

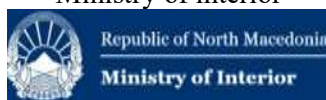
SECURITY HORIZONS

ISSN 2671-3624

DOI 10.5281/zenodo.22637980



Under the patronage of the
Ministry of interior



S2C2 2026: XVII Security Challenges Scientific Conference

SECURITY 2030: AN ERA OF SMART TECHNOLOGIES, RESILIENT SOCIETIES, RISKS, AND CHALLENGES

September 24-25, 2026

Ohrid, Republic of North Macedonia

Volume XI, No. 4

Skopje, 2026

БЕЗБЕДНОСНИ ХОРИЗОНТИ

ISSN 2671-3624

DOI 10.5281/zenodo.22637980



Под покровителство на
Министерство за внатрешни работи



S2C2 2026:

**XVII Научна конференција за безбедносни
предизвици**

**БЕЗБЕДНОСТ 2030: ЕРА НА ПАМЕТНИ
ТЕХНОЛОГИИ, ОТПОРНИ ОПШТЕСТВА,
РИЗИЦИ И ПРЕДИЗВИЦИ**

24-25 Септември 2026

Охрид, Република Северна Македонија

Година XI, Број 4

Skopje, 2026

Издавачи:

Универзитет „Св. Климент
Охридски“ Битола
Факултет за безбедност – Скопје

За издавачите:

Проф. д-р Игор Неделковски,
Ректор на Универзитетот „Св.
Климент Охридски“ – Битола
Проф. д-р Јонче Ивановски,
Декан на Факултетот за
безбедност – Скопје

Уредник на изданието:

Проф. д-р Јонче Ивановски

За сите трудови објавени во
овој зборник е извршена
јазична проверка.

Авторите дадоа согласност на
издавачката политика на
издавачот.

Компјутерска обработка:

м-р Оливера Трајанова Ѓорѓијовски
Кемал Рушид

Печати:

„Графо пром“ - Битола

Адреса на издавачите:

Факултет за безбедност
1000 Скопје
П. Фах 103
тел: 022546211
www.fb.uklo.edu.mk
Универзитет „Св. Климент
Охридски“
1ви Мај б.б. 7000 Битола,
тел: 047223788

Publishers:

University “St. Kliment Ohridski”- Bitola
Faculty of Security- Skopje

For the Publishers:

Prof. Dr. sc. Igor Nedelkovski,
Rector of the University “St.
Kliment Ohridski”- Bitola
Prof. Dr. sc. Jonce Ivanovski,
Dean of the Faculty of Security - Skopje

Editor in Chief:

Prof. Dr. sc. Jonce Ivanovski

All papers in these proceedings have
been copy-edited and proofread.

The authors have accepted the
publication policy of the proceedings.

Computer Processing:

M.Sc. Olivera Trajanova Gjorgijovski
Kemal Rusid

Print:

“Grafo prom” - Bitola

Address of the Publishers:

Faculty of Security
1000 Skopje
P.O. Box 103
tel: +389(0)22546211
www.fb.uklo.edu.mk
University “St. Kliment
Ohridski”
1 Maj b.b. 7000 Bitola
tel: +389(0) 47223788

INTERNATIONAL PROGRAMME COMMITTEE

1. **Prof. Dr. sc. Vlado Kambovski**, Academician, Macedonian Academy of Sciences and Arts;
2. **Prof. Dr. sc. Jonce Ivanovski**, Dean of the Faculty of Security, Skopje, Chair;
3. **Prof. Dr. Katerina Krstevska Savovoska**, Vice-Dean for Teaching, Faculty of Security, Skopje, Member;
4. **Prof. Dr. sc. Snezana Mojsoska**, Vice-Dean for Finance, Faculty of Security, Skopje, Member;
5. **Prof. Dr. sc. Igor Nedelkovski**, Rector of the University “St. Kliment Ohridski”, Bitola, Member;
6. **Prof. Dr. sc. Svetlana Nikoloska**, Vice-Rector of the University “St. Kliment Ohridski”, Bitola, Member;
7. **Prof. Dr. sc. Zlate Dimovski**, Faculty of Security, Skopje, Member;
8. **Prof. Dr. sc. Natasha Pelivanova**, Faculty of Security, Skopje, Member;
9. **Prof. Dr. sc. Marjan Nikolovski**, Faculty of Security, Skopje, Member;
10. **Prof. Dr. sc. Vesna Stefanovska**, Faculty of Security, Skopje, Member;
11. **Prof. Dr. sc. Tatjana Spaseska**, Dean of the Faculty of Economics, Prilep, Member;
12. **Prof. Dr. sc. Gabor Kovacs**, President of the Association of European Police Colleges and Dean of the Faculty of Law Enforcement at the University of Public Service Ludovika, Budapest, Hungary;
13. **Prof. Dr. sc. Ivan Vidolov**, Rector of the Academy of the Ministry of Interior, Sofia, Bulgaria;
14. **Prof. Dr. sc. Simo Mihov**, Vice-Rector of the Academy of the Ministry of Interior, Sofia, Bulgaria;
15. **Prof. Dr. sc. Zoran Đurđević**, Rector of the University of Criminal Investigation and Police Studies, Belgrade, Serbia;
16. **Prof. Dr. sc. Aleksandar Bošković**, Vice-Rector of the University of Criminal Investigation and Police Studies, Belgrade, Serbia;
17. **Prof. Dr. sc. Darko Trifunović**, Director of the Institute for National and International Security, Belgrade, Serbia;
18. **Prof. Dr. sc. Damir Piplica**, Dean of the Faculty of Forensic Sciences, University of Split, Croatia;
19. **Prof. Dr. sc. Ante Orlović**, Dean of the University of Applied Sciences in Criminal Investigation and Public Security;
20. **Assoc. Prof. Dr. sc. Mirzo Selimić**, Director of College “Centre for Business Studies”, Kiseljak, Bosnia and Herzegovina;
21. **Assoc. Prof. Ing. Stanislav Šišulák**, Rector of the Academy of Police Forces, Bratislava, Slovakia;
22. **Prof. Dr. sc. Branislav Kováčik**, Dean of the Faculty of Political Science and International Relations, Matej Bel University, Banská Bystrica, Slovakia;
23. **Prof. Dr. sc. Rastislav Kazanský**, Faculty of Political Science and International Relations, Matej Bel University, Banská Bystrica, Slovakia;
24. **Prof. Dr. sc. Julius Piwowarski**, Rector of Aperiion University, Kraków, Poland;
25. **Prof. Dr. sc. Igor Bernik**, Dean of the Faculty of Criminal Justice and Security, Ljubljana, University of Maribor, Slovenia.

ORGANIZING COMMITTEE

1. **Ass. Prof. Dr. sc. Stevco Jolakoski, Chairman**
2. **Ass. Prof. Dr. sc. Kire Babanoski, Deputy Chairman**
3. **Prof. Dr. sc. Marina Malis Sazdovska**
4. **Prof. Dr. sc. Snezana Mojsoska**
5. **Prof. Dr. sc. Nikola Dujovski**

SECRETARIAT

1. **M.Sc. Elena Dimovska**
2. **M.Sc. Olivera Trajanova Gjorgjieski**
3. **Kemal Rushid**

CONTENT

HYBRID WARFARE AS COGNITIVE CONFLICT: STRATEGIC COMMUNICATIONS, ARTIFICIAL INTELLIGENCE, AND THE FUTURE OF INFLUENCE OPERATIONS.....	9
Miroslav Mitrović.....	9
METAMORPHOSES OF RUSSOPHILIA AS A HYBRID THREAT TOOL IN SLOVAKIA'S POLITICAL DEVELOPMENT: FROM HISTORICAL SENTIMENT TO 21ST-CENTURY SECURITY RISKS	35
Rastislav Kazanský.....	35
COGNITIVE-INFORMATIONAL WARFARE IN RUSSIA'S INTERFERENCE IN ELECTIONS.....	43
Iulian Chifu.....	43
HYBRID ACTIVITY INDICATOR SYSTEM (HAIS): A SECURITY SCIENCE FRAMEWORK FOR EARLY DETECTION AND ASSESSMENT	61
Darko Trifunović	61
AUTHENTICATION TO SURVEILLANCE: LEGAL AND SECURITY IMPLICATIONS OF BIOMETRIC TECHNOLOGIES IN HYBRID THREAT ENVIRONMENTS	81
Emir Kremić	81
Haris Muminovic	81
Edita Dapo	81
CYBER-OFFENDERS AND TRADITIONAL OFFENDERS IN THE ERA OF SMART TECHNOLOGIES: CRIMINOLOGICAL PERSPECTIVES ON MOTIVATION, SOCIAL NETWORKS, AND RISK FACTORS	103
Olga Tešović.....	103
DIGITAL REALITY AND CYBER CRISIS FORMATION: COMMUNICATION UNDER UNCERTAINTY IN POLITICALLY SALIENT PUBLIC DIGITAL SYSTEMS	123
Ivan Ristov	123
INFORMATION SECURITY AND CHALLENGES IN PUBLIC INSTITUTIONS OF THE REPUBLIC OF NORTH MACEDONIA	147
Pullumb Abdullai.....	147
BUILDING INSTITUTIONAL RESILIENCE AGAINST HYBRID THREATS IN THE REPUBLIC OF NORTH MACEDONIA: IMPLEMENTATION, CHALLENGES, AND PERSPECTIVES.....	171
Besnik Emini	171
Igor Gjoreski.....	171

CONCEPTUALIZING RADICALIZATION IN SECURITY SCIENCES: THEORETICAL APPROACHES, METHODOLOGICAL CRITIQUES AND MUU IDEOLOGY.....	189
Frosina Tashevska-Remenski	189
Zlate Dimovski.....	189
DIGITAL EXCLUSION AS A RESILIENCE GAP: AGE-STRATIFIED POLITICAL DISENGAGEMENT ACROSS WESTERN BALKAN ENLARGEMENT ECONOMIES.....	207
Besa Shahini	207
MIGRATION, SECURITY AND LOCAL IDENTITY: THE CASE OF DUBROVNIK.....	227
Marijana Musladin	227
SECONDARY LABOUR MIGRATION, LABOUR GOVERNANCE AND EMERGING VULNERABILITIES IN NORTH MACEDONIA.....	245
Sandra Tomovska	245
Stojanka Mircheva	245
MEASURING CITIZENS' PERCEPTIONS OF SECURITY: CONSTRUCTING THE 2025 SECURITY INDEX IN NORTH MACEDONIA	265
Marjan Nikolovski	265
Cane Mojanoski	265
Aleksandar Ivanov	265
Kire Babanoski	265
PREVENTING VIOLENT EXTREMISM: MULTI-LEVEL GOVERNANCE APPROACHES FOR BUILDING RESILIENT AND SUSTAINABLE SOCIETIES.....	283
Agus Demirovski	283
BATTERED WOMAN SYNDROME AND SPOUSAL HOMICIDE: A QUALITATIVE STUDY OF INCARCERATED WOMEN WITH HISTORIES OF INTIMATE PARTNER VIOLENCE.....	301
Dragana Batic	301
Dimitrinka Jordanova Peshevska.....	301
CYBERBULLYING AS A PREDICTOR FOR DEPRESSION AND SUICIDAL BEHAVIORS IN PRIMARY AND SECONDARY SCHOOL STUDENTS.....	313
Dimitrinka Jordanova Peshevska	313
Dragana Batic	313

HYBRID WARFARE AS COGNITIVE CONFLICT: STRATEGIC COMMUNICATIONS, ARTIFICIAL INTELLIGENCE, AND THE FUTURE OF INFLUENCE OPERATIONS

Miroslav Mitrović

Faculty of Project and Innovation Management, Educons University

mitrovicmm@gmail.com

ORCID: 0000-0001-7541-8565

Abstract

This article examines how the convergence of hybrid warfare, strategic communications, and artificial intelligence reshapes contemporary influence operations as a form of cognitive conflict. It situates this question within the transformation of warfare from predominantly kinetic confrontation toward strategic competition over perception, legitimacy, trust, and decision-making. Methodologically, the paper applies theoretical synthesis and a qualitative comparative design based on analytically selected cases: Moldova, Germany, and Ukraine-related Russian influence operations. The analysis shows that contemporary hybrid campaigns operate through the interaction of narratives, pseudo-local and cross-platform infrastructures, emotional triggers, institutional vulnerabilities, and wider political objectives. It also finds that artificial intelligence primarily enables faster, more scalable, adaptive, and personalised influence activity, while not independently guaranteeing strategic effect. The article contributes by proposing an integrated framework for analysing cognitive battlespaces, linking structural vulnerability, actor infrastructure, cognitive-emotional mechanisms, amplification dynamics, effect assessment, and resilience capacity. It concludes that cognitive security is becoming an important component of national defence and democratic resilience, especially in the Security 2030 context of smart technologies, resilient societies, and emerging digital risks.

