Doctrine for Neutralizing the Russian Federation's Cognitive Aggression and Establishing Echeloned Safeguards for Institutional Resilience of Public Administration

Volodymyr Hurkovskyi, Serhii Ilnytskyi, Rena Marutian, Oleksandr Hrynchuk, Jana Antalíková

https://doi.org/10.33542/VSS2026-1-2

Abstract

The contemporary global security environment is characterized by a transition to a state of permanent information confrontation, where the cognitive domain is formally defined as the fifth theater of operations. The relevance of this study is driven by the rapid intensification of systemic hybrid aggression, particularly the qualitative leap in the utilization of artificial intelligence technologies for the mass production of synthetic manipulative content. Such threats target fundamental democratic institutions and erode public trust in objective facts.

The aim of this article is to formulate and theoretically substantiate a comprehensive cognitive security doctrine focused on a strategic shift from passive, reactive defense to a model of active deterrence of cognitive threats. The theoretical and methodological approach of the study is based on an interdisciplinary integration of military strategy, strategic communication tools, political psychology, and the norms of international law. The work applies the Kill Chain model for the deconstruction of hostile influence operations, as well as an analytical framework for the behavior-centric analysis of foreign information manipulation.

The primary results of the research include a complex classification of the systemic vulnerabilities of liberal societies across three levels: systemic-legal, socio-psychological, and technological. The study substantiates the strategic advantage of proactive methods, such as pre-bunking, over traditional methods of debunking disinformation. A key theoretical contribution is the development of safeguards against the drift of democracy toward tyranny, which involves strengthening institutional immunity and protecting the sphere of privacy instead of implementing state censorship tools.

The practical significance of these results lies in the creation of a coherent operational strategy that allows for the identification and neutralization of cognitive attacks at the early stages of their preparation. The proposed approaches ensure a significant increase in the cost of aggression for the adversary and create conditions for the preservation of national cognitive sovereignty while simultaneously protecting fundamental liberal values and human rights.

Keywords: *Disinformation, FIMI (Foreign Information Manipulation and Interference), Cognitive Domain, Engineered Polarization, Cognitive Security.*

Introduction

Since 2022, Russian information and psychological operations (PSYOPS) in Europe have undergone a significant adaptation. Instead of the mass dissemination of primitive propaganda narratives, Moscow increasingly employs fragmented cognitive campaigns aimed at undermining trust in state institutions, electoral procedures, and collective security systems. In their 2025–2026 reports, analysts from the NATO Strategic Communications Centre of Excellence and the European External Action Service (EEAS AI, 2026) noted a shift in focus from direct pro-Russian agitation to the stimulation of internal polarization within European societies. In practical terms, this manifests through the synchronized use of Telegram networks, X (formerly Twitter) accounts, TikTok content, and pseudo-alternative media to promote themes such as "Ukraine fatigue," "ineffectiveness of sanctions," or the "crisis of democracy in the EU" (NATO, 2025. 2026; EEAS AI, 2026]. Russian intelligence services increasingly integrate elements of information influence with cyber operations and influence operations via local political actors, creating an effect of organic internal distrust rather than external interference. Consequently, the modern model of Russian IPSO in Europe functions not as classical propaganda, but as a technology for the erosion of social cohesion. In Central Europe energy dependence on Russian gas has also become an important informational narrative. Concerns about energy prices, supply security, and sanctions can be exploited to fuel distrust in state institutions, EU policies, and support for Ukraine (Sinicyn, Denciová and Eštok, 2026).

In the Baltic States, this strategy exhibits a different intensity and a significantly higher level of targeting. For the Russian Federation, this region remains a testing ground for cognitive operations due to its geographical proximity to Russia, the historical context of Soviet occupation, and the presence of Russian-speaking communities. Reports from the Estonian Foreign Intelligence Service (EFIS) and the European Council on Foreign Relations (ECFR) indicate that between 2024 and 2026, the Kremlin systematically exploited themes of historical memory, language policy, and social discrimination to foster a sense of "alienation" among specific population groups in Estonia and Latvia (EFIS, 2026; ECFR, 2025). In Estonia, one of Europe’s most digitized nations, Russian influence networks have adapted to the high level of media literacy among the population. Instead of crude fakes, they disseminate blended information constructs where real facts are combined with manipulative interpretations regarding cybersecurity, electronic voting, or the presence of NATO forces in the Baltic region. Concurrently, there is an observed intensification of agent structures and affiliated media platforms attempting to create the impression that Western security integration itself allegedly increases the risk of military escalation in the region.

At the pan-European level, Russian cognitive operations are increasingly less focused on convincing the audience of the "correctness" of the Kremlin's position. Their primary function is to destroy the ability of societies to form a shared vision of reality. Analysts from the Center for Strategic and International Studies (CSIS), the RAND Corporation, and the Institute for the Study of War (ISW) note that in 2025–2026, the Russian Federation was particularly active in using decentralized

networks of influencers, anonymous channels, and content with low levels of overt politicization to penetrate youth environments (CSIS, 2025; RAND, 2026; ISW, 2026). Compared to Central or Southern Europe, the Baltic States demonstrate higher institutional resilience to direct propaganda but remains vulnerable to long-term trust attrition campaigns. Therefore, the key threat to Estonia and the entire Baltic region is not an individual fake or information attack, but the gradual formation of an environment of strategic uncertainty, in which society begins to doubt the ability of democratic institutions to ensure security, stability, and control over the information space. These cognitive operations also exploit the political limits of the European Union's soft power. By amplifying Euroscepticism, institutional distrust, and perceptions of European weakness, hostile narratives can undermine the EU's capacity to sustain support for integration and collective action in the face of external challenges (Dubóczi and Dvorský, 2025).

Relevant data from recent years indicates an unprecedented intensification of cognitive aggression. Ukraine remains the primary testing ground for Russian FIMI (Foreign Information Manipulation and Interference) technologies, accounting for nearly half of all documented incidents worldwide. In 2025, a qualitative leap in the threat was recorded: the use of artificial intelligence to create synthetic content and the mass dissemination of disinformation increased by 2.5 times compared to the previous year. These operations have become more covert, combining digital manipulation with physical acts of sabotage and intimidation.

