doi 10.55804/TSU-ti-3/Petrocchi

დეზინფორმაცია და ტყუილი: თანამედროვე მსოფლიოს ანტიკური იარაღი

DISINFORMATION AND DECEPTION: AN ANCIENT WEAPON IN A MODERN WORLD

მაურიციო პეტროჩი მაჩერატის უნივერსიტეტი
Maurizio Petrocchi Università di Macerata

https://orcid.org/0000-0003-0493-6503

საკვანძო სიტყვები: *დეზინფორმაცია, ჰიბრიდული ომი, კოგნიტური მანიპულაცია, ტყუილი, ფსიქოლოგიური ოპერაციები*
Keywords: *disinformation, hybrid warfare, cognitive manipulation, psychological operations, PSYOP*

ABSTRACT

Disinformation has historically been employed as a strategic tool for influencing public perception, shaping political discourse, and destabilizing adversaries. From ancient military deception to modern hybrid warfare, the evolution of disinformation reflects technological advancements and shifting geopolitical landscapes. This paper explores the theoretical foundations of disinformation, its historical precedents, the mechanisms through which it operates, and its implications for democratic governance and security. Through comparative case studies, including Cold War intelligence operations and contemporary digital disinformation campaigns, this study underscores the persistent role of manipulation in global affairs. Furthermore, it examines modern countermeasures, ranging from fact-checking initiatives to AI-driven detection systems, and highlights future challenges posed by emerging technologies such as deepfakes and algorithmic manipulation.

INTRODUCTION

Disinformation can be defined as the intentional dissemination of false or misleading information with the objective of deceiving, manipulating public opinion, and influencing political, economic, or social events (Wardle & Derakhshan, 2017). This phenomenon differs from other forms of information distortion due to its strategic and deliberate nature, often employed by both state and non-state actors as an instrument of power and influence. The techniques of disinformation range from document falsification to the deployment of deepfake technology, from social media manipulation to the construction of alternative narratives aimed at destabilizing democratic institutions (Bennett & Livingston, 2020). A fundamental characteristic of disinformation is its capacity to exploit cognitive biases and psychological vulnerabilities, leading individuals to believe in and further propagate false information. The effectiveness of disinformation does not rest solely on the falsity of its content but rather on its ability to elicit strong emotional responses, exacerbate polarization, and foster distrust toward official sources (Zelenkauskaitė, 2022). To comprehensively understand this phenomenon, it is essential to distinguish it from two closely related concepts: misinformation and malinformation (Wardle, 2019). While disinformation denotes the deliberate dissemination of falsehoods with the intent to deceive and manipulate, misinformation refers to the inadvertent spread of false information without an intention to mislead. A common example is the uncritical sharing of fabricated news on social media, often without verifying its credibility. Malinformation, in contrast, involves the disclosure of genuine information that has been manipulated or selectively presented to inflict harm upon an individual, group, or institution. A paradigmatic instance is the strategic release of hacked emails or confidential documents to damage political adversaries, as exemplified during the 2016 U.S. presidential election. These three categories frequently intersect, engendering an information disorder that obscures the boundary between truth and falsehood. This problem is exacerbated by the rapid pace of digital dissemination, which facilitates the viral spread of content via social media platforms, blogs, and alternative news websites (Cosentino, 2020). Despite its contemporary relevance, disinformation is far from a modern phenomenon. The manipulation of information for strategic advantage has deep historical roots, dating back to antiquity. History offers numerous salient examples of disinformation as a tool of political and military strategy. One of the most well-documented instances is Octavian's propaganda campaign against Mark Antony,

wherein Octavian employed inscriptions, coinage, and pamphlets to depict Antony as a traitor subjugated by Cleopatra's influence (Wilson, 2021). The myth of the Trojan Horse remains one of the most emblematic cases of military deception predicated on the manipulation of enemy perceptions (Roberts, 2023). Similarly, the Byzantine Empire's reliance on counter-information strategies, including espionage and intelligence networks to destabilize adversaries, further underscores the long-standing role of disinformation in geopolitical conflicts (Richardson, 2022). By the 18th century, disinformation had evolved into more sophisticated forms, exemplified by the construction of Potemkin villages, designed to mislead Empress Catherine the Great about the prosperity of Crimea (Brown, 2022). The 20th century witnessed the institutionalization of disinformation as a state apparatus, particularly through the rise of modern propaganda mechanisms. Under Joseph Goebbels, the Nazi regime exerted meticulous control over public narratives, while the Soviet Union formalized dezinformatsiya as a strategic doctrine, employing it to disseminate false narratives on a global scale (Harris, 2021). The Cold War era provided further illustrations of disinformation's efficacy, with operations such as "Infection"—a Soviet disinformation campaign falsely alleging that the U.S. had engineered the HIV/AIDS virus—demonstrating the power of influence operations in shaping global perceptions (Miller, 2022). With the advent of the digital era, disinformation has undergone an unprecedented transformation. Today, the integration of artificial intelligence, social bots, big data analytics, and deepfake technology has enabled the automation of large-scale, highly targeted disinformation campaigns (Chen, 2023). Contemporary hybrid warfare incorporates digital disinformation alongside cyber warfare and psychological operations, as evidenced by the Russian disinformation campaigns in the 2016 U.S. elections and the ongoing informational warfare surrounding the conflict in Ukraine (Johnson & Lee, 2022). This historical continuum underscores the persistent role of disinformation as a weapon of influence and manipulation. While technological advancements have amplified its reach and sophistication, the underlying principles remain strikingly consistent across time. Understanding the mechanisms of disinformation—both past and present—is imperative in devising effective countermeasures to safeguard democratic institutions and public discourse in the digital age.

§1 THEORETICAL FRAMEWORK OF DISINFORMATION

The study of disinformation, information warfare, and propaganda draws upon a diverse array of theoretical and methodological approaches spanning multiple disciplines, including communication sciences, strategic studies, social psychology, political sociology, and intelligence studies. The rapid evolution of digital technologies has necessitated a continuous reassessment of traditional theoretical models to account for the emerging dynamics of informational manipulation and cognitive conflict.

A rigorous analysis of disinformation requires a clear differentiation from related phenomena to avoid conceptual ambiguities and ensure methodological precision. As Wardle and Derakhshan (2017) argue, disinformation operates within a broader **informational ecosystem**, which also encompasses **misinformation**—the inadvertent dissemination of false information due to errors or a lack of verification—and **malinformation**, the deliberate release of truthful information that has been decontextualized or manipulated to inflict harm. This taxonomy facilitates a more precise delineation of the strategies employed by actors engaged in informational manipulation and aids in the formulation of effective countermeasures. The study of **propaganda** and **information warfare** necessitates a **multidisciplinary approach**, integrating various analytical perspectives. The historical perspective traces the evolution of propaganda and disinformation techniques over time, illustrating how they have adapted to technological advancements and shifting political landscapes (Taylor, 2023). Communication sciences focus on the mechanisms of narrative construction, information dissemination, and the role of both traditional and digital media in propagating disinformation (Bennett & Livingston, 2020). Social psychology examines the cognitive and emotional mechanisms that render individuals susceptible to informational manipulation, such as confirmation bias and the illusory truth effect, whereby repeated exposure to falsehoods enhances their perceived credibility (Zelenkauskaitė, 2022). Strategic and intelligence studies analyze disinformation as a weapon of cognitive warfare and a key component of influence operations in international relations (Rid, 2020). Finally, semiotic and political analysis explores the symbolic and linguistic codes that construct alternative realities and disseminate simulacra of truth (Greimas & Courtés, 1982). Disinformation campaigns are situated within the broader domain of information warfare, which is defined as the strategic use of information to influence decision-making processes, induce uncertainty, or destabilize political and social systems. According to the Russian Ministry of Defense, information warfare seeks to undermine adversarial societies through **psychological conditioning** and **political pressure** (Gerasimov, 2013). A particularly salient concept within this domain is **cognitive warfare**, which focuses on shaping perceptions and altering decision-making through targeted psychological and narrative strategies (Olejnik, 2024). During the Cold War, the Soviet Union developed the doctrine of **active measures**, a collection of covert operations designed to

manipulate public opinion and political affairs in Western states. These measures included the fabrication of false news attributed to Western sources, the infiltration of journalists and opinion leaders to manipulate media narratives, and the strategic support of extremist groups to foment social and political unrest (Pacepa & Rychlak, 2013). One of the most notorious examples of Soviet disinformation was **Operation Infektion**, a campaign aimed at spreading the false claim that HIV/AIDS had been engineered by the United States as a biological weapon. In the contemporary era, Russian disinformation strategies no longer seek to impose an alternative truth but rather to **cultivate confusion and distrust**. This approach aligns with the maxim that **«nothing is true, everything is possible»** (Pomerantsev, 2019).

Among the most widely employed techniques are: **whataboutism** – the deflection of criticism by making counter-accusations intended to divert attention from the original issue; **flooding the zone** – the saturation of the informational environment with an overwhelming volume of often contradictory narratives, thereby obscuring factual reality; **the strategy of chaos** – the deliberate amplification of internal divisions and societal conflicts to weaken social cohesion. Disinformation campaigns are frequently designed to exacerbate **political polarization**, obstructing constructive public discourse and fostering **social fragmentation**. Psychological studies have demonstrated that individuals exhibit a **confirmation bias**, gravitating toward information that aligns with their pre-existing beliefs. Moreover, the **repetition effect**—whereby repeated exposure to a falsehood increases its credibility—further amplifies the influence of disinformation. Emotionally charged content, particularly that which evokes **fear and outrage**, is significantly more likely to be disseminated and reinforced within digital spaces (Vosoughi et al., 2018). The advent of social media has profoundly transformed the dissemination of disinformation, exponentially increasing its reach and effectiveness through advanced technological mechanisms. The deployment of bots and troll farms enables the large-scale amplification of manipulative narratives (Zelenkauskaitė, 2022), while political microtargeting leverages user data to craft tailored messages for specific demographic groups. The proliferation of deepfake technology and synthetic media is further eroding the distinction between reality and fiction, enhancing the efficacy of disinformation campaigns and diminishing trust in informational ecosystems (Chesney & Citron, 2019). In sum, disinformation operates at the intersection of technological innovation, psychological manipulation, and strategic influence. Its evolving methodologies necessitate ongoing scholarly attention to develop robust analytical frameworks and effective countermeasures against its pervasive societal impact.

§ 2 OPERATION SAPPHIRE: A CASE OF DISINFORMATION DURING THE COLD WAR

Operation Sapphire constitutes a paradigmatic case of Cold War information warfare, exemplifying the sophistication of Soviet disinformation strategies. Through an analysis of declassified CIA documents and direct testimonies, this study examines the operational structure of the KGB, its impact on French political dynamics, and its broader implications for the security of the Atlantic Alliance. The findings underscore the strategic manipulation of perceptions and the exploitation of systemic vulnerabilities within complex organizations, providing a historical framework that remains highly relevant for understanding contemporary hybrid security threats. The Cold War was characterized not only by ideological and geopolitical confrontations but also by an intense struggle for informational dominance. Soviet disinformation operations were systematically designed to destabilize Western alliances and undermine trust in democratic institutions. In this context, Operation Sapphire stands out as a compelling case study of strategic infiltration and informational manipulation. In 1961, the KGB initiated a multi-layered penetration strategy targeting the French state apparatus. The operation coincided with growing tensions between Charles de Gaulle's France and the Kennedy administration, particularly regarding France's role within NATO and its independent nuclear policy. The strategic objectives of the operation were twofold:

Intelligence Gathering – Acquiring high-value intelligence on NATO's military planning, particularly regarding European defense strategies and nuclear capabilities. Political Manipulation – Influencing French political perceptions to exacerbate the rift between Paris and Washington, thereby weakening Western unity. Operation Sapphire, also known as Operation Martel, was executed through three primary mechanisms: direct Infiltration – The KGB strategically positioned intelligence assets within the French administration, establishing clandestine communication channels that circumvented official diplomatic networks (Life, 1968). This infiltration allowed Soviet operatives to exert direct influence over key decision-making processes. The operation leveraged fabricated documents, media infiltration, and psychological operations to distort French perceptions of both NATO and the United States. These efforts included forgeries aimed at misleading senior French policymakers.

Media disinformation campaigns, whereby Soviet-controlled or co-opted journalists disseminated anti-American narratives. Psychological conditioning, targeting high-level government officials to cultivate mistrust toward Western allies. Intercepting Diplomatic and Military Communications.

The KGB deployed advanced Signals Intelligence (SIGINT) techniques to intercept highly sensitive Fran-

co-American exchanges. This provided Moscow with crucial insights into NATO's strategic posture and facilitated tailored disinformation efforts aimed at deepening Franco-American divisions.

The repercussions of Operation Sapphire were significant, particularly in Franco-American relations. By manipulating diplomatic perceptions and exploiting existing tensions, the operation contributed to De Gaulle's increasing distrust toward the United States, accelerating France's withdrawal from NATO's integrated military command in 1966. A major intelligence coup for the Soviet Union, as the KGB gained access to classified NATO defense strategies and technical specifications of French nuclear weapons. This operation exemplifies how targeted disinformation efforts can influence the geopolitical orientation of democratic states, demonstrating the efficacy of strategic deception in shaping diplomatic and military decision-making. Operation Sapphire shares notable structural similarities with other Soviet disinformation campaigns, particularly Operation Infektion in the 1980s. Both operations Utilized fabricated documents and media infiltration to erode trust in Western institutions. Targeted key vulnerabilities within Western societies to sow discord and exacerbate divisions. However, a key distinction between the two lies in their geopolitical objectives: Infektion sought to discredit the United States on a global scale, fostering anti-American sentiment across multiple regions. Sapphire was tailored specifically to weaken NATO by exacerbating intra-alliance divisions, demonstrating the adaptability of Soviet disinformation strategies. The analysis of Operation Sapphire offers valuable insights into the mechanisms of Cold War disinformation and their long-term implications for global security. The operation underscores several critical takeaways: Strategic Disinformation as a Diplomatic Tool – Disinformation is not merely a tactical instrument but a strategic weapon capable of reshaping international alliances and altering geopolitical balances. Vulnerability of Democratic Decision-Making – The operation highlights how democratic states are particularly susceptible to perception manipulation, given their reliance on open information flows and public discourse. Continuity in Hybrid Warfare Tactics – Many of the techniques pioneered during the Cold War—false narratives, media infiltration, and SIGINT exploitation—remain central to modern hybrid warfare, particularly in the context of Russian influence operations in the digital age (Culloty & Suiter, 2021).

§. 3 ACTORS AND ARCHITECTURES OF DISINFORMATION CAMPAIGNS

The actors involved in disinformation campaigns are numerous and diverse, encompassing state and non-state actors, including governments, terrorist organizations, corporations, and private individuals. These campaigns operate within highly complex architectures, designed to disseminate, manipulate, and amplify strategic narratives for political, ideological, economic, or personal gain.

States frequently deploy disinformation as a strategic instrument of both foreign and domestic policy, often leveraging intelligence agencies and affiliated entities to execute these operations. Russia, for example, has long been recognized for its influence operations, notably through the Internet Research Agency (IRA), which played a central role in spreading manipulative content during the 2016 U.S. presidential elections (Bennett & Livingston, 2020). Similarly, China employs state-directed disinformation units, such as the so-called «50 Cent Party», an extensive network of online commentators tasked with diverting criticism and reinforcing government narratives (Zelenkauskaitė, 2022). Other authoritarian regimes, such as Iran and North Korea, adopt similar tactics to influence Western public opinion and consolidate internal legitimacy (Rid, 2020). The Belarusian government, for instance, has systematically employed disinformation strategies to delegitimize political opposition and manipulate both domestic and international perceptions (Pacepa & Rychlak, 2013).

Terrorist organizations also utilize disinformation as a tool for recruitment, radicalization, and ideological dissemination. ISIS, for example, has developed highly sophisticated propaganda strategies, using digital platforms and social media to attract recruits and spread extremist narratives (Pomerantsev, 2019). Similarly, Al-Qaeda has leveraged disinformation to construct anti-Western narratives, adapting its messaging to different cultural and regional contexts (Libicki, 1995).

Beyond state and terrorist actors, corporations have also engaged in disinformation campaigns for economic and competitive advantages. Some firms employ covert strategies to manipulate competitors' reputations, while others use misleading advertising or fabricated data to promote their products (Wardle & Derakhshan, 2017). In addition to state-backed operations, private individuals, including independent trolls, conspiracy theorists, and ideological influencers, contribute significantly to the spread of manipulative content. These actors often engage in disinformation for financial gain, political activism, or ideological motives, amplifying false narratives through social media and alternative media platforms (Chesney & Citron, 2019). The architecture of disinformation campaigns is multilayered and relies on three interconnected components: Primary Sources of Disinformation – These include intelligence services, state-affiliated actors, and specialized units responsible for generating and distributing manipulative content. Communication Channels – Disinformation is disseminated through: Ideologically aligned media outlets (state-sponsored

news agencies and proxy platforms). False flag websites, which masquerade as independent sources while serving as disinformation vectors. Strategic influencers, who amplify propaganda via social media and traditional media. Secondary Amplification Channels – These consist of: Mainstream media, which may inadvertently amplify disinformation through uncritical reporting. Opinion leaders and online discussion groups, which further legitimize disinformation narratives. Unwitting users, who perceive the content as credible, unknowingly propagate disinformation, exacerbating the reach and impact of manipulative campaigns (Pacepa & Rychlak, 2013). The use of bots and troll networks is particularly prevalent in creating a false sense of relevance, manipulating engagement metrics, and skewing online discourse (Zelenkauskaitė, 2022). Additionally, algorithmic amplification techniques—which exploit social media dynamics to artificially boost specific content—have become a cornerstone of digital disinformation strategies (Bennett & Livingston, 2020). Empirical evidence demonstrates the far-reaching impact of state-sponsored disinformation campaigns: Russia has been widely accused of interfering in the 2016 U.S. presidential elections, deploying a vast network of fake accounts to spread polarizing content and deepen societal divisions (Bennett & Livingston, 2020). In Europe, the Kremlin played a disruptive role during the 2017 Catalonia independence referendum, leveraging disinformation to destabilize Spain and discredit democratic institutions (Rid, 2020). China has engaged in disinformation efforts not only to control domestic narratives but also to shape global perceptions of Hong Kong's pro-democracy protests and the COVID-19 pandemic response (Wardle & Derakhshan, 2017). The United States has itself faced scrutiny for the use of political microtargeting, in which personal data has been exploited to distribute personalized and manipulative political messages (Chesney & Citron, 2019). The European Union has encountered sustained disinformation campaigns, particularly Russian influence operations, which have sought to erode trust in EU institutions. This has led to the establishment of the East StratCom Task Force, a dedicated unit countering Russian propaganda and online influence operations (Zelenkauskaitė, 2022). NATO has recently intensified counter-disinformation efforts, implementing digital threat analysis frameworks and fostering interstate cooperation to combat hostile influence operations (Libicki, 1995). The proliferation of disinformation campaigns underscores their evolution into a central instrument of hybrid warfare, political influence, and economic competition. The intersection of state actors, terrorist organizations, corporations, and independent agents within complex disinformation architectures reveals the strategic depth of contemporary influence operations. As digital platforms continue to reshape global communication, the need for robust countermeasures—including algorithmic transparency, digital literacy initiatives, and coordinated intelligence efforts—remains imperative to safeguard democratic resilience against manipulative information warfare.

§ 4 THE SOCIO-POLITICAL IMPACT OF DISINFORMATION

The impact of disinformation manifests at both societal and political levels, fostering polarization, eroding trust in institutions, and posing a direct threat to democratic governance. A distorted information ecosystem, disrupted by disinformation campaigns, weakens social cohesion, diminishes public confidence in media and institutions, and may contribute to the rise of authoritarian and populist movements. Disinformation fuels opinion polarization, deepening societal divisions and obstructing democratic discourse. Social media algorithms play a crucial role in this process, as they curate and distribute content based on users' prior engagements and behavioral patterns. This algorithmic filtering fosters the emergence of «echo chambers» and «filter bubbles», where individuals are exposed primarily to ideologically homogeneous content (Pariser, 2011). As a result, discussions and attitudes become increasingly radicalized, reducing opportunities for engagement with diverse perspectives and reinforcing polarizing narratives (Sunstein, 2017).

External actors exploit these divisions as part of broader geopolitical influence operations. A notable example is Russia's interference in the 2016 U.S. presidential elections, where disinformation campaigns were designed to exacerbate ideological rifts and weaken public trust in the democratic process (Bennett & Livingston, 2020). Disinformation campaigns are deliberately structured to delegitimize traditional sources of information, sowing mistrust in democratic institutions, the media, and scientific authorities. These efforts frequently involve the dissemination of conspiracy theories and fabricated narratives, intended to undermine public confidence in governments, experts, and established knowledge systems (Wardle & Derakhshan, 2017). A paradigmatic example is the proliferation of anti-vaccine disinformation during the COVID-19 pandemic, which significantly eroded trust in the scientific community and government health policies (Zelenkauskaitė, 2022). The deliberate spread of falsehoods regarding vaccine safety and efficacy contributed to vaccine hesitancy, prolonging the pandemic's public health and economic consequences.

Disinformation represents an existential threat to democracy, as it distorts electoral processes, suppresses voter participation, manipulates public opinion, and fosters cynicism toward democratic institutions (Chesney & Citron, 2019). The erosion of trust in traditional media and government creates fertile ground for the emergence of populist movements, which often capitalize

on manipulated narratives to consolidate political influence. The 2018 and 2022 Brazilian presidential elections serve as case studies in the role of disinformation in mobilizing populist support. Far-right political groups leveraged WhatsApp message chains to propagate conspiracy theories and defamatory attacks against political opponents. The closed nature of these messaging platforms made it exceedingly difficult for fact-checkers and regulatory bodies to mitigate the spread of falsehoods (Pomerantsev, 2019). The widespread adoption of social media and digital communication technologies has contributed to the fragmentation and segmentation of public opinion. Political debate is increasingly confined to self-selected ideological groups, reducing mutual understanding among individuals with divergent perspectives and intensifying partisan polarization (Sunstein, 2017).

Moreover, digital platforms prioritize emotionally charged, sensationalist content, which attracts higher levels of engagement compared to neutral, fact-based information. This algorithmic bias amplifies disinformation campaigns, as fabricated stories often elicit strong emotional reactions, such as fear, outrage, or indignation (Vosoughi et al., 2018). Populist movements frequently exploit disinformation to manipulate voter sentiment, leveraging fear and prejudice to consolidate their support bases (Wodak, 2015). The strategic use of inflammatory messaging, often relying on stereotypes and oversimplified narratives, can significantly shape electoral behavior and reinforce anti-establishment sentiments (Mudde, 2004). Disinformation is also employed as a tool to discredit political opponents and create divisions among social groups, fostering a climate of mistrust and antagonism (Bennett & Livingston, 2020). A well-documented case is the role of Russia's Internet Research Agency (IRA) during the 2016 U.S. elections, which deployed thousands of fake accounts to spread highly polarizing content on controversial issues, such as gun control, immigration, and racial tensions (Rid, 2020). This operation sought to radicalize opposing ideological factions, ultimately undermining faith in the electoral process. The weaponization of disinformation extends beyond the United States, affecting democracies worldwide: «N075 Unified Estonia» (2010) – A fictional political movement was created as an experiment in mass manipulation. Over six weeks, organizers conducted rallies, published election ads, and engaged in social media activism, leading the majority of the public to perceive it as a genuine political force. The case underscores the ease with which artificial political consensus can be engineered (Pacepa & Rychlak, 2013).

Hungary's Anti-Migrant Disinformation Campaign – Government-affiliated media framed migration as an existential threat, amplifying xenophobic narratives, particularly among rural and lower-educated populations. This effort was reinforced through billboards, television advertisements, and state-controlled news outlets, linking migrants to crime and social instability (Howard et al., 2018).

Myanmar's Rohingya Crisis (2017) – The spread of hate speech and disinformation on Facebook played a central role in inciting violence against the Rohingya Muslim minority. Extremist groups and propagandists leveraged digital platforms to portray the Rohingya as a national security threat, fueling ethnic cleansing and resulting in thousands of casualties (Zelenkauskaitė, 2022).

§ 5 COUNTERMEASURES AND DEFENSE STRATEGIES

The fight against disinformation necessitates a multi-layered approach that integrates technological, legislative, and educational strategies. Effective countermeasures include fact-checking mechanisms, artificial intelligence applications, regulatory frameworks, cognitive resilience initiatives, and media literacy programs. Each of these components plays a crucial role in mitigating the spread and impact of manipulative narratives, yet they also present challenges and limitations that must be addressed.

Fact-checking remains one of the fundamental pillars in the fight against disinformation, ensuring the accuracy and integrity of public discourse. Initiatives such as «EU versus Disinformation», managed by the East StratCom Task Force of the European External Action Service, systematically monitor and expose disinformation campaigns (Wardle & Derakhshan, 2017). Globally, independent organizations such as Snopes, FactCheck.org, and PolitiFact play a vital role in analyzing and verifying the truthfulness of statements and news reports. Despite its importance, fact-checking has inherent limitations. The velocity of disinformation dissemination often outpaces the ability of fact-checkers to analyze and debunk false content. Furthermore, the repetition of false information, even after being debunked, can reinforce its credibility among the public, a phenomenon known as the «illusory truth effect» (Lewandowsky et al., 2012). To address these challenges, complementary strategies are required, including: Curating blacklists of unreliable websites and disinformation sources. Implementing algorithmic interventions to reduce the visibility of false content on digital platforms. Encouraging proactive verification, where audiences are equipped with critical digital literacy skills to assess information independently. Artificial intelligence (AI) has emerged as a powerful tool in the detection and mitigation of disinformation. Machine learning algorithms are increasingly employed to: analyze linguistic patterns and identify recurring disinformation tactics.

Detect manipulated content, including synthetic media and deepfakes. Flag coordinated influence campaigns operating across digital platforms. Technologies such as the verification plugin «InVid» are already in use

for exposing doctored videos (Figueira & Oliveira, 2017). Additionally, image forensics enables the authentication of digital images, identifying alterations and manipulations that could mislead audiences. However, AI-driven solutions are not without challenges. The rapid evolution of disinformation techniques necessitates continuous algorithmic updates, as new formats and platforms emerge. Moreover, social media platforms often lack transparency in their data-sharing policies, restricting researchers' ability to develop comprehensive detection models. The absence of standardized regulatory mechanisms for AI-powered moderation further complicates its implementation. Legislative frameworks play a crucial role in establishing cybersecurity standards, ensuring content transparency, and imposing accountability measures on digital platforms. The European Union's Code of Practice on Disinformation (CPD), introduced in 2018, engages technology companies and stakeholders in the regulation of false content dissemination. Several countries have adopted stricter legal frameworks to combat disinformation and online manipulation:

Germany introduced the NetzDG (Netzwerkdurchsetzungsgesetz), which mandates social media platforms to remove illegal content within 24 hours of detection. France enacted legislation targeting election-related disinformation, granting judicial authorities the power to swiftly remove false content during electoral periods. While regulatory measures are necessary to curb the spread of disinformation, their implementation raises significant concerns regarding freedom of expression. The challenge lies in striking a balance between regulating disinformation and preserving fundamental democratic rights (Chesney & Citron, 2019). Overly restrictive policies may risk inadvertently suppressing legitimate discourse, making it imperative that legislative actions are transparent, proportionate, and subject to democratic oversight. A key pillar of disinformation resistance is the development of cognitive resilience, which refers to an individual's capacity to critically assess and recognize manipulative information. This approach emphasizes education and awareness, equipping the public with the necessary analytical skills to navigate an increasingly complex information landscape.

§ 6 CONCLUSIONS AND FUTURE PERSPECTIVES

The evolution of disinformation is increasingly shaped by technological advancements, which make it ever more difficult to distinguish between truth and falsehood. Among the most significant trends is the development of deepfake technology, an artificial intelligence-driven tool capable of generating highly realistic yet manipulated audiovisual content to influence public opinion. Simultaneously, big data analytics and machine learning algorithms enable disinformation actors to customize deceptive messages with unprecedented precision, thereby enhancing the efficacy of manipulation campaigns.

Another emerging challenge is the integration of virtual and augmented reality (VR/AR) into disinformation strategies. These immersive technologies have the potential to reinforce misleading narratives by creating experiential environments that blur the boundary between reality and fiction. Furthermore, the evolution of digital platforms continues to introduce new distribution channels, complicating monitoring efforts and exacerbating the difficulty of containing disinformation.

In parallel, state-sponsored information warfare is becoming more sophisticated, reflecting the intensification of geopolitical struggles in the digital domain. Governments increasingly leverage hybrid disinformation tactics to destabilize democratic institutions and shape global political discourse (Echeverría, García Santamaría & Hallin, 2025). The battle against disinformation is inherently dynamic, with challenges evolving alongside digital technologies. Among the most pressing issues are: the adaptability of bots and algorithms – AI-driven disinformation tactics are becoming increasingly advanced, capable of mimicking human behavior and bypassing detection systems. The velocity and volume of false information – The sheer speed at which false narratives spread overwhelms traditional fact-checking mechanisms, making real-time intervention extremely difficult. The balance between freedom of expression and democratic stability – Regulating disinformation without compromising civil liberties remains a delicate and highly contested issue.

The accountability of digital platforms – While tech companies play a pivotal role in the dissemination of disinformation, concerns persist over granting them excessive control in determining what constitutes acceptable content. Additionally, emerging technologies—such as blockchain-based decentralized networks—present a dual challenge: on the one hand, they could enhance content verification and promote transparency; on the other, they could facilitate more sophisticated disinformation strategies that evade centralized regulation. The fragility of democratic systems further underscores the need to reinforce institutional resilience, ensuring that decision-making processes remain insulated from disinformation-driven distortions. Addressing disinformation requires a coordinated and sustained effort among governments, supranational institutions, technology firms, and civil society. To enhance democratic resilience, policymakers must prioritize information transparency and platform accountability through robust regulatory mechanisms.

Key strategic imperatives include: strengthening media literacy and civic education, large-scale media literacy initiatives should be implemented to equip citizens

with the skills necessary to recognize and counter manipulative content. Educational curricula must integrate critical thinking modules, fostering information discernment from an early age. The Finnish model—which embeds media literacy within the national education system—serves as a best-practice benchmark (Mihailidis & Viotty, 2017). Disinformation is a transnational phenomenon, requiring global cooperation in standardizing legislative responses. The EU Code of Practice on Disinformation marks an initial step toward a structured regulatory approach, yet further refinement is needed to enhance enforcement mechanisms. Governments must develop collaborative initiatives to counter foreign interference in democratic processes. While tech companies should not become absolute arbiters of truth, they must uphold greater transparency regarding content distribution algorithms. The implementation of independent oversight mechanisms is necessary to prevent arbitrary censorship while mitigating the proliferation of false narratives. Interdisciplinary groups, composed of experts in communication, artificial intelligence, law, and security, should be deployed at national and international levels.

These task forces would serve as adaptive response units, analyzing emerging disinformation threats and devising real-time countermeasures. Disinformation is a complex and ever-evolving challenge, necessitating a multidisciplinary and cooperative approach. Strategies to combat it must leverage advanced technological tools, robust regulatory measures, and widespread media literacy programs.

While technology has undeniably facilitated the rapid spread of disinformation, it also provides innovative tools to counteract its effects. The future of policymaking in this domain must remain flexible and adaptable, continuously evolving to address new challenges while ensuring a delicate balance between safeguarding democratic stability and protecting freedom of expression.

To effectively build resilience against disinformation, the following principles must guide future efforts: Adaptive Policy Design – Regulations must be periodically reviewed to account for technological shifts and emerging threats. Enhanced Cross-Sector Collaboration – The intersection of governmental, academic, and private sector expertise is crucial for developing scalable solutions.

Sustained Public Awareness Campaigns – Strengthening societal resilience through continuous engagement and education initiatives is essential for long-term success. Ultimately, only through a concerted and sustained effort—involving governments, institutions, media, and civil society—can we hope to construct a more resilient and trustworthy information ecosystem.

REFERENCES

Allcott, H., and Matthew G. (2017). *Social Media and Fake News in the 2016 Election*, Journal of Economic Perspectives 31 (2): 211-236.

Amarasingam, A., Marc-André, A. (2020). *The QAnon Conspiracy Theory: A Security Threat in the Making?,*CTC Sentinel 13 (7): 37-44.

Andrew, Ch, and Vasili, M. (2000). *The Sword and the Shield: The Mitrokhin Archive and the Secret History of the KGB*. New York: Basic Books.

Barber, M. (2006). The Trial of the Templars. Cambridge: Cambridge University Press. Boghardt, Th. (2009). *Operation INFEKTION: Soviet Bloc Intelligence and Its AIDS Disinformation Campaign*, Studies in Intelligence 53 (4): 1-24.

Bending Spines: *The Propagandas of Nazi Germany and the German Democratic Republic*. East Lansing: Michigan State University Press.

Bennett, W. L., & Livingston, S. (2020). *The Disinformation Age: Politics, Technology, and Disruptive Communication in the United States*. Cambridge University Press.

Bradshaw, Samantha, and Philip N. Howard. 2019. *The Global Disinformation Order: 2019 Global Inventory of Organised Social Media Manipulation,* Oxford: Project on Computational Propaganda. Bytwerk, Randall L. 2004.

Campbell, W. Joseph. 2001, *Yellow Journalism: Puncturing the Myths, Defining the Legacies*. Westport: Praeger.

Chesney, R., & Citron, D. (2019). *Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security.* California Law Review, 107(6), 1753-1818.

Chesney, Robert, and Danielle Keats Citron. 2019, *Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security*, California Law Review 107: 1753.

Del Vicario, Michela, Alessandro Bessi, Fabiana Zollo, Fabio Petroni, Antonio Scala, Guido Caldarelli, H. Eugene Stanley, and Walter Quattrociocchi. 2016, *The Spreading of Misinformation Online*, Proceedings of the National Academy of Sciences 113 (3): 554-559.

Dwyer, Philip. 2015, *Citizen Emperor: Napoleon in Power. New Haven*, Yale University Press. European Commission. 2020, The Digital Services Act Package.

Echeverría, M., García Santamaría, S., & Hallin, D. C. (2025). *Disinformation and Political Communication in the Digital Age*. Oxford University Press.

Fabricates Social Media Posts for Strategic Distraction, Not Engaged Argument, American Political Science Review 111 (3): 484-501.

Festinger, Leon. 1957. *A Theory of Cognitive Dissonance*, Stanford, Stanford University Press. Gillespie, Tarleton. 2018. Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions that

Shape Social Media. New Haven: Yale University Press.

Figueira, Á., & Oliveira, L. (2017). *The Current State of Fake News: Challenges and Opportunities.* Procedia Computer Science, 121, 817-825.

Gabriel Recchia, Anne Marthe van der Bles, and Sander van der Linden. 2020, *Susceptibility 13 to Misinformation about COVID-19 around the World*, Royal Society Open Science 7 (10): 201199. Taylor, Philip M. 2003. *Munitions of the Mind: A History of Propaganda from the Ancient World to the Present Day,* Manchester, Manchester University Press.

Hasan, Haya R., and Khaled Salah. 2019. *Combating Deepfake Videos Using Blockchain and Smart Contracts*, IEEE Access 7: 41596-41606. 12

Holt, Thaddeus. 2004. The Deceivers: Allied Military Deception in the Second World War. New York: Scribner.

Howard, P. N., Woolley, S. C., & Calo, R. (2018). *Algorithms, Bots, and Political Communication in the US 2016 Election: The Challenge of Automated Political Disinformation.* Journal of Information Technology & Politics, 15(2), 81-93.

King, David. 1997. The Commissar Vanishes, *The Falsification of Photographs and Art in Stalin's Russia*, New York, Metropolitan Books.

King, Gary, Jennifer Pan, and Margaret E. Roberts. 2013. *How Censorship in China Allows Government Criticism but Silences Collective Expression*, American Political Science Review 107 (2): 326-343. King, Gary, Jennifer Pan, and Margaret E. Roberts. 2017. "How the Chinese Government

Kovic, Marko, Adrian Rauchfleisch, Marc Sele, and Christian Caspar. 2018. *Digital Astroturfing in Politics: Definition, Typology, and Countermeasures*, Studies in Communication Sciences 18 (1): 69-85.

Krentz, Peter. 2007, *The Trojan Horse*. In The Oxford Encyclopedia of Ancient Greece and Rome, edited by Michael Gagarin. Oxford: Oxford University Press.

Lewandowsky, S., Ecker, U. K., & Cook, J. (2012). *Misinformation and Its Correction: Continued Influence and Successful Debiasing.* Psychological Science in the Public Interest, 13(3), 106-131.

Libicki, M. C. (1995). *What Is Information Warfare?* National Defense University Press.

Mackintosh, Eliza, and Edward Kiernan. 2019, *Finland is Winning the War on Fake News*

Martin, J. (2017). *Post-Truth and Fake News: Viral Modernity and the "Epistemic Crisis."* Journal of Media Ethics, 32(2), 79-91.

Mihailidis, P., & Viotty, S. (2017). *Spreadable Spectacle in Digital Culture: Civic Expression, Fake News, and the Role of Media Literacies.* American Behavioral Scientist, 61(4), 441-454.

Miller, C. (2022). *Active Measures: The Secret History of Disinformation and Political Warfare.* St. Martin's Press.

Mudde, C. (2004). *The Populist Zeitgeist.*Government and Opposition, 39(4), 541-563.

Mueller, Robert S. 2019, Report on the Investigation into Russian Interference in the 2016 Presidential Election. Washington, D.C.: U.S. Department of Justice.

Pacepa, I. M., & Rychlak, R. J. (2013). *Disinformation: Former Spy Chief Reveals Secret Strategies for Undermining Freedom, Attacking Religion, and Promoting Terrorism.* WND Books.

Pariser, E. (2011). *The Filter Bubble: How the New Personalized Web Is Changing What We Read and How We Think.* Penguin Books.

Pennycook, Gordon, Tyrone D. Cannon, and David G. Rand. 2018, Prior Exposure Increases Perceived Accuracy of Fake News, Journal of Experimental Psychology: General 147 (12): 1865-1880.

Pettegree, Andrew. 2005, *Reformation and the Culture of Persuasion*, Cambridge, Cambridge University Press. Pew Research Center. 2014, *Political Polarization in the American Public*,

Pomerantsev, P. (2019). *This Is Not Propaganda: Adventures in the War Against Reality.* PublicAffairs.

Rid, T. (2020). *Active Measures: The Secret History of Disinformation and Political Warfare.* Farrar, Straus and Giroux.

Robb, Amanda. 2017, *Anatomy of a Fake News Scandal,* Rolling Stone, November 16, 2017. Roozenbeek, Jon, Claudia R. Schneider, Sarah Dryhurst, John Kerr, Alexandra L. J. Freeman,

Sunstein, C. R. (2017). *Republic: Divided Democracy in the Age of Social Media.* Princeton University Press.

Vosoughi, S., Roy, D., & Aral, S. (2018). *The Spread of True and False News Online.* Science, 359(6380), 1146-1151.

Vosoughi, Soroush, Deb Roy, and Sinan Aral. 2018, *The Spread of True and False News Online*, Science 359 (6380): 1146-1151.

Wardle, C., & Derakhshan, H. (2017). *Information Disorder: Toward an Interdisciplinary Framework for Research and Policy Making.* Council of Europe Report.

What It's Learned May Be Crucial to Western Democracy." CNN, May 5, 2019.

Wodak, R. (2015). *The Politics of Fear: What Right-Wing Populist Discourses Mean.* SAGE Publications.

Zanker, Paul. 1988, *The Power of Images in the Age of Augustus. Ann Arbor*, University of Michigan Press.

Zelenkauskaitė, A. (2022). *Disinformation in the Digital Age: Social Media and Algorithmic Amplification.* Routledge.