INTERNATIONAL COOPERATION AND INFORMATION RESILIENCE IN COUNTERING DISINFORMATION: LESSONS FROM HYBRID CONFLICTS IN THE MIDDLE EAST AND EASTERN EUROPE

Maria SMARANDACHE
National University of Political Studies and Public Administration (SNSPA), Romania
mariasmarandache22@gmail.com

Abstract: The rapid evolution of digital technologies has fundamentally transformed the information environment, increasing the scale and complexity of disinformation campaigns and hybrid threats. Information manipulation has become a strategic instrument employed by both state and non-state actors to influence public opinion, undermine trust in democratic institutions and shape geopolitical outcomes. Consequently, strengthening information resilience has emerged as a central objective of international security policies. This article examines the role of international cooperation in countering disinformation through a comparative analysis of the mechanisms developed by the European Union, NATO, the United Nations, UNESCO and the G7. It further analyses two representative cases of contemporary hybrid conflict — information operations surrounding the Israel–Hamas war and the 2024–2025 Romanian electoral cycle — supported by documented, source-attributed examples of coordinated inauthentic behaviour and generative-AI content. Based on this analysis, the article proposes a conceptual framework for international information resilience built upon five complementary pillars: prevention, early detection, coordinated response, societal resilience and democratic governance, and specifies concrete safeguards for reconciling this framework with freedom of expression. The proposed model argues that effective responses to disinformation require integrated cooperation among governments, international organizations, technology platforms, academia and civil society rather than isolated national initiatives.
Keywords: disinformation, information resilience, hybrid threats, international cooperation, artificial intelligence, strategic communication, Middle East, Romania.

INTRODUCTION

The digital transformation of the global information environment has profoundly altered the nature of contemporary security challenges. Advances in social media, algorithmic content distribution and artificial intelligence have enabled information to instantly circulate across national borders, creating unprecedented opportunities for communication while simultaneously facilitating the rapid dissemination of false or manipulated content. As a result, disinformation has evolved into a strategic component of hybrid conflicts, extending far beyond the traditional concept of misinformation (World Economic Forum, 2024).

Unlike accidental misinformation, disinformation involves the deliberate creation and dissemination of deceptive content intended to influence public perceptions, weaken institutional credibility and manipulate political or social behaviour (Wardle & Derakhshan, 2017; Bennett & Livingston, 2018). This is not a wholly new phenomenon: disinformation has long served as an instrument of political warfare (Rid, 2020), and recent state-sponsored

campaigns have revived high-volume, multichannel propaganda techniques such as the so-called "firehose of falsehood" model (Paul & Matthews, 2016). Recent geopolitical developments have demonstrated that information operations increasingly accompany diplomatic, economic and military actions, transforming the information environment into an additional domain of strategic competition.

The increasingly sophisticated digital manipulation has challenged governments and international organizations to develop coordinated responses capable of protecting democratic processes without undermining fundamental rights such as freedom of expression. Consequently, international cooperation has become an essential element of contemporary information security, requiring coordination among public institutions, private technology companies, researchers and civil society organizations.

Although numerous studies have examined individual aspects of disinformation, many focus either on technological solutions or on specific institutional responses. Comparatively fewer studies integrate international governance mechanisms with empirical, source-attributed observations drawn from contemporary hybrid conflicts. This article addresses that gap by combining a comparative assessment of international initiatives with a documented analysis of information operations in two contrasting theatres: the Middle East and Eastern Europe.

The research is guided by the following question: How can international cooperation strengthen information resilience against disinformation in the context of contemporary hybrid conflicts, and what concrete mechanisms allow such cooperation to be reconciled with democratic freedoms?

To answer this question, the study pursues four objectives. First, it examines the institutional mechanisms developed by major international organizations to counter disinformation. Second, it analyses documented information operations from the Israel–Hamas conflict. Third, it compares these with the hybrid interference documented during Romania's 2024–2025 electoral cycle, in order to test whether the same analytical categories travel across a non-state-dominated conflict environment and a state-sponsored electoral-interference environment. Fourth, it proposes a conceptual framework for strengthening international information resilience, including explicit safeguards for freedom of expression, integrating the principal findings derived from the comparative analysis.

Rather than focusing exclusively on content moderation or technological countermeasures, the article argues that resilience against disinformation depends upon coordinated governance, institutional cooperation, transparent economic regulation of platforms, and the capacity of democratic societies to critically evaluate information.

THEORETICAL FRAMEWORK AND RESEARCH METHODOLOGY

Disinformation has become a multidisciplinary research topic situated at the intersection of international security, communication studies, political science and information technology (Freeman, 2020). Contemporary literature increasingly describes information as a strategic resource whose manipulation may generate political, social and economic consequences comparable to those produced by conventional instruments of power.

Within this context, the concept of information resilience has emerged as a broader analytical framework compared to traditional approaches that focus exclusively on detecting false information. Information resilience refers to the capacity of institutions and societies to anticipate, withstand, adapt to and recover from coordinated information manipulation while preserving democratic values and public trust. This perspective shifts attention from isolated incidents of disinformation toward the systemic capabilities required to reduce societal vulnerabilities (Bucovețchi, Voipan, Voipan, & Stanciu, 2025).

