on the logic of attention economy and surveillance capitalism incentivizes the creation and dissemination of illegal and harmful content and prioritizes

¹⁸ Article 19. (2021) Taming Big

Tech.<https://www.article19.org/wp-content/uploads/2023/02/Taming-big-tech-UPDATE-Jan2023-P05.pdf>

¹⁹ Forum on Information and Democracy. (2023). Pluralism of News and Information in Curation and Indexing Algorithms. <https://informationdemocracy.org/wp-content/uploads/2023/08/Report-on-Pluralism-Forum-on-ID.pdf>

²⁰ Forum on Information and Democracy. (2023). Pluralism of News and Information in Curation and Indexing Algorithms. <https://informationdemocracy.org/wp-content/uploads/2023/08/Report-on-Pluralism-Forum-on-ID.pdf>

²¹ UN Global Principles for Information Integrity. (2024). <https://www.un.org/en/information-integrity/global-principles>

engagement and profit over truth and reasoned debate.²² Hence, to safeguard information integrity, it is vital to make a shift away from the current extractive business model of online platforms. The recent UN Guiding Principles on Information Integrity also underlines this imperative.²³ State platform regulations should focus on reorienting platform incentives so that they promote diverse, mindful, and quality interactions over sensationalist, abusive, and impulsive interactions, by requiring changes in techno–design aspects and values driving their content curation and content moderation algorithms, and developing strong liability frameworks for platforms.

8. Require Government actions to strengthen information integrity to adhere to international human rights standards (V. 1. P. 10):

The Recommendations on upgrading institutional architecture and open government practices should expressly require adherence to international human rights standards of legitimacy, necessity, and proportionality in government measures aimed at regulating speech on online platforms and media.²⁴

Rationale: There are many instances where states have resorted to excessive and arbitrary measures to curb disinformation and misinformation, such as internet shutdown, opaque content takedown orders made to platforms, and criminalizing various kinds of speech.²⁵ This coupled with state-sponsored disinformation and collusion with platform companies to control public narratives raises serious concerns about the government’s control over the information ecosystem.²⁶ Therefore, it is vital to expressly require government measures to strengthen information integrity to meet the triple test of legality, necessity, and proportionality as laid down in the International Covenant on Civil and Political Rights under its guarantee of freedom of expression and opinion.²⁷

9. Call for accountability of AI developers, deployers, platforms carrying AI-generated content, and users: Recommendations should urge governments to institute regulatory frameworks to hold AI developers, deployers, platforms carrying AI-generated content, and users accountable for

²² Vosoughi, Soroush et al. (2018). The Spread of True and False News Online. *Science*.

<https://www.science.org/doi/10.1126/science.aap9559>; Zuboff, Shoshana. (2018). *The Age of Surveillance Capitalism*.

²³ UN Global Principles for Information Integrity. (2024). <https://www.un.org/en/information-integrity/global-principles>

²⁴ OHCHR. 2021. Disinformation and freedom of opinion and expression. Report of the Special Rapporteur on the promotion and protection of the right to

freedom of opinion and expression. <https://documents.un.org/doc/undoc/gen/g21/085/64/pdf/g2108564.pdf>

²⁵ Id; Dara, Rishabh. 2017. “Intermediary Liability in India: Chilling Effects on Free Expression on the Internet.” CIS.

https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2038214

²⁶ OHCHR. 2021. Disinformation and freedom of opinion and expression. Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression. <https://documents.un.org/doc/undoc/gen/g21/085/64/pdf/g2108564.pdf>

²⁷ Article 19, ICCPR.

threats to information integrity and democracy stemming from the deployment and use of AI. This should include ex-ante and ex-post assessment of the impact of AI systems on fundamental rights, particularly, freedom of expression, access to information, right to human dignity, and privacy; mandating transparency and explainability of AI systems;²⁸ and instituting responsibility and accountability regime for different AI actors corresponding to the systemic risk of the AI system to the information space and the level of involvement/control of the actor in perpetuating that risk.²⁹

Rationale: AI systems and recent advancements in the form of generative AI (Gen AI) and Large Language Models (LLMs) have had a disruptive effect on the information and communication space. AI models can be potent in terms of the risks they pose for information integrity not only because they increase the possibilities for generating disinformation, misinformation, and hate speech, but also since they facilitate the rapid and targeted dissemination of such content, that too, at scale, by malicious actors.³⁰ Hence, the Recommendations should not overlook the importance of regulating AI systems and the associated actors to truly strengthen information integrity.

- 10. The government's duty to disseminate information in non-electoral contexts (V. 5.a. p.11):** The duty to provide timely and reliable information to citizens should extend beyond electoral contexts. The government should also proactively disseminate accurate information and debunk false information about public health, climate change, and during situations of natural disasters, civil wars, etc.

Rationale: There is extensive documentation showing that the spread of disinformation and misinformation has posed significant challenges to the right to health and responses to the COVID-pandemic.³¹ As in the case of global health, with respect to climate change too, scientific information has been discredited.³² Ideological and identity-based disinformation has fomented

²⁸ Forum on Information and Democracy. (Feb 2024). AI as a Public Good: Ensuring Democratic Control of AI in the Information Space; <https://informationdemocracy.org/wp-content/uploads/2024/03/ID-AI-as-a-Public-Good-Feb-2024.pdf>

²⁹ Council of Europe (2023). Draft Framework Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law. <https://rm.coe.int/cai-2023-28-draft-framework-convention/1680ade043>; Forum on Information and Democracy. (Feb 2024). AI as a Public Good: Ensuring Democratic Control of AI in the Information Space. <https://informationdemocracy.org/wp-content/uploads/2024/03/ID-AI-as-a-Public-Good-Feb-2024.pdf>.

³⁰ OHCHR (2023). Taxonomy of Human Rights Risks Connected to Generative AI. www.ohchr.org/sites/default/files/documents/issues/business/b-tech/taxonomy-GenAI-Human-Rights-Harms.pdf.

³¹ OHCHR. (2020). Disease pandemics and the freedom of opinion and expression. Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression; <https://www.ohchr.org/en/documents/thematic-reports/ahrc4449-disease-pandemics-and-freedom-opinion-and-expression-report>.

³² Bradshaw, S., & Howard, P. (2019). The global disinformation order. 2019 Global inventory of organized social media manipulation. Oxford Internet Institute. <https://demotech.oii.ox.ac.uk/wp-content/uploads/sites/12/2019/09/CyberTroop-Report19.pdf>

discrimination and hatred against minorities, migrants, and other marginalized communities, generating ethnic or religious tensions that have culminated, at times, in violence offline.³³ These challenges are compounded by some governments withholding or falsifying information.³⁴

³³ OHCHR. 2021. Disinformation and freedom of opinion and expression. Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression. <https://documents.un.org/doc/undoc/gen/g21/085/64/pdf/g2108564.pdf>

³⁴ Id.