International Journal of Sub-Saharan African Research (IJSSAR)
Vol. 4, Issue 2, pp. 406-422, June 2026, ISSN: 3043-4467 (Online), 3043-4459 (Print)
https://doi:10.5281/zenodo.21065727

Emerging Frontiers of Insecurity: Implications of AI-Driven Deepfake News for National Security in Nigeria

Chidinma Felicia Nwosu

Department of Mass Communication, Ogbonnaya Onu Polytechnic Aba, Abiạ State, Nigeria.

https://orcid.org/0009-0007-1502-2201

***Corresponding Author**: chidinma.nwosu@abiastatepolytechnic.edu.ng

ABSTRACT

Background: Nigeria's digital information environment between 2020 and 2026 has been increasingly disrupted by AI-driven deepfake technology, which poses significant threats to national security, democratic governance, and public trust in institutions. Deepfakes have become more accessible, technically sophisticated, and easier to produce and disseminate. Despite these developments, scholarly research on their security implications within African political contexts remains limited.
Objective: This study examined the implications of AI-driven deepfake news for national security in Nigeria. Specifically, it investigated the role of the media in mitigating the effects of deepfake disinformation, identified the challenges confronting media organisations in combating synthetic content, and proposed a strategic media framework for addressing emerging deepfake threats.
Method: A qualitative exploratory research design was adopted. Primary data were generated through structured brainstorming sessions involving eight purposively selected participants comprising senior journalists, digital verification specialists, and political communication stakeholders with direct experience of Nigeria's 2023 general election. Secondary data were obtained from peer-reviewed literature, policy documents, and credible journalistic reports. Data were analysed thematically using Braun and Clarke's (2006) six-phase thematic analysis framework.
Results: The findings reveal that AI-driven deepfake news constitutes a multidimensional threat to Nigeria's national security. Beyond facilitating deception, deepfakes generate widespread uncertainty, undermine public confidence in the media and state institutions, reinforce confirmation bias within politically polarised environments, and create the *Liar's Dividend* effect, whereby authentic evidence is dismissed as potentially fabricated. The study further found that Nigerian media organisations face significant capacity, technological, financial, and regulatory constraints in responding effectively to these threats.
Conclusion: Addressing the national security risks posed by AI-driven deepfakes requires a coordinated, multi-stakeholder strategy that integrates institutional deepfake detection capabilities, public media literacy initiatives, strengthened regulatory frameworks, and proactive crisis communication protocols. The agenda-setting, investigative, and fact-checking functions of the media remain indispensable for safeguarding national security in the era of synthetic media.
Unique Contribution: This study provides an empirically grounded, Africa-centred analysis of AI-driven deepfake news as an emerging national security challenge. By integrating the concepts of the *Liar's Dividend*, confirmation bias, and structural media capacity deficits within the Nigerian context, it extends the deepfake governance literature beyond predominantly Western perspectives. The study also offers a replicable qualitative framework for examining deepfake-related security challenges in other emerging democracies.

International Journal of Sub-Saharan African Research (IJSSAR)
Vol. 4, Issue 2, pp. 406-422, June 2026, ISSN: 3043-4467 (Online), 3043-4459 (Print)
https://doi:10.5281/zenodo.21065727

Key Recommendation: The study recommends a four-pillar strategic response framework comprising: (1) institutional capacity building through AI-assisted deepfake detection units; (2) sustained public media and digital literacy programmes; (3) the development and enforcement of comprehensive deepfake legislation and regulatory guidelines; and (4) transparent, proactive crisis communication protocols. Effective collaboration among media organisations, regulatory agencies, technology companies, civil society organisations, and government institutions is essential to mitigating the national security risks posed by AI-generated disinformation.

Keywords: AI-driven deepfake news, national security, synthetic media, disinformation, media, information disorder, Nigeria.

INTRODUCTION

Deepfake news represents a new frontier in contemporary media with significant implications for national security. The term *deepfake* combines *deep learning* and *fake* to describe both the artificial intelligence (AI) technology and the synthetic audiovisual content it generates (Heidari et al., 2023; Mubarak et al., 2023). Technically, deepfakes are created using Generative Adversarial Networks (GANs), a dual-algorithm architecture in which a generator produces synthetic content while a discriminator evaluates its authenticity. Through iterative competition, the system progressively generates highly realistic audio, image, and video content that is often indistinguishable from genuine media (Mubarak et al., 2023).

What distinguishes deepfakes from conventional forms of media manipulation is not only their technical sophistication but also their accessibility. The proliferation of open-source AI tools and pre-trained models has significantly lowered the technical barriers to creating convincing synthetic media, enabling individuals with limited expertise to generate and disseminate fabricated content rapidly (Vaccari & Chadwick, 2020). Coupled with the speed and reach of social media platforms, this democratisation of deception allows deepfake news to spread widely before verification mechanisms can intervene. Consequently, deepfakes have emerged as powerful instruments in information warfare, capable of manipulating public opinion, undermining democratic institutions, disrupting crisis communication, and intensifying political and social conflicts (Bonfanti, 2020; Görgen & Saygıner, 2025).

The weaponisation of deepfake technology in political communication is now well documented. Across different geopolitical contexts, including the Russia–Ukraine conflict and electoral processes in several democracies, synthetic media has been deployed to fabricate evidence, discredit political leaders, influence public perception, and distort information ecosystems (Mubarak et al., 2023; Shoaib et al., 2023). Twomey et al. (2023), through a thematic analysis of social media discourse during the Russia–Ukraine war, found that the circulation of deepfake content generated widespread scepticism, leading audiences to question not only fabricated materials but also authentic visual evidence, including documentation of alleged war crimes. This phenomenon demonstrates that the consequences of deepfakes extend beyond deception to the erosion of trust in information itself.

