



**CCDCOE**  
NATO COOPERATIVE  
CYBER DEFENCE  
CENTRE OF EXCELLENCE

---

# Ontological Foundations of Cognitive Warfare

Dr Fedir Korobeynikov, Dr Andrii Davydiuk, Prof Volodymyr Mokhor

NATO Cooperative Cyber Defence Centre of Excellence  
G.E. Pukhov Institute for Modelling in Energy Engineering National Academy of Sciences of Ukraine

---

## About the authors

Fedir Korobeynikov is a PhD in Information Security, director of digital technologies and information security at System Capital Management, founder of the Security Studies and Research Center (Kyiv, Ukraine), and doctoral student (D.Sc. program) at the G.E. Pukhov Institute for Modelling in Energy Engineering, National Academy of Sciences of Ukraine

Andrii Davydiuk is a PhD in Cybersecurity, a P&C branch head at NATO CCDCOE, deputy branch head in the State Cyber Protection Centre State Service of Special Communications and Information Protection of Ukraine, senior scientific research staff, doctoral student (D.Sc. program) at the G.E. Pukhov Institute for Modelling in Energy Engineering National Academy of Sciences of Ukraine.

Volodymyr Mokhor is a DrSc, Professor in Math Modelling, CyberSecurity and Risk Science, corresponding member National Academy of Sciences (NAS) of Ukraine, deputy chair of the Division of the Energy and Energy Technologies NAS of Ukraine, Director the G.E. Pukhov Institute for Modelling in Energy Engineering NAS of Ukraine

## CCDCOE

The NATO Cooperative Cyber Defence Centre of Excellence (CCDCOE) is a NATO-accredited knowledge hub offering a unique interdisciplinary approach to the most relevant issues in cyber defence. The heart of the CCDCOE is a diverse group of international experts from military, government, academia, and industry, currently representing 39 nations.

The CCDCOE maintains its position as an internationally recognised cyber defence hub, a premier source of subject-matter expertise and a fundamental resource in the strategic, legal, operational, and technical aspects of cyber defence. The Centre offers thought leadership on the cutting edge of all aspects of cyber defence and provides a 360-degree view of the sector. The Centre encourages and supports the process of mainstreaming cybersecurity into NATO and national governance and capability, within its closely connected focus areas of technology, strategy, operations, and law.

The Tallinn Manual, prepared at the invitation of the CCDCOE, is the most comprehensive guide for policy advisers and legal experts on how international law applies to cyber operations carried out between and against states and non-state actors. Since 2010, the Centre has organised Locked Shields, the biggest and most complex technical live-fire cyber defence challenge in the world. Each year, Locked Shields gives cybersecurity experts the opportunity to enhance their skills in defending national IT-systems and critical infrastructure under real-time attacks. The focus is on realistic scenarios, cutting-edge technologies, and simulating the entire complexity of a massive cyber incident, including strategic decision-making and legal and communication aspects.

The CCDCOE hosts the International Conference on Cyber Conflict, CyCon, a unique annual event in Tallinn, bringing together key experts and decision makers from the global cyber defence community. The conference, which has taken place in Tallinn since 2009, attracts more than 600 participants each spring.

The CCDCOE is responsible for identifying and coordinating education and training solutions in the field of cyber defence operations for all NATO bodies across the Alliance. NATO-accredited centres of excellence are not part of the NATO Command Structure.

[www.ccdcoe.org](http://www.ccdcoe.org)

[publications@ccdcoe.org](mailto:publications@ccdcoe.org)

## Disclaimer

This publication is a product of the CCDCOE. It does not necessarily reflect the policy or the opinion of the Centre or NATO. The Centre may not be held responsible for any loss or harm arising from the use

of information contained in this publication and is not responsible for the content of the external sources, including external websites referenced in this publication.

Digital or hard copies of this publication may be produced for internal use within NATO and for personal or educational use when for non-profit and non-commercial purpose, provided that copies bear a full citation.

# 1. Table of Contents

|      |   |    |
|------|---|----|
| 2.   | Abstract.....   | 5  |
| 3.   | Introduction .....  | 6  |
| 4.   | Conceptual landscape: approaches to understanding cognitive warfare .....                         | 8  |
| 4.1. | Effect-Centric Approaches .....   | 8  |
| 4.2. | Actor-Centric Approaches .....  | 8  |
| 4.3. | Technology-Centric Approaches.....  | 9  |
| 4.4. | Towards a Structure-Centric Approach.....   | 10 |
| 5.   | Theoretical framework: invariants and the multiplex architecture of cognitive vulnerability ..... | 11 |
| 5.1. | Invariants as Objects of Cognitive Warfare.....   | 11 |
| 5.2. | The Multiplex Architecture of Cognitive Vulnerability.....  | 12 |
| 5.3. | Cognitive Decoherence as a Systemic Effect .....  | 13 |
| 5.4. | Invisibility as an Ontological Property of Cognitive Warfare .....                                | 14 |
| 6.   | Discussion.....   | 17 |
| 7.   | Conclusion .....  | 18 |
| 8.   | Bibliography .....  | 19 |

## 2. Abstract

This study examines the ontological foundations of cognitive warfare and introduces a structure-centric conceptualisation that shifts the analytical focus from observable effects, actors, and technological instruments to the systemic conditions of cognitive vulnerability. Central to this approach is the concept of systemic invariants – epistemic, axiological, identificatory, social, and teleological structures that sustain the coherence, identity, and adaptive capacity of complex socio-technical systems.

It is argued that these systemic invariants constitute the ontological scaffolding of cognitive architecture – securing the connectivity of its components, defining the boundary conditions of adaptive transformation, and enabling coherent meaning-making and strategic self-determination under uncertainty. Cognitive warfare is reconceptualised as the deliberate targeting of systemic invariants through the exploitation of inter-layer linkages within a system's multiplex architecture, with the strategic aim of inducing cognitive decoherence.

Within this framework, decoherence is understood as a structurally conditioned and potentially irreversible erosion of cognitive sovereignty, whereby a system loses its capacity for coherent perception and analysis of reality, development, adaptation, and self-protection as an integrated cognitive order. This multiplex perspective enables the differentiation of vulnerability logics according to the type of socio-technical system, which carries practical implications for the development of tailored cognitive protection strategies. By distinguishing cognitive warfare from information warfare at the ontological level, the proposed framework establishes a foundation for diagnosing systemic vulnerabilities and advancing proactive strategies for cognitive resilience.

**Keywords:** cognitive warfare, cognitive decoherence, systems invariants, multiplex networks, structural vulnerability, cognitive resilience, transmorphance, socio-technical systems, hybrid threats, information security.

### Acknowledgements

The authors would like to express their gratitude to the Estonian Military Academy and the NATO Cooperative Cyber Defence Centre of Excellence (CCDCOE) for their invaluable support throughout the development of this research.