A critical review of contemporary literature and doctrinal documents, specifically EEAS reports (2023–2026) and RAND Corporation studies, points to a paradigm shift in defense (EEAS, 2023-2026). While previous approaches focused on the "shield"–reactive debunking and resilience building–newer concepts, such as the "FIMI Deterrence Playbook," emphasize the necessity of the "sword": active deterrence by increasing the cost of aggression for the adversary (EEAS, 2026). The problem of "Truth Decay," researched by RAND experts, highlights that cognitive biases and polarization are systemic vulnerabilities that the aggressor uses as force multipliers (Kavanagh, Rich, 2018; Paul, Matthews, 2016; Williams, McCulloch, 2023). At the same time, Russian "active measures" have evolved from Soviet-era disinformation methods into complex ecosystems that include networks of proxy actors, hacktivists, and diplomatic cover (FireEye, 2017).

The novelty and significance of this study lie in the development of an operational strategy that moves from threat diagnosis to active neutralization through echeloned safeguards. The unresolved issue of the "paradox of liberty" remains pertinent: how to ensure a country's cognitive security without drifting toward tyranny and censorship. Protecting the democratic order requires a balance between state control and the preservation of fundamental freedoms (EEAS, 2024; Fukuyama, 2023).

The aim of this article is to formulate a cognitive security doctrine based on three objectives: 1) Classification of systemic vulnerabilities in a liberal society; 2) Implementation of a "Kill Chain" model to disrupt cognitive attack cycles; 3) Establishment of institutional safeguards against

excessive restriction of citizens' rights in the process of countering IPSOs (EEAS, 2026). The strategy views the cognitive domain as a space where institutional resilience and active deterrence are the sole means of preserving national sovereignty.

1 Literature Review

The modern architecture of cognitive security requires the deep integration of multidisciplinary research to neutralize aggression in the sixth theater of war, where information space is recognized as a full-scale battlefield. Andrii Novik argues that in the large-scale, all-encompassing conflict in which Ukraine finds itself, "success is guaranteed exclusively by an active position." He emphasizes the need for a strategic transition to the "preemptive application of active measures," which involves constant monitoring and forestalling the Kremlin's disinformation efforts. This correlates with Russian doctrinal assertions regarding "information confrontation," where information influence has become a permanent front across a state's entire territory, and "non-military means have exceeded traditional military force in their effectiveness" (Novik, 2023).

Novik highlights that the Russian propaganda machine is an integrated ecosystem–a complex, multi-level structure capable of self-regulation, rapid recovery, and growth. Using disinformation as a weapon allows Russia to wage a unique form of war "against practically the entire civilized world," not limited to zones of direct kinetic engagement. At the core of the aggressor's strategy lies the conviction that "a nuclear power cannot be defeated in open confrontation, but it can be destroyed from within." In this context, "intentional agents" of influence, such as affiliated religious structures, political forces, and NGOs, alongside the systemic use of Soviet propaganda narratives adapted to modern conditions, pose a particular danger (Novik, 2023).

Thomas Rid has explored the evolution of these methods in detail, defining them as "professional, organized, aggressive lying," the strategic goal of which is to weaken the very foundation of liberal democracy–trust in objective facts. In his works, he argues that modern political warfare operations are effectively a computerized assault on violence itself through three classic types of activity: sabotage, espionage, and subversion. Rid expands the understanding of the cognitive threat with the following propositions: subversion, unlike classical espionage, is becoming increasingly independent of weaponry, as its direct "targets are human minds, not machines" (Rid, 2013).

The researcher identifies a technological paradox where digital technologies have significantly lowered the entry barrier for conducting subversive activities while simultaneously "raising the threshold for achieving lasting success" due to the difficulty of sustaining audience attention. Modern active measures create a state of a "hall of mirrors" or "engineered polarization," where the aggressor cynically exploits the openness, liberal values, and trust of a liberal society as entry vectors for its destruction. Furthermore, the attribution problem–the difficulty of identifying the

true origin of an attack–is a fundamental change in digital conflict that favors the attacker, allowing them to operate with impunity within the "gray zone" (Rid, 2013).

Thus, while Novik calls for a transition from reactive defense to "active counteraction" of a preventive nature to avoid the strategic dead end of "passive defense" (Novik, 2023), Rid cautions that in the process of combating disinformation, a democratic state must be careful: "the cure must not be worse than the disease" (Rid, 2020), so as not to destroy the values of the free exchange of ideas and freedom of speech. The synthesis of these two approaches forms our operational strategy, where technological preemption through Kill Chain analysis and OSINT monitoring is combined with the protection of institutional trust, the strengthening of the "constitution of knowledge," and the preservation of social bonds as the primary safeguards against cognitive occupation (Rid, 2013).

Within the development of an operational doctrine for cognitive security, we classify the fundamental vulnerability of the liberal order as a strategic design of the aggressor based on the cynical use of a free society's fundamental rights against itself. Z. Gershberg and S. Illing define this state as the "paradox of democracy": a situation where the openness of the media space and freedom of speech become entry vectors for "perilous persuasion" that threatens the very survival of the system (Ophir, 2023). Researchers argue that the digital revolution has made the world "more democratic but less liberal," as free platforms, stripped of editorial control, have become the ideal environment for the spread of conspiracy theories and disinformation (Ophir, 2023).

This concept is complemented by the works of K. Loewenstein, who in 1937, analyzing the collapse of European democracies under the pressure of fascism, introduced the term "militant democracy." Loewenstein warned of the "legalistic blindness" of democratic fundamentalism, which allows anti-democratic forces to use democratic mechanisms as a "Trojan Horse" to legally dismantle the state order from within. He argued that democratic institutions must not become instruments of their own destruction, and in a state of "internal war," legality may require temporary restrictions to preserve the very foundations of freedom (Loewenstein, 1937).

The strategic depth of this vulnerability is revealed by Thomas Rid, who points out that it is precisely the "community of ideas, a vivid and open dialogue" that makes democracy critically vulnerable to "active measures" (Rid, 2020). Rid argues that the goal of modern subversion is not to destroy machines but to target human minds, turning a society's openness into a tool of "engineered polarization" (Rid, 2013, 2020). Rid's ideas correlate with the findings of P. Pomerantsev, who describes a "war against reality," where the aggressor's goal is not merely to convince of a certain idea but to create a state of "cognitive chaos" where "nothing is true and everything is possible," rendering any rational democratic discussion impossible (Bourke, 2025).

Francis Fukuyama details the challenges to the liberal order in this context by analyzing the state of "democratic recession." He notes that liberal institutions (courts, independent media) become the primary targets of leaders with autocratic tendencies who use legal mandates to dismantle the system of checks and balances. Fukuyama also points to the development of

"vetocracy"–a situation where the procedural complexity and institutional inertia of democracies become a brake on effective responses to hybrid threats, creating a critical gap in reaction speed compared to the aggressor (Fukuyama, 2023).

The operational significance of these studies lies in the recognition that Russian cognitive aggression is not a random collection of fakes but a planned campaign of political warfare. As Novik notes, Russian strategy is based on the belief that "a nuclear power cannot be defeated in open confrontation, but it can be destroyed from within" (Novik, 2023). Consequently, the response must be based, according to J. Rauch, not on passive defense but on the active strengthening of the "constitution of knowledge" and institutional immunity, allowing democracy to remain "militant" in its self-defense (Fukuyama, 2023).

Furthermore, Fukuyama explored the institutional inertia of democratic societies, using the term "vetocracy" to describe political paralysis where complex constitutional procedures and numerous "veto points" make rapid crisis response impossible. In modern liberal democracies, these foundational institutions become the primary targets of aggressive leaders. This state of vetocracy leads to the "ossification" of institutions, turning procedural rigor into a strategic vulnerability under conditions of hybrid aggression (Fukuyama, 2023).

The complexity of the legal classification of cognitive attacks in the "gray zone" is confirmed by the *Tallinn Manual 2.0*. Experts led by M. Schmitt note that most information operations do not reach the threshold of an "armed attack" under Article 51 of the UN Charter, as international law traditionally links the concept of "attack" exclusively to consequences such as death, injury, or physical damage. This creates a legal lacuna, as cognitive operations cause "merely inconvenience, irritation, stress, or fear" and are thus not classified as threats justifying collective self-defense (Schmitt, 2017).

Moreover, developing an operational response is complicated by the "sovereignty dilemma." While most experts agree that remote physical damage violates sovereignty, there is no consensus on whether operations that merely manipulate data or perception without causing a loss of system functionality constitute such a violation (Schmitt, 2017).

Another factor is the threshold of intervention and coercion. Under international law, prohibited "intervention" requires an element of coercion. The aggressor successfully masks IPSOs as "diplomacy," "criticism," or "propaganda," which are not inherently illegal if they merely influence a state's will rather than forcing it into specific actions (Schmitt, 2017).

A favored tool of the Russian Federation is the exploitation of diplomatic immunity. Russia has fully integrated its diplomatic channels into its general ecosystem of foreign interference. Official accounts of diplomatic missions are used as networks for coordination and the dissemination of disinformation, benefiting from the legal protection afforded to diplomatic missions (Schmitt, 2017).

Thomas Rid emphasizes that the difficulty of attribution is a fundamental change in digital conflict, allowing an attacker to conduct subversion for years while remaining outside the bounds of legal responsibility (Rid, 2013).

At the core of the counter-strategy on the systemic-legal echelon must lie Loewenstein's concept of "militant democracy." He argued that democratic institutions must be "steeled and strengthened" when facing movements aimed at their destruction. If ordinary legislative channels are blocked by sabotage or procedural inertia, the state is obliged to use extraordinary powers, as "government is meant to govern," not to be a powerless witness to its own ruin (Loewenstein, 1937). It is necessary to adapt the legal framework to neutralize the "non-linear" actions of the enemy, which, according to General Gerasimov, exceed traditional military force in effectiveness today (Rid, 2020).

Within the cognitive security doctrine, we classify the state of the civilian population as the "democratic rear," whose resilience is undermined by a systemic crisis of perception. The analysis of this crisis is based on the concept of "Truth Decay," developed by RAND experts J. Kavanagh and M. Rich. They define this phenomenon as the "diminishing role of facts and analysis in public life." Truth Decay includes four interrelated trends: increasing disagreement about objective facts; a blurring of the line between opinion and fact; the increasing relative volume and influence of opinion over data; and declining trust in formerly respected sources of information. This state acts as a "threat multiplier" that paralyzes political decision-making (Kavanagh, Rich, 2018; Williams, McCulloch, 2023).

The strategic goal of this aggression in the sixth theater of war is detailed by Peter Pomerantsev in his study of the "war against reality." The aim is to "overload our capacity to make sense of information" and provoke a nihilistic skepticism, which A. Arif compares to "listening to static noise through headphones," eventually causing citizens to abandon the search for truth altogether (Starbird, 2019). I. Kalpokas and T. Deryugina add that this "dizzying whirlpool" of reality leads to the erosion of identity and the search for artificially created "pop-up identities" (Kalpokas, 2019).

The psychological mechanism of our society's vulnerability is explained through the phenomenon of "motivated reasoning." As Fukuyama notes, people do not begin with a neutral observation of reality; instead, they use their cognitive abilities to select only the data that confirms their prior preferences or fears (Fukuyama, 2023). RAND research confirms that people are "cognitively lazy" and use heuristics, preferring information that causes emotional arousal, even if it is knowingly false (Paul, Matthews, 2016). This makes debunking nearly impossible, as it can trigger a "backfire effect," strengthening belief in disinformation (Kavanagh, Rich, 2018).

The critical consequence for national defense is that Truth Decay leads to political paralysis and social atomization. M. Naím warns that when power loses control over narratives and society fragments, it becomes impossible to build the consensus necessary to repel an aggressor (Naím, 2013). As J. Rauch concludes, a liberal society cannot survive without establishing a "hierarchy of

factual truths," which the aggressor purposefully destroys, replacing rational analysis with "affective polarization" and hatred. Thus, defense requires not just fighting fakes, but strengthening the psychological immunity of the democratic rear (Fukuyama, 2023).

The effectiveness of the Russian "Firehose of Falsehood" model is scientifically grounded by RAND researchers C. Paul and M. Matthews, who characterize it by "high-volume and multi-channel" messaging. This approach "entertains, confuses, and overwhelms the audience." Its psychological resilience is based on the principle that "quantity has a quality of its own": repeated exposure increases perceived credibility, and first impressions are extremely difficult to correct (Paul, Matthews, 2016).

EEAS reports for 2024–2026 record a qualitative shift in AI use from experimental to "fully integrated." In 2025, AI-related incidents increased by 259%, allowing the mass production of deepfakes and synthetic audio at minimal cost. A particular danger is "LLM grooming" (e.g., the Portal Kombat network), which involves flooding the internet with low-quality content to "poison" AI training data so that future user queries yield manipulative claims as factual results (EEAS, 2026).

Digital mimicry networks demonstrate the industrial scale of aggression. The *Doppelgänger* campaign uses 228 domains and 25,000 CIB (Coordinated Inauthentic Behavior) networks to mimic authoritative Western media like *Der Spiegel*, creating a "hall of mirrors" (EEAS, 2025). Simultaneously, the *Portal Kombat* network demonstrates a "saturation" strategy, publishing nearly 10,000 articles a day in 35 languages. The *False Façade* (Storm-1516) operation has refined "information laundering," moving content from sham YouTube channels and imitator sites to official Russian diplomatic accounts (EEAS, 2025).

The mathematical basis of digital vulnerability was revealed by S. Vosoughi in *Science*, proving that false news spreads 70% faster than the truth and penetrates "deeper and wider" into social graphs (Ebute, n.d.). This creates "cognitive chaos." As Rid notes, modern subversion is increasingly independent of weapons, as its "targets are human minds" (EEAS, 2026).

For the systemic deconstruction of enemy operations, it is necessary to implement the "Kill Chain" model, which decomposes a complex FIMI operation into stages from strategic planning to execution. According to the EEAS, understanding the attack as an end-to-end process allows defenders to develop specific countermeasures at each stage, effectively "killing" the operation before it achieves its strategic goals (EEAS, 2023, 2024).

This approach is grounded in the works of Bruce Schneier and Ben Nimmo, who adapted military Kill Chain models to the cognitive domain. Nimmo proposes tactical analysis by phases to identify specific "Tactics, Techniques, and Procedures" (TTPs). Early detection of behavioral patterns–such as registering inauthentic domains–allows for neutralization before the threat reaches a mass audience (EEAS, 2023, 2024).

Active deterrence strategy, according to the EEAS's "FIMI Deterrence Playbook," involves shifting focus from reactive debunking to proactive influence on the aggressor's vulnerabilities. The

goal is to increase the enemy's operational costs by striking key nodes of their infrastructure. Deterrence in the cognitive sphere is a structured sequence of actions creating cumulative pressure (EEAS, 2026).

Crucially, "pre-bunking" holds an advantage over traditional debunking. RAND researchers argue that the "first impression" of disinformation is highly resilient. Pre-bunking allows for "intercepting" the narrative before the campaign gains momentum. The 2022 US experience of strategically disclosing intelligence on Russian "false flag" operations is a classic example of successful pre-bunking (Williams, McCulloch, 2023).

The necessity of transitioning to active methods is dictated by realities Moises Naím describes as the "End of Power," where power becomes more diffuse and fragmented. In a situation where the "power of large units is diminishing" and "consensus-building is becoming increasingly difficult," only flexible, networked, and preemptive defense strategies are effective (Naím, 2013). We must recognize that in "liquid modernity," as described by Z. Bauman, the only guarantor of security is the ability of the state to act faster than the aggressor (Bauman, 2008).

To prevent democracy from drifting toward tyranny, it is necessary to analyze the risk of defense methods becoming so repressive that they destroy the state's liberal identity. As Fukuyama warns, liberal institutions are "canaries in the coal mines." Democracies die when power attacks the zone of privacy, which is a prerequisite for democratic discussion. Fukuyama emphasizes that "privacy is a norm... that must be viewed as an extension of the virtue of tolerance" (Fukuyama, 2023).

Protecting cognitive sovereignty by strengthening the "constitution of knowledge," according to Rauch, is the only effective means of preserving truth without sliding into state censorship. Rauch's system relies on two rules: "no one has the last word" and "knowledge must rely on empirical evidence, not authority." State attempts to monopolize truth only increase cognitive chaos.

Zygmunt Bauman warns that the erosion of privacy and trust destroys a society's agency, stripping people of the "protective armor of citizenship" (Bauman, 2008). Thomas Rid reinforces these warnings, stating that "the cure must not be worse than the disease." Limiting the free exchange of ideas can erode the "community of ideas and open dialogue" that makes a free society viable. The strategic goal of "active measures" is precisely to provoke democracies into self-destructive steps (Rid, 2020).

Loewenstein's "militant democracy" asserts that constitutions must be "steeled and strengthened," but defense must aim to preserve the spirit of freedom, not destroy it under the pretext of security (Loewenstein, 1937).

Thus, the literature review demonstrates that an active defense strategy is a scientifically grounded response to the systemic threat to cognitive sovereignty. It is based on strengthening institutional immunity and the "constitution of knowledge," allowing the democratic state to protect its cognitive borders while remaining true to its fundamental values. Recognizing the cognitive

domain as the sixth theater of war requires an asymmetrical response: not narrowing the space of freedom, but making it operationally unsuitable for the aggressor's manipulations.

2 Methodological Framework

The development of a theoretical basis for protecting cognitive sovereignty and countering foreign interference is based on an interdisciplinary approach that integrates tools from strategic communications, military theory, political psychology, and international law.

A fundamental principle of this research is the shift from content analysis (disinformation) to manipulative behavior analysis (EEAS, 2025). According to the FIMI (Foreign Information Manipulation and Interference) concept developed by the EEAS, a threat is defined not merely by the falsity of content, but by the intentional, coordinated, and manipulative nature of the actions taken by foreign actors. This allows for the identification of threats in the "gray zone," where information may be formally legal but strategically harmful (EEAS, 2026).

To deconstruct the aggressor’s operations, the Kill Chain model is applied. This model views an attack as an end-to-end process consisting of distinct phases: planning, preparation, execution, and impact assessment. Such an approach enables the detection of the aggressor's "behavioral footprints" at the early stages of infrastructure preparation (EEAS, 2024).

The analysis of democratic vulnerabilities is rooted in the Truth Decay concept developed by the RAND Corporation. The methodology involves examining four trends: increasing disagreement over objective facts, the blurring of the line between opinion and fact, the rising influence of opinion over data, and the decline in trust in expert sources. Particular emphasis is placed on studying motivated reasoning as a driver of cognitive paralysis (Kavanagh, Rich, 2018; Fukuyama, 2023).

The theoretical justification for protective measures relies on Jonathan Rauch’s idea of the "Constitution of Knowledge." The methodology posits that protecting sovereignty should be based not on censorship, but on strengthening a decentralized social system for establishing objective truth through the scientific method, the legal system, and professional journalism (Fukuyama, 2023).

The methodology for transitioning from reactive to proactive defense is based on the FIMI Deterrence Playbook. It involves altering the aggressor’s "cost-benefit" calculus through public exposure mechanisms and the disruption of the logistical and financial chains of "FIMI-as-a-Service."

3 Results

It has been established that the Russian strategy of "information confrontation" views the information space as an integrated ecosystem of warfare, where information is simultaneously a weapon and an environment. The research finds that the aggressor implements a "Firehose of Falsehood" model. Characterized by high speed, massive scale, and a total lack of commitment to objective truth, this model creates a "first impression" advantage that democratic institutions find extremely difficult to overcome due to their procedural inertia (EEAS, 2025).

The study substantiates that the democratic system is vulnerable due to its "legalistic blindness." The adversary freely exploits fundamental rights–such as freedom of speech, press, and assembly–as vectors for legally dismantling the state order from within. It is proven that the Kremlin's strategic goal is to undermine the "Constitution of Knowledge" as the social system for establishing truth, without which liberal democracy cannot function (Loewenstein, 1937).

A transition is proposed from a defensive model of passive debunking to a strategy of active pre-bunking and deterrence. This approach involves exposing the enemy's tactics before an operation begins, thereby increasing the "cognitive immunity" of society and making the conduct of IPSOs (Information-Psychological Operations) operationally costly and strategically ineffective for the aggressor (EEAS, 2025).

The research identifies three echelons of systemic vulnerabilities in democratic countries that require immediate neutralization: Systemic-Legal, Socio-Psychological, and Technological.

The Systemic-Legal Echelon

The study reveals that the enemy conducts "legal sabotage" by exploiting the "Paradox of Openness," using liberal values as attack vectors. The aggressor successfully infiltrates disinformation into democratic discourse by masking it as "alternative views" or "investigative results," forcing free media and legal institutions to protect the enemy's right to destroy those very institutions (Novik, 2023). Russian entities utilize the status of journalists or human rights defenders for direct and indirect advocacy of Kremlin interests, generating coverage in mainstream media because liberal standards demand the representation of diverse perspectives.

Liberal democracies are "highly procedural," with complex rules of coordination and legal hierarchies designed to prevent the abuse of power. However, in a cognitive war, these procedures become "brakes," creating a critical gap in reaction time (Fukuyama, 2023). An operational asymmetry exists between the time of attack and the speed of response. Statistics show that the average duration of a disinformation incident from the first publication to the end of the amplification phase is only 37 hours. Meanwhile, the procedure for collective decision-making or the introduction of sanctions at the EU level requires unanimous political consensus, which can take weeks (EEAS, 2026; Fukuyama, 2023).

The Socio-Psychological Echelon

The fundamental threat is the phenomenon of "Truth Decay," defined by four interrelated trends: 1) increasing disagreement over objective facts; 2) the blurring of the line between opinion and fact; 3) the increasing volume and influence of opinion over data; and 4) the decline in trust in authoritative information sources (Kavanagh, Rich, 2018). This state leads to a "cognitive vacuum" where society loses its capacity for rational democratic discussion (Fukuyama, 2023).

The aggressor uses a "threat multiplier" strategy, where real social challenges are amplified through disinformation to provoke chaos. For example, the migration crisis is artificially inflamed through "facts" disguised as investigations into spikes in violence to increase xenophobia and

political destabilization in the EU. In 2025, the active use of the "Ukrainian refugee-terrorists" theme was recorded during election campaigns in Poland and Romania (Williams, McCulloch, 2023).

Russia implements "divide and rule" tactics, widening societal "cracks" by fueling cultural conflicts. The aggressor instrumentalizes "identity politics," pushing marginalized groups toward radical demands and the conservative majority toward a paranoid fear of "replacement," paralyzing national governments in a state of "vetocracy" (Fukuyama, 2023).

The Technological Echelon

The research substantiates the high efficiency of the "Firehose of Falsehood" model (Paul, Matthews, 2016).

The Portal Kombat (Pravda) network "floods" local information spaces with automated republications of Russian content to dominate search engine results.

The Doppelgänger campaign implements digital mimicry, instantly registering new domains and using "Operation Matryoshka" tactics to inject disinformation directly into the comments sections of fact-checking organizations (EEAS, 2025, 2026).

LLM Grooming: A new form of aggression involving the poisoning of Large Language Model training data (EEAS, 2026).

Deepfakes and Voice Cloning: Routinely used by operations *Storm-1516* and *Overload* to provoke panic or simulate political crises.

Information Laundering: Described through the *False Façade* operation, where content moves from "informant" actors on YouTube to imitation local press sites (e.g., *londoncrier.co.uk*) and is finally cited by official Russian diplomatic accounts (EEAS, 2026).

Operational Recommendations: The Kill Chain and Deterrence

Implementing the Kill Chain mechanism allows for the decomposition of a FIMI operation into phases. Neutralizing a threat during the preparation stage–for example, by identifying "sleeper" networks or funding sources–is significantly more effective as it "breaks the chain" before the attack reaches the target audience (EEAS, 2024).

Effective deterrence in the cognitive domain is achieved through:

Increasing operational costs via sanctions and digital regulation.

Isolating the operational space by blocking distribution channels and demonetizing manipulative networks.

Delegitimizing actors through the public exposure of links between proxy resources and Russian state structures (EEAS, 2026).

Preserving the Liberal Order

The research concludes that the strategic victory of democracy is impossible if the zones of privacy and free discussion are dismantled. Protecting cognitive sovereignty must be based on strengthening the "Constitution of Knowledge" and the immunity of democratic institutions. A drift

toward tyranny during the defense process would signify a total strategic victory for the Kremlin (Humisky, 2020).

The protection of the cognitive domain must rely on a decentralized system of establishing objective truth (courts, scientific communities, professional journalism). Attempts by the state to monopolize the truth through censorship only exacerbate Truth Decay and destroy institutional trust. The primary risk of defense lies in "overshooting the target," where liberal states inadvertently push legitimate civic actions into the realm of illegal subversion (Rid, 2013).

Our response must be asymmetrical: we do not narrow the space of freedom; instead, we strengthen society's ability to exercise it. Protecting sovereignty is, above all, protecting the right of a free individual to remain the subject of their own will.

4 Discussion

The strategy for the active defense of cognitive sovereignty developed in this study resonates with leading global advancements in countering hybrid threats. A comparative analysis of works from the last five years (2020–2026) allows for a deeper appreciation of the significance of the proposed solutions.

The identified qualitative leap in the threat–specifically the surge in AI usage within FIMI operations–correlates with the findings of the EEAS 4th Report (2026), which states that AI has become a "fully integrated" tool for the mass production of manipulative content. As noted by experts from VIGINUM and CheckFirst (2025), the poisoning of AI training data poses a long-term threat of "polluting" the global knowledge base, necessitating not only technical filters but also institutional protection of truth.

The balance between security and liberty remains a point of contention. Our formulated directive to prevent "friendly fire" builds upon Francis Fukuyama's (2022) warnings that democracies perish when the state assaults the zone of privacy and free discussion in an attempt to protect itself from manipulation. This engages in a polemic with Karl Loewenstein's concept of "militant democracy," which permits the restriction of rights to preserve order. We argue that defense must be rooted in Jonathan Rauch's "Constitution of Knowledge"–a decentralized system for establishing objective truth–which serves as the only viable alternative to state censorship or managed chaos.

Conclusion

This study substantiates that the contemporary information space has definitively transformed into a full-scale sixth theater of war, where a state's cognitive sovereignty is a critical element of national security. It is established that the Russian model of aggression in this domain is based on the concept of "non-linear warfare," where information influence has become a permanent front, and non-military means often exceed traditional military force in effectiveness.

The primary results of the research are defined by the following points:

Implementation of the "Kill Chain" Model: Decomposing FIMI operations into discrete phases allows for the identification of the aggressor's specific Tactics, Techniques, and Procedures (TTPs) at early stages. This enables the neutralization of threats before they achieve mass dissemination by focusing on the disruption of the technical and financial infrastructure ("FIMI-as-a-Service") rather than merely debunking content.

Strategic Shift to Active Deterrence: It is established that the traditional reactive model is insufficient against the "Firehose of Falsehood." Instead, the study proves the superiority of pre-bunking–proactively informing the public about manipulative tactics and disclosing intelligence–to foster cognitive resilience among citizens.

Preservation of Liberal Values: The research confirms that protecting the cognitive domain cannot be achieved by dismantling liberal values. Democracies die when the state attacks the very possibility of open communication. Preserving the zone of privacy and the "Constitution of Knowledge" is the foundation of democratic resilience; excessive repressive measures risk becoming a "cure worse than the disease."

AI-Escalation Threat: The rapid growth of AI in influence operations requires a revision of attribution approaches, as these technologies have significantly lowered the entry barrier for subversive activities.

Prospects for further research lie in detailing the legal frameworks for the "gray zone," where cognitive attacks do not reach the threshold of an "armed attack" under *Tallinn Manual 2.0* but still threaten political stability. There is a need to develop new metrics for measuring the rate of Truth Decay and creating machine-learning systems for the automated detection of manipulative behavior.

Recommendations

Based on the analysis, the following recommendations are proposed to neutralize systemic aggression in the cognitive domain:

Integrate the Kill Chain Model into the practice of Strategic communications and cybersecurity units to deconstruct FIMI into discrete phases (planning, infrastructure preparation, execution). Focus should be placed on detecting technical and behavioral indicators during the early stages of network creation.

Implement the FIMI Deterrence Playbook with focus shifted toward increasing operational costs for the aggressor. This includes:

Expanding sanctions against entities providing financial and technical logistics for "FIMI-as-a-Service."

Prosecuting proxies and intermediaries involved in activities bordering on criminal offenses (cybercrime, illicit financing).

Cooperating with platforms for the rapid blocking of Coordinated Inauthentic Behavior (CIB) infrastructure.

Counter AI-Driven Influence priority must be given to developing tools for the automated detection of synthetic content and implementing safeguards against LLM grooming.

In order to modernize Civil Education, media literacy programs should focus on understanding manipulation mechanisms and platform algorithms. Support for independent fact-checking networks and investigative journalism is essential to restore trust in expert institutions.

Cognitive sovereignty requires a multi-agency approach. Information Sharing and Analysis Centers (FIMI-ISAC) should be created to coordinate efforts between the government, the intelligence community, the private sector, and civil society.

To avoid a drift toward tyranny, the state must ensure that defense focuses on neutralizing the manipulative behavior of foreign actors rather than introducing state censorship. Moderation and self-restraint of power are necessary conditions for preserving liberal identity during wartime.

The implementation of these recommendations will transform the cognitive space from a zone of vulnerability into an echeloned system of active defense, where the cost of aggression for the enemy far outweighs any potential benefits.

References

BAUMAN, Z. 2008. *Liquid modernity.* (translated by Y. V. Asochakov) Saint Petersburg: Piter. Available at: https://www.prostranstvo.media/wp-content/uploads/2018/10/bauman_liquid_modernity_2008.pdf

BOURKE, L. 2025. *Review on This Is Not Propaganda: Adventures in the War Against Reality by Peter Pomerantsev.* A Garden from the Libraries. Available at: https://lizbourke.wordpress.com/2025/02/25/89-the-war-against-reality-peter-pomerantsevs-this-is-not-propaganda/

DERYUGINA, T. 2024. *Lessons from "This is Not Propaganda".* Human stories & ideas. Available at: https://tatyana-57116.medium.com/lessons-from-this-is-not-propaganda-ec3fc88efc72

DUBÓCZI, P. and DVORSKÝ, T. 2025. *The limits of the EU's soft power vis-à-vis challengers to European integration.* In: BURGONSKI, P., SKOLIMOWSKA, A., TROJANOWSKA-STRZEBOSZEWSKA, M. and ZENDEROWSKI, R. (eds.). *Europa w stosunkach międzynarodowych: polityka – rola – wyzwania.* Warszawa: Wydawnictwo Naukowe Uniwersytetu Kardynała Stefana Wyszyńskiego, pp. 23–47.

EBUTE, M. E. 2024. *The manipulation of public opinion: Examining the role of intelligence, media and political propaganda in global affairs.* SSRN. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=6576618

Estonian Foreign Intelligence Service (EFIS). 2026. *International security and Estonia 2026.* Available at: https://raport.valisluureamet.ee/2026/en/

EŠTOK, G. 2020. *(Ne)dávne hľadanie strednej Európy*. Košice: Univerzita Pavla Jozefa Šafárika v Košiciach. 192 pp.

European Council on Foreign Relations (ECFR). 2025. *Baltic resilience assessment: Countering the hybrid pressure from the East.*

European External Action Service (EEAS). 2023. *1st EEAS report on foreign information manipulation and interference threats: Towards a framework for networked defense.* Available at: https://www.eeas.europa.eu/eeas/1st-eeas-report-foreign-information-manipulation-and-interference-threats_en

European External Action Service (EEAS). 2024. *2nd EEAS report on foreign information manipulation and interference threats.* Available at: https://www.eeas.europa.eu/eeas/2nd-eeas-report-foreign-information-manipulation-and-interference-threats_en

European External Action Service (EEAS). 2025. *3rd EEAS report on foreign information manipulation and interference threats: Exposing the architecture of FIMI operations.* Available at: https://www.eeas.europa.eu/eeas/3rd-eeas-report-foreign-information-manipulation-and-interference-threats-0_en

European External Action Service (EEAS). 2026. *4th EEAS report on foreign information manipulation and interference threats: Dismantling the FIMI house of cards.* Available at: https://www.eeas.europa.eu/sites/default/files/2026/documents/EEAS%204th%20Threat%20Report_web%20version_1.pdf

European External Action Service (EEAS). 2026. *EEAS special report on foreign information manipulation and interference: The impact of generative AI on electoral integrity.*

EUvsDisinfo. 2024. *Building a false façade.* Available at: https://euvsdisinfo.eu/building-a-false-facade/

EUvsDisinfo. 2024. *Pro-Kremlin disinformation narratives about MH17.* Available at: https://euvsdisinfo.eu/uploads/2024/07/NEW-MH17-2024-almost-final.png

FireEye. 2017. *Anatomy of Russia's 2016 influence operations.* Available at: https://archive.org/details/2017-10-fireeye-anatomy

FUKUYAMA, F. 2023. *Liberalism and its discontents (translated by Budko Y).* Kyiv: Nash Format.

GREGORY, S. 2025. *Russia's Shadow War Against the West.* Center for Strategic and International Studies (CSIS). Available at: https://www.csis.org/analysis/russias-shadow-war-against-west

HUMISKY, J. 2020. *The secret history of disinformation and political warfare.* Diplomatic Courier. Available at: https://www.diplomaticourier.com/posts/the-secret-history-of-disinformation-and-political-warfare

Institute for the Study of War (ISW). 2026. *Russian information operations assessment:* Strategic deception and the erosion of Western cohesion. ISW Reports.

IVANČÍK, R., NEČAS, P. 2022. On disinformation as a hybrid threat spread through social networks In: *Entrepreneurship and Sustainability Issues. Vilnius Entrepreneurship and sustainability center*, ISSN 2345-0282. Vol. 10, no. 1 (2022), pp. 344-357. Available at: https://jssidoi.org/jesi/uploads/articles/37/Ivancik_On_disinformation_as_a_hybrid_threat_spread_through_social_networks.pdf

JAKABOVIČ, L, NEČAS, P. 2023 Political and military implications of private military contractors in Russia In: *Journal for political science, modern history, international relations, security studies* Banská Bystrica, Mateja Bel University, Belianum, 2023. - ISSN 1335-2741. Vol. 26, No. 4 (2023),

pp. 36-54. Available at https://politickevedy.fpvmv.umb.sk/25467/political-and-military-implications-of-private-military-contractors-in-russia

KALPOKAS, I. 2019. *Book review: This is not propaganda: Adventures in the war against reality by Peter Pomerantsev.* The London School of Economics and Political Science. Available at: https://researchonline.lse.ac.uk/id/eprint/107981/

KAVANAGH, J., Rich, M. D. 2018. *Truth decay: An initial exploration of the diminishing role of facts and analysis in American public life.* RAND Corporation. Available at: https://www.rand.org/pubs/research_reports/RR2314.html

LEONARDSON, J. 2020. *Intelligence in public media. Review of the books: Active measures: The secret history of disinformation and political warfare, by Rid T. (2020) and Information wars: How we lost the global battle against disinformation, by Stengel R. (2019).* CIA website. Unclassified Extracts from Studies in Intelligence vol. 64 No.1. Available at: https://www.cia.gov/resources/csi/studies-in-intelligence/volume-64-no-1/active-measures-the-secret-history-of-disinformation-and-political-warfare/

LOEWENSTEIN, K. 1937. *Militant democracy and fundamental rights, I.* The American Political Science Review, Vol. 31, No. 3 (Jun., 1937), pp. 417-432. https://doi.org/10.2307/1948164

NAÍM, M. 2013. *The end of power: From boardrooms to battlefields and churches to states, why being in charge isn't what it used to be.* New York: Basic Books.

NATO Strategic Communications Centre of Excellence (NATO StratCom COE). 2025. *The evolution of Russian information operations: From mass propaganda to targeted polarization.*

NATO Strategic Communications Centre of Excellence (NATO StratCom COE). 2026. *Cognitive warfare and social cohesion: Adapting to decentralized influence operations.*

NOVIK, A. 2023. *Active measures to counter Russian disinformation. Countering disinformation in the context of Russian aggression against Ukraine: Challenges and prospects.* Abstracts of participants of the international scientific and practical conference (Ann Arbor – Kharkiv, December 12-13 2023), pp. 269-272). https://doi.org/10.32782/PPSS.2023.1.70

OPHIR, Y. 2023. *Review of the book: The paradox of democracy: Free speech, open media, and perilous persuasion, by Gershberg Z. and Illing S.* International Journal of Communication, Vol. 17, pp. 338-341. Available at: https://ijoc.org/index.php/ijoc/article/view/20752

PAUL, C., MATTHEWS, M. 2016. *The Russian "firehose of falsehood" propaganda model: Why it might work and options to counter it.* RAND Corporation. Available at: https://www.rand.org/pubs/perspectives/PE198.html

RAND Corporation. 2026. *Cognitive warfare studies: Multi-domain threats and the future of social resilience.* RAND Publishing.

RID, T. 2013. *Cyber war will not take place.* Journal of Strategic Studies, Vol. 35 Issue 1 pp. 5-32, Available at: https://www.tandfonline.com/doi/abs/10.1080/01402390.2011.608939

RID, T. 2020. *Active measures: The secret history of disinformation and political warfare.* Farrar, Straus and Giroux. 513 pp.

SCHMITT, M. (Ed.). 2013. *Tallinn manual on the international law applicable to cyber warfare.* Cambridge University Press. Available at: https://www.onlinelibrary.iihl.org/wp-content/uploads/2021/05/2017-Tallinn-Manual-2.0.pdf

SCHMITT, M. (Ed.). 2017. *Tallinn manual 2.0 on the international law applicable to cyber operations.* Cambridge University Press. https://doi.org/10.1017/9781316822524

SINICYN, S., DENCIOVÁ, M. and EŠTOK, G. 2026. *An informational narrative on Slovakia's energy dependence on Russian gas. Szellem és Tudomány*, Vol. 17 Issue 1–2, pp. 189–207. ISSN 2062-204X. https://doi.org/10.32966/szellem.2026.011

STARBIRD, K. 2019. *Disinformation's spread: Bots, trolls and all of us. Nature website.* Available at: https://www.nature.com/articles/d41586-019-02235-x

VOSOUGHI, S., Roy, D., Aral, S. 2018. *The spread of true and false news online.* Science, Vol. 359 Issue 6380, pp. 1146-1151. Available at: https://www.science.org/doi/10.1126/science.aap9559

WILLIAMS, H., McCULLOCH, C. 2023. *Truth decay and national security: Intersections, insights, and questions for future research.* RAND Corporation. Available at: https://www.rand.org/pubs/perspectives/PEA112-2.html

WOOLLEY, S. (Ed), HOWARD, P. (Ed). 2018. *Computational propaganda: Political parties, politicians, and political manipulation on social media.* Oxford University Press. Available at: https://academic.oup.com/book/25859

Contact addresses

Volodymyr Hurkovskyi
ORCID ID: 0000-0003-2021-5204
Central Research Institute of the Armed Forces of Ukraine, Kyiv, Ukraine
Povitroflotskyi avenue, 28B, Kyiv, Ukraine, 03049
Email: Vladimir.gurkovskyi@gmail.com

Serhii Ilnytskyi
ORCID ID: 0009-0009-3000-9034
Central Research Institute of the Armed Forces of Ukraine, Kyiv, Ukraine
Povitroflotskyi avenue, 28B, Kyiv, Ukraine, 03049
Email: anzhx175wf@icloud.com

Rena Marutian
ORCID ID: 0000-0001-9184-1590
Taras Shevchenko National University of Kyiv, Kyiv, Ukraine
Volodymyrska St., 60, Kyiv, Ukraine, 01033
Email: renata_kiev@i.ua

Oleksandr Hrynchuk
ORCID ID: 0009-0004-6576-1342
The Main Directorate of the Military Law and Order Service
of the Armed Forces of Ukraine
Beresteyskyi Avenue, 55/2, Kyiv, Ukraine, 03113
Email: greenyaa120@gmail.com

Jana Antalíková
ORCID ID: 0009-0004-5259-4733
The University of Security Management in Košice
Košťova 1, Košice, Slovakia 04001,
Email: jana.antal1004@gmail.com