International Journal of Sub-Saharan African Research (IJSSAR)
Vol. 4, Issue 2, pp. 406-422, June 2026, ISSN: 3043-4467 (Online), 3043-4459 (Print)
https://doi:10.5281/zenodo.21065727

A particularly significant consequence of deepfake proliferation is the Liar's Dividend (Chesney & Citron, 2019), whereby the public's awareness of synthetic media enables individuals, especially political actors, to dismiss authentic audio or video evidence as AI-generated fabrications. Rather than merely creating false information, deepfakes foster an environment in which genuine evidence loses credibility, thereby weakening accountability, judicial processes, and public confidence in institutions. This effect is further reinforced by confirmation bias, whereby individuals are more likely to accept fabricated content that aligns with their existing political or ideological beliefs while rejecting contradictory evidence.

Nigeria has become increasingly vulnerable to these emerging threats due to a combination of structural, political, and technological factors. High social media penetration among a youthful population, persistent ethno-religious and political divisions, ongoing security challenges, and the absence of comprehensive legislation regulating AI-generated synthetic media have created favourable conditions for the proliferation of deepfake disinformation. During Nigeria's 2023 general election, fabricated audio recordings falsely attributed to presidential candidates and AI-generated videos depicting international celebrities endorsing particular candidates circulated extensively on social media, exploiting voters' existing political biases and sensitivity to international validation (Punch Newspapers, 2026). Confirmation bias amplified the impact of these fabricated materials, as audiences predisposed to particular political narratives readily accepted synthetic content that reinforced their existing beliefs.

Government institutions have increasingly acknowledged the severity of the threat posed by AI-generated disinformation. The Federal Ministry of Information and National Orientation (FMINO, 2025) warned that deepfake videos were being deployed to damage reputations, undermine governance, and distort public discourse. Similarly, the Chief of Army Staff identified synthetic disinformation as an emerging frontline security challenge, prompting the military to strengthen its strategic communication and media response capabilities (Punch Newspapers, 2025). In May 2026, the Presidency issued an official advisory cautioning Nigerians against organised efforts to weaponise religion through AI-manipulated content ahead of future elections (The Guardian Nigeria News, 2026a). Subsequently, in June 2026, the Grassroots Mobilisation Initiative (GMI) reported that sophisticated actors were cloning the voices of public officials, forging government documents, and producing deepfake videos of religious leaders to incite communal violence (Punch Newspapers, 2026). Reflecting the transnational nature of the challenge, Nigeria has also joined more than sixty countries in international collaborative efforts aimed at addressing the proliferation of deepfakes and AI-generated disinformation (The Journal Nigeria, 2026). Meanwhile, indigenous fact-checking organisations, including PRNigeria, have intensified forensic verification efforts while advocating stronger AI-content labelling systems and improved moderation of indigenous-language content on digital platforms (Vanguard News, 2025).

Despite growing international scholarship on deepfakes, existing research has focused predominantly on Western democracies, with relatively limited empirical attention devoted to African political and security contexts. Furthermore, few studies have examined deepfake news

International Journal of Sub-Saharan African Research (IJSSAR)
Vol. 4, Issue 2, pp. 406-422, June 2026, ISSN: 3043-4467 (Online), 3043-4459 (Print)
https://doi:10.5281/zenodo.21065727

from a national security perspective by integrating media institutions, public trust, confirmation bias, the Liar's Dividend, and the structural capacity of the media to respond effectively. This gap limits understanding of how AI-driven synthetic media interacts with Nigeria's complex information environment and evolving security landscape.

Against this background, this study examines the implications of AI-driven deepfake news for national security in Nigeria. Specifically, it investigates the media's role in mitigating the effects of deepfake disinformation, identifies the structural challenges confronting Nigerian media organisations in combating synthetic content, and proposes a strategic media response framework for addressing emerging deepfake threats. The study draws on structured expert consultations involving media and political communication stakeholders, alongside evidence from recent peer-reviewed literature, policy documents, and verified reports from the Nigerian context.

STATEMENT OF THE PROBLEM

Deepfake technology poses a layered, escalating threat to national security, democratic governance, and public trust. Open-source software and pre-trained AI models enable rapid production and dissemination of synthetic content, reducing the technical barrier for non-experts (Vaccari & Chadwick, 2020). In Nigeria, where mobile internet penetration among youth is high, the structural conditions for viral deepfake dissemination are firmly in place. Yet scholarly literature on deepfake impacts in African political and security contexts remains thin. Detection and mitigation research originates largely in Western institutional settings and may not account for the specific media literacy deficits, regulatory gaps, and political vulnerabilities present in Nigeria. These gaps collectively motivate this study.

RESEARCH QUESTIONS

1. To what extent does AI-driven deepfake news constitute a threat to national security?
2. What roles can the media play in mitigating AI-driven deepfake news?
3. What are the major challenges facing the media in combating deepfake news?

LITERATURE REVIEW

This section examines conceptual, theoretical, and empirical scholarship on deepfake technology, its national security consequences, and the media's role in detection and mitigation, with particular attention to the Nigerian context.