### 3. Introduction

The concept of cognitive warfare has emerged in response to an observable shift in the nature of interstate confrontation. Traditional forms of conflict, including kinetic, economic, diplomatic, and informational factors are increasingly being supplemented by systematic interventions targeting the cognitive processes of adversary populations. Whereas kinetic warfare destroys material infrastructure, economic warfare undermines resource bases and reproductive capacity, diplomatic warfare destabilises alliance systems and international legitimacy, and information warfare distorts content and disrupts data availability, cognitive warfare targets the very mechanisms of interpretation, specifically targeting not what people know, but how they construct knowledge, perceive reality, and locate themselves within it.

In 2021, the NATO Science and Technology Organization identified the cognitive domain as a key area of strategic research [1], an initiative that by 2022 had expanded significantly [2]. This very act of institutional recognition signals that the phenomenon has transcended the bounds of tactical instrumentation and demands conceptual treatment at the level of strategic theory.

Yet existing conceptualisations of cognitive warfare [3–9] exhibit a marked asymmetry, as they describe effects in considerable detail, ranging from societal polarisation, erosion of institutional trust, narrative manipulation, and disruption of decision-making cycles, yet while offering no genuinely systemic model capable of explaining the mechanisms that generate these outcomes. The NATO ACT definition characterises cognitive warfare as 'activities conducted in synchronization with other Instruments of Power, to affect attitudes and behaviours by influencing, protecting, and/or disrupting individual and group cognitions to gain an advantage' [3]. This definition captures the objective, but it leaves open the question of what, precisely, within the structure of a socio-technical system renders it vulnerable to such influence, and through what pathways that influence translates into systemic effects.

Despite growing scholarly interest in expanding the theoretical field of cognitive warfare as developed within NATO, existing approaches have yet to yield a coherent model of the structural vulnerability of socio-technical systems to cognitive attack. A key obstacle lies in methodological limitations, articulated with particular clarity by Deppe and Schaal [10]. Analysing the NATO ACT concept, the authors point to the problem of 'conceptual stretching', the blurred boundaries with adjacent concepts (such as hybrid threats, Foreign Information Manipulation and Interference [FIMI], and information warfare), and the difficulty of operationalising the term in analytically rigorous and empirically tractable ways.

As an attempt to systematise this field, the NATO Human Factors and Medicine Panel Exploratory Team (HFM-356) proposed the so-called 'House Model' [11]. It encompasses seven domains of knowledge relevant to understanding cognitive warfare, ranging from cognitive neuroscience to socio-cultural studies. Nevertheless, this taxonomy remains predominantly descriptive in character. While charting the disciplinary landscape and delineating significant domains, the model fails to capture the dynamics of the process itself, in particular regarding how precisely cognitive influences propagate within a system to generate observable effects.

In contrast to existing conceptual frameworks, which focus primarily on the descriptive level by asking 'what happens under conditions of cognitive warfare'. The present study shifts attention to the analytical-ontological dimension. The aim of this research is to identify the structural foundations that render such influences effective, and to characterise the elements of socio-technical systems upon which they are brought to bear. The central argument advanced here is that cognitive warfare constitutes a deliberate targeting of the invariants of a socio-technical system (systemic invariants). Those stable structural elements that sustain the system's coherence and its capacity for adaptive response to disruptive influences. The concept of invariance is here adapted from the theory of security in complex evolving systems [37] and transposed to the context of cognitive warfare. A systemic invariant is understood as a property of a socio-technical system that is preserved across transformations of structure, function, and

environment, thereby ensuring continuity of identity, the reproduction of core processes, and the capacity for coherent adaptive response to perturbations.

In the context of cognitive security, invariants include epistemic structures (shared criteria of truth and modes of verification), axiological hierarchies (value priorities governing collective choice), identity constructs (conceptions of group belonging and the boundaries of the in-group), social trust architectures (models of institutional legitimacy and interpersonal cooperation), and, finally, teleological projections – visions of the future through which the system conceives its own development. These encompass normative orientations, strategic expectations, and ontological assumptions regarding admissible and inadmissible scenarios; taken together, they define both the direction of transformation and the boundaries of the possible.

To operationalise the approach proposed in this study, the concept of a multiplex architecture of cognitive vulnerability is introduced. Socio-technical systems, whether they are states, communities, organisations, or transnational corporations, are described as multiplex networks [12–13], in which the same actors (nodes) are connected by different types of relations corresponding to distinct layers of the network. Each layer reflects a particular type of invariant linkage, epistemic, axiological, identificatory, social, or teleological, and constitutes a topologically distinct, yet functionally coupled, circuit of interactions.

Within this model, it becomes possible to analyse how a perturbation initiated in one layer, for instance, within the epistemic layer through the discrediting of knowledge sources, or within the identity layer through the fragmentation of conceptions of group belonging, may propagate into adjacent layers, producing cascading and nonlinear effects. Structural vulnerabilities are thereby localised in nodes exhibiting high interlayer connectivity, points within the multiplex structure through which perturbations are most likely to propagate across the system as a whole. The result is cognitive decoherence, a state characterised by a loss of coherence among invariants, whereby the system forfeits its capacity for integrated self-description and coordinated response.

The proposed approach addresses three interrelated tasks. First, it offers a theoretical explanation of the efficacy of cognitive warfare, not by cataloguing tactics, but by identifying the systemic conditions under which those tactics become effective. Second, it provides an analytical toolkit for identifying vulnerabilities prior to their exploitation, shifting the focus from reactive response to the proactive strengthening of resilience and transmorphance [37] in socio-technical systems. Third, it establishes a conceptual bridge between military and academic discourses on cognitive warfare, contributing to a shared framework that can traverse the divide between strategic reasoning and scientific analysis.

## 4. Conceptual landscape: approaches to understanding cognitive warfare

### 4.1. Effect-Centric Approaches

The most prevalent mode of conceptualising cognitive warfare is to define it by reference to observable effects. NATO doctrinal documents [3] identify five categories of anticipated outcomes, which include the disruption of the OODA (Observe – Orient – Decide – Act) cycle and decision-making processes [14], social polarisation and fragmentation, the instrumentalisation of identity, narrative manipulation, and the undermining of the will to resist. The 2025 NATO STO Chief Scientist Research Report [15] frames cognitive warfare as a contest for cognitive superiority and delineates three functional vectors that include the degradation of an adversary's cognitive capabilities, the enhancement of one's own cognitive and technological capacities, and the cultivation of resilience to external influence.

Admittedly, an effect-centric approach offers considerable practical utility, providing a solid foundation for the cataloguing of threats, the construction of monitoring architectures, and the development of early-warning indicators. Yet its explanatory potential remains limited, for the mere registration of an effect is not tantamount to an understanding of the mechanism by which it arises. The claim that cognitive warfare leads to societal polarisation, for instance, leaves unresolved a critical question, which underscores why some socio-technical systems display vulnerability to such interventions, whilst others retain their stability? What structural properties determine a system's susceptibility to particular modes of influence?

Critics of this approach [10] point to an epistemological problem often described as 'conceptual stretching'. When a phenomenon is defined exclusively by its outcomes, the boundaries of the concept tend to become diffuse, insofar as a broad range of informational interventions that alter recipients' attitudes may be subsumed under the rubric of cognitive warfare. This impedes rigorous scientific operationalisation and complicates demarcation from adjacent constructs, including hybrid threats, information warfare, and FIMI.

