



INFORMATION DISORDER AND RESILIENCE IN THE GLOBAL SOUTH: STRUCTURAL DRIVERS, GOVERNANCE, MEDIA LITERACY, AND FACT-CHECKING

**EDITED BY MELI M NCUBE, HERMAN WASSERMAN,
AND RUHIYA KRISTINE SEWARD**

Information Disorder and Resilience in the Global South



CCIJ Center for Collaborative
Investigative Journalism

INTERNETLAB



**Resisting Information Disorder
in the Global South**



Authors and Affiliations

Ruhiya Seward (IDRC)

Herman Wasserman and Meli M Ncube (Stellenbosch University)

Helani Galpaya and Isuru Samaratunga (LIRNEasia)

Ester Borges and Isabelle Fernanda dos Santos (InternetLab)

Scott Timcke and Zara Schroeder (Research ICT Africa (RIA))

Arwa Kooli, Ahmed Ashour, Saja Mortada, Fatima Bani Ahmad, and Menna

Elhosary (Arab Reporters for Investigative Journalism (ARIJ))

Acknowledgements:

The research presented in this publication was carried out with the financial assistance of Canada's International Development Research Centre (IDRC). The views expressed herein do not necessarily represent those of IDRC or its Board of Governors.

This project would not have been possible without the support, collaboration, and intellectual contributions of many individuals and institutions. We extend our sincere gratitude to the research team whose dedication and expertise shaped the development of this work: Herman Wasserman, Ruhiya Seward, Meli M Ncube, Ester Borges, Isabelle Fernanda dos Santos, Scott Timcke, Zara Schroeder, Isuru Samartunga, Helani Galpaya, Arwa Kooli, Ahmed Ashour, Saja Mortada, Fatima Bani Ahmad, and Menna Elhosary. Their scholarship and commitment to advancing research on information disorder in the Global South have been invaluable to this project.

We also wish to thank the peer reviewers of the chapters – Dr Catherine Musuva, Dr Marenet Jordaan, Dr Anthony Kaziboni, and Deshni Govender – whose thoughtful feedback and critical insights greatly strengthened the quality and rigour of this publication. Special thanks to Zandri van Zyl for designing this publication.

Our sincere appreciation also goes to the partner organisations, researchers, and practitioners who contributed their knowledge, perspectives, and engagement throughout the project. Their collaboration has been central to the broader goal of strengthening understanding of, and responses to, information disorder in the Global South.

This work is licensed under a **Creative Commons Attribution 4.0 (CC-BY 4.0) International License**.

Further information about the project can be found on the project website: <https://cinia.africa/idrc/>, and more information about IDRC is available at <https://idrc-crdi.ca/en/about-idrc>.

TABLE OF CONTENTS

CHAPTER 1

Information Disorder and Resilience in the Global South:
An Overview of the Issues and Challenges

1

.....
Meli M Ncube, Herman Wasserman, Ruhiya Seward

CHAPTER 2

When Journalism Can't Pay the Bills: Economic Precarity and
Information Integrity in Latin America

19

.....
Ester Borges and Isabelle Fernanda dos Santos

CHAPTER 3

Beyond Content Moderation: The Political Economy of
Information Disorder and Regulatory Pathways

42

.....
Scott Timcke and Zara Schroeder

CHAPTER 4

Empowering Children Against Misinformation: A Review of
Media and Information Literacy (MIL) Interventions

70

.....
Isuru Samartunga and Helani Galpaya

CHAPTER 5

The Facts About Fact-checking: Struggles and Possibilities in
Countering Information Disorder

96

.....
*Arwa Kooli, Ahmed Ashour, Zara Schroeder, Saja Mortada, Fatima
Bani Ahmad, Menna Elhosary*

CHAPTER 6

Conclusion: Strengthening Information Integrity in the
Global South

120

.....
Meli M Ncube

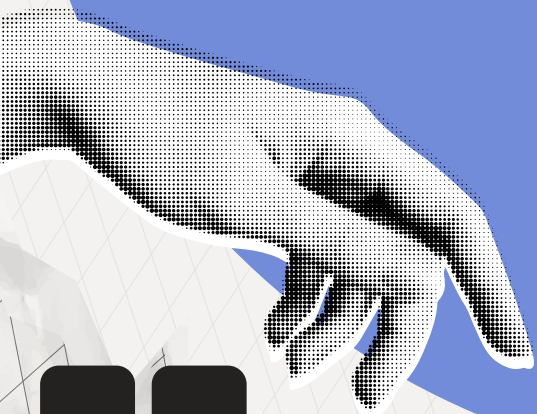
INFORMATION DISORDER AND RESILIENCE IN THE GLOBAL SOUTH:

An Overview Of The Issues And Challenges

Meli M Ncube,

Herman Wasserman,

Ruhiya Kristine Seward



INTRODUCTION

“Information disorder” is one of the most defining challenges of the digital era . As a framing for a broad set of experiences and content in the digital information ecosystem, “information disorder” encompasses the intentions and harms associated with misleading or harmful content and the complexities associated with navigating truth and harassing speech online. By comparison with the more popular, but problematic, term ‘fake news’, information disorder is one way to understand the processes by which unintentionally or strategically created narratives can manipulate social, political and economic outcomes (Wasserman et al, 2022; Wardle and Derakhshan 2017; Monsees, 2023).

Yet to understand the prevalence and impact of information disorder, particularly in the Global South, the phenomenon should be understood more broadly. Information disorder is not only about falsehoods or harmful speech in isolation but also how these connect to the deeper structural and

“ Information disorder is not only about falsehoods or harmful speech in isolation but also how these connect to the deeper structural and geopolitical vulnerabilities in global communication systems.

”

geopolitical vulnerabilities in global communication systems: Global platforms dominating local digital economies and allowing algorithms to drive the ‘attention economy’ that prioritizes engagement over accuracy; fragile regulatory systems that not only are unable to keep up with the pace of change but also lack the will or power to exert pressure on platforms; weak media institutions where governments limit access to quality or verifiable information or suppress freedom of expression; and political-economic structures where media is

concentrated into a few vectors of power. All these vulnerabilities exacerbate the risks of false or harmful content when intersecting with political and economic volatility. Countries in the Global South have historically been particularly socially, politically and economically volatile because of the forces of colonial subjugation, post-colonial tumult and contemporary global economic inequalities. But increasingly, the problems of information disorder can be seen across countries and regions, and across the spectrum of economic wealth, regulatory environments, and political systems.

Given the historical and contemporary challenges to the information environment in the majority world, there is a need for more scholarship rooted

in Global South contexts (Wasserman and Madrid-Morales, 2022; Tumber and Waisbord, 2024). This publication is a summative output from a project which aimed to do just that – bring together scholarship across the global South to unpack, document and analyse the drivers of and responses to information disorder. It is building on locally contextual research that shows the

- 1. structural drivers of information disorder, and the challenges associated with**
- 2. governance and platform regulation,**
- 3. media and information literacy, and**
- 4. fact-checking practices.**

It represents a collaborative effort across Sub-Saharan Africa, Latin America, the Middle East and North Africa, and Asia to understand and compare different approaches to the problems and responses.

“ Given the historical and contemporary challenges to the information environment in the majority world, there is a need for more scholarship rooted in Global South contexts.

”

This introductory chapter synthesizes the research across these four regions and areas of inquiry. It brings together the descriptive, normative, conceptual and practical approaches to understanding and countering mis- and disinformation. Synthesizing the research and discussions from across the regions, we can see

that while new technological tools (such as generative AI and sophisticated chatbots) are rapidly entering the information ecosystems, it remains crucial to understand the human and cultural dimensions around trust, perceptions, and local media legitimacy that contribute to the encroaching disorder and may offer possibilities to counter it. Moreover, these dimensions need to be understood within contexts and lived experiences. Diving further into the data generated by the project, comparing across regions can inform global and regional efforts.

STRUCTURAL DRIVERS OF INFORMATION DISORDER

MEDIA CONCENTRATION AND INEQUALITY

Across the Global South, information disorder is shaped by structural and historical conditions that undermine media independence and curtail the diversity of voices in public life. A good demonstration of these conditions can be found

in Latin America. Traditional media in the region has long been dominated by a small number of powerful conglomerates, with leading families such as the Marinhos in Brazil, the Claríns in Argentina, the Cisneros in Venezuela, and the Ardila Lülles in Colombia, which have exercised disproportionate control over media landscapes with close ties to political and economic elites (Becerra and Mastrini, 2019).

Media concentration manifests horizontally (multiple outlets of the same type owned), vertically (both production and distribution controlled), and cross-media (different media sectors and industries owned at once). In Brazil, for example, media diversity is superficial with 26 conglomerates composed of telecommunications and agribusiness interests dominating the media landscape (Intervozes, 2023). These corporate entanglements reduce journalistic independence and according to the 2025 Edelman Trust Barometer also fuel distrust: only 39% of Brazilians say they trust the media and 72% believe journalists mislead the public (Edelman, 2025).

Across the region, historical patterns of concentrated media ownership drive editorial agendas and elite influence. The elite and corporate alignment marginalizes grassroots movements and alternative narratives and entrench information asymmetries.

DIGITALIZATION DISRUPTION

Despite these entrenched ownership interests, global digital media platforms have shifted media consumption habits across Latin America as well, perhaps primed by the above dynamics. Now across the region, 65% of media consumers rely on social media for news with diminishing influence of legacy outlets (Vectors and Implications of Information Disorder in Latin America, 2025); in Brazil, 80% of the population consumes news digitally largely through WhatsApp and YouTube (Reuters Institute, 2024).

This shift in consumption is fragmenting the media landscape, segmenting audiences into ‘echo chambers’ that reduce collective news experiences and create individualized ‘information diets’ (Katz and Blumler, 1974), transferring local informational sovereignty to global corporations (Gingras 2022). Here algorithms replace local editors as key gatekeepers of visibility and prioritize engagement and virality over public relevance. As elsewhere, this emphasis on engagement may drive content that highlights conflict, polarisation and hate. Historical and traditional media brands are being replaced by digital influ-

encers and non-journalistic voices who can “appear” more credible than the legacy media (Reuters Institute, 2024). This last effect in particular highlights the profound disruption: local content creators taking priority in information ecosystems based on algorithmic choices in global platforms. The downstream impacts - divergent narratives, sometimes manipulated from the outside - make consensus and democratic deliberation on public issues harder to achieve.

The disruption of advertising-based revenue models by big digital platforms have created huge sustainability challenges for journalism worldwide. This can also be seen in Latin America, where advertising revenues have also migrated from traditional media to digital platforms. Corporations like Google and Meta are now capturing 60–80% of digital advertising markets in Latin America (Rolnik et al, 2019; Clara, 2025). Programmatic advertising channels funding toward engagement-driven content, undermining editorial independence. This dependence on platform metrics weakens quality journalism and reinforces the dominance of sensational, click-driven material.

DIGITAL COLONIALISM AND REGULATORY GAPS

Regulation can potentially mitigate some of the problems created by rapid and widespread digitalisation of the media. However, developing regulatory systems that are appropriate for context, and do not contribute to information disorder when abused by authoritarian governments, is a challenge in the Global South. The situation in Sub-Saharan Africa is illustrative of these problems. Across the African continent, the dynamics of information disorder are shaped by weak and reactive regulation, and dependence on foreign platforms and infrastructures. The rules and monetization systems are set externally but experienced locally, in what is sometimes referred to as new type of ‘digital’ colonialism (Kwet, 2019; Young, 2019; Nothias, 2025). In practice, it means that, like Latin America, foreign platforms dominate advertising markets, but their investments in content moderation or local language support are limited (Sambuli, 2021). Moderation, especially when managed externally or through automation, is poorly equipped to identify problematic content in local languages or which draws on local idiom and context. This asymmetry creates vulnerabilities to disinformation and undermines local media sustainability.

Regulatory frameworks in many African states are underdeveloped or highly politicized. While governments justify regulation to tackle harmful content, many “fake news” laws are vague and open to abuse. In Nigeria, Kenya, and Tanzania, such laws have been used to silence journalists and activists rather

than to protect citizens, undermining public trust and restricting civic space (Article 19, 2019). Linguistic diversity across the continent also complicates content moderation (when it exists) because platforms lack the capacity to detect harmful content in indigenous languages, leaving large populations vulnerable to inaccurate information.

The [Praia Declaration on Information Integrity](#), adopted in September 2025, as a result of the Regional Conference on Information Integrity in West Africa and the Sahel, is one of the latest African-led declarations aiming to strengthen information ecosystems across the region. The declaration is part of a broader, but perhaps also reactive, effort to build greater trust in the information ecosystem and advance digital platform governance. By addressing the increasing threats of disinformation, misinformation, hate speech, and misuse of digital technologies, the declaration sets out guiding principles for protecting the integrity of information environments while safeguarding human rights, ensuring transparency and accountability, and promoting freedom of expression and access to information.

FRAGMENTED REGULATION AND POPULIST POLITICS

The problems of regulation also manifest in other regions of the Global South. Asia's diverse regulatory and media environment similarly struggles with tackling the multiple challenges and are turning to local regulatory fixes. In India, regulations have tightened around platform traceability, requiring encrypted messaging services to reveal the origins of viral content that if unchecked, stand to undermine privacy and free expression (Udupa et al, 2021). Indonesia and Malaysia have passed broad "fake news" bills that are increasingly understood as tools for political control. In Africa, governments often use the smokescreen of "fake news" to clamp down on criticism and dissent, including shutting down the internet during periods of heightened political contestations like elections (Marchant and Stremmlau, 2020; Mare, 2020). Yet fact-checking networks, community-based literacy programs, and other grassroots responses to mis- and disinformation are also playing an important role across the region in filling governance gaps, while also operating in highly polarized environments where trust is low and sustainable funding is lacking.

STRUCTURAL IMPACTS ON JOURNALISM AND TRUST

Information disorder presents systemic vulnerabilities far beyond the chal-

lenges of false content. The problem is multi-faceted, including platform dependence, fragile regulatory systems, historically concentrated ownership and eroding trust in journalism. Audiences across the global South increasingly rely on platforms where incentives favour sensational and polarizing content, while traditional outlets struggle financially. These conditions highlight the limitations of technical fixes, but also the uphill challenges of how to address underlying structural drivers and weakened media environments. Media pluralism, economic sustainability, and institutional trust are all critically important and their absence or weakness represent durable vulnerabilities in information ecosystems.

GOVERNANCE AND PLATFORM REGULATION

DEFINING GOVERNANCE APPROACHES

State regulation, platform self-regulation, and co-regulation are the three main approaches to governance, referring to laws and policies enacted by governments (state), voluntary measures taken by platforms and media organizations (self-regulation), or a blend of state oversight and private compliance (co-regulation) (Gorwa, 2019). Each approach raises issues of legitimacy, effectiveness, and human rights.

“
In Africa, governments often use the smokescreen of “fake news” to clamp down on criticism and dissent, including shutting down the internet during periods of heightened political contestations like elections.

STATE REGULATION

Governments across the Global South have passed laws to address disinformation, but many are criticized as vague at best, or repressive at worst. Singapore’s Protection from Online Falsehoods and Manipulation Act (POFMA) (2019), for example, allows ministers to order corrections or removals of content. While justified as protecting democracy, it has been criticized for enabling censorship (Tan, 2020). Across the Middle East and North Africa, governments have adopted similar cybercrime laws that aim in

theory to counter fake news with harsh penalties but also impact dissent and legitimate expression. Jordan’s Cyber Crime Law, for instance, passed in September 2023, has already had a chilling effect on speech, civic engagement and media integrity. Poor judicial oversight, harsh penalties that include prison and heavy fines, and the laws’ vague and widely interpretable definitions of terms like ‘hate speech’ have already been used to silence legitimate criticism (Kooli et al, 2026; Halasa, 2025).

Across the African continent, fake news laws have been selectively applied against journalists and opposition voices (Article 19, 2019). In Latin America, Brazil's proposed fake news bill aimed to improve transparency in messaging apps, but critics warn it could undermine privacy and free speech (Fenwick et al, 2021).

These examples highlight the central governance tension: ensuring information integrity while also protecting free expression. Some governments will frame regulation as necessary to protect democracy, but in the contexts without robust protections/rule of law, civil society groups warn that many of these laws also risk shrinking civic space and silencing dissent (Kaye, 2019; Garbe et al, 2023; Wahutu, 2019). Mitigating these risks is one of the most difficult challenges for information integrity. Legal frameworks must be transparent, narrowly defined, and subject to judicial review, and civil society needs to be engaged in drafting legislation to enhance legitimacy and safeguard rights. The political, social and economic conditions in the Global South requires multi-stakeholder engagement around regulation. Regulatory frameworks cannot merely be imported from other regions. These are not straightforward propositions in many contexts.

PLATFORM SELF-REGULATION

Digital platforms such as Meta, Google, X (formerly Twitter) and TikTok are central to the contemporary information ecosystem. Prior to 2025, many had measures such as fact-checking partnerships, algorithmic moderation, and transparency reports. However, by 2025, many of these measures, already uneven, were being rolled back. This is happening at the same time as governments around the world are finding ways to limit encryption and to expand legal tools for surveillance and control (Alphonso et al, 2025). This demonstrates the accountability issues inherent in self-regulation: platforms act as regulators but lack democratic legitimacy. Platform's policies are often opaque, and commercial interests frequently outweigh public interest considerations (Gillespie, 2018), and sometimes their reliance on automated tools do not capture cultural nuance, leading to both under- and over-enforcement. This underscores the limits of self-regulation and the need for human oversight informed by local expertise.

CO-REGULATION

Co-regulation could offer a compromise between state and private power, albeit with a range of complexities that speak to the structural drivers. In this instance, governments can set the frameworks while requiring companies to implement them under independent oversight. The European Union's Code of

Practice on Disinformation (European Commission, 2018) is one prominent model, as is the Praia Declaration. The EU's approach will often inform debates across the majority about adapting co-regulation to local contexts, in particular Latin America and Africa (Napoli and Caplan, 2017).

CIVIL SOCIETY, JOURNALISM AND GOVERNING TENSIONS

Civil society organizations, journalists, and fact-checkers play crucial roles in countering disinformation. Networks such as the Arab Fact Checkers Network (AFCN), Africa Check and LatamChequea provide verification services and promote literacy (ARIJ 2025, Graves and Mantzarlis, 2020). Fact-checking networks show that governance cannot rely solely on top-down regulation but also needs digitally literate citizens, independent media, and transparent governmental and corporate practices. Increasingly, the limits of fact-checking as a countermeasure to information disorder are also becoming clear (Graves, 2018; Vinhas and Bastos, 2022), serving as a reminder that debunking alone is not sufficient as a countermeasure to stem the tide of information disorder.

Yet the tensions over democratic values of transparency and expression versus control, global versus local, and public versus private power and goods remain the key. Overly restrictive laws threaten free expression, while weak regulation leaves societies vulnerable; stakeholders need skills to critically engage with information environments, making media and information literacy one type of future-proofing response.

MEDIA AND INFORMATION LITERACY

FOUNDATIONS OF MEDIA AND INFORMATION LITERACY

Media and information literacy (MIL) is increasingly recognized as another part of the response to information disorder to equip people with the skills to access, analyse, evaluate, and create information critically and responsibly (Seo et al, 2019; Cohen et al, 2020). It can be adapted locally and help foster a filter of scepticism, encouraging readers to cross-check sources and resist manipulation (LIRNEasia, 2025, Dumitru, 2020; Ng, 2012). Global platforms have invested in this approach: Meta's *We Think Digital* and Google's *Be Internet Awesome* are being implemented in Sri Lanka through Sarvodaya Fusion, using interactive content like the educational game Interland (Sarvodaya Fusion, 2025), illustrating a role for public-private partnerships in scaling literacy efforts. While scholars have pointed to the importance of expanding MIL programmes to

include a specific focus on misinformation, they also warn against a one-size-fits-all approach to curriculum development that is not sensitive to contextual specificities in Global South contexts (Madrid-Morales and Wasserman, 2025).

MYTHS OF THE DIGITAL NATIVE

Prensky (2001) popularized the idea that children are digital natives, inherently skilled at using technology. However, subsequent research (Kirschner and De Bruyckere, 2017) has shown that this narrative can be misleading because it often results in young people using technology passively, without cultivating critical evaluation skills. Despite growing up in digital environments, young people remain vulnerable to misinformation. Assuming digital fluency can lead to underinvestment in foundational literacy programs, particularly in underserved communities, which is why early education should begin integrating AI and data literacy to address both cognitive and emotional vulnerabilities (World Economic Forum, 2025). Assuming that such programmes can be uniformly rolled out across different contexts is a mistake – structural factors such as access to devices, provision of data and even availability of electricity, as well as social factors such as supportive family and home environments, must be considered in each specific case (Madrid-Morales and Wasserman, 2025).

CASE STUDY: MEDIA AND INFORMATION LITERACY AT THE GRASSROOTS IN SRI LANKA

Between 2023 and 2025, LIRNEasia conducted a 12-month longitudinal study on Sarvodaya Fusion's MIL program, adapted from Google's Be Internet Awesome. The program targeted children aged 11–14 across diverse socio-economic regions of Sri Lanka. Ninety children and 91 parents participated, engaging with five digital safety principles: Be Smart, Be Alert, Be Strong, Be Kind, and Be Brave.

Findings showed several positive outcomes. Children reported greater awareness of online risks, improved skills such as verifying sources and creating strong passwords, and more cautious sharing behaviour. Many participants shared their knowledge with peers and families, extending the reach of the intervention. However, retention varied. Simple skills like password creation were remembered, while complex practices like triangulating sources faded over time. Without integration into school curricula, the long-term impact of MIL programs was limited.

This suggests that while media and information literacy can transform children from passive consumers into informed citizens, it would benefit from being reinforced at systemic levels. Short-term interventions risk losing impact without curricular integration and parental involvement.

STRENGTHS AND LIMITS OF MEDIA LITERACY

Media literacy alone is not the answer to information disorder. It can deepen inequalities, especially where systemic issues like algorithmic bias and platform dominance remain unaddressed and must be embedded in institutional and regulatory reforms (World Economic Forum, 2025). But it can also foster resilience when combined with other governance reforms, like fact-checking and civic education to help foster critical thinking on governance, rights, and responsibilities.

FACT-CHECKING IN THE GLOBAL SOUTH

FACT-CHECKING AS EDITORIAL AND CIVIC PRACTICE

Fact-checking benefits from and reinforces media literacy by offering real-world verification practices and is another response to information disorder. Fact-checkers identify, analyse, and publicly correct misleading claims circulating online (Graves, 2016). Traditionally, verification was conducted pre-publication within newsrooms to ensure credibility, but the flood of online information has made post-publication fact-checking equally important. It can serve as both an editorial function and civic practice in public life.

STRUGGLES AND POSSIBILITIES OF FACT-CHECKING

Deficits in trust undermine effectiveness, particularly in polarized environments, where corrections can be dismissed as partisan (Lewandowsky et al, 2012). Distrust of both state-aligned and independent media throughout the world mean fact-checkers can struggle to build credibility.

False information also spreads faster and wider than corrections, creating an asymmetry between disinformation producers and fact-checkers and limited resources constrain fact-checkers' reach (Vosoughi et al, 2018). Furthermore, in authoritarian or proto-authoritarian contexts there are political and legal pressures – from surveillance to censorship to legal threats - that restrict fact-checkers' ability to challenge powerful actors. There are alliances that can serve as models for resilience under pressure, and in some contexts, fact-checkers turn to these coalitions and legal advocacy groups to protect their rights and ensure operational continuity (Grbeša Zenzerović and Nenadic, 2022; Eze and Zacky-Eze, 2025). But sustainability is always precarious. Many, if not most, fact-checking organizations rely on funding from inter-

national or bilateral donors, leaving them vulnerable to shifting priorities – something which is particularly highlighted in 2025 with the disappearance of USAID and other sources of US’s funding (AFCN, 2024).

In combination with this vulnerability, the technological challenges are growing with the rise of artificial intelligence and auto-generated synthetic content, including deepfakes and fabricated text; it can be difficult for fact-checkers to keep up. Yet AI represents both a challenge and an opportunity. On the one hand, generative AI produces synthetic content that can erode trust in what

“ In combination with this vulnerability, the technological challenges are growing with the rise of artificial intelligence and auto-generated synthetic content, including deepfakes and fabricated text.

is real and what is generated (Chesney and Citron, 2019). On the other hand, AI tools support verification by automating detection tasks and identifying patterns in disinformation networks. AI tools such as reverse image search and forensic software help identify manipulated content (Paris and Donovan, 2019). The key challenges are ensuring the use of AI is transparent, accountable, and adapted to local contexts. Furthermore, training datasets should reflect regional languages and cultural references to ensure relevance and accuracy.

”

Regional networks like the Arab Fact-Checkers Network (AFCN) and the International Fact-Checking Network (IFCN) facilitate collaboration, training, and resource-sharing. These networks are testing how best to use AI tools and experimenting with public engagement strategies, such as short videos, memes, and gamified verification which could resonate with younger audiences and expand reach.

Despite the struggles, fact-checking, like media and information literacy, holds significant potential as an intervention. Integrating fact-checking into journalism training and education strengthens verification at both institutional and individual levels. Linking fact-checking with media literacy initiatives can foster long-term resilience (Ireton and Posetti, 2018) – but these are most effective with governance and sustainable media systems; otherwise, fact-checking is a short-term patch in a structurally weak information environment.

CONCLUSION

Information disorder in the Global South is a complex problem shaped by historical inequalities, changes in media systems, and the growing influence of

CASE STUDY: LESSONS IN ARABIC FACT CHECKING

Research by the Arab Fact-Checkers Network (AFCN) and the Arab Reporters for Investigative Journalism (ARIJ) in 2023 and 2024 reveals both persistent constraints and emerging innovations in the region. Political hostility toward independent media, limited financial resources, and the complexity of linguistic diversity continue to pose significant obstacles to fact-checking initiatives. These challenges are compounded by varying levels of digital literacy among audiences, which affect the reach and impact of verification efforts.

Despite these hurdles, promising developments are evident. Cross-border collaborations among fact-checking organizations have strengthened knowledge-sharing and resource pooling, while the integration of AI-driven tools for content verification signals a shift toward more scalable and efficient practices. Importantly, the experiences of Arabic-language fact-checking underscore a critical lesson: such efforts cannot succeed in isolation. They require systemic support through robust media ecosystems, comprehensive literacy programs, and governance reforms that promote transparency and accountability.

digital platforms. Evidence from Latin America, Africa, Asia and the Middle East regions show that technical fixes alone cannot solve this issue because its roots lie in political, economic, and cultural contexts, as well as in the design of digital technologies. Solutions must combine structural reforms with efforts to build citizens' skills, while recognizing that platform dominance, weak journalism, and uneven regulation limit the quality and trustworthiness of public information.

To strengthen information integrity, both systemic and educational approaches are needed. Media and information literacy (MIL), especially when it includes functional, critical, and rhetorical skills, can empower citizens and support healthier democratic dialogue. Research on children's media and information literacy programs shows that well-designed initiatives can improve digital habits, critical thinking, and resilience to misinformation, though advanced analytical skills require ongoing support from schools, families, and broader education systems. Fact-checking remains vital, but its impact is limited by political pressure, scarce resources, and declining public trust. New literacies, such as AI and algorithmic literacy, are also essential for navigating today's fast-changing information environment.

At a structural level, the collapse of traditional media revenue and the rise of platform-driven models have weakened independent journalism at a time when democracy needs it most. Yet this disruption has also created opportunities for innovative, community-based, and digital-first journalism that challenge old patterns and promote diversity and accountability. Sustainable solutions will require coordinated action across policy, civil society, and new institutions to support media diversity, protect editorial independence, and ensure fair access to reliable information.

Overall, building strong and credible information ecosystems calls for an integrated approach: promoting media plurality, regulating platform power, taxing digital corporations fairly, and banning technologies that harm labour rights or democracy. It also means investing in literacy, fact-checking, and public-interest media as interconnected pillars of democratic information systems. Innovations across the Global South, from education programs to new journalism models, show that while challenges are serious, there is growing capacity to protect citizen rights and strengthen democracy in the digital age.

As the research in this volume will show, information disorder is a global challenge shaped by many factors - structural, political, and historical - in the Global South. Responses are multi-layered and cannot rely on technical solutions alone; they must be attentive to the political, economic, and cultural contexts in which information circulates, and the interplay between digital architecture and socio-political dynamics.

