

ADALAH

Buletin Hukum & Keadilan

Cognitive Warfare and the Law of Armed Conflict: Legal Gaps, Normative Challenges, and the Regulation of Algorithmic Influence Operations in Hybrid Conflicts

Firman Hamdani^{*}, Gilang Rizki Aji Putra^{*}

Universitas Islam Negeri Syarif Hidayatullah Jakarta



[10.15408/adalah.v10j4./54223](https://doi.org/10.15408/adalah.v10j4./54223)

Abstract:

The rapid advancement of digital technologies and artificial intelligence has fundamentally transformed the nature of contemporary conflict, enabling influence operations that target human cognition as a strategic domain of hostilities. This article examines how international law, particularly the Law of Armed Conflict (LOAC), should regulate AI-driven algorithmic influence

^{*} Head of the Education Section of School Enterprises and Doctoral Students, Syarif Hidayatullah State Islamic University, Jakarta. Email: firmanhamdani@mpuin-jkt.sch.id

^{*} Researcher at the Center for Constitutional and National Legislation Studies (Poskolegnas), Syarif Hidayatullah State Islamic University, Jakarta. Email: gilang.rizkiajiputra19@uinjkt.ac.id.

operations that function as instruments of cognitive warfare within hybrid conflicts. Employing a doctrinal legal research method, this study undertakes a normative and analytical examination of existing international legal principles, complemented by critical evaluation of emerging scholarly debates and policy developments. The findings demonstrate that current LOAC frameworks are structurally designed for kinetic warfare and are therefore ill-equipped to capture the legal relevance of algorithmic influence, resulting in significant gaps concerning attribution, civilian protection, and sovereignty. The analysis further reveals that existing jus ad bellum and jus in bello doctrines fail to adequately regulate non-kinetic cognitive harm and systematic narrative manipulation, allowing strategic exploitation below traditional conflict thresholds. The article formulates normative and institutional proposals for a revised legal framework grounded in cognitive protection, due diligence obligations, and enhanced corporate responsibility. It concludes that international law must evolve towards formally recognising cognitive warfare as a regulated domain of armed conflict to ensure the effective protection of cognitive autonomy, democratic stability, and international security.

Keywords: Cognitive Warfare, Law of Armed Conflict, Algorithmic Influence, Hybrid Conflict, International Humanitarian Law, Cognitive Liberty, Disinformation Regulation

INTRODUCTION

Background and Problem Statement

The law of armed conflict (LOAC) remains the essential legal framework for regulating hostilities. Its conduct-of-hostilities rules rest on familiar but demanding principles: distinction, proportionality, and precaution. These principles seek to protect civilian populations and to limit harm to persons and objects that do not make an effective contribution to military action (Le & Trinh, 2026, p. 1). Contemporary conflict, however, increasingly includes cyber operations and

cognitive warfare: forms of coercion that do not fit neatly into conventional assumptions about battlefields, visible weapons, identifiable combatants, or geographical boundaries. These practices help produce a grey zone in which coercion is intentionally designed to avoid clear legal classification and to complicate lawful response (Le & Trinh, 2026, p. 1).

The rapid expansion of artificial intelligence in information operations has fundamentally transformed contemporary conflict by enabling algorithmic influence capable of shaping geopolitical narratives and cognitive perceptions beyond traditional battlefields (Manitra & Prabandari, 2026, p. 1). Information flows now travel across borders at unprecedented speed, reshaping public opinion, political identities, and collective decision-making processes. In this evolving digital ecosystem, algorithms increasingly mediate what individuals see, read, and believe, thereby influencing social narratives and political attitudes in subtle yet powerful ways (Manitra & Prabandari, 2026, p. 1). While these developments have expanded access to information and democratic participation, they have also created new vulnerabilities, particularly in relation to manipulation, disinformation, and external interference in domestic political processes (Manitra & Prabandari, 2026, p. 1).

Cognitive warfare represents a qualitative shift from traditional psychological operations. Unlike conventional information operations that seek to persuade—to change what people think about specific issues—cognitive warfare targets the mental processes through which individuals and collectives perceive, interpret, and respond to their

environment. This includes targeting attention, memory, emotional regulation, and the epistemic frameworks through which individuals determine what constitutes reliable knowledge (NATO Science & Technology Organization, 2020). The strategic significance of this domain has been increasingly recognized by military organizations and policy institutions worldwide, with NATO formally adopting a cognitive warfare concept in June 2025 that lists concrete objectives of adversaries and recommended measures for member states.