Evolution and Mechanics of AI-Driven Deepfake Technology

Deepfake technology entered public discourse around 2017 when online actors began sharing AI-generated face-swap videos. Since then, the capability has evolved from rudimentary facial substitutions into a sophisticated multi-modal synthetic media infrastructure (Heidari et al., 2023). The primary generative architecture remains the GAN: two competing algorithms one generating synthetic output, the other evaluating its plausibility drive iterative fidelity

International Journal of Sub-Saharan African Research (IJSSAR)
Vol. 4, Issue 2, pp. 406-422, June 2026, ISSN: 3043-4467 (Online), 3043-4459 (Print)
https://doi:10.5281/zenodo.21065727

improvements until outputs become visually indistinguishable from authentic material (Mubarak et al., 2023). Modern production additionally draws on convolutional neural networks for facial tracking, autoencoders for attribute mapping, and natural language processing for voice synthesis, with high-performance computing making near-real-time generation increasingly feasible (Heidari et al., 2023). Progressive open-sourcing of these tools has placed professional-grade deepfake production within reach of political operatives and individual bad actors without specialised technical backgrounds, fundamentally altering the threat landscape (Shoaib et al., 2023).

Deepfake News, National Security, and Democratic Stability

National security scholarship has historically centred on kinetic threats. Synthetic media introduces a distinct category: information-based operations requiring no physical force yet capable of destabilising governance, corroding public trust, and manipulating strategic decision-making (Mubarak et al., 2023). Mubarak et al. (2023) document that deepfakes threaten democratic institutions by enabling fabricated evidence, sophisticated impersonation fraud, and hoaxes capable of deceiving both mass publics and institutional decision-makers simultaneously.

Görgen and Saygıner (2025) analyse a deepfake falsely depicting the German Federal Chancellor, demonstrating that synthetic political content can generate diplomatic confusion and public alarm within hours of circulation. Twomey et al. (2023) establish, through thematic analysis of social media discourse during the Russia-Ukraine conflict, that deepfake content generated pervasive epistemic scepticism leading users to distrust all footage from the conflict zone, including genuine atrocity documentation with evidentiary significance for international accountability mechanisms.

Vaccari and Chadwick (2020), in a large-scale experimental study, found that exposure to deepfake political videos increased audience uncertainty rather than producing outright deception, and that this induced uncertainty significantly eroded trust in social media news more broadly. Hameleers et al. (2026) extended this through multi-country experimental evidence demonstrating that exposure to politically targeted deepfakes reduced support for targeted figures even after fact-check corrections indicating delegitimising effects that persist beyond debunking. Momeni (2024) establishes that citizens frequently cannot identify deepfakes and that political opinions are susceptible to synthetic influence, a finding with direct implications for Nigeria2019s volatile electoral environment.

Wittenberg et al. (2020) introduce an important qualification: deepfakes are not always more persuasive than text-based disinformation, as credibility frequently depends on narrative plausibility rather than production sophistication alone. Hameleers (2024) corroborates this, finding comparably effective delegitimising outcomes from low-production manipulations. Murphy et al. (2025) similarly demonstrate that accessible homemade deepfakes retain significant persuasive potential, cautioning against exclusively technical framings and underscoring the primacy of social and contextual factors.

International Journal of Sub-Saharan African Research (IJSSAR)
Vol. 4, Issue 2, pp. 406-422, June 2026, ISSN: 3043-4467 (Online), 3043-4459 (Print)
https://doi:10.5281/zenodo.21065727

The Liar's Dividend (Chesney & Citron, 2019) represents a secondary and arguably more durable destabilising effect. As deepfake capability becomes widely known, political actors acquire a reflexive instrument for denying authentic evidence. Vaccari and Chadwick (2020) demonstrate that this epistemic corrosion operates independently of direct deception2014uncertainty induced by deepfakes erodes trust in all social media news, including authentic content.

Deepfake News in the Nigerian Context

Nigeria occupies a position of distinctive vulnerability within the global deepfake threat landscape. The convergence of high mobile internet penetration, intense multi-party competition, deep ethno-religious cleavages susceptible to targeted provocation, active insurgencies in which information operations constitute a recognised strategic tool, and a regulatory environment that has not produced deepfake-specific legislation creates structurally exploitable conditions.

Beyond electoral deployment, Nigerian authorities have documented deepfake threats across multiple security domains. The GMI specifically flagged the circulation of AI-manipulated audio and video in indigenous languages Hausa, Yoruba, and Igbo as a deliberate strategy for evading English-language fact-checking infrastructure while inciting community-level violence (Punch Newspapers, 2026). Fabricated footage depicting military personnel improperly protecting combatants was identified as a tactic for undermining public confidence in security forces and providing insurgent groups with propaganda material (Punch Newspapers, 2025). The Presidency's May 2026 advisory named religious disinformation as a specific target vector, accusing organised political actors of deploying AI to weaponise inter-communal sensitivities for electoral purposes (The Guardian Nigeria News, 2026a). Separately, an SK Usman (2026) professional development session delivered to TETFund and NUC public affairs staff under NIPR auspices in Abuja reflected growing institutional recognition of the deepfake competency gap within Nigerian public communications.

International patterns contextualise the Nigerian case. Twomey et al. (2023) document deepfake deployment in the Russia-Ukraine conflict context, while Moyo et al. (2026) identify multi-stakeholder governance frameworks as structurally necessary for any durable deepfake governance architecture. Over 60 nations, including Nigeria, have joined international coalitions targeting deepfake proliferation, though detection tools continue to lag generation capabilities and jurisdictional gaps persist in cross-border enforcement (The Journal Nigeria, 2026).