Thus, notwithstanding its effectiveness for operational diagnostics, the effect-centric lens remains insufficient for elucidating causal mechanisms. It requires supplementation by a more structured analytical framework capable of explaining the internal dynamics of influence as a systemic process.

### 4.2. Actor-Centric Approaches

A second mode of conceptualisation shifts the focus from effects to the actors behind influence, state and non-state agents, their doctrines, motivations, and strategic cultures. Within this perspective, cognitive warfare is treated as one instrument within a broader repertoire of geopolitical contestation, embedded in historically specific traditions of strategic thought.

A paradigmatic example is the Russian concept of reflexive control [16–17]. Reflexive control entails the transmission of specially crafted information to an adversary so as to induce decisions advantageous to the initiator of influence. Its central premise is that the target retains a subjective sense of autonomous choice, while remaining unaware that the space of available alternatives has been pre-structured. Splidsboel Hansen [18] traces the evolution of this concept from Soviet military-theoretical developments to contemporary practices of information confrontation.

The Chinese approach, conceptualised as the 'Three Warfares'; psychological, public opinion, and legal warfare are analysed by Lee [19] and Aukia [20]. In contrast to the Russian model, which emphasises manipulation of the decision-making process, the Chinese doctrine is oriented toward shaping a favourable strategic environment through long-term influence on perception, legitimacy, and normative frameworks.

This logic traces back to the strategic thought of Sun Tzu, for whom 'what is of supreme importance in war is to attack the enemy's strategy' [21]. Here, the cognitive dimension is integrated into a broader strategy aimed at achieving objectives without direct confrontation.

A cognate approach is found in Basil Liddell Hart's strategy of indirect action [22]. Its essence lies in undermining the adversary's capacity for strategic initiative not through frontal assault, but by shaping the adversary's perception of the environment, the logic of response, and the range of available choices.

Particular attention is warranted by research into, and the exploitation of, cognitive biases and suggestibility in the practice of cognitive influence. The work of Tversky and Kahneman [23] laid the foundation for understanding systematic deviations of human thought from rational models; Cialdini [24] systematised principles of persuasion, such as social proof, authority, scarcity, reciprocity, and others, that are exploited in practices of mass influence. The classical 'doctrine of suggestion' [25], originating in early twentieth-century research, is experiencing a renaissance in the context of algorithmically mediated communication [26]. Within the framework of cognitive warfare, cognitive biases and suggestibility are treated as 'entry points' for manipulative influence.

Undeniably, actor-centric approaches make a substantial contribution to our understanding of cognitive warfare, while highlighting that what is at stake is neither a spontaneous social reflex nor a process dictated by technology, but rather a repertoire of purposive strategies, embedded in distinct traditions of thinking about conflict. Yet their limitation is, in a sense, the reverse of that encountered in effect-centric accounts. Whereas the latter register what happens without explaining why, actor-centric approaches elucidate who acts, and with what intentions, yet remain silent on the more fundamental question of what general principles of cognitive influence underwrite its efficacy, irrespective of the particular doctrine or strategic culture in which it is instantiated?

Why do the same principles of cognitive warfare, such as indirect influence, the manipulation of interpretation, the structuring of the decision space, the undermining of trust in sources, the erosion of teleological orientations, and so forth, prove decisive in some socio-technical contexts, yet fail to generate comparable effects in others? What properties of a target system determine its susceptibility to cognitive influence per se, rather than merely to its specific doctrinal expression? The absence of satisfactory answers to these questions marks the methodological limit of actor-centric analysis, and underscores the necessity of moving towards a structural-systemic account of cognitive vulnerability.

### 4.3. Technology-Centric Approaches

A third mode of conceptualisation attends to the instrumental dimension of cognitive warfare, specifically the technologies, platforms, and vectors through which influence is exerted. Within this perspective, cognitive warfare appears chiefly as a function of technological affordances, driven by social networks and algorithmic amplification; big data and microtargeting; synthetic media and deepfakes; and, prospectively, neurotechnologies and brain-computer interfaces.

The 'House Model' resonates, in part, with this logic, identifying 'technology enablers' as one of seven domains of knowledge relevant to understanding cognitive warfare. NATO researchers have sought to systematise technological vectors of influence by distinguishing three broad categories, categorised into traditional (kinetic means, mass media), existing technologies (social networks, big data, augmented reality), and emergent technologies (synthetic media, generative AI, metaverses, and prospective neuro-technical means) [11]. Such a classification reflects the dynamics of technological development and facilitates the anticipation of an expanding arsenal of cognitive influence.

A distinct strand of technology-centric literature is devoted to the study of so-called 'echo chambers' and 'filter bubbles'. Sunstein and Pariser [27–29] advanced the influential thesis that personalisation algorithms may sequester users within ideologically homogeneous environments, thereby amplifying polarisation and

eroding the conditions of democratic discourse. Systematic reviews of the empirical literature, however, disclose a considerably more complex, and, indeed, contested, picture, noting that most users encounter heterogeneous content, and algorithmic filtering appears to account for only a limited share of observed polarisation [30–32]. The implication is a fundamental limitation of the technology-centric approach, as it suggests that technology operates, more often than not, as a modulator rather than a generator of dynamics whose preconditions reside in the structural properties of social systems themselves.

In recent years, synthetic media and deepfake technologies have developed with particular rapidity. Vaccari and Chadwick [33] show that deepfakes affect cognitive processes not merely through deception, but also through the production of uncertainty, insofar as even an unsuccessful deepfake may heighten generalised distrust towards media content. This phenomenon, commonly termed the 'liar's dividend', enables actors to discredit authentic recordings by casting them as probable fabrications [34]. Chinese military theorists have conceptualised adjacent capabilities in terms of 'algorithmic cognitive warfare', involving the use of algorithms to profile target audiences and to optimise the timing, targeting, and delivery of influence content [35–36].

The technology-centric approach possesses evident strengths, specifically by facilitating the tracking of the evolution of the instrumental repertoire, the development of technical countermeasures, and the assessment of risks associated with the emergence of new technologies. Yet its explanatory potential is constrained by a characteristic reductionism. Technology is here treated as an independent variable determining the character and scale of influence. Crucially, however, the selfsame technological means, such as recommendation algorithms, viral content, targeted advertising, produce divergent effects across different systems. Social networks do not polarise societies in and of themselves; rather, they amplify and accelerate processes whose preconditions reside in the structural properties of those very societies.

The technology-centric approach, then, answers the question of 'how technically' whilst bypassing the question of 'why systemically'. It describes the means of delivery, but fails to explain what, precisely, in the architecture of the target system renders it susceptible to the influence being delivered. Technology emerges as a modulator, essentially acting as an amplifier or accelerator of processes whose generative mechanisms lie beyond the purview of technological analysis.