Keywords: *hybrid warfare; cognitive conflict; strategic communications; artificial intelligence; influence operations.*

1. INTRODUCTION

The character of warfare is changing in ways that challenge traditional assumptions about power, coercion, and strategic success. Although kinetic force remains important, contemporary conflict increasingly unfolds in the cognitive domain, where influence over perception, interpretation, and collective behaviour can generate strategic effects without large-scale military confrontation (Deppe & Schaal, 2024). In this context, hybrid warfare has moved beyond the simple combination of military and non-military means and has become an adaptive mode of conflict aimed at shaping political outcomes, institutional legitimacy, and social cohesion. Strategic communications and artificial intelligence now stand at the centre of this evolution, enabling influence operations that are faster, more targeted, more scalable, and more difficult to attribute (Mumford & Carlucci, 2023; Burdette et al., 2026).

This transformation is particularly important in environments marked by social fragmentation, digital dependence, institutional distrust, and intensified geopolitical competition. Under such conditions, the battlespace no longer ends with territory, forces, or infrastructure, but extends into cognition, where narratives, symbols, emotions, and mediated interpretations of reality shape both public behaviour and elite decision-making. Hybrid warfare therefore increasingly operates through the management of perception and legitimacy rather than through overt escalation alone (Deppe & Schaal, 2024; Mitrović, 2025a).

Against this background, this article examines the convergence of hybrid warfare, strategic communications, and artificial intelligence in contemporary influence operations. Its central argument is that hybrid warfare is increasingly evolving into a form of cognitive conflict in which legitimacy, perception management, and behavioural influence become more decisive than kinetic superiority alone. On that basis, the article advances the analytical claim that the integration of artificial intelligence into strategic communications and influence operations can increase the capacity of hybrid actors to shape domestic discourse, erode institutional trust, and constrain the strategic autonomy of target states, particularly smaller and institutionally vulnerable ones (Mitrović, 2025a; Rauchfleisch et al., 2025; Zhou et al., 2025). This argument speaks directly to the Security 2030 agenda because AI-enabled influence operations show how smart technologies can simultaneously support resilience and create new vulnerabilities in democratic and security systems.

Existing scholarship has examined hybrid warfare as a mode of ambiguous and multidomain coercion (Mumford & Carlucci, 2023; Genini, 2025), cognitive warfare as a struggle over perception, emotion, trust and decision-making (Deppe & Schaal, 2024; Zilincik, 2026; Ćosić et al., 2026), and artificial intelligence as an accelerator of disinformation and influence operations (Stockwell, 2024; Bentzen, 2025; Hanhijärvi, 2026). However, these strands of literature are often treated separately. The hybrid warfare literature tends to emphasise strategic ambiguity and cross-domain coercion; the cognitive warfare literature focuses on perception, emotion and trust; while the AI and disinformation literature highlights automation, synthetic media and scalable manipulation. This article addresses the gap between these debates by analysing how AI-enabled strategic communications connect the operational logic of hybrid warfare with the cognitive mechanisms through which influence produces political and security effects.

The article addresses the following research question: how does the integration of artificial intelligence into strategic communications and influence operations reshape hybrid warfare as a form of cognitive conflict? It argues that artificial intelligence changes the operational conditions of influence activity by accelerating production, localisation, adaptation, and distribution, while strategic outcomes remain dependent on political intent, organisational infrastructure, and target vulnerability. The article contributes to the literature by connecting hybrid warfare, cognitive conflict, and AI-enabled strategic communications within a single analytical framework for studying cognitive battlespaces.

Methodologically, the article employs a theoretically informed qualitative comparative design, using analytically selected cases rather than a representative sample (George & Bennett, 2005). The analysis proceeds in two steps. First, it synthesises scholarship on hybrid warfare, cognitive conflict, strategic communications and AI-enabled influence operations in order to clarify the conceptual relationship between these fields (Mumford & Carlucci, 2023; Deppe & Schaal, 2024; Nilsson & Ekman, 2024; Hanhijärvi, 2026). Second, it compares three illustrative cases (Moldova, Germany and Ukraine-related Russian influence operations) selected for their variation in target

environment and operational logic. Moldova illustrates small-state vulnerability under layered hybrid pressure (OSCE/ODIHR, 2024; OSCE/ODIHR, 2026); Germany shows the use of AI-enabled pseudo-local media infrastructure in a major European democracy (Pfaller, Schmid & Sadeghi, 2025); and Ukraine-related operations demonstrate narrative laundering and cross-platform coordination in a transnational wartime information environment (Smith et al., 2026). The cases reveal recurring mechanisms through which AI-enabled strategic communications can shape cognitive vulnerability, institutional trust and strategic autonomy.

The article makes three contributions. First, it clarifies the relationship between hybrid warfare and cognitive conflict without reducing all strategic communication to warfare. Second, it explains how artificial intelligence changes the speed, scale and adaptability of influence operations while remaining dependent on political intent, organisational infrastructure and target vulnerabilities. Third, it proposes an integrated framework for analysing cognitive battlespaces by linking structural vulnerability, actor infrastructure, cognitive-emotional mechanisms, amplification dynamics, effect assessment and resilience capacity.

The article proceeds as follows. It first examines hybrid warfare beyond the kinetic paradigm and then analyses cognitive conflict as an emerging battlespace. It next discusses strategic communications and influence operations as instruments of hybrid power, before assessing the role of artificial intelligence in cognitive warfare. The comparative section draws insights from Moldova, Germany and Ukraine-related Russian influence operations. The final sections propose an integrated framework for analysing cognitive battlespaces and discuss selected policy and defence implications.

2. METHODOLOGY AND RESEARCH DESIGN

This article applies a theoretically informed qualitative comparative design. The research is not based on a statistically representative sample, but on analytically selected cases that illustrate different configurations of AI-enabled influence operations and cognitive vulnerability. The cases were selected according to four criteria: documented exposure to recent hybrid or influence activity; variation in target environment; variation in operational infrastructure; and relevance for the Security 2030 agenda. Moldova was selected because it illustrates small-state vulnerability under layered hybrid pressure; Germany because it demonstrates AI-enabled pseudo-local media infrastructure in a major European democracy; and Ukraine-related Russian influence operations because they reveal narrative laundering, cross-platform amplification, and attribution challenges in a wartime information environment.

The analysis proceeds in two stages. First, the article synthesises recent scholarship on hybrid warfare, cognitive conflict, strategic communications, and AI-enabled influence operations in order to clarify the conceptual relationship between these fields. Second, it compares the selected cases across four analytical dimensions: structural vulnerability of the target environment, actor and infrastructure configuration, cognitive-emotional mechanisms, and observable limits or effects of influence activity. The purpose of comparison is therefore analytical rather than statistical: it identifies recurring mechanisms and structured similarities, but it does not estimate prevalence or claim direct causal effects between AI-enabled content and political outcomes.

The author's analytical perspective is that the appropriate unit of analysis is not an isolated message or synthetic artefact, but the wider influence ecosystem in which narratives, actors, digital infrastructures, emotional triggers, and target vulnerabilities interact. This perspective guides the comparison and underpins the integrated framework developed later in the article.

The study is subject to several methodological limitations. It relies primarily on open-source academic literature, institutional reports, platform disclosures, and investigative reporting. Although these sources provide a valuable basis for comparative analysis, their availability and level of detail vary across cases and may reflect differences in monitoring and detection capacities rather than differences in the actual intensity of influence activity. Attribution in the context of influence operations also remains inherently probabilistic, since technical, behavioural, and contextual indicators can increase analytical confidence but cannot entirely eliminate uncertainty regarding actor intent, coordination, and operational reach. Moreover, because exposure to manipulative content cannot be reliably linked to changes in political behaviour or strategic outcomes, the study limits its conclusions to the identification of recurring mechanisms, patterns of vulnerability, and operational capacities rather than claiming direct causal effects on electoral or broader strategic outcomes.

3. HYBRID WARFARE BEYOND THE KINETIC PARADIGM

Hybrid warfare should be understood not as the disappearance of kinetic force, but as a reordering of the instruments through which strategic effects are produced. In earlier military paradigms, armed force functioned as the principal means of breaking an adversary's capacity and will. In hybrid conflict, by contrast, kinetic action is increasingly subordinated to broader political, informational, economic, cyber, legal, and psychological lines of effort. The objective is less immediate battlefield destruction than the gradual shaping of the adversary's strategic environment, the exploitation of internal vulnerabilities, and the manipulation of thresholds of response. Hybrid warfare is therefore best understood as competition conducted under conditions of ambiguity, in which coercion is dispersed across domains and calibrated to remain below the threshold of decisive military retaliation (Mumford & Carlucci, 2023). This interpretation is also consistent with the view that contemporary hybrid conflict seeks control rather than conquest and targets societal vulnerabilities, cognitive landscapes, and institutional legitimacy as much as physical battlespace (Mitrović, 2025a).

What distinguishes hybrid warfare from a merely expanded conception of irregular conflict is not the coexistence of military and non-military tools, but their synchronised use within a unified strategic design. Ambiguity is not incidental to hybrid warfare; it is one of its defining operational principles. Actors rely on deniability, calibrated escalation, and blurred attribution in order to advance political objectives without provoking a clear and unified counter-response. In that sense, hybrid warfare occupies the space between grand strategy and tactical execution. It is neither a wholly new type of war nor merely a rhetorical label for everything short of invasion. Rather, it is a historically specific mode of operational art suited to an environment defined by strategic competition, legal constraint, technological acceleration, and the high costs of overt interstate war (Mumford & Carlucci, 2023).