Media Role in Combating Deepfake News

Watchdog and Verification Functions

The media's investigative and fact-checking capacity constitutes the most direct institutional check on deepfake dissemination. Mubarak et al. (2023) identify consistent debunking as foundational to crisis communication resilience. A structural asymmetry nevertheless persists: fabrications typically achieve viral reach before corrections can be deployed, and delegitimising effects on targeted political figures often survive subsequent debunking (Hameleers et al., 2026). Vanguard News (2025) documented the role of Nigerian fact-checking organisations in forensically analysing viral synthetic content, while experts continue to press technology

International Journal of Sub-Saharan African Research (IJSSAR)
Vol. 4, Issue 2, pp. 406-422, June 2026, ISSN: 3043-4467 (Online), 3043-4459 (Print)
https://doi:10.5281/zenodo.21065727

platforms for improved content moderation and AI-generated content labelling within Nigerian cyberspace.

Media Literacy and Public Education

Scholarly consensus identifies media literacy as indispensable alongside technical detection capabilities (Hwang et al., 2021; Roozenbeek et al., 2022). Hwang et al. (2021) establish that media literacy education reduces deepfake susceptibility, though carelessly designed interventions risk generating generalised distrust in legitimate news. Roozenbeek et al. (2022) demonstrate that psychological inoculation—pre-bunking through attenuated exposure to manipulation techniques at scale—improves audience resilience. Deng and Ahmed (2025) confirm that news literacy skills improve deepfake identification while cautioning against overcorrection that reduces engagement with authentic information.

Agenda-Setting and Collaborative Responses

The media's agenda-setting function determines how deepfake threats are understood by publics and policymakers. Sustained coverage creates the awareness environment within which regulatory reform and literacy initiatives become politically feasible (McCombs & Shaw, 1972). Dobber et al. (2020) document that micro-targeted deepfake exposure amplifies attitude change within susceptible audience subgroups, underscoring the importance of media framing that contextualises targeted effects for broader audiences. Individual media organisations cannot address the deepfake threat unilaterally; Moyo et al. (2026) identify multi-stakeholder governance frameworks integrating media, technology platforms, fact-checkers, and civil society as structurally necessary.

The Media Challenges

Deepfake detection is an adversarial domain: each advance in detection methodology stimulates corresponding evasion innovation, creating an arms race that systematically disadvantages defenders (Westerlund, 2019). Current AI-based detection tools perform well on known manipulation types but degrade significantly against novel generative architectures or low-resolution mobile-optimised output (Heidari et al., 2023; Mubarak et al., 2023). Most African media organisations lack the financial capacity and specialist personnel to deploy AI-assisted verification workflows, relying instead on human visual detection that generation technology has already partially rendered unreliable.

Politically, journalists engaging in deepfake-related verification face targeted harassment and platform de-amplification from actors whose interests the synthetic content serves. Legally, the absence of Nigeria-specific deepfake legislation leaves media organisations without institutional protection and limits accountability mechanisms against producers and distributors of harmful synthetic content (Chesney & Citron, 2019). The Liar2019s Dividend extends into media practice itself: when any recording can be credibly challenged as a possible fabrication, the evidentiary value of authentic documentation declines, complicating investigative journalism (Chesney & Citron, 2019).

International Journal of Sub-Saharan African Research (IJSSAR)
Vol. 4, Issue 2, pp. 406-422, June 2026, ISSN: 3043-4467 (Online), 3043-4459 (Print)
https://doi:10.5281/zenodo.21065727

THEORETICAL FRAMEWORK

This paper is anchored in two complementary theoretical frameworks: Agenda-Setting Theory and the Information Disorder Framework. Agenda-Setting Theory, developed by McCombs and Shaw (1972) from their foundational study of the 1968 United States presidential election, posits that media organisations, through selective issue emphasis, substantially shape what audiences regard as important. Applied to deepfake news, the theory operates bidirectionally: media prioritisation of deepfake threats generates public salience enabling regulatory and institutional responses, while simultaneously deepfake content itself functions as an agenda-setting instrument by making fabricated events appear credible and prominent.

The Information Disorder Framework, proposed by Wardle and Derakhshan (2017) and adopted by the Council of Europe, distinguishes misinformation (false content shared without harmful intent), disinformation (false content deliberately created and distributed to cause harm), and malinformation (accurate content shared to cause harm). AI-driven deepfake news deployed against Nigerian security institutions, political candidates, or religious communities constitutes disinformation in the strictest sense. However, once initially distributed, deepfake content frequently circulates as misinformation among secondary audiences who are themselves deceived. The framework's attention to agents, messages, and interpreters provides vocabulary for designing responses calibrated to the full threat architecture rather than targeting content in isolation.

METHODOLOGY

Research Design

This study adopted a qualitative exploratory research design, which is appropriate for investigating the emerging phenomenon of AI-driven deepfake news within African political and national security contexts, where empirical evidence remains limited. The design is grounded in the interpretivist paradigm, which views knowledge as socially constructed and context-dependent, with meanings shaped by participants' professional experiences and perspectives (Denzin & Lincoln, 2018). An exploratory qualitative approach was considered suitable because the study sought to generate in-depth, expert-informed insights into the implications of deepfake news for national security in Nigeria, rather than to test predetermined hypotheses or establish statistical relationships (Creswell & Poth, 2018).

