



Received: 03-05-2026
Accepted: 13-06-2026

International Journal of Advanced Multidisciplinary Research and Studies

ISSN: 2583-049X

Regulating the Proliferation of False Information on X (Formerly Twitter) in Nigeria: A Reform-Based Analysis

Richard Moses Peace

Member of the Nigerian Bar Association/Barrister and Solicitor of the Supreme Court of Nigeria, Nigeria

DOI: <https://doi.org/10.62225/2583049X.2026.6.3.6500>

Corresponding Author: **Richard Moses Peace**

Abstract

This article examines the proliferation of false information on X and critically assesses the effectiveness of existing legal and regulatory mechanisms in Nigeria in mitigating its spread. Through a doctrinal and comparative analysis, the article reviews relevant provisions of the Cybercrimes (Prohibition, Prevention, etc.) Act 2015, the Nigerian Communications Act 2003, the National Information Technology Development Agency's Code of Practice for Interactive Computer Service Platforms/Internet Intermediaries 2022 (NITDA Code of Practice 2022), the Nigerian Communications Commission Internet Code of Practice 2026 and the Criminal Code Act 1916. It further analyses the effectiveness of X's self-regulatory

mechanisms, particularly the Community Notes programme, and highlights their limitations in curbing the rapid spread of false content. Drawing lessons from the regulatory approaches adopted in the European Union, the United Kingdom, and China, the article argues that Nigeria requires a balanced co-regulatory framework that promotes transparency, platform accountability, AI-assisted content verification, independent fact-checking, and media literacy while safeguarding constitutionally guaranteed freedom of expression. The article concludes by proposing reform-oriented recommendations aimed at strengthening digital information integrity and combating the harmful effects of false information on Nigeria's social media ecosystem.

Keywords: Social Media Regulation, False Information, Digital Governance, Freedom of Expression, Nigeria

Introduction

In recent years, content creation on social media platforms like X (formerly Twitter), TikTok, Facebook and others has gradually become a profitable venture.¹ The rapid advancement of artificial intelligence has transformed social media into an increasingly powerful medium for the dissemination of information; however, it has simultaneously accelerated the proliferation of misinformation, disinformation, and other forms of unverified content.² With the introduction of engagement-based monetisation on platforms such as X, the incentive to publish sensational or AI-generated content for financial gain has increased significantly.

The problem of social media monetisation on X began in 2023, when the executive chairman, Elon Musk, announced that content creators would receive a share of advertising revenue from ads appearing in replies to their posts.³ The first payouts, which commenced in July 2023, amounted to approximately \$5 million.⁴ The prospect of such significant financial rewards generated widespread interest among content creators, thereby stimulating a notable increase in platform engagement and content creation. By late 2023, the programme had been further expanded, allowing verified users to earn revenue directly from interactions and engagement generated by their posts.

The requirements for users to be eligible to earn on X are steep. For instance, according to X's Help Centre, for a user to be

¹ Sunny D. Nigerian creators are turning to rage-bait to cash in on x. TechPoint Africa, 2026 [Cited 2026 Mar. 1]. Retrieved from: <https://techpoint.africa/news/rage-bait-for-cash-on-x/>

² *ibid.*

³ Perez S. Twitter, now x, opens up its ad revenue sharing program with global creators. Tech Crunch, 2023 [Cited 2026 Mar. 1]. Retrieved from: <https://techcrunch.com/2023/07/28/twitter-now-x-opens-up-its-ad-revenue-sharing-program-with-global-creators/>

⁴ *ibid.*

eligible, the account must have an active Premium, Premium Business, or Premium Organisations subscription; have at least 5M organic impressions within the last 3 months; have at least 500 verified followers; be in a supported country; and be compliant with the X User Agreement.⁵ Payouts are processed every two weeks, with a minimum of \$30. This means that the higher the impression generated from a user's posts, the higher the earnings.

Undoubtedly, X has put in place some measures and policies to mitigate content that promotes violent and hateful speech, abuse, harassment and hateful conduct, as well as the spread of false information.⁶ Some of these measures include Community Notes, which allows specific users to propose and vet fact-checking notes for potentially misleading posts.⁷ Although studies have shown that, despite these policies, misinformation, especially AI-generated or unverified content, is prolific on the platform.⁸

Nigeria has, over the years, grappled with social media regulation, particularly given the absence of a clear legal and regulatory framework. Attempts to regulate social media in Nigeria began in 2019, when the Protection from Internet Falsehood and Manipulation Bill 2019 (also known as the Anti-Social Bill) was introduced in the National Assembly.⁹ This bill, which sought to criminalise the peddling of false or malicious information through social media, scaled second reading at the Nigerian Senate.¹⁰ Although it was vehemently criticised for hindering freedom of speech and freedom of the press.¹¹ Pushback and public sentiment discouraged the Federal Government from further sponsoring the Bill, and it was later withdrawn. Nevertheless, the lacuna in social media-specific laws in Nigeria has further intensified the peddling of false information on the X platform, leading some users to fabricate false and provocative content simply to garner impressions and engagement. Users who deliberately publish misleading, unverified content are rarely sanctioned on the X platform; there are no real legal penalties. Furthermore, even where accounts are suspended or disabled for violating the platform's policies, the affected users create new accounts and continue to 'farm' engagements unscrupulously.

Section 39(1) of the Constitution of the Federal Republic of Nigeria (As Amended) 1999 (CFRN 1999) guarantees freedom of expression, including the right to receive and share information, freedom of the press and media, and artistic and scientific expression. However, the right to

⁵ See X Help Centre. Creator revenue sharing. [Cited 2026 Mar. 1]. Retrieved from: <https://help.x.com/en/using-x/creator-revenue-sharing#:~:text=from%20the%20program>

⁶ Slaughter I, Peytavin A, Ugander J, Saveski M. Community notes reduce engagement with and diffusion of false information online. *PNAS*. 2025; 122(38):1-8.

⁷ *ibid*, 2.

⁸ Vallance C. Disinformation most active on x, formerly known as twitter, EU says, 2023 [Cited 2026 Mar. 3]. Retrieved from: <https://www.bbc.com/news/technology-66926080>

⁹ Coker O, Adeniran O. Nigeria and the struggle to regulate social media, 2022 [Cited 2026 Mar. 3]. Retrieved from: <https://www.dentonsacaslaw.com/en/insights/articles/2022/january/26/regulation-of-social-media-in-nigeria>

¹⁰ *ibid*.

¹¹ *ibid*.

freedom of expression is not absolute. It does not protect hate speech (speech that incites violence or discrimination), defamation (false statements that harm a person's reputation), or incitement of violence.¹² For instance, under section 39(3) and 45 of the of the CFRN 1999, the right to freedom of expression could be restricted by a law reasonably justifiable in a democratic society for the purpose of preventing the disclosure of information received in confidence, maintaining the authority and independence of the courts and in the interest of defence, public safety, public order¹³ and public morality, public health, for the protection of the rights and freedom for other persons.¹⁴ These constitutional limitations recognise that freedom of expression, although fundamental to democratic governance, must coexist with competing societal interests and the rights of other individuals. Accordingly, expression that threatens public order, undermines national security, infringe privacy rights, or cause demonstrable harm may legitimately attract regulatory intervention.

From the foregoing, this article critically assesses the extent to which Nigeria's existing legal and regulatory regime effectively mitigates the dissemination of misinformation and other forms of false information on social media platforms, particularly X. This article is structured into ten parts. Following this introduction, Part Two examines the concept of false information by distinguishing between misinformation, disinformation, and fake news, and analyses their implications for public trust and social stability. Part Three investigates the prevalence of false information on X, drawing on empirical studies and practical examples from Nigeria and other jurisdictions. Part Four evaluates the effectiveness of X's self-regulatory mechanisms, particularly the Community Notes programme, in addressing the dissemination of misleading content. Part Five examines the competing arguments for and against social media regulation, with particular emphasis on freedom of expression and human rights concerns. Part Six critically analyses Nigeria's existing legal and regulatory framework, including the Cybercrimes (Prohibition, Prevention, etc.) Act 2015 (CA 2015), the Nigerian Communications Act 2003(NC Act 2003), the NITDA Code of Practice 2022, the Criminal Code Act 1916 (CC Act 1916) and the Nigerian Communications Commission Internet Code of Practice 2026(NCC Code 2026). Part Seven undertakes a comparative examination of regulatory approaches adopted in the European Union, the United Kingdom, and China. Part Eight discusses the need for regulatory intervention in Nigeria and identifies existing regulatory gaps. Part Nine proposes reform-oriented recommendations to strengthen digital information integrity while safeguarding constitutional rights. Finally, Part Ten concludes the article by summarising the key findings and advocating a balanced co-regulatory framework for addressing the proliferation of false information on X in Nigeria.

Methodology

This article adopts a doctrinal legal research methodology complemented by a comparative analytical approach. The

¹² Udofa I. Right to freedom of expression and the law of defamation in Nigeria. *International Journal of Advanced Legal Studies and Governance*. 2011; 2(1):78-79.

¹³ CFRN 1999, s 45(1) (a).

¹⁴ CFRN 1999, s 45(1) (b).

doctrinal aspect involves a critical examination of relevant primary and secondary legal sources relating to the regulation of false information on social media platforms. Primary sources include the Constitution of the Federal Republic of Nigeria (as amended), 1999, the Cybercrimes (Prohibition, Prevention, etc.) Act 2015, the Nigerian Communications Act 2003, the Nigerian Communications Commission Code 2026, the NITDA Code of Practice 2022, and other relevant legislative and regulatory instruments. Secondary sources comprise journal articles, books, policy reports, official publications, and empirical studies on misinformation, disinformation, artificial intelligence, and social media governance.

The article further employs a comparative approach by examining regulatory frameworks adopted in the European Union, the United Kingdom, and China in combating the proliferation of false information on social media platforms. Through this comparative analysis, the study identifies best practices, regulatory gaps, and lessons that may be adapted to the Nigerian context. The research also draws on selected case studies and empirical evidence relating to the spread of false information on X (formerly Twitter) to evaluate the effectiveness of existing regulatory measures. The objective is to assess the adequacy of Nigeria's current legal framework and propose reform-oriented recommendations that balance effective regulation with the protection of constitutional rights, particularly freedom of expression.

Result and Discussion

1. Concept of False Information and Its Consequences

False information on social media platforms broadly includes misinformation, disinformation, and fake news, which are essential concepts that require dissecting. According to Ireton and Posetti, misinformation is misleading information, typically disseminated without malicious intent.¹⁵ On the other hand, disinformation is the malicious and intentional spread of misleading or inaccurate information with the intent to deceive or manipulate and influence public opinion or conceal the truth.¹⁶ While both concepts differ in form and intent, they have a similar tendency to erode public trust and mislead the general public. Both are harmful to society. The European Commission described disinformation as intentional misinformation that poses a serious threat to democratic processes and various industries.¹⁷ Fake news, in contrast, is news articles that are purposely and verifiably false, with the intention to mislead readers.¹⁸ In essence, fake news is deceitful stories disguised as news to influence public opinion. While the intent behind misinformation, disinformation and fake news differ, the commonalities of all three concepts are predicated on the content: inauthenticity and deception. All three concepts deceive and

mislead the general public with the intention to distort and manipulate the truth. The dangers of false information are that since the recipients cannot identify it as false, their opinion and actions are wrongly influenced from the onset, leading to uninformed decisions, which could result in negative consequences. It has been argued that false information on social media can erode the unity and peace of a polity, thereby creating distrust, violence and discord.¹⁹ Furthermore, false information also makes the citizens of a nation ill-informed and weakens the legitimacy of journalism, while facilitating anti-democratic tendencies that can sabotage political processes. This view is in line with the position posited by Berghel, who asserted that fake news can be modified and utilised as a political weapon of choice by politicians.²⁰ In Nigeria, Apuke and Omar observed that false information, particularly fake news, has for many decades escalated conflicts, political hostility and social panic and in some cases occasioned the death of several citizens.²¹

In Nigeria, false information and fake news on X is rife. For instance, in 2017, false content was circulated on social media alleging that five students of the College of Education Gidan Waya were ambushed and killed by Fulani herdsmen in southern Kaduna.²² This news stirred up tensions and conflict between Christians and Fulani Herdsmen.²³ The report was later found to be fake as no student was killed.²⁴ In February 2026, a TikTok user known as Abigail Nsuka, aka Mirabel (with username @mirab351), published and disseminated a video in which she alleged that she had been burgled and sexually assaulted by an unknown individual within her apartment.²⁵ The video rapidly gained traction on various social media platforms, generating widespread public reaction and provoking significant controversy and conflict among social media users and influencers.²⁶ Subsequent investigations reportedly revealed that the allegations were fabricated and that the video had been deliberately created and circulated to attract public sympathy, increase online engagement, and solicit financial contributions from members of the public through crowdfunding initiatives.²⁷ The incident illustrates the capacity of social media platforms to facilitate the rapid dissemination of false information capable of misleading the public and distorting public discourse. Notwithstanding the

¹⁹ *ibid.*

²⁰ Berghel H. Lies, damn lies and fake news. *Computer*. 2017; 50(1):80-85.

²¹ Apuke O, Omar B. Fake news proliferation in Nigeria: consequences, motivations and prevention through awareness strategies. *Humanities and Social Sciences Reviews*. 2020; 8(2):318-327.

²² Ogbette A, Idam M, Kareem A, Ogbette D. Fake news in Nigeria: causes, effect and management. *Information and Knowledge Management*. 2019; 9:96-99.

²³ *ibid.*, 98.

²⁴ *ibid.*, 96-99.

²⁵ Akasike C, Uthman I, Busari B, Adetunji T, Olasupo A, Taiwo B. Mirabel: How false rape alarm turned into money-making venture. *Punch News Paper*, 2026 Feb. 21 [Cited 2026 Mar. 25]. Retrieved from: <https://punchng.com/mirabel-how-false-rape-alarm-turned-into-money-making-venture/>

²⁶ *ibid.*

²⁷ *ibid.*

¹⁵ Ireton C, Posetti J. Journalism, fake news and disinformation: Handbook for journalism education and training. United Nations Digital Library, 2018, 7.

¹⁶ *ibid.*, 7.

¹⁷ European Commission. Disinformation: A threat to democracy – brochure, 2021 [Cited 2026 Mar. 18]. Retrieved from: <https://digital-strategy.ec.europa.eu/en/library/disinformation-threat-democracy-brochure>

¹⁸ Allcott H, Gentzkow M. Social media and fake news in the 2016 elections. *J Econ Perspect*. 2017; 31:211-236.

apparent falsity of the allegations and the significant public attention generated by the incident, there is, as of the date of this study, no publicly available evidence indicating that the user was criminally charged or prosecuted for the dissemination of false information or for allegedly inducing members of the public to part with money under false pretences. This raises important questions regarding the adequacy of existing legal and platform-based mechanisms for addressing the deliberate spread of false information online, particularly where such conduct is intended to deceive the public and obtain financial gain.

2. Prevalence of False Information on X Platform

Fake news and false content are considered highly prevalent on X, with multiple studies suggesting that the platform has a higher proportion of such content compared to other major social media platforms. For instance, in 2023, the European Commission study, which analysed over 6,000 social media posts, discovered X as having the highest 'Ratio of discoverability' of disinformation in comparison to the other 5 major platforms, Facebook, Instagram, LinkedIn, TikTok, and YouTube.²⁸ The study covered Spain, Poland and Slovakia, which are countries deemed particularly at risk of disinformation due to elections and proximity to the war in Ukraine.²⁹ Furthermore, another 2026 study conducted by the London School of Economics showcased how generative AI misinformation played a major part in fueling the UK civil unrest between July and August 2024, following the stabbing attack and horrifying killing of three girls in Southport, England.³⁰ The study comprised an analysis of several comments, views, shares and likes from two prominent accounts, one of which was managed by a UK-based far-right political party that supports white nationalism and the other account, categorised as Media and News, was used to share false information about the perpetrator of the attack.³¹ Overall, the study showed that X platform was generally used to amplify racist conspiracy theories, reproduce racist generative AI content with high viral potential, spread fake news and legitimise racist conspiracy theories through verified accounts.³² Another study carried out by Corsi, Marino and Wong from Harvard Kennedy School in June 2024 also revealed AI generated synthetic media (deepfakes, fake images) increased on X between 2022 and 2023 after the launch of Midjourney V5, a text-to-image AI tool.³³ A recent study published by Pew Research also revealed that X has the highest rate of

information when compared to other mainstream social media platforms.³⁴ According to Pew Research Centre, despite X being more of a news outlet than other platforms, about 86% of users in the study reported seeing inaccurate news; about 37% reported seeing it often.³⁵

In 2020, Nigerians witnessed how the #EndSARS and #EndBagGovernanceinNigeria movement, which started as a peaceful online protest against police brutality, was unfortunately hijacked by unscrupulous individuals who exploited social media to distribute false, harmful, and incendiary content.³⁶ This incident led to the loss of properties and lives, breakdown of law and order, while underlining the dangers of unregulated social media content.³⁷ Fake news and misinformation usually create confusion, tension, and, in certain cases, promote suicide, depending on the recipient; it undermines serious media coverage and renders legitimate journalism difficult.

3. How Effective are X Policies Against the Spread of False Information?

On the 6th of October 2022, X introduced the community Notes Programme, which enables enrolled users to flag posts which they believe are misleading and provide textual notes about the source of the post. Following the submission of the community notes, the fact-check becomes available for other enrolled users to rate while remaining hidden from the public users.³⁸ The programme, which utilises a bridging-based rating system, calculates the usefulness results for every community note based on the rating provided by the contributors.³⁹ Only the notes that are rated useful by various contributors with heterogeneous rating histories are displayed to the public users on the platform.⁴⁰ This method remains the platform's primary method for

²⁸ Vallance C. Disinformation most active on x, formerly known as twitter, EU says, 2023 [Cited 2026 Mar. 3]. Retrieved from: <https://www.bbc.com/news/technology-66926080>

²⁹ Ibid.

³⁰ Buarque BL. The x factor: Four ways one social media platform undermined democracy during the 2024 summer riots, 2026 [Cited 2026 May 8]. Retrieved from: <https://www.lse.ac.uk/research/research-for-the-world/society/x-undermined-democracy-uk-riots>

³¹ Ibid.

³² Ibid.

³³ Corsi G, Marino B, Wong W. The spread of synthetic media on x. Harvard Kennedy School Misinformation Review. 2024; 5(30):1. [Cited 2026 May 8]. Retrieved from: <https://misinforeview.hks.harvard.edu/article/the-spread-of-synthetic-media-on-x/>

³⁴ Kirkland C. X has highest rate of misinformation as a news source, study finds. Media Daily News, 2024 [Cited 2026 May 8]. Retrieved from: <https://www.mediapost.com/publications/article/396846/x-has-highest-rate-of-misinformation-as-a-news-sou.html>

³⁵ Shearer E, Naseer S, Liedke J, Matsa K, How Americans get news on tiktok, x, facebook and Instagram. Pew Research Centre, 2024 [Cited 2026 May 10]. Retrieved from:

<https://www.pewresearch.org/journalism/2024/06/12/how-americans-get-news-on-tiktok-x-facebook-and-instagram/> Accessed 10 May 2026

³⁶ National Information Technology Development Agency. An analysis of compliance with laws and misinformation management by social media platforms in Nigeria, 2024 [Cited 2026 May 10]. Retrieved from: <https://nitda.gov.ng/wp-content/uploads/2024/12/NITDA24-COP-REPORT-JNT-compressed.pdf>

³⁷ Ibid.

³⁸ X Community Notes. Signing up. [Cited 2026 May 8]. Retrieved from: <https://communitynotes.x.com/guide/en/contributing/signing-up>

³⁹ Wojcik S, Hilgard S, Judd N, Mocanu D, Regain S, Hunzaker M, Coleman K, Baxter J. Birdwatch: Crowd wisdom and bridging algorithms can inform understanding and reduce the spread of misinformation. 2022 [Cited 2026 May 10]. Retrieved from: <https://arxiv.org/abs/2210.15723>.

⁴⁰ X Note ranking algorithm. [Cited 2026 May 12]. Retrieved from:

crowdsourcing content moderation and providing context to potentially misleading or false posts.⁴¹

In 2024, Chuai *et al.* analysed the effectiveness of the Community Notes programme in mitigating the spread of misleading posts on X.⁴² The study disclosed that the programme, on average, had a 62.0% chance of reducing the spread of misleading posts. Nevertheless, it was also discovered that the community notes might be too slow in intercepting the early (and most viral) stage of the dissemination.⁴³ In this regard, it was observed that while 75.9% of all notes considered useful were displayed to the general public within 36 hours of creation, there was a considerable time lag (averaging 61.4 hours) between post-creation and note display. In comparison, posts on X spread considerably faster, with the half-life of reposts over 36 hours being 5.75 hours, with only 13.5% of all helpful notes displayed before this time point.⁴⁴ This further suggests that community notes might be inefficient in mitigating the spread of misleading posts at the viral stage of dissemination.⁴⁵ From a regulatory perspective, this further suggests the need for a co-regulatory framework in which social media platforms collaborate with external regulatory authorities to combat the proliferation of false information while preserving fundamental rights. Under such an approach, platforms would retain primary responsibility for day-to-day content moderation, enforcement of community standards, and the deployment of technological measures for detecting misleading or harmful content. However, these internal mechanisms would operate within a broader framework of statutory obligations, independent oversight, and transparency requirements established by law.

4. Debates for and Against Social Media Regulation

Irrespective of the merits of social media platforms in facilitating the effortless dissemination of news and useful information to a wider audience speedily, there have been agitations for their regulation by individuals and governments around the world. As earlier demonstrated, the downside of social media includes cyberbullying, manipulation of content, and dissemination of false or fake news with the aim of influencing public opinions. According to Chin, the UK's Defamation Act 2013 (DA 2013) was prompted by defamation activities by individuals on social media.⁴⁶ Issues about social media, such as data protection

<https://communitynotes.twitter.com/guide/en/under-the-hood/ranking-notes> (2024)

⁴¹ X Help Center. About x community notes on x. [Cited 2026 May 12]. Retrieved from: [https://help.x.com/en/using-x/community-](https://help.x.com/en/using-x/community-notes#:~:text=Community%20Notes%20aim%20to%20create,%2C%20and/or%20other%20remediations.)

[notes#:~:text=Community%20Notes%20aim%20to%20create,%2C%20and/or%20other%20remediations.](https://help.x.com/en/using-x/community-notes#:~:text=Community%20Notes%20aim%20to%20create,%2C%20and/or%20other%20remediations.)

⁴² Chuai Y, Pilarski M, Renault T, Restrepo-Amariles D, Troussel-Clement A, Lenzini G, Prollochs N. Community-based fact-checking reduced the spread of misleading posts on social media. arXiv, 2024 [Cited 2026 May 13] Retrieved from: <https://arxiv.org/pdf/2409.08781>

⁴³ *ibid*, 7

⁴⁴ *ibid*, 7

⁴⁵ *ibid*, 16.

⁴⁶ Chin Y. Regulating social media, regulating life (and lives): A report on the workshop 'social media, regulation and freedom of expression' in May at Hong Kong Baptist University. Columbia Journalism Review, 2013, 33-37.

and privacy, hate speech, cyber-bullying and copyright are driving regulation. In this vein, some researchers have argued that self-regulation is not adequate and further regulation is necessary to curb the proliferation of false content on social media.⁴⁷ In view of this, Baccarella *et al.*, argued that an independent, reliable and universal regulatory framework that is applicable to every social media company and platform is necessary.⁴⁸

On the contrary, the regulation of social media is usually discouraged on the premise that it will undermine freedom of speech, which is a fundamental constitutional right in many jurisdictions, including Nigeria. In this regard, Egbunike contends that regulation is abhorrent because it puts the right to free speech in jeopardy and could prevent people from speaking out.⁴⁹ Egbunike further argued that while regulations aim to mitigate the spread of false content on social media platforms, including fake news and outright misinformation, their real aim is to outlaw freedom of expression and criticism.⁵⁰ In line with this, Mhaka also opined that it is becoming apparent that the regulations of social media in Nigeria are aimed at silencing dissenting voices and keeping them permanently incarcerated; unlike the military regime, the law is being utilised to achieve this purpose.⁵¹ There is a general concern that while regulation can help mitigate illegal or harmful activities and the spread of misinformation on social media platforms, they can also be misused by authorities to constrain social media's influence on citizens or, in extreme cases, silence online criticism of the government.⁵² For instance, in May 2024, Daniel Ojukwu, an investigative journalist in Nigeria, was arrested and detained for a week without a formal charge over an investigative report revealing corruption implicating senior Nigerian Officials.⁵³ His arrest was carried out pursuant to the provisions of the Cybercrimes (Prohibition,

[Cited 2026 May 15]. Retrieved from: https://www.academia.edu/25901348/Regulating_social_media_regulating_life_and_lives_A_report_on_the_workshop_Social_Media_Regulation_and_Freedom_of_Expression_in_May_at_Hong_Kong_Baptist_University

⁴⁷ Baccarella C, Wagner T, Kietzmann J, McCarthy I. Averting the rise of the dark side of social media: The role of sensitisation and regulation. European Management Journal. 2020; 38(1):3-6.

⁴⁸ *ibid*.

⁴⁹ Egbunike N. Nigeria's social media bill will obliterate online freedom of expression. Global Voices, 2020 [Cited 2026 May 21]. Retrieved from: <https://advoc.globalvoices.org/2019/11/29/nigerias-social-media-bill-will-obliterate-online-freedom-of-expression/>

⁵⁰ *ibid*.

⁵¹ Mhaka T. How social media regulations are silencing dissent in Africa. Aljazeera, 2020 [Cited 2026 May 21]. Retrieved from: <https://www.aljazeera.com/opinions/2020/11/12/how-social-media-regulations-are-silencing-dissent-in-africa>

⁵² Muruaru D, Kombo S. Thumbs up or down: Regulating social media is a double-edged sword. Busara, 2026 [Cited 2026 May 24]. Retrieved from: available at <https://busara.global/blog/thumbs-up-or-down-regulating-social-media-is-a-double-edged-sword/>

⁵³ Akinwotu E. Nigeria has detained a journalist who reported on corruption in a widening crackdown. NPR, 2024 [Cited 2026 May 24]. Retrieved from:

Prevention, etc.) Act 2015, a law that grants the Nigerian government broad powers to regulate perceived online offences.

In view of the concerns for human rights violations of freedom of speech, international legal experts believe that international human rights law can serve as a standard in the legal regulation of social media.⁵⁴ For instance, Professor McGoldrick from the University of Nottingham proposes that relevant domestic statutory provisions should be adequately modernised to ensure their compatibility with international human rights standards in view of the increasing growth of social media.⁵⁵ In support of this view, the author acknowledges the significance of human rights law jurisprudence in the domestic regulation of social media. Therefore, it is argued here that social media regulations and laws need to be carefully drafted in a way that aligns with human rights principles.

5. Social Media Regulatory Landscape in Nigeria: How Effective is it in Curbing False Content?

Nigeria's social media regulatory landscape is evolving and currently in its embryonic stages. There is currently no dedicated social media legislation to curtail the spread of false content or information in Nigeria. However, relevant provisions are scattered across various statutes and subsidiary regulations. Provisions that indirectly proscribe false content on social media in Nigeria can be found in the Cybercrimes (Prohibition, Prevention) Act 2015, NITDA Code of Practice 2022, the Nigerian Criminal Code Act, and, more recently, the Communication Commission Code 2026. This section aims to critically examine the effectiveness of the provisions of these various laws in terms of curtailing the spread of false content on social media platforms, particularly on X.

Cybercrimes (Prohibition, Prevention) Act 2015(CA 2015)

In Nigeria, the regulation of false content on social media platforms such as X is implicitly governed by the CA 2015, which makes it an offence to knowingly or intentionally send a message via a computer system or network that is known to be false.⁵⁶ In 2024 the Act underwent a crucial amendment, further redefining the scope of the offence in section 5 of the Cybercrimes (Prohibition, Prevention, ETC)(Amendment) Act 2024 as follows: 'Any person who knowingly or intentionally sends a message or other matter by means of computer systems or network that - (b) he knows to be false, for the purpose of causing a breakdown of law and order, posing a threat to life, or causes such a message to be sent: commits an offence under this Act and shall be liable on conviction to a fine of not more than N7,000,000.00 or imprisonment for a term of not more than 3 years or to both such fine and imprisonment. This provision appears to criminalise the intentional dissemination of false information through computer systems. However, its scope is limited to situations where such false information is deliberately and maliciously circulated with the intention of causing a breakdown of law

and order or causing a threat to life. Consequently, the provision primarily addresses disinformation, which involves the intentional spread of falsehoods for harmful purposes. It does not, however, extend to misinformation, where false information is shared without malicious intent or knowledge of its falsity. The issue with this is that when false information is posted or shared recklessly on social media without verification, and such information happens to prejudice or defame an individual, it would not be considered an offence if the initial intention was not to cause a breakdown of law and order or cause a threat to life.

In 2024, Olumuyiwa Adejobi, the Nigerian Police Force Public Relations Officer, summarised the scope of section 24 of the CA 2024 in relation to false content as follows: "Whether you are a content creator, blogger or journalist, you must make sure you verify your stories and possibly, balance your stories before you come online to peddle falsehood. We arrested and prosecuted those who were out there peddling falsehood, fake news, disinformation and misinformation to cause problems and discredit individuals and institutions, to malign people and to kill people's characters".⁵⁷ Despite the prohibition of false content under the provisions of the CA 2024, the dissemination of false information has continued to thrive predominantly through social media networks and platforms. This raises important questions regarding the effectiveness of existing legal frameworks and the self-regulatory mechanisms adopted by social media companies in combating the spread of misinformation and disinformation.

Nigerian Communications Act 2003

According to section 146(1), of the NC Act 2003, a licensee shall use his best endeavour to prevent the network facilities that he owns or provides or the network service, application service or content application service that he provides, from being used in, or in relation to, the commission of any offence under any law in operation in Nigeria. This provision does not expressly regulate social media activities nor criminalise the spread of false information on social media per se. It generally prevents licensees from permitting their computer networks or services to be used as a platform for committing an offence prescribed by Nigerian legislation. Therefore, since the offence of false information on computer networks is already prescribed by the CA 2015, it can be inferred that by virtue of section 146(1) of the NC Act, 2003, social media platforms must put in place measures to mitigate or prevent the spread of false information. However, as highlighted above, the Act does not expressly regulate misinformation, disinformation, or fake news on social media platforms. Rather, its provisions are framed broadly to impose general obligations on licensees to prevent their networks and services from being utilised for unlawful activities already recognised under Nigerian law. Consequently, the NC Act 2003, which largely operates by reinforcing and complementing existing statutory offences such as those contained in the CA 2015, suffers from similar regulatory limitations.

<https://www.npr.org/2024/05/09/1250008596/nigeria-detained-daniel-ujukwu-investigative-journalist-corruption>

⁵⁴ Chin Y. Regulating social media, regulating life (and lives) Columbia Journalism Review, 2013, 33-37.

⁵⁵ *ibid.*

⁵⁶ CA 2015, s 24(b).

⁵⁷ Ayeni V. How nigerian authorities use cybercrimes act to harass, detain journalists, activists. Punch Nigeria Newspaper, 2025 [Cited 2026 May 20]. Retrieved from: <https://punchng.com/how-nigerian-authorities-use-cybercrime-act-to-harass-detain-journalists-activists/>

The Nigerian Communication Commission Internet Code, 2026(NCC Code 2026)

This NCC Code 2026 is issued as guidelines in line with the provisions of Section 70 of the NC Act 2003, which empowers the Commission to issue subsidiary legislation on all matters pertaining to the provisions of the Act and its effective implementation. In relation to social media regulation, NCC Code 2026 tackles matters such as the spread of harmful content, including hate speech and disinformation. In this regard, online service platforms are required to establish robust community rules and guidelines that are in congruence with the provisions in section 146(1) of the NC Act 2003 in terms of protection of the national interest.⁵⁸ Particularly, online and digital platforms are also mandated to establish a channel of engagement with the Designated Online Governance Officer (DOGO), who is appointed by the Commission, in which the channel shall be the main engagement point for managing harmful content, disinformation, fraudulent activities and unlawful content.⁵⁹ This provision promotes self-regulation and ensures that social media platforms include in their policies mechanisms and measures to mitigate the spread of disinformation. The implication of this is that it gives social media platforms the sole discretion to decide the nature and type of mechanisms needed to curtail harmful and false content, leaving room for little to no regulatory oversight by the Nigerian government. The issue with this, as can be gleaned from the X's Community Notes as analysed above, is that the various content moderation mechanisms introduced on social media platforms may be ineffective in swiftly tackling the proliferation of false content. Another issue with this provision is that, despite providing a clear definition of the term 'misinformation'⁶⁰ in Regulation 8.1., the scope of the code only applies to disinformation, which, as examined above, is a lacuna that allows reckless and harmful false information to be posted on social media platforms without consequences, provided there is no intention of causing harm or threat to life.

NITDA Code of Practice for Interactive Computer Service and Internet Intermediaries 2022

In terms of prohibiting misinformation and disinformation, the NITDA has introduced the NITDA Code of Practice 2022, aimed at defining guidelines for interactions in the digital space, and further protecting the digital rights of all citizens in Nigeria, and foreigners that are domiciled in Nigeria. The Code requires online platforms to acquaint themselves with indigenous and contextual manifestations causing disinformation and/or misinformation, and factors motivating its spread in Nigeria. Particularly, social media platforms are to take steps to prevent, curtail and take down content capable of misinformation and disinformation.⁶¹ Platforms are to trace, expose, penalise and close accounts/sources that amplify disinformation and misinformation.⁶² Platforms are also mandated to inform

users of their services through terms of service, not to create, publish, promote, modify, transmit or share content or information that is false or misleading.⁶³ Undoubtedly, the principles of the NITDA Code of Practice 2022 are commendable, as its guidelines would ensure that social media platforms adopt best practices to prevent cyberbullying, harmful content and spread of false information. However, some of the provisions of the NITDA Code 2022 appear to be extreme and may likely infringe on the right to freedom of speech and expression⁶⁴ and privacy⁶⁵ as encapsulated in the CFRN 1999. Platforms are granted wide powers to set rules for access or usage of their platforms and to terminate the users' access where such rules are not complied with.⁶⁶ There is the concern that such a provision can be utilised by social media platforms to restrict users' access arbitrarily, thereby infringing on their right to freedom of expression.⁶⁷ Furthermore, part 2 of the NITDA Code 2022, which allows platforms to preserve disabled or removed content and any information concerning a person who is no longer a user of a platform, is also considered to infringe on an individual's right to privacy.⁶⁸

Criminal Code Act 1916

The Criminal Code Act 1916 (CCA 1916), while outdated, can be considered instrumental in the regulation of false information on social media in Nigeria, as it criminalises the publication of false news with the intent to cause fear and alarm to the public. According to Section 59 (1) of the CCA 1916, any person who publishes or reproduces any statement, rumour or report which is likely to cause fear and alarm to the public or to disturb the public peace, knowing or having reason to believe that such statement, rumour or report is false, is guilty of a misdemeanour and liable on conviction to imprisonment for three years.⁶⁹ However, unlike the Cybercrimes Act, the CCA 1916 does not contain provisions specifically targeted at misinformation, disinformation, or fake news on social media. Its relevance arises indirectly through offences relating to false publications, defamatory matter, public mischief, and acts capable of disturbing public peace. Furthermore, the CCA suffers certain shortcomings in addressing contemporary challenges presented by social media misinformation. First, the Act was enacted long before the emergence of social media platforms and therefore contains no provisions

⁶³ Part 2(2)(f) NITDA Code of Practice, 2022.

⁶⁴ CFRN 1999, s 39.

⁶⁵ CFRN 1999, s 37.

⁶⁶ NITDA Code 2022, Part 2(8).

⁶⁷ Ajibola M, Social media regulation: An overview of the national information technology development agency's (NITDA) code of practice for interactive computer service platforms/internet intermediaries. Lexworth Legal, 2024 [Cited 2026 May 28]. Retrieved from: <https://www.lexworthlegal.com/social-media-regulation-an-overview-of-the-national-information-technology-development-agencys-nitda-code-of-practice-for-interactive-computer-service-platforms-internet-intermediaries/>

⁶⁸ *ibid.*

⁶⁹ Similar provisions can be found in section 418 of the Penal Code Act 1960, which applies to the Northern part of Nigeria. The Criminal Code Act applies to the southern part of Nigeria.

⁵⁸ NCC Code 2026, Regulation 6.1.1.

⁵⁹ NCC Code 2026, Regulation 6.1.3.

⁶⁰ NCC Code 2026 defines misinformation as false, inaccurate, or misleading information that is shared without the intent to deceive the public. See Regulation 8.1. NC Code 2026.

⁶¹ NITDA Code of Practice 2022, Part 5.

⁶² Part 5(10), NITDA Code of Practice, 2022.

specifically dealing with digital communications, algorithmic amplification, anonymous accounts, deepfakes, synthetic media, or AI-generated content. Secondly, most of its provisions are reactive in nature and require proof of specific harm, public disturbance, or reputational damage before liability can arise. Thirdly, the Act does not impose any obligations on social media platforms to monitor, remove, label, or mitigate false content circulating on their services. Consequently, while the CCA 1916 may provide limited remedies against certain forms of harmful false information, it is largely inadequate as a comprehensive mechanism for regulating misinformation and disinformation on modern social media platforms such as X. This deficiency further underscores the need for a dedicated and technologically responsive legal framework capable of addressing the unique challenges posed by digital communication and AI-driven misinformation in Nigeria.

6. Global Trends on Social Media Regulations

With the advent of artificial intelligence and the evolution of technology as a means of communication and content modification, social media has become a suitable platform for creating, disseminating, and popularising content at incredible speed. The conventional statutory regulations, which were once deprecated on the premise of freedom of expression, are now being re-evaluated to protect public and national interests. Lately, the idea of self-regulation, which was once the golden standard for social media oversight, is being modified in various jurisdictions to favour a system of co-regulation as the future generation's panacea for affording desired freedom of expression and transparency simultaneously. This section looks at efforts made by the EU, UK and other prominent countries in combating false content on social media.

Europe

Recently, Europe has intensified legislative efforts to combat the spread of false news and disinformation on social media, with the enactment of the EU Digital Services Act 2022(DSA 2022) and, more recently, the Artificial Intelligence Act 2024(EU AI Act 2024), which both govern content moderation practices of social media platforms to address illegal content, misinformation and disinformation. The AI Act specifically tackles fake news, AI-generated false content, including audio or video, deep fakes, by mandating creators to clearly label all content that is generated using AI.⁷⁰ In so doing, deployers who use AI systems to generate or manipulate image, audio, or video content are required to state that such content is artificially created while clearly indicating the AI output, and its artificial origin. Although, to safeguard against the infringement of freedom of expression and right of freedom of art, the EU AI Act went further to state that transparency obligation, as far as fictional, satirical, and creative work are concerned, is only limited to disclosing the existence of such generated or manipulated content in a manner that does not impede its enjoyment, display, including its exploitation, use and quality.⁷¹ This provision is significant because it ensures that the freedom of expression of citizens is preserved, and there is a clear balance between safeguarding false content and protecting the rights of citizens to freely

express themselves.

Additionally, despite no express references to disinformation and misinformation, the DSA 2022 specifically emphasises a transparency obligation on digital platforms to tackle and minimise the spread of illegal content or information on their systems. The notion of illegal content, which is the main focus of the DSA 2022, is framed in a very broad term by Article 3(h), which includes any information that in itself or in relation to an activity is not in compliance with Union Law or the Law of any Member State. By interpretation, it means that any information, including disinformation, which is proscribed by EU laws, including member states, will be considered illegal content which must be curtailed. By so doing, online platforms must ensure a safe, precise, and trusted online environment, including tackling societal risk or harm that the dissemination of disinformation or other false content may occasion.⁷² According to Article 35, Very Large Online Platforms (VLOPs) must then implement effective measures to mitigate systemic risks, which includes ensuring that an item of information, whether it constitutes a generated or manipulated image, audio or video that resembles an existing person, object, or place and falsely appears to a person to be authentic or truthful is distinguishable through prominent marking when published on their platforms. While online platforms are not expressly obligated to actively seek out illegal activity, Article 35 underlines their responsibilities concerning, *inter alia*, the spread of disinformation.

United Kingdom

In the United Kingdom, the Online Safety Act 2023 (OSA 2023) was introduced to combat false communication online. In this regard, section 179 provides that a person commits an offence if— (a)the person sends a message (b)the message conveys information that the person knows to be false, (c)at the time of sending it, the person intended the message, or the information in it, to cause non-trivial psychological or physical harm to a likely audience, and (d)the person has no reasonable excuse for sending the message. The *mens rea* of the offence of false communication provisions is in two parts. Firstly, the message sent must be knowingly false.⁷³ Secondly, it must be sent with the intent to cause non-trivial psychological or physical harm.⁷⁴ A cursory look at the provisions shows that the offence, which requires intention to disseminate false content and cause harm with it, is limited only to disinformation and not misinformation. Misinformation, as examined above, refers to information that is fully or partially false but is disseminated without intending to cause harm to the recipient.