References

- Alphonso, A.; Barbosa, S., Clifford, C., Garimella, K.; Hickok, E.; Riedl, M., Saka, E.; Wasserman, H. & Udupa, S. (2025). *Encrypted Messaging and Extreme Speech: Policy Directions*. <https://doi.org/10.5282/ubm/epub.127250>
- ARIJ's AFCN team. 2024. *AI-Generated Challenges for Fact-Checkers in the Arab World*. IDRC, Stellenbosch University. Available at: <https://idrc.sun.ac.za/> (accessed 27 Nov. 2025).
- ARIJ's AFCN team. 2024. *Enhancing the Organisational Framework of Arab Fact-Checking Initiatives*. IDRC, Stellenbosch University, 8 Aug. Available at: <https://idrc.sun.ac.za/enhancing-the-organisational-framework-of-arab-fact-checking-initiatives/> (accessed 27 Nov. 2025).
- ARIJ's AFCN team. 2024. *Studying Information Disorder in the Arab World: Drivers and Fact-Checking Response*. IDRC, Stellenbosch University. Available at: <https://idrc.sun.ac.za/> (accessed 27 Nov. 2025).
- ARIJ's AFCN team. 2024. *Transforming Literacy for the Digital Age: Strategies to Boost Reading and Learning in the Arab World*. IDRC, Stellenbosch University. Available at: <https://idrc.sun.ac.za/> (accessed 27 Nov. 2025).
- Borges, E. & dos Santos, I.F. 2026. 'When journalism can't pay the bills: Economic precarity and information integrity in Latin America', in Ncube, M.M., Wasserman, H. & Seward, R.K. (eds.) *Information Disorder and Resilience in the Global South: Structural Drivers, Governance Challenges, and the Role of Media Literacy and Fact-Checking*.
- Damasceno, C.S. (2021). Multiliteracies for combating information disorder and fostering civic dialogue. *Social media+ Society*, 7(1). <https://doi.org/10.1177/2056305120984444>
- Eze, U. Z., & Zacky-Eze, C. J. (2025). Combatting disinformation: Analyzing effective communication strategies and tactics for organizational resilience. *Tripodos*, (57), 01. <https://doi.org/10.51698/tripodos.2025.57.01>
- Garbe, L., Selvik, L., & Lemaire, P. 2023. How African countries respond to fake news and hate speech. *Information, Communication & Society* 26(1):86-103
- Graves, L. (2018). Understanding the promise and limits of automated fact-checking (RISJ Factsheets). Reuters Institute for the Study of Journalism. <https://ora.ox.ac.uk/objects/uuid:f321ff43-05f0-4430-b978-f5f517b73b9b>
- Graves, L. and Mantzarlis, A., 2020. Amid political spin and online misinformation, fact checking adapts. *The Political Quarterly*, 91(3), pp.585-591.
- Grbeša Zenzerović, M., & Nenadić, I. (2022). *Strengthening resilience to disinformation: The state of affairs and guidelines for action* (Second supplemented edition, Zagreb, September 2022). Agency for Electronic Media.

- Galpaya, H.; Jayasinghe, G.; Bandaranayake, R.; Ababakirov, A.; de Silva, R. (LIR-NEAsia Research Team). 2024. *Exploring AI in Fact-Checking: Perspectives from Fact-Checkers*. IDRC, Stellenbosch University. Available at: <https://idrc.sun.ac.za/> (accessed 27 Nov. 2025).
- Halasa, Ayman (2025) *The Cybercrimes Law – Legal Analysis*. King Hussein Foundation Information Research Center.
- Kooli, A., Ashour, A., Schroeder, Z., Mortada, S., Bani Ahmad, F. & Elhosary, M. 2025. ‘The facts about fact-checking: Struggles and possibilities to countering information disorder in the Arab world’, in Ncube, M.M., Wasserman, H. & Seward, R.K. (eds.) *Information Disorder and Resilience in the Global South: Structural Drivers, Governance Challenges, and the Role of Media Literacy and Fact-checking*.
- Kwet, M. (2019). Digital colonialism: US empire and the new imperialism in the Global South. *Race & Class*, 60(4), 3–26. <https://doi.org/10.1177/0306396818823172>
- Madrid-Morales, D. & Wasserman, H. (2025). Cynical or Critical Media Consumers? Exploring the Misinformation Literacy Needs of South African Youth. *African Journalism Studies* 45(4):327-345
- Marchant, E. & Stremlau, N. (2020). Internet Shutdowns in Africa: The Changing Landscape of Internet Shutdowns in Africa. *International Journal of Communication* 14 (2020). <https://ijoc.org/index.php/ijoc/article/view/11490>
- Mare, A. (2020). Internet Shutdowns in Africa: State-Ordered Internet Shutdowns and Digital Authoritarianism in Zimbabwe. *International Journal of Communication* 14 (2020). <https://ijoc.org/index.php/ijoc/article/view/11494>
- Monsees, L. (2021). Information disorder, fake news and the future of democracy. *Globalizations*, 20(1), 153–168. <https://doi.org/10.1080/14747731.2021.1927470>
- Ncube, M. M. 2025. *The Language of Change: Exploring Sustainable Transitions*. IDRC, Stellenbosch University. Available at: <https://idrc.sun.ac.za/> (accessed 27 Nov. 2025).
- Nothias, T. (2025). An intellectual history of digital colonialism. *Journal of Communication*. Advance online publication. <https://doi.org/10.1093/joc/jqaf003>
- Samartunga, I. & Galpaya, H. 2025. ‘Empowering children against misinformation: A review of media and information literacy (MIL) interventions in Sri Lanka’, in Ncube, M.M., Wasserman, H. & Seward, R.K. (eds.) *Information Disorder and Resilience in the Global South: Structural Drivers, Governance Challenges, and the Role of Media Literacy and Fact-checking*.
- Schroeder, Z. 2025. *Navigating the Intersection of AI, Media, and Democracy: Insights from the CJID Conference 2024*. IDRC, Stellenbosch University. Available at: <https://idrc.sun.ac.za/> (accessed 27 Nov. 2025).

Schroeder, Z. & Haller, D. 2025. *Grok AI: A Conduit for Misinformation in the Digital Age*. IDRC, Stellenbosch University, 26 June. Available at: <https://idrc.sun.ac.za/> (accessed 27 Nov. 2025).

Selber, S.A., (2004). *Multiliteracies for a digital age*. Southern Illinois University Press.

Timcke, S. 2024. *AI Regulation Paradigms and the Struggle over Control Rights*. IDRC, Stellenbosch University. Available at: <https://idrc.sun.ac.za/> (accessed 27 Nov. 2025).

Timcke, S. & Rens, A. 2024. *Using Regulation to Curtail Platform Decay and Information Disorders*. IDRC, Stellenbosch University. Available at: <https://idrc.sun.ac.za/> (accessed 27 Nov. 2025).

Timcke, S. & Schroeder, Z. 2024. *Engagement at What Cost? Examining the Intersection of Social Media, Generative AI and Gender-Based Violence*. IDRC, Stellenbosch University. Available at: <https://idrc.sun.ac.za/> (accessed 27 Nov. 2025).

Timcke, S. & Schroeder, Z. 2025. 'Beyond content moderation: A political economy of information disorders and regulatory pathways', in Ncube, M.M., Wasserman, H. & Seward, R.K. (eds.) *Information Disorder and Resilience in the Global South: Structural Drivers, Governance Challenges, and the Role of Media Literacy and Fact-checking*.

Vinhas, O. & Bastos, M. (2022). Fact-checking misinformation: Eight notes on consensus reality. *Journalism Studies* 23(4): 448-468. <https://doi.org/10.1080/1461670X.2022.2031259>

Wahutu, J.S. (2019). Fake News and Journalistic "Rules of the Game" *African Journalism Studies* 40(4): 13-26

Wardle, C., & Derakhshan, H. (2017). *Information disorder: Toward an interdisciplinary framework for research and policymaking* (DGI(2017)09) (pp. 1–109). Strasbourg: Council of Europe. <https://rm.coe.int/information-disorder-report-version-august-2018/16808c9c77>

Wasserman et al. 2022. Meeting the challenges of information disorder in the Global South. IDRC - International Development Research Centre. <https://idl-bnc-idrc.dspacedirect.org/items/5aad1667-cf56-44bf-b85f-8695cb5f11dd>

Young, J. C. (2019). The new knowledge politics of digital colonialism. *Environment and Planning A: Economy and Space*, 51(7), 1424–1441. <https://doi.org/10.1177/0308518X19858998>

Zwane-Siguqa, M. 2025. *Increasingly, people in positions of power spread blatant falsehoods without fear of any consequence – Andrew Thompson*. IDRC, Stellenbosch University. Available at: <https://idrc.sun.ac.za/> (accessed 27 Nov. 2025).

Endnotes

¹ Wardle and Derakhshan (2017) conceptualize information disorder as comprising misinformation, disinformation, and malinformation. Misinformation refers to false information shared without intent to harm, disinformation describes deliberately misleading or fabricated content intended to cause harm, and malinformation involves the strategic use of genuine information to damage individuals or groups. Although this framework has been contested, partly because the conceptualisation is premised on the assumption that intent can be ascertained, it provides a more nuanced understanding of information disorder than the term ‘fake news’ because it includes a broader range of harmful information, and because ‘fake news’ has often been weaponised to dismiss critical journalism.

²A notable part of this declaration is that it was informed by the research conducted by scholars contributing to this book. The declaration is reactive in the sense that it is consistent with the paradigm many Southern scholars have been pushing for the last 15 years, with evidence to adjust platform policies. It is proactive in the sense that there is regional integration and buy in for regulator testing, proactive disclosure by states on data access, data as a public good.

³ In January 2025, UNESCO awarded Research ICT Africa (RIA) a no-bid contract to serve as technical advisors for the Praia Model Policy Framework for Information Integrity. This role arose from RIA’s outputs from the IDRC Resisting Information Disorder project. RIA’s team also provided technical advisory services for the accompanying regulator guide, which aims to operationalize the framework. The framework was workshopped with West African civil society stakeholders, where RIA contributed inputs and guided discussions, incorporating direct feedback from states, and in early September, the framework was signed by 14 West African ministers.

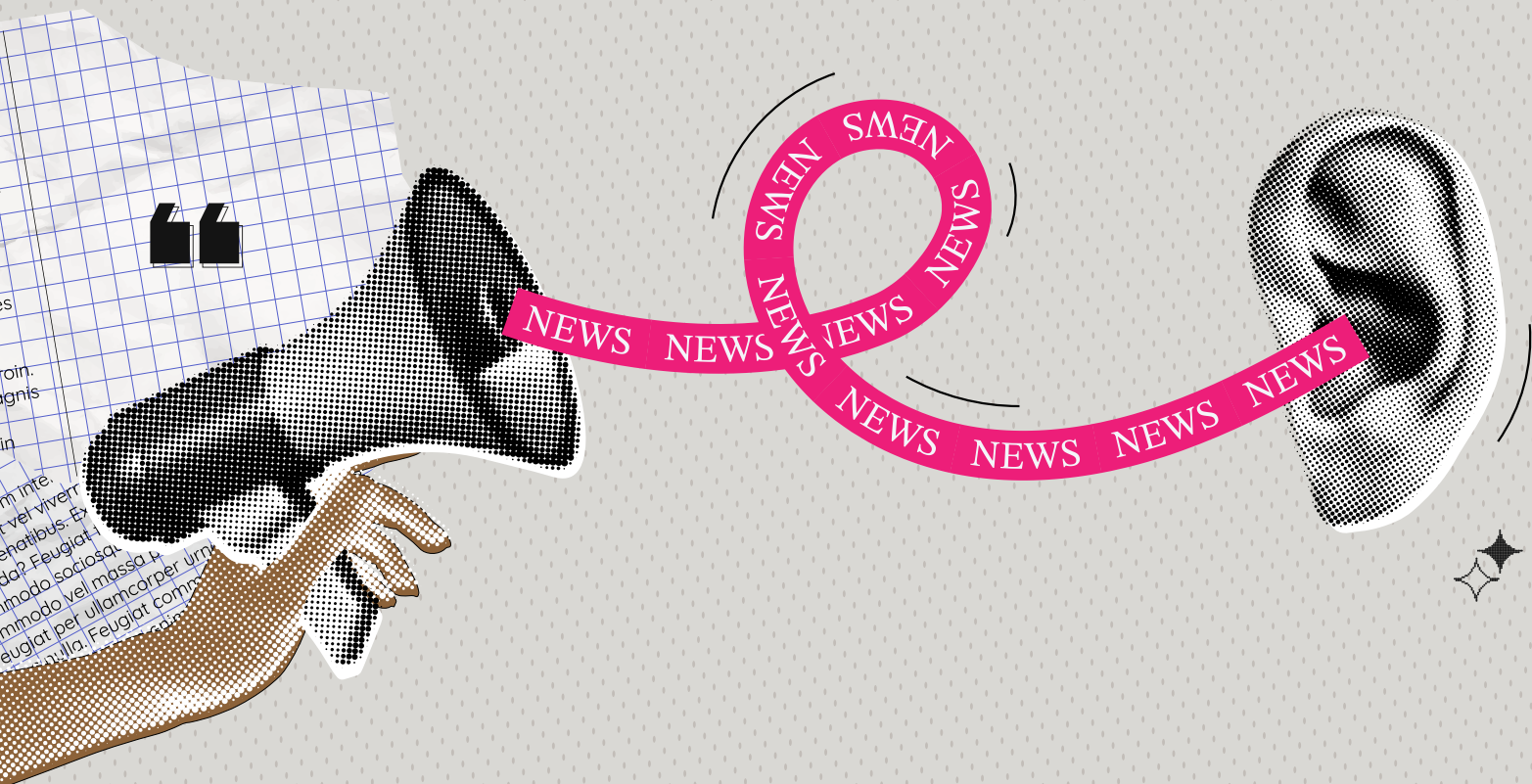
⁴ Research shows that content moderation resources are already heavily concentrated in the Global North, with far fewer resources for African and Asian languages (Gorwa, 2019; Sambuli, 2021). Facebook’s failure to moderate hate speech in Myanmar, for instance, contributed to violence against the Rohingya and illustrates the dangers of neglecting local contexts (Mozur, 2018).

WHEN JOURNALISM CAN'T PAY THE BILLS:

Economic Precarity And Information Integrity In Latin America

Ester Borges

Isabelle Fernanda dos Santos



INTRODUCTION

Latin America has undergone a rapid shift in its media consumption, marked by the expansion of digital news, primarily through smartphones, social media, and search engines. More than 70% of this population uses social media, messaging apps or websites in general as a primary source of information (InternetLab and Rede Conhecimento Social, 2025). While this shift has expanded access, it has not necessarily resulted in greater diversity or higher-quality information.

In Latin America, the digital transformation process has not always broken with historical patterns of media concentration; in many cases, it has reinforced them. The combination of traditional family oligopolies and the penetration of transnational capital has resulted in new configurations of informational power, heightening the risks to plurality. Countries like Brazil, Argentina, and Colombia have undergone accelerated transitions to digital consumption, but with uneven impacts (Borges, Pavarin e Vilela, 2023). In Argentina and Brazil, news pages on social media and online portals dominate a substantial portion of the digital ecosystem. However, this increased access has not led to greater diversity or improved financial sustainability for the communications sector. In Brazil, the growth of digital consumption has expanded funding possibilities. Still, it has also intensified dependence on models driven by virality, to the detriment of editorial quality. As highlighted

“ This chapter departs from the question of how digitalization has impacted journalism business models and, consequently, transformed the media landscape in the region.

in the *Intervozes* report (2023), formerly dominant groups have succeeded in adapting their strategies and maintaining positions of power in the digital environment, contributing to the formation of “new informational monopolies”.

”

This chapter departs from the question of how digitalization has impacted journalism business models and, consequently, transformed the media landscape in the region. We begin with the hypothesis that digitalization, by interacting with historically concentrated media systems, has deepened existing inequalities and generated new informational vulnerabilities. Through descriptive analysis, we argue that the transformation of journalism’s economic foundations, which has reshaped incentives, production models, and the mechanisms that sustain journalism

in the region, must be understood as a critical component when discussing information disorder in Latin America.

Our objective is not to establish direct causal links between economic transformations and specific disinformation phenomena, but rather to demonstrate that any comprehensive discussion of information integrity in the region must account for these structural changes in the media ecosystem. The economic destabilization of quality journalism, the migration of advertising to digital platforms, and the subordination of editorial decisions to algorithmic metrics create conditions in which reliable information becomes increasingly scarce and economically unsustainable, a scenario that demands urgent attention in debates about democratic information flows.

To develop this argument, the chapter is divided into three parts. The first outlines a historical overview of the media structure in Latin America, marked by high levels of concentration and persistent challenges to diversity, independence, and informational plurality. The second section analyses how digitalization has reconfigured the economic foundations of journalism, focusing on three central dimensions: changes in media consumption patterns, the disruption caused by digital platforms to traditional models of distribution and editorial control, and the redesign of the advertising market. This section concludes by positioning these transformations as critical elements that must be considered in discussions of information disorder, showing how the economic destabilization of journalism creates structural vulnerabilities for information integrity. Finally, the third part explores the paths currently being contested in the region to reconfigure this market: from the emergence of new funding models and the barriers introduced by paywalls to public policy proposals aimed at supporting journalism. Recognizing the potential unlocked by digitalization, including the emergence of historically marginalized voices, we seek to reflect on how to structure a more plural, sustainable, and resilient information ecosystem.

HISTORICAL MEDIA CONCENTRATION IN LATIN AMERICA

The media ownership structure in Latin American countries is marked by high concentration and dominance of a small number of conglomerates, often controlled by families with close ties to political and economic elites. Groups such as the Marinho family in Brazil, the Clarín family in Argentina, the Cisneros family in Venezuela, and the Ardila Lülle family in Colombia

exemplify how family-based structures operate as true informational dynasties, influencing the public agenda and shaping political debate.

In the region, media concentration takes three primary forms: horizontal, vertical, and cross-ownership. Horizontal concentration occurs when a group controls several media outlets of the same type (such as multiple newspapers). Vertical concentration occurs when a company dominates different stages of media production (from creation to distribution). Cross-ownership concentration, the most concerning, involves holding assets in different types of media, as well as in other economic sectors, such as telecommunications and agribusiness (Becerra and Mastrini, 2019).

In Brazil, for example, a survey by Intervezes (2023) showed that the apparent diversity of the national media is concentrated in 26 conglomerates, many of which hold cross-sector interests in different sectors of the economy, such as telecommunications, agribusiness, and infrastructure. These conglomerates not only own TV and radio channels, but also print newspapers, digital portals, and streaming platforms, reinforcing their power to vertically integrate information.

The opacity surrounding the links between media groups and other spheres of power represents an additional challenge. The link between media ownership and the exercise of political office is a structural characteristic of the region. Media-based clientelism manifests in the approval of radio and television licenses to political allies, often in exchange for electoral support. Specifically in Brazil, families such as the Sarneys (Maranhão), the Barbalhos (Pará), and the Collors (Alagoas) exemplify this arrangement, which confers disproportionate advantages on politicians with direct control over media outlets, distorting electoral contests and limiting the plurality of voices. (Intervezes, 2023; Valente, 2013:102).

The high concentration of media outlets, controlled by a few families and conglomerates with close ties to political elites, directly contributes to public distrust in the media. Studies show that the presence of family conglomerates with cross-sector interests in areas such as telecommunications, agribusiness, and infrastructure undermines independent journalism and creates perceptions of bias and hidden political agendas (Intervezes et al, 2017). According to a 2025 report by the Reuters Institute (Newman, et al, 2025), only 42% of Brazilians

with internet access say they trust major news outlets - the lowest level in a decade. In another survey, 67% of respondents say that government leaders also lie frequently, further reinforcing the association between media and politics as a source of institutional distrust (Contado, 2025). Research conducted by InternetLab and Rede Conhecimento Social (2025) shows that distrust also arises from the widespread perception that much of the information circulating about politics, elections, and public security is either false, exaggerated, or taken out of context, fuelling scepticism towards both traditional and digital media.

“
Research conducted by InternetLab and Rede Conhecimento Social (2025) shows that distrust also arises from the widespread perception that much of the information circulating about politics, elections, and public security is either false, exaggerated, or taken out of context, fuelling scepticism towards both traditional and digital media.”

This disconnect between journalistic practice and public perception was analysed by Friderichsen and Lobos (2018) in the Chilean case, where opinion polls showed that the public views the media as submissive to external powers, even though internal studies show that journalists do not feel directly pressured. To explain this disconnect, the authors propose a theoretical model that articulates the concepts of framing, cultural industry, and ideological apparatuses, demonstrating how the concentration of media ownership, its dependence on external funding, and the homogenization of content reinforce the sense of external control over

information. This widespread sense of manipulation weakens the legitimacy of the media as a public sphere and fuels public distrust.

Cardoso (2021), analysing Brazil between 2014 and 2018, similarly concludes that the reproduction of polarized and selective narratives by mass media was causally linked to the erosion of public trust in democratic institutions. Taken together, these findings show how structural concentration and political connections not only shape media systems but also fuel a vicious cycle: the less plural and more politically entangled the media landscape, the lower its perceived legitimacy, ultimately weakening the very social function of journalism.

DIGITALIZATION AS A DISRUPTION OF THE MEDIA'S ECONOMIC BASE

The historical context of Latin America, marked by the high concentration of

media ownership and the close ties between conglomerates, political elites, and strategic economic sectors, has acquired a new layer of complexity with the popularization of the internet. In recent decades, two simultaneous transformations have shaken the foundations of regional journalism: profound changes in how the population consumes information, and the massive migration of advertising investments to social media platforms.

These transformations did not occur in isolation. They intertwined and reinforced each other, creating a scenario in which the old dynamics of concentration and editorial control coexist tensely with new forms of fragmentation and technological dependence. The current dilemmas facing Latin American journalism have close ties with how these two changes, in consumption and advertising, have redefined the conditions for the production and circulation of public information in the region.

THE REVOLUTION IN INFORMATION CONSUMPTION HABITS

The way Latin Americans consume news has undergone a radical change. Data from the Reuters Institute (2024) shows that 80% of the Brazilian population now consumes news online, with WhatsApp and YouTube ranking among the primary sources of information. This shift has been enabled by the widespread expansion of internet access: in Brazil, for instance, according to the TIC Domicílios 2024 survey by Cetic.br (2024), 83% of households now have internet, with 85% in urban areas and 74% in rural areas, up from only 13% of urban households in 2005. Increased connectivity has enabled a growing share of the population to rely on online platforms as their primary source of information. This trend is mirrored across the region, where the study “Drivers and Implications of Information Disorder in Latin America” (2025), conducted by InternetLab and Rede Conhecimento Social, shows that 59% of the population uses social media and messaging apps as their regular primary source for news, surpassing by far the 26% who still rely TV or Radio as their primary source.

This digital transformation has generated three structural changes in the region’s information ecosystem. The first is the fragmentation of the audience, which now accesses content in an individualized and disaggregated manner, abandoning collective rituals of news consumption such as the morning newspaper or the evening newscast (Elihu Katz and Jay Blumler, 1974). Now,

each person constructs their own “information diet” across multiple platforms and sources, often without necessarily being aware of the origin of the news (Internetlab and Rede Conhecimento Social, 2022). This fragmentation has eliminated the shared experience that was historically seen as a foundation for public debate, replacing a common agenda with multiple personalized micro-agendas.

The second transformation lies in the deinstitutionalization of news consumption and the erosion of the symbolic value of traditional journalistic brands. For decades, major newspapers functioned as intermediaries between facts and the public, but also as instruments of power tied to local political and economic interests. The media concentration characteristic of the pre-digital period was closely linked to these regional power networks. With digitalization, we observe not only the decline in the legitimizing power of these traditional brands, but also a shift in informational control: from local political elites to transnational technology companies.

This erosion is clear in the levels of trust. In Brazil, Argentina, and Colombia, overall trust in news remains below the average observed in G7 countries (Reuters Institute, 2024). The study “Drivers and Implications of Information Disorder in Latin America” (2025) reveals that, although 68% of respondents report seeking multiple sources of information, the majority place greater trust in those they see as experts than in journalists when it comes to sensitive topics such as politics, public health, or security. In parallel, for checking information, searching engines and influencer channels are preferred over fact-checking websites. What we are observing is a transfer of legitimacy from professional journalists to figures who build authority through proximity and perceived authenticity on social media.

The third structural change is the rise of platform-based intermediation, which replaces editorial judgment with systems designed around engagement and monetization. Whereas mediation was once concentrated in newspaper editors - historically tied to political and economic elites and guided by professional norms such as newsworthiness and public relevance, albeit with their own biases, it is now shaped by transnational corporations whose algorithms reward virality, attention, and commercial incentives rather than journalistic value (Napoli, 2019). This transformation redefines agenda-setting and framing: what gains visibility is no longer determined solely by local elites,

but also by platform logics that amplify content capable of triggering reactions and sustaining user attention. Research on Reddit and other platforms shows that even default settings and algorithmic curation systematically privilege certain voices while marginalizing others, underscoring the profound influence of digital intermediaries over public discourse (Jürgens and Stark, 2017; Stark et al, 2020).

The result is a transformation that goes far beyond a simple technological shift. It represents a reconfiguration of the structures of informational power in Latin America: from concentration solely in the hands of local political and economic elites to a hybrid system where control is also exercised by global corporations. Agenda-setting, framing, and editorial decisions are no longer the exclusive domain of traditional media, influenced by political and economic elites, but also respond to the logics of relevance and virilization on the internet and social media. This shift has profound implications for the quality of information available to support democratic citizenship, raising questions about informational sovereignty and democratic control over the information flows that sustain public debate in the region.

THE MIGRATION OF ADVERTISING: FROM PRINT TO DIGITAL

With audiences shifting to digital and personalized media consumption, a parallel and equally disruptive transformation was taking place in the advertising market, following the movement of the audience. The classic model of news media, sustained by the combination of direct sales and advertising, based on the logic of bundling diverse content into a single product, has collapsed.

Historically, the bundling model has served as an effective strategy for enabling the financing of public interest journalism, as the economic sustainability of content for which readers were less willing to pay was made possible by indirect subsidies from other types of content that held greater commercial value (Rolnik, et al, 2019). Newspapers operated, then, as intermediary platforms in two-sided markets, connecting advertisers and readers through a subsidy structure generated by advertising revenue (Macedo and Moreira, forthcoming).

Digitization disrupted this dynamic by allowing the “unbundling” of news: readers could now access specific content directly, without needing to purchase

the complete newspaper bundle (Macedo and Moreira, forthcoming). At the same time, the advertising revenue that once played an important role for broadcasters and printed media, did not automatically transitioned to digital media. Advertising money has not disappeared; but it became significantly concentrated in digital platforms, with Google and Meta controlling a major share of the digital advertising market in Latin America (Clara, 2025) offering advertisers targeting tools that traditional media outlets cannot replicate. This concentration represents more than a mere technological shift. As noted by Rolnik, et al (2019). The rise of digital platforms has created an asymmetric interdependence between news organizations and technology companies, marked by a complex coexistence in the competition for advertising revenues. Platforms now operate as multilateral markets, connecting users, advertisers, and content producers within the same system of intermediation, whereas traditional news outlets continue to function as two-sided markets that link

“

The rise of digital platforms has created an asymmetric interdependence between news organizations and technology companies, marked by a complex coexistence in the competition for advertising revenues.

”

readers and advertisers. The result is a new form of layered intermediation, in which journalism continues to produce public-interest information, but the access, visibility, and monetization of that information are increasingly governed by the algorithms and business models of digital platforms.

The programmatic advertising system, dominated by these platforms, directs investments to where there is the most significant promise of engagement and clicks, rather than necessarily to where there is the highest quality or informational relevance. As a result, there is “growing competition among the various agents that make up the media ecosystem,” where public interest journalism outlets “see their funding space shrinking, both due to the loss of bundling capacity due to this atomization, and due to the difficulty of funding/demand restrictions on their supply” (Rolnik et al, 2019).

For news outlets, this means adapting or dying. They need to adjust headlines, formats, and even editorial lines to please the algorithms that control reach and monetization. In countries like Brazil, where the average daily time spent on social media is extremely high, this pressure becomes even more intense. The quest for visibility and revenue, tied to opaque and volatile platform metrics, undermines editorial autonomy and jeopardizes the sustainability of independent media outlets. Newspapers are forced to compete not only with

one another, but also with the full spectrum of digital entertainment, in an unequal battle for users' fragmented attention.

In short, digitalization has not only destabilized the traditional economic model of Latin American journalism, based on advertising and content bundling, but has also shifted the centres of informational power beyond the journalistic field and national borders. These transformations demand a careful analysis of their implications for information integrity in the region.

IMPLICATIONS FOR INFORMATION INTEGRITY

This structural reconfiguration of Latin American journalism, characterized by audience fragmentation, the erosion of traditional brands' authority, platform intermediation, and the collapse of conventional funding models, has created an environment highly susceptible to information disorder. As highlighted in 2025, 22 of the 28 countries in the region experienced significant declines in economic indicators for journalism, signalling a historic crisis in the financial stability of media outlets (RSF, 2025). This economic fragility directly affects editorial independence and diminishes the capacity of news organizations to produce quality journalism, transforming the crisis from a purely business issue into a disruption of an already fragile infrastructure that is crucial for democratic information flows in the region.

It is important to clarify the scope of this argument: we are not asserting a direct causal link between these economic shifts and specific disinformation phenomena. Rather, we contend that the economic destabilization of journalism is a critical factor in understanding information integrity in the region. When quality journalism lacks sustainable funding, when editorial choices are increasingly driven by platform metrics, and when producing public-interest information becomes economically unviable, the result is not merely a business crisis but a breakdown in the structures that sustain democratic information circulation. This is especially critical in a region whose information ecosystem was already fragile. The migration of advertising revenue to platforms, coupled with the historical erosion of media credibility due to ownership concentration, generates a context in which reliable information and trust in it is scarce precisely when it is most needed.

This is not to suggest that digitalization is solely detrimental. On the contrary, it has enabled historically marginalized voices and fostered new forms of

independent journalism (Gingras, 2025). Yet these opportunities coexist with profound structural vulnerabilities. The challenge for Latin America is therefore not limited to combating false content or improving platform moderation, nor is it returning to the prior structures of concentrated media. But, to build the economic and institutional foundations necessary to sustain a pluralistic, independent, and financially viable journalism ecosystem, one capable of fulfilling its democratic role in an increasingly complex information environment.

DISPUTED PATHS TOWARDS A MORE PLURAL INFORMATION ECOSYSTEM

PATHWAYS UNDER CONSTRUCTION

In this process of reconstruction, digitally native media outlets have adopted hybrid revenue models that combine subscriptions, individual donations, crowdfunding, international philanthropy, and, to a lesser extent, programmatic advertising.