Data Collection

Primary data were gathered through purposive expert consultations with eight (8) participants, selected on the principle of information richness rather than statistical representativeness (Patton, 2015). The sample comprised two subgroups: four (4) senior journalists and media practitioners with demonstrable experience in digital verification and fact-checking; and four (4) political stakeholders including campaign managers and communications professionals with direct operational involvement in the 2023 Nigerian presidential election cycle. Structured brainstorming sessions were facilitated using an open-ended topic guide organised around the research questions. Sessions were audio-recorded with participants' informed consent and transcribed verbatim. Secondary data were drawn from peer-reviewed academic literature, policy

International Journal of Sub-Saharan African Research (IJSSAR)
Vol. 4, Issue 2, pp. 406-422, June 2026, ISSN: 3043-4467 (Online), 3043-4459 (Print)
https://doi:10.5281/zenodo.21065727

documents, and verified journalistic reports, identified through systematic searches on Google Scholar, JSTOR, and ResearchGate using search terms including: deepfake, synthetic media, national security Nigeria, AI disinformation, and information disorder.

DATA ANALYSIS

Qualitative data were analysed using the six-phase thematic analysis framework of Braun and Clarke (2006): data familiarisation through repeated transcript reading; systematic generation of preliminary codes; organisation of codes into candidate themes; review of themes against the coded data; definition and naming of final themes; and production of the analytical report. Theme development was primarily inductive, grounded in participants' own framings, while theoretical frameworks guided interpretive inferences. Analytical rigour was strengthened through peer debriefing and member-checking with two participants, who reviewed interpretive summaries for accuracy and representational fairness.

Scope and Limitations

The study's geographic focus is Nigeria, with comparative reference to selected African and international contexts where relevant. The purposive sample of eight participants, appropriate for qualitative depth, does not support statistical generalisation. Findings are presented as contextually grounded, expert-informed insights with potential transferability to comparable emerging-economy contexts. A further limitation is the rapidly evolving technological landscape: specific technical assessments reflect conditions as of early-to-mid 2026.

RESULTS

Findings from participant consultations, read against the empirical literature, are organised around four thematic clusters corresponding to the research questions.

Theme 1: AI-Driven Deepfake News as a Multidimensional National Security Threat

All eight participants characterised AI-driven deepfake news as an immediate rather than future security concern. Media professionals consistently identified the speed-verification asymmetry as the defining operational challenge: fabricated content achieves viral reach before institutional correction processes are complete, and corrections rarely recover comparable audience penetration. This finding aligns with Hameleers et al. (2026), who establish that deepfake-induced delegitimisation of political figures persists even after correction exposure, and with Vaccari and Chadwick (2020), who demonstrate that induced epistemic uncertainty erodes trust in authentic media alongside fabricated content.

Political stakeholder participants provided granular accounts of deepfake deployment during the 2023 presidential election. The Liar's Dividend mechanism (Chesney & Citron, 2019) was empirically observable: as deepfake incidents accumulated, the rhetorical space for dismissing genuine recordings as potential fabrications correspondingly expanded, raising the evidentiary threshold beyond what fact-checkers could routinely meet. Momeni (2024) corroborates these accounts, finding that political opinions are measurably shaped by synthetic content exposure even when audiences are uncertain about authenticity.

International Journal of Sub-Saharan African Research (IJSSAR)
Vol. 4, Issue 2, pp. 406-422, June 2026, ISSN: 3043-4467 (Online), 3043-4459 (Print)
https://doi:10.5281/zenodo.21065727

Journalists described deepfake content circulated specifically in Hausa, Yoruba, and Igbo to bypass English-language fact-checking infrastructure a deliberate linguistic targeting strategy corresponding to the GMI's documented concerns (Punch Newspapers, 2026). Military and security dimensions were equally prominent: fabricated footage depicting military personnel acting improperly in operational contexts was identified as a tactic for undermining public confidence in security forces and providing insurgent groups with propaganda material (Punch Newspapers, 2025). The Guardian Nigeria News (2026b) separately documented the GMI's formal warning that fake news and deepfakes posed a serious and growing threat to national security infrastructure.

Theme 2: The Media's Role in Mitigating Deepfake Threats

Participants identified four primary media functions in deepfake response: verification and fact-checking; public media literacy education; agenda-setting and institutional framing; and collaborative platform governance. These functions correspond closely to roles identified in the empirical literature, affirming the media's centrality to any sustainable counter-deepfake governance architecture.

Verification was identified as most critical and most resource-constrained. Participants described growing reliance on technical artefact identification visual inconsistencies in blinking patterns, edge blurring, and audio-visual desynchronisation while acknowledging that these heuristics were becoming unreliable against high-fidelity outputs from current generative architectures. This aligns with Heidari et al. (2023), who establish that detection methods degrade significantly against novel deepfake architectures. Several participants noted the absence of institutional access to AI-assisted forensic tools, leaving their organisations dependent on approaches the technology had already partially rendered obsolete.

Media literacy was unanimously endorsed as an indispensable long-term component. Participants found Roozenbeek et al.'s (2022) pre-bunking approach theoretically compelling but practically underdeveloped in the Nigerian context. Vaccari and Chadwick (2020) and Twomey et al. (2023) collectively underscore that empathetic, transparency-oriented crisis communication is an additional media competency—one addressing the emotional and social processes through which synthetic content shapes collective perception, beyond factual rebuttal. Participants corroborated this from professional experience: cold empirical corrections to emotionally resonant deepfakes frequently failed to dislodge already-formed impressions.

