

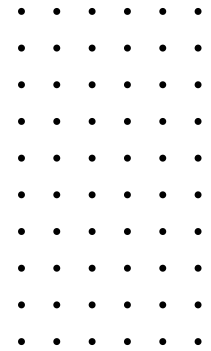


NATIONAL COALITION
ON FREEDOM OF EXPRESSION &
CONTENT MODERATION IN KENYA



Contributions to *Our Common Agenda*

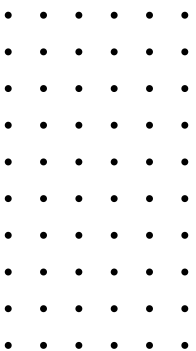
Policy Brief 8: Information Integrity on Digital Platforms



Recommendations on the proposed UN Code of Conduct on Information Integrity on Digital Platforms



in preparation for the
United Nations Summit of the Future 2024



Funded by
the European Union



TABLE OF CONTENTS



Background	03
Executive Summary	06
Recommendations	09
Acknowledgements	19

Prepared by:
John Okande and Sharmaine Koh
UNESCO Regional Office for Eastern Africa

Published March 2024



BACKGROUND

On 5th December 2023, the **National Coalition on Freedom of Expression and Content Moderation in Kenya (FECoMo)** participated in an Expert Consultation on the principles of a voluntary UN Code of Conduct on Information Integrity on Digital Platforms.[1]

The UN Code of Conduct will help to guide Member States, digital platforms and other stakeholders in their efforts to ensure a safer and more inclusive digital space for all, while vigorously defending the right to freedom of expression and access to information.

Organised by the United Nations Information Service (UNIS) and UNESCO Regional Office for Eastern Africa, the Expert Consultation provided a platform for members of FeCoMo to contribute Kenyan and African perspectives to the drafting of the global Code of Conduct for Information Integrity. This was part of several global consultations aimed at seeking feedback on the proposed principles, in preparation for the **United Nations Summit of the Future 2024** that will take place from 22-23 September 2024 in New York. The Summit of the Future is a high-level intergovernmental meeting that will agree on multilateral solutions for a better tomorrow and strengthen global governance for present and future generations.

[1] These proposed principles can be accessed in **Policy Brief 8: Information Integrity on Digital Platforms** in Our Common Agenda, a series of policy briefs published by the Office of the UN Secretary-General.

BACKGROUND

The Kenyan Expert Consultation convened 26 experts from a variety of professional fields, including regulators, peacebuilders, civil society, media professionals, technology companies, academia, development practitioners, and law among others. Majority of the experts who engaged in the consultation were members of the FECoMo, established under UNESCO's EU-funded **Social Media 4 Peace** project.[2] The Coalition is composed of multidisciplinary stakeholders committed to addressing potential harmful content online, particularly mis/disinformation and hate speech in Kenya.

This paper distils the key recommendations presented by the Coalition and other experts during the consultations in Kenya. It outlines FECoMo's position on protecting information integrity, digital platform governance, and user empowerment—particularly in a Kenyan and African context. These interventions are consistent with FECoMo's mission to provide thought leadership, expertise, and strengthen cross-sectoral partnerships in advancing information integrity on digital platforms in Kenya.

It is hoped that the recommendations provide opportunities for reflection, consideration, and fruitful dialogue with policy makers and citizens, to collaboratively realise a safe, inclusive and empowering online environment in Kenya.

Social Media 4 Peace



The UNESCO Social Media 4 Peace project, funded by the EU, is a multi-year project implemented in Kenya, Indonesia, Bosnia and Herzegovina and Colombia. It aims to build the resilience of societies to potentially harmful content online.

Learn More

- [Social Media 4 Peace Project Page](#)
- [Publication: Local Lessons for Global Practices](#)
- [Social Media 4 Peace Handbook](#)

Our policy brief on information integrity on digital platforms puts forward a framework for a [concerted] international response.

Its proposals are aimed at creating guardrails to help governments come together around guidelines that promote facts, while exposing conspiracies and lies, and safeguarding freedom of expression and information;

And to help tech companies navigate difficult ethical and legal issues and build business models based on a healthy information ecosystem.

António Guterres

Secretary-General, United Nations

12 June 2023, at the Secretary-General's [Press Briefing on Policy Brief on Information Integrity on Digital Platforms](#)

05



EXECUTIVE SUMMARY

FECoMo’s commitment to safeguarding information integrity is deeply rooted in its fundamental defence of democracy and human rights, particularly that of freedom of expression and access to information. Without appropriate safeguards, our information space has and will continue to facilitate and enable the proliferation of mis/disinformation and hate speech on digital platforms—phenomena that are rapidly undermining key democratic functions and human rights today. A strong, coordinated, open and participatory whole-of-society response is needed to effectively address prevailing threats to information integrity.