Digital subscriptions offer a promising alternative by enabling a direct connection between the outlet and its audience, reducing dependence on platforms and advertisers. However, they face structural obstacles in the Latin American context. High socioeconomic inequality limits the number of people both willing and able to pay for informative content. According to the Reuters Digital News Report (Newman, et al, 2025), the proportion of people paying for online news remains relatively low across the region: Brazil (17%), Peru (18%), Colombia (14%), Mexico (14%), Argentina (11%), and Chile (10%). This limited penetration reflects not only economic constraints but also the culture of free access that has been consolidated over recent decades, making it challenging to convert readers into subscribers. Moreover, the implementation of paywalls can restrict access to information, particularly among the most vulnerable groups, thereby deepening information inequalities (UNESCO, 2023). These figures reveal a fundamental tension: while subscriptions may strengthen the financial sustainability and editorial independence of individual outlets, they risk creating a two-tiered information system where quality journalism becomes accessible primarily to those with sufficient economic resources.

Crowdfunding has proven effective in contexts of community mobilization and the cultivation of a sense of belonging, particularly for local media

outlets or those focused on specific causes. One-off campaigns can mobilize resources and engagement, demonstrating the symbolic value of journalism for attentive and organized communities. In Latin America, several initiatives exemplify the characteristics identified by the European Parliament as success cases of crowdfunding (Brogi and Sjøvaag, 2023): diversity according to political, economic, and cultural context; greater success of ad hoc campaigns; association with political and ideological commitments; and the need to function as complementary income to ensure resilience. In Brazil, Agência Pública has raised over R\$231,000 through biennial crowdfunding campaigns to finance investigative reporting, building a growing base of engaged readers. The Intercept Brasil has achieved even greater scale, reaching over 10,000 supporters who contribute approximately R\$300,000 monthly through recurring donations, sustained by continuous communication strategies and audience proximity. In El Salvador, El Faro mobilizes its community around sensitive investigations into violence and corruption, while in Colombia, La Silla Vacía has integrated crowdfunding into its political journalism strategy, emphasizing planning, transparency, and constant communication with supporters (Ronderos, 2016). However, the sustainability of this model remains fragile.

It depends on recurring fundraising cycles, ongoing communication efforts, and a financially stable base of supporters—constraints that hinder long-term planning and the maintenance of robust journalistic structures (UNESCO, 2023). Even successful cases like The Intercept Brasil and Agência Pública function as part of broader hybrid models that combine multiple revenue streams (Ronderos, 2016).

International philanthropy has played a key role in incubating journalism projects focused on investigation, community engagement, and social causes. Foundations such as Ford, Open Society, and Luminate have supported important initiatives that would struggle to find market support. Between 2016 and 2022, 40 independent journalism outlets in Latin America received over \$27 million in funding from the Open Society Foundations (OSF) and the Ford Foundation, according to research by Skoknic (2024). Although the Ford Foundation granted fewer than half the number of grants compared to OSF, its donations were significantly larger, making the total amounts nearly equivalent. OSF has been fundamental in providing establishment funds to new outlets and fostering their initial growth, while Ford tends to

support already consolidated organizations, seeking long-term contributions. Luminate, created by eBay founder Pierre Omidyar, distributed 11 grants among the 40 outlets analysed, with an average donation value substantially higher than the other two foundations, approximately \$360,000 compared to Ford's \$240,000 and OSF's \$135,000 (Skoknic, 2024).

However, this ecosystem is highly concentrated around a small number of major funds, which poses multiple risks. First, external dependence can create vulnerabilities when funding priorities shift or foundations redirect resources

“ While international philanthropy has been essential for supporting public interest journalism that markets alone would not sustain, the concentration of funding sources and the limited development of domestic philanthropic infrastructure remain structural constraints on the resilience of independent journalism in Latin America.

”

the concentration of funding sources and the limited development of domestic philanthropic infrastructure remain structural constraints on the resilience of independent journalism in Latin America.

to other regions or issues. Second, potential misalignment between editorial priorities and institutional agendas may subtly shape what gets investigated and reported, even in the absence of explicit pressure. Third, the scarcity of structured journalistic philanthropy in the region, particularly the lack of local or regional foundations with deep knowledge of national contexts, limits the potential for these initiatives to scale and endure (Sambrook, 2024). The absence of consistent data on grant duration further complicates assessment of long-term sustainability, suggesting that annual analyses may reveal different patterns than multi-year aggregates. While international philanthropy has been essential for supporting public interest journalism that markets alone would not sustain,

LOCAL JOURNALISM: ADDRESSING NEWS DESERTS AND REBUILDING TRUST

The funding models mentioned above whether subscriptions, crowdfunding, or philanthropy, tend to concentrate in urban centres and national outlets with greater visibility and technical capacity. This raises a critical question: what happens to journalism in territories that lack the infrastructure, audience scale, or institutional connections to access these alternative models? It is precisely in these locations: small towns, peripheral urban areas, and marginalized communities, that the information crisis manifests most acutely, and where

local journalism emerges as both a pressing need and a promising response. It is within this context that local journalism has established itself as one of the most relevant responses to the information crisis. The closure of newsrooms in small towns and low-income urban areas has expanded the so-called “news deserts”—regions where there is little or no regular journalistic coverage (Abernathy, 2018). This absence not only undermines access to information but also affects public trust, local accountability, and civic engagement. As the Pew Research Center (2019) highlights, local journalism is perceived as the most trustworthy by a large portion of the population, precisely because it is embedded in everyday life and in decisions that directly affect the community.

The lack of news outlets in these territories is exacerbated by the algorithmic architecture of digital platforms, which prioritizes national-level media with greater technical capacity, marginalizing smaller initiatives even when relevant to their local audiences (Fischer, Jaidka and Lelkes, 2021). This highlights that the crisis in local journalism is not only economic, but also structural and technological. The same digital transformation that enabled new funding models for national outlets has, paradoxically, made it harder for local journalism to survive algorithms favour scale, subscription models require sufficient population density, and philanthropic funding gravitates toward outlets with greater visibility and impact metrics. The absence of quality local journalism creates a vacuum that is quickly filled by other actors. Research on information disorder in Latin America (InternetLab e Rede Conhecimento Social, 2025) reveals that trust in information sources is strongly mediated by proximity: while only 17% of respondents consider journalists among the most trustworthy voices on key issues, 25% trust family members and 53% trust “professionals in the field” or “experts” - categories often filled by influencers and local figures who build authority through perceived closeness rather than journalistic credentials. This pattern becomes more acute in news deserts, where the absence of local reporting leaves communities vulnerable to actors who may lack journalistic training or ethical commitments.

Without credible local sources embedded in community life, residents turn to whatever information is available, often sensationalized content optimized for engagement rather than accuracy, or politically motivated narratives that exploit the information vacuum. Moreover, the same research shows that when people perceive information as relevant to their immediate context, such as public health alerts or local security issues, they are significantly more likely to share it,

with 40% reporting that they regularly forward information considered to be of public utility (InternetLab e Rede Conhecimento Social, 2025).

This underscores a critical dimension of local journalism's value: it produces information that communities recognize as directly relevant to their lives and therefore actively circulate. When this function is absent, the informational needs don't disappear, they are simply met by less reliable sources. In response to this situation, concrete initiatives are emerging across Latin America. In 2023, civil society organizations launched the Alliance for Local Journalism

in Latin America, bringing together efforts from Brazil, Mexico, Colombia, and Argentina. The coalition aims to promote the sustainability of local media outlets and mitigate news deserts through regional coordination, professional training, advocacy, and the development of alternative business models (Atlas da Notícia, 2023). The goal is to build a journalistic ecosystem that is more connected to its territories, more diverse in its voices, and more

resilient in the face of economic and political pressures, recognizing that information integrity at the national level depends fundamentally on the health of journalism at the local level.

“
This underscores a critical dimension of local journalism's value: it produces information that communities recognize as directly relevant to their lives and therefore actively circulate.

”

REGULATORY FRAMEWORKS: REBALANCING POWER BETWEEN PLATFORMS AND JOURNALISM

Beyond alternative funding models and local journalism initiatives, the debate on regulatory frameworks to rebalance the relationship between digital platforms and news outlets is gaining momentum globally and beginning to resonate across Latin America. The Australian experience with the News Media and Digital Platforms Mandatory Bargaining Code, approved in 2021, has served as an international benchmark. The legislation establishes a framework for direct negotiation between platforms and news outlets, with provisions for binding arbitration if agreements cannot be reached, explicitly recognizing that journalistic content generates value for platforms while those same platforms have captured the advertising revenue that once sustained journalism (Australia, 2021). Canada followed with similar legislation in 2023, the Online News Act, which led to agreements between Google and Canadian news organizations worth approximately CAD \$100 million annually.

These international models have sparked significant debate in Latin America about whether and how similar frameworks might address the region's specific challenges. In Brazil, the debate on rebalancing the relationship between platforms and journalism has advanced through multiple legislative and judicial initiatives. Bill 2338/2023 introduces provisions for the remuneration of journalistic material used in training artificial intelligence systems, aiming to compensate for the value of professional reporting in the digital ecosystem (Brazil, 2023). While Bill 4.675/2025 updates Brazil's competition law to address the market power of Big Tech companies, empowering the national antitrust authority (CADE) to designate gatekeeper platforms and impose specific obligations, including mechanisms that can ensure economic responsibilities toward news outlets (Brazil, 2025). Additionally, platform regulation proposals are a recurring topic, often including provisions for platform transparency and accountability, with direct implications for the distribution of journalistic content. Together, these initiatives illustrate Brazil's leading role in Latin America as a laboratory for regulatory experiments that seek to strengthen the sustainability of journalism while balancing platform power and the protection of fundamental rights.

At the same time, these regulatory innovations raise critical questions about implementation and equity. One recurring concern is that mandatory remuneration mechanisms could disproportionately benefit large media conglomerates, the same groups historically associated with media concentration, while leaving independent and local outlets at a disadvantage. Moreover, concerns about government overreach loom large in a region with recent histories of authoritarian media control. Any regulatory framework must navigate the challenge of constraining platform power without creating new mechanisms for state interference in editorial decisions. Alternative approaches are also being discussed. Some proposals focus not on direct remuneration but on requirements for platforms to increase transparency about algorithmic curation, provide fairer distribution to local and independent outlets, or contribute to public funds that support journalism broadly rather than through one-to-one negotiations. Mexico's 2024 proposed reforms to telecommunications law, for example, include provisions requiring platforms to disclose algorithmic criteria and provide preferential treatment to Mexican news sources, though implementation remains uncertain.

These diverse approaches reflect a growing recognition across Latin America that platform power cannot be addressed through market mechanisms alone. However, they also underscore the complexity of designing regulation that genuinely serves democratic pluralism rather than entrenching existing power structures, whether of states, legacy media conglomerates, or transnational platforms. The challenge is to create frameworks that channel resources toward journalism as a public good, with particular attention to supporting independent, local, and diverse voices that have historically been marginalized, while avoiding both state censorship and the consolidation of media power in the hands of a few dominant players.

The pathways explored in this section: hybrid funding models, local journalism initiatives, and regulatory frameworks, illustrate both the creativity and the fragility of responses to journalism's economic crisis in Latin America. Subscriptions, crowdfunding, and philanthropy offer promising alternatives to platform-dependent models but remain difficult to scale in contexts marked by inequality and limited domestic philanthropic infrastructure. Local journalism addresses critical gaps in coverage and trust yet struggles against algorithmic marginalization and resource constraints. Regulatory interventions, such as the Australian bargaining code, propose structural corrections to platform power but raise concerns about implementation and unintended consequences in Latin American institutional contexts.

What these diverse initiatives share is a recognition that journalism's economic sustainability cannot be separated from its democratic function. The challenge is not simply to find new revenue streams, but to build funding models that preserve editorial independence, support diverse voices, and ensure that quality information remains accessible to all, not just those who can pay for it. This requires moving beyond market-based solutions alone and recognizing journalism as infrastructure that demands collective investment, whether through regulation, public funding, or community support.

Yet significant questions remain unresolved. How can alternative funding models reach sufficient scale without reproducing existing inequalities or creating new forms of dependence? How can local journalism survive in territories that lack the population density or economic base to sustain it through subscriptions or advertising? How can regulation rebalance power dynamics without strengthening the hand of either authoritarian states or

entrenched media conglomerates? These questions underscore that the construction of Latin American journalism is not a technical problem with straightforward solutions, but a political and social challenge that will require sustained experimentation, coordination across civil society and policy actors, and ongoing commitment to the principles of plurality, independence, and democratic access to information.

CONCLUSION

This chapter has demonstrated that information disorder in Latin America cannot be adequately understood without accounting for the structural transformation of journalism's economic foundations. By tracing the trajectory from historical media concentration through digital disruption to emerging alternative models, we have shown how the region's information ecosystem has undergone a profound reconfiguration. The shift from oligopolistic control by local political and economic elites to algorithmic intermediation by transnational platforms has not merely changed who controls information flows – it has fundamentally destabilized the economic model that sustained journalism, leaving quality information production precarious precisely when democratic societies need it most. At the same time, this disruption has opened space for new actors, including digitally native outlets, local journalism initiatives, and community-based models that challenge both the old concentration and the new platform dominance. And those new initiatives cannot be overlooked when moving forward.

The challenges revealed by this analysis operate at multiple levels. Economically, the collapse of traditional funding models and the capture of advertising revenue by platforms have left journalism searching for sustainability through hybrid combinations of subscriptions, crowdfunding, and philanthropy. These models show promise but remain fragile and difficult to scale in contexts marked by inequality and limited philanthropic infrastructure. In terms of legitimacy, journalism faces a double bind: historical concentration eroded trust, while current precarity threatens quality, creating a vicious cycle where the institutions meant to sustain democratic information flows lack both credibility and resources. Structurally, the algorithmic prioritization of engagement over accuracy, and of national outlets over local ones, means that market forces alone will not produce the diverse, independent journalism that democracy requires. As international and regional experiments have demonstrated, addressing these challenges demands coordinated action

across civil society, policy frameworks, and new institutional arrangements, recognizing that isolated initiatives, however innovative, cannot substitute for systemic transformation.

The reorganization of the media ecosystem should not aim to restore a past marked by centralization and scarcity, but rather to build a future grounded on plurality, autonomy, and innovation. This means recognizing that the traditional model, despite its economic stability, also carried significant democratic shortcomings: ownership concentration, homogenization of voices; high entry barriers, and excessive reliance on advertising.

A more just and pluralistic information ecosystem must rest on a set of core principles: diversity of voices and perspectives; economic sustainability that

“
The paths to this future are not predetermined and involve legitimate contestation between different visions regarding the role of the market, the State, and civil society in organizing the information ecosystem.
”

preserves editorial independence; democratic access to high-quality information; continuous innovation in formats and methodologies; and transparency in both funding and production processes.

This implies recognizing journalism as an essential infrastructure of democracy, capable of sustaining public debate in plural, complex, and unequal societies. More than adjusting algorithms or regulating content, it's about redefining the economic and political foundations of information in the digital age.

The paths to this future are not predetermined and involve legitimate contestation between different visions regarding the role of the market, the State, and civil society in organizing the information ecosystem. What seems clear is that sustainable solutions will require creative combinations of financing models, well-designed regulatory frameworks, and the strengthening of initiatives that have already demonstrated potential to renew journalism democratically.

Latin America, with its rich cultural diversity and tradition of resistance and social innovation, is well-positioned to lead this reconstruction, provided it can overcome the pitfalls of both conservative nostalgia and technological determinism. The future of journalism in the region will be shaped by the tensions and potential syntheses among these competing forces.

References

- Abernathy, P. M. 2018. *The expanding news desert*. Chapel Hill: Center for Innovation and Sustainability in Local Media, University of North Carolina.
- Australia. 2021. *Treasury Laws Amendment (News Media and Digital Platforms Mandatory Bargaining Code) Act of 2021*. Canberra: Parliament of Australia, [Online]. Available: <https://www.legislation.gov.au/Details/C2021A00021> [2025, July 31].
- Atlas of News. 2023. *Latin American organizations launch the Alliance for Local Journalism: Latin America to reduce news deserts*, [Online]. Available: <https://atlas.jor.br/institucional/organizacoes-latino-americanas-lancam-a-alianca-pelo-jornalismo-local-na-america-latina-para-reduzir-os-desertos-de-noticias/> [2025, July 31].
- Becerra, M. & Mastrini, G. 2019. *Infocommunication concentration in Latin America 2000–2015: new media and technologies, fewer actors*. Montevideo: Observacom [Online]. Available: at: <https://www.observacom.org/wp-content/uploads/2019/09/La-concentracion-de-infocomunicacion-en-Am%C3%A9rica-Latina-2000-2015.pdf> [2025, July 31].
- Blumler J.G. & Katz, E. 1974. *The uses of mass communications: Current perspectives on gratifications research*. Sage Annual Reviews of Communication Research.
- Borges, E., Pavarin, V., Vilela, C. 2024. *The Evolution of Journalism and the Spread of Disinformation in Latin America*. IDRC *Resisting information disorder in the Global South* [Online]. Available: <https://idrc.sun.ac.za/the-evolution-of-journalism-and-the-spread-of-disinformation-in-latin-america/> [2025, July 31].
- Brazil. 2023. *Projeto de Lei N° 2338*. Brasília: Senado Federal, [Online]. Available: <https://tinyurl.com/ycx8a44z> [2025, July 31].
- Brazil. 2025. *Projeto de Lei N° 4675*. Brasília: Câmara dos Deputados, [Online]. Available: <https://www.camara.leg.br/proposicoesWeb/fichadetramitacao?idProposicao=2562481> [2025, July 31].
- Brogi, E. & Sjøvaag, H. 2023. *Good practices for sustainable news media financing*. Committee of Experts on Media Resilience (MSI-RES): Council of Europe report [Online]. Available: <https://www.unav.edu/documents/10174/11264174/informe-sostenibilidad-medios.pdf>. [2025, July 31].
- Cardoso, E. A. S. 2021. *Media, trust and democracy: the role of the press in the weakening of Brazilian institutions (2014–2018)*. Published doctoral dissertation. Porto Alegre: Federal University of Rio Grande do Sul [Online]. Available: <https://lume.ufrgs.br/handle/10183/225478>. [2025, July 30].

Cetic.br & NIC.br. 2025. *Pesquisa sobre o uso das tecnologias de informação e comunicação nos domicílios brasileiros: TIC Domicílios 2024*. São Paulo: Comitê Gestor da Internet no Brasil (CGI.br). Available: <https://cetic.br/pt/tics/domicilios/2024/domicilios/> [2025, July 30].

Clara. 2025. *Clara Ads Report 2025*. São Paulo: Clara, [Online]. Available: <https://www.clara.com/pt-br/conteudos-adicionais/clara-mkt-ads-report-2025> [2025, July 30].

Contado, V. 2025. Brazilians resent the media and government, but trust companies. *Meio & Mensagem* [Online], 13 March. Available: <https://www.meioemensagem.com.br/marketing/brasileiros-se-ressentem-com-a-midia-e-governo-mas-confiam-nas-empresas> [2025, July 30].

Fischer, S., Jaidka, K. & Lelkes, Y. 2021. Auditing local news presence on Google News. *Nature Human Behaviour*, 5: 1465–1475. DOI: 10.1038/s41562-020-00954-0 (no full-stop after DOI).

Friderichsen, C. R. & Lobos, P. M. 2018. The framing of media distrust: a proposal to understand media discredit. *Communication and Media Magazine*, 27 [Online]. Available: <https://go.gale.com/ps/i.do?p=AONE&u=anon~52beb9cc&id=GALE|A717080618&v=2.1>. [2025, July 30].

Gingras, R. 2025, March 20. The evolution of media and democracy: how we got here, how we might move forward. *Richard Gingras' Medium* [Web log post]. Available: <https://richardgingras.medium.com/the-evolution-of-media-and-democracy-how-we-got-here-how-we-might-move-forward-534776e1bde2>. [2025, July 31].

Internetlab & Rede Conhecimento Social. 2025. *Drivers and Implications of Information Disorder in Latin America*. São Paulo: InternetLab; Rede Conhecimento Social [online]. Available: <https://internetlab.org.br/wp-content/uploads/2025/10/Vetores-e-implicacoes-da-Desordem-informacional-na-America-Latina-versao-completa-07.pdf> [2025, July 31].

Internetlab & Rede Conhecimento Social. 2022. *Investigando os vetores de disseminação de conteúdo eleitoral*. São Paulo: InternetLab; Rede Conhecimento Social [online]. Available: <https://internetlab.org.br/en/news/new-research-indicates-that-the-impact-of-information-in-messaging-apps-has-a-deep-connection-with-other-media-habits/> [2025, July 31].

Intervozes, 2023. *Who Controls the Media in Brazil? From Old Oligopolies to Digital Monopolies*. New York: Routledge.

Intervozes; Global Media Registry. 2017. *Media Ownership Monitor Brazil 2017*. São Paulo: Intervozes; Global Media Registry [Online]. Available: <https://brazil.mom->

gmr.org/en/media/ [2025, July 30].

Jürgens, P. & Stark, B. 2017. The Power of Default on Reddit: A General Model to Measure the Influence of Information Intermediaries. *Policy & Internet*, 9(4): 395-419. DOI: <https://doi.org/10.1002/poi3.166>

Macedo, B. G. & Moreira, M. Digital transformation and the economic sustainability dynamics of public interest journalism: considerations on the provision of social interest journalistic content based on economic theory and on case studies of Argentina, Brazil and Colombia. [forthcoming]. Collaboration between InternetLab and LCA Consultoria. Forthcoming in the journal *Internet & Sociedade*.

Napoli, P. M. 2019. Social media and the public interest: media regulation in the disinformation age [Press release]. New York City: Columbia university press. July 30.

Newman, N., Ross Arguedas, A., Robertson, C. T., Nielsen, R. K. & Fletcher, R. 2025. *Digital News Report 2025*. Oxford: Reuters Institute for the Study of Journalism of University of Oxford [Online]. Available: <https://static.poder360.com.br/2025/06/Reuters-Institute-Digital-News-Report-2025-1.pdf> [2025, July 30].

Pew Research Center. 2019. *For local news, Americans embrace digital but still want a strong community connection*. [Washington, D.C.]: Pew Research Center. [Online]. Available: <https://www.pewresearch.org/journalism/2019/03/26/for-local-news-americans-embrace-digital-but-still-want-strong-community-connection/> [2025, July 30].

Reporters Without Borders. 2025. *World Press Freedom Index 2025: over half the world's population in red zones*. Reporters Without Borders [Online]. Available: <https://rsf.org/en/world-press-freedom-index-2025-over-half-worlds-population-red-zones> [2025, July 30].

Reuters Institute. 2023. *Digital News Report 2023*. Oxford: University of Oxford [Online] Available: <https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2023> [2025, July 30].

Reuters Institute. 2024. *Digital News Report 2024*. Oxford: University of Oxford, [Online]. Available: <https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2024> [2025, July 30].

Rolnik, G., Cage, J., Gans, J., Goodman, E. & Pickard, V. 2019. *Protecting journalism in the age of digital platforms*. Oxford: Reuters Institute, [Online]. Available: https://sciencespo.hal.science/hal-03947806/file/2019_gans_cage_goodman_et_al_protecting_journalism_in_the_age_of_digital_platforms.pdf [2025, July 30].

Ronderos, M. T. 2016. Why crowdfunding can keep journalism true to its promise.

Global Investigative Journalism Network [Online], 4 March . Available: <https://gijn.org/stories/why-crowd-funding-can-keep-journalism-true-to-its-promise> [2025, July 30].

Sambrook, R. 2024. *Sustaining journalism in democracy: from crisis to possibility*. Oxford: Reuters Institute, [Online]. Available: <https://reutersinstitute.politics.ox.ac.uk/sustaining-journalism-democracy> [2025, July 30].

Skoknic, F. 2024. *The looming threat (and possible solutions) to Latin America's foreign donor-funded journalism*. Oxford: Reuters Institute, [Online]. Available: <https://reutersinstitute.politics.ox.ac.uk/looming-threat-and-possible-solutions-latin-americas-foreign-donor-funded-journalism#header--2> [2025, July 30].

Stark, B., & Stegmann, D. 2020. *Are Algorithms a Threat to Democracy? The Rise of Intermediaries: A Challenge for Public Discourse*. [Berlin]: AW AlgorithmWatch GmbH [Online]. Available: <https://algorithmwatch.org/en/wp-content/uploads/2020/05/Governing-Platforms-communications-study-Stark-May-2020-AlgorithmWatch.pdf> [2025, July 30].

United Nations Educational, Scientific and Cultural Organization. 2023. *Finding the funds for journalism to thrive: policy options to support media viability*. Paris: UNESCO, Report [Online. Available: <https://unesdoc.unesco.org/ark:/48223/pf0000386247> [2025, July 30].

Valente, J. 2013. *Regulação democrática dos meios de comunicação [Democratic regulation of the media]*. São Paulo: Fundação Perseu Abramo (Coleção O que saber, 2) [Online]. Available: https://siac.fpabramo.org.br/uploads/acervo/FPA_ColecaoOQS_LVR_0007.pdf [2025, July 30].



BEYOND CONTENT MODERATION:

**A Political Economy Of
Information Disorders And
Regulatory Pathways**

Scott Timcke

Zara Schroeder



INTRODUCTION

Information disorder manifest across many facets of life, and are most threatening when they intersect with politics, electoral integrity, discrimination and undesirable norms that reproduce illegitimate, unjustified inequalities of power. From our perspective, the main objective of coordinated information disorder is to disempower people, robbing them of the key information they need to navigate their worlds, or pushing them out of venues where they otherwise have the right (or should have the right) to participate. Essentially, information disorder are phenomena that undermine democratic agency. This chapter explores how these issues appear in the Global South considering the abusive exploitation of the political economy of information disorder on platforms. We also cover selected regulatory elements that can be used to better curtail this abuse.

Digital technologies are fundamental to most societies. However, their current trajectory promotes economic concentration and monopolies, which is generally detrimental to the public good. Moreover, these monopolies provide mechanisms and vectors that coordinated campaigns can abuse to promote information disorder. Given the significant expense, importance, and ecological consequences of platforms, government intervention is necessary for a different approach. We define platforms not just as technological entities but also inclusive of user practices and collective behaviours, best understood from a sociocultural perspective (Jordan, 2023).

Some of this policy intervention can be undertaken at the national level, given that the state is responsible for governing markets. Even so, being mindful of the transnational character of platforms requires creating independent, multilateral organizations and social pacts to either govern or provide vital digital infrastructure and services. The Global Digital Compact is an example of the former, while the International Telecommunication Union is an example of a multilateral organization that provides vital digital infrastructure and services like coordinating global radio spectrum allocation and satellite orbits. Another option to consider curtailing opportunities for information disorder is the development of public utilities, digital public goods, digital public infrastructure or other shared resources to advance development through global cooperation.

There is overwhelming scholarly evidence that the typical business model of Very Large Online Platforms prioritizes data extraction and engagement over user well-being, concentrating immense power in a few hands (United Nations, 2024; Research ICT Africa, 2023; Timcke, 2024, Timcke and Schroeder, 2024). An example of this is Facebook’s algorithm which is designed to maximize user engagement by promoting content that elicits strong emotional reactions, often outrage, fear, or anger (Washington Post, 2021). Internal research (e.g., the Facebook Files leaked by whistleblower Frances Haugen) revealed that the platform knowingly amplified divisive or misleading content because it increased time spent on the platform, despite potential harms to mental health, public trust, and democratic discourse (Hao, 2021). This approach prioritizes engagement metrics and ad revenue over user well-being.

Another example of this is how TikTok’s current business model relies heavily on collecting extensive user data to feed its recommendation algorithm (Hollyland, 2025). This results in a highly addictive user experience designed to keep users scrolling for as long as possible. The platform has faced scrutiny for its opaque data practices and its centralization of control over content moderation and algorithm design, which further entrenches its power over information flow (Chen, 2022).

Transforming the business models of Very Large Online Platforms requires coordinated intervention on multiple fronts. These include restructuring incentive systems, establishing enforceable minimum standards that

“ Transforming the business models of Very Large Online Platforms requires coordinated intervention on multiple fronts. ”

place human rights above commercial interests, and implementing participatory governance mechanisms that empower users through media literacy and democratic platform oversight. These actions can close the mechanisms and vectors that coordinated campaigns abuse to promote information disorder agendas, shape public narratives and suppress dissent.

Publishers, platforms, and policymakers have shared, albeit differently orientated, responsibilities to counter and prevent information disorder. Collectively, their actions have pursued content moderation cognizant of the protection of expression, independent journalism and human rights, to

draw attention to a few specific areas. However, there remains a need to more explicitly address the role of platforms not merely as technological intermediaries but as political and economic actors with substantial power over information flows. Any effective approach requires navigation of trade-offs to ensure governments, private companies, or powerful groups cannot effectively have unilateral control of public information or become ‘arbiters of truth.’ Heeding the transnational character of the problem of information disorder, national legislative frameworks can nevertheless empower their independent regulatory bodies with transparent oversight mechanisms to thwart the abuse of platforms, their business models, or algorithmic designs and operation.

In addition to general problems of information disorder, there are some special attributes that emerge in the Global South. Many Very Large Online Platforms headquartered in the Global North are reliant on labour based in the Global South. This dependency presents a potential point of leverage, particularly if regulatory efforts are coordinated in such a way that prevents Very Large Online Platforms from easily relocating outsourced operations across jurisdictions in response to emerging regulations. Due to limited institutional capacity and resources, many Global South countries, including middle-income countries, can struggle to identify emerging information disorder. Detection challenges and the lack of concrete data often lead regulators to depend on evidence collected in other jurisdictions.