Theme 3: Structural Challenges Confronting the Media

Resource constraints emerged as the most consistently reported challenge. The gap between the technical sophistication required for reliable deepfake detection and the capabilities accessible to Nigerian newsrooms particularly regional and community outlets was characterised as systemic rather than incidental. AI-assisted verification infrastructure, specialist training, and detection database subscriptions represent expenditures most Nigerian media organisations cannot sustain within current commercial models.

Political pressure and regulatory uncertainty constituted the second major challenge cluster. Multiple participants described incidents in which deepfake-related investigative work attracted

International Journal of Sub-Saharan African Research (IJSSAR)
Vol. 4, Issue 2, pp. 406-422, June 2026, ISSN: 3043-4467 (Online), 3043-4459 (Print)
https://doi:10.5281/zenodo.21065727

targeted harassment from actors whose interests were served by the synthetic content's credibility. The absence of Nigeria-specific deepfake legislation was identified as a dual failure: it leaves media organisations unprotected and limits accountability mechanisms against producers and distributors of harmful synthetic content.

The arms race dynamic received prominent attention. Published detection research functions inadvertently as an evasion development roadmap, creating structural asymmetry in which defensive capacity is perpetually reactive rather than anticipatory (Westerlund, 2019; Mubarak et al., 2023). Participants noted that homemade deepfakes of limited technical quality what Murphy et al. (2025) characterise as accessible fabrications were often harder to debunk publicly than sophisticated outputs, because their imperfections were attributable to video quality rather than manipulation artefacts.

DISCUSSION

This study set out to examine the implications of AI-driven deepfake news on national security in Nigeria. Findings from participant consultations and the empirical literature converge on a coherent picture: deepfakes constitute a real, present, and structurally embedded threat operating through mechanisms that are simultaneously technical, sociological, and political.

First, the threat is fundamentally sociological. Deepfakes are operationally effective in Nigeria not because Nigerian audiences are uniquely credulous, but because they are deployed with acute awareness of existing fault lines—ethno-religious tensions, political polarisation, linguistic community boundaries, and documented susceptibilities such as sensitivity to international validation. The GMI's documented concern about indigenous-language targeting (Punch Newspapers, 2026) confirms that technical detection alone cannot address this dimension: it requires complementary investment in indigenous-language media literacy and platform moderation.

Second, the Liar's Dividend (Chesney & Citron, 2019) may represent a more durable threat than individual deepfake incidents. When generalised awareness of deepfake capability enables routine dismissal of authentic evidence as potential fabrication empirically documented through the 2023 Nigerian election cycle (Punch Newspapers, 2026) and corroborated experimentally by Vaccari and Chadwick (2020) the informational power balance shifts structurally in favour of bad actors. No individual fact-check can restore the epistemic authority of authentic documentation once AI fabrication's general plausibility is normalised. This reinforces the case for pre-bunking strategies (Roozenbeek et al., 2022) that build audience resilience before specific fabrications circulate, rather than relying exclusively on reactive debunking.

Third, the institutional response remains structurally reactive and insufficiently coordinated. Government advisories from the Presidency (The Guardian Nigeria News, 2026a), military statements from the COAS (Punch Newspapers, 2025), civil society warnings from the GMI (Punch Newspapers, 2026), and fact-checker debunks all operate after disinformation has achieved initial circulation. Consistent with the Wardle and Derakhshan (2017) Information Disorder Framework, effective response must address agents, messages, and interpreters

International Journal of Sub-Saharan African Research (IJSSAR)
Vol. 4, Issue 2, pp. 406-422, June 2026, ISSN: 3043-4467 (Online), 3043-4459 (Print)
https://doi:10.5281/zenodo.21065727

simultaneously a systemic challenge requiring governance innovation rather than piecemeal reactive communication.

The finding that media capacity building constitutes a national security investment not merely a professional development concern deserves particular emphasis. When Nigerian newsrooms lack detection infrastructure, when journalists face political harassment for deepfake-related reporting, and when regulatory frameworks provide neither investigator protection nor producer accountability, the media's structural capacity to discharge its indispensable verification, literacy, and agenda-setting functions is fundamentally compromised.

A STRATEGIC MEDIA FRAMEWORK FOR COMBATING DEEPFAKE NEWS

Drawing on study findings, participant recommendations, and the convergent policy literature, this paper proposes a four-pillar strategic media framework.

Pillar 1: Institutional Detection Capacity

Public broadcasters, national newspapers of record, and digital verification organisations should invest in dedicated digital forensics units staffed with AI-assisted detection-trained personnel. Partnerships with technology companies, international fact-checking networks, and academic AI research centres can supplement in-house capacity, providing access to continuously updated detection tools and shared incident databases. Moyo et al. (2026) identify AI-driven detection infrastructure as operationally essential to any durable national security deepfake response architecture.

Pillar 2: Coordinated Public Media Literacy

A sustained, multi-channel public education campaign should equip Nigerian news consumers with practical deepfake identification skills, calibrated to distinct audience segments including populations with limited formal education reached through community radio and religious institutions. Content should encompass recognition of common generation artefacts, source verification practices, and critical engagement with emotionally provocative viral content. Pre-bunking approaches validated by Roozenbeek et al. (2022) should complement reactive media literacy to build advance resilience, while Moyo et al. (2026) recommend formal curriculum integration for long-term structural impact.