This move beyond the kinetic paradigm has also been institutionalised in Euro-Atlantic security doctrine. NATO's 2022 Strategic Concept frames the contemporary

environment as one in which strategic competitors interfere in democratic processes, conduct malicious cyber activities, spread disinformation, instrumentalise migration, manipulate energy dependencies, and employ economic coercion, often through hybrid tactics and proxies (NATO, 2022). NATO's more recent guidance on hybrid threats likewise defines them as combinations of military and non-military, covert and overt means designed to blur the line between war and peace and destabilise target societies (NATO, 2026). In parallel, the EU Strategic Compass presents the strategic environment as one in which non-military tools, including data and technology standards, are increasingly weaponised in geopolitical competition (European Union, 2022).

This shift has practical consequences. Hybrid warfare can no longer be treated as a concern of the armed forces or intelligence services alone. It increasingly engages the wider architecture of state and societal resilience. Research on Sweden and Norway shows that prolonged hybrid pressure pushes security thinking away from a narrow focus on armed confrontation and towards broader models of total defence, civil preparedness, and cross-sectoral coordination (Ljungkvist, 2024; Skare & Jore, 2024). For smaller states especially, this means that critical infrastructure, information environments, public trust, and societal preparedness become integral elements of national defence.

For that reason, the phrase 'beyond the kinetic paradigm' should not be mistaken for a post-military condition. Force remains present, but often as background pressure, latent escalation potential, or selective support to non-kinetic campaigns. The decisive shift lies in the fact that strategic effects are increasingly sought before, around, and beneath open armed confrontation. In this sense, the non-kinetic turn is not a replacement for war, but a reordering of how coercion is organised and how political objectives are pursued in an era of permanent strategic competition (Genini, 2025; Mumford & Carlucci, 2023).

4. COGNITIVE CONFLICT AS THE EMERGING CORE OF FUTURE WARFARE

If hybrid warfare has moved beyond a primarily kinetic paradigm, the next analytical step is to identify what increasingly occupies the strategic centre of gravity in that expanded battlespace. A growing body of literature suggests that the answer lies in the cognitive domain. This does not mean that material power, force, or coercion has disappeared. Rather, it means that the shaping of perception, interpretation, emotion, trust, and decision-making is becoming increasingly decisive in determining whether broader political and military objectives can be achieved. In this sense, cognitive conflict can be understood as the struggle to influence how individuals, groups, and institutions perceive reality, assign legitimacy, process uncertainty, and choose among strategic alternatives (Deppe & Schaal, 2024; Zilincik, 2026).

Cognitive conflict should not be reduced to a synonym for information warfare, disinformation, or propaganda. Those remain important instruments, but the cognitive perspective shifts analytical attention from the message itself to its effects on interpretation and behaviour. Recent scholarship suggests that the concept is useful precisely because it highlights the growing importance of targeting mental processes through which trust, meaning, and political judgment are formed, while also warning against conceptual overstretch if all persuasive activity is labelled as warfare (Deppe & Schaal, 2024; Zilincik, 2026). The most productive approach, therefore, is to treat cognitive conflict not as a replacement for existing concepts, but as an integrative lens for understanding the strategic targeting of cognition, affect, and collective interpretation.

This article therefore uses the concept of cognitive conflict in a restricted sense, in line with recent warnings against conceptual overstretch in the cognitive warfare literature (Deppe & Schaal, 2024; Zilincik, 2026). It does not treat all persuasion, public diplomacy, public affairs, political campaigning, or ordinary strategic communication as cognitive warfare. Rather, the concept refers to coordinated, adversarial, and strategically designed efforts to manipulate perception, trust, emotion, and decision-making in pursuit of political or security objectives.

Under this lens, the strategic object of conflict extends beyond territory, military formations, or infrastructure to include the social and psychological processes that mediate political action. Cognitive conflict seeks to influence whether citizens trust institutions, whether elites interpret crises accurately, whether societies remain cohesive under pressure, and whether political communities retain the ability to distinguish manipulation from legitimate persuasion. This is why the literature increasingly links cognitive conflict to epistemic security, emotional regulation, and social resilience (Casino, 2025; Crum et al., 2025).

Technological change has intensified these dynamics. Recent research shows that manipulated content can be analysed through machine-assisted examination of language, readability, emotion, and psychological cues, indicating that the cognitive battlespace is becoming increasingly measurable and scalable. At the same time, influence operations are also becoming more adaptive and more systematically optimised (Gaeta et al., 2025). Yet the most important point is not simply that technology expands capability, but that it enables more sustained attempts to shape how audiences interpret events, evaluate credibility, and respond emotionally.

Emotion is especially important in this context. Contemporary cognitive warfare often operates not primarily through rational persuasion, but through affective pathways, symbolic framing, and emotionally resonant narratives that shape fear, anger, humiliation, solidarity, or confidence. Recent work on defensive cognitive warfare is particularly useful here, as it argues that cognition, belief, trust, emotion, and decision-making are deliberate targets of contemporary influence strategies (Ćosić et al., 2026). For small and vulnerable states, these processes are especially consequential. Limited institutional depth, fragile media ecosystems, political polarisation, and dependence on external digital platforms increase exposure to manipulation that can weaken strategic autonomy without open military escalation.

For that reason, cognitive conflict should be treated not as a marginal supplement to hybrid warfare, but as one of its most consequential dimensions. Its significance lies in the fact that future conflict will depend not only on who can destroy more effectively, but also on who can more successfully shape perception, structure emotional climates, influence trust, and orient collective interpretation. The next step, therefore, is to examine how much of this influence is operationalised in practice: through strategic communications and influence operations.

4.1 Comparative Indicators of Cognitive Vulnerability and AI-Enabled Influence Operations

One of the main methodological difficulties in analysing cognitive conflict is that direct causal effects on political outcomes are rarely demonstrable with high confidence. For that reason, a more defensible approach is to compare observable indicators across cases rather than infer strategic effectiveness from isolated incidents. Particularly useful indicators include the number of confirmed AI-enabled disinformation incidents, the infrastructural scale of influence operations, levels of institutional trust in the target

environment, and platform-reported disruptions of covert influence networks (Stockwell, 2024).

The comparative evidence suggests that AI-enabled influence operations should not be evaluated solely by counting viral items. Their strategic significance depends on the broader organisational and institutional environment in which they operate. CETaS identified sixteen confirmed viral cases of AI-enabled disinformation or deepfakes during the 2024 United Kingdom general election, and eleven comparable cases in the 2024 European Union and French electoral processes (Stockwell, 2024). These figures do not prove overwhelming technological disruption, but they do show that AI-enabled manipulation has entered democratic contestation in visible and measurable form.

The German case broadens the picture. In early 2025, NewsGuard and Correctiv identified a network of 102 AI-generated German-language websites linked to a pro-Russian disinformation operation targeting the German electoral environment (Pfaller, Schmid & Sadeghi, 2025). This matters because it shifts attention from single pieces of content to the infrastructural depth of manipulation. The key issue is not merely whether one deepfake goes viral, but whether an actor possesses the sustained capacity to generate, distribute, reinforce, and normalise manipulative narratives across an interconnected information ecosystem.

Institutional trust is another essential variable. According to the OECD (2025), average trust in national governments across surveyed countries stands at 39 %, while trust in national parliaments stands at 37 %, compared with 54 % for courts and judicial systems. Standard Eurobarometer 103 similarly reports 52 % trust in the European Union, but only 36 % in national governments in spring 2025 (European Commission, 2025c). These figures matter because influence operations do not operate in a social vacuum. Their strategic potential depends partly on pre-existing levels of legitimacy, alienation, and informational receptivity. Where trust is fragile, hostile narratives may encounter less resistance and gain broader resonance.

A further distinction concerns the difference between attempted influence, detected misuse, and demonstrated political effect. OpenAI reported disrupting five covert influence operations by May 2024 and more than twenty deceptive operations and networks by October 2024, while also noting that these did not show evidence of significant audience reach or engagement through its services (OpenAI, 2024a, 2024b). This distinction is central to the argument of this article. Artificial intelligence expands capability in terms of speed, scale, adaptability, and content production, but it does not automatically guarantee strategic success. The available evidence suggests that AI more readily enhances operational capacity than it secures political effect.

Taken together, these findings support a more nuanced understanding of cognitive conflict in hybrid warfare. The importance of AI-enabled influence operations lies not in the simplistic claim that they directly determine political outcomes, but in the way they interact with structural vulnerabilities, especially low trust, social fragmentation, and uneven institutional legitimacy. For that reason, the comparative framework used here proceeds across four dimensions: the scale of confirmed AI-enabled content, the operational infrastructure, the trust profile of the target environment, and the available evidence of effect or limitation. In this sense, Table 1 does more than summarise empirical material; it provides a compact analytical model for understanding how cognitive vulnerability and AI-enabled influence capacity intersect in contemporary hybrid conflict.

Table 1. Comparative indicators of cognitive vulnerability and AI-enabled influence operations

Case / indicator	Exact comparative data	Analytical significance
United Kingdom general election, 2024	16 confirmed viral cases of AI-enabled disinformation or deepfakes	Demonstrates visible but limited confirmed scale of AI-enabled manipulation in an electoral setting (Stockwell, 2024).
European Union and French elections, 2024	11 confirmed cases of AI-enabled disinformation	Useful for comparing density and visibility of AI-enabled influence across parallel democratic contexts (Stockwell, 2024).
Germany snap election, 2025	102 AI-generated German-language websites linked to a pro-Russian influence network	Indicates infrastructural breadth and sustained narrative capability rather than isolated content production (Pfaller, Schmid & Sadeghi, 2025).
OpenAI threat reporting, May 2024	5 covert influence operations disrupted	Shows verified attempts to operationalize generative AI in coordinated deceptive activity (OpenAI, 2024a).
OpenAI threat reporting, October 2024	More than 20 operations and deceptive networks disrupted since the beginning of 2024	Suggests acceleration in attempted misuse, while also illustrating the role of platform countermeasures (OpenAI, 2024b).
OECD institutional trust benchmark, 2025	39% trust in national government; 37% trust in national parliament; 54% trust in courts/judiciary	Provides a comparative baseline for assessing cognitive and institutional resilience (OECD, 2025).
Eurobarometer, Spring 2025	52% trust in the European Union; 36% trust in national governments	Reveals trust asymmetry across levels of governance that may shape narrative receptivity and strategic influence (European Commission, 2025c).

Source: Author's synthesis based on selected contemporary sources.

This table advances the central argument by showing that AI-enabled influence capacity becomes strategically relevant when it intersects with institutional vulnerability, low trust, and weak resilience. It also supports the distinction between attempted misuse, detected reach, and demonstrated political effect, which is essential for avoiding technological determinism.

5. STRATEGIC COMMUNICATIONS AND INFLUENCE OPERATIONS IN HYBRID CONTEXTS

Strategic communications in hybrid conflict should not be reduced to the mere dissemination of messages. In a stronger analytical sense, it refers to the deliberate alignment of narratives, actions, symbols, timing, and institutional behaviour in order to shape perceptions, legitimacy, and decision-making. What matters is not only what an actor says, but whether communication is integrated with broader political and operational conduct. This is why recent work on wartime communication emphasises coherence between message, messenger, and action rather than treating communication as an autonomous media layer detached from events on the ground (Nilsson & Ekman, 2024).

From this perspective, influence operations are among the principal operational expressions of strategic communications in the hybrid environment. Recent research defines them as coordinated activities designed to shape the perceptions or behaviour of target audiences in order to achieve political objectives, while foreign information manipulation and interference is understood as intentional, coordinated, and manipulative activity capable of undermining values, procedures, and political processes (Smith et al., 2026). These definitions are important because they move the discussion beyond vague references to propaganda and frame influence operations as structured, purposive, and politically consequential interventions in the information environment.

In hybrid contexts, strategic communications performs both offensive and defensive functions. Offensively, it can frame events, assign blame, create emotional salience, and normalise preferred interpretations of reality before audiences form independent judgments. Defensively, it can expose manipulation, strengthen interpretive resilience, and reduce the strategic utility of ambiguity for the adversary. NATO's current approach to countering information threats reflects this logic. Its Information Environment Assessment capability combines people, repeatable processes, and technological tools in order to analyse actors, behaviour, content, degree, and effect, while also emphasising pre-bunking through open and anticipatory communication (NATO, 2025). Strategic communications is thus no longer merely reactive public affairs, but part of the active shaping of the information environment.

This broader understanding also helps explain why strategic communications can no longer be treated as a niche function confined to ministries, public diplomacy offices, or military spokespersons. Hybrid influence operations target the wider ecology of public meaning: legacy media, digital platforms, civil society, expert networks, diasporas, and transnational advocacy circuits (Bolt & Lange-Ionatamishvili, 2026). The European External Action Service reflects this institutional adjustment by combining communication campaigns, partner engagement, and resilience-building across multiple regions in response to foreign information manipulation and interference (European External Action Service, 2025). This is analytically important because it shows that strategic communications in hybrid settings increasingly operates as a resilience and governance function, not simply as state messaging.

The connection between strategic communications and influence operations becomes even clearer when attribution is considered. In many hybrid campaigns, effectiveness depends not only on spreading content, but on concealing origin, laundering narratives through intermediaries, and exploiting uncertainty about authorship and intent. Recent work on attribution proposes multi-source approaches that combine technical, behavioural, contextual, and legal-ethical evidence in order to identify hostile actors with

greater confidence (Smith et al., 2026). This matters because countering influence operations is not only a matter of rebuttal. It also requires credible evidentiary standards that can justify proportionate legal, diplomatic, and political responses.

Artificial intelligence intensifies these dynamics by increasing the speed, adaptability, and scale of influence activity. Influence networks linked to authoritarian actors have already used AI tools to strengthen content production, audience adaptation, and operational persistence, while democratic states are being forced to combine technological responses with more traditional countermeasures in order to protect the integrity of the information environment (Hanhijärvi, 2026). AI does not overturn the logic of strategic communications, but it makes narrative production, localisation, and cross-platform manipulation more scalable and more difficult to counter through fragmented institutional responses.

For that reason, strategic communications should be treated as a core component of hybrid power rather than an auxiliary instrument. It links public diplomacy, political signalling, information operations, resilience-building, attribution, and narrative competition into a single strategic field. Its effectiveness depends not only on message design, but also on credibility, institutional consistency, timing, and audience understanding. In hybrid conflict, the struggle is therefore not merely to speak louder than the adversary, but to shape the conditions under which audiences decide what is believable, who is legitimate, and which futures appear politically plausible (Nilsson & Ekman, 2024; NATO, 2025; Smith et al., 2026).

6. ARTIFICIAL INTELLIGENCE AND THE OPERATIONAL ACCELERATION OF COGNITIVE WARFARE

Artificial intelligence should be understood in this context not as an autonomous substitute for strategy, but as an operational accelerator that expands the speed, scale, adaptability, and granularity of cognitive and influence operations. Its significance lies less in replacing traditional instruments of strategic communications than in enhancing their efficiency and operational reach. AI systems enable rapid content generation, audience segmentation, predictive targeting, translation, synthetic voice and image production, automated amplification, and large-scale analysis of sentiment and behavioural patterns. As a result, they increase the ability of both state and non-state actors to shape the informational and cognitive environment in ways that were previously more costly, slower, and more labour-intensive (Burdette et al., 2026).

Recent strategic analysis supports this interpretation. AI has the potential to alter several key military and informational competitions, including those between quantity and quality, hiding and finding, centralised and decentralised command, and cyber offence and defence (Burdette et al., 2026). Although such assessments are not limited to influence operations, their implications are directly relevant to the cognitive domain. If AI lowers the cost of mass, improves the processing of large volumes of information, and supports more adaptive forms of coordination, then it also changes the conditions under which cognitive warfare is planned and executed.

The specific relevance of AI to information manipulation has become increasingly visible in European security analysis. Generative AI makes deceptive influence campaigns cheaper, faster, and potentially more persuasive by facilitating automated content creation, synthetic audio, deepfake video, and deceptive websites (Bentzen, 2025). Related research also shows that influence networks linked to China and Russia have used both foreign and

domestic AI tools to reinforce existing operational tactics, especially through greater scale, cost-efficiency, and personalisation (Hanhijärvi, 2026). The significance of this development lies not simply in the introduction of new tools, but in the changing tempo and economics of manipulation.

This operational acceleration is especially important because the decisive variable in cognitive warfare is often not the mere existence of hostile narratives, but the ability to adapt them continuously to different audiences, emotional climates, and political moments. AI enables hostile actors to test, refine, localise, and reproduce narratives at a speed difficult to match through purely human-centred systems. It also lowers the barrier to entry for complex manipulation. Activities that once required larger teams, stronger linguistic capacity, and greater production resources can now be conducted more quickly and at lower cost. In that sense, AI changes the scale, persistence, and rhythm of cognitive competition (Bentzen, 2025; Hanhijärvi, 2026).

At the same time, the available evidence cautions against technological determinism. OpenAI's threat reporting is useful in this regard. In May 2024, the company reported disrupting five covert influence operations that attempted to use its models in support of deceptive activity, while also stating that these campaigns did not appear to have achieved significant audience reach or engagement through its services (OpenAI, 2024a). By October 2024, it had disrupted more than twenty deceptive operations and networks, again without evidence that technical access alone translated into meaningful political effect (OpenAI, 2024b). These findings suggest that AI-enabled capability and strategic success should not be conflated.

This point is reinforced by the comparative evidence discussed earlier. Studies of the 2024 United Kingdom, European Union, and French electoral cycles identified confirmed cases of AI-enabled influence activity, but did not establish that such content decisively altered electoral outcomes (Stockwell, 2024). The strategic importance of AI therefore lies less in any single deepfake or synthetic post than in the cumulative transformation of the information environment. AI enables more experimentation, more adaptation, more plausible mimicry, and more persistent intervention, but its effectiveness still depends on the vulnerabilities of the target environment, including low trust, political polarisation, weak institutional credibility, and the availability of amplification channels. Table 2 summarises the principal AI-enabled capabilities relevant to cognitive warfare together with the main constraints on their strategic effectiveness.

Table 2. AI-enabled capabilities, strategic advantages and limitations in cognitive warfare

AI-enabled capability	Relevance for cognitive warfare	Strategic advantage	Limitation / vulnerability	Defensive implication
Rapid content generation	Enables large-scale production of narratives, talking points, and manipulative content	Speed, volume, and low-cost output	Quality inconsistencies, repetition, and detectability	Content authentication, monitoring, and rapid verification
Audience segmentation	Supports tailoring of messages to specific social, political, or emotional groups	Personalisation and behavioural precision	Requires reliable user data and accurate audience insight	Data protection, platform oversight, and segmentation transparency
Multilingual translation and localisation	Expands cross-border influence operations and narrative adaptation	Wider reach across linguistic environments	Cultural mismatch, contextual errors, and lower persuasive credibility	Local fact-checking, contextual analysis, and culturally grounded rebuttal
Synthetic audio and video	Increases deceptive capacity and emotional impact through deepfakes	Strong persuasive or shock effect	Growing forensic detection and often short-lived impact	Deepfake detection, rapid response protocols, and media literacy
Automated amplification	Enhances repetition and visibility across platforms	Low-cost scaling and sustained narrative exposure	Platform takedowns, bot detection, and network disruption	Platform coordination and takedown capacity
Sentiment and behavioural analysis	Helps adapt messages to emotional and political conditions in real time	Greater precision and dynamic adjustment	Depends on access to large datasets and accurate interpretation	Early-warning systems and resilience mapping
Continuous narrative testing	Enables rapid refinement of framing, tone, and persuasive messaging	Adaptability and iterative optimisation	Many tested narratives fail to gain traction	Pre-bunking and proactive strategic communication

Source: Author's synthesis based on Bentzen (2025), Burdette et al. (2026), Hanhijärvi (2026), OpenAI (2024a, 2024b), and Stockwell (2024).

The comparison indicates that the most consequential AI contribution is not any single synthetic artefact, but the combination of rapid production, localisation, automated amplification, and continuous testing. Its effect remains limited where attribution, platform disruption, content authentication, and resilient public trust reduce the conversion of operational capacity into strategic influence.

For that reason, AI should be conceptualised as an enabling layer within cognitive warfare rather than as its independent driver. It amplifies pre-existing strategic logics of deception, persuasion, narrative competition, and behavioural targeting. It does not eliminate the need for political intent, operational design, organisational infrastructure, or favourable social conditions. Nor does it remove the continuing importance of credibility, timing, and strategic coherence. What it does is make manipulation faster, cheaper, more personalised, and more continuous. In the context of hybrid warfare, that is a significant strategic development: the struggle over perception and legitimacy can now be pursued with greater speed and persistence while remaining below traditional thresholds of armed confrontation.

This is precisely why future defence planning must move beyond narrow understandings of cyber resilience or media literacy alone. AI-enabled influence systems should be treated as a distinct security challenge requiring governance frameworks, strategic awareness, and institutional responses capable of addressing the interaction between technological acceleration and cognitive vulnerability (Bentzen, 2025; OpenAI, 2024a; OpenAI, 2024b).

7. COMPARATIVE QUALITATIVE INSIGHTS FROM RECENT HYBRID CAMPAIGNS

Recent hybrid campaigns suggest that cognitive conflict is best understood not through isolated incidents, but through recurring operational patterns across different strategic environments. For that reason, this section adopts a selective qualitative comparison rather than an exhaustive case catalogue. The selected cases vary in scale and geopolitical setting: Moldova illustrates a small-state environment exposed to dense hybrid pressure around electoral processes; Germany shows how AI-enabled and pseudo-localised disinformation infrastructures can be deployed against a major European democracy; and Ukraine-related Russian influence operations demonstrate how narrative laundering, cross-platform coordination, and attribution challenges operate in a transnational information battlespace. Taken together, these cases allow comparison not only of what influence actors say, but also of how they organise dissemination, exploit vulnerabilities, and attempt to shape cognitive and political outcomes.

7.1 Moldova: Hybrid Pressure in a Small-State Electoral Environment

Moldova offers one of the clearest recent examples of how hybrid pressure can converge on a politically vulnerable and geopolitically exposed state. International observation findings on the 2024 presidential election and constitutional referendum described an environment marked by illicit foreign interference and active disinformation campaigns that affected the integrity of the process (OSCE/ODIHR, 2024). Subsequent reporting on the 2025 parliamentary elections showed that these pressures had not diminished, but had instead become more explicit and operationally layered, including foreign interference, illicit financing, cyberattacks, widespread disinformation, and extensive vote-buying schemes (OSCE/ODIHR, 2026).

The Moldovan case is especially revealing because it blurs any simple distinction

between cognitive and material interference. The same operational environment included coordinated support networks, financial inducement, cyber pressure, disinformation, and law-enforcement action against organised structures linked to foreign-backed political influence (OSCE/ODIHR, 2026). In this setting, the cognitive dimension of hybrid warfare was not merely symbolic. It was embedded in a broader coercive architecture aimed at shaping both belief and behaviour. Moldova therefore illustrates how hybrid influence can constrain the strategic autonomy of a small state by combining narrative manipulation with organisational, financial, and cyber mechanisms.

7.2 Germany: AI-Enabled Infrastructure and Pseudo-Local Credibility

If Moldova highlights layered coercion in a small-state environment, the German case illustrates a different but equally important pattern: the industrialisation of influence through AI-enabled infrastructure and simulated local legitimacy. In early 2025, NewsGuard and Correctiv identified a network of 102 AI-generated German-language websites spreading pro-Russian narratives ahead of Germany's snap elections (Pfaller, Schmid & Sadeghi, 2025; Schmid, 2025). The importance of this case lies not only in the number of false claims involved, but in the way the operation mimicked authentic local and national media identities in order to make hostile narratives appear socially native rather than externally imposed.

This case is significant because it shows that the central issue in cognitive conflict is not only falsehood, but the manipulation of interpretive authority. By imitating local journalistic forms, the network sought to intervene in the conditions under which audiences decide what appears plausible, legitimate, and socially anchored. Germany also illustrates how the reach of manipulative narratives is conditioned by broader social factors such as affective polarisation, populism, and the political centrality of the target environment (Okholm, 2025). In this sense, the strategic value of the operation lay in infrastructural depth, pseudo-local credibility, and the capacity to contaminate the information environment over time.

7.3 Ukraine-Related Influence Operations: Narrative Laundering and Multi-Channel Coordination

A third pattern emerges from recent Russian influence operations related to Ukraine. Here the key issue is not simply disinformation content, but the coordinated management of dissemination pathways. Recent reporting identifies a range of tactics including sanctions workarounds, coordinated Telegram amplification, source laundering, and cross-platform circulation of fabricated stories (Smith et al., 2026). One particularly revealing example involved a false story claiming that Georgian and Ukrainian soldiers had attacked one another. According to the report, the story circulated across multiple Kremlin-linked outlets, with publication timing and source substitution suggesting coordinated dissemination rather than organic spread (Smith et al., 2026).

This case matters because it demonstrates that influence operations increasingly depend on shaping epistemic pathways rather than merely inserting false claims into public discourse. The objective is to create the appearance that a narrative emerged organically, was independently confirmed, or gained legitimacy through repetition across different channels. This is analytically important because it shows that cognitive conflict depends not only on persuasive content, but also on the strategic organisation of amplification, source laundering, and apparent verification. Ukraine-related influence operations therefore illustrate the importance of coordination, infrastructure reuse, and evidentiary convergence in contemporary hybrid campaigns. Table 3 compares the principal patterns

across the selected hybrid campaigns.

Table 3. Comparative patterns across recent hybrid influence campaigns

Case	Target environment	Main instruments	Cognitive objective	Operational mechanism	Observed effect / strategic significance
Moldova (2024–2025)	Small, politically vulnerable, geopolitically exposed state	Disinformation, illicit financing, vote-buying, cyberattacks	Erode trust, influence voter behaviour, weaken strategic autonomy	Layered hybrid pressure combining cognitive and material interference	Demonstrates how cognitive conflict in small-state environments is embedded in broader coercive architecture (OSCE/ODIHR, 2024, 2026).
Germany (2025)	Large European democracy with broader regional influence	AI-generated pseudo-local websites, election disinformation, narrative mimicry	Manipulate interpretive authority, undermine institutional trust, normalise hostile narratives	Scaled narrative infrastructure simulating local credibility and journalistic legitimacy	Illustrates the industrialisation of influence through AI-enabled infrastructure and pseudo-local legitimacy (Pfaller, Schmid, & Sadeghi, 2025; Schmid, 2025).
Ukraine-related influence operations	Transnational information battlespace	Telegram coordination, source laundering, synchronized reposting, fabricated stories	Shape legitimacy, simulate verification, influence cross-border perception	Cross-platform repetition, source substitution, and timing coordination	Highlights the importance of epistemic pathway management rather than false content alone (Smith et al., 2026).

Source: Author’s synthesis based on selected contemporary sources.

This table shows that different cases share a common strategic logic despite variation in target environments: influence becomes consequential when narratives are embedded in organisational, financial, technological, and platform infrastructures. This supports the article’s claim that cognitive conflict is relational and depends on the interaction between actor capability and target vulnerability.

7.4 Comparative Insights

Despite their differences, the selected cases reveal several common patterns. First, recent hybrid campaigns are layered rather than modular. Moldova demonstrates the

interaction of disinformation, illicit financing, cyber incidents, and vote-buying; Germany shows the infrastructural depth and pseudo-localisation of AI-enabled manipulation; and the Ukraine-related cases reveal the importance of source laundering, repetition, and multi-channel dissemination. In each case, the strategic object is not only the circulation of falsehood, but the manipulation of the cognitive conditions under which audiences interpret events and assign legitimacy.

Second, the effectiveness of hybrid influence is strongly conditioned by the target environment. Manipulative techniques do not produce identical effects across different settings. Their impact depends on existing political fractures, institutional trust, media ecology, and the geopolitical value of the target. In that sense, cognitive conflict is always relational: it depends as much on the vulnerabilities of the target as on the capabilities of the actor conducting the campaign.

Third, the comparative evidence reinforces the broader argument of this article that cognitive conflict is not reducible to digital deception alone. It is a broader strategic effort to shape perception, erode trust, and constrain decision-making through the interaction of communicative, institutional, financial, and technological mechanisms. The future of influence operations therefore lies less in isolated spectacular manipulations than in the cumulative interaction of narrative control, infrastructural coordination, target-specific adaptation, and societal vulnerability. It is this interaction that increasingly defines hybrid warfare as cognitive conflict.

The author's synthesis of the three cases leads to a more specific analytical proposition: the decisive variable in contemporary cognitive conflict is not the sophistication of AI-generated content by itself, but the campaign's capacity to convert technological scale into socially credible and politically consequential influence. That conversion depends on three conditions operating together: an infrastructure capable of sustained dissemination, pre-existing vulnerabilities in the target environment, and mechanisms that translate exposure into changes in trust, interpretation, or institutional behaviour. Where one of these conditions is absent, operational activity may remain visible without producing strategic effect.

8. TOWARD AN INTEGRATED FRAMEWORK FOR ANALYSING COGNITIVE BATTLESPACES

The preceding sections have shown that contemporary hybrid warfare increasingly operates through the interaction of cognitive mechanisms, strategic communications, and AI-enabled influence systems. What remains necessary, however, is an analytical framework capable of connecting these elements in a way that is both conceptually rigorous and operationally useful.

Existing approaches often illuminate only one dimension of the problem. Some focus on narratives and media effects, others on technological infrastructure, attribution, or societal resilience. Yet the contemporary cognitive battlespace is constituted precisely by the interaction of these dimensions. For that reason, this article proposes an integrated framework for analysing cognitive battlespaces that combines structural vulnerability mapping, actor and infrastructure analysis, cognitive-emotional assessment, amplification dynamics, effect evaluation, and resilience capacity.

The need for such a framework is well supported by recent literature. Current scholarship shows that the concept of cognitive warfare is analytically useful only when treated as more than a rhetorical extension of information warfare, while also warning against conceptual overstretch.

At the same time, frameworks such as DISARM and VIGINUM's OpenCTI-oriented guidance underline the importance of standardised documentation, interoperability, and the distinction between technical imputation and political attribution. Recent work on foreign information manipulation and interference likewise reinforces the need for structured and transparent analytical models, particularly where direct impact remains difficult to prove conclusively (Deppe & Schaal, 2024; Zilincik, 2026; DISARM Foundation, n.d.; VIGINUM, 2025; Sicurella & Morača, 2025).

The framework proposed here builds on these insights but also extends them. It is broadly compatible with the Clark–Heuer integrative logic developed in recent security analytics research, which combines target-centric systemic mapping with safeguards against cognitive bias and premature closure. The present adaptation extends that logic into the cognitive battlespace by integrating strategic communications, emotional climate, AI-enabled adaptation, and effect assessment into a single analytical model (Mitrović, 2025b).

Its original contribution lies in combining variables that are usually assessed separately and in placing effect assessment between amplification and resilience. This sequencing prevents analysts from treating detected activity, audience reach, and strategic success as equivalent outcomes. It also makes the framework suitable for comparative application across small states, major democracies, and transnational wartime information environments.

The proposed framework consists of six analytically connected layers. Table 4 summarises the integrated framework for analysing cognitive battlespaces.

Table 4. Integrated framework for analysing cognitive battlespaces

Analytical layer	Core question	Main indicators	Typical sources	Expected analytical output
Structural vulnerability mapping	What makes the target susceptible?	Institutional trust, political polarisation, media ecology, platform dependence, civic fragmentation	Surveys, institutional data, V-Dem, OECD, Eurobarometer, national reports	Vulnerability profile
Actor and infrastructure identification	Who produces, launders, amplifies, and legitimizes influence?	State actors, proxies, pseudo-media, platforms, coordinated networks, technological infrastructure	Attribution reports, OSINT, DISARM, VIGINUM, platform investigations	Actor-infrastructure map
Cognitive-emotional mechanism analysis	How are perception, trust, identity, and emotion targeted?	Fear, anger, humiliation, legitimacy frames, identity triggers, symbolic narratives	Narrative analysis, content analysis, emotional coding, discourse analysis	Mechanism map
Amplification and adaptation dynamics	How does the campaign scale and	Algorithmic amplification, AI-assisted content	Platform analysis, technical reports, campaign tracing,	Diffusion and adaptation profile

Analytical layer	Core question	Main indicators	Typical sources	Expected analytical output
	evolve?	production, localisation, cross-platform migration, narrative adjustment	digital forensics	
Effect assessment	What effects are actually observable?	Interpretive, emotional, behavioural, institutional, and strategic effects	Election reports, trust data, media analytics, case studies, audience research	Graded effect assessment
Resilience and governance capacity	How does the target detect, absorb, attribute, and respond?	Strategic communication capacity, fact-checking, attribution mechanisms, regulation, cyber hygiene, public trust reserves	Policy documents, institutional reports, resilience assessments, strategic communication reviews	Resilience profile

Source: Author's synthesis based on OECD (2025), European Commission (2025c), V-Dem Institute (2025), and national/institutional reports.

This table operationalises the article's conceptual contribution by translating cognitive conflict into a sequence of observable analytical tasks. It can guide future empirical research by specifying which indicators should be collected and how they can be connected to effect assessment and resilience evaluation.

First, structural vulnerability mapping. The point of departure must be the target environment rather than the message alone. This includes the political, military, economic, social, informational, and infrastructural characteristics of the target, but also institutional trust, media ecology, affective polarisation, platform dependence, and civic fragmentation. The same manipulative technique may produce very different effects depending on the receptivity and resilience of the target context.

Second, actor and infrastructure identification. Cognitive campaigns should be analysed not only in terms of content, but in terms of who produces, launders, amplifies, and legitimises it. This requires attention to state actors, proxies, pseudo-media outlets, platform affordances, reusable digital infrastructures, and influence networks. At this level, the key question is how hostile information activity is organised and sustained over time.

Third, cognitive-emotional mechanism analysis. The central issue here is how a campaign seeks to shape interpretation, trust, identity, emotion, and decision-making. This means going beyond narrative frames alone and identifying the emotional triggers and symbolic forms through which campaigns generate fear, anger, humiliation, fatigue, or false confidence. Emotional dynamics should therefore be treated as a central component of analysis, not a secondary supplement.

Fourth, amplification and adaptation dynamics. At this level, the framework examines how campaigns gain scale, speed, and persistence. Relevant mechanisms include

algorithmic curation, automated amplification, multilingual localisation, AI-assisted content production, source laundering, cross-platform migration, and iterative narrative adjustment. The key analytical point is that AI alters the tempo and cost structure of manipulation, even when it does not determine outcomes on its own.

Fifth, effect assessment. A robust framework must distinguish between attempted influence, operational reach, and strategic effect. Visibility does not automatically imply transformation. Effect assessment should therefore proceed across several registers: interpretive effects, emotional effects, behavioural effects, institutional effects, and strategic effects. This distinction is essential if analysts are to avoid exaggerating technological significance or conflating exposure with success.

Sixth, resilience and governance capacity. The final layer evaluates the target's ability to detect, absorb, attribute, counter, and recover from cognitive attacks. This includes strategic communications capabilities, public trust reserves, civil-society participation, investigative capacity, regulatory tools, cyber hygiene, and coordinated institutional response mechanisms. Resilience, in this sense, is not merely corrective communication, but the broader capacity to preserve democratic legitimacy and shared interpretive stability under pressure.

Taken together, these six layers form a cyclical rather than linear model. Structural vulnerabilities condition actor strategy; actor infrastructures shape the available mechanisms of influence; cognitive-emotional mechanisms are amplified through organisational and technological adaptation; effects emerge unevenly across audiences and institutions; and resilience capacity then influences the success or failure of future campaigns.

The cognitive battlespace should therefore be understood as a dynamic system of interaction between hostile intervention and target-specific vulnerability, rather than as a static arena populated by messages and counter-messages.

The analytical value of this framework lies in its ability to connect macro-level strategic competition with meso-level influence infrastructures and micro-level cognitive effects. It also provides a more disciplined basis for empirical work.

Instead of asking only whether false content circulated, analysts can ask: what vulnerabilities were exploited, through which actors and infrastructures did the campaign travel, what emotional and interpretive mechanisms did it activate, how was it amplified, what effects were observable, and which resilience mechanisms mitigated or failed to mitigate those effects? These questions make the analysis of cognitive battlespaces more transparent, more comparable, and more useful for policy and defence planning.

Finally, this framework reinforces the article's central claim that hybrid warfare is increasingly becoming a form of cognitive conflict. This does not mean that kinetic force or territorial control has disappeared. It means that strategic success increasingly depends on the ability to shape perception, legitimacy, trust, and interpretation before, during, and after open confrontation. The proposed framework is intended as one step toward analysing that reality with greater conceptual discipline and operational usefulness.

9. POLICY AND DEFENCE IMPLICATIONS

If hybrid warfare is increasingly becoming a form of cognitive conflict, then policy responses cannot remain limited to conventional defence planning, reactive fact-checking, or ad hoc crisis communication. The preceding analysis suggests that the central challenge is systemic: hostile actors exploit structural vulnerabilities, cognitive-emotional

mechanisms, digital infrastructures, and AI-enabled adaptation in order to shape interpretation, weaken trust, and constrain strategic autonomy.

The appropriate response is therefore not a single instrument, but an integrated architecture linking deterrence, resilience, regulation, attribution, strategic communications, and AI governance. In practical terms, this requires permanent monitoring of the information environment, clear attribution procedures, rapid inter-institutional coordination, and proactive communication before hostile narratives become socially normalised.

A first implication is that cognitive resilience should be treated as a pillar of national security rather than a peripheral communication concern. NATO's current approach to information threats frames them as intentional, harmful, manipulative, and coordinated activities designed to create confusion, sow division, destabilise societies, and influence perceptions and behaviours. It also emphasises the need for continuous monitoring, structured assessment, and anticipatory response capacities rather than episodic reactions to individual disinformation incidents (NATO, 2024). In practical terms, this means that cognitive defence must become a permanent component of security planning.

A second implication is that resilience must be organised as a whole-of-society function. Recent Euro-Atlantic approaches show that countering hybrid threats requires the involvement not only of military and intelligence institutions, but also of election authorities, media regulators, public broadcasters, schools, civil society, digital platforms, and research communities (NATO, 2026; European Commission, 2025b). This matters because hybrid campaigns target the wider ecology of public meaning and institutional legitimacy, not only isolated technical systems. The defence of democratic resilience therefore depends on broad coordination across state and societal actors.

A third implication concerns election protection and platform accountability. Recent European regulatory developments, including the Digital Services Act framework and related guidelines on electoral processes, show a movement away from voluntary platform commitments and towards auditable obligations concerning transparency, risk mitigation, incident response, and generative AI-related safeguards (European Commission, 2025a; European Commission, 2026). This is significant because it introduces a more structured accountability architecture for the protection of civic discourse and electoral integrity.

A fourth implication is that AI-enabled influence systems must be governed as risk-bearing sociotechnical systems rather than treated as neutral innovation tools. European regulatory initiatives and the NIST AI Risk Management Framework both point in this direction by emphasising responsible development, deployment, and risk mitigation for AI systems that shape communication, recommendation, and content generation (European Commission, 2024; European Commission, n.d.; NIST, 2024). For cognitive security, the key issue is not technology alone, but the interaction between technological capability and societal vulnerability.

A fifth implication concerns attribution, deterrence, and cost imposition. Democratic resilience cannot rely solely on exposure and rebuttal. It must also include credible mechanisms for attribution, sanctioning, diplomatic signalling, network disruption, and the reduction of hostile actors' operational sustainability. Recent European responses to foreign information manipulation and interference are important precisely because they seek to make hostile activity more costly and less sustainable over time (European External Action Service, 2026a; European External Action Service, 2026b). In

the cognitive domain, deterrence therefore depends not only on counter-narratives, but also on whether manipulative campaigns encounter friction, exposure, and strategic cost.

These implications are especially important for small and vulnerable states. Where institutional depth is limited, political polarisation is high, and external dependency is pronounced, hybrid campaigns can combine disinformation, illicit financing, cyber incidents, and broader political pressure in particularly damaging ways.

Under such conditions, a purely platform-centric or communication-centric response is insufficient. What is needed is a more integrated model of democratic total resilience combining trust measurement, protection of electoral infrastructure, strategic communications capacities, legal instruments against covert foreign influence, cyber defence, and rapid inter-institutional coordination.

The broader lesson is that defence planning in the age of cognitive conflict cannot be limited to protecting territory, networks, and infrastructure in a narrow sense. It must also protect the conditions under which a political community can interpret reality, sustain trust, make decisions, and preserve legitimacy under pressure.

This does not require the securitisation of all political communication or the treatment of dissent as hybrid aggression. It does, however, require recognition that the struggle over perception, legitimacy, and shared factual foundations has become a core theatre of strategic competition.

10. CONCLUSION

This article has argued that the future security environment toward 2030 will be shaped not only by the diffusion of smart technologies, but also by their integration into cognitive and influence operations. Hybrid warfare is therefore increasingly evolving into a form of cognitive conflict in which the decisive terrain extends beyond territory, force deployment, or physical destruction to include perception, legitimacy, trust, and decision-making. The argument does not imply the disappearance of kinetic power, but rather a reordering of strategic priorities. Military force remains relevant, yet it is increasingly embedded within broader campaigns that seek to shape the interpretive environment before, during, and after overt confrontation.

The article has made three principal contributions. First, it has clarified that cognitive conflict should not be treated as a loose synonym for information warfare, propaganda, or disinformation. Rather, it should be understood as the strategic targeting of the processes through which individuals and institutions interpret reality, assign legitimacy, regulate emotion, and make political judgments (Deppe & Schaal, 2024; Zilincik, 2026). Second, it has shown that artificial intelligence is best understood as an enabling and accelerating layer rather than an autonomous driver of strategic success. AI expands the speed, scale, adaptability, and personalisation of manipulation, but its effects remain conditioned by institutional trust, political polarisation, media ecology, and target vulnerability (Burdette et al., 2026; Hanhijärvi, 2026). Third, through the analysis of Moldova, Germany, and Ukraine-related Russian influence operations, it has demonstrated that recent hybrid campaigns converge in a common strategic logic: they shape the cognitive environment of the target through the interaction of narratives, infrastructures, coordination, and societal vulnerability.

A further contribution has been methodological. The integrated framework proposed here connects structural vulnerability mapping, actor and infrastructure identification, cognitive-emotional mechanism analysis, amplification dynamics, effect

assessment, and resilience capacity within a single analytical model. Its value lies in reducing conceptual vagueness, improving comparability across cases, and supporting more disciplined analysis of cognitive battlespaces. At the same time, the framework should be treated as a basis for future empirical testing rather than as a completed measurement model. Further research should apply it to a larger number of cases, combine open-source evidence with platform data and audience research, and examine more precisely how exposure to influence activity translates into interpretive, behavioural, or institutional effects.

The broader implication is that future defence and security planning must move beyond reactive, platform-centric, or narrowly technical approaches. States and democratic institutions can no longer assume that protecting infrastructure, rebutting falsehoods, or responding after a crisis has begun will be sufficient. What is required is a more integrated conception of security that includes cognitive resilience, institutional trust, strategic communications capacity, AI governance, and coordinated mechanisms for attribution, deterrence, and societal preparedness. This is especially important for smaller and more vulnerable states, whose limited strategic depth makes them particularly exposed to hybrid campaigns operating below the threshold of armed conflict.

At the same time, analytical discipline remains essential. Not every political crisis is externally engineered, not every digital controversy is a coordinated information operation, and not every persuasive act, public diplomacy effort, or political campaign should be labelled as cognitive warfare. For precisely that reason, the study of cognitive conflict must remain conceptually precise, empirically grounded, and normatively cautious. Yet the overall conclusion remains clear: cognitive security is becoming a central pillar of contemporary defence planning. The future of conflict will be shaped not only by who can strike harder, but by who can better influence how reality is perceived, how legitimacy is assigned, and how collective decisions are made under conditions of uncertainty and manipulation. In that sense, the struggle for cognitive advantage is increasingly becoming one of the defining characteristics of hybrid warfare.

DECLARATION OF GENERATIVE AI AND AI-ASSISTED TECHNOLOGIES

During the preparation of this manuscript, the author used generative AI tools for language refinement, structural review, and consistency checking. The author subsequently reviewed, verified, and edited all content, references, interpretations, and conclusions and takes full responsibility for the final manuscript.

REFERENCES

- Bentzen, N. (2025, December 11). Information manipulation in the age of generative artificial intelligence (PE 779.259). European Parliamentary Research Service. [https://www.europarl.europa.eu/RegData/etudes/BRIE/2025/779259/EPRS_BRI\(2025\)779259_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/BRIE/2025/779259/EPRS_BRI(2025)779259_EN.pdf)
- Bolt, N., & Lange-Ionatamishvili, E. (2026, February 6). *The NextGen information environment*. NATO Strategic Communications Centre of Excellence. <https://stratcomcoe.org/publications/the-nextgen-information-environment/339>
- Burdette, Z., Phillips, D., Heim, J. L., Geist, E., Frelinger, D. R., Heitzenrater, C., & Mueller, K. P. (2026). *How artificial intelligence could reshape four essential*

- competitions in future warfare* (RR-A4316-1). RAND Corporation. <https://doi.org/10.7249/RR-A4316-1>
- Casino, F. (2025). Unveiling the multifaceted concept of cognitive security: Trends, perspectives, and future challenges. *Technology in Society*, 83, Article 102956. <https://doi.org/10.1016/j.techsoc.2025.102956>
- Ćosić, K., Popović, S., & Wiederhold, B. K. (2026). Emotionally based strategic communications as a new tool in defensive cognitive warfare. *Frontiers in Psychology*, 17, Article 1751406. <https://doi.org/10.3389/fpsyg.2026.1751406>
- Crum, J., Allred, A., Bostrom, S., Doherty, E., McLain, M., Spencer, C., Niemeyer, R. E., Anderson, A. P., Tossell, C., Čeko, M., & Hirshfield, L. (2025). Understanding the neurocognitive mechanisms of cognitive security. *Neuroscience & Biobehavioral Reviews*, 179, Article 106448. <https://doi.org/10.1016/j.neubiorev.2025.106448>
- 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>
- DISARM Foundation. (n.d.). *DISARM red framework*. <https://www.disarm.foundation/framework>
- European Commission. (2024, August 1). *AI Act enters into force*. https://commission.europa.eu/news-and-media/news/ai-act-enters-force-2024-08-01_en
- European Commission. (2025a, February 13). *The code of conduct on disinformation*. <https://digital-strategy.ec.europa.eu/en/library/code-conduct-disinformation>
- European Commission. (2025b, November 12). *Defending democratic values in the digital age*. <https://digital-strategy.ec.europa.eu/en/policies/democracy-digital>
- European Commission. (2025c, May 27). *Standard Eurobarometer 103 – Spring 2025*. <https://europa.eu/eurobarometer/surveys/detail/3372>
- European Commission. (2026, March 18). *Guidelines under the Digital Services Act*. <https://digital-strategy.ec.europa.eu/en/policies/dsa-guidelines>
- European Commission. (n.d.). *Artificial intelligence*. https://commission.europa.eu/topics/artificial-intelligence_en
- European External Action Service. (2025). *EEAS strategic communication task forces*. https://www.eeas.europa.eu/eeas/eeas-strategic-communication-task-forces_en
- European External Action Service. (2026a). *Information integrity and countering foreign information manipulation and interference (FIMI)*. https://www.eeas.europa.eu/eeas/information-integrity-and-countering-foreign-information-manipulation-interference-fimi_en
- European External Action Service. (2026b, March 12). *4th EEAS annual report on foreign information manipulation and interference threats*. https://www.eeas.europa.eu/eeas/4th-eeas-annual-report-foreign-information-manipulation-and-interference-threats_en
- European Union. (2022, March 24). *A strategic compass for security and defence: For a European Union that protects its citizens, values and interests and contributes to international peace and security*. European External Action Service. https://www.eeas.europa.eu/eeas/strategic-compass-security-and-defence-0_en
- Gaeta, A., Loia, V., Lorusso, A., Orciuoli, F., & Pascuzzo, A. (2025). Computational analysis of information disorder in cognitive warfare. *Online Social Networks and Media*, 48, Article 100322. <https://doi.org/10.1016/j.osnem.2025.100322>
- Genini, D. (2025). Countering hybrid threats: How NATO must adapt (again) after the war

- in Ukraine. *New Perspectives*, 33(2), 122–149. <https://doi.org/10.1177/2336825X251322719>
- George, A. L., & Bennett, A. (2005). *Case studies and theory development in the social sciences*. MIT Press.
- Hanhijärvi, H. (2026). Artificial intelligence and foreign information manipulation: Chinese and Russian approaches (Hybrid CoE Paper 29). European Centre of Excellence for Countering Hybrid Threats. <https://www.hybridcoe.fi/publications/artificial-intelligence-and-foreign-information-manipulation-chinese-and-russian-approaches/>
- Ljungkvist, K. (2024). The military-strategic rationality of hybrid warfare: Everyday total defence under strategic non-peace in the case of Sweden. *European Journal of International Security*, 9(4), 533–552. <https://doi.org/10.1017/eis.2024.18>
- Mitrović, M. (2025a). *Invisible fronts: Hybrid warfare and the future of conflict*. Independently published.
- Mitrović, M. (2025b). Hybrid threat assessment through security analytics: A Clark–Heuer integrative model. *Security Science Journal*, 6(2), 195–217. <https://doi.org/10.37458/ssj.6.2.9>
- Mumford, A., & Carlucci, P. (2023). Hybrid warfare: The continuation of ambiguity by other means. *European Journal of International Security*, 8(2), 231–249. <https://doi.org/10.1017/eis.2022.19>
- NATO. (2022, June 29). *NATO 2022 strategic concept*. <https://www.nato.int/content/dam/nato/webready/documents/publications-and-reports/strategic-concepts/2022/290622-strategic-concept.pdf>
- NATO. (2024, October 18). *NATO's approach to counter information threats*. <https://www.nato.int/en/about-us/official-texts-and-resources/official-texts/2024/10/18/natos-approach-to-counter-information-threats>
- NATO. (2025, February 3). *NATO's approach to counter information threats*. <https://www.nato.int/en/what-we-do/wider-activities/natos-approach-to-counter-information-threats>
- NATO. (2026, January 29). *Countering hybrid threats*. <https://www.nato.int/en/what-we-do/deterrence-and-defence/countering-hybrid-threats>
- Nilsson, P.-E., & Ekman, I. (2024). “Be brave like Ukraine”: Strategic communication and the mediatization of war. *National Security and the Future*, 25(1), 21–64. <https://doi.org/10.37458/nstf.25.1.2>
- NIST. (2024). *Artificial intelligence risk management framework: Generative artificial intelligence profile (AI RMF 1.0)*. National Institute of Standards and Technology.
- OECD. (2025). *Government at a glance 2025*. OECD Publishing.
- Okholm, C. S. (2025). Conditions of subversive reach: Comparing societal and strategic factors for Russian propaganda outlets’ reach among Western European fringe communities. *European Journal of International Security*, 1–23. <https://doi.org/10.1017/eis.2025.10018>
- OpenAI. (2024a, May 30). *Disrupting deceptive uses of AI by covert influence operations*. <https://openai.com/index/disrupting-deceptive-uses-of-ai-by-covert-influence-operations/>
- OpenAI. (2024b, October 9). *An update on disrupting deceptive uses of AI*. <https://openai.com/global-affairs/an-update-on-disrupting-deceptive-uses-of-ai/>
- OSCE/ODIHR. (2024). *Republic of Moldova – Presidential election and constitutional referendum, 20 October 2024: Statement of preliminary findings and conclusions*.

- OSCE Office for Democratic Institutions and Human Rights.
- OSCE/ODIHR. (2026, February 18). Republic of Moldova – Parliamentary elections, 28 September 2025: Final report. OSCE Office for Democratic Institutions and Human Rights. <https://odihr.osce.org/odihr/662251>
- Pfaller, L., Schmid, R., & Sadeghi, M. (2025, January 23). Russian propagandist turns his sights to German election. *NewsGuard*. <https://www.newsguardtech.com/special-reports/russian-propagandist-john-mark-dougan-german-election>
- Rauchfleisch, A., Jungherr, A., & Wuttke, A. (2025). Explaining public preferences for regulating artificial intelligence in election campaigns: Evidence from the U.S. and Taiwan. *Telecommunications Policy*, 49(10), Article 103072. <https://doi.org/10.1016/j.telpol.2025.103072>
- Schmid, R. (2025, February 24). German Elections Misinformation Tracking Center. *NewsGuard*. <https://www.newsguardtech.com/special-reports/german-elections-misinformation-tracking-center/>
- Sicurella, F. G., & Morača, T. (2025). *Analysing enablers and incentives of election-related foreign information manipulation and interference: A global methodology*. International Institute for Democracy and Electoral Assistance. <https://doi.org/10.31752/idea.2025.48>
- Skare, E., & Jore, S. H. (2024). Hybrid threats in the Norwegian petroleum sector: A new category of risk problems for safety science? *Safety Science*, 180, Article 106521. <https://doi.org/10.1016/j.ssci.2024.106521>
- Smith, V., Pamment, J., Dikhtiarenko, S., Heap, B., Tsurtssumia, D., & Maunder, A. (2026). Attributing Russian information influence operations: Testing the information influence attribution framework with real-world case studies. NATO Strategic Communications Centre of Excellence and Ukrainian Centre for Strategic Communications. <https://stratcomcoe.org/publications/attributing-russian-information-influence-operations-testing-the-information-influence-attribution-framework-with-real-world-case-studies/340>
- Stockwell, S. (2024). *AI-enabled influence operations: Threat analysis of the 2024 UK and European elections*. Centre for Emerging Technology and Security, The Alan Turing Institute.
- V-Dem Institute. (2025). *V-Dem country-year dataset v15 [Data set]*. Varieties of Democracy (V-Dem) Project. <https://doi.org/10.23696/vdemds25>
- VIGINUM. (2025). *Leveraging OpenCTI to combat foreign information manipulation: Doctrine for threat knowledge capitalization*. SGDSN/VIGINUM.
- Zhou, J., Zhang, D., Zhu, J., Wang, F., & Bi, C. (2025). Speaking like humans, spreading like machines: A study on opinion manipulation by artificial-intelligence-generated content driving the internet water army on social media. *Information*, 16(10), Article 850. <https://doi.org/10.3390/info16100850>
- Zilincik, S. (2026). Cognitive warfare: Some sceptical observations from general strategic theory and affective science. *European Journal of International Security*, 1–22. <https://doi.org/10.1017/eis.2026.10057>

METAMORPHOSES OF RUSSOPHILIA AS A HYBRID THREAT TOOL IN SLOVAKIA'S POLITICAL DEVELOPMENT: FROM HISTORICAL SENTIMENT TO 21ST-CENTURY SECURITY RISKS

Rastislav Kazanský

Faculty of Political Science and International Relations, Matej Bel University, Banská
Bystrica, Slovakia
rastislav.kazansky@umb.sk

Abstract

The study analyses the changing political and security meanings of Russophilia in Slovakia. It approaches Russophilia neither as a uniform political doctrine nor as an automatic indicator of hostility toward the Euro-Atlantic order, but as a historically layered phenomenon composed of cultural memory, Slavic reciprocity, ideological inheritance and geopolitical imagination. The main argument is that, in the current environment of hybrid conflict and strategic communication, selected fragments of Russophile memory can be reactivated as instruments of political polarization, distrust and societal vulnerability. The article follows this transformation from the nineteenth-century Slovak national movement, through twentieth-century reinterpretations connected with war memory, Soviet power and the post-1989 transition, to the security context after 2014 and 2022. Special attention is devoted to the intersection of Russophilia with anti-Americanism, suspicion toward NATO and the European Union, conspiracy thinking and narratives of civilizational decline. The research is based on qualitative conceptual and historical analysis, supported by secondary scholarship and policy-oriented security literature. The conclusion is that the security relevance of Russophilia does not lie in cultural affinity itself, but in its selective use as a cognitive and emotional frame capable of weakening institutional trust, alliance cohesion and democratic resilience. Slovakia therefore offers an instructive case for studying how historical sentiment may become a resource for contemporary hybrid influence.