Research Gap and Contribution

Despite growing policy attention, the academic literature on the legal regulation of cognitive warfare remains fragmented and underdeveloped. Three specific gaps emerge from this fragmentation.

First, theoretical integration between international law and cognitive warfare scholarship remains incomplete. While significant advances have been made in understanding the technical and psychological dimensions of cognitive warfare, the legal implications have received comparatively less systematic attention. The existing LOAC framework was designed for kinetic warfare and struggles to accommodate non-kinetic forms of harm that characterize cognitive operations (Manitra & Prabandari, 2026, p. 1).

Second, the grey zone problem—where operations are deliberately designed to fall below the threshold of armed conflict—creates a legal vacuum. As Le and Trinh (2026) argue, LOAC remains applicable once operations are sufficiently linked to an armed conflict, but it is strained by

dual-use digital infrastructure, contested attribution, and non-kinetic forms of harm. Meanwhile, operations that remain below the conflict threshold may escape legal regulation entirely.

Third, the intersection of algorithmic decision-making and international legal doctrine has not been adequately theorized. The use of AI-driven systems to conduct influence operations raises novel questions about attribution, state responsibility, and the threshold for classifying conduct as an “attack” under international law (Manitra & Prabandari, 2026, p. 1).

This article addresses these gaps by providing a comprehensive legal analysis of cognitive warfare within the framework of LOAC, proposing normative and institutional adaptations to ensure that international law remains protective in the digital age. The article pursues four objectives: (1) to analyze the legal status of cognitive warfare under existing LOAC frameworks; (2) to identify specific normative gaps and structural limitations; (3) to examine emerging legal and policy responses across international, regional, and national levels; and (4) to formulate proposals for legal reform.

Research Questions

This article is guided by the following research questions:

- RQ1: How does existing LOAC apply to cognitive warfare operations, and what are the principal legal gaps?

- RQ2: What normative and institutional adaptations are necessary to regulate algorithmic influence operations effectively?
- RQ3: How can the protection of civilian cognitive autonomy be operationalized within international humanitarian law?

LITERATURE REVIEW

The Law of Armed Conflict and Non-Kinetic Harm

The LOAC, codified in the four Geneva Conventions of 1949 and their Additional Protocols, provides the primary normative framework for regulating armed conflict. Its core principles—distinction, proportionality, and precaution—are designed to protect civilians and civilian objects during hostilities. However, these principles are fundamentally structured around kinetic assumptions of harm. As Jibril et al. (2026) observe, contemporary armed conflicts increasingly extend into cognitive and digital domains, where social media platforms are strategically used to influence perception, behaviour, and the informational environment of hostilities. The scale, speed, and algorithmic amplification enabled by social media create new regulatory challenges that existing frameworks struggle to address.

The Tallinn Manual 2.0, the most authoritative expert commentary on the international law applicable to cyber operations, defines an “attack” for the purposes of LOAC as “a cyber operation, whether offensive or defensive, that is reasonably expected to cause injury or death to persons or

damage or destruction to objects” (Tallinn Manual 2.0, Rule 92). Under this definition, non-violent psychological operations—including propaganda broadcasts and disinformation campaigns—do not qualify as attacks, even when directed at civilians. The Tallinn Manual’s International Group of Experts agreed that psychological operations such as making propaganda broadcasts are permitted even if against civilians, and that non-violent operations do not qualify as attacks (Tallinn Manual 2.0, Rule 123 commentary).

This definitional framework creates a significant gap. As Jibril et al. (2026) argue, social media-driven influence operations frequently fall below the legal threshold of an “attack” despite producing significant humanitarian consequences. These harms are often cumulative, diffuse, and non-kinetic, challenging legal frameworks traditionally centred on physical destruction. The article further identifies difficulties relating to proportionality, attribution, and accountability, concluding that LOAC requires contextual interpretation to address emerging forms of harm caused by the manipulation of information, perception, and cognition.

Cognitive Warfare as a Legal Category

The conceptual shift from psychological operations to cognitive warfare has profound legal implications. Manitra and Prabandari (2026) demonstrate that algorithmic cognitive warfare represents a profound and multidimensional challenge to existing international humanitarian law. By systematically mapping normative gaps against targeted LOAC instruments, they highlight both the urgency and the

feasibility of integrating cognitive hostilities into the legal framework of armed conflict.

Their analysis reveals that current LOAC frameworks are structurally designed for kinetic warfare and are therefore ill-equipped to capture the legal relevance of algorithmic influence, resulting in significant gaps concerning attribution, civilian protection, and sovereignty. Existing *jus ad bellum* and *jus in bello* doctrines fail to adequately regulate non-kinetic cognitive harm and systematic narrative manipulation, allowing strategic exploitation below traditional conflict thresholds.

The article argues that international law must evolve towards formally recognising cognitive warfare as a regulated domain of armed conflict to ensure the effective protection of cognitive autonomy, democratic stability, and international security. This requires the legal classification of algorithmic influence operations as conduct capable of producing strategic harm comparable to kinetic attacks, thus subjecting them to the principles of distinction, proportionality, military necessity, and accountability.

Effects-Based Approaches to Civilian Harm

A central difficulty identified in the legal analysis is the indirect nature of harm within the digital battlespace. Disinformation does not often affect people directly or immediately, but rather shapes behaviour and decision-making, which may generate harmful downstream consequences (Cogent Social Sciences, 2026). International Humanitarian Law provides a coherent normative framework

for regulating armed conflict; however, its application to disinformation as a form of cognitive warfare exposes not merely technical gaps but a structural bias within LOAC toward material conceptions of harm.

An effects-based approach to civilian harm offers a potential solution. As Jibril et al. (2026) explain, an effects-based approach brings these dynamics into sharper focus by asking not how an operation is carried out, but what consequences it produces. From this perspective, the legal analysis shifts from the modality of the operation to its material effects on civilian populations. Under LOAC, proportionality requires that the anticipated military advantage of an operation be weighed against the expected incidental harm to civilians. If cognitive operations produce effects functionally equivalent to physical harm—such as large-scale behavioural manipulation or disruption of civilian decision-making—they may fall within the scope of regulated conduct (Cogent Social Sciences, 2026).

This effects-based approach aligns with broader developments in international humanitarian law, where the focus has increasingly shifted from the means of warfare to its consequences for protected persons. The ICRC's Oxford Statement on the Protection of Cyberspace in International Law affirms that “the conduct of information operations or activities in armed conflict is subject to the applicable rules of International Humanitarian Law”. This position supports the extension of IHL protections to cognitive harm, even when the mechanisms of harm differ from traditional kinetic operations.

The Grey Zone Problem and Legal Ambiguity

The grey zone problem is central to understanding the legal challenges posed by cognitive warfare. As Le and Trinh (2026) explain, cyberspace is transboundary by design, and much of it is predominantly owned, operated, and managed by private-sector actors. This creates a grey zone in which coercion is intentionally designed to avoid clear legal classification and to complicate lawful response.

Legislation itself can become an instrument of cognitive warfare. Pijpers (2024) demonstrates that the use of law as an instrument of power to affect perception and cognition is possible because of ongoing legal disputes about how to apply international law to cyberspace. States can cherry-pick or even assertively exploit variations in interpretations of international law to pursue or defend their national interests as a means of cognitive warfare. This phenomenon—lawfare as cognitive warfare—creates a meta-level challenge: the legal framework intended to constrain cognitive warfare can itself be weaponized in service of cognitive objectives.

The grey zone problem is exacerbated by the difficulty of attribution in cyberspace. Unlike kinetic attacks, which typically leave physical evidence that can be traced to specific actors, cognitive operations can be conducted through proxy servers, bot networks, and anonymizing technologies that obscure the identity of the perpetrator. This attribution challenge complicates both the application of state responsibility doctrines and the practical enforcement of any legal norms (Manitra & Prabandari, 2026, p. 1).

Cognitive Liberty as an Emerging Human Right

The concept of cognitive liberty has emerged as a potential normative foundation for protecting individuals against cognitive warfare. Cognitive liberty encompasses the right to self-determination over one's own brain and mental experiences, and the right to be free from interference, manipulation, or coercion by others. As legal scholars increasingly advocate, the recognition of cognitive liberty as a fundamental human right would encompass mental privacy, freedom of thought, and self-determination (Manitra & Prabandari, 2026, p. 1).

This understanding finds support in the UN Human Rights Committee's interpretation of Article 19 in its General Comment No. 34 of 2011, which emphasises that freedom of expression includes the freedom to seek, receive, and impart information and ideas of all kinds, and that any restriction must be strictly necessary and proportionate to a legitimate aim (Manitra & Prabandari, 2026, p. 1). The extension of civilian protection principles under LOAC to include epistemic autonomy would represent a natural evolution of humanitarian norms, aligning legal doctrine with technological realities. Such recognition would strengthen civilian immunity by prohibiting algorithmic practices that intentionally distort cognitive environments or exploit psychological vulnerabilities (Manitra & Prabandari, 2026, p. 1).

The operationalisation of these principles requires institutional mechanisms that preserve meaningful human control over AI systems. Scholars consistently warn against automation bias and the erosion of human moral agency in

algorithmically mediated decision-making (Manitra & Prabandari, 2026, p. 1). The Prediction Management Framework and structured ethical assessment models provide practical tools for embedding transparency, auditability, and interpretability into military AI systems, ensuring that human oversight remains central throughout the deployment cycle (Manitra & Prabandari, 2026, p. 1).

Emerging Legal and Policy Responses

International-Level Developments

At the international level, several developments signal growing recognition of the cognitive warfare challenge. NATO's Allied Command Transformation has developed a cognitive warfare concept adopted by the NATO Military Committee in June 2025, which lists concrete shortcomings, objectives of adversaries, and recommended measures for member states to plan and coordinate their responses. The concept recommends that militaries integrate information and influence experts into their command structures, creating a dedicated J10 staff branch for cognitive effects, and encourages more aggressive peacetime information operations rather than only countering disinformation defensively.

The ICRC's Oxford Statement represents another significant development, affirming that information operations and activities in armed conflict are subject to applicable rules of International Humanitarian Law. This position provides a normative foundation for extending IHL protections to cognitive harm, even when the mechanisms of harm differ from traditional kinetic operations.

However, these developments remain largely declaratory. As Manitra and Prabandari (2026) note, existing *jus ad bellum* and *jus in bello* doctrines fail to adequately regulate non-kinetic cognitive harm and systematic narrative manipulation, allowing strategic exploitation below traditional conflict thresholds. Binding legal instruments specifically addressing cognitive warfare remain absent.

Regional-Level Responses

The European Union has taken the most proactive approach to regulating disinformation and cognitive threats. The Digital Services Act (DSA), adopted in 2022, establishes a comprehensive framework for platform accountability and risk mitigation. The DSA introduces strict requirements for online platforms, including accountability for illegal content and fines for non-compliance, risk assessments on harmful information, transparency in how algorithms work, user reporting tools, stricter ad rules, and crisis response mechanisms to limit fake information during emergencies.

A significant development occurred in February 2025 when the European Commission and the European Board for Digital Services endorsed the integration of the voluntary Code of Practice on Disinformation into the framework of the Digital Services Act as a code of conduct on disinformation. Since 1 July 2025, the Code of Practice on Disinformation has been officially integrated into the DSA framework, marking a shift from self-regulation to co-regulation.

The DSA requires providers of Very Large Online Platforms (VLOPs) and Very Large Online Search Engines

(VLOSEs) to analyse whether and how risks are influenced by manipulation of their services, such as the use of bots or fake accounts. While the DSA does not prescribe which content is illegal—this is determined by national law or other EU laws—it establishes procedural obligations that enhance platform accountability and transparency.

National-Level Approaches

At the national level, approaches to regulating cognitive warfare vary significantly. France's Law on the Freedom of the Press of 1881 remains the cornerstone of its legal framework, reflecting a balance between freedom of expression and responsibility that has gradually evolved under the influence of new technologies but whose unchanged objective remains to reconcile freedom of expression with the need to address information manipulation.

Indonesia has taken a different approach, with the Electronic Information and Transactions Law (UU ITE) providing the primary legal basis for addressing disinformation. However, Article 40 of the UU ITE, which states that the government has the duty to protect public safety and the public interest from disruptions caused by misinformation and disinformation, has been criticized for potential overreach and for being used to restrict legitimate expression. Comparative analysis suggests that Indonesia's fragmented regulatory framework results in limited platform compliance—approximately 15% during electoral periods—whereas the EU's DSA achieves up to 85% compliance with valid disinformation reports actioned within 24 hours.

Brazil's Bill No. 2630, known as the "Fake News Law," seeks to combat misinformation by requiring platforms to store records of mass-forwarded messages for three months and banning external tools used for spam or misinformation dissemination. The bill represents an attempt to impose accountability obligations on platforms while respecting constitutional protections for freedom of expression.

RESEARCH METHOD

Research Design

This article employs a doctrinal legal research method, complemented by normative and analytical examination of existing international legal principles. Doctrinal legal research is appropriate for this study because it enables systematic analysis of primary legal sources—treaties, customary international law, and general principles—as well as secondary sources including scholarly commentary and policy documents. This approach is consistent with the methodology employed in recent legal scholarship on cognitive warfare (Manitra & Prabandari, 2026; Le & Trinh, 2026; Jibril et al., 2026).

The research design combines three analytical approaches:

1. Normative analysis of primary legal instruments governing armed conflict, including the four Geneva Conventions of 1949, Additional Protocol I of 1977, the Tallinn Manual 2.0, and relevant human rights instruments including the International Covenant on Civil and Political Rights.

2. Critical evaluation of emerging scholarly debates on the legal regulation of cognitive warfare, including recent journal articles, policy reports, and institutional publications from NATO, the ICRC, and the European Union.
3. Comparative analysis of legal and policy responses across international, regional, and national levels, identifying patterns, divergences, and best practices.

Data Collection and Sources

Primary legal sources include:

- The four Geneva Conventions of 1949 and their Additional Protocols
- The Tallinn Manual 2.0 on the International Law Applicable to Cyber Operations
- The International Covenant on Civil and Political Rights (ICCPR)
- General Comment No. 34 of the UN Human Rights Committee on Article 19 of the ICCPR
- European Union Digital Services Act (Regulation (EU) 2022/2065)
- EU Code of Practice on Disinformation
- NATO Cognitive Warfare Concept (2025)

Secondary sources include peer-reviewed journal articles, institutional reports, and policy analyses from

recognized academic publishers including Taylor & Francis (Cogent Social Sciences), Oxford University Press, Cambridge University Press, and university-affiliated research centers (GCSP, NATO STO). All sources are analyzed for relevance, verifiability, and scholarly rigor.

Analytical Framework

The analytical framework integrates four conceptual dimensions:

1. **Threshold analysis:** Determining when cognitive operations cross legal thresholds—the threshold for “attack” under LOAC, the threshold for armed conflict under jus ad bellum, and the threshold for prohibited intervention under the principle of non-intervention.
2. **Harm analysis:** Examining whether cognitive operations produce effects functionally equivalent to physical harm, using an effects-based approach to civilian harm (Jibril et al., 2026).
3. **Attribution analysis:** Assessing the challenges of attributing cognitive operations to state or non-state actors and the implications for state responsibility.
4. **Normative analysis:** Evaluating the adequacy of existing legal norms and identifying opportunities for normative evolution, including the recognition of cognitive liberty as a protected legal interest.

Limitations

This study has several limitations. Doctrinal legal analysis relies on interpretation of existing legal instruments and scholarly commentary; it does not generate empirical data on the effectiveness of cognitive warfare or the impacts of specific legal measures. The rapid pace of technological change means that specific techniques and countermeasures may evolve quickly. Additionally, the analysis is predominantly focused on Western legal traditions; cognitive warfare dynamics and legal responses may differ in other legal systems and cultural contexts.

RESULTS AND DISCUSSION

1. The Structural Bias of LOAC Toward Material Harm

The analysis reveals that the most fundamental challenge for regulating cognitive warfare under LOAC is the structural bias toward material conceptions of harm. As the literature demonstrates, disinformation does not often affect people directly or immediately, but rather shapes behaviour and decision-making, which may generate harmful downstream consequences. This indirect nature of harm exposes not merely technical gaps but a structural bias within LOAC toward material conceptions of harm (Cogent Social Sciences, 2026).

This bias is evident in the Tallinn Manual's definition of "attack," which requires injury or death to persons, or damage or destruction to objects. Under this definition, non-violent psychological operations do not qualify as attacks, even when directed at civilians. The Tallinn Manual's International Group of Experts explicitly agreed that

“psychological operations such as making propaganda broadcasts and dropping leaflets are permitted even if against civilians,” and that non-violent operations do not qualify as attacks.

The consequence is that cognitive warfare operates in what Manitra and Prabandari (2026) characterize as a legal grey zone, where strategic exploitation below traditional conflict thresholds is possible because existing *jus ad bellum* and *jus in bello* doctrines fail to adequately regulate non-kinetic cognitive harm and systematic narrative manipulation. States and non-state actors can conduct large-scale cognitive operations that produce significant societal and operational effects without triggering the legal consequences that would attach to kinetic attacks.

2. Attribution Challenges and Accountability Gaps

The analysis identifies attribution as a critical obstacle to the effective regulation of cognitive warfare. Unlike kinetic attacks, which typically leave physical evidence that can be traced to specific actors, cognitive operations can be conducted through proxy servers, bot networks, and anonymizing technologies that obscure the identity of the perpetrator. This attribution challenge complicates both the application of state responsibility doctrines and the practical enforcement of any legal norms (Manitra & Prabandari, 2026, p. 1).

The problem is compounded by the involvement of corporate actors. As Manitra and Prabandari (2026) observe, AI developers and digital platform providers increasingly function as strategic actors within contemporary conflict

ecosystems, yet remain only partially bound by international humanitarian obligations. Extending the UN Guiding Principles on Business and Human Rights to encompass algorithmic warfare contexts would impose enhanced human rights due diligence requirements on corporate entities. Ethical governance frameworks further recommend traceability mechanisms, transparency audits, and operational safeguards designed to prevent corporate complicity in cognitive manipulation campaigns.

The attribution challenge also creates strategic incentives for the use of disinformation as a low-risk, high-impact instrument of warfare. As the literature notes, the non-forceful, non-dictatorial, and secretive nature of cyber operations necessitates re-examining or expanding the traditional concept of intervention. Standards centered exclusively on overt coercion fail to capture wrongful cyber interventions carried out through manipulation, deception, disruption, and disinformation, creating a legal grey zone that major cyber power states exploit (International Law Studies, 2026).

3. The Effects-Based Approach as a Legal Pathway

The effects-based approach to civilian harm offers the most promising pathway for extending LOAC protections to cognitive warfare. As Jibril et al. (2026) demonstrate, an effects-based approach brings the dynamics of cognitive harm into sharper focus by asking not how an operation is carried out, but what consequences it produces. From this perspective, the legal analysis shifts from the modality of the operation to its material effects on civilian populations.

This approach finds support in the ICRC's Oxford Statement on the Protection of Cyberspace in International Law, which affirms that “the conduct of information operations or activities in armed conflict is subject to the applicable rules of International Humanitarian Law”. The Oxford Statement's framework, which defines regulated conduct as “any coordinated or individual application of digital resources for cognitive purposes,” provides a normative foundation for extending IHL protections to cognitive harm (ICRC International Review, 2025).

The effects-based approach also has implications for the proportionality analysis under LOAC. If cognitive operations produce effects functionally equivalent to physical harm—such as large-scale behavioural manipulation or disruption of civilian decision-making—the proportionality calculation should account for these cognitive harms alongside physical harms. As Manitra and Prabandari (2026) argue, the extension of civilian protection principles under LOAC to include epistemic autonomy would represent a natural evolution of humanitarian norms, aligning legal doctrine with technological realities.

However, the effects-based approach faces practical challenges. Determining whether a cognitive operation produces “functionally equivalent” harm requires developing metrics for assessing cognitive harm that do not currently exist. The causal chain between exposure to disinformation and behavioural change is complex and mediated by numerous factors, making it difficult to attribute specific harms to specific operations. These challenges, while

significant, are not insurmountable; they parallel similar challenges in environmental law and public health, where causation is often probabilistic and indirect.

4. Cognitive Liberty and the Protection of Mental Autonomy

The analysis suggests that the emerging concept of cognitive liberty provides a normative foundation for protecting individuals against cognitive warfare. As recognized in the literature, cognitive liberty encompasses the right to self-determination over one's own brain and mental experiences, and the right to be free from interference, manipulation, or coercion by others. The recognition of cognitive liberty as a fundamental human right would encompass mental privacy, freedom of thought, and self-determination (Manitra & Prabandari, 2026, p. 1).

This understanding finds support in the UN Human Rights Committee's interpretation of Article 19 in General Comment No. 34, which emphasises that freedom of expression includes the freedom to seek, receive, and impart information and ideas of all kinds, and that any restriction must be strictly necessary and proportionate to a legitimate aim (Manitra & Prabandari, 2026, p. 1). The extension of these protections to encompass cognitive autonomy would represent a logical evolution of human rights doctrine in response to technological developments that enable unprecedented interference with mental processes.

The operationalisation of cognitive liberty requires institutional mechanisms that preserve meaningful human control over AI systems. As Manitra and Prabandari (2026)

observe, scholars consistently warn against automation bias and the erosion of human moral agency in algorithmically mediated decision-making. The Prediction Management Framework and structured ethical assessment models provide practical tools for embedding transparency, auditability, and interpretability into military AI systems, ensuring that human oversight remains central throughout the deployment cycle.

5 Comparative Analysis of Regulatory Approaches

The analysis reveals significant variation in how different jurisdictions have approached the regulation of cognitive warfare and related threats.

European Union: The EU has developed the most comprehensive regulatory framework through the Digital Services Act and the Code of Practice on Disinformation. The integration of the Code into the DSA framework in 2025 marked a shift from voluntary self-regulation to binding co-regulation. The DSA requires Very Large Online Platforms to conduct risk assessments on harmful information, provide transparency in algorithm functioning, and implement crisis response mechanisms during emergencies. This approach achieves significantly higher compliance rates than fragmented national frameworks, with up to 85% of valid disinformation reports actioned within 24 hours under the DSA compared to approximately 15% under Indonesia's fragmented framework.

Indonesia: Indonesia's approach relies primarily on the Electronic Information and Transactions Law (UU ITE), which provides government authority to address misinformation but

has been criticized for potential overreach and restriction of legitimate expression. The fragmented nature of Indonesia's regulatory framework limits platform compliance and creates uncertainty for both platforms and users.

International level: NATO has adopted a cognitive warfare concept and recommended measures for member states, including the integration of information and influence experts into command structures and more aggressive peacetime information operations. However, NATO's approach remains at the level of doctrine and recommendations rather than binding legal obligations. The ICRC's Oxford Statement provides normative guidance but is not legally binding.

This comparative analysis suggests that effective regulation of cognitive warfare requires a combination of binding legal obligations, institutional mechanisms for enforcement, and international cooperation. The EU model, while not without limitations, provides the most developed template for comprehensive regulation.

6. Normative and Institutional Proposals

Based on the analysis, the article formulates several normative and institutional proposals for a revised legal framework.

Proposal 1: Formal recognition of cognitive warfare as a regulated domain. International law should formally recognise cognitive warfare as a legitimate domain of hostilities, extending established humanitarian principles into

the informational and psychological spheres of contemporary conflict (Manitra & Prabandari, 2026, p. 1). This requires the legal classification of algorithmic influence operations as conduct capable of producing strategic harm comparable to kinetic attacks, thus subjecting them to the principles of distinction, proportionality, military necessity, and accountability.

Proposal 2: Adoption of an effects-based interpretation of “attack.” The definition of “attack” under LOAC should be expanded to encompass operations that produce effects functionally equivalent to physical harm, including large-scale cognitive and behavioural manipulation. This interpretation would enable the application of LOAC protections to cognitive warfare while respecting the principle of distinction.

Proposal 3: Recognition of cognitive integrity and informational sovereignty as protected legal interests. LOAC must evolve to incorporate cognitive integrity, informational sovereignty, and algorithmic accountability as protected legal interests, supported by strengthened state due diligence obligations and expanded civilian protection norms (Manitra & Prabandari, 2026, p. 1).

Proposal 4: Extension of corporate responsibility norms. AI developers and digital platform providers should be subject to enhanced human rights due diligence requirements. Extending the UN Guiding Principles on Business and Human Rights to encompass algorithmic warfare contexts would impose binding responsibilities on corporate entities, ensuring transparency, traceability, and human rights

due diligence throughout the lifecycle of AI systems (Manitra & Prabandari, 2026, p. 1).

Proposal 5: Development of institutional mechanisms for enforcement. A complementary enforcement model should be developed, whereby the International Court of Justice addresses broader questions of state responsibility and legal interpretation, while specialised tribunals or independent investigatory bodies examine operational conduct, attribution, and compliance in specific conflict situations (Manitra & Prabandari, 2026, p. 1). Such tribunals could incorporate multidisciplinary expertise in artificial intelligence, cyber operations, military strategy, and international humanitarian law.

Proposal 6: Adoption of multilateral governance mechanisms. Through multilateral governance mechanisms and emerging international standards, international law can establish a coherent normative framework capable of constraining the weaponisation of algorithmic systems while preserving strategic stability and humanitarian values within hybrid conflict environments (Manitra & Prabandari, 2026, p. 1). The establishment of a “Digital Geneva Convention” has been proposed as a potential framework for governing state behaviour in cyberspace, including cognitive operations (International Journal of Law Management & Humanities, 2026).

CONCLUSION

This article has examined the legal challenges posed by cognitive warfare within the framework of the Law of Armed

Conflict and proposed normative and institutional adaptations to ensure that international law remains protective in the digital age.

The analysis demonstrates that current LOAC frameworks are structurally designed for kinetic warfare and are ill-equipped to capture the legal relevance of algorithmic influence, resulting in significant gaps concerning attribution, civilian protection, and sovereignty. Existing *jus ad bellum* and *jus in bello* doctrines fail to adequately regulate non-kinetic cognitive harm and systematic narrative manipulation, allowing strategic exploitation below traditional conflict thresholds.

The article argues that international law must evolve towards formally recognising cognitive warfare as a regulated domain of armed conflict. This requires the legal classification of algorithmic influence operations as conduct capable of producing strategic harm comparable to kinetic attacks, the adoption of an effects-based interpretation of "attack," and the recognition of cognitive integrity and informational sovereignty as protected legal interests. These developments must be supported by strengthened state due diligence obligations, expanded civilian protection norms, and binding responsibilities on corporate AI providers.

The article further argues that effective regulation requires institutional mechanisms for enforcement, including a complementary model whereby the International Court of Justice addresses broader questions of state responsibility

while specialised tribunals examine operational conduct. Multilateral governance mechanisms and emerging international standards can provide a coherent normative framework capable of constraining the weaponisation of algorithmic systems while preserving strategic stability and humanitarian values.

The stakes are high. Democratic societies depend on the capacity of citizens to form reliable beliefs, engage in reasoned deliberation, and make informed decisions. Cognitive warfare threatens these capacities in ways that existing institutions and frameworks are not fully prepared to address. Meeting this challenge requires the combined efforts of researchers, policymakers, technology companies, civil society, and citizens themselves—a collective response commensurate with a collective threat.

REFERENCE:

- Cogent Social Sciences. (2026). Law in the grey zone: In the age of cyber operations and cognitive warfare. *Cogent Social Sciences*, 12(1), 2725424. <https://doi.org/10.1080/23311886.2026.2725424>
- Cogent Social Sciences. (2026). Disinformation as a weapon: Legal limits of cognitive warfare under International Humanitarian Law. *Cogent Social Sciences*, 12(1), 2687180. <https://doi.org/10.1080/23311886.2026.2687180>
- Cogent Social Sciences. (2026). Weaponisation of social media in Nigeria's armed conflicts: Cognitive warfare and its implications for the law of armed conflict. *Cogent*

- Social Sciences, 12(1), 2717916.
<https://doi.org/10.1080/23311886.2026.2717916>
- Geneva Centre for Security Policy. (2025). Enhancing cognitive security and societal resilience to counter cognitive warfare. GCSP Publication.
<https://www.gcsp.ch/publications/enhancing-cognitive-security-and-societal-resilience-counter-cognitive-warfare>
- International Review of the Red Cross. (2025). The Oxford Statement on the protection of cyberspace in international law. International Review of the Red Cross. <https://international-review.icrc.org>
- Jibril, A. M., Faizin, M., & Pillai, A. V. (2026). Weaponisation of social media in Nigeria's armed conflicts: Cognitive warfare and its implications for the law of armed conflict. Cogent Social Sciences, 12(1), 2717916.
<https://doi.org/10.1080/23311886.2026.2717916>
- Le, M. H., & Trinh, D. T. (2026). Law in the grey zone: In the age of cyber operations and cognitive warfare. Cogent Social Sciences, 12(1), 2725424.
<https://doi.org/10.1080/23311886.2026.2725424>
- Manitra, R. R. M., & Prabandari, A. P. (2026). AI in hybrid conflicts: Towards a legal framework for cognitive warfare. Cogent Social Sciences, 12(1), 2681304.
<https://doi.org/10.1080/23311886.2026.2681304>
- NATO Science & Technology Organization. (2020). Cognitive warfare: A new conflict domain? STO Technical Report. <https://www.sto.nato.int>

- Pijpers, B. M. J. (2024). Legislation as an instrument of cognitive warfare. *The Defence Horizon Journal*. <https://doi.org/10.5281/zenodo.14212637>
- Tallinn Manual 2.0. (2017). Tallinn Manual 2.0 on the International Law Applicable to Cyber Operations. Cambridge University Press.
- United Nations Human Rights Committee. (2011). General Comment No. 34: Article 19: Freedoms of opinion and expression. CCPR/C/GC/34.
- European Commission. (2025). Code of Practice on Disinformation. <https://digital-strategy.ec.europa.eu>
- European Union. (2022). Regulation (EU) 2022/2065 of the European Parliament and of the Council on a Single Market For Digital Services (Digital Services Act). Official Journal of the European Union.
- International Committee of the Red Cross. (2025). Oxford Statement on the Protection of Cyberspace in International Law. <https://international-review.icrc.org>
- NATO Allied Command Transformation. (2025). NATO Cognitive Warfare Concept. NATO Military Committee.