Pillar 3: Regulatory Advocacy and Legislative Reform

Media organisations and professional associations should coordinate advocacy for clear, enforceable legislation governing the production and distribution of harmful synthetic AI-generated media, while preserving protections for satire and legitimate creative expression. International regulatory models including the EU's Digital Services Act and US state-level deepfake legislation provide applicable reference points for Nigerian policy development (Görgen & Saygıner, 2025; Shoaib et al., 2023). The National Information Technology Development Agency (NITDA) and the National Broadcasting Commission (NBC) should receive specific mandates and enforcement instruments addressing AI-generated disinformation.

International Journal of Sub-Saharan African Research (IJSSAR)
Vol. 4, Issue 2, pp. 406-422, June 2026, ISSN: 3043-4467 (Online), 3043-4459 (Print)
https://doi:10.5281/zenodo.21065727

Pillar 4: Transparent Pre-emptive Crisis Communication

Drawing on Mubarak et al. (2023) and Moyo et al. (2026), media organisations and government communication units should develop pre-positioned response protocols for deepfake incidents involving public figures, electoral information, military operations, and critical infrastructure announcements. These protocols should prioritise speed of verified response, transparency about evidence standards, and empathetic framing addressing the emotional dimensions of synthetic media exposure rather than restricting response to factual rebuttal alone. Pre-positioning such protocols before anticipated threat windows particularly election periods is structurally essential to effective crisis communication.

CONCLUSION

This paper has documented that AI-driven deepfake news constitutes a real, present, and escalating threat to national security, democratic governance, and institutional trust in Nigeria. The threat operates through direct deception of individual audiences, systemic epistemic uncertainty corroding trust in authentic information, confirmation bias activation through socially calibrated synthetic content, and the Liar's Dividend that furnishes political actors with generalised deniability for authentic evidence. Nigeria's specific socio-political landscape characterised by ethno-religious tensions, active insurgencies, intense electoral competition, and linguistic diversity creating verification blind spots renders it particularly vulnerable.

The media occupies an indispensable structural position in the response architecture, exercising verification, literacy, agenda-setting, and collaborative governance functions that no other institutional actor can substitute. Yet Nigerian media organisations face profound resource, technical, political, and regulatory obstacles substantially limiting their capacity to discharge these functions effectively. Closing this gap requires investment from state, civil society, and international partners that treats media institutions as national security infrastructure rather than merely commercial or cultural actors.

The epistemic corrosion produced by deepfakes the normalisation of generalised scepticism toward all audiovisual evidence may ultimately represent a more durable threat to democratic culture than any individual fabricated video. Reversing this corrosion requires not only technical solutions and regulatory frameworks but a sustained, society-wide recommitment to evidentiary standards and critical media engagement. The media's intensified pursuit of its agenda-setting, investigative, and fact-checking functions is both a professional obligation and a national security imperative.

Ethical Clearance

Ethical consent was sought and obtained from all participants. They were informed that participation was purely for academic purposes and entirely voluntary, and that no personally identifying information would be disclosed in any publication arising from the study.

Acknowledgements

The author acknowledges the Association of Communication Researchers of Nigeria (AMCRON) conference, where this paper was first presented, and the media professionals and political communications stakeholders who contributed their time and expertise to the study.

Sources of Funding

The study was not funded.

Conflict of Interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Author's Contribution

Chidinma Felicia Nwosu conceived the study, designed the data collection instruments, conducted the expert consultations, performed the thematic analysis, and authored the full manuscript. The author has reviewed and approved the final draft and takes responsibility for the content and accuracy of the manuscript.

Availability of Data and Materials

The data supporting the conclusions of this study are available from the corresponding author on reasonable request.

Citation

Nwosu, C. F. (2026). New frontiers in insecurity: The implications of AI-driven deepfake news on national security. International Journal of Sub-Saharan African Research.

REFERENCES

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative Research in Psychology, 3(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Chesney, R., & Citron, D. K. (2019). Deep fakes: A looming challenge for privacy, democracy, and national security. California Law Review, 107(6), 1753–1819.

Creswell, J. W., & Poth, C. N. (2018). Qualitative inquiry and research design: Choosing among five approaches (4th ed.). Sage.

Deng, R., & Ahmed, S. (2025). Breaking the illusion: Impact of news literacy on deepfake identification and sharing. Online Information Review, 49, 1211–1230. https://doi.org/10.1108/OIR-10-2024-0613

Denzin, N. K., & Lincoln, Y. S. (Eds.). (2018). The SAGE handbook of qualitative research (5th ed.). Sage.

Dobber, T., Metoui, N., Trilling, D., Helberger, N., & de Vreese, C. H. (2020). Do (microtargeted) deepfakes have real effects on political attitudes? The International Journal of Press/Politics, 26(1), 69–91. https://doi.org/10.1177/1940161220944364

Federal Ministry of Information and National Orientation. (2025, May 23). Information minister decries use of deepfakes to undermine public trust. https://fmino.gov.ng/latest-news

Gürgen, A., & Saygıner, C. (2025). Deepfake technology, media, and national security: The case of the German Chancellor's deepfake video. Güvenlik Stratejileri Dergisi. https://doi.org/10.17752/guvenlikstrtj.1706990

Hameleers, M. (2024). Cheap versus deep manipulation: The effects of cheapfakes versus deepfakes in a political setting. International Journal of Public Opinion Research. https://doi.org/10.1093/ijpor/edae004