Broad Principles

The Coalition’s recommendations on the Code of Conduct are aligned with and respond to UNESCO’s recently published **Guidelines for the Governance of Digital Platforms** (2023), which espouses five key elements for digital platforms to adhere to:

1. Conduct of human rights due diligence
2. Adherence to international human rights standards
3. Transparency
4. Accessibility of information
5. Accountability to relevant stakeholders.

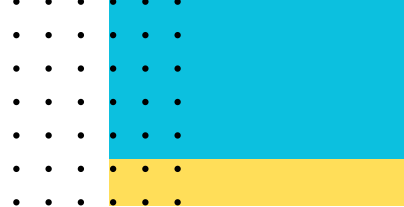
Sensitivity to local, regional and temporal contexts

While the UN Code of Conduct is designed to be applicable for a broad, global context, the guidelines should explicitly acknowledge the digital divide, demonstrate sensitivity to diverse contexts, and acknowledge that local and regional governing bodies—including many jurisdictions in Africa—still lack capacity to formulate or enforce effective policies to safeguard information integrity on digital platforms. Thus, while the guidelines in the UN Code of Conduct are aspirational for all, they may be feasible or applicable for some African countries to varying degrees. Care should be taken by governments to evaluate the appropriateness or relevance of certain recommendations made at a global-scale before scaling-up implementation within their national contexts.

At the same time, while models elsewhere should not be taken wholesale and implemented in different contexts like Kenya, similar frameworks implemented elsewhere, such as Europe's Digital Service Act, could provide valuable lessons for the Code of Conduct and other local adaptations. Some examples include Finland's approach to MIL education, Taiwan's approach to digital literacy and civic engagement, South Korea's fact-checking organisation, and Germany's Network Enforcement Act.

Provisions for periodic reviews and updating of the Code of Conduct should also be established for long-term sustainability, so that it remains relevant amidst rapid changes in the technological landscape.



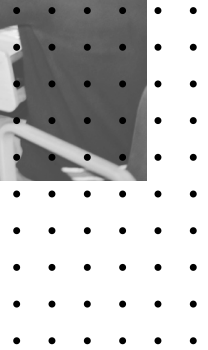


Support and Compliance

In designing the Code of Conduct, mechanisms to encourage support and compliance need to be strategically and intentionally included. Since the Code is voluntary in nature, it should ensure strong incentives to encourage buy-in from governments, digital platforms, and other stakeholders.

Governments have a natural stake in supporting the Code of Conduct, given their mandate and responsibility to demonstrate concern and commitment to defending the digital rights and safety of citizens. Governments also wield the power to shape policy environment or economic incentives that will nudge digital platforms to reform their business models. However, decision makers will have to reflect on how to incorporate local enforcement and implementation mechanisms into their own adaptations of the UN Code of Conduct, as the Code cannot feasibly or desirably influence enforcement across different jurisdictions worldwide.

Platforms, meanwhile, are more likely to offer support and compliance when presented with high-stakes issues that carry reputational risks and/or are unequivocally harmful to societal well-being (e.g. terrorism, child pornography). The Code of Conduct may benefit from greater specificity in defining issues around which it hopes to rally cooperation and compliance from stakeholders. Stronger links to the [Internet For Trust](#) Guidelines and positive framing of recommendations and obligations will also strengthen the Code's universality.



RECOMMENDATIONS

by the National Coalition on Freedom and Expression and Content Moderation in Kenya on the voluntary UN Code of Conduct on Information Integrity on Digital Platforms

The recommendations are categorised according to the 9 proposed principles in the UN Code of Conduct on Information Integrity on Digital Platforms, which are: (1) Commitment to Information Integrity; (2) Respect for Human Rights; (3) Support for independent media; (4) Increased Transparency; (5) User Empowerment; (6) Strengthened Research and Data Access; (7) Scaled-up Responses; (8) Stronger Disincentives; and (9) Enhanced Trust and Safety.

01

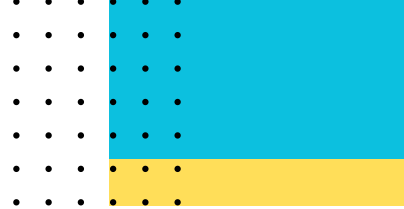
Commitment to Information Integrity

- **Review legal protections that usually shield platforms from being held liable for online harmful content**, to influence stronger commitments to information integrity.
- **Foster a bottom-up approach to securing intergovernmental commitments**, given the region's challenges in speaking with one voice on issues relating to information integrity. A common position should first be secured at the sub-regional level (e.g. East African Community) and then at regional level (e.g the African Union) on a unified approach to information integrity.
- **Emphasise the importance of global cooperation and the need for system interoperability in the implementation of the voluntary code**, given the cross-boundary challenges of cybersecurity, terrorism, and online safety in the digital environment. For example, legal decisions based on local laws are limited when they cannot be enforced elsewhere. Hence, baseline harmonisation of judicial positions across jurisdictions will be beneficial.

02

Respect for Human Rights

- **Governments should bring together legal regulations against harmful content online and constitutional provisions for the rights to freedom of expression and access to information under a single, coherent legislation.** At present, fragmented legislation related to mis/disinformation and hate speech online creates affordances for abuse. *For example, while the rights to freedom of expression and access to information are provided for in the Kenyan Constitution (i.e. Articles 33 and 34), the same rights are frequently limited through the use of other laws (i.e. Computer Misuse and Cybercrimes Act).*
- **Human-centric principles should be mainstreamed in design governance and the development and implementation of technology,** especially concerning generative AI and issues of privacy, security, and data protection.
- **All attempts to mainstream and implement the Code of Conduct should explicitly communicate that it will firmly respect and uphold, not interfere or curtail, human rights,** especially the right to freedom of expression. Balance should be encouraged especially with the adoption of policy tools such as content moderation, where “overzealous” implementation can inadvertently amplify certain narratives while silencing others. Platforms users should be given equal opportunity to be heard.



03

Support for Independent Media

- **Strengthen support for and defence of the viability and sustainability of media pluralism, diverse and independent journalism.** The Code of Conduct can be more intentional in advocating for plurality of voices, fair compensation of media, considering the value that media systems/entities create for digital platforms.
- **Advocate for multistakeholder cooperation between governments, digital platforms, and the media to share and publish truthful and factual information,** especially against the backdrop of adversarial state-media relations in certain jurisdictions.
- **Encourage more investment in continuous professional development for media professionals above and beyond ad hoc training,** so that the media is equipped to adapt to rapid changes that continue to take place in our information environments.
- **Enhance capacities of mainstream and community media houses to fact-check information** so that they are empowered to play a bigger role in fighting information disorder in the ecosystem.

04

Increased Transparency

- **Consider provisions for the protection of rights of whistleblowers** who report what they reasonably believe to be evidence of waste, fraud, abuse or mismanagement within the operations and processes of digital platforms.
- **Highlight the importance of meaningful transparency reporting by digital platforms.** Transparency reports should not be published for the sake of reporting but must be meaningful to be useful.
- **Make conversations about algorithmic transparency understandable, available, and accessible to the public** for more effective oversight and transparent governance of algorithmic systems.
- **Communicate Community and User Agreements on digital platforms in clear and accessible language** so that users are sensitized and educated to protect themselves while also empowered with knowledge about their rights, risks and community standards. Encourage transparency on the use of (paid) “verified” user accounts, especially during politically sensitive periods like elections.
- **Encourage digital platforms to publish risk assessments,** which will enhance user awareness and encourage platforms to demonstrate and invest in mitigative measures to protect users online.

05

User Empowerment

- **Enhance grassroots engagement, participation, and contribution to discussions about user empowerment.** A large population around the world still lacks access to basic instruments to access information.
- **Emphasise that MIL education should be tailored to the needs of different audiences and their specific contexts,** cognisant of existing inequalities that drive the digital divide.
- **Strengthen recommendations for governments to mainstream Media and Information Literacy (MIL) and Digital Competencies programmes in educational curriculum,** while recognising that national MIL programmes enjoy different levels of maturity.
- **Emphasise the need for effective and efficient mechanisms for reporting and harmful content and seeking redress on digital platforms.** Users should be clearly and systematically educated on how to use existing mechanisms for reporting and redress.
- **Consider reparations for victims of online harms** or other mechanisms to provide avenues of restorative justice for individuals affected by online harmful content.
- **Leverage technological innovations to provide tools and solutions for user empowerment,** including nurturing and investing in local technology companies to design context-specific local solutions for local challenges.



06

Strengthened Research and Data Access

- **Emphasise the need to increase collection and foster open access to disaggregated data by digital platforms and other actors**, concerning the use and circulation of potential harmful content on digital platforms and especially in Africa where regional data remains scarce.
- **Encourage all stakeholders to invest more resources in credible research** to inform evidence-based decision and policy making on information integrity on digital platforms, especially in Africa.
- **Encourage greater knowledge sharing and dissemination of research findings** so that users and other stakeholders are kept abreast with cutting-edge insights from research activities.

07

Scaled-up Responses

- **Explore co-regulatory models in addressing mis/disinformation and hate speech on digital platforms**, by empowering independent regulators to support government efforts in holding digital platforms accountable.
- Beyond intersectoral/interdisciplinary collaboration, language on multi-stakeholderism should also **strongly communicate meaningful inclusion of diverse groups** including women and girls, youth and children, persons with disabilities and other traditionally marginalised perspectives.
- **Include specific recommendations regarding the rights of Civil Society Organisations (CSOs) to operate freely and independently**, in recognition of the critical role of CSOs in tackling mis/disinformation and hate speech online and holding actors accountable.
- **Include specific language to engage the developer community in responsible, safe and ethical design of digital platforms**, so that developers place human rights and user empowerment at the forefront of system design considerations.
- **Encourage systems-wide interoperability and collaboration between big-tech digital platforms** (e.g. leverage mechanisms of collaborative flagging and content take-down, so that similar harmful content across several digital platforms can be taken down across the board).



08

Stronger Disincentives

- **Develop stronger approaches to hold big-tech companies accountable**, given the monopolistic influence they hold over digital users worldwide and the expected resistance from companies and the countries in which they are based.
- **Incentivise production and dissemination of quality information and peaceful messaging.** Digital platforms should consider amplifying positive voices, in addition to disincentivizing or blacklisting harmful content and repeat offenders.



09

Enhanced Trust and Safety

- **Mainstream the language of “Design Governance”** to advocate for platform and algorithmic design that prioritises individual safety, respects user preferences, blacklists politically divisive content and contexts. This also works to reduce unsustainable and ineffective dependence on content moderators to regulate harmful content online.
- **Recognise linguistic diversities and increase representation of minority languages and cultures in the content moderation policies and practices of digital platforms.** Currently, there is a lack of monitoring on hate speech and disinformation in local languages due to shortage of content moderators and lack of language data.
- **Set clear and prescriptive standards for effective content moderation systems** that digital platforms should adopt and implement.



ACKNOWLEDGEMENTS

FECOMo Kenya

Rachel Olpengs, Africa Centre for People Institutions and Society (ACEPIS)
Angela Minayo, Article 19 Eastern Africa
Patience Nyange, Association of Media Women in Kenya
Kelvin Nyangweso, Association of Kenya Community Media Operators
Angela Koki, Communications Authority of Kenya
Rosemary Mwangi, Communications Authority of Kenya
Levi Obonyo, Daystar University
Ruth Owino, Kabarak University
Victor Kapiyo, KICTANet
Jacob Nyongesa, Media Council of Kenya
Terence Minshi, Media Council of Kenya
Lyndcey Oriko, National Cohesion & Integration Commission
Susan Wekesa, National Council for Persons with Disabilities
Immaculate Onyango, National Steering Committee of Peacebuilding and Conflict Management

Contributing Experts

Brian Kae, Spin Digital
Catherine Muya, GIZ Kenya
Cheryl Akinyi, Open Society Foundations
Churchill Otieno, Eastern Africa Editors Society
Cynthia Chepkemai, Association of Private Lawyers in Africa
George Nyabuga, Aga Khan University
Joy Mwaniki, Shujaaz Inc.
Kennedy Wandera, Foreign Press Association, Kenya
Lena Slachmuislder, Council on Tech and Social Cohesion
Nancy Booker, Aga Khan University
Risper Onyango, Lawyers Hub

United Nations

Misako Ito, UNESCO Regional Office for Eastern Africa
John Okande, UNESCO Regional Office for Eastern Africa
Sharmaine Koh, UNESCO Regional Office for Eastern Africa
Sandra Macharia, UNIS (Kenya, Seychelles, Uganda)
Stella Vuzo, UNIS Nairobi
Janice Nduati, UNIS
Tirus Wainaina, UNIS
Haji Abdikareem, UNIS
Caroline Khamala, UNIS
Kathryn Porteous, UNHCR
Rana Taha, UN Kenya Resident Coordinator's Office
Jeshta Connery, UN Seychelles
Michael Wangusa, UN Uganda



**NATIONAL COALITION
ON FREEDOM OF EXPRESSION &
CONTENT MODERATION IN KENYA**



fecomo.org



FECoMo_Kenya



FECoMoKE



fecomo_kenya