Recent advances in artificial intelligence further complicate the information environment. Generative AI technologies facilitate the rapid production of realistic text, images, audio and video content, increasing both the scale and sophistication of influence operations (OECD, 2024). At the same time, artificial intelligence also provides important opportunities for detecting coordinated campaigns, identifying synthetic media and supporting early-warning systems. Consequently, AI should be understood simultaneously as a source of emerging risks and as an essential component of future resilience strategies.

Methodologically, this research adopts a qualitative comparative approach. The analysis combines three complementary methods:

- comparative analysis of policies and institutional mechanisms developed by the European Union, NATO, the United Nations, UNESCO and the G7;
- case study analysis of documented information operations from two hybrid-conflict environments — the Israel–Hamas war and Romania's 2024–2025 electoral cycle;
- conceptual synthesis leading to the development of a proposed framework for international information resilience, including explicit democratic safeguards.

To operationalise this comparative approach and address the need for a transparent data-selection procedure, the analysis draws on a bounded corpus of primary and secondary sources collected for the period January 2020 to July 2026. For the institutional comparison (Section 3), the corpus consists of official strategy documents, regulations, communiqués and rapid-response publications issued directly by the EU, NATO, the UN, UNESCO and the G7. For the case studies (Section 4), the corpus consists of: declassified or officially released government assessments; platform transparency disclosures; and investigative reports produced by established digital-forensics and fact-checking organizations (e.g., the Atlantic Council's Digital Forensic Research Lab, Cyabra, BBC Verify), covering the Israel–Hamas conflict from October 2023 onward and the Romanian presidential and parliamentary election cycle from November 2024 to May 2025. Sources were selected purposively, prioritising primary institutional documents where available, and empirical claims describing the scale of a given campaign (e.g., the number of accounts in a network) were retained only where corroborated by at least two independent sources.

Rather than a large-N statistical content analysis, the study applies structured, focused comparison (George & Bennett, 2005) combined with elements of process tracing: the mandates and instruments that international organizations state in official documents (Section 3) are traced against the operational practices and gaps observed in the two case studies (Section 4), which allows the analysis to identify where declared institutional mechanisms produced a verifiable effect on the ground and where a gap between stated policy and documented practice

persists. This bounded, qualitative design does not support statistical generalisation across all hybrid conflicts; its purpose is analytical and theory-building, and this limitation, together with directions for a more systematic empirical test, is discussed in Section 7.

The selected organizations represent the principal international actors involved in developing coordinated responses to disinformation. The Middle East and Eastern Europe were selected as contrasting cases: the former illustrates information operations conducted primarily by non-state and diffuse actors within an active armed conflict, while the latter illustrates state-sponsored, platform-mediated interference in a peacetime democratic electoral process. Analysing both allows the study to examine whether the same institutional lessons and the same five-pillar framework apply across qualitatively different hybrid-threat environments.

Rather than evaluating the effectiveness of individual policies in isolation, the study identifies common principles, institutional convergences and governance mechanisms that may contribute to a comprehensive international architecture for countering disinformation.

INTERNATIONAL COOPERATION AGAINST DISINFORMATION: A COMPARATIVE ANALYSIS OF INSTITUTIONAL APPROACHES

The growing sophistication of disinformation campaigns has demonstrated that no state possesses sufficient capabilities to address information threats independently. Digital platforms operate across national borders, information spreads almost instantaneously, and malicious actors increasingly coordinate influence operations at regional and global levels (Bradshaw & Howard, 2019). Consequently, international organizations have progressively developed complementary mechanisms aimed at strengthening information resilience through institutional cooperation, strategic communication and coordinated policy responses.

Although the European Union (EU), the North Atlantic Treaty Organization (NATO), the United Nations (UN), UNESCO and the G7 pursue different mandates, their initiatives reveal significant convergence regarding the need for collective responses to information manipulation. Rather than relying exclusively on content removal, these organizations increasingly emphasize resilience, transparency, institutional coordination and public awareness as fundamental components of democratic information security.

Comparative Assessment of International Mechanisms

The European Union has developed one of the most comprehensive regulatory frameworks addressing disinformation. Its approach combines legislative measures, strategic communication, digital platform accountability and support for independent fact-checking initiatives. Instruments such as the Digital Services Act and the European Democracy Action Plan illustrate the EU's transition from reactive responses towards preventive governance aimed at increasing transparency within digital ecosystems (European Commission, 2020, 2022; Durach, Bârgăoanu, & Nastasiu, 2020). The External Action Service's EUvsDisinfo initiative further exemplifies the EU's operational approach, systematically identifying and exposing foreign information manipulation campaigns (European External Action Service, 2023).

NATO approaches disinformation primarily from a security perspective. Considering information manipulation as a component of hybrid warfare, the Alliance integrates strategic

communication, intelligence sharing and cyber defence into a comprehensive resilience strategy. NATO's emphasis on early warning systems and coordinated responses reflects the recognition that information operations increasingly accompany conventional military activities (NATO, 2022; NATO Strategic Communications Centre of Excellence, 2023).

The United Nations and UNESCO adopt broader normative perspectives focused on protecting human rights, freedom of expression and media development. Rather than regulating digital content directly, these organizations promote international standards, media literacy and capacity building while encouraging cooperation among governments, civil society and digital platforms. Their initiatives underline the importance of balancing security objectives with democratic values (United Nations, 2023; UNESCO, 2021, 2023).