Hameleers, M., van der Meer, T. G. L. A., & Dobber, T. (2022). You won't believe what they just said! The effects of political deepfakes embedded as vox populi on social media. Social Media + Society, 8(3). https://doi.org/10.1177/20563051221116346

Hameleers, M., van der Meer, T. G. L. A., & Dobber, T. (2023). They would never say anything like this! Reasons to doubt political deepfakes. European Journal of Communication, 39(1), 56–70. https://doi.org/10.1177/02673231231184703

Hameleers, M., van der Meer, T. G. L. A., Tulin, M., & Dobber, T. (2026). Radical right-wing political deepfakes can successfully delegitimize targeted political actors: Evidence from three-wave experiments in the US and the Netherlands. Communication Research. https://doi.org/10.1177/00936502261421437

Heidari, A., Navimipour, N. J., Dağ, H., & Ünal, M. (2023). Deepfake detection using deep learning methods: A systematic and comprehensive review. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 14. https://doi.org/10.1002/widm.1520

Hwang, Y., Ryu, J., & Jeong, S. (2021). Effects of disinformation using deepfake: The protective effect of media literacy education. Cyberpsychology, Behavior, and Social Networking, 24(3), 188–193. https://doi.org/10.1089/cyber.2020.0174

McCombs, M. E., & Shaw, D. L. (1972). The agenda-setting function of mass media. Public Opinion Quarterly, 36(2), 176–187. https://doi.org/10.1086/267990

Momeni, M. (2024). Artificial intelligence and political deepfakes: Shaping citizen perceptions through misinformation. Journal of Creative Communications, 20(1), 41–56. https://doi.org/10.1177/09732586241277335

Mubarak, R., Alsboui, T. A. A., Alshaikh, O., Inuwa-Dutse, I., Khan, S., & Parkinson, S. (2023). A survey on the detection and impacts of deepfakes in visual, audio, and textual formats. IEEE Access, 11, 144497–144529. https://doi.org/10.1109/ACCESS.2023.3344653

Murphy, G., Ching, D., Meehan, E., Twomey, J., Bolger, A., & Linehan, C. (2025). An average Joe, a laptop, and a dream: Assessing the potency of homemade political deepfakes. Applied Cognitive Psychology. https://doi.org/10.1002/acp.70061

Moyo, B. V., Tuyikeze, T., Matsebula, F., & Obagbuwa, I. (2026). An AI-driven conceptual framework for detecting fake news and deepfake content: A systematic review. Frontiers in Artificial Intelligence, 9. https://doi.org/10.3389/frai.2026.1737790

Patton, M. Q. (2015). Qualitative research and evaluation methods (4th ed.). Sage.

Punch Newspapers. (2025, October 7). Fake news, deepfakes now threaten national security—COAS. https://punchng.com

Punch Newspapers. (2026, June 5). AI-driven fake news, deepfakes threaten national security. https://punchng.com

Roozenbeek, J., van der Linden, S., Goldberg, B., Rathje, S., & Lewandowsky, S. (2022). Psychological inoculation improves resilience against misinformation on social media. Science Advances, 8(34). https://doi.org/10.1126/sciadv.abo6254

Shoaib, M. R., Wang, Z., Ahvanooey, M. T., & Zhao, J. (2023). Deepfakes, misinformation, and disinformation in the era of frontier AI, generative AI, and large AI models. In Proceedings of the 2023 International Conference on Computer and Applications (ICCA) (pp. 1–7). https://doi.org/10.1109/ICCA59364.2023.10401723

The Guardian Nigeria News. (2026a, May 29). Presidency warns Nigerians against deepfake videos, religious disinformation. https://guardian.ng

The Guardian Nigeria News. (2026b, June 5). Fake news poses security threat, GMI warns. https://guardian.ng

The Journal Nigeria. (2026, March 4). Over 60 nations, including Nigeria, target deepfake proliferation. https://thejournalnigeria.com

Twomey, J., Ching, D., Aylett, M., Quayle, M., Linehan, C., & Murphy, G. (2023). Do deepfake videos undermine our epistemic trust? A thematic analysis of tweets that discuss deepfakes in the Russian invasion of Ukraine. PLOS ONE, 18. https://doi.org/10.1371/journal.pone.0291668

Usman, S. K. (2026, May). Sharing insights on deepfake threats, misinformation and crisis communication with TETFund and NUC public affairs staff in Abuja [Facebook post]. https://facebook.com

Vaccari, C., & Chadwick, A. (2020). Deepfakes and disinformation: Exploring the impact of synthetic political video on deception, uncertainty, and trust in news. Social Media + Society, 6(1). https://doi.org/10.1177/2056305120903408

Vanguard News. (2025, July 7). Combating deepfake videos: A fight for Nigeria's soul. https://vanguardngr.com

Wardle, C., & Derakhshan, H. (2017). Information disorder: Toward an interdisciplinary framework for research and policymaking (Council of Europe Report DGI(2017)09). https://rm.coe.int/information-disorder-report/168076277c

Westerlund, M. (2019). The emergence of deepfake technology: A review. Technology Innovation Management Review, 9(11), 39–52. https://doi.org/10.22215/TIMREVIEW/1282

Wittenberg, C., Tappin, B. M., Berinsky, A. J., & Rand, D. G. (2021). The (minimal) persuasive advantage of political video over text. Proceedings of the National Academy of Sciences of the United States of America, 118. https://doi.org/10.1073/pnas.2114388118Note: