BULLETIN
OF "CAROL I" NATIONAL DEFENCE UNIVERSITY

https://buletinul.unap.ro/index.php/en/

Social Media as a Disinformation Infrastructure: Tactics, Strategies, and National Security

Captain Assistant Lecturer George-Adrian AIONESEI*

*"Henri Coandă" Air Force Academy, Brașov, Romania
e-mail: aioneseiadrian11@gmail.com

Abstract

For the last two decades, the global information environment has become a place for strategic competition and superiority, where boundaries between communication, propaganda, and psychological operations have become much more blurred. The evolution of social media platforms has amplified this transformation, switching from mere channels of civic expression into solid influence infrastructures capable of shaping perceptions, behaviors, and political attitudes. In this context, disinformation is no longer a marginal phenomenon but rather a central instrument for hybrid and cognitive warfare, used against social cohesion and trust in democratic institutions. Thus, this paper analyzes the role of social media in undermining national security, focusing on disinformation tactics and strategies used in the current digital environment. Through this study, we manage to use a combined sequence model to establish links and relationships between the tactics (micro-level) used in disinformation campaigns and the strategies (macro-level) used in order to affect society and the way people think and behave. This article is part of a larger PhD research effort that focuses on the impact of social media instruments in modern conflicts, serving as a means to analyze the role of disinformation in the current digital environment.

Keywords:
Disinformation; Tactics; Strategies; Social Media; National Security; Online.

Article info
Received: 11 April 2026; Revised: 30 April 2026; Accepted: 4 June 2026; Available online: 30 June 2026

Citation: Aionesei, G.A. 2026. "Social Media as a Disinformation Infrastructure: Tactics, Strategies, and National Security." *Bulletin of "Carol I" National Defence University*, 15(2): 27-50. https://doi.org/10.53477/2284-9378-26-17

Introduction

Changes in the communication environment have exceeded the technological levels and generated a structural reconfiguration of modern social power. The creation and development of social media networks and platforms have fundamentally modified the way individuals, institutions, and states communicate, share information, and build realities. As Castells (2009, 38-42) mentions, the shift towards a society of networks has transformed the way communication works in an environment where information becomes a strategic currency.

Once social media platforms, such as Facebook (2004), YouTube (2005), Twitter (2006), or later Telegram, Instagram, and TikTok, have become more popular, everyone has changed their status from a passive consumer to an active agent by creating and disseminating information. This democratic process of communication was initially perceived as a vector of freedom and civic mobilization, very well manifested during the Arab Spring. However, the new openness to free expression involved systemic vulnerabilities, mainly due to a lack of editorial filters, speed of dissemination, and dependence on the algorithmic logic of attention (Chadwick 2017, 19-22).

Social networks cannot be seen anymore as mere cultural and societal interaction spaces, but rather as strategic infrastructures, where information and cognitive confrontations become the centerpieces of modern conflicts. They allow both social mobilization and political manipulation. As Benkler, Faris, and Roberts (2018, 14-23) mention, the digital platforms' architecture favors emotional polarization and distorted narratives dissemination, thus creating the ideal environment for coordinated disinformation campaigns. After 2014, events such as the conflict in Ukraine, electoral implications in the United States, or medical conspiracies during COVID-19, altogether made it clear that social media has grown into a new global threat that might seriously affect national security. By using these platforms, different actors can influence perceptions, undermine public trust in governmental institutions, and destabilize democracy without any kind of military intervention. As a response, organizations such as NATO, the European Union, or the Council of Europe officially recognized information manipulation and external interferences as forms of hybrid threats (EEAS 2025b, 7-11, NATO 2022, 3-6).

Through this article, we aim to explore the way in which social media instruments have become central tools in shaping modern conflicts, focusing on the tactical and strategic use of disinformation that affects national security and public trust. Also, it seeks to identify the mechanisms through which digital ecosystems enable the dissemination and normalization of manipulative narratives. The overarching objective is to provide an analytical framework for understanding the causal chain of disinformation in the digital environment, from disinformation tactics to cognitive effects and ultimately to security implications affecting social cohesion, institutional legitimacy, or state-citizen relationships.

1. Methodology

This article employs a deductive, qualitative literature synthesis, combined with a conceptual framework-building approach. The analytical framework was not meant to conduct a statistical meta-analysis, but rather synthesized peer-reviewed literature, institutional reports, and policy documents addressing disinformation, hybrid warfare, cognitive warfare, and platform governance, published primarily between 2016 and 2025. The synthesis helped develop an analytical model explaining how disinformation tactics can be combined with social media platforms to obtain strategic results with implications for national security. Sources were identified through systematic searches of academic databases (Scopus, Web of Science) and grey literature repositories (RAND, EEAS, NATO), using the search terms "disinformation tactics", "social media manipulation", "cognitive warfare", "hybrid threats", and "information operations". Inclusion criteria required that sources address either the mechanisms of disinformation production and the dissemination or their security-level effects, ensuring analytical relevance to both the tactical and strategic dimensions of the framework.

The sources selection consisted of a thematic and conceptual coding procedure. First, disinformation mechanisms were identified and coded as tactical categories, including content fabrication, source deception, coordination and amplification, infrastructure exploitation, discourse disruption, and AI-enabled production tactics. Second, broader objectives regarding disinformation campaigns were coded as strategic categories, including delegitimization of institutions, polarization, confusion, attention control, normalization, crisis exploitation, and erosion of democratic resilience. Each of the two categories was the subject of further codification by three criteria: recurrence across multiple analyzed sources or documented campaigns, functional role within the disinformation process, and strategic significance in producing measurable security-level effects.

The two-tier analytical framework – distinguishing between disinformation tactics (micro-level operational mechanisms) and disinformation strategies (macro-level objectives) – was developed deductively, drawing on established theoretical models including the information disorder framework (Wardle and Derakhshan 2017, 23-32), affordance theory as applied to social media platforms (Wu, Wu and Xiao 2025, 1-5), and RAND studies on strategic influence operations (Paul and Matthews 2016, 2-9; Mazarr, et al. 2019, 11-27). The categorization and interdependency logic presented in Table 1 represent the synthesized output of this analytical process. The framework is conceptual rather than empirical, but to strengthen the empirical grounding of the framework, the article shows an application of the model on three cases related to Russia's disinformation campaigns in the war against Ukraine, such as Zelensky's surrender deepfake, Bucha massacre disinformation, and Operation Overload. The three cases are not intended as a statistical validation method, but rather as a concrete illustration of the model's analytical applicability.

2. Theoretical Framework – Disinformation, Hybrid warfare, and Cognitive warfare

If during the 20th century, the state had the power to broadcast filtered and manipulated information through governmental channels in order to control the population, the information flow nowadays has become completely decentralized, being interactive, participatory, driven by algorithms, and orchestrated by state and non-state actors. This shift in standards has led to a democratization of communication, and at the same time, has opened a gate to new forms of systemic manipulation.

Disinformation, considered once an auxiliary part of hybrid warfare, has become a central instrument of modern warfare. In the current context, it can be seen as the most sophisticated expression of combining technology with psychology and geopolitics. In other words, it can be defined as the deliberate dissemination of false or distorted information to influence perceptions, behaviors, and decisions (Wardle and Derakhshan 2017, 20-21; Baines, O'Shaughnessy and Snow 2019, 56-59). Unlike misinformation, which represents an unintended error, disinformation involves strategic intentionality, planning, and coordination. In this article, we will focus specifically on disinformation because it involves this intentional aspect that is manipulative and has a strategic finality. In current European terminology, these phenomena are part of a new concept – FIMI – *Foreign Information Manipulation and Interference*, as defined by the European External Action Service (EEAS 2025a, 4-8), which consists of a set of coordinated actions meant to alter information and undermine democratic activities.

Through its nature, disinformation acts simultaneously on three complementary dimensions: communication, psychology, and institution. For the first dimension, disinformation is used to intentionally distort the narrative framework, using real content in false contexts, or create and disseminate persuasive messages to target emotions. The second dimension is based on cognitive biases, such as confirmation bias or motivational thinking, exploiting an individual's natural predisposition to accept information that confirms their personal identity and values (Lewandowsky, Ecker and Cook 2017, 353-369). For the institutional dimension, the main role of disinformation is to diminish trust in authorities, mass media, and the capacity of institutions to distinguish between what is true and what is false.

When we think about the institutional dimension mentioned above, we also think of institutional security. But if we expand it further, we can refer to national security. This concept of national security cannot be limited anymore to only territorial and military actions, taken to protect the critical physical infrastructures in the traditional way. Recent literature shows that security needs to be understood also by reference to information resilience, and the institutions' capacity to maintain the

public trust of the population, to adapt the information dissemination mechanisms to current threats, and to help society respond coherently to information distortions (Dragomir, Ruas-Araujo and Horowitz 2024, 1-10; Uusikylä, et al. 2024, 1-18). From this perspective, the vulnerability of a state does not come only from direct external constraints, but also from weakening the internal functions that are able to affect social coordination, institutional legitimacy, or democratic processes.

This multidimensional approach does not necessarily start and end with these three extents, but rather it explains why disinformation is such an essential element in *hybrid warfare*. The concept of hybrid warfare describes how state and non-state actors employ a combination of conventional and non-conventional means (military, cyber, economic, information, diplomatic) to achieve political objectives without escalating to open armed conflict. Within this spectrum, disinformation is the element with a great capacity to act psychologically and cognitively, because it does not affect physical infrastructures, but rather takes over public perception.

After the annexation of Crimea and pro-Kremlin campaigns targeting the European information space, it has become clear that hybrid warfare is supported by a strong cognitive component. Through disinformation, actors do not just aim to convince, but also to create confusion amongst people. By flooding the public space with multiple and contradictory versions of the "truth", which most of the time cannot be fact-checked fast enough, the trust in public institutions and official information sources starts to erode, ultimately making citizens perceive reality as an unstable construct. Pomerantsev (2019, 123, 164) refers to this strategy as "post-truth era", where manipulation and content control are not made of clear lies, but through constant relativity of truth.

This recent evolution in the strategic thinking of affecting people's minds has been conceptualized as *cognitive warfare*, defined as the most sophisticated type of modern conflict where the human mind becomes the operational domain. In the strategic literature, cognitive warfare depicts the combination of actions that aim to influence cognitive processes, such as perception, attention, emotion, and rational thinking, factors that can be manipulated to obtain political objectives without physical contact (Bernal, et al. 2020, 9-11). RAND analysis shows that this approach is an extension of influential operations (information/influence), switching from persuasion to affecting the decision-making process of the target, including information overload, strategic ambiguity, or bias exploitation (Paul and Matthews 2016, 2-9, Mazarr, et al. 2019, 11-27). Therefore, the main objective of cognitive warfare is to make people act voluntarily against their will. Here is where the reflexive control takes shape (de Goeij 2023, 97-108), as the mechanism through which an actor provides filtered and apparently neutral information, but conceived in such a way as to manipulate the adversary to make decisions that are beneficial to the offender. This strategy is most of the time amplified by algorithms, microtargeting, and influential networks.

If the cyber domain involves technological infrastructures, the cognitive domain aims for mental infrastructures (perceptions and behaviors). A main role in this equation is played by social networks, because of their technological design, algorithms, and manipulative content (Vosoughi, Roy and Aral 2018, 1146-1153). Unlike traditional propaganda, which was dependent on centralized control of mass media, current digital dissemination allows dispersed control, based on the voluntary participation of users. The European Union has tried to respond to these challenges by adopting the Digital Service Act, meant to raise transparency about social media platforms and each user's responsibility. However, there is a structural asymmetry between the efforts required to fact-check and the ease of spreading false information. Truth demands time, expertise, and validation, whilst falsehood only needs channels and networks through which it can circulate freely and instantaneously. This asymmetry represents the main strategic advantage for modern actors that use disinformation as a weapon.

If we consider the flow of (dis)information through social media in society nowadays, as mentioned above, we can follow a sequential escalation process through which disinformation grows from platform-level tactics into strategic security effects. Starting from opportunities offered by social media platforms, such as algorithmic recommendations, network virality, or low effort visibility, actors have the structural conditions to apply and shape manipulation. By using this environment, they can deploy disinformation tactics to maximize dissemination and influence. These tactics, combined in sequence or simultaneously, can generate cognitive effects at the individual or small group level, involving confusion, misperception, or emotional activation, which later will be translated into behavioral responses, engaging people in sharing content, blaming the authorities or institutions, or manifesting outrage towards governments. By using repetition and algorithmic dissemination through networks, these actions form social amplification loops that transform individual encounters into collective ones. Over time, these processes enable the consolidation of disinformation strategies, understood as the macro-level objectives, including polarization or delegitimization of institutions. In this sense, the strategies do not develop on their own but are implemented through the repeated and coordinated use of tactical mechanisms, which, over time, are able to weaken social cohesion, reduce trust in institutions, and affect national security.

Although this sequence captures the big picture of the escalation from platform dynamics to national security implications, this paper focuses specifically on the disinformation tactics and strategies and the relationship between them. To analyze these dynamics systematically, we will address a two-tier analytical model that distinguishes between the tactics and strategies used in the dissemination of disinformation. **Tactics** refer to the micro-level mechanisms through which disinformation operates in social media environments, including practices such as narrative engineering, coordinated inauthentic behavior (CIB), automated amplification, or algorithmic exploitation to increase visibility and engagement

(Bradshaw and Howard 2018, 11-15; Metzler and Garcia 2024, 735-748). **Strategies**, on the other hand, refer to the macro-level objectives that these mechanisms are intended to advance, including delegitimization, polarization, or deterrence by confusion. Although there are no separate frameworks that capture the evolution of disinformation from tactics to strategies that affect national security, there are several models that support individual phases. The information disorder framework (Wardle and Derakhshan 2017, 23-32) conceptualizes disinformation as a process involving agents, messages, and interpreters, while affordance-based models (Wu, Wu and Xiao 2025, 1-5) show how platform features influence cognitive, affective, and behavioral responses. Also, studies on strategic communication, hybrid threats, and cognitive warfare explain how sustained manipulation can produce political and security effects (Dov Bachmann, Putter and Duczynski 2023, 1-10). By considering these models, this article proposes an integrative model, illustrated below in Figure 1.

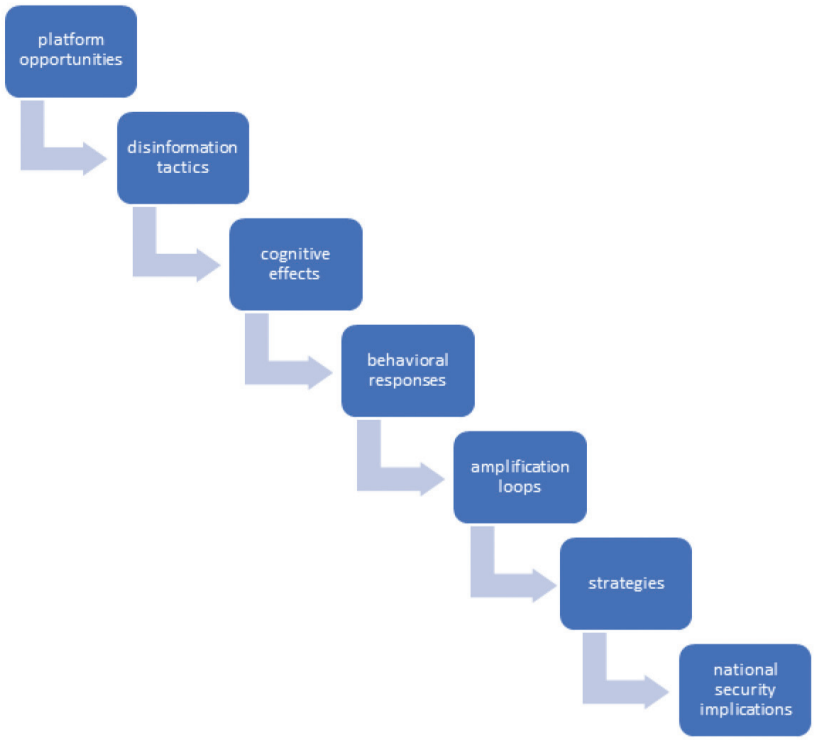

Figura 1 The process through which disinformation evolves from social media platforms to national security threats
Source: Author's interpretation.

3. Tactics and Mechanisms of Disinformation on Social Media

As we mentioned before, disinformation is not a random phenomenon but rather a deliberate, structured process used as a key pillar in hybrid and cognitive warfare strategies. Its success derives from blending technologies such as social media platforms with their algorithms, psychological manipulation, and coordinated amplification of messages. To understand how disinformation operates at the micro-level, we will analyze it through the main categories of tactics. These are isolated techniques, but can be combined in different ways across content,

communication channels, sources, distribution, visibility, or reception. Studies on this subject suggest that disinformation needs to be analyzed from the perspective of the process involving the interaction between content, network amplification, and user engagement on social media platforms in various forms (Chadwick and Stanyer 2022, 1-17). Based on this starting point and other studies (Bradshaw and Howard 2018, 9-15; Kruijver, et al. 2025, 6-20), we can distribute tactics in different categories, each corresponding to a distinct dimension of manipulation.

3.1. Content fabrication and manipulation tactics

This category focuses on the information substance of messages. From production of the content (text, video, audio), to more subtle forms of distortion, disinformation does not rely solely on false information, but also combines facts with manipulative framing, emotional narratives, or conspiracy-based descriptions that provide a simplified and more acceptable interpretation of reality (Egelhofer and Lecheler 2019, 97-111).

Fully fabricated content is one of the most visible mechanisms of this category, by promoting entirely invented facts, under different forms (posts, articles, breaking news, rumors, or pseudo reports). There are also other subtle tactics, such as placing real images or facts in falsified contexts, omitting information that might change users' interpretation, or using narrative engineering to rearrange and organize information related to an event in such a manner that it gives a different meaning, such as betrayal, corruption, censorship, or "us versus them" conflict. In these cases, manipulation lies less in inventing reality and more in distorting the conditions under which reality is interpreted. Other tactics manipulate credibility and evidence by invoking "experts" from different fields (politicians, medics, fabricated witnesses), pseudo-documents, or so-called leaks from the government, big companies, or influential people. More advanced and modern forms include deepfakes and synthetic media, by using AI to impersonate public figures or fabricate events (Farid 2025, 1-9), and false flag content, which is designed to appear as if it originated from another actor, group, or community, thereby avoiding legal measures that can be taken against them (Ferreira 2022, 1537-1540). These two tactics are strategically significant because they involve an increase in uncertainty around the authenticity and attribution, making it hard for audiences to know exactly who is responsible for a specific piece of content, being closely related to the next category of tactics.

Content fabrication and manipulation tactics have the role of distorting the relationship among information, evidence, and interpretation. Their effectiveness does not rely solely on the production of falsehood but also on their capacity to alter context, manipulate credibility, and intensify emotion, making them central to strategies such as delegitimization, polarization, confusion, or crisis exploitation.

3.2. Content identity or source tactics

Another important category of tactics that manipulate perceptions of credibility and authenticity of the content refers to the source or sources of the content. Going beyond the content alteration, these tactics target the perceived origin of messages,

such as fake personas or pseudo-sources that mimic legitimate ones. In a social media environment, users rely on elements such as the identity of an account, popularity, repetition, or apparent credibility to interact with it. Therefore, manipulating who appears to speak, what and how many voices support a narrative, and whether they seem authentic, is a great mechanism of influence used by actors.

One of the most common tactics used in this category is the use of fake personas, or sock puppet accounts, which are fake online identities created to praise, defend, support, or denigrate a person, group, or entity, to manipulate the population's decisions. This tactic is used to disseminate specific content online, while the source might be blocked or restricted. It is effective through the simulation of ordinary users, making coordinated manipulation look like a spontaneous public opinion. Closely related are bots – automated accounts used for posting and sharing content artificially, according to predefined instructions. They can reduce the cost of dissemination and increase the speed and volume of engagement, allowing actors to grow the visibility of a selected narrative (Alkathiri and Slhoub 2025, 1-7). If bots are automated mechanisms, trolls are usually human-operated accounts that take part in the disinformation process aggressively, ironically, or provocatively to push specific narratives, or to discourage the opposition (Hameleers 2023, 4-6). Another complex form of source deception is represented by coordinated inauthentic behavior (CIB), a tactic based on networks of authentic accounts, which are duplicated, falsified, and operated automatically or manually, coordinated in such a manner that they hide the real author and simulate organic behaviors (Murero 2023, 3-4). Astroturfing, close to CIB, involves an organized activity that gives the impression of an organic and spontaneous behavior in favor of or against an idea or opinion, to give it popular consensus, but that does not exist. CIB and astroturfing are relevant for disinformation because they do not distort information, but rather falsify the social media environment in which information is engaged, evaluated, and utilized (Hameleers 2023, 4-6).

3.3. Coordination and amplification tactics

By using these types of tactics, disinformation is scaled beyond the initial point of appearance. They exploit ways in which visibility and repetition can improve, giving the impression that certain contents are organically widespread, urgent, or socially validated. Paul and Mathews (2016, 1-10) depicted this logic through the firehose of falsehood model, highlighting how high volume and repetitive content can overwhelm audiences and their rational thinking.

These tactics include automated amplification, where bots increase visibility through artificial posts, shares, and reactions; coordinating posting where multiple accounts push the same narrative across groups to create an initial impulse; and hashtag or keyword hijacking, where actors insert manipulative narratives in already popular groups and environments, to increase visibility and credibility (Mustafa, Luczak-Roesch and Johnstone 2025, 48-54). Other forms include comment flooding, which overwhelms users with an avalanche of comments and replies, to confuse, intimidate, and finally incapacitate them from deciding; and influencers laundering, where the

content is taken over by recognizable figures so it appears to be more organic and believable.
Socially, these tactics distort the apparent distribution of opinion by bringing manipulated narratives to the front, to appear mainstream. Users' engagement with this content translates into more visibility, creating feedback loops, and ultimately into shaping perception, credibility, and group behavior.

3.4. Platform infrastructure exploitation tactics

Rather than focusing on audience persuasion, these tactics are used to manipulate the algorithms, recommendation systems, and engagement metrics to increase visibility and persistence. They help place strategic content within platform ecosystems, based on preferences, thereby maximizing exposure among specific audiences (Clemons, et al. 2024, 5-11). By using these tactics, actors enable the use of social media platforms as a force multiplier for influence. There are a few tactics, such as optimizing messages which help the recommendation systems of the algorithms to give users specific content for their interests; search flooding, where actors push results, stories, or narratives to dominate specific searches associated with names, events, or topics; and microtargeting, which are used by actors to deliver tailored messages to specific audience segments, based on their interests, vulnerabilities, identities, or behavioral patterns (Kruijver, et al. 2025, 15-16).

These tactics have a great impact through visibility manipulation and audience segmentation. By using algorithmic exploitation, these tactics make selected narratives appear more popular, relevant, or more important than they are. Repeated exposure can increase familiarity and perceived credibility, while personalized messaging can exploit pre-existing fears or vulnerabilities. Also, the impact of microtargeting, as exposing different groups to different versions of reality, can end up in a faulty collective understanding and interpretation of an event, representing a starting point for polarization in society.

3.5. Tactics of suppression and discourse disruption

In this category of tactics, the objective consists of weakening opposing voices to respond effectively. Coordinated harassment, intimidation, or agenda diversion are just a few of those tactics aimed to silence critics and opposing views, discrediting experts, or shifting attention away from subjects that matter. Such practices help develop an environment in which misleading narratives become a normality and can circulate without much resistance.
Harassment, one of the most visible tactics in this category, involves coordinated attacks on experts, or individuals and organizations that contest disinformation narratives. The effect is not focused only on the direct target, but also signals to others that contesting the dominant narrative can generate social punishment, reputational damage, or psychological stress. A complementary tactic is silencing through intimidation, which, on the same logic, discourages others from publicly correcting

misleading claims, even when they possess relevant knowledge. The impact is that manipulated narratives become more accepted, popular, and less contested, whilst credible voices become less visible. Report brigading complements these practices, at the account level, allowing actors to reduce the visibility of adversarial accounts, or even suspend users who oppose their narratives (Wardle 2024, 12-17).
What these tactics do is weaken the counter-discourse. Intimidation discourages experts and ordinary people from intervening, while repeated attacks can diminish trust in those who provide correct information. For the disinformation process, these tactics make it more resilient, not necessarily because of their persuasive character, but because resisting it makes it more costly for the users.

3.6. Artificial Intelligence-enabled tactics

A distinct category of tactics regarding disinformation is the use of Artificial Intelligence (AI) and advanced automation, as an extension of traditional ways to influence the population. The development of technology helps AI to support the phenomenon of disinformation with increasing content production speed, its volume, realism, and personalization capacity of digital messages. Recent literature shows that generative AI does not necessarily change the objectives of disinformation, but contributes to costs, pace, and scalability (Romanishyn, Malytska and Goncharuk 2025, 3-5; Park and Nan 2024, 1502-1504).

One of the most important tactics based on AI is the generation of synthetic content and personas. This includes texts, images, videos, audios, or apparent authentic online identities created or modified by AI. Actors can generate large volumes of content and create credible figures to disseminate that specific material. This can help support or denigrate any agendas or contexts in favor of different actors. Complementarily, AI-assisted microtargeting can help actors to adapt narratives, tone, emotional framing, or claims to specific audience segments to make disinformation more persuasive by aligning messages with users' identities, fears, or ideological orientations.

These categories present different tactics that might be considered innovative and effective, but they cannot work in isolation. For a disinformation campaign to reach its objectives, these tactics must be combined in such a way that their impact derives from coordination, repetition, and adaptation to platform-specific dynamics. Therefore, the significance of these tactics lies in their capacity to be combined, scaled, and sustained over time, generating results that go beyond their individual effects. Also, the identified mechanisms above do not rely solely on absolute falsehood, but rather work through distortion, recontextualization, and manipulation, indicating the necessity of focusing on multiple dimensions (framing, repetition, algorithms), not just on the true-false value. Understanding these mechanisms is essential to both analyzing how disinformation operates at the micro level and explaining how they contribute to broader specific objectives, as well as strategies that derive from them.

4. Disinformation strategies in social media ecosystems

By shifting from the micro level to the bigger picture, we can use, guide, and coordinate these tactical interventions to create disinformation strategies, to influence perceptions, disrupt social cohesion, and alter political and institutional dynamics (Chadwick and Stanyer 2022, 10-14). With these types of objectives in mind, we can refer to strategies as goals that cannot be isolated, but rather as sustained patterns of action, obtained from the coordinated involvement of tactics across time, platforms, or audiences. We are now going to discuss the main categories of strategies that are currently used to support disinformation in social media. Just like in the case of tactics, it is not an exhaustive list of strategies, but it comprises the ones that are used the most and have a great impact on society nowadays.

4.1. Delegitimization of institutions and authorities

This category focuses on eroding trust in governmental institutions, media organizations, scientific experts, or the democratic process; therefore, being a great threat to a few foundation pillars of society. Trust is nowadays replaced with falsehood alternatives, repetitive messages, and large volumes of information, which generate skepticism and questionable thoughts in identifying credible content. Research shows that persistent exposure to disinformation, by using the tactics above, significantly reduces institutional trust and undermines trust in democratic processes (2016 US presidential elections), public health communications (COVID-19), and governance overall (Surjatmodjo, et al. 2024, 1-12).

The causal mechanisms of delegitimization operate through repetition, emotional framing, and diminishing credibility in epistemic authorities. Narratives that present governments as incompetent, corrupt, or not acting in the interest of citizens weaken people's willingness to accept official communications, therefore discrediting those institutions (Lukavska, et al. 2025, 1-11), while attacks on traditional media frame journalism as biased, manipulated, or controlled. On the same page, there are the attacks on scientific expertise that question professional knowledge in areas such as public health or security, replacing fact-checked content with pseudo-science (Lindberg and Denniss 2025, 1-7). Electoral legitimacy can also be the target of promoting and supporting the idea that elections are fraudulent or unfair, or structurally controlled, diminishing confidence in democratic procedures.

As a result, this category of strategies involves suspicion, cynicism, and distrust towards sources that should be guidance and support for social wellbeing. If citizens no longer trust legitimate institutions, expert figures, or electoral procedures, disagreement becomes harder to resolve through institutional channels. The impact of these strategies is not reflected only in the low credibility in institutions, but also in a broader destabilization of the relationship among citizens, knowledge, and authority.

4.2. Polarization and social fragmentation

Considered as central objectives in current disinformation campaigns, the strategies from this category exploit social, political, or cultural divisions and amplify them

through emotional narratives that trigger individual or collective identities. By focusing on the "us versus them" strategy, disinformation increases the erosion of shared reality and reduces the idea of compromise or reaching a consensus. Empirical studies have shown that misinformation related to politics amplifies polarization and fosters ideological approaches (Tucker, et al. 2018, 30-49). What this strategy does is that it affects social cohesion and diminishes people's capacity to react collectively to threats, either internal or external.

Existing social tensions are used by actors to amplify identity conflicts through emotional activation combined with hardening political, ethnic, linguistic, religious, gendered, or other cultural boundaries. The related narratives encourage communities to interpret facts through increasingly antagonistic frames, which, over time, can contribute to radicalization, as users are pushed toward more extreme positions. A plus to this strategy can come from the creation and reinforcement of echo chambers, limiting exposure to alternative perspectives and keeping users within like-minded communities.

When groups of people do not share a standard reference point anymore, social cohesion and collective response capacity start to become more fragile. Public debates become more hostile and move away from solving social problems, therefore diminishing society's ability to react to internal or external threats. Polarization functions both as a political and social effect of disinformation and also as a security strategy, by dividing public opinion, reducing trust between groups, and lowering the capacity to respond to daily challenges.

4.3. Confusion

These strategies shift the focus of disinformation from persuasion towards disruption. Actors do not promote one single coherent narrative that supports a specific agenda, but they disseminate contradictory, ambiguous, or overwhelming information to generate uncertainty and cognitive overload. The firehose of falsehood model clearly illustrates the way how volume, repetition, rapidity, and inconsistency contribute to destabilizing information environments (Paul and Matthews 2016, 2-9). This type of action helps in diminishing the possibility of acquiring strong beliefs and discourages engagement with public discourse, as people are uncertain of what is true and what is not.

The central mechanism of this category is deterrence by confusion, involving enough contradiction and ambiguity spread over the internet, so it makes users disengage from the main issues and institutions struggle with coordination (Hedling and Ördén 2025, 969-974), presenting a mechanism that works on multiple layers. Closely related to this is information flooding, an effect of the firehose of falsehood model, where the information space is saturated with such a high volume of information that users become confused and struggle to distinguish valid information from noise. In this case, the strategic effect, both socially and psychologically, is not persuasion, but paralysis.

4.4. Attention control

These strategies shape what the public notices, discusses, and prioritizes. Instead of confusion, flooding the information environment with competitive narratives,

sensational content, or irrelevant controversies will result in users' attention being redirected from the main issues or inconvenient facts (Loru, et al. 2025, 1-10). Currently, when engagement is more important than relevance, these strategies take advantage of timing and visibility to shift attention as unnoticed as possible.

The functional process of this category often begins with agenda diversion, where actors use alternative topics to redirect attention away from policy issues, failures, and damaging events. Later, this is combined with the sensational element, where content becomes emotionally charged to generate engagement and algorithmic visibility. Also, if content is released at key moments, such as elections or crises, the effect is maximized, as people are continuously searching for explanations and quick reactions. The objective does not always refer to convincing audiences of a specific narrative, but to controlling what becomes visible, what can be treated as urgent, and what disappears from collective attention. Therefore, the effect lies in society's reaction to shift its attention to new, outrageous, and fearful content, making it more fragmented and unstable.

4.5. Normalization

Repeated exposure to misleading content, or out-of-order manipulative content, can gradually shift perceptions of what is acceptable, credible, or plausible. In time, narratives that would normally be excluded can become normal through enough repetition, familiarity, or social context reinforcement. The *illusory truth effect* best portrays this process, as repeated statements are more likely to be perceived as true, regardless of their factual value (Pennycook, Cannon and Rand 2018, 2-7). These strategies are important as they reduce resistance to manipulation and enable disinformation to be accepted as normal in everyday actions.

When actors use narratives in a repeated way, users become more familiar with the information. This strategy supports the normalization of distrust, where suspicion towards institutions and experts seems like common sense. For a better effect, this strategy can work together with confusion strategies. Once audiences become more familiar with a subject, more radical or ideological claims can be introduced gradually, being received as less disruptive than the first ones. What these mechanisms bring to the table is that they reduce resistance to surprising information, shifting the limits of acceptable discourse.

4.6. Crisis exploitation and opportunistic manipulation

A further important strategy involves the exploitation of crises and moments of uncertainty. Nowadays, elections, pandemics, or geopolitical conflicts are the events that get the most attention. During such periods, the demand for reliable information increases abruptly, while verification mechanisms are slow or not trained enough to keep up. Disinformation actors take advantage of these vulnerabilities by introducing misleading narratives that can influence ways of thinking and decision-making processes, affecting on a bigger scale the trust in institutions or the response to a crisis. As Vosoughi et al. have shown, misinformation spreads up to six times faster during crises and can significantly affect public behavior negatively towards

health measures or trust in official communication more than fact-checked content (Vosoughi, Roy and Aral 2018, 1146-1153).
Introducing misleading narratives to harden or redirect attention from political debates, contradict or question official information during emergencies (COVID-19), or spreading false interpretations of an event before reliable official information stabilizes, are just a few strategies that work by combining uncertainty with fear and urgency. The main effect is that disinformation can influence perception and behavior at precise moments when trust, coordination, and institutional communication are most needed.

4.7. Erosion of democratic resilience
The main strategy that encompasses all those mentioned above is represented by the erosion of democratic resilience and national security. This is a cumulative strategic objective of many disinformation campaigns that can weaken trust, polarize societies, or disrupt the collective decision-making process. Whether a threat is internal or external, there is no need for a massive disruption of society, because even partial disruption of trust and coordination can have significant strategic consequences in time by applying the tactics and strategies above (Chadwick and Stanyer 2022, 1-17).

5. Discussions

While tactics and strategies can be analyzed separately, their relationship within the disinformation process is fundamentally interconnected and non-linear. There is no strict formula linking isolated tactical interventions to strategic outcomes; instead, disinformation operates through repeated, coordinated, and adaptive interactions among multiple tactics, which together advance macro-level objectives over time. Critically, this relationship is many-to-many: the same tactic may support multiple strategies simultaneously, and the same strategy may be operationalized through different tactical configurations depending on social context, platform dynamics, and actors' interests. For example, delegitimization depends on the interaction of narrative framing, source deception, amplification, and suppression of corrective voices, while polarization emerges from combinations of emotional activation, coordinated inauthentic behavior, and information saturation. Table 1 maps these interdependencies in detail.

The framework presented in Table 1 can be illustrated through the case of Russian disinformation operations targeting Ukraine and Western European audiences, particularly in the period following the full-scale invasion of February 2022. This case has been extensively documented and provides a concrete example of how tactical mechanisms escalate into strategic security effects. At the tactical level, Russian state-affiliated actors deployed a combination of fabricated content (including doctored images attributed to Ukrainian forces), false flag narratives (presenting Ukrainian defensive actions as aggression), coordinated inauthentic behavior through networks of amplifier accounts across Telegram and Twitter/X, and firehose of falsehood techniques involving simultaneous dissemination of contradictory narratives –

TABLE 1. Strategies and tactics configuration in the social media disinformation ecosystem

Strategy	Tactics core configuration	Supporting tactics	Interdependency logic
Delegitimization of institutions	fabricated content + framing manipulation + conspiracy narratives + content identity and source tactics (fake experts, impersonating, astroturfing)	Coordinated amplification; influencer laundering; harassment of journalists, experts, or officials; algorithmic amplification	Institutions are weakened when negative narratives are repeatedly amplified, appear to originate from credible, authentic sources, and are reinforced by attacks on legitimate voices.
Polarization	Framing manipulation and identity-based narratives + coordinated inauthentic behavior + bots/trolls	Microtargeting; engagement-driven amplification; hashtag manipulation; synthetic personas	Polarization appears when emotionally charged identity narratives are amplified through coordinated networks and reinforced within specific online communities, increasing resentment.
Confusion	Firehose of falsehood + contradictory narratives + misleading or fabricated content	bots; algorithmic exploitation; deepfakes; agenda diversion	Confusion is generated by overwhelming audiences with high volume, repetitive, or contradictory information, making it difficult to distinguish between credible and false content.
Attention control	Platform infrastructure exploitation tactics (algorithmic exploitation) + hashtag manipulation + coordination and amplification tactics	Influencer laundering; coordinating posting; agenda diversion; selective omission	Public attention is redirected when actors manage to exploit platform visibility and algorithms to artificially increase interest in selected narratives while real issues and topics are overshadowed.
Normalization	Amplification tactics, especially repetition + framing manipulation + pseudo-journalistic sources	Influencer laundering; synthetic personas; bots; troll networks; algorithmic recommendation systems	Repeated exposure to specific narratives gradually increases familiarity with the content, making it more acceptable and setting it as the new normal, which helps introduce distrust and misleading frames into everyday discourse.
Crisis exploitation	Fabricated or misleading content related to crisis + emotional framing + fast coordinated amplification	Deepfakes; microtargeting; hashtag manipulation; automated synthetic content	During crises, uncertainty and urgency lower verification processes, allowing narratives to spread rapidly and influence public perceptions, emotions, and behaviors, before official, reliable information is released.
Erosion of democratic resilience	most of the tactics regarding delegitimization, polarization, confusion, or amplification tactics	Suppression and discourse disruption tactics; AI-enabled content; manipulative narrative saturation	Democratic resilience is gradually diminished when multiple strategies and tactics are combined, reducing institutional legitimacy, contributing to polarization, and collective response to crises and internal or external threats.

Source: The author.

denying atrocities, attributing them to other parties, and claiming they were staged (Dov Bachmann , Putter and Duczynski 2023, 1-10). Deepfake content, including a fabricated video of Ukrainian President Zelensky reportedly calling for surrender, demonstrated the deployment of AI-enabled tactics (Farid 2025, 1-9). At the strategic level, these combined operations pursued the delegitimization of Ukrainian governmental and military institutions, the polarization of Western publics around support for Ukraine, the confusion created through information overload that made fact-checking difficult to sustain at scale, and the erosion of trust in NATO and EU institutional communications (EEAS 2025b, 7-11). The interdependency logic visible in this case aligns directly with the many-to-many structure proposed in Table 1: the same coordinated inauthentic behavior networks simultaneously served delegitimization, confusion, and polarization objectives, while the polarization strategy itself drew on framing manipulation, emotional content, and amplification tactics applied across different audience segments in different languages. This case thus validates the analytical utility of the framework proposed in this article, while also confirming that real-world disinformation campaigns are far more fluid and adaptive than any static taxonomy can fully capture.

The Ukraine case discussed above illustrates this logic empirically, showing how the same networks of amplifier accounts simultaneously served delegitimization, confusion, and polarization objectives. This convergence confirms that analytical frameworks focusing on individual tactics or single causal chains will systematically underestimate the scope and adaptability of coordinated disinformation. The key analytical insight is that strategic effects emerge from the sustained, overlapping deployment of multiple tactics – not from any single mechanism acting in isolation.

To strengthen the empirical foundation of the proposed model above, the Ukrainian case can be divided into more specific micro-cases, allowing the seven stages to be illustrated in a clearer manner. Therefore, we chose three well-documented cases, such as President Zelensky's surrender deepfake, Bucha massacre disinformation, and Operation Overload (Matryoshka), compared below in Table 2. The first case refers to March 2022, after the start of the Russian invasion, when a deepfake video circulated on multiple platforms, portraying Volodymyr Zelensky asking the Ukrainians to surrender and go back to their families (Allyn 2022, Bohacek and Farid 2022, 1-3). The Bucha case was based on confusion, denial of identity, and contradictions. After evidence emerged regarding the killing of civilians in Bucha, pro-Kremlin sources negated Russian implications and promoted ideas of fabricated, staged, or falsely attributed content to the aggressor. The "Denying Bucha" study (Fredheim, Ahonen and Pamment 2023, 4-22) showed that these sources posted contradictory and framed information to undermine Western analyses and statements regarding the massacre. The third case, different from the first two, focused more on the amplification loop phase, aiming for journalists, researchers, and fact-checking organizations. It was considered a propaganda effort from the Kremlin meant to undermine the war efforts of Ukraine and destabilize Western democracies (Atanasova, Poldi and Kuster 2025, 8-9).

TABEL nr. 2. Comparative application of the proposed model to three Russian disinformation cases

Phase of the model	Zelensky's surrender deepfake	Bucha massacre disinformation	Operation Overload/ Matryoshka
Platform opportunities	Wartime uncertainty and confusion, public need for coordinated information, speed and visibility offered by platforms, and favorable conditions for fabricated leadership content to gain attention	The impact produced by reports and images from Bucha created an environment where audiences needed visual evidence, continuous updates, and official answers	Platform speed, cross-platform circulation, visibility of journalists, experts, and fact-checkers, and low cost of AI content creation
Disinformation tactics	AI-enabled synthetic videos, Impersonating President Volodymyr Zelensky, Crisis-time exploitation	Denial narratives, false-flag framing, contradictory explanations, firehose messages from different media sources	Fabricated content (images, videos), AI-generated content, impersonation of public figures, fake news formats, coordinated dissemination
Cognitive effects	Uncertainty, panic, and doubt regarding leadership and message credibility	Confusion about attribution and responsibility, doubt regarding visual evidence, and uncertainty about credibility in media sources	Doubt about authenticity, fact-check fatigue, confusion about sources, authorities, and the overload of authorities in fact-checking large volumes of content
Behavioral responses	Rapid cross-platform sharing, Debunking/ fact-checking, public discussion	Online debates, engagement with contradictory claims, public arguments over attribution and responsibility	Public figures, journalists, and fact-checkers were forced to debunk, verify, and react to the overflow of manipulated content
Amplification loops	The video became viral due to platform sharing, media coverage, analysis activity, and public discussion about deepfakes during wartime	Repetition by pro-Kremlin channels, same agenda for all media channels, social media accounts, and global analysis of the event	Content was covered multiple times for debunking and gaining more visibility
Strategies	Crisis exploitation, delegitimization of Ukrainian leadership, and attempted disruption of morale and command credibility	Deterrence by confusion, Delegitimization of Ukrainian and Western institutions, and agenda diversion	Attention control, agenda manipulation, and normalization of anti-Ukraine and pro-Kremlin narratives
National security implications	Potential weakening of trust in leadership, morale, and credibility in public institutions	Public outrage, accountability, and doubt of Western support for Ukraine	Weakening of media and fact-checking entities, increased public uncertainty, and erosion of trust in legitimate institutions

Source: The author.

This table does not validate the model statistically, but rather shows that each stage can be applied to different cases. We chose the three instances related to the Russian invasion in Ukraine because they illustrate different forms of escalation: from a rapid episode of AI manipulation in a crisis context, to a campaign of denial and contradictions regarding an atrocity, and a recent information overload operation exploiting the fact-checking ecosystem. As stated above, we can observe that the process is not a perfect recipe, such as a linear process, but rather adaptive, depending on the context, platform, and audience.

All these tactics and strategies above highlight the fact that disinformation operates through interconnected and mutually reinforcing processes. Tactics are the operational tools through which manipulation occurs at the micro level, whilst strategies represent the broader objectives that these mechanisms aim to achieve over time. They are dependent on each other, but they are not related in a linear or one-dimensional way. Regarding the situations encountered, we can see that there are tactics that can be used to multiple strategic objectives, whilst a single strategy can be implemented through the coordinated deployment of multiple tactics across various platforms, for different audiences, and in different time contexts.

The tactics and strategies analyzed above represent the most operationally significant mechanisms documented in the current literature on digital disinformation. Several cross-cutting insights emerge from the analysis:

- An important aspect is that disinformation on social media should be understood as a multi-level escalation process, and not as a simple flow of false or misleading content. Disinformation grows through a sequence of interconnected levels, from possibilities offered by platforms that enable tactic manipulation, to cognitive and behavioral responses, then amplification through algorithms, and finally, if sustained over time, to applying strategies at the societal level. This shows the shift from content-centered approaches to a process in which manipulation at the micro level can obtain macro-level strategies.
- A second aspect refers to social media functioning as a connective infrastructure between tactics and strategies. Platforms offer different features, such as visibility, rapidity, algorithm recommendation, or low-cost dissemination, which do not necessarily facilitate the dissemination of information, but they shape the conditions under which tactical mechanisms can be scaled and sustained.
- A third important characteristic is that the relationship between tactics and strategies is not straightforward as one link, but rather as many-to-many. One single tactic can be part of multiple strategies, while a strategy can depend on multiple tactics. As an example, CIB might have an important contribution to delegitimization, polarization, normalization, or agenda manipulation, whereas, for example, polarization might require fake experts, fabricated

content, amplification, or framing manipulation. Therefore, disinformation campaigns can be constructed through an overlapping and adaptive relationship between tactics and strategies.

- Also, for specific strategies, there are multiple tactical configurations, depending on the context, objectives, or societal vulnerabilities. There is no straight way of constructing a strategy from a specific tactic. This suggests that the analytical focus should be placed on linking specific tactics in order to obtain the most efficient result, instead of focusing on only one tactic.
- By focusing on the last category of strategies mentioned, regarding the erosion of democratic resilience, we could indicate that disinformation on social media is cumulative rather than immediate. Disinformation rarely undermines trust, cohesion, or institutional legitimacy through a single video, audio, or text, or in a manner of short time. Instead, its effects appear through repetition, coordination, amplification, and persistence. In time, repeated narratives can normalize distrust, intensify polarization, or increase confusion. Following this logic makes it easier to understand disinformation as a process, rather than a collection of messages.
- The last important aspect is represented by the role of social media in undermining social security. Social media acts indirectly towards achieving goals, being an enabler of conditions through which disinformation tactics can become strategic processes. Cognitive and behavioral loops, amplification through algorithms, consistency, or repetition are a few features of social media platforms that work behind the scenes to help actors reach their goals.

Together, these observations confirm that social media functions not merely as a communication environment but as a strategic infrastructure – one that enables disinformation tactics to be coordinated, amplified, and sustained until they produce macro-level security effects.

The original contribution of this article lies in the integrative framework that shows the evolution of disinformation campaigns from platform capabilities to national security implications, going through the intermediate phases mentioned earlier. While existing frameworks approach the steps of the disinformation process, they rarely clarify how platform tactical mechanisms evolve into larger strategic objectives. The proposed model addresses this issue by illustrating how social media platforms enable micro-level manipulation to be enhanced and sustained to the point that it supports strategies such as delegitimization, polarization, deterrence by confusion, agenda manipulation, normalization, and crisis exploitation. Its primary value lies in offering a structured instrument for examining the evolution of disinformation from isolated tactical actions into persistent strategic influence with potential security relevance.

Conclusion

This article has examined the role of social media in undermining national security by mapping the relationship between disinformation tactics and strategies. Rather

than treating social media as a passive channel through which false content circulates, the analysis has demonstrated that disinformation operates as a multi-level escalation process: platform affordances enable tactical manipulation, tactical mechanisms generate cognitive and behavioral responses, and sustained coordinated deployment transforms these responses into strategic security effects. Social media is thus a strategic infrastructure that enables society-level manipulation with measurable political and security consequences.

The article distinguished two analytical levels: disinformation tactics (micro-level mechanisms including content fabrication, framing manipulation, source deception, coordinated amplification, and AI-enabled production) and disinformation strategies (macro-level objectives including delegitimization of institutions, polarization, confusion, normalization, crisis exploitation, and erosion of democratic resilience). As the framework and the Ukraine case study confirm, these levels are connected through adaptive, overlapping, many-to-many relationships rather than linear causal chains. One tactic may serve multiple strategies; one strategy may draw on multiple tactical configurations depending on context and target vulnerabilities. The security implications of disinformation are therefore cumulative: strategic effects – weakened institutional trust, intensified polarization, eroded democratic resilience – accumulate through the sustained, coordinated application of tactical mechanisms over time, affecting the information, cognitive, and institutional foundations on which democratic societies depend.

These results also suggest a few several practical implications. First, counter-disinformation policies should switch their focus from the correction of individual false claims and address the broader process through which this information distortion is amplified, repeated, and made socially credible. To achieve this, there is a need for closer cooperation between legitimate institutions, platform companies, researchers, fact-checkers, and social organizations, especially during important events, which necessitates more resources to prevent and debunk disinformation. Second, media literacy and digital literacy programs should extend their focus from identifying false content to understanding trends, manipulation techniques, source deception, or coordinated amplification for a better understanding of where users should be more alert. Third, by analyzing the proposed model, democratic resilience should be improved by finding the elements of the model where it necessitates more attention and development, so the link to national security implications is broken.

References

Alkathiri, Nasser, and Khaled Slhoub. 2025. "Challenges in machine learning-based social bot detection: a systematic review." *Discover Artificial Intelligence* 5 (214): 1-40. https://doi.org/10.1007/s44163-025-00448-w.

Allyn, Bobby. 2022. *Deepfake video of Zelenskyy could be 'tip of the iceberg' in info war, experts warn.* 03 16. https://www.npr.org/2022/03/16/1087062648/deepfake-video-zelenskyy-experts-war-manipulation-ukraine-russia.

Atanasova, Aleksandra, Francesco Poldi, and Guillaume Kuster. 2025. "Operation Overload, More Platforms, New Technology, Powered by AI." Analysis report.

Baines, Paul, Nicholas O'Shaughnessy, and Nancy Snow. 2019. *The SAGE Handbook of Propaganda*. SAGE Publication Ltd.

Benkler, Yochai, Robert Faris, and Hal Roberts. 2018. *Network Propaganda - Manipulation, Disinformation, and Radicalization in American Politics.* Oxford University Press.

Bernal, Alonso, Cameron Carter, Ishpreet Singh, Kathy Cao, and Olivia Madreperla. 2020. *Cognitive warfare an attack on truth and thought.* Johns Hopkins University.

Bohacek, Matyas, and Hany Farid. 2022. "Protecting world leaders against deep fakes using facial, gestural, and vocal mannerisms." *Proceedings of the National Academy of Sciences of the United States of America* (National Academy of Sciences) 119 (48): 1-3. https://doi.org/10.1073/pnas.2216035119.

Bradshaw, Samantha, and Philip N. Howard. 2018. "Challenging Truth and Trust: A Global Inventory of Organized Social Media Manipulation." *Computational Propaganda Research Project.* Oxford Internet Institute, University of Oxford. 11-15.

___. 2018. "Challenging Truth and Trust: A Global Inventory of Organized Social Media Manipulation." *Computational Propaganda Research Project.* Oxford Internet Institute, University of Oxford. 11-15.

Castells, Manuel. 2009. *Communication Power.* New York: Oxford University Press.

Chadwick, Andrew. 2017. *The Hybrid Media System: Politics and Power.* 2nd Edition. New York: Oxford University Press.

Chadwick, Andrew, and James Stanyer. 2022. "Deception as a Bridging Concept in the Study of Disinformation, Misinformation, and Misperceptions: Towards a Holistic Framework." *Communication Theory* 32 (1): 1-24. https://doi.org/10.1093/ct/qtab019.

Clemons, Eric K., Andrej Savin, Maximilian Schreieck, Stina Teilmann-Lock, Jan Trzaskowski and Ravi Waran. 2024. "A face of one's own: The role of an online personae in a digital age and the right to control one's own online personae in the presence of digital hacking." *Electronic Markets* 34 (1): Article 31. https://doi.org/10.1007/s12525-024-00713-3.

de Goeij, Maria W.R. 2023. "Reflexive control: Influencing Strategic Behavior." *Parameters: The US Army War College Quarterly* 53 (4): 97-108. https://doi.org/10.55540/0031-1723.3262.

Dov Bachmann, Sascha-Dominik, Dries Putter, and Guy Duczynski. 2023. "Hybrid warfare and disinformation: A Ukraine warperspective." *Policy Insights* 858-869. https://doi.org/10.1111/1758-5899.13257.

Dragomir, Marius, Jose Ruas-Araujo, and Minna Horowitz. 2024. "Beyond online disinformation: assessing national information resilience in four European countries." *Humanities and Social Sciences Communications* 11: 101. https://doi.org/10.1057/s41599-024-02605-5.

EEAS. 2025a. *2024 Report on EEAS Activities to Counter Foreign Information Manipulation and Inference (FIMI).* FIMI, Brussels: European External Action Service, 4-8. https://www.eeas.europa.eu/sites/default/files/2025/documents/2024%20Report%20on%20EEAS%20Activities%20to%20Counter%20FIMI.pdf.

___. 2025b. "EEAS Report on Foreign Information Manipulation and Interference Threats." Report on FIMI Threats, 7-11.

Egelhofer, Jana Laura and Sophie Lecheler. 2019. "Fake news as a two-dimensional phenomenon: a framework and research agenda." *Annals of the International Communication Association* 43 (2): 97-116. https://doi.org/10.1080/23808985.2019.1602782.

Farid, Hany. 2025. "Mitigating the harms of manipulated media: Confronting deepfakes and digital deception." *PNAS Nexus* 4 (7): pgaf194. https://doi.org/10.1093/pnasnexus/pgaf194.

Ferreira, Ricardo Ribeiro. 2022. "Liquid Disinformation Tactics: Overcoming Social Media Countermeasures through Misleading Content." *Journalism Practice* 16 (8): 1537-1558. https://doi.org/10.1080/17512786.2021.1914707.

Fredheim, Rolf, Anneli Ahonen, and James Pamment. 2023. *Denying Bucha - The Kremlin's Influence tactics in the aftermath of the 2022 Bucha atrocity.* Research report, Lund University.

Hameleers, Michael. 2023. "Disinformation as a context-bound phenomenon: toward a conceptual clarification integrating actors, intentions and techniques of creation and dissemination." *Communication Theory* 33 (2): 1-10. https://doi.org/10.1093/ct/qtad004.

Hedling, Elsa, and Hedvig Ördén. 2025. "Disinformation, Deterrence and the Politics of Attribution." *International Affairs* 101 (3): 967-986. https://doi.org/10.1093/ia/iiaf012.

Kruijver, Kimberley, Neill Bo Finlayson, Beatrice Cadet, and Sico van der Meer. 2025. "The disinformation lifecycle: an integrated understanding of its creation, spread and effects." *Discover Global Society* 3 (1): 1-26. https://doi.org/10.1007/s44282-025-00194-5.

Lewandowsky, Stephan, Ullrich K.H. Ecker, and John Cook. 2017. "Beyond Misinformation: Understanding and Coping with the' Post-Truth' Era." *Journal of Applied Research in Memory and Cognition* 6 (4): 353-369. https://doi.org/10.1016/j.jarmac.2017.07.008.

Lindberg, Rebecca, and Emily Denniss. 2025. "Social media and the spread of misinformation: infectious and a threat to public health." *Health Promotion International* 40 (2): daaf023. https://doi.org/10.1093/heapro/daaf023.

Loru, Edoardo, Alessandro Galeazzi, Anita Bonetti, Emanuele Sangiorgio, Niccolò Di Marco, Matteo Cinelli, Max Falkenberg, Andrea Baronchelli, and Walter Quattrociocchi. 2025. "Ideology and polarization set the agenda on social media." *Scientific Reports* 15 (35816): 1-13. https://doi.org/10.1038/s41598-025-19776-z.

Lukavska, K., R. Gabrhelík, M. Miovský, N. Hynek, B. Gavurova, L. Stastna, M. Bartak, B. Petruzelka, and V. Moravec. 2025. "Exploring Disinformation: The interplay of exposure, trust, and sharing." *Computers in Human Behavior Reports* 18: 100686. https://doi.org/10.1016/j.chbr.2025.100686.

Mazarr, Michael, Abigain Casey, Alyssa Demus, Scott Harold, Luke Mathews, Nathan Beaucham-Mustafaga, and James Sladden. 2019. *Hostile Social Manipulation.* Santa Monica: RAND Corporation. https://www.rand.org/pubs/research_reports/RR2713.html.

Metzler, Hannah, and David Garcia. 2024. "Social Drivers and Algorithmic Mechanisms on Digital Media." *Perspectives on psychological science: a journal of the Association for Psychological Science* 19 (5): 735-748. https://doi.org/10.1177/17456916231185057.

Murero, Monica. 2023. "Coordinated inauthentic behavior: An innovative manipulation tactic to amplify COVID-19 anti-vaccine communication outreach via social media." *Frontiers in Sociology* 8: 1141416. https://doi.org/10.3389/fsoc.2023.1141416.

Mustafa, Hassan, Markus Luczak-Roesch, and David Johnstone. 2025. "Conceptualizing the Evolving Nature of Computational Propaganda: A Systematic Literature Review." *Annals of the International Communication Association* 49 (1): 45-60. https://doi.org/10.1093/anncom/wlaf001.

NATO. 2022. *Strategic Concept.* Brussels: NATO. https://www.nato.int/content/dam/nato/webready/documents/publications-and-reports/strategic-concepts/2022/290622-strategic-concept.pdf.

Park, Seyeon, and Xiaoli Nan. 2024. "Generative AI and misinformation: a scoping review of the role of generative AI in the generation, detection, mitigation, and impact of misinformation." *AI & Society* 41 (2): 1501-1515. https://doi.org/10.1007/s00146-025-02620-3.

Paul, Christopher, and Miriam Matthews. 2016. "The Russian `Firehose of Falsehood` Propaganda Model: Why It Might Work and Options to Counter It." *RAND Corporation.* July 11. https://doi.org/10.7249/PE198.

Pennycook, Gordon, Tyrone Cannon, and David Rand. 2018. "Implausibility and illusory truth: Prior exposure increases perceived accuracy of fake news but has no effect on entirely implausible statements." *Journal of Experimental Psychology General* 147 (12): 2-7. https://doi.org/10.1037/xge0000465.

Pomerantsev, Peter. 2019. *This is Not Propaganda.* London: Faber & Faber.

Romanishyn, Alexander, Olena Malytska, and Vitaliy Goncharuk. 2025. "AI-driven disinformation: policy recommendations for democratic resilience." *Frontiers in Artificial Intelligence* 8: 1569115. https://doi.org/10.3389/frai.2025.1569115.

Surjatmodjo, Dwi, Andi Alimuddin Unde, Hafied Cangara, and Febri Alem Sonni. 2024. "Information Pandemic: A Critical Review of Disinformation Spread on Social Media and Its Implications for State Resilience." *Social Sciences* 13 (8): 418. https://doi.org/10.3390/socsci13080418.

Tucker, Joshua, Andrew Guess, Pablo Barberá, Cristian Vaccari, Alexandra Siegel, Sergey Sanovich, Denis Stukal, and Brendan Nyhan. 2018. "Online Content and Political Polarization." *Social Media, Political Polarization, and Political Disinformation: A Review of the Scientific Literature.* William and Flora Hewlett Foundation. 30-49.

Uusikylä, Petri, Harri Jalonen, Valdemar Kallunki, Anssi Keinänen, and Silvia Sommarberg. 2024. "Introduction to Information Resilience in the Context of National Preparedness." In *Informationa Resilience and Comprehensive Security*, by Petri Uusikylä, H Jalonen, and A Jokipii, 1-18. Palgrave Macmillan, Cham.

Vosoughi, Soroush, Deb Roy, and Sinan Aral. 2018. "The Spread of True and False News Online." *Science* 359: 1146-1151. https://doi.org/10.1126/science.aap9559.

Wardle, Claire. 2024. *A Conceptual Analysis of the Overlaps and Differences between Hate Speech, Misinformation, and Disinformation.* New York: United Nations.

Wardle, Claire, and Hossein Derakhshan. 2017. *Information Dissorder: Toward an Interdisciplinary Framework for Research and Policy Making.* Strasbourg: Council of Europe, 20-21.

Wu, Manli, Tailai Wu, and Yushan Xiao. 2025. "Why people share misinformation on social media? An integration of affordance and flow theories." *Humanities and Social Sciences Communications* 12: 1129. https://doi.org/10.1057/s41599-025-05511-6.