The exclusion of misinformation from the offence of false communication under the OSA 2023 is considered a key

⁷² Coli M. The role of the European Union in protecting democracy through legislation: the case of disinformation. *European Papers - A Journal on Law and Integration*. 2026; 11(1):81-111.

⁷³ Higson-Bliss L. Legislative comment: The online safety act 2023 and the sending of threatening or false communications. *State Law Review*. 2025; 46(2). [Cited 2026 May 28]. Retrieved from: <https://academic.oup.com/slr/article/46/2/hmaf021/8203212>.

⁷⁴ *ibid*.

⁷⁰ EU AI Act 2024, Reg. 134.

⁷¹ EU AI Act 2024, Reg. 134.

safeguard against the criminalisation of users who share content, lacking knowledge of its falsehood.⁷⁵ However, this also poses a major problem in that users who share false content recklessly would not have committed an offence even if such content misleads the public or causes harm to the general public. The intent of the provisions in section 179 of OSA 2023 also fails to take into account the fact that several false communications on social media platforms are not shared to harm the audience but rather posted to incite others to cause harm. In view of this, it is argued that knowledge of falsity itself is difficult to ascertain as some content may be true, partially true or false.⁷⁶ This problem is further exacerbated by the second aspect of the offence, which requires ascertaining the defendant's intention to cause non-trivial psychological harm. The OSA lacks a clear definition of the term non-trivial psychological harm, which means one cannot really decipher the kind of harm that a false post must cause to fall under the scope of the offence. Coe argued that, since it is difficult to define, it may lead to uncertainty and lack of clarity in the law.⁷⁷ For instance, what one officer or prosecutor may consider non-trivial, another may consider trivial. As Coe rightly observed, since the offence relates to the intention to cause harm rather than to actual harm, there is a concern that it will be applied and interpreted too broadly, thereby leaving room for arbitrary application and abuse.⁷⁸

In view of this lacuna, the Report from the Science, Innovation and Technology Committee (SITC) warns that the OSA 2023 is outdated and does not tackle misinformation in light of the 2024 summer riots, which were driven in part by misinformation and hateful content that was amplified on social media platforms by algorithmic recommenders.⁷⁹ The report further recommends the imposition of a set of duties on the companies to deprioritise content found to be misleading by fact-checkers.⁸⁰

China

In China, measures to curb fake news and the spread of misinformation on social media platforms are drastic and rigorous, owing to the radical steps taken by the Cyberspace Administration of China Regulation (CAC), which acts as the primary agency responsible for enforcing the country's comprehensive legal framework, regulating social media platforms and digital content. One of the prominent laws regulating misinformation on social media in China is the

Cybersecurity Law of the People's Republic of China 2017 (CL 2017), which imposes severe legal penalties, platform liability, and qualification mandates for users of social media platforms. For instance, CL 2017 introduced a Real-Name Registration for social media users. The law requires users to verify their accounts using their real identity, identification numbers and phone numbers, making individual users traceable and accountable for their online content.⁸¹ The real purpose of real-name registration is to improve online accountability, transparency and curb the proliferation of misinformation, cyberbullying and illegal activities.⁸² While the policy contributes to a safer online environment by enhancing accountability, real-name registration is considered to raise ethical and legal issues related to privacy, government control and freedom of expression. It is argued that it could be used by the government as a tool to suppress dissenting voices and individuals critical of the incumbent government regime.⁸³ Furthermore, the fear of potential criminalisation for infractions or surveillance may compel some users to self-censor their opinions and limit their participation in online discussions, thereby impeding freedom of expression.⁸⁴ The task for many jurisdictions in regulating social media is finding the right balance between security and free speech. There is a need to carefully consider the possible harms related to anonymous online activities as well as the significance of safeguarding the rights of individuals to express themselves freely. Thus, any restriction on freedom of expression and privacy must be in accordance with established fundamental legal principles and avoid unwarranted violation.

In October 2025, CAC also introduced a policy prohibiting social media influencers from commenting on sensitive subjects such as medicine, law, education, or finance without official qualifications.⁸⁵ According to this policy, social media influencers shall be required to provide a degree, professional license, and certificate in the respective field before producing content on the aforementioned topics.⁸⁶ The purpose of the regulation is to tackle misinformation and protect users from false advice and recipes circulating online. Chinese authorities believe that unqualified advice on sensitive topics can cause serious harm to society and undermine the accuracy of professional information. These steps mirror the Chinese government's wider efforts to strengthen digital oversight and guarantee that information posted on social media platforms is credible and reliable.

⁷⁵ *ibid.*

⁷⁶ Law Commission. Modernising communications offences: A final report. 2021, HC 547, Law Com 399.

⁷⁷ Coe P. Tackling online false information in the United Kingdom: The online safety act 2023 and its disconnection from free speech law and theory. *Journal of Media Law*. 2023; 15(2):213-242.

⁷⁸ *ibid.*, 232.

⁷⁹ UK Parliament. UK's online safety regime unable to tackle the spread of misinformation and cannot keep users safe online, mps warn, 2025 [Cited 2026 May 28]. Retrieved from:

<https://committees.parliament.uk/committee/135/science-innovation-and-technology-committee/news/208296/uks-online-safety-regime-unable-to-tackle-the-spread-of-misinformation-and-cannot-keep-users-safe-online-mps-warn/>

⁸⁰ *ibid.*

⁸¹ CL 2017, Art. 24.

⁸² Li W, Wang M, Chen Y. Regulation of real-name registration requirements on chinese social media platforms and its impact on freedom of expression. *Law and Economy*. 2023; 2:49.

⁸³ *ibid.*, 51.

⁸⁴ *ibid.*, 51.

⁸⁵ Boufous K. China regulates content creation to curb online misinformation, 2025 [Cited 2026 Jun. 2]. Retrieved from:

<https://www.misbar.com/en/editorial/2025/11/05/china-regulates-content-creation-curb-online-misinformation>

⁸⁶ Osah F. China's new influencer law sparks debate over online freedom. *Vanguard Nigeria*, 2025 [Cited 2026 Jun. 3]. Retrieved from:

<https://www.vanguardngr.com/2025/10/chinas-new-influencer-law-sparks-debate-over-online-freedom/>

However, others argue that the policy gags independent voices and grants the regulator excessive powers to decide who decides who is an expert and capable of giving opinions on the aforementioned sensitive subjects.⁸⁷ Furthermore, it is also argued that the policy could also suppress critical and balanced public debates on key sensitive matters.⁸⁸

7. The Way Forward for Nigeria: The Significance of Regulatory Intervention

A 2019 study by Ogbette *et al.* found that in Nigeria, there is a regulatory gap between the government and social media in areas of monitoring and supervision, and as a result, bad actors exploit loopholes to spread false information and cause tension, in hopes of generating traffic and relevance.⁸⁹ Undoubtedly, mitigating the spread of false information on social media platforms requires a proactive regulatory approach rather than the reactive approach adopted by the X's Community Note, which as stated above is slow in intercepting the early viral stages of posts. The challenge of social media regulation in many jurisdictions has been striking a delicate balance between combating false information and protecting freedom of expression. Consequently, countries such as Nigeria have constitutional limits and laws promoting the freedom of expression, which impede the regulation of content moderation. On the other hand, the European Union is considered a leader in regulating social media platforms, with its efforts to introduce the DSA 2022, aimed at regulating the conduct of large tech platforms such as X, Facebook, TikTok, and several others, while promoting fact-checking and transparent practices. Rather than attempting to replicate China's highly restrictive regulatory model, the Nigerian government could adopt a context-specific approach that prioritises media literacy, supports independent fact-checking initiatives, and enhances transparency in digital governance. This strategy would equip citizens with the skills to critically assess online information, combat the spread of misinformation, and foster responsible digital engagement, while simultaneously safeguarding freedom of expression and preserving an open, inclusive, and balanced online environment. Drawing on the regulatory approaches adopted in the European Union, the United Kingdom, and China, the following recommendations are proposed to address Nigeria's unique social media space on the X platform.

8. Recommendations for Nigeria:

1. Enactment of a Dedicated Social Media and Digital Information Act

Nigeria's current regulatory framework is fragmented, with provisions relating to false information dispersed across the CA 2015 (the Nigerian Communications Act 2003, the Criminal Code Act, the NCC Code 2026, and the NITDA Code of Practice 2022. This fragmented approach creates uncertainty regarding enforcement standards, regulatory responsibilities, and the rights and obligations of users and platforms. There is, therefore, a need to enact a dedicated

⁸⁷ *ibid.*

⁸⁸ Boufous. China regulates content creation to curb online misinformation, 2025.

⁸⁹ Ogbette, Idam, Kareem Ogbette. Fake news in Nigeria: Causes, effect and management. 99.

social media and digital Information Act to regulate the dissemination of false information on digital platforms. Such legislation should provide clear statutory definitions of misinformation, disinformation, fake news, synthetic media, and AI-generated content. The legislation should also establish transparent procedures for content moderation, user appeals, platform accountability, and regulatory oversight. Unlike the abandoned Protection from Internet Falsehood and Manipulation Bill 2019, any future legislation must be carefully drafted to comply with constitutional guarantees of freedom of expression and international human rights standards. The legislation should target harmful false information without criminalising legitimate criticism, investigative journalism, satire, political dissent, or public-interest reporting.

2. Implementation of AI-Based Fact-Checking and Content Authentication Systems

Given the unprecedented rise of artificial intelligence and technological advancement, there is a need for major social media platforms operating within Nigeria to be legally mandated to deploy robust AI-driven content verification mechanisms. The effectiveness of X's Community Notes system is limited by its reactive nature and the significant delay between the publication of misleading content and the eventual display of corrective notes. By the time contextual information is added, the false content may already have reached millions of users. Accordingly, legislation should require large digital platforms to implement proactive AI-powered systems capable of identifying manipulated media, deepfakes, coordinated disinformation campaigns, bot-generated content, and demonstrably false viral information before such content reaches mass circulation. However, automated systems should not operate in isolation. Human oversight, independent auditing, and transparent review procedures must accompany automated moderation processes to prevent over-censorship and algorithmic bias.

3. Mandatory Labelling of AI-Generated and Synthetic Content

Drawing inspiration from the European Union AI Act 2024, Nigeria should require clear disclosure and labelling of AI-generated content. Social media platforms should be mandated to ensure that AI-generated images, videos, audio recordings, and synthetic media are prominently identified as artificially created or manipulated content. This obligation would significantly reduce the deceptive use of deepfakes and fabricated digital media. At the same time, a reasonable degree of exceptions should be made for artistic, satirical, fictional, and creative works, provided that users are adequately informed of the artificial nature. Such a balanced approach would protect freedom of artistic expression while promoting transparency and accountability.

4. Introduce a Co-Regulatory Framework for Platform Oversight

The shortcomings of both complete self-regulation and excessive state control suggest that Nigeria should adopt a co-regulatory model. Under this framework, social media companies would retain responsibility for day-to-day content moderation while remaining subject to regulatory oversight by an independent authority. This model would avoid the dangers associated with unrestricted government

mentorship while simultaneously addressing the inadequacies of purely voluntary moderation systems such as Community Notes. A co-regulatory system would further allow regulators to establish minimum standards for content moderation, transparency, appeals processes, and misinformation management while permitting platforms flexibility in implementation.

5. Introducing a Rights-Based Digital Verification System

Drawing lessons from China's regulatory framework, Nigeria should consider introducing a human rights-compliant digital verification regime for social media users who operate accounts or regularly disseminate information on sensitive subjects such as law, medicine, finance, and public health. Such users could be required to verify their identities with social media platforms and, where they hold themselves out as professionals, disclose their relevant qualifications. This would enhance accountability, reduce the spread of false or misleading information, and discourage anonymous engagement farming. However, unlike the Chinese model, any such framework must be accompanied by robust safeguards protecting privacy, freedom of expression, and due process, ensuring that verification requirements are not used as a tool for censorship or the suppression of legitimate public discourse.

6. Promote Media Literacy and Digital Education

Regulation alone cannot eliminate misinformation. The Federal Government, educational institutions, media organisations, and civil society groups should collaborate to implement nationwide media literacy programmes. These initiatives should equip citizens with skills to: verify information sources; identify manipulated media; recognise AI-generated content; evaluate online credibility; distinguish between facts, opinions, misinformation, and disinformation. Such educational initiatives would strengthen societal resilience against false information while reducing dependence on state intervention.

Conclusion

The proliferation of false information on X presents a significant challenge to Nigeria's democratic development, national security, social cohesion, and information ecosystem. The rise of artificial intelligence, engagement-driven monetisation models, and increasingly sophisticated methods of digital manipulation have amplified the speed, reach, and impact of misinformation and disinformation in unprecedented ways. While X has implemented self-regulatory mechanisms such as Community Notes, empirical evidence demonstrates that these measures remain largely reactive and insufficient to curb the rapid viral spread of misleading content.

Nigeria's existing legal framework, consisting primarily of the Cybercrimes Act 2015, the Nigerian Communications Act 2003, the NCC Code 2026, and the NITDA Code of Practice 2022, provides only fragmented and indirect responses to the problem. These instruments are mostly reactive and not proactive. They neither comprehensively regulate misinformation nor adequately address emerging threats posed by AI-generated content, deepfakes, and engagement-driven disinformation campaigns. At the same time, concerns regarding freedom of expression caution against adopting overly restrictive approaches that could

facilitate censorship or suppress legitimate democratic discourse.

The comparative experiences of the European Union, the United Kingdom, and China demonstrate that effective regulation requires a careful balance between safeguarding information integrity and protecting fundamental rights. While China's model prioritises control and accountability, it raises substantial concerns regarding privacy and freedom of expression. Conversely, the European Union's emphasis on transparency, platform accountability, and risk mitigation provides a more balanced and rights-conscious framework. The United Kingdom's approach further illustrates both the possibilities and limitations of criminalising harmful false communications.

Accordingly, Nigeria should adopt a context-specific co-regulatory framework that combines statutory oversight, platform accountability, independent fact-checking, AI-assisted content verification, transparency obligations, and comprehensive media literacy programmes. Such an approach would strengthen public trust in digital information, deter malicious actors, and promote responsible online engagement while preserving the constitutional freedoms that remain indispensable to a democratic society. Ultimately, the goal of regulation should not be to control speech, but to foster a digital environment in which truth, accountability, and freedom coexist in a manner that protects both individual rights and the collective interests of Nigerian society.

References

1. Ajibola M. Social media regulation: An overview of the national information technology development agency's (NITDA) code of practice for interactive computer service platforms/internet intermediaries. *Lexworth Legal*, 2024 [Cited 2026 May 28]. Retrieved from: <https://www.lexworthlegal.com/social-media-regulation-an-overview-of-the-national-information-technology-development-agencys-nitda-code-of-practice-for-interactive-computer-service-platforms-internet-intermediaries/>
2. Akasike C, Uthman I, Busari B, Adetunji T, Olasupo A, Taiwo B. Mirabel: How false rape alarm turned into money-making venture. *Punch News Paper*, Feb 21, 2026 [Cited 2026 Mar. 25]. Retrieved from: <https://punchng.com/mirabel-how-false-rape-alarm-turned-into-money-making-venture/>
3. Akinwotu E. Nigeria has detained a journalist who reported on corruption in a widening crackdown. *NPR*, 2024 [Cited 2026 May 24]. Retrieved from: <https://www.npr.org/2024/05/09/1250008596/nigeria-detained-daniel-ujukwu-investigative-journalist-corruption>
4. Allcott H, Gentzkow M. Social media and fake news in the 2016 elections. *J Econ Perspect*. 2017; 31:211-236.
5. Ayeni V. How nigerian authorities use cybercrimes act to harass, detain journalists, activists. *Punch Nigeria Newspaper*, 2025 [Cited 2026 May 20]. Retrieved from: <https://punchng.com/how-nigerian-authorities-use-cybercrime-act-to-harass-detain-journalists-activists/>
6. Apuke O, Omar B. Fake news proliferation in Nigeria: Consequences, motivations and prevention through awareness strategies. *Humanities and Social Sciences Reviews*. 2020; 8(2):318-327.
7. Baccarella C, Wagner T, Kietzmann J, McCarthy I.

- Averting the rise of the dark side of social media: The role of sensitisation and regulation. *European Management Journal*. 2020; 38(1):3-6.
8. Berghel H. Lies, damn lies and fake news. *Computer*. 2017; 50(1):80-85.
 9. Boufous K. China regulates content creation to curb online misinformation, 2025 [Cited 2026 Jun. 2]. Retrieved from: <https://www.misbar.com/en/editorial/2025/11/05/china-regulates-content-creation-curb-online-misinformation>
 10. Buarque BL. The x factor: Four ways one social media platform undermined democracy during the 2024 summer riots, 2026 [Cited 2026 May 8]. Retrieved from: <https://www.lse.ac.uk/research/research-for-the-world/society/x-undermined-democracy-uk-riots>
 11. Chin Y. Regulating social media, regulating life (and lives): A report on the workshop 'social media, regulation and freedom of expression' in May at Hong Kong Baptist University. *Columbia Journalism Review*, 2013, 33-37. [Cited 2026 May 15]. Retrieved from: https://www.academia.edu/25901348/Regulating_social_media_regulating_life_and_lives_A_report_on_the_workshop_Social_Media_Regulation_and_Freedom_of_Expression_in_May_at_Hong_Kong_Baptist_University
 12. Chuai Y, Pilarski M, Renault T, Restrepo-Amariles D, Troussel-Clement A, Lenzi G, *et al.* Community-based fact-checking reduced the spread of misleading posts on social media. *arXiv*, 2024 [Cited 2026 May 13] Retrieved from: <https://arxiv.org/pdf/2409.08781>
 13. Coe P. Tackling online false information in the United Kingdom: The online safety act 2023 and its disconnection from free speech law and theory. *Journal of Media Law*. 2023; 15(2):213-242.
 14. Coker O, Adeniran O. Nigeria and the struggle to regulate social media, 2022 [Cited 2026 Mar. 3]. Retrieved from: <https://www.dentonsacaslaw.com/en/insights/articles/2022/january/26/regulation-of-social-media-in-nigeria>
 15. Coli M. The role of the european union in protecting democracy through legislation: The case of disinformation. *European Papers - A Journal on Law and Integration*. 2026; 11(1):81-111.
 16. Corsi G, Marino B, Wong W. The spread of synthetic media on x. *Harvard Kennedy School Misinformation Review*. 2024; 5(30):1. [Cited 2026 May 8]. Retrieved from: <https://misinforeview.hks.harvard.edu/article/the-spread-of-synthetic-media-on-x/>
 17. Egbunike N. Nigeria's social media bill will obliterate online freedom of expression. *Global Voices*, 2020 [Cited 2026 May 21]. Retrieved from: <https://advox.globalvoices.org/2019/11/29/nigerias-social-media-bill-will-obliterate-online-freedom-of-expression/>
 18. European Commission. Disinformation: A threat to democracy – brochure, 2021 [Cited 2026 Mar. 18]. Retrieved from: <https://digital-strategy.ec.europa.eu/en/library/disinformation-threat-democracy-brochure>
 19. Higson-Bliss L. Legislative comment: The online safety act 2023 and the sending of threatening or false communications. *State Law Review*. 2025; 46(2). [Cited 2026 May 28]. Retrieved from: <https://academic.oup.com/slr/article/46/2/hmaf021/8203212>
 20. Ireton C, Posetti J. Journalism, fake news and disinformation: Handbook for journalism education and training. United Nations Digital Library, 2018, 7.
 21. Kirkland C. X has highest rate of misinformation as a news source, study finds. *Media Daily News*, 2024 [Cited 2026 May 8]. Retrieved from: <https://www.mediapost.com/publications/article/396846/x-has-highest-rate-of-misinformation-as-a-news-source.html>
 22. Law Commission. Modernising communications offences: A final report, 2021, HC 547, Law Com 399.
 23. Li W, Wang M, Chen Y. Regulation of real-name registration requirements on chinese social media platforms and its impact on freedom of expression. *Law and Economy*. 2023; 2:49.
 24. Mhaka T. How social media regulations are silencing dissent in Africa. *Aljazeera*, 2020 [Cited 2026 May 21]. Retrieved from: <https://www.aljazeera.com/opinions/2020/11/12/how-social-media-regulations-are-silencing-dissent-in-africa>
 25. Muruaru D, Kombo S. Thumbs up or down: Regulating social media is a double-edged sword. *Busara*, 2026 [Cited 2026 May 24]. Retrieved from: available at <https://busara.global/blog/thumbs-up-or-down-regulating-social-media-is-a-double-edged-sword/>
 26. National Information Technology Development Agency. An analysis of compliance with laws and misinformation management by social media platforms in Nigeria, 2024 [Cited 2026 May 10]. Retrieved from: <https://nitda.gov.ng/wp-content/uploads/2024/12/NITDA24-COP-REPORT-JNT-compressed.pdf>
 27. Ogbette A, Idam M, Kareem A, Ogbette D. Fake news in Nigeria: Causes, effect and management. *Information and Knowledge Management*. 2019; 9:96-99.
 28. Osah F. China's new influencer law sparks debate over online freedom. *Vanguard Nigeria*, 2025 [Cited 2026 Jun. 3]. Retrieved from: <https://www.vanguardngr.com/2025/10/chinas-new-influencer-law-sparks-debate-over-online-freedom/>
 29. Perez S. Twitter, now x, opens up its ad revenue sharing program with global creators. *Tech Crunch*, 2023 [Cited 2026 Mar. 1]. Retrieved from: <https://techcrunch.com/2023/07/28/twitter-now-x-opens-up-its-ad-revenue-sharing-program-with-global-creators/>
 30. Shearer E, Naseer S, Liedke J, Matsa K. How Americans get news on tikTok, x, facebook and instagram. *Pew Research Centre*, 2024 [Cited 2026 May 10]. Retrieved from: <https://www.pewresearch.org/journalism/2024/06/12/how-americans-get-news-on-tiktok-x-facebook-and-instagram/> Accessed 10 May 2026
 31. Slaughter I, Peytavin A, Ugander J, Saveski M. Community notes reduce engagement with and diffusion of false information online. *PNAS*. 2025; 122(38):1-8
 32. Sunny D. Nigerian creators are turning to rage-bait to cash in on x. *TechPoint Africa*, 2026 [Cited 2026 Mar. 1]. Retrieved from: <https://techpoint.africa/news/rage-bait-for-cash-on-x/>

33. Udofa I. Right to freedom of expression and the law of defamation in Nigeria. *International Journal of Advanced Legal Studies and Governance*. 2011; 2(1):78-79.
34. UK Parliament. UK's online safety regime unable to tackle the spread of misinformation and cannot keep users safe online, mps warn, 2025 [Cited 2026 May 28]. Retrieved from: <https://committees.parliament.uk/committee/135/science-innovation-and-technology-committee/news/208296/uks-online-safety-regime-unable-to-tackle-the-spread-of-misinformation-and-cannot-keep-users-safe-online-mps-warn/>
35. Vallance C. Disinformation most active on x, formerly known as twitter, EU says, 2023 [Cited 2026 Mar. 3]. Retrieved from: <https://www.bbc.com/news/technology-66926080>
36. Vallance C. Disinformation most active on x, formerly known as twitter, EU says, 2023 [Cited 2026 Mar. 3]. Retrieved from: <https://www.bbc.com/news/technology-66926080>
37. Wojcik S, Hilgard S, Judd N, Mocanu D, Regain S, Hunzaker M, *et al.* Birdwatch: Crowd wisdom and bridging algorithms can inform understanding and reduce the spread of misinformation, 2022 [Cited 2026 May 10]. Retrieved from: <https://arxiv.org/abs/2210.15723>
38. X Community Notes. Signing up. [Cited 2026 May 8]. Retrieved from: <https://communitynotes.x.com/guide/en/contributing/signing-up>
39. X Help Center. About x community notes on x. [Cited 2026 May 12]. Retrieved from: <https://help.x.com/en/using-x/community-notes#:~:text=Community%20Notes%20aim%20to%20create,%2C%20and/or%20other%20remediations>
40. X Help Centre. Creator revenue sharing. [Cited 2026 Mar. 1]. Retrieved from: <https://help.x.com/en/using-x/creator-revenue-sharing#:~:text=from%20the%20program>
41. X Note ranking algorithm. [Cited 2026 May 12]. Retrieved from: <https://communitynotes.twitter.com/guide/en/under-the-hood/ranking-notes> (2024)