#### 4.4. Towards a Structure-Centric Approach

The foregoing review discloses a systematic lacuna in prevailing conceptualisations of cognitive warfare. Effect-centric approaches register observable consequences, yet do not account for their genesis. Actor-centric approaches illuminate the intentions and doctrines of the agents of influence, yet leave unresolved the question of the systemic conditions under which such influence proves efficacious. Technology-centric approaches map the instrumental repertoire, yet exhibit a tendency to reduce complex social dynamics to technological determinants.

Common to all three perspectives is the absence of an answer to a key question of what, precisely, in the structure of a socio-technical system renders it vulnerable to cognitive influence? Addressing this question necessitates a shift in analytical focus, moving from describing effects, actors, and technologies to analysing the structural foundations of cognitive vulnerability.

The present study introduces a structure-centric approach to the analysis of cognitive warfare. Its foundational premise is that the efficacy of cognitive influence is determined not solely by the characteristics of the intervention itself, but also by the architecture of the target system, specifically regarding the configuration of its invariants, the topology of its inter-layer linkages, and the presence of latent vulnerabilities through which cascading disruption of cognitive integrity becomes possible.

A detailed explication of this approach constitutes the substance of the following section.

# 5. Theoretical framework: invariants and the multiplex architecture of cognitive vulnerability

## 5.1. Invariants as Objects of Cognitive Warfare

Cognitive warfare is waged by, and among, complex socio-technical systems, encompassing states and international organisations, financial groups and industrial corporations; legal and religious institutions; national collectivities and supranational communities. It is, in this sense, most productively understood as a mode of inter-systemic strategic contestation, rather than as a conflict between states or organisations in any narrowly institutional sense, insofar as influence targets systemic invariants rather than institutional forms [37]. Such systems are autopoietic [38–39] and dynamically evolving, insofar as they are grounded in principles, laws, norms, and values, as well as in symbolically articulated images of the future, which represent higher-order structuring commitments that delineate the ultimate contours of systemic identity. It is precisely these higher-order commitments that function as invariants, serving as relatively stable foundations that sustain the unity, integrity, and reproducibility of the complex system.

Invariants are not merely stable parameters; they are the constitutive grounds of systemic existence including epistemological assumptions, axiological hierarchies, normative expectations, symbolic constructions of identity, and teleological vectors. Their function is twofold. On the one hand, they constitute an integrative matrix that binds heterogeneous components into a coherent whole; on the other, they configure the system's modes of external coupling with other systems, thereby shaping the parameters of cooperation, conflict, or strategic disengagement. Invariants form a semantic infrastructure, providing an internal order through which a system reproduces itself across historical and structural change. Their disruption signifies not merely a loss of function, but an erosion of ontological continuity involving the forfeiture of the system's capacity to remain itself through adaptive transformation.

An analysis of the aforementioned NATO strategic and doctrinal documents, as well as the academic literature, permits the identification of a direct correspondence between the declared objectives of cognitive warfare and the category of invariants. Within the NATO ACT conceptualisation, as summarised in the scholarly literature, five key vectors of cognitive influence are commonly distinguished, comprising disruption of the OODA cycle, polarisation and fragmentation of society, the weaponisation of identity, the weaponisation of narratives, and undermining the will to fight [10]. Each of these vectors, when translated into the language of systems analysis, constitutes an attack on a particular class of invariants, targeting epistemic structures of verification, social architectures of trust, identity constructs, narrative matrices of historical memory, and teleological projections of the future.

A key conceptual distinction between cognitive and information warfare concerns the locus of effect, in that cognitive warfare targets not merely the acceptance or rejection of particular informational content, but the manipulation of 'emotional and subconscious processes of the human mind' [10]. Backes and Swab [40] define cognitive warfare as a strategy that focuses on altering how a target population thinks, and through that, how it acts. Du Cluzel [41] observes that cognitive warfare is methodically exploited as a component of a global strategy aimed at weakening, interfering with and destabilising targeted populations, institutions, and states in order to influence their choices. Across these formulations, a common pattern emerges, wherein the object of influence is not individual cognitive acts or information flows, but the stable structures that determine the mode of processing any information, which the present study conceptualises as invariants.

The objectives of cognitive warfare, specifically destabilisation and influence, are realised precisely through the degradation of invariant structure. NATO Innovation Hub documents [41] emphasise that the aim is 'not

to attack what individuals think but rather the way they think', and that such influence 'has the potential to unravel the entire social contract that underpins societies'. A condition of systemic disintegration of this kind cannot be produced by the mere distortion or concealment of data (as in information warfare); it requires the subversion of the very foundations upon which cognitive coherence is constructed, achieved by targeting shared criteria of truth, normative hierarchies, identificatory models, architectures of trust, and images of a possible future.

Thus, the concept of invariants is not exogenous to the discourse on cognitive warfare; it performs the function of an epistemological explication of what is already present there in nuce. Its principal advantage lies in its capacity to translate a multiplicity of fragmentary descriptions and effects into a unified analytical framework, enabling the systematic identification of objects of influence, the assessment of their interdependence, and the anticipation of cascading mechanisms of destabilisation.

## 5.2. The Multiplex Architecture of Cognitive Vulnerability

The interconnectedness of invariants, established in the preceding section, requires an adequate analytical apparatus. Classical network models, confined to single-layer representations, are ill-suited to capturing the specificity of cognitive systems in which the same actors, including individuals, groups, and institutions, are simultaneously bound by a multiplicity of qualitatively distinct relations spanning epistemic, axiological, identificatory, and institutional. To model such structures, the present study employs the apparatus of multiplex networks [12–13], adapted for the analysis of social interactions [42] and cascading vulnerabilities in interdependent systems [43].

A multiplex network is a system in which a single set of nodes is connected by several types of edges, thereby forming topologically distinct yet functionally coupled layers [44]. Each layer reflects a particular type of invariant linkage including epistemic, axiological, identificatory, social, and teleological. The key property of multiplex networks is inter-layer connectivity, defined as structural dependencies between layers by virtue of which cognitive influence exerted upon one layer may cascade into others, producing systemic effects irreducible to any merely local perturbation.

The multiplex approach permits, for example, the formalisation of a fundamental distinction between cognitive warfare strategies directed against different types of political system.

This distinction can be further clarified by analogy with the typology of socio-technical systems as functionally stable versus evolutionarily dynamic, as formalised within a multiplex-network perspective [45]. In this reading, highly centralised regimes tend to approximate functionally stable architectures (low multiplexity and strong vertical coupling), whereas pluralistic democracies exhibit higher irreducible multiplexity and a correspondingly different vulnerability logic.

Cognitive warfare against highly centralised regimes, including autocracies, authoritarian regimes, and other closed political systems, is directed towards eroding the monopoly over the interpretation of reality. Such systems are characterised by a low degree of multiplexity as their cognitive architecture collapses into a single dominant layer, controlled by a vertical structure of authority. An effective strategy of influence consists in the deliberate injection of high informational entropy into this layer – the introduction of uncertainty that cannot be resolved without the regime compromising its own claim to epistemic authority. Conspiracy theories, ambiguous narratives, and ambivalent claims concerning 'deep states' or 'hidden centres of decision-making' generate alternative explanatory frameworks that compete with the official position. The regime can neither confirm nor refute such constructions without conceding the incompleteness of its own control. As a result, the authoritative narrative fragments, and the system loses its capacity for effective cognitive compression, namely the reduction of complexity.

Cognitive warfare against democracies and pluralistic systems is predicated upon an inverse logic. Democratic systems are characterised by a high degree of irreducible multiplexity in which their resilience

is secured not by any monopoly over narrative, but by coherence across multiple layers - axiological consensus, institutional trust, shared procedures of verification, and a common image of the future. An effective strategy of cognitive influence in this context is directed not towards increasing entropy, but towards the disruption of inter-layer coherence. The polarisation of society, the erosion of trust in institutions, the fragmentation of identity, and the discrediting of epistemic authorities, which all widely noted within NATO's conceptual discussion of cognitive warfare, constitute attacks upon the linkages between layers of the multiplex. The system then fragments into incoherent substructures, losing its capacity for a coordinated response whilst retaining, *mutatis mutandis*, the local functioning of individual components.

A particular threat to democratic systems is posed by the implantation of parasitic substructures, which are closed, functionally stable formations embedded within the multilayered architecture. Corrupt vertical networks, shadow influence structures, and foreign-controlled proxy networks may operate as 'systems within the system', reproducing themselves through a single type of linkage whilst bypassing legitimate procedures and normative layers. Such structures can be cultivated by external actors as instruments of long-term cognitive influence. Their implantation leads to the progressive degradation of adaptive capacity where the democratic multiplex retains its outward complexity, yet forfeits its functional irreducibility, as critical decisions become increasingly determined by concealed single-layer circuits.

The operational diagnosis of vulnerability within a multiplex architecture rests upon an analysis of inter-layer topology. Of particular importance are those layers whose destabilisation yields the greatest reduction in the system's integral connectivity. Nodes that participate intensively in inter-layer linkages, such as opinion leaders, institutions of epistemic authority, and mediators of inter-group communication, function as structural connectors. Their delegitimisation, capture, or neutralisation ruptures cognitive connectivity and may initiate cascading decoherence representing a transition from functional complexity to fragmented incoherence.

The multiplex model thus furnishes an operational framework for analysing cognitive vulnerability, one capable of differentiating influence strategies according to system type, identifying critical nodes and linkages, and anticipating scenarios of cascading destabilisation. It provides a foundation for the development of adaptive defence strategies oriented not towards the mere preservation of structure, but towards the maintenance of cognitive coherence and the resilience of the system's invariant organisation.

### 5.3. Cognitive Decoherence as a Systemic Effect

Building on the multiplex architecture introduced in the preceding section, cognitive decoherence can be specified as the systemic outcome of effective cross-layer disruption. In highly centralised regimes, decoherence is typically induced by destabilising the dominant interpretative layer through informational entropy, thereby fracturing the regime's epistemic monopoly. In pluralistic democracies, by contrast, decoherence is more often produced by eroding inter-layer coherence, effectively undermining the connective tissue between epistemic verification, institutional trust, identity integration, and teleological orientation. In both cases, the mechanism is mediated by inter-layer dependencies because perturbations propagate through connector nodes and translate local disruption into system-wide loss of cognitive integrity.

The concept of cognitive decoherence denotes a specific systemic effect arising from the exertion of cognitive influence upon a multiplex architecture. Decoherence consists in a loss of coherence among a system's invariants, whereby the layers of the multiplex may continue to function locally, yet cease to constitute a coherent whole. The system retains its components such as beliefs, values, identities, and institutions, but forfeits their integration so that they no longer coalesce into a unified semantic order capable of orienting coordinated action.

The mechanism of decoherence is cascading in character. A primary perturbation initiated in one layer – for instance, the discrediting of a key epistemic authority, or the introduction of a competing identity narrative

– is transmitted, via nodes of inter-layer connectivity, into adjacent layers. The erosion of trust in sources of knowledge (the epistemic layer) generates uncertainty in the evaluation of alternatives (the axiological layer), which, in turn, undermines the grounds of collective choice and coordination (the social layer) and progressively dissolves shared representations of the desired future (the teleological layer). Each transition intensifies destabilisation, since the disruption of one invariant deprives others of their semantic grounding. The non-linearity of this process entails that a relatively weak initial perturbation may, given a particular topology of linkages, produce a disproportionately large-scale effect, which is a phenomenon known in complex systems theory as a ‘critical transition’ [46].

Phenomenologically, cognitive decoherence manifests through a range of characteristic symptoms. At the level of decision-making, it appears as paralysis, or as chaotic oscillation between incompatible alternatives, owing to the absence of a shared evaluative framework. At the level of identity, it appears as fragmentation into mutually hostile groups, each claiming to embody the ‘true’ identity of the system. At the level of temporal orientation, it appears as the loss of a common horizon of the future manifesting as the disintegration of a unified teleological vector into a multiplicity of incompatible projections. The system becomes incapable of answering the questions ‘Who are we?’, ‘What is true?’, and ‘What do we strive for?’, not because answers are absent, but because there are too many, and they are mutually exclusive.

Cognitive decoherence constitutes the antithesis of resilience and transmorphance, which are two key properties of an adaptive system. Resilience denotes the capacity of a system to absorb perturbations whilst preserving its invariant structure. Transmorphance [37] denotes the capacity for profound structural transformation without forfeiture of identity. Both capacities presuppose the coherence of invariants given that a system can adapt only insofar as it retains the internal coherence that permits it to distinguish adaptive change from disintegration. Decoherence thus represents a form of irreversible attenuation of a system’s cognitive sovereignty, understood as its capacity for strategic self-determination, development, adaptation, and self-defence. In a state of decoherence, the system is not physically destroyed, but is deprived of the capacity for meaningful self-governance; lacking a coherent vector of its own, it becomes susceptible to any external influence capable of temporarily structuring its chaos.

It is precisely the induction of cognitive decoherence that constitutes the strategic objective of the adversary in cognitive warfare. The condition in which a target system destroys itself from within, forfeiting its capacity for resistance, is not a metaphor, but a description of a specific systemic effect towards which cognitive influence is directed [41]. NATO ACT-related discussions register this threat and emphasise the necessity of cultivating cognitive resilience as a strategic priority of defence. The theoretical framework proposed in the present study permits the operationalisation of this task where defence against cognitive warfare is not the protection of particular beliefs or narratives, but the maintenance of the coherence of the invariant structure that underwrites the system’s capacity for meaningful self-governance and adaptive transformation. To illustrate the practical application of this framework, see Supplementary Information [48] (Examples A and B) for two scenario-based worked examples illustrating pathways from invariant disruption to cognitive decoherence across pluralistic democratic and highly centralised authoritarian systems.

## 5.4. Invisibility as an Ontological Property of Cognitive Warfare

Building on the foregoing analysis, we argue that cognitive warfare exhibits a fundamental ontological distinction from other forms of confrontation defined by its constitutive invisibility. At present, NATO officially recognises five operational domains namely land, maritime, air, space, and cyberspace [47]. The cognitive domain is discussed as a potential sixth domain of operations [9], although its institutionalisation within official doctrine remains a matter of debate [10].

An analysis of the five recognised domains reveals a common feature in that an attack within any of these domains is, in principle, detectable. Kinetic action in the land, maritime, or air domains produces physical destruction that cannot be ignored. Space operations, though less visible to the general public, are

registered by specialised monitoring systems. Cyber-attacks may remain latent for a time (for instance, in the form of Advanced Persistent Threats), yet their consequences, ranging from system malfunction, data exfiltration, or the failure of physical infrastructure, eventually become manifest, rendering the fact of attack discernible.

Cognitive warfare, by contrast, represents a qualitatively distinct phenomenon. Its essential characteristic is a constitutive invisibility, namely not merely as a tactical advantage (concealment), but as an ontological property. The target of cognitive influence may never become aware that it has been attacked. Moreover, the very concept of 'attack' in the traditional sense loses much of its conceptual purchase here precisely because the target experiences the induced changes, such as doubt, the re-evaluation of values, the transformation of identity, and the loss of trust in institutions, as its own organic processes, as the outcome of the ostensibly 'natural' evolution of its views.

This ontological configuration engenders a distinct set of implications that differentiate cognitive warfare sharply from other modalities of confrontation. First, it entails the absence of a clear *casus belli*. In the conventional domains, however blurred in practice, there remains a discernible threshold between a state of peace and a state of war: bombardment, invasion, or a large-scale cyber-attack against critical infrastructure is registered as an act of aggression demanding response. Cognitive warfare rarely crosses a clear, event-like threshold that would be registered as an act of war; it unfolds in the grey zone below armed conflict, even where it may amount to coercive intervention in a State's *domaine réservé*. It is waged continuously in the grey zone below the threshold of armed conflict, and the target receives no unambiguous signal that 'war' has begun; accordingly, it cannot mobilise its defensive mechanisms in a timely fashion.

Secondly, cognitive warfare effects an inversion of agency. In kinetic conflict, the victim knows itself to be the object of external violence. In cognitive warfare, by contrast, the target experiences itself as the subject of its own decisions whereby it 'arrives' at conclusions induced from without, and 'forms' convictions whose architecture has been pre-structured by manipulation. This inversion, which involves the transformation of the object of influence into a putative subject, is precisely the mechanism that underwrites the invisibility of the attack.

Thirdly, cognitive warfare is distinguished by the latency of its effects. Kinetic influence produces an immediate outcome; a cyber-attack often yields effects delayed by hours or days. Cognitive influence may manifest after months or years. The erosion of trust in institutions, the fragmentation of identity, and the dissolution of axiological consensus are processes unfolding on generational timescales. This renders attribution, defined as the establishment of a causal link between intervention and effect, exceptionally difficult, and, as a rule, politically contested.

Finally, cognitive warfare lacks conventional markers of termination. Traditional wars end in surrender, armistice, or treaty. Cognitive warfare, as Du Cluzel observes, is 'potentially endless since there can be no peace treaty or surrender for this type of conflict' [40]. Once induced, a state of decoherence may reproduce endogenously since a system that has lost coherence among its invariants may continue to disintegrate by inertia, without any further external stimulus.

In the terms of the proposed multiplex model, the invisibility of cognitive warfare is explained by the fact that influence is directed not primarily at the nodes of the network (actors), but at inter-layer linkages, or more specifically, at the very architecture of meaning-making. The actor continues to function, retains a subjective sense of autonomy, and registers no intervention. What changes is not the content of consciousness (a particular thought), but the structure through which that content is organised into a coherent picture of the world.

This ontological distinction bears direct implications for defensive strategy. If the attack is invisible, the conventional sequence 'detection → attribution → response' becomes inapplicable. Defence against cognitive warfare therefore requires a shift from reactive threat-response to the proactive strengthening of

invariant coherence, which is precisely that which NATO's strategic discourse increasingly conceptualises as cognitive resilience.

## 6. Discussion

The conceptual framework proposed in the present study permits a re-examination of prevailing definitions of cognitive warfare and the articulation of an alternative formulation that integrates a structure-centric perspective.

The most widely cited definition of cognitive warfare is that offered by Claverie and Du Cluzel: “Cognitive warfare is thus an unconventional form of warfare that uses cyber tools to alter enemy cognitive processes, exploit mental biases or reflexive thinking, and provoke thought distortions, influence decision making and hinder action, with negative effects, both at the individual and collective levels” [7]. This definition captures a key distinction between cognitive and information warfare insofar as the cognitive effect is not a by-product of action, but its immediate objective.

NATO’s working definition foregrounds the instrumental dimension: “Cognitive Warfare includes activities conducted in synchronization with other Instruments of Power, to affect attitudes and behaviours, by influencing, protecting, or disrupting individual, group, or population level cognition, to gain an advantage over an adversary” [3]. Backes and Swab offer a more concise formulation: “Cognitive warfare is a strategy that focuses on altering how a target population thinks, thereby transforming how it acts” [40].

Notwithstanding their heuristic utility, these definitions share a common limitation since they foreground processes (cognitive distortions, decision-making, behaviour) without explicitly articulating the objects of attack. What, precisely, is targeted when “cognitive processes” are attacked? To this question, existing definitions fail to proffer a systematic answer.

The structure-centric approach developed in the present study permits an alternative definition.

*Cognitive warfare is the deliberate targeting of a socio-technical system’s invariants – the epistemological, axiological, identificatory, social, and teleological structures that underwrite coherent functioning (shared standards of truth, value hierarchies, identity boundaries, architectures of trust, and images of the future). It is conducted by exploiting inter-layer linkages within a system’s multiplex architecture, with the aim of inducing cognitive decoherence – a systemic loss of coherent self-understanding, coordinated action, and adaptive transformation. Its essential characteristic is constitutive invisibility meaning that the target experiences induced changes as endogenous and ‘organic’, rather than as the outcome of an external intervention.*

This conceptualisation affords several distinct analytical advantages. It explicates the object of attack (invariants as stable structural foundations), specifies the mechanism (cascading propagation across inter-layer linkages), and articulates the objective in systemic terms (cognitive decoherence). It also differentiates cognitive warfare from information warfare where the latter primarily contests the veracity of content, whilst the former targets the structures through which any content is interpreted. Finally, it supports the operational differentiation of influence strategies across distinct types of socio-technical system, by focusing analysis on topology, connector nodes, and cross-layer dependencies.

The limitations of the proposed approach principally reside in its level of abstraction. The multiplex model requires further operationalisation for empirical application, including indicators of inter-layer connectivity, methods for identifying connector nodes, and metrics of cognitive coherence. These tasks lie beyond the scope of the present study and delineate a clear agenda for future research.

## 7. Conclusion

The present study has articulated a structure-centric conceptual framework for cognitive warfare, shifting the analytical focus from effects, actors, and technologies to the structural foundations of cognitive vulnerability. At the centre of this analysis stand the invariants of socio-technical systems and the architecture of their interconnections, specifically the configuration that underwrites cognitive coherence, strategic identity, and the capacity for adaptive transformation.

The principal findings may be summarised as follows. First, cognitive warfare has been reconceptualised as the deliberate targeting of invariant structures, namely the epistemological, axiological, identificatory, social, and teleological foundations that constitute the ontological load-bearing framework of a system's cognitive architecture and condition its capacity for meaningful self-governance under uncertainty and external pressure.

Second, the proposed multiplex model of cognitive architecture enables the differentiation and operationalisation of influence strategies across rigidly centralised and adaptively multiplex systems. For the former, the key vector of attack lies in elevating informational entropy within the dominant layer; for the latter, it lies in the subversion of inter-layer coherence. This distinction carries not merely theoretical, but practical significance for the development of differentiated strategies of cognitive defence.

Third, the study has introduced the concept of cognitive decoherence as the systemic effect towards which the adversary's influence is directed. Decoherence is understood as the loss of a system's capacity for coherent self-description, strategic action, and adaptive reconfiguration, whilst the fragmentary functioning of individual components may remain intact. It signifies not destruction, but internal disintegration marking a systemic shift from governing to being governed.

Fourth, the study has identified an ontological distinction between cognitive warfare and other forms of confrontation owing to its constitutive invisibility and its capacity to simulate the organic self-development of the target system. Influence is effective precisely insofar as it remains unrecognised; its consequences are experienced as endogenous evolution. This renders the traditional triad of 'detection – attribution – response' inapplicable and necessitates a reorientation towards the proactive preservation of cognitive coherence as a condition of systemic security.

The practical significance of the framework developed herein lies in enabling a transition from reactive responses to discrete informational threats towards the systemic diagnosis of cognitive vulnerability. Such diagnosis presupposes the development of indicators of inter-layer connectivity, methodologies for identifying structural connector nodes, and metrics of cognitive coherence capable of assessing a system's resilience and transmorphance. These directions delineate a future research agenda at the intersection of complex systems theory, social psychology, cybernetics, neurobiology, strategic analysis, and information security.

## 8. Bibliography

1. NATO Science & Technology Organization. (2022). *2021 Highlights*. NATO Collaboration Support Office. [https://www.nato.int/content/dam/nato/legacy-wcm/media\\_pdf/2022/4/pdf/2021-NATO-STO-Highlights-web.pdf](https://www.nato.int/content/dam/nato/legacy-wcm/media_pdf/2022/4/pdf/2021-NATO-STO-Highlights-web.pdf)
2. NATO Science & Technology Organization. (2023). *2022 Highlights*. NATO Collaboration Support Office. <https://www.sto.nato.int/wp-content/uploads/2023-NATO-STO-Highlights-Web.pdf>
3. NATO Allied Command Transformation. (2023, May 9). *Cognitive warfare: Beyond military information support operations*. <https://www.act.nato.int/article/cognitive-warfare-beyond-military-information-support-operations/>
4. NATO Joint Warfare Centre. (2025). Cognitive warfare: Special report: The battlespace of the mind: Command, control, and the cognitive frontier. *The Three Swords*, 41. [https://www.jwc.nato.int/wp-content/uploads/2025/12/issue41\\_Art3\\_SpecialReport\\_COGWAR.pdf](https://www.jwc.nato.int/wp-content/uploads/2025/12/issue41_Art3_SpecialReport_COGWAR.pdf)
5. Hung, T.-Z., & Hung, T.-W. (2022). How China's cognitive warfare works: A frontline perspective of Taiwan's anti-disinformation wars. *Journal of Global Security Studies*, 7(4), Article osac016. <https://doi.org/10.1093/jogss/ogac016>
6. Hellström, J., Kallioniemi, P., Kytöneva, S., & Puranen, M. (2024). *Are Russian narratives amplified by PRC media? A case study on narratives related to Sweden's and Finland's NATO applications*. NATO Strategic Communications Centre of Excellence. <https://stratcomcoe.org/publications/are-russian-narratives-amplified-by-prc-media-a-case-study-on-narratives-related-to-swedens-and-finlands-nato-applications/298>
7. Claverie, B., & du Cluzel, F. (2022). The cognitive warfare concept. In B. Claverie, B. Prébot, N. Buchler, & F. du Cluzel (Eds.), *Cognitive warfare: The future of cognitive dominance* (pp. 6–11). NATO Collaboration Support Office. <https://hal.science/hal-03635889/document>
8. Smiljanic, D. (2025). Cognitive warfare - the human mind as the new battlefield. *Proceedings of the Defence and Security Conference 2025*, 1(1), 84–114. [https://www.researchgate.net/publication/391704397\\_Cognitive\\_warfare\\_-\\_the\\_human\\_mind\\_as\\_the\\_new\\_battlefield](https://www.researchgate.net/publication/391704397_Cognitive_warfare_-_the_human_mind_as_the_new_battlefield)
9. Le Guyader, H. (2022). Cognitive domain: A sixth domain of operations. In B. Claverie, B. Prébot, N. Buchler, & F. du Cluzel (Eds.), *Cognitive warfare: The future of cognitive dominance* (pp. 1–5). NATO Collaboration Support Office. <https://hal.science/hal-03635898v1>
10. Deppe, C., & Schaal, G. S. (2024). Cognitive warfare: A conceptual analysis of the NATO ACT cognitive warfare exploratory concept. *Frontiers in Big Data*, 7, Article 1452129. <https://doi.org/10.3389/fdata.2024.1452129>
11. Masakowski, Y. R., & Blatny, J. M. (Eds.). (2023). Mitigating and responding to cognitive warfare (STO Technical Report STO-TR-HFM-ET-356). NATO Science and Technology Organization. <https://www.sto.nato.int/document/mitigating-and-responding-to-cognitive-warfare-2/>
12. Kivelä, M., Arenas, A., Barthelemy, M., Gleeson, J. P., & Moreno, Y. (2014). Multilayer networks. *Journal of Complex Networks*, 2(3), 203–271. <https://doi.org/10.1093/comnet/cnu016>
13. Boccaletti, S., Bianconi, G., Criado, R., Del Genio, C. I., Gómez-Gardenes, J., Romance, M., Sendina-Nadal, I., Wang, Z., & Zanin, M. (2014). The structure and dynamics of multilayer networks. *Physics Reports*, 544(1), 1–122. <https://doi.org/10.1016/j.physrep.2014.07.001>