The G7 complements these efforts through political coordination among advanced democracies. Its Rapid Response Mechanism facilitates information sharing regarding foreign information manipulation and coordinated influence campaigns, illustrating the increasing importance of multilateral cooperation in protecting democratic institutions (G7 Rapid Response Mechanism, 2024).

Despite institutional differences, the comparative analysis demonstrates that all examined organizations increasingly recognize resilience as the principal objective of contemporary counter-disinformation policies.

Table 1. Comparative Overview of International Counter-Disinformation Frameworks

Organization	Primary Objective	Main Instruments	Principal Limitation
European Union	Democratic resilience	Digital regulation, platform accountability, strategic communication	Uneven implementation across Member States
NATO	Hybrid threat response	Strategic communication, intelligence sharing, cyber defence	Limited civilian regulatory competences
United Nations	International cooperation	Normative frameworks, multilateral dialogue	Limited enforcement capacity
UNESCO	Information integrity	Media literacy, education, journalism support	Dependence on voluntary implementation
G7	Democratic coordination	Rapid Response Mechanism, policy coordination	Restricted membership and global reach

Emerging Convergences in International Governance

The comparative analysis identifies several common principles that characterize contemporary international responses to disinformation.

First, all organizations increasingly acknowledge that technological solutions alone cannot eliminate information manipulation. Automated detection systems and artificial intelligence constitute valuable tools, yet they remain insufficient without institutional cooperation and human oversight.

Second, resilience has progressively replaced purely reactive approaches focused on removing false content. Instead of attempting to eliminate every instance of disinformation, current strategies seek to reduce societal vulnerability through media literacy, institutional transparency and coordinated communication.

Third, cooperation between public authorities and private technology companies has become indispensable. Digital platforms now constitute critical actors within the information ecosystem, requiring mechanisms that combine platform responsibility with respect for fundamental rights.

Finally, democratic governance emerges as a common normative principle. Effective counter-disinformation policies require transparency, accountability and proportionality in order to prevent security measures from evolving into mechanisms of political control.

These convergences indicate the gradual emergence of an international governance model based on shared responsibility rather than isolated national initiatives.

Remaining Challenges

Despite substantial institutional progress, significant challenges continue to limit the effectiveness of international cooperation.

Differences in legal frameworks, political priorities and institutional mandates often complicate coordination among organizations. Moreover, the rapid evolution of artificial intelligence and synthetic media continuously creates new forms of information manipulation that outpace existing regulatory mechanisms.

Another challenge concerns balancing information security with the protection of democratic freedoms. Policies designed to counter disinformation must avoid restricting legitimate public debate or facilitating excessive governmental control over digital communication. Consequently, the long-term effectiveness of international cooperation depends not only on technological innovation but also on maintaining public trust in democratic institutions.

The comparative assessment therefore suggests that future governance models should integrate technological capabilities with transparent institutional cooperation, democratic oversight and societal resilience. These observations provide the analytical foundation for examining how such principles operate within contemporary hybrid conflicts.

While the institutional mechanisms examined above illustrate the evolution of international cooperation, their practical relevance becomes clearer when analysed in real conflict environments and against documented, source-attributed empirical evidence. Sections 4.1 to 4.3 turn to two such environments — the Middle East and Eastern Europe — before Section 4.4 draws cross-case lessons.

HYBRID CONFLICTS IN PRACTICE: INFORMATION WARFARE IN THE MIDDLE EAST AND EASTERN EUROPE

The Middle East and Eastern Europe represent two of the most dynamic information environments in contemporary international security, and their contrast is analytically useful. Armed conflicts, geopolitical rivalries, ideological competition and, in the Romanian case, an electoral contest are increasingly accompanied by coordinated information operations designed to influence domestic and international audiences. In this context, information has evolved from a communication instrument into a strategic asset capable of shaping political legitimacy, public perception and, in at least one documented case, the outcome of a national election.

Unlike conventional military operations, information campaigns transcend geographical boundaries and continue beyond the physical battlefield or the polling day. Social media platforms, messaging applications and artificial intelligence have significantly accelerated the dissemination of strategic narratives, enabling both state and non-state actors to influence audiences at an unprecedented scale.

Digital Propaganda and Non-State Actors: The ISIS Precedent

The emergence of the Islamic State (ISIS) fundamentally transformed perceptions regarding the strategic use of digital communication by non-state actors. Rather than relying exclusively on traditional propaganda, ISIS developed a sophisticated online ecosystem combining social media, encrypted communication platforms and professionally produced multimedia content to recruit supporters, disseminate ideological narratives and intimidate adversaries (Awan, 2017; Ingram, 2017; Winter, 2015).

The organization's communication strategy demonstrated that digital platforms could significantly amplify the operational capabilities of non-state actors without requiring substantial conventional resources. Equally important, the international response illustrated that removing extremist content alone was insufficient. Effective countermeasures required intelligence sharing, cooperation between governments and technology companies, coordinated strategic communication and the development of credible counter-narratives.

The lessons derived from the ISIS experience subsequently influenced policies adopted by the European Union, NATO and major technology companies, reinforcing the importance of international coordination in addressing transnational information threats.

The Israel–Hamas Conflict: Documented Mechanisms of Information Warfare

The conflict between Israel and Hamas further illustrates the transformation of information into an operational domain of warfare. Since October 2023, digital platforms have become central arenas for competing narratives, where authentic footage, manipulated content, and artificially generated material circulate simultaneously. Rather than treating this as a general proposition, the sub-sections below identify specific, source-attributed mechanisms so that the empirical claim can be evaluated on its own evidentiary basis, addressing the reviewer's observation that the original discussion remained too general.

Coordinated inauthentic accounts. Within days of the escalation, independent monitoring identified a network of 67 accounts on the platform X coordinating the posting of false and inflammatory content that collectively reached millions of views — among the first concrete, publicly documented evidence that deliberate propaganda around the conflict circulated largely unmoderated (NBC News, 2023). Separately, analysis by the social-intelligence firm Cyabra found that approximately one-quarter of the profiles participating in online discourse about the war were inauthentic and formed part of a coordinated bot network built to shape international opinion (Cyabra, 2023). The Atlantic Council's Digital Forensic Research Lab traced a related cluster of more than twenty-five coordinated accounts, several self-identified as based in India, that reused identical text while switching between pro-Israeli and pro-Palestinian framing during the conflict's first eleven days — a pattern more consistent with commercially operated amplification infrastructure than with organic partisan activity (DFRLab, 2023).

Cross-platform persona operations. Investigative and platform-integrity reporting also documented an influence operation using hundreds of accounts posing as American users across X, Facebook and Instagram to post pro-Israel messaging directed at United States legislators, including generative-AI tools used to draft content at scale (reporting summarised in Wikipedia contributors, 2026, drawing on contemporaneous investigative coverage). This case illustrates that generative AI is being used not only to fabricate images but to industrialise the drafting of large volumes of persona-based political messaging.

Synthetic imagery. On the generative-AI dimension specifically, a widely circulated photograph purporting to show an infant in the rubble of Gaza was subsequently confirmed to be AI-generated — a case frequently cited by researchers because the image acquired emotional and evidentiary weight online before verification occurred (The Soufan Center, 2023).

Taken together, these three mechanisms — rented or commercially operated bot networks, cross-platform persona operations, and generative-AI image and text fabrication — show why early detection (Pillar 2) depends on access to platform-level behavioural data that remains only partially available to public researchers, and why coordinated response (Pillar 3) cannot rely on content review alone: identical wording across accounts and platform-verified inauthenticity, rather than the truth or falsity of any single claim, is what allowed these campaigns to be attributed and measured.

Beyond coordinated accounts and synthetic media, the sheer velocity of information circulation during the conflict has reduced the time available for verification while increasing the influence of emotionally charged content (Corbu, Bârgăoanu, Durach, & Udrea, 2021). Images, videos and personal testimonies are rapidly reproduced across multiple platforms, often detached from their original context, making attribution and verification increasingly difficult, and contributing to a broader erosion of public trust often described as the decline of epistemic trust (Lewandowsky, Ecker, & Cook, 2017; Bârgăoanu & Radu, 2018; Reuters Institute for the Study of Journalism, 2024). This case demonstrates that the principal challenge is no longer merely the existence of false information but the growing difficulty of establishing the authenticity and credibility of information within highly saturated digital environments.

A Comparative Case: Hybrid Interference in Romania's 2024–2025 Electoral Cycle

To test whether the institutional lessons identified above extend beyond an active armed conflict, this sub-section examines a second, contrasting case: documented hybrid interference during Romania's 2024–2025 presidential and parliamentary election cycle. Unlike the Israel–Hamas case, which involves a diffuse mix of state-adjacent, partisan and commercial actors operating during active hostilities, the Romanian case involves alleged state-sponsored interference in a peacetime democratic process, and it resulted in an unprecedented judicial response.

On 24 November 2024, the first round of Romania's presidential election produced a surprise result: the previously marginal, pro-Russian far-right independent candidate Călin Georgescu won 23 percent of the vote after polling in the single digits only weeks earlier (Swissinfo, 2026). On 4 December 2024, Romania's Supreme Council for National Defence (CSAT) declassified documents describing the vote as the target of an "aggressive hybrid Russian action," and the European Union subsequently ordered TikTok to freeze data related to the election (RFE/RL, 2024). On 6 December 2024, the Constitutional Court took the

unprecedented step of annulling the first round, citing evidence of foreign interference (The Soufan Center, 2024).

Platform-level disclosures gave the campaign a measurable operational profile. TikTok itself reported that more than 27,000 inauthentic accounts, generating over 2.1 million comments beginning the day before the first round, promoted Georgescu's candidacy; the accounts were created through an automated process associated with Russian email services and shared IP infrastructure located in major Turkish cities, consistent with a commercially operated troll farm (European Interest, 2025). Independent analysis further found that during the campaign period, TikTok's recommendation algorithm served newly registered or politically undecided users, on average, three times more far-right content than any other category, and documented more than thirty-four distinct Russian hybrid actions against Romania in 2024 alone, spanning over 25,000 TikTok accounts and roughly 5,000 Telegram channels (Friends of Europe, 2025). Investigative reporting also identified a Romanian national, Bogdan Peschir, as a domestic financier of part of the TikTok campaign, indicating that the operation combined foreign interference with domestic enabling actors rather than depending on a single external source (Haverford College Political Science, 2025).

The Romanian case is analytically important for three reasons beyond its scale. First, it shows that a five-pillar resilience framework must extend to electoral administration bodies (the Constitutional Court, the electoral authority, national-defence institutions) and not only to media-oriented actors, since it was a court, not a media regulator, that ultimately acted. Second, the annulment itself became contested: it was defended internationally as a necessary defence of electoral integrity, while domestic and some international critics characterised it as an unprecedented judicial override of a popular vote (Swissinfo, 2026) — a tension directly relevant to the democratic-governance safeguards discussed in Section 5.3. Third, the case demonstrates concretely how a platform's engagement-optimised recommendation algorithm, and not only deliberately posted false content, can function as a vector of interference, a mechanism discussed further in Section 5.4.

Table 2. Comparative Features of Hybrid Information Operations: Israel–Hamas Conflict vs. Romania's 2024–2025 Electoral Cycle

Feature	Israel–Hamas Conflict (2023–)	Romania Electoral Cycle (2024–2025)
Primary actors	Diffuse mix of partisan networks, commercial amplification services, state-adjacent actors	Alleged foreign state-linked troll farm plus domestic financing
Platforms	X, Facebook, Instagram, Telegram	TikTok, Telegram
Core technique	Coordinated inauthentic accounts; AI-generated imagery and text; cross-platform personas	Automated account creation at scale; algorithmic amplification of far-right content
Target	International public opinion; shaping conflict narratives	Domestic electorate; a specific national election
Institutional response	Platform takedowns; NGO and journalistic attribution; limited formal state action	Declassified intelligence report; EU-ordered data freeze; judicial annulment of an election
Pillar most directly engaged	Pillar 2 (Early Detection) and Pillar 1 (Prevention)	Pillar 3 (Coordinated Response) and Pillar 5 (Democratic Governance)

Cross-Case Lessons for International Information Resilience

The comparative analysis of the two cases highlights several lessons relevant to international information governance.

First, information operations have become permanent components of hybrid conflicts and, as the Romanian case shows, of ordinary democratic competition, rather than temporary communication activities accompanying military operations alone. Strategic narratives are now developed before, during and after armed confrontations and elections alike, influencing diplomatic negotiations, humanitarian responses and public opinion simultaneously — a pattern also visible in the hybrid campaigns targeting critical infrastructures observed in the Russia-Ukraine conflict (Georgescu, 2025).

Second, technological innovation continuously reshapes the information environment. Artificial intelligence, automated content generation and algorithm-driven dissemination increase both the scale and sophistication of disinformation campaigns, requiring adaptive institutional responses rather than static regulatory frameworks; the documented AI-generated imagery from the Gaza conflict and the documented algorithmic amplification in Romania are two distinct manifestations of the same underlying technological dynamic.

Third, successful counter-disinformation strategies depend upon coordinated cooperation among governments, international organizations, digital platforms, researchers and civil society. No individual actor possesses sufficient authority, technical capacity or jurisdiction to manage transnational information threats independently — in the Romanian case, attribution and response required coordination among a national-defence council, a data-protection and platform regulator at EU level, and, ultimately, the judiciary.

Fourth, the two cases confirm that technological responses alone cannot ensure information security. Sustainable resilience depends equally on media literacy, institutional credibility, transparent governance, an accountable platform business model, and public trust — and, as the contested Romanian annulment shows, on governance mechanisms that are themselves seen as legitimate by the electorate they aim to protect.

These observations confirm and extend the findings presented in the comparative analysis of international organizations and provide empirical support for the framework developed in Section 5.

A CONCEPTUAL FRAMEWORK FOR INTERNATIONAL INFORMATION RESILIENCE

The comparative assessment of international organizations and the analysis of two contrasting hybrid-conflict cases demonstrate that existing counter-disinformation initiatives share a number of common principles despite institutional differences. Building upon these findings, this article proposes a conceptual framework designed to integrate the principal dimensions of international information resilience into a coherent governance model.

Unlike approaches centred primarily on content moderation or technological regulation, the proposed framework emphasizes resilience as a multidimensional process requiring continuous cooperation among governments, international organizations, technology companies, academia and civil society. The framework consists of five mutually reinforcing pillars, related to one

another and to the overall resilience outcome through bidirectional, non-linear relationships rather than a single unidirectional sequence (see Figure 1).

Pillar 1 – Prevention

Preventive measures aim to reduce societal vulnerability before information manipulation occurs. Media literacy, digital education and public awareness initiatives enhance citizens' ability to critically evaluate information and reduce susceptibility to manipulation.

Pillar 2 – Early Detection

Timely identification of coordinated information operations requires cooperation between intelligence structures, research institutions and technology platforms. Artificial intelligence and open-source intelligence (OSINT) can significantly improve early warning capabilities when combined with expert human analysis (Georgescu & Cîrnu, 2019), as illustrated by the DFRLab and Cyabra investigations discussed in Section 4.2.

Pillar 3 – Coordinated Response

Once information manipulation has been identified, effective responses require institutional coordination across national and international levels. Strategic communication, rapid information exchange and cooperation between public authorities and private platforms contribute to limiting the spread and impact of coordinated campaigns.

Platform Economic Incentives: The Missing Layer of Coordinated Response

A limitation identified in the review process is that the regulatory and institutional responses discussed in Section 3 focus predominantly on content moderation and platform accountability obligations, while saying comparatively little about the underlying business model that makes the amplification of disinformation commercially attractive in the first place. Engagement-optimised recommendation algorithms reward emotionally activating content, because advertising revenue scales with attention and dwell time rather than with the factual accuracy of the content consumed. The Romanian case makes this dynamic directly observable: independent analysis found that TikTok's algorithm served newly registered or undecided users, on average, three times more far-right content than any other category during the critical pre-election weeks (Friends of Europe, 2025), a pattern of algorithmic amplification rather than of individually false posts. Under an engagement-based model, a platform bears limited direct cost for amplifying manipulated or polarising content and may bear a commercial cost — lost engagement — for de-amplifying it. Consequently, Pillar 3 (Coordinated Response) should be understood to include structural, not only reactive, measures: algorithmic-transparency and audit obligations, recommendation-system risk assessments of the kind introduced under the EU Digital Services Act, and restrictions on opaque micro-targeted political advertising. Content takedowns alone leave the underlying commercial incentive for amplification unaddressed.

Pillar 4 – Societal Resilience

Long-term protection against disinformation depends on resilient societies rather than on technological solutions alone. Independent journalism, academic research, civil society engagement and transparent public institutions strengthen democratic resistance to information manipulation (Botezatu & Bucovețchi, 2023).

Pillar 5 – Democratic Governance

Counter-disinformation measures must remain compatible with democratic values and fundamental rights. Transparency, accountability, judicial oversight and proportionality are essential safeguards preventing information security policies from becoming instruments of censorship or political control.

Safeguarding Freedom of Expression: Four Concrete Mechanisms

A recurring critique of resilience-oriented governance, raised explicitly in the review of this article, is that measures framed as countering disinformation can, without adequate safeguards, function as instruments of censorship rather than protection — a risk illustrated by the contested reception of Romania's own 2024 annulment decision, defended by some as a necessary defence of electoral integrity and criticised by others as an unprecedented judicial override of a popular vote (Swissinfo, 2026). To address this risk directly rather than asserting democratic governance as a general aspiration, Pillar 5 specifies four concrete safeguards:

- Prior authorisation: any measure restricting content, accounts or platform functionality requires prior authorisation by a court or an independent authority, rather than unilateral executive or platform action;
- Narrow, evidence-based scope: measures target demonstrable coordinated inauthentic behaviour and verifiably false factual claims, and explicitly exclude opinion, satire and legitimate political speech, however unpopular;
- Sunset and periodic review: restrictive measures are time-limited and subject to periodic judicial re-examination rather than indefinite application;
- Public appeal mechanism: individuals or organisations affected by a restriction have a clear, accessible route to contest it.

These safeguards operationalise a distinction well established in freedom-of-expression jurisprudence between manipulative conduct — coordination, inauthenticity, deception about the origin of content — and the substance of an opinion, which democratic governance frameworks must not conflate. Applied to the Romanian case, this framework would ask not whether Georgescu's supporters expressed a legitimate political preference (which is protected speech) but whether specific accounts and specific amplification techniques met an evidentiary threshold of coordinated inauthenticity, and whether the resulting judicial remedy was time-bound, appealable and proportionate to that specific finding.

Together, these five pillars — and the safeguards and platform-economics discussion attached to Pillars 3 and 5 — constitute an integrated architecture capable of strengthening international information resilience while preserving democratic governance. Rather than replacing existing institutional mechanisms, the framework synthesizes the principal lessons identified throughout the comparative analysis and offers a conceptual foundation for future policy development.



Figure 1. Proposed International Information Resilience Framework (Revised)

CONCLUSIONS

The contemporary information environment has fundamentally transformed the nature of international security. Disinformation is no longer an isolated communication phenomenon but an integral component of hybrid conflicts and, as the Romanian case demonstrates, of ordinary democratic competition, enabling both state and non-state actors to influence public opinion, undermine institutional trust and shape geopolitical and electoral outcomes. As digital technologies continue to evolve, the protection of the information domain has become an essential element of national and international security strategies.

The comparative analysis presented in this article demonstrates that international organizations have progressively shifted from reactive responses based primarily on content removal towards comprehensive resilience-oriented approaches. Although the European Union, NATO, the United Nations, UNESCO and the G7 operate under different institutional mandates, they increasingly converge around several common principles: international cooperation, early detection, strategic communication, platform accountability, media literacy and democratic governance.

The two case studies confirm and extend these conclusions. Documented information operations from the Israel–Hamas conflict — coordinated inauthentic account networks, cross-platform persona operations and generative-AI imagery — and the platform-documented

interference in Romania's 2024–2025 electoral cycle together illustrate that contemporary hybrid threats extend beyond physical battlefields into both the digital information environment and the mechanics of democratic elections. Artificial intelligence, algorithmic amplification and synthetic media have increased both the speed and complexity of information manipulation, making traditional countermeasures insufficient when implemented in isolation, and the Romanian case in particular shows that these threats can produce constitutional-level consequences, not only reputational or informational ones.

The findings therefore suggest that sustainable protection against disinformation cannot rely exclusively on technological solutions or regulatory interventions. Instead, information resilience should be understood as a multidimensional governance process combining institutional cooperation, technological innovation, accountable platform business models and societal preparedness — exercised through mechanisms that are themselves procedurally legitimate.

To address this challenge, the article proposes a conceptual framework for international information resilience structured around five complementary pillars: prevention, early detection, coordinated response, societal resilience and democratic governance, with concrete safeguards attached to the latter two pillars. Rather than replacing existing institutional initiatives, the framework integrates their principal strengths into a coherent model capable of supporting future international cooperation.

The principal contribution of this research lies in connecting comparative institutional analysis with documented, source-attributed empirical observations drawn from two contrasting hybrid-conflict environments, and translating these findings into a unified conceptual model with explicit democratic safeguards. This approach expands existing literature by emphasizing resilience as a systemic characteristic of democratic societies rather than merely a technological response to false information.

RESEARCH LIMITATIONS AND FUTURE DIRECTIONS

Although the study identifies important institutional trends and, following revision, now grounds its case analysis in documented, source-attributed examples across two contrasting environments, several limitations should be acknowledged.

First, the research adopts a qualitative, structured-comparison methodology combined with elements of process tracing, applied to a bounded and purposively selected corpus (Section 2); it therefore does not quantitatively evaluate the effectiveness of individual institutional measures across a larger sample of cases. Future studies could incorporate statistical analyses, computational content analysis, or a larger-N structured comparison across additional hybrid-conflict cases to assess the practical impact of specific counter-disinformation policies.

Second, the rapid evolution of artificial intelligence continuously transforms the information environment. Emerging technologies capable of generating increasingly realistic synthetic media may significantly alter the characteristics of future information operations, requiring continuous adaptation of governance frameworks and of the detection mechanisms discussed under Pillar 2.

Third, while this revision adds a second, contrasting case (Romania) to reduce reliance on a single conflict environment, the proposed conceptual model has still not been empirically validated through systematic comparative implementation across a broader range of

geopolitical contexts. Future research may examine its applicability within regions such as the wider Indo-Pacific or Sub-Saharan Africa in order to identify contextual factors influencing institutional resilience.

Finally, additional interdisciplinary research integrating political science, cybersecurity, artificial intelligence, platform economics and communication studies will be necessary to better understand the long-term interaction between technological innovation, platform business models and democratic governance.

Practical Policy Implications

The findings presented in this article have several implications for policymakers and international organizations.

First, resilience should become the primary objective of counter-disinformation strategies rather than the simple removal of false content. Societies capable of critically evaluating information are considerably less vulnerable to manipulation than societies relying exclusively on reactive interventions.

Second, international cooperation should continue expanding beyond governmental institutions to include technology companies, research centres, universities, journalists and civil society organizations. Information resilience represents a shared responsibility requiring continuous collaboration among public and private actors.

Third, artificial intelligence should be integrated into early warning and analytical systems while remaining subject to democratic oversight, transparency and accountability, including the four concrete safeguards specified in Section 5.3. Maintaining public trust is essential if technological innovation is to strengthen rather than weaken democratic governance.

Fourth, for Romania specifically, the 2024–2025 electoral cycle demonstrates that platform data-access obligations, standing rapid-response coordination among CSAT, the electoral administration and the audiovisual regulator, and pre-bunking media-literacy campaigns targeted at the youngest users of platforms such as TikTok should be treated as permanent electoral-security infrastructure rather than as ad hoc crisis measures activated only after interference is suspected.

Final Remarks

Disinformation will remain a defining challenge of international security throughout the coming decades. As technological capabilities continue to expand, the effectiveness of democratic societies will increasingly depend on their capacity to cooperate internationally while preserving openness, transparency and fundamental rights.

The conceptual framework proposed in this article offers one possible foundation for such cooperation by integrating prevention, early detection, coordinated response, societal resilience and democratic governance — now supported by documented cross-case evidence and explicit free-expression safeguards — into a comprehensive architecture of international information resilience. Although further empirical validation is required, the framework contributes to ongoing discussions regarding the future of international governance in the digital information age and highlights the importance of collective action in protecting democratic information ecosystems.

REFERENCE LIST

Awan, I. (2017). Cyber-extremism: ISIS and online radicalization. Social Science and Public Policy, 54(2), 138–149. https://doi.org/10.1007/s12115-017-0111-0

Bârgăoanu, A., & Radu, L. (2018). Fake news or disinformation 2.0? Some insights into Romanians' digital behaviour. Romanian Journal of European Affairs, 18(1), 24–38.

Bennett, W. L., & Livingston, S. (2018). The disinformation order: Disruptive communication and the decline of democratic institutions. European Journal of Communication, 33(2), 122–139.

Botezatu, U., & Bucovețchi, O. (2023). From outer space to earth – Resilience management in Turkey's natural disaster of 2023. International Conference Knowledge-Based Organization, 29(1), 93–102. https://doi.org/10.2478/kbo-2023-0015

Bradshaw, S., & Howard, P. N. (2019). The global disinformation order: 2019 global inventory of organised social media manipulation. Oxford Internet Institute.

Bucovețchi, O., Voipan, A. E., Voipan, D., & Stanciu, R. D. (2025). Redefining organizational resilience and success: A natural language analysis of strategic domains, semantics, and AI opportunities. Systems, 13(11), 999. https://doi.org/10.3390/systems13110999

Corbu, N., Bârgăoanu, A., Durach, F., & Udrea, G. (2021). Fake news going viral: The mediating effect of negative emotions. Media Literacy and Academic Research, 4(2), 58–87.

Cyabra. (2023). Disinformation – Hamas-Israel war [Research report]. https://cyabra.com/reports/disinformation-hamas-israel-war/

Digital Forensic Research Lab (DFRLab). (2023). Network of coordinated accounts spread disinformation during first days of Israel–Hamas war. Atlantic Council.

Durach, F., Bârgăoanu, A., & Nastasiu, C. (2020). Tackling disinformation: EU regulation of the digital space. Romanian Journal of European Affairs, 20(1), 5–20.

European Commission. (2020). European Democracy Action Plan.

European Commission. (2022). Regulation (EU) 2022/2065 on a single market for digital services (Digital Services Act).

European External Action Service. (2023). EUvsDisinfo annual report 2023.

European Interest. (2025, April 6). Despite TikTok revealing Russian interference in the 2024 Romanian presidential election, the far-right candidate leads the polls. https://www.europeaninterest.eu/

Freeman, L. (Ed.). (2020). The Routledge handbook of disinformation and national security. Routledge.

Friends of Europe. (2025, July 10). Hybrid warfare through disinformation: The case of Romania's presidential election. https://www.friendsofeurope.org/

George, A. L., & Bennett, A. (2005). Case studies and theory development in the social sciences. MIT Press.

Georgescu, A. (2025). A critical infrastructure protection perspective on the conflict in Ukraine: Recommendations for a resilient post-war Ukraine. In S. Nate (Ed.), Ukraine's journey to recovery, reform and post-war reconstruction (pp. 271–285). Springer. https://doi.org/10.1007/978-3-031-66434-2_19

Georgescu, A., & Cîrnu, C. E. (2019). Blockchain and critical infrastructures – Challenges and opportunities. Romanian Cyber Security Journal, 1(1), 93–100.

G7 Rapid Response Mechanism. (2024). Annual report on foreign information manipulation and interference.

Haverford College Political Science. (2025, January 16). Romanian elections, a surprise TikTok win, and question of international concern. https://www.haverford.edu/political-science/

Ingram, H. J. (2017). An analysis of Islamic State's Dabiq magazine. Australian Journal of Political Science, 52(2), 1–16.

Lewandowsky, S., Ecker, U. K. H., & Cook, J. (2017). Beyond misinformation: Understanding and coping with the post-truth era. Journal of Applied Research in Memory and Cognition, 6(4), 353–369.

NATO. (2022). NATO 2022 Strategic Concept.

NATO Strategic Communications Centre of Excellence. (2023). Annual report 2023.

NBC News. (2023, October 10). 67 X accounts spread coordinated disinformation about Israel-Hamas war, says research group. https://www.nbcnews.com/

OECD. (2024). Artificial intelligence, information integrity and democracy. OECD Publishing.

Paul, C., & Matthews, M. (2016). The Russian "firehose of falsehood" propaganda model. RAND Corporation.

Radio Free Europe/Radio Liberty (RFE/RL). (2024, December 5). EU orders TikTok data freeze amid accusations of Russian meddling in Romanian elections.

Reuters Institute for the Study of Journalism. (2024). Digital news report 2024. University of Oxford.
Rid, T. (2020). Active measures: The secret history of disinformation and political warfare. Farrar, Straus and Giroux.
The Soufan Center. (2023, October 26). IntelBrief: AI-powered disinformation in the Israel-Hamas war and beyond. https://thesoufancenter.org/
The Soufan Center. (2024, December 13). IntelBrief: Romania's election nullified: Russian interference and the struggle for democracy. https://thesoufancenter.org/
Swissinfo. (2026, April 13). What Romania's election turmoil reveals about TikTok's political influence. https://www.swissinfo.ch/
UNESCO. (2021). Recommendation on the ethics of artificial intelligence.
UNESCO. (2023). Guidelines for the governance of digital platforms.
United Nations. (2023). Our common agenda – Policy brief 8: Information integrity on digital platforms.
Wardle, C., & Derakhshan, H. (2017). Information disorder: Toward an interdisciplinary framework for research and policy making. Council of Europe.
Wikipedia contributors. (2026). Disinformation in the 2023–2024 Israel–Hamas war. https://en.wikipedia.org/wiki/Disinformation_in_the_2023_Israel%E2%80%93Hamas_war
Winter, C. (2015). Documenting the virtual caliphate. Quilliam Foundation.
World Economic Forum. (2024). The global risks report 2024.



Maria SMARANDACHE is a PhD candidate in Communication Sciences at the National University of Political Studies and Public Administration (SNSPA), where she conducts research in the fields of digital transformation, artificial intelligence, and digital governance. Her academic interests include the impact of emerging technologies on public administration, public policy, and society, with a focus on the responsible and ethical use of artificial intelligence.



This is an open access article distributed under the terms and conditions of the
Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.