In our view, legislation and regulation should ideally be based on a sound understanding of the political economy of platforms, given that these entities have become the load-bearing infrastructure for everyday life. By largely controlling information flows, platforms’ governance structures, market organization, incentives, fiduciary obligations, and ownership motivations are significant factors in shaping the extent to which information disorder affect everyday life, including the social reshaping of society, perceptions to the outside world and potential economic instability of a region. Policy approaches must therefore account for the market dynamics, network effects, and power relationships inherent in platform economies, ensuring that regulatory interventions strengthen the structural foundations on which millions now depend for their social and economic activities.

THE POLITICAL ECONOMY OF INFORMATION DISORDER ON PLATFORMS

Over the course of nearly three decades the internet has transformed contentious politics by enabling decentralized mobilization, cross-border solidarity, and real-time coordination (Timcke et al, 2024). From anti-authoritarian protests to climate justice movements, online platforms serve as ‘critical infrastructure’ for organizing, storytelling, and amplifying dissent. Critical infrastructure are vital assets and systems that are of a physical or virtual nature, the incapacity or destruction of such systems and assets would have a debilitating impact on security, national economic security, national public health or safety (Harašta, 2018). In the Global South, digital activism often emerges in response to structural injustice, state repression, or exclusion from mainstream media. The platforms that empower movements can also constrain them, through algorithmic demotion, surveillance, or content takedowns (Feldstein, 2021).

It is important to recognize that platforms do not only impose constraints; they can create opportunities and affordances for adversaries of such movements, actively amplifying the attack surfaces through which these movements can be targeted, undermined, or co-opted. While online spaces offer new tools for resistance, they are also deeply entangled with the interests of states and corporations (Timcke and Schroeder, 2024). Understanding these competing interests is required to analyse and understand how contentious politics operates in digitally mediated environments as well as understanding the possible sites for policy interventions, given the social organisation of any given society. However, it is important to note that the implementation of policy interventions would be feasible in a state where platforms would promote democracy of information, however, it would not work in other countries, for example.

The interplay between weak state actors and strong non-state actors in the arena of platform governance has also given rise to significant imbalances in technological power and informational control. This creates a governance landscape where conventional regulatory approaches are no longer effective. For example, Silicon Valley’s influence on the Global South has reshaped digital economies, governance, and information ecosystems, often reinforcing neo-colonial dynamics (Timcke, 2024). Big Tech giants like Meta, Google, and X export platform models that do not account for the Global South’s

sociopolitical contexts, leading to the unchecked spread of misinformation and disinformation, surveillance, and data extraction (Obi, 2024).

In Latin America, a comparative study of Argentina, Brazil and Colombia shows that digital platforms have not only disrupted the informational ecosystem but also contributed to a structural reorganization of the information production market in the region. By introducing new logics of distribution, consumption and advertising revenue, these platforms have reshaped the relationships between content producers and audiences, intensifying power concentration and the challenges faced by public interest journalism. In media environments historically marked by high concentration and resource asymmetries, such as those in the Global South, this transformation has deepened the exclusion of independent and community-based voices. In Brazil, for instance, data from the Association of Digital Journalism (AJOR), which brings together more than 110 native digital news organizations, indicating a high dependency on search platforms for traffic and on philanthropic foundations for funding. This reflects how platformization imposes new structural and financial challenges on already fragile media ecosystems.

Facebook's delayed response to hate speech in Ethiopia during the Tigray conflict, for example, exacerbated real-world violence, highlighting the harms of algorithmic governance without local accountability (Muia, 2023). These platforms profit from user data while offering limited reinvestment or content moderation resources (Cipesa, 2025). Moreover, infrastructure projects like Google's Equiano cable or Facebook's 2Africa initiative, while expanding connectivity, risk deepening dependencies on foreign-owned digital backbones (Francois and George, 2019; Munga, 2025). In response, organizations like Paradigm Initiative in Nigeria and Mozilla Africa Mradi advocate for digital rights, data sovereignty, and platform accountability (Paradigm Initiative, 2024; Mozilla Foundation, 2022). In the African case, advocates argue that resisting information disorder requires governance models and technology built for, with, and by the continent's diverse populations (Timcke, Orembo and Hlomani, 2023; Brown and Marsden, 2023).

Digital inequality simultaneously drives and results from economic underdevelopment in the Global South, where structural imbalances in global trade and limited technological infrastructure have and continue to hinder

equitable growth. As the Global South confronts the dual challenge of technological advancement and resisting dependency-reinforcing models, for example, addressing this divide requires not only increased investment but decolonized approaches to innovation that prioritize local agency, sustainable development, and participation in shaping global digital futures. Investment is limited by the cost of acquiring capital, national indebtedness, and weakened local capital markets.

The uneven exchange and export of technological infrastructure from the Global North and China to the Global South often reinforces dependency. Western and Chinese corporations install data centers, undersea cables and

“
The uneven exchange and export
of technological infrastructure
from the Global North and China to
the Global South often reinforces
dependency.

”

cloud services across the African continent, like Google’s Equiano cable or Amazon Web Services’ data hubs in South Africa, while retaining ownership and control of critical digital infrastructure (Francois and George, 2019; Amazon Web Services, 2020). Other dependencies include advertisers who are often compelled to use dominant gatekeeping platforms that extract rents from the buyers and sellers of ads and frequently engage in self-preferencing. Especially where payment for services is enforced in USD, these

dynamics are intensified. Following the Nigerian revolt in Lagos, AWS suddenly began accepting payments in Naira, highlighting how political and economic resistance can pressure even dominant tech firms to localize operations or modify financial practices (Dosunmu, 2025; Ashiru, 2025).

Among a series of other mechanisms, long-term lease and operating contracts typically ensure operational control remains with the foreign corporation for decades, for example. These kinds of contracts limit the Global South’s ability to define its digital future and subject it to foreign surveillance, economic extraction and policy influence. By contrast, homegrown efforts like Nigeria’s MainOne cable and Senegal’s national data center represent steps toward digital self-determination (MainOne, 2010; Diarniadio National Datacenter, 2021). Investing in locally owned infrastructure and regional data governance frameworks is essential to break free from the colonial logic of imported technology. As difficult as it may be, the Global South must seek to build, control and govern digital tools on its own terms.

THE POLITICAL ECONOMY OF CONTENT CREATION

The political economy of content creation examines how digital content is produced, monetized, and controlled within capitalist systems. It encompasses the production of information landscapes refers to how narratives, knowledge, and content are shaped, circulated, and prioritized across digital platforms. In Africa, these landscapes are increasingly influenced by external actors, tech companies, foreign governments, and global media, which frame how people access and interpret information.

Platforms like YouTube, TikTok, and Instagram incentivize content creation through advertising revenue, attention economies, and algorithmic reward structures. In the Global South, creators often enter these economies under precarious conditions, facing unequal access to monetization, censorship, and lack of platform support (Duffy et al., 2021; Arora, 2024; Duffy and Meisner, 2023). Global digital capitalism extracts value from the labour of content creators while retaining control over visibility, circulation and rates of payment. Content creators must navigate not only audience expectations but also opaque algorithms and platform governance structures that increasingly privilege virality and engagement metrics over other forms of value (Timcke and Schroeder, 2024).

The dynamic reinforces the commodification of cultural expression; whereby creative content is disciplined by market logic and platform imperatives. Consequently, this logic does not only affect individual creators but also disrupts broader media ecosystems. Traditional creative and media agencies, once central intermediaries in content production, branding, and marketing, are being dis-intermediated. For example, platforms like Meta are beginning to offer generative AI tools and in-house marketing solutions that effectively bypass the need for and expertise of external creative agencies. This shift further consolidates platform power, extending from distribution to content creation and strategic planning, at the expense of longstanding media and advertising infrastructures skilled at cultural interpolation of western or foreign concepts, for local market context.

Content production suffers from digital infrastructure disparities and platform governance favouring Global North creators, even for users in the Global South. Intentional algorithmic bias limits the visibility of African content unless it aligns with global trends and a platform's judgement

of ‘good content’, as South African TikTok creators experience with culturally specific videos receiving fewer views than Western-aesthetic English content (Malinga, 2025). Unequal access to monetization tools like the YouTube Partner Program and the TikTok Creator Fund further disadvantages African creators, creating digital dependency where their creativity requires non-African systems for legitimacy and profit (Matthew, 2021). Meanwhile, Facebook’s prioritization of engagement has spread hate speech in Ethiopia and electoral disinformation in Nigeria, while inadequate language recognition leaves communities vulnerable to both over-censorship and neglect. Platforms prioritize sensational and emotive content, as seen during Uganda’s 2021 elections, where coordinated disinformation campaigns dominated while opposition voices were suppressed (Nanfuka, 2021).

Limited local content production capacity means many African digital spaces are saturated with imported narratives, resulting in epistemic dependency and cultural dislocation. African user data, extracted without meaningful consent, fuels engagement strategies while returning little benefit to communities. Solutions include emerging platforms like Mkito and Boomplay, decentralized African-owned platforms, regional creator funds, initiatives like Africa No Filter, community media projects, Nigeria’s Dubawa fact-checking network, and the AU’s Digital Transformation Strategy. To reclaim narrative power, Africa needs homegrown algorithmic research, regional AI ethics frameworks, platform accountability, stronger local media ecosystems, digital and AI literacy, fair labour conditions, algorithmic transparency, and equitable platform governance (Schroeder et al, 2025).

VALUE EXTRACTION AND PLATFORM DECAY

AA socio-technical approach to political-economic analysis recognizes that digital platforms are shaped by market forces, institutional regulations, their technical architectures and end user interactions (Timcke, 2021). Platforms are not neutral intermediaries; they embed power asymmetries through design choices, algorithms, and data policies (Timcke and Schroeder, 2024). An effective critique must, therefore, integrate the material infrastructures of technology, such as code, data flows, accessibility and end user interfaces, with broader political and economic systems. This enables governance strategies that address both structural exploitation and the technical means through which it is enacted.

Platform decay is a special kind of abuse of market dominance that leverages prior, sustained investment into that model of transaction costs (Timcke and Rens, 2024). Put differently, when business users invest in platform-specific relationships, the platform can exploit these sunk costs during a shift in strategy, often called the ‘treacherous turn’ (Bloodstein, 2019; Doctorow, 2024). This behaviour distorts market efficiency and can lead to poor capital allocation across the broader economy.

One reason dominant platforms maintain their market power is the high switching costs and inconvenience end users face. Moving to a new platform often means losing social connections and having to manually transfer data and content, an inconvenience that effectively locks end users in. This lock-in discourages competition, facilitates price gouging, and stifles the emergence of new players over time. To reinforce this effect, platforms use various strategies. One common tactic is the use of long, complex “clickwrap” agreements, or dense legal documents, containing convoluted technical terms that most end users do not fully read or understand. End users accept extensive terms and conditions that the platform can later change unilaterally.

To prevent such decay, platforms should not be allowed to unilaterally alter their terms of service (Khan, 2016). Any changes should require the informed consent of end users, with oversight from national competition or information regulators. Regulators should also have the authority to impose a ‘terms freeze’, preventing or reversing changes until platform decay is demonstrably halted or the platform is dissolved (Timcke and Rens, 2024). This freeze would lift only once the platform stabilizes and harmful behaviour ceases. Such safeguards would discourage opportunistic conduct that undermines trust, while still permitting pro-competitive platform evolution, if changes are transparent and accepted by the market. In the Global South, platform decay disproportionately harms smaller businesses and local knowledge actors, who lack bargaining power or technical resilience. Protecting users and businesses from exploitation in its many forms can support a healthier digital ecosystem.

POWER DYNAMICS, REPRESENTATION AND ALGORITHMIC NEO-COLONIALISM

Another productive way to view Very Large Online Platforms is as epistemological infrastructures. Epistemological infrastructures refer to the systems and norms that govern how knowledge is produced, validated, and

circulated within society (Carlson, 2020). These infrastructures are shaped by dominant platforms whose algorithms and moderation policies determine visibility and de facto credibility. In the Global South, these infrastructures are often fragile, shaped by historical inequalities and limited institutional capacity (Gorwa, Binns and Katzenbach, 2020). When platforms engage in exploitative practices, such as unilateral policy changes or algorithmic manipulation, they distort local knowledge ecosystems, privileging certain voices while silencing others (Treré, 2016).

This inequality can manifest in biased search results, discriminatory content moderation, and unequal algorithmic amplification (Gorwa, Binns, and

“

These technologies do not simply reflect the world, they actively construct it, shaping public perception and political discourse.

”

Katzenbach, 2020). In the Global South, colonial narratives are often replicated through platform design, data labelling, and AI training sets dominated by Western norms (Timcke and Schroeder, 2024).

These technologies do not simply reflect the world, they actively construct it, shaping public perception and political discourse. Control over representation thus becomes a form of power, determining who is seen, heard, and deemed credible. Strengthening epistemological infrastructures requires

democratizing access to digital tools, supporting pluralistic media, and embedding local knowledge systems into platform governance. Regulatory responses must, therefore, consider not just content moderation, but the deeper question of who decides what counts as truth in a global, digital economy.

The politics of exclusion in knowledge platforms can manifest through algorithmic gatekeeping, content moderation biases, enabled intimidation of vulnerable groups (such as women journalists) and design choices that prioritize dominant cultural and linguistic norms (Yu, Bekerian and Osback, 2024). For example, Wikipedia’s notability criteria often disqualify entries about African scholars, oral histories, or community leaders, due to a lack of citations from “reliable” sources, usually defined through Western academic or media institutions (Ovesen, 2014; Ford, 2011). This can result in epistemic erasure, where valuable local knowledge is rendered invisible. Platform governance rarely includes Global South actors, leading to a lack of contextual sensitivity in moderation policies or algorithmic prioritization

(Gorwa, Binns and Katzenbach, 2020). To counteract this, initiatives like Whose Knowledge? (n.d.) work to reclaim space for marginalized voices and democratize representation in global knowledge systems. Resisting exclusion requires centering diverse epistemologies and challenging the metrics that determine what counts as credible or worthy of preservation online.

Epistemological violence in digital ecosystems occurs when certain ways of knowing, particularly those rooted in indigenous, feminist, or non-Western traditions, are systematically erased, distorted, or delegitimized through biased algorithms, culturally insensitive content moderation, and colonial-like information architectures (Gorwa, Binns and Katzenbach, 2020). This digital power imbalance not only determines what (and whose) knowledge becomes visible, searchable, and legitimate but also requires systemic reforms in platform governance and data justice to restore narrative agency and embrace more pluralistic understandings of truth.

Epistemic minorities, that is communities whose knowledge systems face structural marginalization, can be harmed by algorithmic systems that suppress indigenous languages, dismiss oral traditions, and misclassify cultural content as misinformation. When platforms like YouTube demonetize indigenous-language videos or when data voids in less digitized regions allow local narratives to be overwhelmed by dominant discourses, these communities experience both silencing and delegitimization of their knowledge practices (Makananise and Madima, 2025). Projects like Masakhane (2019) which supports natural language processing for African languages, represent necessary counter-efforts. Even so, broader structural shifts in algorithm design and policy remain essential.

Arguably algorithmic neo-colonialism imposes Western-centric knowledge systems by embedding dominant values and taxonomies into digital infrastructures, determining how information is organized, discovered, and legitimized (Muldoon, 2023). This manifests when search engines prioritize Euro-American sources for topics on indigenous traditions or when machine learning models trained primarily on Western datasets perpetuate biases at scale. Resistance requires both technological and political will and strategies, developing indigenous AI frameworks, establishing local data commons, and implementing inclusive design processes that centre local ontologies to challenge the epistemic violence of Western universalism.

Moreover, these imported models tend to crowd out indigenous innovation ecosystems. Resistance is emerging through African tech startups like SafeBoda (2017) in Uganda, which adapt platform models to local conditions, and policy initiatives pushing for data sovereignty and fair digital taxation. Resistance requires building platforms, policies, and practices that recognize diverse ways of knowing. Efforts like AfroCuration, the Masakhane NLP project, and community-driven Wikimedia edit-a-thons promote African-led content creation in local languages (AfroCuration, 2019; Masakhane, 2019; and Newman, 2019). To better resist information disorder, the Global South must reclaim epistemological space to ensure cultural relevance in digital spaces and to shape global knowledge ecosystems on its own terms.

REGULATORY RESPONSES

Government regulation is profit-shaping. It seeks to alter economic incentives to modify corporate platform behaviour in service of public interest objectives that markets alone may not prioritize. However, the multi-layered nature of digital ecosystems often outpaces the ability of legal frameworks to adapt to rapid technological change. Regulatory actions must navigate jurisdictional limits, international technology networks, and the quickly changing algorithms that can instantly spread misinformation globally. These interdependencies render policy approaches that are used in other sectors as ill-suited to platform governance; regulators must be willing to and capable of responding to technological shifts with exceptional speed and accuracy.

The changing nature of information disorder makes them difficult to address with existing disinformation legislation. Current regulatory frameworks, which are often reactive and not equipped to handle the complexities of contemporary digital misinformation. Effective regulation necessitates a shift away from simplistic solutions like content removal or platform moderation, and towards more sophisticated approaches that tackle the root causes of information manipulation, such as algorithmic and economic incentives. This requires the development of legislative tools that can thoroughly evaluate technological architectures, establish dynamic accountability mechanisms, and set nuanced standards for information quality and platform responsibility. Central to this task is understanding the increasingly pivotal role of AI in shaping platform dynamics. From algorithmic content curation to automated moderation and personalized advertising, AI systems are embedded in the core functions of platforms, making them not just tools, but active agents in

influencing public discourse, user behaviour, and information visibility. As such, any regulatory approach must directly engage with how AI technologies structure power, labour, and knowledge within digital ecosystems.

The ability of governments to effectively address information disorder can be severely limited by a lack of technological expertise, computational resources, and adaptability. This creates a critical vulnerability to sophisticated information manipulation, especially as many state institutions rely heavily on private technology companies to understand the complexities of the digital landscape. To counter these threats, significant investment in public sector digital literacy, interdisciplinary research, and state-level digital intelligence infrastructure is essential.

Initiatives designed to augment digital and data literacy can empower the public to exercise greater control over data systems. Similarly, programs that empower worker advocacy and AI co-governance between labour and management can cultivate a more equitable distribution of data power. Labor unions and worker organizations can serve as critical counterweights to managerial control in AI governance, ensuring that systems reflect workplace realities and uphold workers' rights. Embedding labour as a co-governor in algorithmic oversight not only enhances transparency but also democratizes decision-making in data-driven environments.

Novel international legal AI regimes, treaties, and global data governance institutions can assist in mitigating asymmetries. However, their legitimacy is contingent upon multi-stakeholder participation, resistance to corporate appropriation, and the empowerment of the Global South. Lastly, open data infrastructure, federated systems, and the open sourcing of training data and models all represent pathways to decentralization and can counteract data concentration.

RIGHTS EXTENSION

From our perspective, there are two key objectives of platform regulation. The first is to preserve and extend fundamental rights in a democratic state. Second, national digital, media, information, data, cloud and AI policies can be aligned to foster pluralism by governing content curation and recommender systems. In parallel, it is also necessary to address how generative AI and AI agents reshape the informational ecosystem, not primarily through platform

distribution mechanisms, but through upstream production dynamics with significant downstream impacts. These technologies raise acute copyright and authorship concerns, as their outputs often constitute direct substitutes for the creative and journalistic work they are trained on, thus intensifying the challenges to media pluralism and content originality.

This involves crafting policies that require or incentivize platforms to prioritize content diversity, transparency, and equitable representation, while respecting freedom of expression, individual privacy, and editorial independence. There is value in these policies being technologically agnostic and forward-looking to anticipate emerging digital communication paradigms and algorithmic biases that could undermine information diversity and democratic discourse.

ANTICIPATORY AND ADAPTIVE REGULATORY FRAMEWORKS

As the Global South navigates digital transformation, the development of forward-looking, context-sensitive policy frameworks is essential to resist information disorder and build resilient digital systems. Regulatory approaches must go beyond reactive measures to anticipate future technological shifts, mitigate risks, and foster equitable innovation. With many African states still building digital regulatory capacity, for example, anticipatory and adaptive frameworks provide a pathway toward sovereignty, accountability, and inclusive governance in the face of Big Tech influence (African Union, 2020; Continental Artificial Intelligence Strategy, 2024). These strategies can help ensure that the digital transition serves public interest goals and strengthens democratic and economic institutions.

Anticipatory frameworks are policy tools designed to forecast and proactively manage the societal impacts of emerging technologies. For African countries, such frameworks are crucial in contexts where rapid digital adoption outpaces regulatory preparedness. Anticipatory governance emphasizes foresight, multi-stakeholder participation, and flexible policy design. Rwanda's AI Policy and Kenya's Digital Economy Blueprint are examples of African states engaging in long-term digital strategy-making that includes ethical considerations, infrastructure planning, and workforce development (Republic of Rwanda, Ministry of ICT and Innovation, 2023; Republic of Kenya, 2019). Additionally, Ghana's Data Protection Commission employs ongoing consultation mechanisms with businesses and civil society to update data

governance rules (Ghanaian Government, Ministry of Communications, 2014). Similarly, the African Union's (2022) policy on cybersecurity and personal data protection promotes a continental framework that can be adapted by individual member states. The downside is that this is not enforceable, meaning member states are encouraged to adapt and adopt but do not have to - thus leading to the existing issues in countries such as Ethiopia, for instance. This flexibility is essential in addressing information disorder, where new forms of manipulation or platform behaviour are constantly emerging.

The effectiveness of anticipatory and adaptive regulatory frameworks ultimately hinges not only on their design but on whether their intended outcomes are realized in practice. Such frameworks are generally implemented using a principles-based approach, which allows for greater flexibility and responsiveness than strictly rule-based systems. This requires systematic, independent research to evaluate whether the expectations set by such frameworks are being met. Essentially, this form of evaluative inquiry is distinct from regulatory monitoring or compliance checks; it involves a deeper assessment of impact, legitimacy, and responsiveness over time which can be achieved through research..

TRANSPARENCY, ACCOUNTABILITY, MONITORING AND EVALUATION

Ensuring transparency around the metrics and algorithms that determine access to core platform infrastructure must be a regulatory priority. To balance this approach, genuine data portability and strong adversarial interoperability standards are essential to counterbalance platform power. Beyond technical measures, regulators must also confront how platforms systematically optimize for investor returns, often at the expense of users and workers. This tension arises because, by sidestepping democratic oversight, platforms exploit regulatory gaps, which deepens existing inequities across the digital ecosystem.

Equally important is pricing transparency from the supplier's standpoint. When platforms implement sudden shifts in incentives or access terms, especially when these changes are designed to entrench market dominance, they can harm innovation and erode trust throughout the value chain. Such extractive tactics, while possibly profitable in the short term, are ultimately self-defeating as they undermine the very ecosystem upon which platforms depend.

Effective policy requires clear tools for measuring the impact of digital interventions, especially in fast-evolving information ecosystems. In Africa, many digital policies can fail to include robust monitoring and evaluation systems, resulting in limited feedback loops and poor policy adaptation (Mapitsa and Churchill, 2024). Measuring interventions in the fight against disinformation or digital exclusion means going beyond surface-level metrics like internet penetration, and instead assessing indicators such as digital literacy, misinformation prevalence, and access to trustworthy content. South Africa’s Information Regulator and Nigeria’s Digital Rights and Freedom Bill represent steps toward institutionalizing metrics that assess user rights and platform accountability.

“ Enforcing digital services policies across Africa faces significant hurdles, as regulatory bodies often lack the technical expertise, resources, and legal authority to hold global platforms accountable. ”

Enforcing digital services policies across Africa faces significant hurdles, as regulatory bodies often lack the technical expertise, resources, and legal authority to hold global platforms accountable. Ghana’s struggles with platform misinformation and the politicized application of regulations demonstrate how enforcement can be inconsistent and ineffective (Kwode and Selekane, 2023). The transnational nature of digital services, encrypted content, and complex platform architectures further complicate accountability efforts. Effective enforcement requires a multi-stakeholder approach combining governmental oversight with platform self-regulation, independent auditing, and regional cooperation. This balanced strategy must both penalize harmful practices and incentivize positive behaviours through constructive engagement, while building digital forensic capacity and ensuring regulatory independence from political interference.

INTEROPERABILITY AND DATA PORTABILITY

Interoperability and data portability can disrupt platform monopolies and empower users (Bietti, 2020). By enabling users to transfer their data and social connections across services, regulators can challenge the current ecosystem of digital information control. However, implementing these principles presents technical challenges.

Additionally, adversarial interoperability offers a potential remedy grounded

in core market principles. This approach, where companies reverse-engineer compatibility with dominant platforms without needing their permission, has historically driven innovation, even in monopolized markets (Doctorow, 2019). Tech giants like Apple, Google, and Meta used this tactic to challenge incumbents by offering better products that worked within existing networks. In this way, a dominant position became a vulnerability that others could leverage. Applied to platforms, adversarial interoperability would give users and businesses the power to leave a platform and transition to alternatives more easily when they are dissatisfied, thereby restoring competition and consumer choice.

Interoperability can combat information disorder by decentralizing information control. Easy user migration between platforms creates market pressure for higher information quality over engagement, discouraging the algorithmic amplification of misinformation. Reduced platform lock-in would foster a diverse ecosystem with varied content governance, allowing users to choose environments aligning with their values on information integrity and transparency.

For example, in the US Department of Justice's antitrust case against Google for the illegal monopolization of the online search market, one proposed remedy was the divestiture of Google Chrome and the prohibition of revenue-sharing agreements with rival browsers (Cooper, 2025). While aimed at reducing Google's dominance, this proposal raised concerns among web developers and open web advocates, the end users who feared that such a move could undermine the financial support Google currently provides to key web infrastructure and open-source projects, effectively draining the fund that supports the web's ongoing development and standardization.

Google, through Chrome, has played a central role in shaping web standards and funding browser innovation, albeit while simultaneously extracting user data and compromising privacy. The proposed divestiture would primarily insulate the web ecosystem from Google's specific dominance and its ability to exert vertical control over and manipulation of the browser, search, and ad ecosystems, rather than from Big Tech influence in general. By disrupting this integration, such a remedy could limit the spread of information disorder or self-serving information gatekeeping that stem from the manipulation of search algorithms, content prioritization, and data-driven behavioural targeting.

UNDERUTILIZED EXISTING REGULATIONS AND CAPACITIES

Countries can use existing laws against defamation, libel, and incitement to violence to combat information disorder online. Updating regulations on electoral campaigns and political advertising can also curb online political disinformation. Consumer protection agencies can combat misinformation by investigating systematic misinformation, establishing guidelines to distinguish between intentional and unintentional errors, and educating consumers as well as significant penalties for malevolent actors. Information and privacy regulators can play a key role in mitigating harmful information ecosystems. Their evolving mandate enables them to develop regulatory frameworks that address modern information challenges through enforcing data transparency, algorithmic accountability, and user protection. These initiatives can limit the spread of harmful content while upholding free expression.

From other areas of public life, academic networks can integrate research from various disciplines to track, analyse, and predict information disorder patterns. Civil society engagement can empower local groups to identify and develop strategies against misinformation, creating resilient responses that are sensitive to local contexts. Competition policy can be leveraged to curb the power of Big Tech platforms and data brokers through updated analytical frameworks, more aggressive merger reviews, and pre-emptive action against anti-competitive acquisitions.

Additionally, it would be useful to explore at greater length how increased competitive attention to platforms with an emphasis on the public interest, rather than so-called consumer welfare, can increase competition amongst platforms thus giving them incentives to seek legitimacy by responding to information disorder. This shift in regulatory focus could create market dynamics where platforms compete not just on user engagement or advertising revenue, but on their ability to demonstrate responsible information stewardship and public benefit.

Tax policies offer another avenue for intervention. New taxes on corporate data stores, AI system outputs, or data labour could disincentivize extractive practices while funding public digital governance resources. Closing tax loopholes and restructuring wealth taxes to account for intangible data assets can also realign incentives. Furthermore, exploring global minimum corporate tax rates, inheritance taxes, and capital gains tax closures can promote a more

equitable distribution of resources. Tax structuring of infrastructure charges, such as landing fees, can recapture value for African countries which can be used to support institutional, regulatory capacity.

If resources allow for investment in regulatory capacity, we recommend prioritizing the creation of AI Safety Audit Systems and deceptive Design Analysis. This is because different countries in the Global South possess juxtaposing resources which impacts their regulatory capacities and the enforceability of these capacities. The merit of an African AI Safety Institute, for instance, has the potential to extend beyond simple compliance or risk mitigation which would enable a better understanding of the ways in which algorithmic systems restructure power dynamics, economic opportunities and social interactions (Timcke, 2024). Additionally, it would be a strategic imperative for preserving self-determination. The institute can present itself as a defence against digital colonialism, where it would assess supply chains, monitor misuse in the military and propaganda contexts, and maintain that any technological developments are ultimately in Africa's best interest. This would foster mechanisms of accountability and understanding by challenging power structures and resisting digital colonialism that is enforced by Big Tech companies in the Global North.

GLOBAL COLLABORATION AND GOVERNANCE

Various international bodies are working to address information disorder and information integrity. The United Nations Special Rapporteur on Freedom of Opinion and Expression monitors and defends global free expression rights through strategic interventions and collaboration. Concurrently, the Open-ended Working Group on ICT Security (2021-2025) advances international cybersecurity cooperation by developing principles and facilitating dialogue among UN member states. The UN High-Level Advisory Body on Artificial Intelligence addresses AI opportunities and risks while emphasizing ethical guidelines and transparency, though it faces criticism for overlooking market-driven social inequalities. Similarly, the Global Partnership on Artificial Intelligence (GPAI) fosters responsible AI development through multi-sector expert collaboration focused on research, ethics, and international cooperation. Meanwhile, the UN Information Integrity Initiative works to safeguard the information ecosystem against disinformation by empowering individuals, promoting media literacy, and advocating transparency.

Given that many major platforms are based in the United States, a U.S. withdrawal or non-participation in international initiatives like UNESCO's Guidelines for the Governance of Digital Platforms would undermine the enforceability and global coherence of such standards. The multitude of significant platforms emanating from one country itself poses a risk to ethical, safe and representative AI. Indeed, platforms may not feel obligated to comply with these guidelines, weakening their global impact. This could lead to fragmented regulation, where platform governance norms vary significantly by region, and may also embolden companies to resist transparency and accountability measures, especially around algorithmic moderation.

CONCLUSION

We propose four high-level measures to regulate dominant technology companies and address information disorder.

First, foster digital pluralism. Support a diverse ecosystem of local platforms (commercial, non-profit, and public) that gives citizens meaningful alternatives to global giants. This can involve declaring essential digital infrastructure like search engines and foundational AI models as a 'commons' governed by new public institutions representing many actors. The creation of publicly accessible computing marketplaces without vendor lock-in, with governments prioritizing procurement from these sources while phasing out contracts with dominant corporations.

Second, prevent digital power consolidation. Rigorously regulate mergers, acquisitions, venture capital investments, and strategic alliances that extend the control of dominant players. Address the monopolization of intellectual assets like data, knowledge, and narratives. Grant regulators access to accumulated intangible assets and promote public use of collectively generated data to limit the concentration of social knowledge.

Third, implement effective taxation. Recognise that large digital corporations capture significant rents and accumulate social knowledge while operating globally. National governments should impose digital services taxes where economic activity actually occurs, where users generate data and engage with platforms. However, unilateral measures alone are insufficient. International frameworks must establish global minimum corporate taxes and consistent rules to prevent profit-shifting and tax avoidance. These revenues can be

directed toward alternative digital infrastructures, public interest technologies, and democratic digital sovereignty.

Fourth, prohibit harmful technologies. Not all technological advancement is beneficial. One must consider banning technologies that displace labour without meaningful alternatives or those that increase workforce surveillance and control. Regulation is toothless if it cannot wholesale prohibit technologies posing systemic risks to democracy, human rights, or public safety. Prioritize labour protection, creative autonomy, and human rights. Pursue multilateral agreements to guarantee autonomous, democratically governed institutions and technological solutions.

References

African Union. 2020. *The Digital Transformation Strategy for Africa* [Online]. Available: <https://au.int/sites/default/files/documents/38507-doc-dts-english.pdf> [2025, May 6].

African Union. 2022. *AU Data Policy Framework* [Online]. Available: <https://au.int/sites/default/files/documents/42078-doc-AU-Data-Policy-Framework-Eng1.pdf> [2025, May 6].

African Union. 2024. *Continental Artificial Intelligence Strategy* [Online]. Available: <https://tinyurl.com/3ek9uv96> [2025, May 6].

Afrocurator. 2019. *Moleskin Foundation*, October [Online]. Available: <https://moleskinfoundation.org/afrocurator/> [2025, May 2].

Alsaleh, A. 2024. The impact of technological advancement on culture and society. *Scientific Reports*, 14(1):32140.

Ashiru, G. 2025. AWS to Accept Naira Payments, Shaping Nigeria's Cloud Services Market. *Tech in Africa* [Online]. Available: <https://www.techinafrica.com/aws-to-accept-naira-payments-shaping-nigerias-cloud-services-market/> [2025, May 2].

BBC. 2019. Facebook 'to be fined \$5bn over Cambridge Analytica scandal', 13 July [Online]. Available: <https://www.bbc.com/news/world-us-canada-48972327> [2025, May 2].

BBC. 2023. Nigerian elections 2023: False claims and viral videos debunked, 28 February [Online]. Available: <https://www.bbc.com/news/world-africa-64797274> [2025, May 5].

Bietti, E. 2020. Consent as a Free Pass: Platform Power and the Limits of the Informational Turn. *Pace Law Review*, 40(1):307-397.

Bloodstein, R. 2019. Amazon and Platform Antitrust. *Fordham Law Review*, 88(1):189.

Brown, I. & Marsden, C.T. 2023. *Regulating code: Good governance and better regulation in the information age*. Cambridge, Mass.: MIT Press.

Carlson, M. 2020. Journalistic epistemology and digital news circulation: Infrastructure, circulation practices, and epistemic contests. *New Media & Society*, 22(2):230-246. doi:10.1177/1461444819856921

Chen, B. 2022. *TikTok and Surveillance Capitalism: A Survey of Users on Power Imbalances and Potential Harms*.

CIPESA. 2025. What Does Meta's About-Turn on Content Moderation Bode for Africa? 14 January. *Collaboration on International ICT Policy for East and Southern Africa (CIPESA)* [Online]. Available: <https://cipesa.org/2025/01/what-does-metas-about-turn-on-content-moderation-bode-for-africa/> [2025, May 2].

Cooper, A. 2025. The True Cost of Browser Innovation: Why Chrome's Divestiture Wouldn't End the Open Web, 24 February. *Tech Policy Press* [Online]. Available: <https://www.techpolicy.press/the-true-cost-of-browser-innovation-why-chromes-divestiture-wouldnt-end-the-open-web/> [2025, May 2].

Diamniadio National Datacenter. 2021. *Senegal Numerique*, 22 June [Online]. Available: <https://www.adie.sn/domaine/entreprise> [2025, May 5].

Doctorow, C. 2024. 'Enshittification' is coming for absolutely everything, 8 February. *Financial Times*.

Dosunmu, D. 2025. Nigerians are building affordable alternatives to AWS and Google Cloud. *Rest of World, Reporting Global Tech Stories* [Online]. Available: <https://restofworld.org/2025/aws-google-cloud-nigeria-alternatives/> [2025, May 2].

Duffy, B.E., Pinch, A., Sannon, S. & Sawey, M. 2021. The nested precarities of platformized creative labour.

Duffy, B.E. & Meisner, C. 2023. Platform governance at the margins: Social media creators' experiences with algorithmic (in) visibility. *Media, Culture & Society*, 45(2):285-304.

- Feldstein, S. 2021. *The rise of digital repression: How technology is reshaping power, politics, and resistance*. Oxford: Oxford University Press.
- Ford, H. 2011. The Missing Wikipedians, in *Critical Point of View, A Wikipedia Reader*. Amsterdam: Institute of Network Cultures.
- Francois, M. & George, C. 2019. Introducing Equiano, a subsea cable from Portugal to South Africa. *Google Cloud* [Online]. Available: <https://tinyurl.com/3rbphbu4> [2025, May 2].
- Gorwa, R., Binns, R. & Katzenbach, C. 2020. Algorithmic content moderation: Technical and political challenges in the automation of platform governance. *Big Data & Society*, 7(1). doi:10.1177/2053951719897945
- Hao, K. 2021. The Facebook whistleblower says its algorithms are dangerous. Here's why. *MIT Technology Review* [Online]. Available: <https://tinyurl.com/2e6r4zpb> [2025, May 2].
- Hollyland. 2025. *What Data Does TikTok Collect and Why It Matters* [Online]. Available: <https://www.hollyland.com/blog/tips/what-data-does-tiktok-collect> [2025, May 2].
- Jordan, T. 2020. *The digital economy*. University of Sussex.
- Khan, L. 2016. Amazon's Antitrust Paradox. *The Yale Law Review*, 126(3):805.
- Kwode, P.A.K. & Sele Kane, N.L. 2023. Fake news and the political economy of the media: A perspective of Ghanaian journalists. *Communicare: Journal for Communication Sciences in Southern Africa*, 42(2):55-63.
- Makananise, F.O. & Madima, S.E. (eds.). 2025. *Decolonising digital media and indigenisation of participatory epistemologies*. London: Routledge.
- Malinga, S. 2025. TikTok insists African creators are 'coining it' on the platform, 11 April. *ITWeb* [Online]. Available: <https://www.itweb.co.za/article/tiktok-insists-african-creators-are-coining-it-on-the-platform/KQenvjyknRqZd2r> [2025, May 2].
- Mapitsa, C. & Churchill, C. 2024. *Monitoring Systems in Africa*. ResearchGate [Online]. Available: https://www.researchgate.net/publication/380205481_Monitoring_Systems_in_Africa [2025, May 6].
- Masakhane. 2019. *Masakhane a grassroots NLP community for Africa, by Africans* [Online]. Available: <https://www.masakhane.io> [2025, May 2].
- Matthew, J. 2021. It might get harder for budding African creators to make money with YouTube's new terms of service, 1 June. *Techpoint Africa* [Online]. Available:

<https://techpoint.africa/news/youtube-terms-of-service-update/> [2025, May 2].
Mozilla Foundation. 2022. *Mozilla Africa Innovation Mradi: In Real Life (IRL) Fund*, October [Online]. Available: <https://foundation.mozilla.org/en/what-we-fund/programs/Africa-Mradi-IRL-Fund/> [2025, April 30].

Muia, B.K.Y. & W. 2023. Facebook's algorithms "supercharged" hate speech in Ethiopia's Tigray conflict, 31 October [Online]. Available: <https://www.bbc.com/news/world-africa-67275219> [2025, May 2].

Muldoon, J. & Wu, B.A. 2023. Artificial intelligence in the colonial matrix of power. *Philosophy & Technology*, 36(4):80. doi:10.1007/s13347-023-00687-8
Munga, J. 2025. *Beneath the Waves: Addressing vulnerabilities in Africa's undersea digital infrastructure*. Carnegie Endowment for International Peace [Online]. Available: https://carnegie-production-assets.s3.amazonaws.com/static/files/Munga_Undersea%20Cables-025.pdf [2025, April 30].

Nanfuka, J. 2021. *Uganda's 2021 Election: A Textbook Case of Disruption to Democracy and Digital Networks in Authoritarian Countries*. CIPESA [Online]. Available: <https://cipesa.org/2021/01/ugandas-2021-election-a-textbook-case-of-disruption-to-democracy-and-digital-networks-in-authoritarian-countries/> [2025, May 2].

Newman, L. 2019. WikiAfrica Education Initiative teaches South African students to decolonise the internet with Wikipedia, 29 January. *Wikimedia Foundation* [Online]. Available: <https://wikimediafoundation.org/news/2019/01/29/wikiafrica-education-initiative-teaches-south-african-students-to-decolonise-the-internet-with-wikipedia/> [2025, May 2].

Obi, P.A. 2024. Labouring and Smiling: Re-Imagining Digital Colonialism in Africa, Silicon Valley Big Techs, and the Politics of Prosumer Capitalism in Nigeria. *tripleC: Communication, Capitalism & Critique. Open Access Journal for a Global Sustainable Information Society*, 22(1):381-395. doi:10.31269/triplec.v22i1.1451

Ovesen, H. 2014. *The political economy of Wikiality: A South African inquiry into knowledge and power on wikipedia*. Unpublished master's thesis. Cape Town: University of Cape Town [Online]. Available: <https://open.uct.ac.za/server/api/core/bitstreams/5ffc3724-25fb-4c78-923a-d5fee8e8d3fb/content> [2025, May 2].

Paradigm Initiative. 2024. Digital Rights - Paradigm Initiative, 7 June [Online]. Available: <https://paradigmhq.org/programs/digital-rights/> [2025, May 2].

Rens, A. 2015. South Africa, Censorship on Demand: Failure of Due Process in ISP Liability and TakeDown Procedures, in N. Rizk, C.A. De Souza & P. Parakesh (eds.). *Global Censorship and Access to Knowledge, International Case Studies*. New Haven, Conn.: Information Society Project, Yale Law School. 65-81.

Republic of Kenya. 2019. *Digital Economy Blueprint, Powering Kenya's Transformation* [Online]. Available: <https://lands.go.ke/wp-content/uploads/2021/04/Kenya-Digital-Economy-Blueprint-2019.pdf> [2025, May 6].

Republic of Rwanda, Ministry of ICT and Innovation. 2023. *The National AI Policy* [Online]. Available: <https://www.minict.gov.rw/index.php?eID=dump-File&t=f&f=67550&token=6195a53203e197efa47592f40ff4aaf24579640e> [2025, May 6].

Research ICT Africa. 2023. *Submission on: What might a just AI global governance framework look like for Africa?* [Online]. Available: <https://researchictafrica.net/research/what-might-a-just-ai-global-governance-framework-look-like-for-africa/> [2025, May 2].

SafeBoda. 2017. [Online]. Available: <https://www.safeboda.com/uganda-website/about-us-ug> [2025, May 2].

Schroeder, Z., Haller, D., Timcke, S. & Rens, R. 2025. *The impact of social media and Generative AI on gender-based violence*. Research ICT Africa [Online]. Available: <https://researchictafrica.net/research/the-impact-of-social-media-and-generative-ai-on-gender-based-violence/> [2025, May 2].

Timcke, S. 2021. *Algorithms and the End of Politics, How Technology Shapes 21st-Century American Life*.

Timcke, S. 2024a. Deficits, Gaps, and Compromises in the UN's "Governing AI for Humanity". *Tech Policy Press* [Online]. Available: <https://www.techpolicy.press/deficits-gaps-and-compromises-in-the-uns-governing-ai-for-humanity/> [2025, May 2].

Timcke, S. 2024b. Towards Tech Self-Determination: The case for an African AI Safety Institute, 14 December. *Afrika Techno Policy* [Online]. Available: https://afrikatechnopolity.blogspot.com/2024/12/towards-tech-self-determination-case_14.html [2025, May 7].

Timcke, S., Orembo, E. & Hlomani, H. 2023. *Information Disorders in Africa*. Research ICT Africa [Online]. Available: <https://tinyurl.com/dt6vhthw> [2025, May 2].

Timcke, S. & Rens, R. 2024. *Using regulation to curtail platform decay and information Disorder*. Research ICT Africa [Online]. Available: <https://tinyurl.com/mr3p76b8> [2025, May 2].

Timcke, S. & Schroeder, Z. 2024. *Engagement at what cost? Examining the intersection of social media, Generative AI and gender-based violence*. Research ICT Africa [Online]. Available: <https://researchictafrica.net/research/engagement-at-what-cost-examining-the-intersection-of-social-media-generative-ai-and-gender-based-violence/> [2025, May 2].

Treré, E. 2016. *The dark side of digital politics: Understanding the algorithmic manufacturing of consent and the hindering of online dissidence* [Online]. Available: <https://hdl.handle.net/20.500.12413/7697> [2025, May 2].

United Nations AI Advisory Body. 2024. *Governing AI for Humanity* (eISBN: 9789211067873). New York: United Nations Publications [Online]. Available: https://www.un.org/sites/un2.un.org/files/governing_ai_for_humanity_final_report_en.pdf [2025, May 5].

Washington Post. 2021. Five Points for Anger, One for a ‘Like’: How Facebook’s Formula Fostered Rage and Misinformation [Online]. Available: <https://www.washingtonpost.com/technology/2021/10/26/facebook-angry-emoji-algorithm/> [2025, May 2].

Wasserman, H. 2022. Introduction, in H. Wasserman (ed.). *Meeting the Challenges of Information Disorder in the Global South*. IDRC [Online]. Available: <https://idl-bnc-idrc.dspacedirect.org/bitstream/handle/10625/60954/IDL-60954.pdf> [2025, May 2].
Whose knowledge? [n.d.]. [Online]. Available: <https://whoseknowledge.org/> [2025, May 2].

Yu, J., Bekerian, D.A. & Osback, C. 2024. Navigating the Digital Landscape: Challenges and Barriers to Effective Information Use on the Internet. *Encyclopedia*, 4(4):1665-1680. doi:10.3390/encyclopedia4040109

EMPOWERING CHILDREN AGAINST MISINFORMATION:

A Review Of Media And
Information Literacy (Mil)
Interventions In Sri Lanka

Isuru Samartunga

Helani Galpaya



INTRODUCTION

Information disorder is a widespread problem in countries in the Global South (Wasserman, 2022). Wardle and Derakshan's (2017) conceptual framework on information disorder distinguishes three types of information: Misinformation, Disinformation and Malinformation based on falseness and intent. Misinformation is false information shared without harmful intent (e.g. a person forwarding a social media message genuinely believing it to be true). Disinformation is false information shared deliberately to mislead (e.g. a knowingly fabricated news article circulated online to influence voters during an election). Malinformation is genuine information shared to cause harm (e.g. leaks of private data that could damage the reputation of a person). Despite these definitional differences, all forms can yield similar harmful societal consequences such as corrupting the integrity of democratic processes and reducing vaccination rates. These outcomes illustrate how information disorder impacts the public sphere regardless of whether falsehoods spread with intent or by accident (Posetti and Cherilyn, 2018).

Approaches to counter information disorder span multiple disciplines. On the technological side, researchers are applying computational tools (network analysis, machine learning, data mining, and sentimental analysis) to detect and model the spread of false content. Complex network models can simulate the propagation of "fake news", and classifiers can be trained to distinguish false from true content. Agent-based simulations have also been proposed to predict how individual behaviours (e.g. sharing habits) lead to large-scale misinformation cascades (Gradoń et al, 2021). While these tech-based methods are capable in early detection or intervention, they are not a standalone solution.

Addressing information disorder requires a combination of strategies implemented by different actors in four broad categories: 1) Law and regulations to deter false content though it raises free speech and jurisdictional concerns, 2) Content related measures including quality journalism, labelling, takedowns, and algorithmic moderation, though they often lack transparency and depends on platforms, 3) Independent fact-checking, which is limited by funding, scale, and reach compared to viral falsehoods, and 4) Educational approaches, such as digital literacy, though there are sustainability and impact

gaps (Day of Information Disorder: Evidence based solutions for a resilient digital age, 2025).

Educators are pushing for media and information literacy (MIL) programs. Such interventions can often be rapidly tailored to local contexts and involve those directly affected by “fake news.” Wasserman (2022) finds that across Africa, Asia, and Latin America, the most common strategies against information

disorder included fact-checking, media literacy education, public awareness campaigns, and policy changes. These efforts leverage local knowledge and encourage critical consumption of information, complementing the more technical solutions.

A World Economic Forum (WEF) report reinforces the idea that a whole-of-society approach is essential (World Economic Forum, 2025). The report presents an integrated model aligning the “disinformation life cycle” (pre-creation to post-consumption) with a

socio-ecological model (individual to policy level), illustrating intervention points throughout the information ecosystem. It argues that putting the burden solely on individual users is inadequate given the scale and sophistication of today’s disinformation tactics.

MEDIA AND INFORMATION LITERACY (MIL) PROGRAMS AS COUNTERMEASURES

A policy report notes that MIL efforts tend to be closely connected to the communities that misinformation affects and thus can innovate rapidly in response to new disinformation tactics (United Nations General Assembly, 2022). Further, several scholars emphasize the promotion of MIL programs for the public to counter information disorder (Seo, 2019; Cohen, 2020). MIL programmes broadly refer to equipping citizens with the competencies to access information, analyse it critically, create content, and act as informed consumers of media, and the approach focuses on empowering users by teaching people the critical thinking skills needed to discern factual content from falsehoods. MIL interventions range from school curricula to public training programmes and gamified online trainings. One advantage of such programs is that they can be highly adaptive and locally relevant. The

“
Educators are pushing for media and information literacy (MIL) programs. Such interventions can often be rapidly tailored to local contexts and involve those directly affected by “fake news.”
”

participants of MIL programs are expected to gain necessary critical thinking skills when using online content to differentiate between factual and fake content, helping to reduce disinformation overtime. MIL empowers citizens with competencies which will enable them to access, critically analyse, create, and consume information. Dumitru (2020), for instance, highlights that a digitally literate individual learns to apply an internal “filter” of scepticism when navigating the web – cross-checking sources, recognizing logical fallacies or emotional manipulation, and forming independent, evidence-based opinions. Ng (2012) proposed a multidimensional model of digital literacy which includes technical skills, cognitive evaluation skills, and social-emotional skills for online engagement. This means MIL exceeds basic computer use; it involves leveraging one’s general knowledge and reasoning abilities to examine online content.

Major technology companies have also invested in digital literacy initiatives as part of countering misinformation. For example, Meta’s “We Think Digital” campaign is a global digital literacy program aiming to create responsible digital citizens. In the Asia-Pacific region, Meta collaborated with experts, NGOs, and governments to roll out localized We Think Digital training modules. In Sri Lanka, the program launched in 2020 via a partnership with Sarvodaya-Fusion , with support from the national Ministry of Education. This campaign educated students, youth, and adults on how to engage positively and safely online. Google supported Sarvodaya-Fusion in implementing the “Be Internet Awesome” program, which teaches school children the fundamentals of digital citizenship and online safety. In 2022–2023 alone, this program reached over 1,600 students (ages 7–14) across multiple districts, using interactive lessons and an educational game (“Interland”) to reinforce skills like recognizing phishing scams, creating strong passwords, and thinking critically about online information. These programs by Meta, Google, and others demonstrate how digital literacy efforts can be scaled through public-private partnerships (Sarvodaya Fusion, n.d.).

Though there is optimism that MIL can empower society against misinformation, scholars notice that its effectiveness is not guaranteed. Emerging evidence on MIL interventions shows mixed results. Some experimental studies have found only modest improvements in participants’ ability to spot fake content, and occasionally unexpected outcomes. For example, one study noted that after a training, students became less successful

at detecting bots on social media, while another found no significant improvement in the quality of sources students chose when researching, despite undergoing a MIL course (Anstead, Edwards, Livingstone, and Stoilova, 2025). These findings suggest that simply exposing people to a short workshop or curriculum might not be sufficient. According to the World Economic Forum (2025), successful MIL initiatives integrate interactive pedagogies—fact-checking exercises, role-play simulations, and gamified learning—and are embedded across formal, informal, and professional settings. MIL should be integrated with institutional support, technological safeguards, and regulatory frameworks that promote transparent algorithms and responsible platform behaviour, the report insists.

borated with experts, NGOs, and governments to roll out localized We Think Digital training modules. In Sri Lanka, the program launched in 2020 via a partnership with Sarvodaya-Fusion, with support from the national Ministry of Education. This campaign educated students, youth, and adults on how to engage positively and safely online. Google supported Sarvodaya-Fusion in implementing the “Be Internet Awesome” program, which teaches school children the fundamentals of digital citizenship and online safety. In 2022–2023 alone, this program reached over 1,600 students (ages 7–14) across multiple districts, using interactive lessons and an educational game (“Interland”) to reinforce skills like recognizing phishing scams, creating strong passwords, and thinking critically about online information. These programs by Meta, Google, and others demonstrate how digital literacy efforts can be scaled through public-private partnerships (Sarvodaya Fusion, n.d.).

Though there is optimism that MIL can empower society against misinformation, scholars notice that its effectiveness is not guaranteed. Emerging evidence on MIL interventions shows mixed results. Some experimental studies have found only modest improvements in participants’ ability to spot fake content, and occasionally unexpected outcomes. For example, one study noted that after a training, students became less successful at detecting bots on social media, while another found no significant improvement in the quality of sources students chose when researching, despite undergoing a MIL course (Anstead, Edwards, Livingstone, and Stoilova, 2025). These findings suggest that simply exposing people to a short workshop or curriculum might not be sufficient. According to the World Economic Forum (2025), successful MIL initiatives integrate interactive

pedagogies—fact-checking exercises, role-play simulations, and gamified learning—and are embedded across formal, informal, and professional settings. MIL should be integrated with institutional support, technological safeguards, and regulatory frameworks that promote transparent algorithms and responsible platform behaviour, the report insists.

MEDIA AND INFORMATION LITERACY (MIL), CHILDREN, AND INFORMATION DISORDER

When considering digital literacy, age is a pivotal factor. Marc Prensky's (2001) concept of “digital natives” versus “digital immigrants” suggests that young people who grew up with the internet are naturally fluent in digital technologies, whereas older generations who adopted tech later will always be playing catch-up. He suggests that digital-native students speak an entirely new “language” of tech, causing a disconnect with immigrant instructors from a pre-digital age. This idea has been influential, but later research offers a more critical perspective. Studies have debunked the “digital native” myth, finding that simply being born in the internet era does not guarantee effective digital literacy. As one review mentions it, there is “no evidence that there is a single new generation of young students” inherently proficient in navigating digital information; many youths use technology in the same shallow way as older people – for example, passively consuming content without verifying it (Nature, 2017). Education systems must not assume children have critical digital skills by default (Kirschner and De Bruyckere, 2017). In other words, a teenager with many hours on social media does not naturally become skilled at spotting misinformation. The digital natives' concept has value in understanding generational differences; however, the concept undermines the need for training for the younger generation.

UNICEF's analysis underscores that children cannot always distinguish truth from falsehood and thus may be harmed by misinformation and propagate it among peers. For example, during health crises like COVID-19 pandemic, children were exposed to conspiracy theories on social media (sometimes directly, other times via parents or teachers) and lacked the experience to judge what is trustworthy (Howard, P., Neudert, L.M., Prakash, N. and Vosloo, S. 2021).

Empirical studies on children's ability to spot misinformation emphasise the need to start digital literacy in early ages. For instance, a study by Loos,

Ivan, and Leu (2018) had children examine a spoof website about a fictional “Pacific Northwest Tree Octopus”, Only two out of 27 Dutch children (11–12 years old) realized the website was a hoax – the rest believed the false information. The authors noted that the classroom context and the students’ lack of cynicism toward a well-presented site made them highly vulnerable to the fake content. The finding underscores that being a ‘digital native’ does not inherently confer critical digital literacy- the evaluative and reflective capacities required to detect and counter misinformation.

The World Economic Forum (2025) advocates for embedding MIL into early education, noting that trust in AI-generated content is high among youth while actual evaluative ability remains low. The report calls for integrating AI and data literacy into school curricula and designing age-appropriate interventions that consider cognitive development stages and emotional literacy.

MEASURING THE IMPACT OF MEDIA AND INFORMATION LITERACY (MIL) PROGRAMS ON CHILDREN

Despite the enthusiasm for digital literacy as a remedy for information disorder, there is a notable gap in understanding its long-term impact. As Wasserman (2022) observes, organizations fighting information disorder often struggle with measuring the reach and efficacy of their interventions, including media literacy, and encourage longer-term collaborations between academics and practitioners to develop better evaluation frameworks. It is relatively straightforward to count how many workshops were delivered or how many students participated (reach), but far more difficult to determine how those interventions affect behaviour and beliefs months or years down the line (impact). To date, many digital literacy initiatives rely on short-term feedback (e.g. pre- and post-training quizzes) which might show immediate knowledge gains. However, whether a one-off training leads to sustained improvement in someone’s ability to identify and ignore misinformation is uncertain.

In the context of Sri Lanka, these measurement challenges are especially pertinent. The country has reasonably high basic digital literacy rates and internet usage is rapidly growing. According to the 2024 Department of Census and Statistics, 64.3% of Sri Lankans (ages 5–69) are “digital literate”. Males show slightly higher levels of digital literacy (66.8%) compared to females (62.0 %). Urban areas lead significantly with 77.5% digital literacy,

reflecting greater infrastructure, resources and access to technology. Rural areas lag with 62.5% digital literacy, while the lowest rate (42.6%) is reported in the estate areas: often referred to as the plantation sector, which reflects some of the country's most pronounced inequalities. Among children aged 5-9, digital literacy is 45.6%, suggesting early exposure to mobile technology rather than computers. Among 10–14-year-olds, digital literacy rises sharply to 76.3% as this stage children begin to use digital tools for schoolwork, communication, and online learning resources. The 15-19 age group demonstrates very high digital literacy at 92.1%, reflecting frequent use of digital technology.

However, the official definition on digital literacy is quite narrow. A “digitally literate” person is defined simply as someone who can operate a computer, laptop, tablet, or smartphone on their own (Department of Census and Statistics, 2025). This does not entail any assessment of internet skill, information literacy, or safe online practices. By such a measure, a child who can open YouTube on a phone would count as digitally literate, even if they have no idea how to verify a news headline or adjust privacy settings. Thus, meaningful MIL (in terms of critical thinking and informed internet use) is much lower. This discrepancy underscores the need to evaluate MIL programs on outcomes that matter (critical evaluation skills, changes in behaviour), not just on access or basic usage stats.

METHODOLOGY

LIRNEasia conducted a longitudinal, quasi-experimental qualitative study over 12 months (November 2023 to January 2025). The study examined the effectiveness of MIL interventions as a countermeasure to the information disorder, among school-aged children in Sri Lanka. This longitudinal approach was required for tracing the durability of knowledge and behavioural shifts by the interventions.

INTERVENTION DESIGN

A MIL programme by Sarvodaya-Fusion was used as a case study for this study. The programme aimed at equipping children with essential skills to navigate the digital world safely, ethically, and responsibly. Recognizing that a classroom-style setting would be most familiar and effective for the target

age group, the sessions were delivered in an environment that mimicked traditional classroom dynamics. This approach fostered a sense of comfort and encouraged active engagement among participants, enabling more effective learning. All the sessions were conducted in the children’s native language:

“
The programme aimed at equipping children with essential skills to navigate the digital world safely, ethically, and responsibly.

”

Sinhala and Tamil. The training was based on an existing digital literacy curriculum that Sarvodaya-Fusion has been conducting over time. For the purpose of this research, they customized the program content to align with the objectives and requirements of the study and their regular trainers were trained by a master trainer to ensure the consistency across training sessions Sarvodaya-Fusion adapted Google’s “Be Internet Awesome” curriculum to suit children between the ages of 11 to 14, taking into account their varying cognitive abilities and levels of digital maturity. The training was conducted for 60 minutes, introduced participants to five key principles of digital safety: Be Internet Smart, Be Internet Alert, Be Internet Strong, Be Internet Kind, and Be Internet Brave.

The training emphasized three core areas. First, Be Internet Smart (Share with Care) encouraged children to think carefully before posting content online and to understand the importance of safeguarding personal information. Second, Be Internet Alert (Don’t Fall for Fake) focused on helping children recognize misinformation, online scams, and fake profiles. Third, Be Internet Strong (Secure Your Secrets) provided practical instruction on how to create strong passwords and manage privacy settings across digital platforms.

Each module featured interactive components to enhance learning and reinforce critical thinking. These included scenario analysis exercises to evaluate which types of information are safe to share, activities aimed at detecting fake online identities, simulations to spot phishing scams and misinformation, practical exercises in creating strong passwords, and guided exploration of tools to verify information content.

PARTICIPANTS AND DEMOGRAPHICS

The research focused on children aged 11 to 14 years, selected from three districts (Colombo, Monaragala, and Nuwara Eliya) to reflect the socio-economic diversity of the country. All participating children were full-time students enrolled in school. Children from Colombo and Monaragala districts

studied in either the Sinhala or English medium of instruction, while those from Nuwara Eliya studied in the Tamil medium. The study sample included 90 children in total, 42 male and 48 female, with 28 respondents from Colombo, 32 from Monaragala, and 30 from Nuwara Eliya.

In addition, a total of 91 parents or guardians of both treatment and control group participants were interviewed to understand parental attitudes toward the training and the perceived value of developing media and information literacy (MIL) in children.

Colombo district represents the urban sector in this study and, according to Sri Lanka's Multidimensional Poverty Index (MPI) 2019 compiled by the Department of Census and Statistics, it recorded the lowest incidence of poverty in the country at just 3.5%. Monaragala district, representing the rural sector, recorded a significantly higher MPI rate of 32.7%, ranking as the third poorest district out of Sri Lanka's 25 districts. Nuwara Eliya district represents the estate sector in this study; according to the 2012 Census of Population and Housing, 53.5% of its population resides in estate areas. Furthermore, the district faces considerable socio-economic challenges, with 44.2% of the population identified as living in multidimensional poverty under the Sri Lankan National MPI.

RISK ASSESSMENT

Given the involvement of children, the research team conducted a risk and ethical assessment before initiating the study. Potential risks identified included the possibility of coercive recruitment, breaches of confidentiality, psychological distress due to the content of discussions, and cultural sensitivity concerns. To address these, recruitment was managed by a trusted community-based organization (Sarvodaya-Fusion) with longstanding ties to the local communities and established child protection policies. Informed consent was obtained from both children and their parents or guardians using clear, accessible local language.

Qualitative data collectors (moderators) and field staff were selected from the same localities and ethnic communities as participants, and were trained in ethical data collection, child-sensitive interviewing, and research protocols to ensure respectful and safe interactions. Special care was taken to ensure

anonymity, non-judgmental engagement, and cultural appropriateness throughout the study. LIRNEAsia implemented a set of organizational risk mitigation strategies to safeguard participant information. All individuals and partner organizations involved in the research process—such as interview transcribers—signed non-disclosure agreements (NDAs) to ensure confidentiality. All data collected for the study, including audio recordings, photographs, and transcripts, were stored securely on password-protected devices. To further minimize the risk of information breaches, all materials were kept in digital format only—transcripts and photographs were not printed at any stage of the research process.

LIRNEAsia obtained research ethics approval for the study from the social, behavioural and educational research ethics committee, at the University of Stellenbosch, South Africa in 2023.

DATA COLLECTION

The study employed a comparative design between a treatment group (TG)—59 children who received the MIL intervention—and a control group (CG)—31 children who did not receive the training. The data collection approach involved FGDs and IDIs across three phases:

PHASE 1: PRE-INTERVENTION AND IMMEDIATE POST-INTERVENTION (MONTH 0)

Prior to the intervention, both TG and CG participants engaged in FGDs to gather baseline data on their socio-economic backgrounds, media consumption patterns, and initial understanding of online content. Each FGD consisted of 5 to 6 children and was designed to encourage open discussion in age-appropriate language. A total of 30 FGDs were conducted prior to the intervention. Following this, the TG children participated in the MIL training session: the intervention.

After the intervention, they took part in follow-up FGDs to reflect on the training experience and express their initial impressions of its usefulness. In addition, 91 in-depth IDIs were conducted with the parents or guardians of both TG and CG participants to understand parental attitudes toward the training and the perceived value of developing MIL in children. All these were conducted at the training centre/intervention location.

Additionally, during the pre-intervention and post-intervention FGDs, all children were shown a series of visuals content to assess their ability to recognize and evaluate online information. Their responses were documented as part of a broader effort to track changes in MIL skills over time.

PHASE 2: 6 MONTHS AFTER THE INTERVENTION

The second phase (six months after the intervention) involved a total of 56 IDIs conducted exclusively with TG participants. These interviews were carried out through household visits to create a familiar, comfortable environment. The interviews explored whether the training had led to sustained changes in online behaviour, critical thinking, and the ability to counter misinformation. Children were asked to reflect on their daily online activities, how they assessed the credibility of content, and whether they continued to apply lessons from the training in real-life digital contexts.

PHASE 3: 12 MONTHS AFTER THE INTERVENTION

The final phase (12 months after the intervention), the research team revisited both TG and CG participants to assess long-term outcomes and retention of learning. A total of 85 IDIs (31 CG and 54 TG participants) at the participants households were completed across the two groups. As in phase one, participants were shown a new set of online visuals and asked to comment on their authenticity. This exercise helped evaluate any progression in their critical thinking and information assessment skills over time.

DATA ANALYSIS

Following each phase of data collection, the audio recordings were transcribed verbatim, with transcripts produced in the relevant local languages (Sinhala or Tamil) and also translated into English. The analysis was a collaborative process for the research team. The team began by reviewing a subset of transcripts to identify recurring themes, ideas, and significant patterns. Based on this preliminary review, a detailed codebook in English was developed, outlining the thematic codes to be applied across the data set. This codebook was shared with all relevant researchers, who used it to code the transcripts systematically within Atlas.ti. At least two researchers were involved in coding each transcript.

The next section explores the outcomes of a quasi-experimental intervention

designed to assess the effectiveness of MIL training. Drawing on longitudinal qualitative data, it examines how the training influenced children’s digital behaviour, knowledge retention, and critical thinking over time, and what gaps remain in ensuring sustained media literacy.

THE IMPACT OF MIL TRAINING ON COUNTERING MISINFORMATION

Children in the study were seen to use the internet for a variety of purposes spanning education, entertainment, and social interaction. Many children accessed online content to support their schoolwork, such as looking up

“
These findings reveal that even at a young age, children engage with digital technologies in multifaceted ways that blend academic needs with entertainment and social bonding.

”

lessons in subjects like science, mathematics, history, and civic education. Beyond academics, the internet also served as a personalized entertainment space—children watched videos of vehicles, sports matches, or accessed novels via social media groups, reflecting individual interests. Additionally, children relied on the internet for socialization, often using it to stay connected with peers and extended family members through messaging apps

and group chats. These findings reveal that even at a young age, children engage with digital technologies in multifaceted ways that blend academic needs with entertainment and social bonding. Parental influence emerged as a factor shaping children’s access to and engagement with digital technologies, functioning as digital gatekeepers. Device access was restricted to educational purposes, with time limits and usage supervised by adults.

Given the diverse ways in which children were already navigating digital spaces and the parental control, the MIL intervention sought to build on these existing practices.

The longitudinal quasi-experimental study demonstrates the impact of the MIL training session (intervention) for children on their digital behaviours, attitudes toward the internet, and ability to counter misinformation. The training transformed children from passive digital users to more informed and responsible ones. Insights from both treatment and control groups highlight the contributions and limitations of the intervention in enhancing digital resilience among children.

ENHANCED DIGITAL AWARENESS

Pre-intervention FGDs revealed that children (in both CG and TG) actively used the internet for educational purposes. They often looked up facts studied at school and searched for relevant content for assignments. As one participant noted, “I searched YouTube for a Grade 6 math lesson... and then scored high on that assignment” (R35, TG, Monaragala, 6 month follow up, IDI_07, Male).

Use of platforms such as YouTube was common, sometimes even for tasks that didn't necessarily require video content. For instance, one child shared, “I used YouTube to write an essay about my sports hero. I searched ‘my sports hero English essay’ and found content on Susanthika Jayasinghe [a Sri Lankan sprinter]” (R36, TG, Monaragala, 6 month follow up, IDI_1, Female). These practices illustrate that websites—YouTube in particular—serve as platforms for inquiry-based learning. However, verifying the credibility of video content proved difficult. One participant mentioned “I learnt how to find fake videos by looking at the colours and whether they are edited” (R45, TG, Monaragala, 6 month follow up, IDI_15, Male). This suggests children need training to identify manipulated videos and the content of those. Children lacked strategies to assess reliability and often depended on adults, such as parents or teachers, for clarification. This revealed a dependency rather than autonomy in navigating digital spaces and raised questions about adults' own capacity to verify information.

Post-intervention discussions, however, showed marked improvements in digital awareness. Participants demonstrated a strong grasp of the training's core messages. All participants expressed confidence in applying the knowledge and skills they had acquired and committed to becoming safe, informed, and responsible digital citizens. They articulated their understanding using the terminology introduced during the training: clever, alert, strong, kind, and brave. One participant stated, “I can use the internet without getting into the wrong things. I can protect my personal information. I won't click on unnecessary links. I will only post correct information on social media” (R46, TG, Monaragala, 6 month follow up, IDI_08, Male) Immediate post-intervention, TG, Colombo). Another said, “We have to be very keen and careful while using the internet—like dealing with a bomb” (R8, TG, Immediate post-intervention, Colombo, Female), reflecting a deeper sense of caution and awareness.

DEVELOPED DIGITAL SKILLS

The enhanced awareness among participants led to improved digital skills. In post-intervention discussions, children described how they now looked for verification indicators such as authenticity badges, large numbers of likes or comments, and inconsistencies in image or video quality to identify manipulated content. One participant shared, “I can find out [verified social media accounts] using the verifying mark as well as millions of likes, comments, and false pictures... Sometimes colours [in an image] differ from the original. Also, lines appear on edited, false videos” (R10, TG, Colombo, Immediate post-intervention, FGD_02, Female). These skills were not mentioned in pre-intervention discussions, indicating they were newly acquired through the training.

“ Children also learned about strong passwords and their role in protecting online privacy—an area they had limited knowledge of before. ”

Children also learned about strong passwords and their role in protecting online privacy—an area they had limited knowledge of before. The training included a practical activity on creating strong passwords, which participants appreciated. One child explained, “I learnt how to use simple letters and characters like @#\$ when making passwords. I will share my knowledge with others” (R12, TG, Colombo, 12 month follow up, IDI_02, Female). They also showed a better understanding of hacking threats associated with weak passwords and demonstrated preparedness to respond to such risks.

None of the above specific skills were mentioned in the 12 months follow up interviews with the control group.

CHANGED BEHAVIOURS IN ONLINE SHARING

The training led to noticeable changes in children’s online sharing behaviour. Pre-intervention discussions revealed that children frequently forwarded sensational messages and hoaxes without much scrutiny, sometimes unknowingly spreading misinformation. In follow-up discussions, however, they reported making more conscious decisions about what to share.

One participant recalled receiving a chain message that promised rewards or made threats and said, “I knew clearly that those are lies” (R40, TG,

Monaragala, 6 month follow up, IDI_20, Female). Another shared, “If we get any message... that says, ‘Touch and open this one [referring to online links] and we [the first sender] will give you gifts,’ we don’t be foolish” (R18, TG, Colombo, 6 month follow up, IDI_10, Female). A participant affirmed, “I will only post correct information on social media” (R7, TG, Colombo, Immediate post-intervention, FGD_01, Female). These examples reflect a shift from indiscriminate to thoughtful and responsible sharing. Additionally, children began verifying content before sharing it, using tools such as reverse image search, which was absent in pre-intervention behaviour. Although control group participants remembered receiving chain messages and AI-generated images, none described a conscious decision to avoid sharing them as part of responsible online behaviour.

INSTILLED SOCIAL RESPONSIBILITY

Beyond personal change, the intervention fostered a sense of responsibility to share learnings with peers and family members. A significant outcome of the training was its influence extending beyond direct participants. This was evident in follow-up discussions conducted within the first year after the intervention.

Before the training, children primarily acted as receivers of information. By the six-month follow-up, many had become proactive in sharing knowledge. One child said, “I told them [friends and family] how they [the trainer] taught us to become an internet hero and how to identify fake accounts” (R12, TG, Colombo, 6 month follow up, IDI_01, Female). Another shared, “I learnt how there are fake videos on the internet and how to identify them... how to tell if texts sent by people are real or fake” (R17, TG, Colombo, Immediate post-intervention, FGD_01).

This shift marks the transformation from passive consumers to empowered guides capable of disseminating digital literacy. Participants shared insights on fake accounts, scams, and password safety with others, thereby amplifying the reach of the intervention. Such peer-to-peer knowledge sharing was not observed in pre-intervention discussions, indicating that the training helped instil a sense of social responsibility. However, it is noteworthy that discussions about misinformation were still infrequent among peer groups, as most children indicated that the topic rarely came up in their circles.

Children reported instances of encountering misinformation on social media and through peers. This included false messages, deceptive online advertisements, prank chain messages, and sensational or mythical content. One child recalled sharing a message that falsely promised snowfall if it was forwarded: “I got a message saying to share it with friends, so that makes snowfall. I did, but no snowfall” (R16, TG, Colombo, 12 month follow up, IDI_14, Male). Another described a misleading online game offer: “I played the game, but I still did not receive the 30GB data package” (R29, TG, Colombo, Pre-Intervention, FGD_01, Female).

RECOGNITION OF MISINFORMATION

The treatment group demonstrated a stronger and more confident understanding of misinformation than the control group. Most children in the treatment group were able to identify and analyse misinformation using strategies taught during the training. These included examining visual inconsistencies, checking for verification badges on profiles, and critically questioning suspicious content. One participant stated: “If someone shares a fake post, I tell them it’s fake and explain why.” (R12, TG, Colombo, 12 Month, IDI_02, Female) “I learnt how to find fake videos by looking at the colours and whether they are edited.” (R39, TG, Colombo, 12 Month, IDI_03, Female). These statements reflect both internalized understanding and a willingness to correct others.

In contrast, the control group did not show a clear understanding of misinformation. They often relied on friends or family for verification and used intuition rather than structured verification methods. In Nuwara Eliya, for example, although some children were vaguely aware of false content online, few could clearly articulate what misinformation was or how to recognize it. Their understanding was often anecdotal or based on hearsay: “I heard from friends,” “Teachers warned us,” “I don’t know,” “I ask others,” “Friends told me it’s fake.” Such responses indicate a lack of structured awareness and limited exposure to verification strategies.

RESPONSES TO SUSPICIOUS CONTENT

Children in the treatment group exhibited more proactive responses to suspicious content. These included deleting chain messages, avoiding dubious links offering prizes or mobile data, and warning peers and family members to be cautious. Some expressed confidence in correcting others:

“I told my friends not to believe that video—it looked edited.” (R47, TG, Monaragala, 6 Month, IDI_10, Female). This reflects increased awareness and assertiveness.

In contrast, children in the control group were more passive. Most said they simply ignored suspicious messages rather than interrogating or addressing them. A few mentioned seeking helps: “I just ask someone if I’m not sure.” (R59, CG, Monaragala, 12 Month, IDI_02, Female) This difference suggests the intervention enhanced children’s confidence in acting when encountering misinformation.

DEVELOPMENT OF CRITICAL THINKING SKILLS

Pre-intervention discussions revealed that children had limited awareness of false information and few tools to verify it. Many recalled being confused when they encountered contradictions between facts taught by schoolteachers and those found online. Faced with conflicting information—whether between textbooks and websites or across different videos—they often turned to adults for clarification.

The idea of being “smart online” or an “internet hero” introduced during the training resonated with many participants. After the intervention, the treatment group displayed emerging critical thinking habits, including questioning content, cross-checking sources, and reflecting on the credibility of online material.

Some children in the control group also showed curiosity and a desire to verify information, often by seeking guidance from adults. A few demonstrated emerging media literacy skills even before the intervention. For instance: “I compare what I read on the internet and what I read in books [school textbooks]” (R17, TG, Colombo, Pre-intervention, FGD_02, Female) “I always ask around or check other sources” (R38, TG, Monaragala, Pre-intervention, FGD_02, Male).

However, such behaviours were mostly limited to verifying educational content. Some children also reported skipping questionable chain messages: “I don’t believe them [chain messages] and skipped that message” (R8, TG, Colombo, Pre-intervention, FGD_02, Male).

Post-intervention, treatment group participants showed significantly enhanced

critical thinking and misinformation detection skills. They no longer acted as passive recipients of content but instead questioned it, verified it using multiple sources, and exercised greater caution. One child stated: “Now I can understand whether the message is true or false... I must be very intelligent while using the internet” (R1, TG, Colombo, Post-intervention, FGD_02, Female).

Participants described practical strategies such as reverse image search, checking multiple websites, and consulting adults. They also referred to checking for visual inconsistencies, verification badges, and analysing comment sections to evaluate the credibility of online content. In contrast, none of the control group children mentioned using digital tools for verification.

Even a year after the intervention, some children could still recall and describe digital deception tactics, such as fake giveaways, phishing links, and emotionally manipulative chain messages. They explained how they now verify online content by checking the authenticity of the sender, avoiding suspicious links, and identifying overly enticing rewards. This lasting behavioural change demonstrates that the MIL training successfully instilled critical thinking skills in participants.

RETENTION OF LEARNING FROM THE TRAINING

The post-intervention discussions—conducted immediately after the MIL training, and again at six and 12 months—examined children’s ability to recall and retain the learnings. Insights from the treatment group highlight both the contributions and limitations of the MIL intervention in enhancing digital resilience among children. While practical digital safety skills, such as password management, were retained over time, deeper critical engagement with misinformation and verification practices diminished in the absence of continued reinforcement. These findings underscore the need for sustained, institutionally embedded digital literacy programs that evolve with learners.

FGDs conducted before the MIL training revealed that children’s technical understanding of digital safety was limited. Few participants mentioned secure passwords or awareness of hacking risks. In contrast, by six months after the intervention, the password training component had left a strong and lasting

impression. One child shared: “My favourite part was when they taught us how to create a strong password... and then they checked how many years it would take to hack it” (R3, TG, Colombo, Post-intervention, FGD_01, Male). Another added: “They said I can use zero instead of the letter ‘O’ so that I can make the password stronger” (R18, TG, Colombo, 6 Month follow up, IDI_10, Female), illustrating how the training incorporated relatable examples that supported long-term learning. Children across all treatment locations retained knowledge about creating strong passwords and described applying it in their own online practices as well as sharing it with peers and family members. They recalled specifics such as using at least eight characters, including numbers and symbols, and employed personalized strategies like replacing the letter “O” with “0” in their names—demonstrating retention not only of the concepts but also of practical application.

Six months after the intervention, children in the treatment group continued to show signs of retention, reflection, and cautious engagement with online content. Household visits revealed that many not only remembered the key components of the training but also shared the content with friends and

“
Six months after the intervention,
children in the treatment group
continued to show signs of retention,
reflection, and cautious engagement
with online content.

”

family, helping to diffuse the knowledge beyond the immediate circle of intervention participants. As one participant said: “I told them how they [the trainer] taught us to become an internet hero and how to identify fake accounts on Instagram and Facebook” (R12, TG, Colombo, 6 Month follow up, IDI_01, Female), highlighting the social value and memorability of the training messages. The training appeared to have a lasting influence on children’s decision-making, especially when confronted with

suspicious messages or scams. In the six-month follow-up interviews, children described deliberately avoiding links that promised free data or prizes such as iPhones. “Even my friend received [scam messages], but we knew that it was a lie after what we learnt from the training program. So, we did not click that link or share our details” (R34, TG, Monaragala, 12 Month follow up, IDI_20, Male). “Once I got a message saying that if I do not share the post someone would die. But I did not care for such messages because I knew they were lies” (R12, TG, Colombo, 6 Month follow up, IDI_01, Female). These responses suggest a strengthened internal filter for detecting deception and resisting emotional manipulation.

However, while the six-month findings showed a cautious and thoughtful approach to suspicious content—especially scams and chain messages—this awareness had not meaningfully evolved by twelve months. Although recognition of fake or edited visual content (e.g., implausible images) remained intact, more complex verification strategies—such as reverse image search, identifying phishing tactics, or assessing source credibility—were rarely recalled. As one child noted: “I think we learned how to check if a video is fake... but I don’t remember how” (R71, TG, Nuwaraeliya, 12 Month follow up, IDI_03, Male). Another said: “I just ask my friends if something looks real” (R79, TG, Nuwaraeliya, 12 Month follow up, IDI 13, Female). While a few children mentioned cross-referencing information on multiple websites (e.g., confirming Buddhist terms across three sources), the broader practice of triangulating or independently verifying content was not widespread. Some admitted they would disengage rather than seek clarity: “If I see different pictures of King Maya Dunne, I just don’t look at them anymore” (R04, TG, Colombo, 12 Month follow up, IDI_09, Male).

At earlier time points, children had articulated reflective decision-making and source-checking behaviors. By twelve months, however, references to techniques such as reverse image search, triangulating sources, or examining metadata had largely disappeared. Instead, children reverted to relying on instinct and visual cues as primary decision-making strategies.

This shift—from active verification to passive avoidance—may indicate cognitive fatigue or the gradual erosion of abstract verification habits in the absence of reinforcement through institutional or social structures.

Across interviews, children noted that schools and tuition classes rarely addressed digital safety or misinformation. Most ICT lessons focused on theoretical knowledge (e.g., identifying computer parts or memorizing definitions) rather than building practical digital literacy skills. This lack of curricular integration suggests a fragile retention environment, particularly for abstract and evolving topics such as misinformation, phishing, and digital footprints—concepts that most participants struggled to recall or explain one year after the intervention.

Findings from the follow-up interviews suggest that temptation and misplaced

trust continue to influence behavior. For instance, some children admitted to clicking on suspicious links shared by parents or enticed by promises of prizes such as iPhones. One child recounted: “My father sent me a link on WhatsApp... we tried it out too... But after we clicked on it, the website didn’t work anymore. So, we realised it is a scam and deleted the message from others’ accounts [as the child had shared the scam via WhatsApp with others]” (R08, TG, Colombo, 6 Month follow up, IDI_03, Male).

Another participant described tapping a photo expecting it to change colour; despite knowing it was likely a trick—indicating that temptation sometimes overrides caution. These examples highlight that, even with improved awareness, children’s decision-making can still be influenced by adults or by the temptation of rewards. This illustrates that digital resilience is still in development and would benefit from continued reinforcement over time.

Additionally, interviews with parents revealed that while their efforts to supervise digital activities were well-intentioned, these efforts were often shaped by a lack of technical confidence among them. This emphasizes the importance of MIL programs that not only focus on children’s skill-building but also enhance caregiver awareness. When both children and caregivers are equipped with digital literacy skills, they can navigate the online environment more effectively and safely together

CONCLUSION

The objective of the MIL training was to strengthen children’s ability to identify false information, and the results suggest meaningful progress in this area. The comparison between pre-intervention and follow-up findings reveals notable shifts in children’s digital practices, awareness of misinformation, and critical thinking. These changes indicate that the intervention led to positive, though varied, transformations in how children engage with online information and digital platforms.

Children moved from believing and spreading hoaxes—such as chain messages or fake news about celebrities—to evaluating visual cues, verifying sources, and resisting suspicious links. They demonstrated enhanced critical thinking and digital verification skills. Participants spoke about the importance of looking for verification indicators—such as authenticity marks, large numbers of likes or comments, and discrepancies in image or video

quality—as cues for identifying manipulated content. Notably, even when misinformation was presented by trusted figures (e.g. parents), some children were able to recognize scams and avoid harm, indicating a maturation in judgment and scepticism. We note, however, that the intervention improved children’s resilience against everyday digital manipulation in the short term (i.e. up to 6 months), but the effects wear out in the longer term (i.e. by 12 months). MIL training is sustainable when centred on practical, immediately applicable content such as creating strong passwords. This hands-on, low-complexity content had the highest retention over time, suggesting the effectiveness of personally relevant training modules. However, the decline of more complex cognitive skills—such as source evaluation and misinformation analysis—when not reinforced through institutional support or regular practice highlights the limits of standalone MIL interventions.

The sustainability of MIL is thus deeply tied to broader educational structures, including school curricula and support from parents. A consistent theme across follow-up discussions was the lack of reinforcement from formal education systems. Children reported that ICT lessons rarely addressed digital safety or media literacy. This absence likely contributed to the decline in retention of higher-order skills.

The MIL intervention had a clear and positive impact on children’s digital lives. It fostered greater awareness, taught valuable technical and critical thinking skills, and encouraged responsible online behaviour. Children emerged not only as safer internet users but also as knowledge multipliers within their communities. The findings demonstrate that even short-term, well-designed literacy programs can produce meaningful shifts in attitudes, skills, and practices—especially when reinforced through interactive content and peer engagement. To deepen this impact, future efforts could include follow-up sessions, targeted content for caregivers, and context-specific support to address trust-based vulnerabilities.

RECOMMENDATIONS

1. **Integrate MIL into formal education curricula (Target group: Educators and teacher training institutes)**

To ensure sustainability, MIL content—especially related to misinformation recognition, digital safety, and critical thinking—

should be embedded within school curricula. Current ICT syllabi should be revised to move beyond theoretical knowledge and include practical, age-appropriate training on navigating online information.

2. Reinforce learning through follow-up sessions (Target group: Institutes working on MIL training programmes)

Given the decline in the retention of complex skills over time, periodic refresher sessions are essential. These could be delivered quarterly or biannually and in person or through online videos. The focus should be on reinforcing misinformation detection, source verification, and ethical online behaviour.

3. Develop targeted MIL content for caregivers (Target group: Parent-teacher associations).

Parents and guardians play a critical role in shaping children's digital behavior. MIL programs should include modules for caregivers to enhance their own digital literacy and support them in guiding their children. This will also help address the issue of children receiving misinformation from trusted adults.

4. Emphasize practical and engaging learning modules (Target group: EdTech developers)

The strongest retention was seen in hands-on, immediately applicable components such as password security. Future training should continue to emphasize interactive, activity-based modules that children can relate to and apply in their daily online activities.

5. Address emotional and trust-based vulnerabilities (Target group: School counsellors)

While technical knowledge improved, children were still susceptible to emotionally manipulative messages and misinformation shared by trusted individuals. MIL programs must address these specific vulnerabilities through scenario-based learning, role plays, and discussion prompts that challenge assumptions and encourage critical reflection.

6. Monitor and evaluate long-term impact (Target group: Evaluators, development funders)

Ongoing monitoring and evaluation are critical to understanding how digital behaviour evolves post-intervention. Future studies should incorporate longitudinal assessments to track which skills persist, which decline, and why—informing future program adjustments.

References

Anstead, N., Edwards, L., Livingstone, S. & Stoilova, M. (2025). Media @LSE blog [Web log post] Available: <https://blogs.lse.ac.uk/medialse/2025/05/22/is-media-literacy-an-effective-tool-against-misinformation/> [2025, October 20].

Cohen, R., Parmentier, A., Melo, G., Sahu, G., Annamalai, A., Chi, S., Clokie, T., Farrag, A., Naik, A., Naseem, S., Sakhuja, S., Wang, J., Clausi, R. & Santin, A. (2020). Digital literacy for secondary school students: using computer technology to educate about credibility of content online. *Creative Education* [Online], 11(5), 674–692. Available: <https://www.scirp.org/journal/paperinformation?paperid=100127>. [2025, October 30].

Day of Information Disorder: Evidence based solutions for a resilient digital age. 2025, July 8. LIRNEasia [Web log post]. Available: <https://lirneasia.net/2025/07/day-of-information-disorder-evidence-based-solutions-for-a-resilient-digital-age/> [2025, August 25].

Department of Census and Statistics, Sri Lanka. (2024). Computer literacy statistics – 2024 (First six months) [Online]. Available: <https://www.statistics.gov.lk/ComputerLiteracy/StaticInformation/Bulletins/2024FirstSixMonths> [2024, December 10].

Dumitru, E.A. 2020. Testing children and adolescents' ability to identify fake news: A combined design of quasi-experiment and group discussion. *Societies* [online], 10(3), 71. Available: <https://www.mdpi.com/2075-4698/10/3/71> [2025, June 29].

Gradoń, K. T., Holyst, J. A., Moy, W. R., Sienkiewicz, J., & Suhecki, K. 2021. Countering misinformation: A multidisciplinary approach. *Big Data & Society*, 8(1) [Online]. Available: <https://journals.sagepub.com/doi/pdf/10.1177/20539517211013848> [2025, May 20].

Howard, P., Neudert, L.M., Prakash, N., Vosloo, S. & UNICEF Office of Global Insight & Policy. 2021. Digital misinformation, disinformation and children [Online]. Available: <https://www.unicef.org/innocenti/media/856/file/UNICEF-Global-Insight-Digital-Mis-Disinformation-and-Children-2021.pdf> [2025, June 14].

Ireton, C., & Posetti, J. (ed.). 2018. Journalism, “Fake News” & Disinformation: Handbook for Journalism Education and Training [Online]. Paris: UNESCO Publishing. Available: https://webarchive.unesco.org/web/20230930104950/https://en.unesco.org/sites/default/files/journalism_fake_news_disinformation_print_friendly_0.pdf [2025, February 14].

Kirschner, P. A., & De Bruyckere, P. 2017. The myths of the digital native and the multitasker. *Teaching and Teacher education* [Online], Available: <https://konstantinus.com/wp-content/uploads/2016/12/kirschner2017.pdf67>, 135-142 [2025, August 30].

Nature: The digital native is a myth [Online]. (2017, July 27). Available: <https://doi.org/10.1038/547380a>

Ng, W. 2012. Can we teach digital natives' digital literacy? *Computers & Education* [Online], 59(3), 1065–1078. Available: <https://doi.org/10.1016/j.compedu.2012.04.016> [2025, May 15]

Prensky, M. 2001. Digital Natives, Digital Immigrants. *On the Horizon*, 9(5), 3-6. SarvodayaFusion: The official site: Projects [Online]. [n.d]. Available: <https://www.sarvodayafusion.org/s-projects-side-by-side> [2025, May 15]

Seo, H., Erba, J., Altschwager, D., & Geana, M. 2019. Evidence-based digital literacy class for low-income African American older adults. *Journal of Applied Communication Research*, 47(2), 130-152 [Online]. Available: <https://tinyurl.com/2etkzr38>

UNICEF. 2023. Sri Lanka's Multidimensional Poverty Index 2019 results: National and child analysis [Online]. Available: <https://tinyurl.com/5dfrefec> [2025, July 14].

United Nations General Assembly. 2022. Report of the Secretary General: Countering disinformation for the promotion and protection of human rights and fundamental freedoms (12 August 2022). Available: <https://docs.un.org/en/A/77/287> [2025, September 12].

Wardle, C., & Derakhshan, H. 2017. Information disorder: Toward an interdisciplinary framework for research and policy making. Council of Europe [Online]. Available: <https://edoc.coe.int/en/media/7495-information-disorder-toward-an-interdisciplinary-framework-for-research-and-policy-making.html> [2025, July 17].

Wasserman, H (eds.). 2022. Meeting the challenges of information disorder in the Global South [Online] Cape Town: University of Cape Town. 191-192. Available: <https://idrc-crds.ca/en/news/novel-study-reveals-efforts-global-south-counter-information-disorder> [2025, January 10].

World Economic Forum. 2025. Rethinking Media Literacy - A New Ecosystem Model for Information Integrity (Insight Report). [Online]. Available: <https://tinyurl.com/3msysn7u> [2025, August 15]

THE FACTS ABOUT FACT-CHECKING:

Struggles And Possibilities
To Countering Information
Disorder In The Arab World

Arwa Kooli,

Ahmed Ashour,

Zara Schroeder,

Saja Mortada,

Fatima Bani Ahmad,

Menna Elhosary



INTRODUCTION

Misinformation, disinformation, and malinformation are often grouped under the umbrella term “information disorder,” which seeks to capture the diverse dimensions and manifestations of contemporary challenges to information quality. While the scholarly literature on information disorder is extensive in the Global North, research on this phenomenon is also steadily expanding in the Global South (see Wasserman and Madrid-Morales, 2022; Tumber and Waisbord, 2024).

Understanding disinformation mechanisms in the Global South is essential for developing a more holistic and comprehensive framework for practical and effective responses. To explore the drivers of information disorder in Arab Speaking countries, we conducted 127 interviews between April and July 2023 with fact-checkers, journalists, academics, experts, and government representatives as part of a study conducted by the Arab Fact-Checkers Network (AFCN)⁷ (a network launched by The Arab Reporters for Investigative Journalism (ARIJ)⁸ in 2020). The interviews were carried out across 10 Arab countries: Egypt, Iraq, Jordan, Lebanon, Libya, Morocco, Palestine, Syria, Tunisia, and Yemen. Three main drivers of information disorder are brought to the discussion in the region.⁹

Conflicts emerge naturally as the main driver in countries like Palestine, Yemen, Syria, and Lebanon; however, political conflicts and disruptions, as well as political disinformation, are prevalent in both stable and unstable states. In these contexts, political disinformation is often used to manipulate public opinion and fuel conspiracy-driven campaigns.

Lack of access to information and lack of information availability create fertile ground for mis- and disinformation to fill the vacuum. What we called the silent communicator syndrome in AFCN’s study reflected the unwillingness but also the dysfunctionality of official communication to provide the public and journalists with needed information, especially when a crisis breaks out. Social media and traditional media business models also have their share in fuelling information disorder, according to the interviewees. Several under-resourced media organizations, striving to revive audience engagement and compete with the speed and volume of social media, have resorted to increasing their content output despite limited staff. This pressure often leads

to the dissemination of unverified information, a trend also observed in African media landscapes (van Wyk, 2024). Social media has often been portrayed by interviewees as an ideal environment for the spread of false and misleading information, due to the absence of embedded verification mechanisms and its susceptibility to inauthentic activities fuelled by both local and foreign actors; the use of social media widely figures among the seven factors of disinformation resilience identified by Humprecht et al (2020). In addition to partisan media, which is controlled by the government and serves in major cases the interests of the powerful.

While the Arab region has its uniqueness, it shares similarities with other regions in the Global South. The role of technology and new media is also omnipresent, with actors involved in trolling and cyberattacks in Africa, Asia, and Latin America (Wasserman, 2022). The impact on political and electoral processes has also been documented in other regions, including Africa and Latin America (Wasserman, 2022). The latter has lower resistance to misinformation compared to countries in the Global North (Rodríguez-Pérez and García-Vargas, 2021).

According to the interviewees, the final major category of drivers stemmed from legal and regulatory issues. Some interviewees emphasized the importance of legal texts in mitigating the impact of disinformation, while others expressed concern that such laws may restrict freedom of expression in the countries where they are enforced. The role of legal texts remains dependent on the institutional status, where weaknesses are varied and dissimilar to the Global North (Salazar Rebolledo, 2024).

The mechanisms and contexts that allow information disorder to further prevail revolve around political instability and populism, lack of access to information, the growing role of information technology, the lack of procedures to mitigate the impacts of information disorder, and partisan media, among others.

In response to the growing challenges of information disorder, post-publication fact-checking has, in the last decade, grown into a global movement (Graves, 2016; Amazeen, 2017; Lauer and Graves, 2024). A growing number of units and organizations verifying public figures' statements, mainstream, and social media claims have also gained momentum in the Global South.

Post-publication fact-checking is the practice of testing claims against

authoritative sources and collecting evidence to eventually issue a verdict on the degree of accuracy of the claim (Amazeen, 2013). Despite its growth globally, fact-checking is characterized by significant heterogeneity (Cazzamatta et al, 2024). On the other hand, when it is practiced internally, pre-publication fact-checking ensures the publication of accurate and verified information by media organizations.

Fact-checking can also help in tackling the deteriorating quality of information. It is still challenged by the limited access to information, oppressive regimes, partisan media, and the growing role of AI in distorting the facts.

In this chapter, we aim to discuss the role of fact-checking as a form of response to information disorder in the Arab world. While fact-checking both internally and externally has its limitations, it remains an important practice for sorting facts from falsehoods and holding politicians and public figures accountable. It remains also a prominent instrument to help pave the road for information integrity (Laroussi and Laalou, 2025; Oladokun et al, 2024).

EXTERNAL FACT-CHECKING: EDUCATING THE AUDIENCE BUT NOT ABOUT POLITICIANS

What we will refer to here as post-publication is the external fact-checking practiced by organizations that dedicate a section or an entire platform to monitor, detect, and verify claims made in the public sphere, typically issuing a verdict on the veracity of these claims.

Even though fact-checking was presented as a development and expansion of the journalism practice (Cheruiyot and Ferrer-Conill, 2018), as its mission involves evaluating the veracity of public claims, post-publication fact-checkers need a sharp set of principles. The framework in which fact-checkers operate must be protected from bias and partisanship. If an institution is partisan, it cannot, according to IFCN (n.d.) and Arab Fact-Checkers Network (AFCN) (2023b) principles, join those networks, as bias undermines the credibility of its verification work and casts doubt on the integrity of any claim it assesses. Collecting evidence, maintaining transparency in methodology and sources, and applying high standards of cross-checking and documentation enable the public to replicate the process and draw their own conclusions. These standards promote internal accountability and transfer the authority to judge facts to the audience.

Like many other regions in the Global South, post-publication fact-checking practices in the Arab region have expanded significantly. As of 2021, 53 fact-checking organizations, platforms, and initiatives were identified (Ababakirov et al, 2022). Early practices, however, date back to 2012 with “Morsi Meter” in Egypt (Fakida, 2021), followed by projects like “Saheh Khabarak” (Correct Your News) in 2013 in Jordan, and “Da Begad” (Is It Correct?) in 2013 in Egypt, among others (Arab Fact-Checkers Network (AFCN), 2023a). The outbreak of COVID-19 further fuelled the demand for fact-checking, the number of fact-checkers expanded further, and new projects started operating. The AFCN community, which is the largest fact-checking community in the region, has grown from 7 fact-checking projects in 2021 to 36 in 2025, gathering

TABLE 1: INSTITUTIONAL TYPOLOGY OF AFCN FACT-CHECKING COMMUNITY

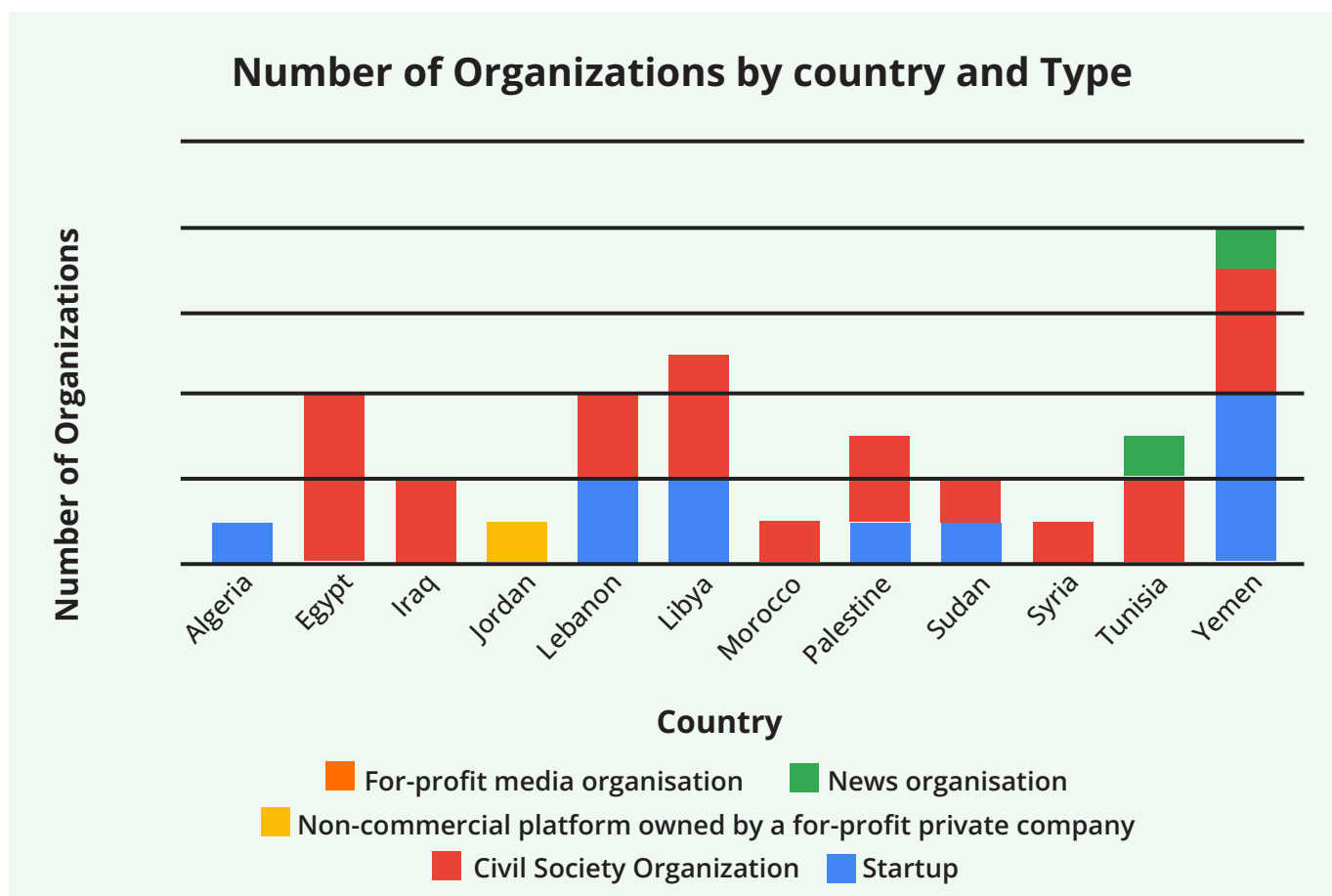
Organization Type	Number of Organizations	Percentage
Civil Society Organization	21	58.33%
Startup	11	30.56%
News organization	2	5.56%
Non-commercial platform owned by a for-profit private company	1	2.78%
For-profit media organization	1	2.78%

(Source: The Authors based on AFCN data)

more than 300 independent fact-checkers from 14 countries in the region. These projects (Table 1) are mostly civil society organizations, followed by fact-checking start-ups. The newsroom model, while dominant in other regions of the globe (see Graves and Cherubini, 2016), is negligible in the Arab region. Only two media organizations created their own fact-checking units.

AFCN’s fact-checking organizations do not fully fit into the classification system of “news organizations, civil society organizations, academic institutions, and independent fact-checking startups” (Westlund et al, 2024, pp. 899-900). There are no organizations affiliated with an academic institution, for instance.

FIGURE 1: AFCN COMMUNITY GEOGRAPHICAL AND ECOSYSTEM DISTRIBUTION



(Source: The Authors based on AFCN data)

Yemen has the highest number of fact-checking organizations or initiatives in the region (Figure 1), where the highest number of fact-checking independent startups were created.

To understand the current state of fact-checking in the Arab region, it is essential to understand the timing of the adoption of post-publication fact-checking. While political fact-checking declined in favour of what Graves et al. (2023, p.1418) call “policing viral misinformation on social networks,” the fact-checking movement in the Global South coincided with the debunking

phase (Graves et al, 2023). The role of Covid-19-related misinformation was vital in the emergence of many fact-checking initiatives and their growing interest in countering misinformation (Atef et al, 2020; Zamit et al, 2020).

The second study, conducted by the AFCN investigating “Information Disorder in the Global South,” investigated both pre-publication and post-publication fact-checking. It surveyed, in 2024, post-publication fact-checkers across the region (n=77) and conducted two focus groups: one with

professional fact-checkers (n=10) and another with managers and editors-in-chief of fact-checking organizations in the region (n=8). According to the survey and focus groups results, participants highlighted their workflow that typically starts with monitoring social media and tracking viral claims to be scrutinized.

However, unlike some of their counterparts around the globe, cleaning’ social media is not driven by a certain business model that relies on the previously well-established third-party agreement with Meta, for instance. Only a handful of organizations in the Arab region are IFCN signatories, which is a requirement in Meta’s phasing-out program, after Meta decided to replace fact-checkers with community notes from the audience and started applying that in the United States (see Kaplan, 2025). In contrast, Arab fact-checkers are not only doing the policing of social media for free, but their content moderation activities, prominent in non-WEIRD countries (Vinhas and Bastos, 2023), would not obviously result in removing or labelling content except on a few occasions when the organization has connections with the giant tech companies.

Unexpectedly, holding politicians accountable is not a priority for fact-checkers in the Arab region. Most respondents underlined promoting media

“

To understand the current state of fact-checking in the Arab region, it is essential to understand the timing of the adoption of post-publication fact-checking.

”

literacy (n=39, 50.7%) as a priority action against information disorder, and fact-checking efforts themselves came second with (n=32, 41.5%). Holding politicians and public figures accountable for spreading mis- and disinformation was considered the least important, with only (n=6, 7.8%) of the respondents indicating it.

This orientation is best understood through contextual lenses. Fact-checking prominent figures rarely goes unnoticed, and as such, it is frequently avoided for safety reasons; prioritizing “(...) human interest topics focusing on rumours, hoaxes, myths and urban tales” (Fakida, 2021, p.135). This is not different from how fact-checkers behave in a similar context in some African countries, where they seek refuge in avoiding sensitive topics and resort to self-censorship (Badji et al., 2024).

The claim selection process also assigns greater weight to virality, as evidenced by the typical sentence that opens most of the fact checks in Arabic: “Many social media accounts circulated...”. Still, there are a few fact-checking organizations in the Arab region that give priority to politicians’ claims and focus on holding public figures accountable.

While they do not significantly educate the public about politicians, public figures, or their accountability, the fact-checking initiatives surveyed make substantial contributions to educating the public on how to consume information and how to verify circulated content. The intersection between media literacy and fact-checking is particularly evident (Mesquita et al., 2024). Media and Information Literacy is a vital practice that helps build defensive mechanisms against mis- and disinformation, compared to the direct targeting approach employed by fact-checking. It extends the life of fact checks, boosting their impact on the long term and audience’s trust (Çömlekçi, 2022). While the efficiency of fact-checking is debatable, media literacy not only complements efforts, but it also enables fact-checking organizations to secure possible additional funds (Tompkins, 2020), especially with the sustainability challenges that fact-checkers face (Zamit et al, 2020; Kooli et al, 2022; Badji et al, 2024).

INTERNAL FACT-CHECKING: NO SHORTCUT TO QUALITY AND PUBLIC TRUST

Pre-publication or internal fact-checking comprises a different logic and

methodology compared to post-publication fact-checking. It is a practice that dates to the 1920s in the United States (Graves and Amazeen, 2019). Due to limited resources and fast-paced decision-making dynamics in newsrooms, it has not been systematically implemented across the journalism industry. Verification and accuracy, on the other hand, are embedded practices in the journalism production process. Both global and internal codes of ethics prioritize accuracy and prohibit journalists from disseminating false or fabricated information through broadcast or publication; still, in today's information-saturated world, verification is more crucial than ever (Kabashi, 2023). With technological advancements and the rise of social media, speed has become a dominant priority in news production. Although journalistic tradition holds that it is better to be late than to be wrong, the digital era has intensified pressure on journalists to favour exclusivity over accuracy. This is further compounded by newsroom understaffing and heightened competition (Korunka et al, 2015).

In-depth journalism, on the other hand, requires more attention to detail and a factual basis. The systematic review of data journalism and investigative pieces before publication helps to ensure the quality and accuracy of produced pieces (Gupta, 2023). For newsrooms and media organizations around the globe, internal fact-checking is a strategic necessity for ensuring high-quality journalism and building public trust.

In response to the need for providing content both quickly and accurately, while ensuring a baseline quality standard for in-depth journalism pieces, some newsrooms have begun hiring dedicated fact-checkers to verify content before it is published (Arab Fact-Checkers Network (AFCN), 2022). While staff cuts and limited resources urged news organizations to cut fact-checkers out, in recent years internal fact-checking and especially in-depth pieces is no longer perceived as the sole responsibility of reporters or editors; the need for an independent fact-checker not involved in writing or editing the story, ensures a sharper focus and a fresh perspective to systematically review the journalistic pieces (Arab Fact-Checkers Network (AFCN), 2022).

In other words, while reporters are still responsible for the accuracy and the quality of their content, as in the case of The Washington Post, a tendency to refer to the services of an independent fact-checker is also adopted by other organizations. For instance, The New Yorker employed 16 full-time fact-

checkers in 2003, who meticulously reviewed the long-form articles before publication (Smith, 2004). Likewise, Der Spiegel had one of the largest fact-checking units in Europe, with about 70 staff members in 2017, dedicated to examining every word and image that appeared in its investigative pieces (Smith, 2004).

Internal fact-checking methodologies also differ from one organization to another. In 2004, the Swedish public broadcaster SVT, through its investigative program “Uppdrag Granskning” (Mission: Investigation), developed a pre-publication verification methodology known as the “Three Verification Points”. This system breaks the fact-checking process down into three main phases: the early investigation phase, the fact-checking phase upon submission of the draft, and a thorough line-by-line verification phase (Hanson, 2021). The line-by-line phase itself is a highly detailed process, divided into three steps: verification: the fact-checker reviews every fact and its source to confirm its accuracy; evaluation: the fact-checker assesses the reliability of the sources and the strength of the information based on those sources; documentation: all sources, documents, and verification methods used are carefully preserved, whether the information is confirmed or corrected. Nevertheless, it is pertinent to draw clear limitations for the role of a fact-checker. While the editor is responsible for overall editorial oversight, the fact-checker must work independently, without intervening in editing the story, changing headlines, or requesting additional sources unless it directly relates to fact verification (Arab Fact-Checkers Network (AFCN), 2022). Drawing the lines between the roles of each team member is left for the organization to determine when drafting its internal fact-checking methodology.

One of the clearest signs of the effectiveness and efficiency of pre-publication fact-checking is a drop in the number of corrections a news outlet has to issue later. Several news organizations have strengthened their fact-checking processes in recent years, even without having a fully independent department, and have witnessed a noticeable decline in both typographical errors and factual mistakes flagged by readers (Smith, 2004). Nonetheless, outlets that rush to publish without sufficient review have witnessed an increase in publicly discussed errors, which has damaged their credibility (Kabashi, 2023).

The effectiveness of pre-publication fact-checking is measurable in trust

levels. Outlets that invested in strong fact-checking procedures have earned higher levels of reader and viewer trust compared to those that did not. As such, verification is an instrument of building trust (Bennke, 2023) Over the past decade, pre-publication fact-checking has gradually become more important in Arab media. In recent years, a growing number of media organizations started to incorporate systematic fact-checking, particularly in investigative journalism and in-depth reporting in their workflow (ARIJ's AFCA Team, 2024).

Traditional media deployed no dedicated fact-checking departments inside newsrooms; it was largely up to individual editors to exercise their judgment (International Development Research Centre (IDRC), n.d.). In many cases, the primary concern was avoiding conflict with authorities rather than ensuring

“
In many cases, the primary concern was avoiding conflict with authorities rather than ensuring every detail was correct.

every detail was correct. State-owned media often simply echoed official government statements. Within this environment, accuracy sometimes took a backseat to censorship and official narratives.

”

The turmoil that swept the region in 2011, driven by regime changes, provided fertile ground for the spread of misinformation, rumours, and propaganda (El Gody, 2023). However, the loosening of restrictions in countries like Tunisia allowed dozens of independent media outlets to emerge, many of which were eager to build their credibility by adhering to professional standards (El Gody, 2023). In other parts of the Arab region, new digital platforms, such as Mada Masr and Raseef22, were launched. Both are committed to publishing accurate, in-depth journalism free from government control. As the political situation in Egypt grew increasingly tense, the team at Mada Masr doubled down on its commitment to fact-checking and professional standards, such as balance and accuracy, seeing it as a critical part of navigating an increasingly repressive environment (Leih, 2021).

Mada Masr's editorial process involved cross-checking information with multiple sources. For editors to independently verify the facts, reporters were required to provide supporting materials, such as documents or recorded interviews (Wordcrunch, 2019).

Around 2019, Arab media began to take pre-publication fact-checking more

seriously. ARIJ, which has trained thousands of journalists in investigative techniques, took a remarkable step in 2020 by advancing its pre-publication fact-checking as a formal part of its editorial workflow (Arab Reporters for Investigative Journalism (ARIJ), 2020). Fact-checkers hired by ARIJ work on ensuring that the content is verified against primary sources (Arab Reporters for Investigative Journalism (Arab Reporters for Investigative Journalism (ARIJ), n.d.a).

Despite these efforts, many traditional Arab media outlets still lack clear and formal protocols for pre-publication fact-checking. In 2024, AFCN explored pre-publication fact-checking and surveyed (n=95) journalists working across (n=79) Arab media organizations, in addition to a focus group gathering (n=3) media organizations. Although 59% of respondents said their institutions conducted some form of pre-publication fact-checking, around 27% did not provide a consistent definition of what the process entailed. The survey also revealed confusion around the true purpose of pre-publication verification, with some journalists mixing it up with other editorial tasks such as proofreading or checking for compliance with internal policies.

These results were not especially surprising, as 60% of respondents said their organizations lacked documented procedures for pre-publication fact-checking. Even among the 39% who claimed to have clear and documented protocols, many struggled to accurately explain them when asked to do so in just a sentence or two.

There was also widespread confusion about who should be responsible for the pre-publication fact-checking: 88% of participants said that it was up to editors-in-chief, journalists, or general editors. Only 28% mentioned that internal fact-checking should primarily be the responsibility of an independent fact-checker. Just 22% of respondents said that every piece of content produced by their organizations went through a detailed, paragraph-by-paragraph fact-checking.

The adoption of pre-publication fact-checking is dependent on the size, resources, and especially the media outlet. Integrating practices of internal verification in different formats and styles still helps organizations to avoid pitfalls. It is notable, nonetheless, that a growing number of Arab media organizations had introduced institutional fact-checking policies, written

protocols, or at least dedicated time and staff to fact-checking before publishing.

Daraj, an independent media digital platform launched in 2017 in Lebanon, operated with a small team that often struggled to deploy pre-publication fact-checking to cross-border stories with limited resources (Oliver, 2023). By 2024, however, Daraj was selected as one of three Arab institutions chosen to participate in the AFCN project, aimed at incorporating pre-publication verification into editorial workflows (Arab Fact-Checkers Network (AFCN), 2024). Over several months, an independent fact-checker worked closely with the Daraj team to develop a detailed, customized methodology for internal verification. Similarly, in 2024, in Egypt, Al-Masry Al-Youm newspaper established a Fact-Checking and Data Journalism Department as part of a similar initiative (Arab Fact-Checkers Network (AFCN), 2024). This signals a gradual shift in newsroom mindsets, moving toward making pre-publication fact-checking a built-in part of journalistic practice rather than a luxury. Transitioning to a systematic verification of content before publication entails several challenges. First, financial resources are often too limited to allow for hiring dedicated fact-checkers to review all journalistic content. In many newsrooms, speed is prioritized over accuracy, with breaking the story quickly taking precedence over carefully verifying details. Many journalists surveyed by AFCN admitted that they lack the know-how to carry out effective pre-publication fact-checking and acknowledged the urgent need to build their skills in this area.

Another major obstacle is the weak enforcement of freedom of information laws across the Arab world, which makes it difficult for journalists and fact-checkers to access reliable sources. Many Arab countries either lack strong legislation granting journalists the right to information, or where laws do exist, they are often poorly implemented due to institutional weaknesses. As a result, journalists are sometimes forced to rely on questionable and unreliable sources, which compromise the accuracy of their reports and undermine the credibility of the media and its role as a watchdog. However, in some cases, media organizations may deliberately publish false information to advance a specific agenda. In such instances, even a well-designed internal verification process cannot fully counteract institutional partisanship.

Regardless of partisan organizations, it is crucial to establish strong

internal fact-checking systems as a core part of countering information disorder. Building the capacities of specialized fact-checkers and the fact-checking skills of journalists, setting up dedicated fact-checking units within newsrooms, and adhering to an internal verification methodology not only reinforces the credibility of the media organizations or protects the journalists from heightened legal and security risks, but they also cut down on misinformation sources.

AI AND FACT-CHECKING: GENERATING PROBLEMS AND SOLUTIONS

The fact-checking movement in the Arab World and other regions of the Global South is challenged by limited access to information, partisan media, and a restrictive political environment in several places. Amid the growing need to navigate information disorder, a reflective pause is needed to further reframe the use of Artificial Intelligence (AI) to address the complexity of AI use and adoption, but also the pitfalls.

When used in journalism, AI introduces new risks and exacerbates existing challenges, particularly in regions such as the Arab World and Africa, where infrastructure, regulatory frameworks, and digital literacy vary significantly (Harb and Arafta, 2024; Sun et al, 2024). On the other hand, AI is increasingly central to combating mis- and disinformation around the world. The gradual adoption of automated fact-checking and the central role of AI in deep-fake detection are inevitable in the fight for truth.

One of the most pressing contradictions is that the same AI tools designed to fight disinformation can also be used to create it. Generative AI models like ChatGPT, Gemini, Midjourney, and voice-cloning technologies can generate convincing fake news articles, images, and audio clips. This has been weaponized in political contexts, such as during the 2024 Indian elections, where AI-generated videos misrepresented political candidates (Bond, 2021). In Nigeria, during the 2023 elections, AI-generated fake news stories and doctored videos spread on WhatsApp and Facebook, influencing public perception in areas with low media literacy (Salako, 2023). In the Arab World, content generated by artificial intelligence has been maliciously employed to deny the genocide in Gaza, contributing to the propagation of false narratives and fuelling propaganda (Klepper, 2023). For instance, hashtags such as #Gazawood and #Playwood have circulated claims that the deaths of civilians,

including children, were fabricated or staged using AI-generated imagery.

AI, however, can help to some extent to correctly depict generated content. With the relative efficiency of these tools, they are often developed in English and for Western contexts, making them less effective in multilingual and culturally nuanced environments (Prat et al., 2024). Additionally, AI presents profound challenges, especially around ethics, transparency, legality, and public trust, challenges that are particularly complex in the Global South contexts, where digital ecosystems are still evolving. A dilemma that particularly journalists face, as journalists are left to navigate AI adoption on their own due to a lack of formal newsroom guidance, resulting in usage often driven by individual journalists rather than agreed organizational strategies and approaches (Radcliffe, 2025).

Media houses in countries like Kenya, South Africa, and Ghana are adopting AI-powered newsroom tools without robust training or understanding of potential biases embedded in these systems (Ogola, 2023). When under-resourced journalists use AI as an investigative shortcut, the line between credible reporting and algorithmically generated misinformation can blur. A few media organizations in the MENA region, for instance, began exploring AI's potential in Arabic. Among them, ARIJ launched an AI strategy in late 2022 aimed at supporting small and medium newsrooms (Arab Reporters

“

When under-resourced journalists use AI as an investigative shortcut, the line between credible reporting and algorithmically generated misinformation can blur.

”

for Investigative Journalism (Arab Reporters for Investigative Journalism (ARIJ), n.d.b). ARIJ focused on raising AI awareness and literacy within its network of journalists and fact-checkers (JournalismAI, n.d.). ARIJ's efforts to integrate AI into journalism in the Arab world faced numerous challenges, including the complexities of the Arabic language, limited resources and expertise, widespread skepticism and resistance toward AI, concerns over its role in spreading misinformation, and ethical dilemmas. Ongoing regional crises also meant AI was often deprioritized by media organizations and journalists.

AI has enabled faster, large-scale fact-checking through tools like Google's Fact Check Explorer, Full Fact, and Africa Check tools (Dudfield, 2020). These systems analyze content, compare it with verified databases, and flag

inaccuracies. In journalism, AI is also assisting in investigative reporting (Hu & Downie, 2024), such as using natural language processing (NLP) to sift through leaked documents or detect coordinated disinformation campaigns (Hu and Downie, 2024). At the same time, AI can be leveraged to automate components of the fact-checking process. Examples from the Global South show a gradual adoption of AI in assisting fact-checkers' daily tasks. One of the solutions introduced by AFCN is localizing and developing Arabic-language tools, including the AFCN fact-checking chatbot and Full Fact's AI tools. Between April 2023 and June 2025, the AFCN chatbot fact-checked 2550 claims, with AI successfully verifying 12.05% of them.

A critical global issue is the “black box” nature of AI systems, which often make decisions without clear human oversight or explainability (Hassija et al., 2024). This lack of transparency creates ethical problems around accountability (Al-Dulaimi & Mohammed, 2025). For instance, identifying who is held responsible when AI systems wrongly flag factual content or fail to catch dangerous misinformation.

Algorithms may misclassify nuanced content, particularly when cultural context is misunderstood. Over-reliance on AI can also reduce journalistic accountability. Media houses in Africa are adopting AI tools without adequate training (Nguyen & Cecchini, 2021), potentially blurring the lines between credible reporting and AI-generated misinformation.

As AI becomes more prevalent in content moderation and fact-checking, public resistance is also rising. Around the world, AI is often seen as a tool for censorship, especially when it operates without human accountability. In African nations, where foreign-developed AI tools dominate the market, scepticism is intensified by colonial and neocolonial histories of information control.

In many citizens in countries, AI-powered content regulation is viewed as another form of external influence, often unaccountable, opaque, and culturally insensitive (Pasipamire & Muroyiwa, 2024). This resistance undermines AI's potential benefits and makes the adoption of trustworthy AI tools more difficult.

AI-driven interventions can be perceived as neocolonial, top-down efforts to control public discourse, particularly when they suppress indigenous languages or fail to recognize cultural context.

Despite technical, linguistic, ethical, and political challenges, several initiatives demonstrate AI's potential in the field. However, risks remain high, especially regarding misinformation, algorithmic bias, and accountability, underscoring the need for localized, ethical, and inclusive AI strategies. Legal frameworks around AI and misinformation are still developing globally (European Parliament, 2023), and enforcement is weak in Global South countries.

CONCLUSION

While the Arab region is the main case study presented in this chapter, we tried to draw on the similarities that it shares with other regions in other regions in the Global South, and particularly Africa. We identified similar perspectives regarding post-publication fact-checking, particularly its tendency to avoid political claims while prioritizing viral social media content. We also examined the practices of pre-publication fact-checking, exploring how it has evolved but remains unevenly adopted across different organizations. Finally, we highlighted the challenges and opportunities that artificial intelligence presents for fact-checkers and journalists in the region and beyond.

In the fight against information disorder, both pre- and post-publication fact-checking stands out as a direct primary attack line. The reliability of content produced by news media, especially in-depth pieces when internal verification is practiced, and the effectiveness of post-publication fact-checking in holding public figures accountable and spreading awareness among the audiences are pertinent.

However, the efficiency of fact-checking is extremely dependent on the context. When the media is partisan and chooses a biased approach to news and information, it is hard for pre-publication fact-checking to be well practiced. This also applies to ecosystems where those in power remain untouchable, forcing fact-checkers to avoid scrutinizing claims made by political leaders, prominent politicians, or influential business figures. Additionally, the overwhelming volume of false and misleading content, AI-generated images, videos, audio, and text, combined with declining public trust in the media and rising news avoidance, compels fact-checkers to seek coping strategies that enable them to engage citizens and support informed decision-making in the Arab region.

The efficiency of fact-checking is dependent on the context, but it is also very

dependent on how the audience consumes and interacts with information. For that, complementing fact-checking with a package of literacies that helps audiences navigate the complexity of mis- and disinformation can include not only media literacy, but also AI literacy and Algorithm literacy, to name a few. Fact-checking is also challenged today to an unprecedented extent. The lack of resources and deteriorating funds endanger the sustainability of fact-checking activities. The term fact-checking has been stripped of its original meaning, politicized, and at times misused. The survival of fact-checking in the Arab region and beyond is not only dependent on financial resources and the international political mood, but also on its internal dynamics and ability to survive, especially in contexts where it was not always welcomed in the first place. It would be of interest in the future to study the resilience of fact-checking in the region, including any coping mechanisms and the impact and possible contribution of fact-checking in flattening the curve of information disorder.

References

- Ababakirov, A., Ashour, A., Bandaranayake, R., De Silva, R., Diouf, D., & Wasserman, H. 2022. *Meeting the challenges of information disorder in the Global South*. Ottawa: International Development Research Centre. Available: <http://hdl.handle.net/10625/60954>
- Al-Dulaimi, A.O.M. & Mohammed, M.A.-A.W. 2025. Legal responsibility for errors caused by artificial intelligence (AI) in the public sector. *International Journal of Law and Management*. Advance online publication. <https://doi.org/10.1108/IJL-MA-08-2024-0295>
- Amazeen, M.A. 2013. *Making a difference: A critical assessment of fact-checking in 2012*. Washington, DC: New America Foundation, Media Policy Initiative Research Paper.
- Amazeen, M.A. 2020. Journalistic interventions: The structural factors affecting the global emergence of fact-checking. *Journalism*, 21(1):95–111. <https://doi.org/10.1177/1464884917730217>

Arab Fact-Checkers Network (AFCN). 2022. *Guidance manual for pre-publication fact-checking* [Dalīl istirshādī li-tadqīq al-ma'lūmāt ma qabl al-nashr]. 10 November. [Online]. Available: <https://sa7.arabfcn.net/prepublication-guide/?utm> [2025, June 14].

Arab Fact-Checkers Network (AFCN). 2023a. *Fact-checking information in times of crises*. [Tadqeeq al-ma'lūmat fi zaman al-azamat]. [Online]. Available: <https://tinyurl.com/3zdfdzdy> [2025, May 12].

Arab Fact-Checkers Network (AFCN). 2023b. *Code of Principles and Guidelines*. [Online]. Available: <https://arabfcn.net/wp-content/code-of-principles-en/> [2025, September 4].

Arab Fact-Checkers Network (AFCN). 2024. *A fact-checker in an independent Arabic newsroom: Pre-publication fact-checking project by AFCN. 11 August*. [Online]. Available: <https://arabfcn.net/en/events/2024/08/11/a-fact-checker-in-an-independent-arabic-newsroom-pre-publication-fact-checking-project-by-afc/> [2025, April 26].

Arab Reporters for Investigative Journalism (ARIJ). n.d.a. *Our approach*. [Online]. Available: <https://en.arij.net/our-approach/> [2025, April 26].

Arab Reporters for Investigative Journalism (ARIJ). n.d.b. *ARIJ Strategy Playbook*. [Online]. Available: <https://arij.net/ai/playbook/arij-strategy/index.html> [2025, July 7].

Arab Reporters for Investigative Journalism (ARIJ). 2020. *ARIJ fact-checking workshop: A new thinking methodology in the Arab world* [Warshat ARIJ litadqiq al-haqā'iq: Manhajyyat tafkir jadida fi al-'alam al-'Arabi]. [Online]. Available: <https://tinyurl.com/23yd7daa> [2025, November 16].

ARIJ's AFCN team. 2024. *Challenges in implementing pre-publication fact-checking in the Arab world*. [Online]. Available: <https://idrc.sun.ac.za/challenges-in-implementing-pre-publication-fact-checking-in-the-arab-world/> [2025, November 16].

Atef, N., El-Mahdi, B. & Mostafa, B. 2020. *Combating media disinformation: Fake news in the Arab world – COVID-19 pandemic as a model*. ARIJ – Arab Reporters for Investigative Journalism Network in partnership with Friedrich Naumann Foundation for Freedom. Available: <https://en.arij.net/materials/arij-publishes-a-policy-paper-combating-disinformation-in-the-arab-world-the-covid-19-pandemic-as-a-case-study/>

Badji, S., Orgeret, K. & Mutsvairo, B. 2024. An exploratory study of fact-checking practices in conflict and authoritarian contexts. *Media and Communication*, 12: article no: 8698. <https://doi.org/10.17645/mac.8698>

- Bennke, J. 2023. Media of verification: An epistemological framework for trust in a digital society. *communication +1*, 10(1). <https://doi.org/10.7275/cpo.1878>
- Bond, S. 2021. How AI deepfakes polluted elections in 2024. In *NPR*, 21 December. [Online]. Available: <https://www.npr.org/2024/12/21/nx-s1-5220301/deepfakes-memes-artificial-intelligence-elections> [2025, April 24].
- Cazzamatta, R., Graves, L. & Lauer, L. 2024. Editorial: Fact-checkers around the world—Regional, comparative, and institutional perspectives. *Media and Communication*. <https://doi.org/10.17645/mac.i458>
- Cheruiyot, D. & Ferrer-Conill, R. 2018. “Fact-Checking Africa”: Epistemologies, data and the expansion of journalistic discourse. *Digital Journalism*, 6(8):964–975. <https://doi.org/10.1080/21670811.2018.1493940>
- Çömlekçi, M.F. 2022. Why do fact-checking organizations go beyond fact-checking? A leap toward media and information literacy education. *International Journal of Communication*, 16:4563–4583. Available: <https://ijoc.org/index.php/ijoc/article/view/19728>
- Dudfield, A. 2020. How we’re using artificial intelligence to scale up global fact-checking. *Africa Check*, 30 July. [Online]. Available: <https://africacheck.org/fact-checks/blog/how-were-using-artificial-intelligence-scale-global-fact-checking> [2025, April 24].
- El Gody, A. 2023. Issue Introduction: Decoding the Echo Chamber—The Proliferation and Impact of Fake News in Arab Media Post-Arab Spring. *Arab Media & Society*, 17 August. [Online]. Available: <https://www.arabmediasociety.com/issue-introduction-decoding-the-echo-chamber-the-proliferation-and-impact-of-fake-news-in-arab-media-post-arab-spring/> [2025, April 26].
- European Parliament. 2023. *EU AI Act: First regulation on artificial intelligence*. 14 June. [Online]. Available: <https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence> [2025, February 19].
- Fakida, A. 2021. Political fact-checking in the Middle East: What news can be verified in the Arab world? *Open Information Science*, 5(1):124–139. <https://doi.org/10.1515/opis-2020-0117>
- Graves, L. 2016. Boundaries Not Drawn: Mapping the institutional roots of the global fact-checking movement. *Journalism Studies*, 19(5):613–631. Available: <https://doi.org/10.1080/1461670X.2016.1196602>

Graves, L. & Cherubini, F. 2016. *The rise of fact-checking sites in Europe*. Oxford: Oxford University, Reuters Institute.

Graves, L. & Amazeen, M. 2019. Fact-Checking as Idea and Practice in Journalism. *Oxford Research Encyclopedia of Communication*. [Online]. Available: <https://oxfordre.com/communication/view/10.1093/acrefore/9780190228613.001.0001/acrefore-9780190228613-e-808> [2025, November 18].

Graves, L., Bélair-Gagnon, V. & Larsen, R. 2023. From Public Reason to Public Health: Professional Implications of the “Debunking Turn” in the Global Fact-Checking Field. *Digital Journalism*, 1–20. Available: <https://doi.org/10.1080/21670811.2023.2218454> [2025, July 22].

Gupta, S.S. 2023. Exploring the investigative journalism: Uncovering truths. In Ghose, S. & Gupta, S.S. (eds.). *History of journalism*. Rev. ed. New Delhi: Wisdom Press, pp. 41–48.

Hanson, N. 2021. A guide to fact-checking investigative stories. In *Global Investigative Journalism Network*, 3 November. [Online]. Available: <https://gijn.org/resource/guide-to-fact-checking-investigative-stories> [2025, April 26].

Harb, Z. & Arafat, R. 2024. The Adoption of Artificial Intelligence Technologies in Arab Newsrooms: Potentials and Challenges. *Emerging Media*, 2(3):371–381. <https://doi.org/10.1177/27523543241291068>

Hassija, V., Chamola, V. & Mahapatra, A. 2024. Interpreting Black-Box Models: A Review on Explainable Artificial Intelligence. *Cognitive Computation*, 16:45–74. <https://doi.org/10.1007/s12559-023-10179-8>

Hu, C. & Downie, A. 2024. How is AI being used in journalism? *IBM*, 25 November. [Online]. Available: <https://www.ibm.com/think/insights/ai-in-journalism> [2025, April 24].

International Development Research Centre (IDRC). n.d. *Challenges in implementing pre-publication fact-checking in the Arab world*. [Online]. Available: <https://idrc.sun.ac.za/challenges-in-implementing-pre-publication-fact-checking-in-the-arab-world> [2025, April 26].

Humprecht, E., Esser, F. & Van Aelst, P. 2020. Resilience to Online Disinformation: A Framework for Cross-National Comparative Research. *The International Journal of Press/Politics*, 25(3):493–516. <https://doi.org/10.1177/1940161219900126>

International Fact-Checking Network (IFCN). n.d. *The commitments of the Code of Principles*. In *IFCN Code of Principles (Poynter)*. [Online]. Available: <https://ifcn-codeofprinciples.poynter.org/the-commitments> [2025, September 4].

JournalismAI. n.d. *JournalismAI Academy MENA*. [Online]. Available: <https://www.journalismai.info/programmes/academy-mena> [2025, July 7].

Kabashi, O. 2023. Exploring information accuracy: The overlooked mandate in newsrooms [At-taḥaqquq min daqqat al-ma'lūmāt: al-farīdah al-ghā'ibah fī ghuraf al-akhbār]. *Al Jazeera Net*, 3 January. [Online]. Available: <https://www.aljazeera.net/blogs/2023/1/3/رابخألأ-فرغ-ي-ف-تبئأغلأ-قضيرفلأ-مؤولعملأ-ةؤد-نم-ق-قحتلأ> [2025, April 26].

Kaplan, J. 2025. More speech and fewer mistakes. *Meta Newsroom*, 7 January. [Online]. Available: <https://about.fb.com/news/2025/01/meta-more-speech-fewer-mistakes/> [2025, September 4].

Klepper, D. 2023. Fake babies, real horror: Deepfakes from the Gaza war increase fears about AI's power to mislead. *AP News*, 28 November. [Online]. Available: <https://apnews.com/article/artificial-intelligence-hamas-israel-misinformation-ai-gaza-a1bb303b637ffbbb9cbc3aa1e000db47> [2025, April 26].

Kooli, A., Toumi, I. & Zamit, F. 2022. Fact-checking in Tunisia: Newsroom Organization and Methods in Action [Les médias de vérification en Tunisie: organisation des rédactions et méthodes à l'œuvre]. *Les Cahiers du journalisme- Recherches*, 2(8-9):R173–R186. [https://doi.org/10.31188/CaJsm.2\(8-9\).2022.R173](https://doi.org/10.31188/CaJsm.2(8-9).2022.R173)

Korunka, C., Kubicek, B., Paškvan, M. & Ulferts, H. 2015. Changes in work intensification and intensified learning: Challenge or hindrance demands? *Journal of Managerial Psychology*, 30(7):786–800. <https://doi.org/10.1108/jmp-02-2013-0065>

Laroussi, Z. & Laalou, A. 2025. The state of fact-checking in Morocco: From information disorder to information integrity. *Studies in Media, Journalism, and Communications*, 4(1):9–XX. <https://doi.org/10.32996/smjc>

Lauer, L. & Graves, L. 2024. How to grow a transnational field: A network analysis of the global fact-checking movement. *New Media & Society*, 27(6):3505–3523. <https://doi.org/10.1177/14614448241227856>

Leihs, N. 2021. A Case Study: Mada Masr—A Progressive Voice in Egypt and Beyond. *Media and Communication*, 9(4):275–285. <https://doi.org/10.17645/mac.v9i4.4492>

Mesquita, L., Maneta, M. & Brites, M.J. 2024. Beyond verification: the evolving role of fact-checking organisations in media literacy education for youth. *Media and Communication*, 12: article 8690. <https://doi.org/10.17645/mac.8690>

Nguyen, T. & Cecchini, S. 2021. Countering COVID-19 Misinformation in Africa. *Think Global Health*, 23 July. [Online]. Available: <https://www.thinkglobalhealth.org/article/countering-covid-19-misinformation-africa> [2025, April 24].

- Ogola, G. 2023. *AI, journalism, and public interest media in Africa*. International Media Support. [Online]. Available: <https://www.mediasupport.org/wp-content/uploads/2023/06/AI-Africa-Report-2023R-Double-Spread-View.pdf> [2025, April 24].
- Oladokun, B.D., Aruwa, J.E., Ottah, G.A. & Ajani, Y.A. 2024. Misinformation and disinformation in the era of social media: The need for fact-checking skills. *Journal of Information and Knowledge*, 61(1):1–7. <https://doi.org/10.17821/srels/2024/v61i1/171266>
- Oliver, L. 2023. How News Site Daraj Reports on Women’s Rights and Corruption across the Middle East. Global Investigative Journalism Network, 6 July. [Online]. Available: <https://reutersinstitute.politics.ox.ac.uk/news/how-news-site-daraj-reports-womens-rights-and-corruption-across-middle-east> [2025, April 26].
- Pasipamire, N. & Muroyiwa, A. 2024. Navigating algorithm bias in AI: Ensuring fairness and trust in Africa. *Frontiers in Research Metrics and Analytics*, 9:1486600. <https://doi.org/10.3389/frma.2024.1486600>
- Prat, L.D., Lucas, O.N., Golias, C. & Lewis, M. 2024. Decolonizing LLMs: An Ethnographic Framework for AI in African Contexts. *Ethnographic Praxis in Industry Conference Proceedings*, 2024(1):46–85. <https://doi.org/10.1111/epic.12196>
- Radcliffe, D. 2025. *Journalism in the AI Era: Opportunities and Challenges in the Global South*. Thomson Reuters Foundation, January. [Online]. Available: <https://www.trust.org/wp-content/uploads/2025/01/TRF-Insights-Journalism-in-the-AI-Era.pdf> [2025, June 23].
- Rodríguez-Pérez, C. & García-Vargas, G.R. 2021. Understanding which factors promote exposure to online disinformation. In López García, G. et al. (eds.). *Politics of disinformation: The influence of fake news on the public sphere*, pp. 173–186. John Wiley & Sons.
- Salazar Rebolledo, G. 2024. Rethinking disinformation for the Global South: Towards a particular research agenda. In Echeverría, M. & García Santamaría, S. (eds.). *State-sponsored disinformation around the globe: How politicians deceive their citizens*, pp. 38–48. New York: Routledge. <https://doi.org/10.4324/9781032632940-4>
- Smith, S.H. 2004. *The fact checker’s bible*. New York: Anchor Books.
- Sun, M., Liu, J. & Lu, J. 2024. Digital literacy in Africa: Exploring its relationship with infrastructure, policy, and social inequality. *African Journalism Studies*. Advance online publication. <https://doi.org/10.1080/23743670.2024.2329705>
- Tompkins, A. 2020. Media and information literacy programs provide fact-checking skills. In *Deutsche Welle*, 17 July. [Online]. Available: <https://www.dw.com/en/media-and-information-literacy-programs-provide-fact-checking-skills/a-54036475> [2025, July 13].

Tumber, H. & Waisbord, S. (eds.). 2024. *The Routledge companion to media disinformation and populism*. London: Routledge.

van Wyk, C. 2024. The Influence of Dedicated Fact-Checking on Journalism Practice in South Africa. *African Journalism Studies*, 45(2):77–93. <https://doi.org/10.1080/23743670.2024.2374401>

Vinhas, O. & Bastos, M.T. 2023. The WEIRD Governance of Fact-Checking and the Politics of Content Moderation. *New Media and Society*. <https://doi.org/10.1177/14614448231213942>

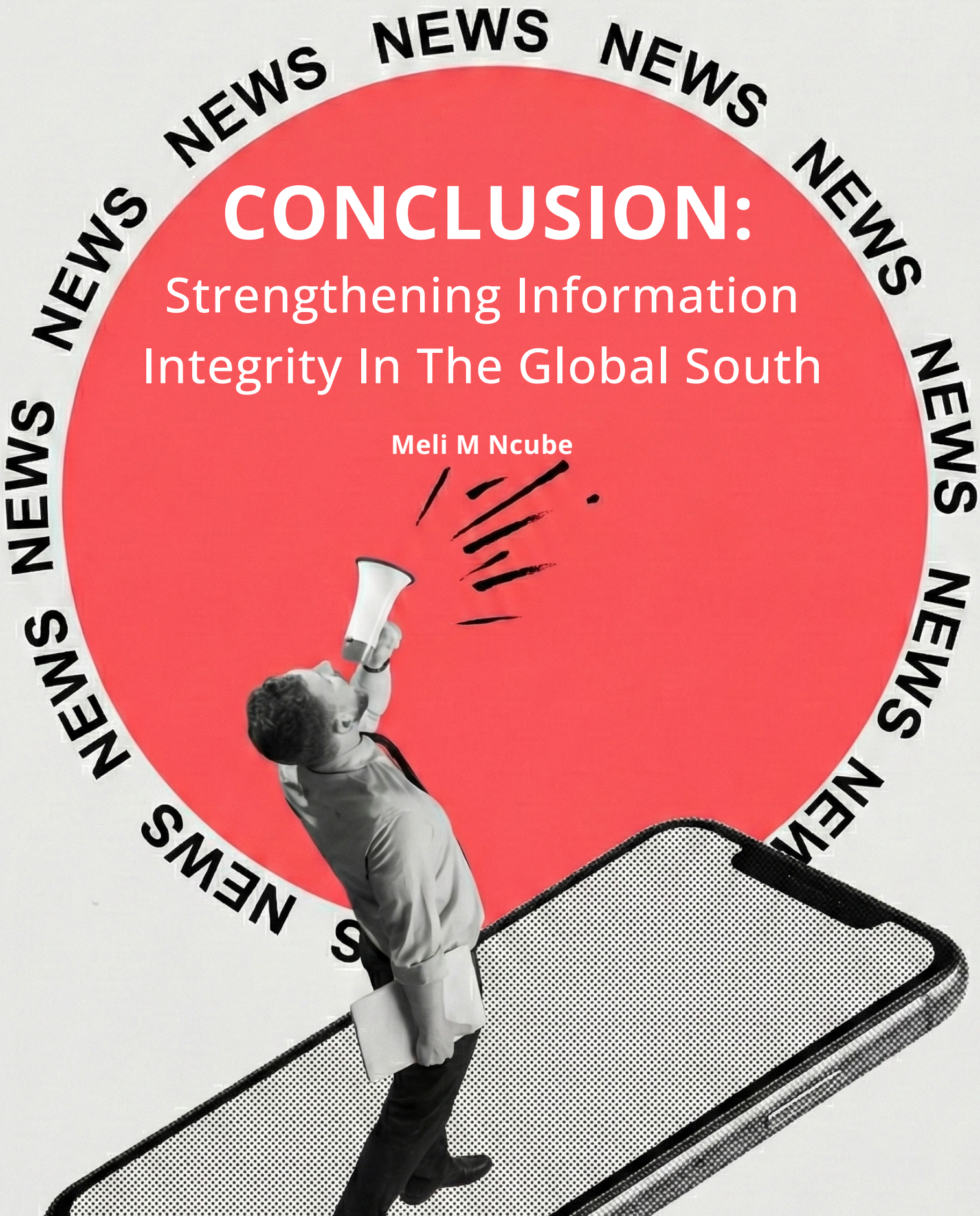
Wasserman, H. 2022. Introduction. In Wasserman, H. (ed.). *Meeting the challenges of information disorder in the Global South*, pp. 6–23. Cape Town: University of Cape Town. Available: <http://hdl.handle.net/10625/60954>

Wasserman, H. & Madrid-Morales, D. (eds.). 2022. *Disinformation in the Global South*. John Wiley & Sons. <https://doi.org/10.1002/9781119714491>

Westlund, O., Belair-Gagnon, V., Graves, L., Larsen, R. & Steensen, S. 2024. What is the problem with misinformation? Fact-checking as a sociotechnical and problem-solving practice. *Journalism Studies*, 25(8):898–918. <https://doi.org/10.1080/1461670X.2024.2357316>

Worldcrunch. 2019. Behind the crackdown on independent Egyptian news outlets. In *Worldcrunch*, 2 December. [Online]. Available: <https://worldcrunch.com/in-the-news/behind-the-crackdown-on-independent-egyptian-news-outlet/> [2025, July 18].

Zamit, F., Kooli, A. & Toumi, I. 2020. An examination of Tunisian fact-checking resources in the context of COVID-19. *JCOM*, 19(07): A04. <https://doi.org/10.22323/2.19070204>



CONCLUSION:

Strengthening Information
Integrity In The Global South

Meli M Ncube

CONCLUSION

This book illustrates the reality that information disorder is not solely a problem of untrue news or toxic speech; it is structural, entwined intrinsically with politics, economics, technology, and culture. The same pattern—including a weak media system with consolidated ownership, fragile regulation, and global platforms orientated toward engagement rather than truth—stands out from Africa through to Asia, Latin America, and the Middle East. These conditions create societies that are ripe for misinformation and disinformation. Solutions, therefore, must reach beyond technical fixes and get to the deeper structures that dictate how information flows.

WHY THE PROBLEM IS STRUCTURAL

The information disorder framework of Wardle and Derakhshan (2017) helps us to understand the various ways harmful content can manifest. However, intent is difficult to prove, and harm usually depends on context. In the Global South, information disorder proliferates in conditions where media operate on shoestring budgets, governments systematically limit access to information, and platforms capture advertising markets. These structural weaknesses set up an environment in which false content can spread easily and trust declines.

MEDIA CONCENTRATION AND PLATFORM DOMINANCE

Latin America provides an example of how the old patterns of media concentration are merged with new digital disruptions. The old oligopolies of traditional media still dominate, but now platforms like WhatsApp, YouTube, and Facebook oversee the traffic and revenue that flow to them. Algorithms reward sensational content, while programmatic advertising siphons money away from independent journalism. This is what is currently called ‘digital colonialism’ in Africa: foreign platforms set the rules while local languages and contexts are ignored. Infrastructure projects expand connectivity but deepen dependency when ownership remains offshore.

REGULATION: BALANCING INTEGRITY AND RIGHTS

Meanwhile, governments have introduced laws against so-called ‘fake news’, but many are based on vaguely worded provisions that have been used to silence critics. Inconsistency and bias towards Global North languages in platform self-regulation persist. Co-regulation, where responsibility is shared by states and companies under independent oversight, is preferable, but

requires robust institutions and civil society participation. Transparency of algorithms, data portability, and competition measures will be necessary to check the power of platforms.

JOURNALISM UNDER PRESSURE

The collapse of the old revenue models and the capture of advertising by platforms have done much to make journalism economically fragile. But this is more than a business problem; it is, importantly, a democratic one. When newsrooms are starved of resources, quality suffers and trust erodes. Hybrid funding models—subscriptions, donations, and philanthropy—help, but these have their limits. Journalism should be considered akin to public infrastructure and supported as such by fair taxation levied on digital platforms, alongside policies to ensure media diversity.

FACT CHECKING, MEDIA LITERACY

Fact-checking is indispensable, but it cannot be the sole panacea. It is fraught with political risks, underfunding problems, and maintaining pace with AI-generated output. Programs on media and information literacy show promise, particularly for children, but their effects quickly fade without reinforcement. The Sri Lanka case study further established that practical skills, such as keeping passwords safe, are retained, while habits for verification involving more complexity slowly deteriorate over time. Media and information literacy must be embedded within school curricula and championed through parenting; it must be placed in a local context.

SIX PRIORITIES FOR ACTION

To construct resilient information ecosystems, what is required is effort at six points:

1. Support independent journalism with sustainable funding and fair regulation.
2. Incentivize platforms, not just their content, to reduce the spread of harmful material.
3. Localize moderation and infrastructure to counter digital colonialism.
4. Integrate MIL into education and assess long-term impacts.
5. Strengthen networks of fact-checking and connect them with literacy programs.
6. Encourage epistemic diversity through promoting local languages and knowledge systems.

LOOKING AHEAD

Success does not mean the eradication of all lies. It means creating conditions under which toxic content cannot easily take hold and in which citizens have access to trustworthy information. To achieve that, it requires plural media markets, transparent platform governance, rights-based regulation, and robust literacy programs. Above all, it requires collaboration among governments, platforms, journalists, educators, and communities.

The Global South is not powerless. Regional initiatives such as the Praia Declaration, fact-checking networks, and community-based media and information literacy programmes demonstrate progress can be achieved. The next challenge is in scaling these programs and embedding them within systemic reforms. Information integrity is a technical problem alone: it is a political, economic, and cultural project. It will take time to construct, but it is imperative for democracy in the digital age.