14. Groestad, P. (2022). *Cognitive Warfare – Hacking the OODA Loop* [Conference presentation]. NATO CCDCOE CyCon, Estonia. <https://www.youtube.com/watch?v=H3RqF5PiqXM>
15. Blatny, J. M., & Søndergaard, S. (2025). *NATO Chief Scientist's reports: Cognitive warfare*. NATO Science & Technology Organization. <https://www.sto.nato.int/wp-content/uploads/chief-scientist-report-cognitive-warfare-4.pdf>
16. Thomas, T. (2004). Russia's reflexive control theory and the military. *The Journal of Slavic Military Studies*, 17(2), 237–256. <https://doi.org/10.1080/13518040490450529>
17. Lefebvre, V. A. (1987). The fundamental structures of human reflexion. *Journal of Social and Biological Structures*, 10(2), 129–175. [https://doi.org/10.1016/0140-1750\(87\)90004-2](https://doi.org/10.1016/0140-1750(87)90004-2)
18. Splidsboel Hansen, F. (2021). When Russia wages war in the cognitive domain. *The Journal of Slavic Military Studies*, 34(2), 181–201. <https://doi.org/10.1080/13518046.2021.1990562>
19. Lee, S. (2014). China's 'Three Warfares': Origins, applications, and organizations. *Journal of Strategic Studies*, 37(2), 198–221. <https://doi.org/10.1080/01402390.2013.870071>
20. Aukia, J. (2023, April). China's hybrid influence in Taiwan: Non-state actors and policy responses (Hybrid CoE Research Report 9). European Centre of Excellence for Countering Hybrid Threats. <https://www.hybridcoe.fi/wp-content/uploads/2023/04/20230406-Hybrid-CoE-Research-Report-9-Chinas-hybrid-influence-in-Taiwan-WEB.pdf> .
21. Sun Tzu. (1963). *The art of war* (S. B. Griffith, Trans.). Oxford University Press. (Original work published ca. 5th century BCE)
22. Liddell Hart, B. H. (1929). *The decisive wars of history: A study in strategy*. G. Bell & Sons.
23. Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157), 1124–1131. <https://doi.org/10.1126/science.185.4157.1124>
24. Cialdini, R. B. (1984). *Influence: The psychology of persuasion*. HarperCollins.
25. Parsons, P. R. (2021). The lost doctrine: Suggestion theory in early media effects research. *Journalism & Communication Monographs*, 23(2), 80–138. <https://doi.org/10.1177/15226379211006119>
26. Woolley, S. C., & Howard, P. N. (Eds.). (2018). *Computational propaganda: Political parties, politicians, and political manipulation on social media*. Oxford University Press. <https://doi.org/10.1093/oso/9780190931407.001.0001>
27. Sunstein, C. R. (2001). *Republic.com*. Princeton University Press.
28. Sunstein, C. R. (2017). *Republic: Divided democracy in the age of social media*. Princeton University Press.
29. Pariser, E. (2011). *The filter bubble: What the Internet is hiding from you*. Penguin Press.
30. Arguedas, A. R., Robertson, C. T., Fletcher, R., & Nielsen, R. K. (2022). *Echo chambers, filter bubbles, and polarisation: A literature review*. Reuters Institute for the Study of Journalism. <https://reutersinstitute.politics.ox.ac.uk/echo-chambers-filter-bubbles-and-polarisation-literature-review>
31. Kitchens, B., Johnson, S. L., & Gray, P. (2020). Understanding echo chambers and filter bubbles: The impact of social media on diversification and partisan shifts in news consumption. *MIS Quarterly*, 44(4), 1619–1649. <https://doi.org/10.25300/MISQ/2020/16371>
32. Bruns, A. (2019). *Are filter bubbles real?* Polity Press.
33. Vaccari, C., & Chadwick, A. (2020). Deepfakes and disinformation: Exploring the impact of synthetic political video on deception, uncertainty, and trust in news. *Social Media + Society*, 6(1), 1–13. <https://doi.org/10.1177/2056305120903408>
34. Schiff, K. J., Schiff, D. S., & Bueno, N. S. (2024). The liar's dividend: Can politicians claim misinformation to evade accountability? *American Political Science Review*, 1–20. <https://doi.org/10.1017/S0003055423001454>

35. Lange, L. (2024, November 21). Decoding China's AI-Powered 'Algorithmic Cognitive Warfare' [White paper]. Special Competitive Studies Project. <https://www.scsp.ai/resource/decoding-chinas-ai-powered-algorithmic-cognitive-warfare/>
36. Special Competitive Studies Project. (2024). *Generative AI: The future of innovation and power*. SCSP. <https://www.scsp.ai>
37. Korobeynikov, F., & Mokhor, V. (2026). Adaptive security: Strategic principles for complex socio-technical systems. *Royal Society Open Science*, 13(1), 251481. <https://doi.org/10.1098/rsos.251481>
38. Maturana, H. R., & Varela, F. J. (1980). *Autopoiesis and cognition: The realization of the living*. D. Reidel.
39. Luhmann, N. (1995). *Social systems* (J. Bednarz, Jr. & D. Baecker, Trans.). Stanford University Press. (Original work published 1984)
40. Backes, O., & Swab, A. (2019). *Cognitive warfare: The Russian threat to election integrity in the Baltic states*. Belfer Center for Science and International Affairs, Harvard Kennedy School. <https://www.belfercenter.org/publication/cognitive-warfare-russian-threat-election-integrity-baltic-states>
41. Du Cluzel, F. (2020). *Cognitive warfare*. NATO ACT Innovation Hub. [https://innovationhub-act.org/wp-content/uploads/2023/12/20210122\\_CW-Final.pdf](https://innovationhub-act.org/wp-content/uploads/2023/12/20210122_CW-Final.pdf)
42. Szell, M., Lambiotte, R., & Thurner, S. (2010). Multirelational organization of large-scale social networks in an online world. *Proceedings of the National Academy of Sciences*, 107(31), 13636–13641. <https://doi.org/10.1073/pnas.1004008107>
43. Buldyrev, S. V., Parshani, R., Paul, G., Stanley, H. E., & Havlin, S. (2010). Catastrophic cascade of failures in interdependent networks. *Nature*, 464(7291), 1025–1028. <https://doi.org/10.1038/nature08932>
44. Bianconi, G. (2018). *Multilayer networks: Structure and function*. Oxford University Press. <https://doi.org/10.1093/oso/9780198753919.001.0001>
45. Korobeynikov, F. (2025). Differentiating socio-technical systems via multiplex network theory [Conference contribution, III Scientific-Practical Conference "Resilience of Dynamic Systems", Kyiv]. Figshare. <https://doi.org/10.6084/m9.figshare.30730316.v1>
46. Scheffer, M., Bascompte, J., Brock, W. A., Brovkin, V., Carpenter, S. R., Dakos, V., Held, H., van Nes, E. H., Rietkerk, M., & Sugihara, G. (2009). Early-warning signals for critical transitions. *Nature*, 461(7260), 53–59. <https://doi.org/10.1038/nature08227>
47. Sălăvăstru, C. M., & Rațiu, A. (2024). Mission command in multi-domain operations. *International Conference KNOWLEDGE-BASED ORGANIZATION*, 30(1), 161–165. <https://doi.org/10.2478/kbo-2024-0022>
48. Korobeynikov, F., Davydiuk, A., & Mokhor, V. (2026). Supplementary Information for "Ontological Foundations of Cognitive Warfare". *figshare*. <https://doi.org/10.6084/m9.figshare.31446277>