Keywords: Russophilia; Slovakia; hybrid threats; societal resilience; political development

1. INTRODUCTION

Russophilia in Slovakia requires careful conceptual treatment. It should not be reduced to a binary opposition between sympathy for Russian culture and political loyalty to the Russian state. In Slovak political development it has appeared as cultural memory, national self-reflection, ideological vocabulary, geopolitical imagination and, in contemporary circumstances, a potential vulnerability within the hybrid-threat environment. Its meanings have been historically contingent, politically contested and repeatedly reinterpreted.

The argument belongs to the wider security debate on hybrid threats, information security, public manipulation, psychological influence, social cohesion and societal resilience. The Security 2030 framework is relevant because the coming decade will not be shaped by digital technologies alone. Older cultural and ideological narratives will continue

to circulate through new communication infrastructures. Hybrid threats therefore cannot be understood as purely technical phenomena; they also operate through memory, identity, emotion and distrust.

The article asks how a historically intelligible cultural-political sentiment could become available for hybrid influence. This does not mean that positive references to Russian culture, literature or history should be securitized. Such a position would be analytically simplistic and politically counterproductive. The security question is more precise: under which conditions can historically rooted Russophile narratives be mobilized to undermine democratic institutions, weaken Euro-Atlantic commitments, relativize aggression and deepen societal polarization?

The discussion proceeds in four parts. It first clarifies the relationship between cultural sentiment and hybrid-threat vulnerability. It then outlines the historical roots of Slovak Russophilia in the nineteenth-century national movement and Slavophile discourse. The third part examines twentieth-century reinterpretations, including the Soviet legacy and the post-1989 Euro-Atlantic reorientation. The final part turns to the post-2014 and post-2022 environment, in which Russophilia, anti-Americanism and anti-Western narratives increasingly intersect within information conflict.

2. CONCEPTUAL APPROACH: FROM SENTIMENT TO HYBRID THREAT

Hybrid threats are usually associated with cyber operations, disinformation campaigns, covert action, proxy actors and the deliberate blurring of boundaries between war and peace. This understanding is necessary, but insufficient. Hybrid influence depends on pre-existing social and cultural conditions. It becomes effective when external narratives resonate with domestic grievances, collective memories, identity conflicts and institutional distrust. Hybrid threats therefore exploit both digital platforms and the cognitive-affective infrastructure of society.

The Russian approach to political warfare has often relied on ambiguity, plausible deniability, historical analogy and the strategic use of identity-based narratives (Galeotti, 2019; Lucas & Pomerantsev, 2016; Rid, 2020; Snegovaya, 2015). In Central and Eastern Europe, these narratives rarely appear as wholly external imports. They attach themselves to domestic traditions: anti-liberalism, anti-Americanism, nostalgia, civilizational conservatism, resentment toward elites or selective memories of liberation and Slavic reciprocity. Their effectiveness rests on this dual character: they are at once internal and external, cultural and geopolitical, emotional and strategic.

For the purposes of this analysis, Russophilia becomes a hybrid-threat tool only when it is politically activated to generate security-relevant effects. These effects may include lower trust in democratic institutions, erosion of confidence in NATO and the European Union, relativization of Russian aggression, delegitimization of support for Ukraine, or the framing of domestic opponents as agents of a hostile West. The distinction matters. Cultural interest in Russia belongs to legitimate pluralism; the instrumental use of Russophile narratives to weaken democratic resilience belongs to the security domain.

Three mechanisms are central to this transformation. The first is historical compression, whereby complex historical relationships are simplified into a binary narrative of Slavic brotherhood or Western betrayal. The second is moral inversion, through which authoritarian power is presented as defensive, traditional and peace-oriented, while liberal-democratic institutions are framed as aggressive, decadent or externally controlled.

The third is institutional displacement, in which distrust is redirected from specific policy failures toward the democratic system and Euro-Atlantic orientation as such.

3. HISTORICAL ROOTS OF SLOVAK RUSSOPHILIA

The origins of Slovak Russophilia are inseparable from the nineteenth-century national revival (Hudek, 2011). Slovak intellectuals sought political recognition and cultural emancipation within the Habsburg Monarchy, and the wider Slavic world provided an important symbolic horizon. Russia appeared to some Slovak thinkers not merely as a foreign state, but as a powerful Slavic reference point and a possible counterweight to Magyarization and German cultural dominance. This was a historical imagination formed under conditions of national insecurity.

The writings and controversies surrounding Ľudovít Štúr are central to this issue. Scholarship has shown that Štúr's late reflections on the Slavic world and Russia were shaped by disappointment with the political prospects of small Slavic nations and by the limitations of the Habsburg framework (van Duin, 2016). Later interpretations often transformed these historically situated reflections into simplified evidence of a continuous Slovak Russophile tradition. Such readings obscure the uncertainty, ambivalence and strategic constraints of nineteenth-century Slovak political thought.

The long-term significance of this period does not lie in direct continuity. Nineteenth-century Slovak intellectuals did not operate in the geopolitical universe of NATO, the European Union, contemporary international law or modern information warfare. Their references to Russia must therefore not be interpreted anachronistically. The relevant point is rather symbolic availability: later political actors could selectively reinterpret the national-revival tradition to legitimize very different agendas. Historical memory became a reservoir from which fragments could be extracted and recontextualized.

This observation is important for security analysis. Small nations often construct political identity through external reference points. When domestic institutional capacity is weak, external protectors or civilizational patrons may be imagined as guarantors of survival. Russophilia in the Slovak case must therefore be understood as part of a broader pattern of small-state political imagination. It expressed not only admiration for Russia, but also anxiety about recognition, sovereignty and survival.

4. TWENTIETH-CENTURY METAMORPHOSES

The twentieth century profoundly altered the meaning of Slovak attitudes toward Russia and later the Soviet Union. The experience of war, the memory of the Red Army's role in the defeat of Nazism, communist rule, the invasion of Czechoslovakia in 1968 and the post-1989 transition all generated competing layers of interpretation. Russophilia was no longer primarily a romantic Slavophile sentiment; it became entangled with ideology, regime legitimacy, anti-fascist memory, coercion and geopolitical dependence.

Communist rule institutionalized a particular vocabulary of Soviet friendship and socialist internationalism. This official version differed from nineteenth-century Slavophilism because it was embedded in a party-state system and supported by ideological control. At the same time, it produced ambivalent social consequences. For some citizens, Russia and the Soviet Union remained associated with liberation and modernization; for others, they symbolized domination, censorship and the suppression of sovereignty. The legacy was therefore internally divided rather than uniform.

The post-1989 transformation and Slovakia's accession to NATO and the European Union represented a strategic reorientation. Yet institutional anchoring in the West did not automatically remove older geopolitical imaginations. Euro-Atlantic integration created a formal security framework, but symbolic narratives about Russia, the West, sovereignty and national authenticity continued to circulate in political and cultural discourse. The result was a dual structure: Slovakia became institutionally Western, while parts of public discourse remained open to anti-Western interpretations. This duality produced a persistent vulnerability. Russophilia could serve different functions in different contexts: a memory of anti-fascist liberation, a residue of communist socialization, a conservative critique of liberal modernity, an expression of disappointment with domestic elites, or a geopolitical alternative to Euro-Atlantic alignment. Its political meaning depended on the actors who mobilized it and the crises in which it was activated.

5. AFTER 2014 AND 2022: RUSSOPHILIA IN THE HYBRID-THREAT ENVIRONMENT

The annexation of Crimea in 2014 and the full-scale invasion of Ukraine in 2022 changed the security meaning of pro-Russian narratives across Europe. In Slovakia, these events intensified the competition between Euro-Atlantic commitments and alternative geopolitical imaginaries. Russophilia increasingly appeared not only as cultural memory, but also as a narrative frame through which events in Ukraine, NATO policy, sanctions, energy security and domestic political conflict were interpreted.

Contemporary pro-Russian influence in Slovakia rarely relies on a simple message of admiration for Moscow. More often it combines several elements: distrust of domestic elites, criticism of liberal democracy, anti-Americanism, suspicion toward NATO, scepticism toward the European Union, conservative moral language and narratives of national sovereignty under threat. This combination allows Russophile references to function as part of a wider anti-Western repertoire rather than as a stand-alone ideology.

Public-opinion research and regional security analyses have repeatedly shown that Central European societies differ in their levels of trust toward NATO, the European Union, the United States and Russia, and that Slovakia has often displayed comparatively high vulnerability to pro-Russian or anti-Western narratives (GLOBSEC, 2024; Marušiak, 2023; Suslov et al., 2023; Šufliarská, 2023). The analytical relevance of these findings is not that Slovak society is uniformly pro-Russian. Rather, they indicate that the Slovak information environment contains social segments in which historical affinity, distrust and geopolitical ambiguity can be politically activated.

The digital environment after 2014 accelerated these processes. Social media and alternative information channels enabled rapid circulation of emotionally charged content, including claims about Western hypocrisy, NATO expansion, Ukrainian nationalism, sanctions as self-harm and Russia as a defender of traditional order. Such content is effective when it combines familiar historical motifs with contemporary grievances. It does not need to persuade everyone; it is sufficient to create doubt, fatigue, cynicism and institutional distrust.

After 2022, the war in Ukraine became a central test of Slovak societal resilience. Support for Ukraine, sanctions, energy policy, refugees and defence assistance became objects of political contestation. In this environment, Russophilia acquired a more explicitly security-relevant function when it served to relativize aggression, depict Ukraine as an instrument of Western interests, or present NATO and the United States as the

principal sources of instability. This was also the point at which Russophilia and anti-Americanism became mutually reinforcing narratives.

The overlap between Russophilia and anti-Americanism is analytically significant. The two phenomena are not identical. Russophilia may arise from cultural memory, Slavophile imagination or historical gratitude, whereas anti-Americanism may draw on critiques of capitalism, military intervention, cultural globalization or perceived hegemonic pressure. In contemporary Slovak political communication, however, the two often converge. Russia is framed as a counterweight to the United States, while the United States is portrayed as the architect of European insecurity. This binary structure simplifies complex geopolitical realities and converts foreign-policy debate into identity conflict.

6. SECURITY RISKS AND RESILIENCE IMPLICATIONS

The principal security risk of instrumentalized Russophilia lies in its effect on societal resilience. A resilient democratic society is not one without disagreement. Political conflict, critical debate and plural interpretations of history are normal and necessary. Resilience is weakened when disagreement becomes systematically organized around distrust, delegitimization and epistemic fragmentation. In such conditions, citizens no longer merely disagree about policy; they inhabit incompatible realities concerning threats, allies, aggression and responsibility.

Several risks are particularly relevant for Slovakia. First, instrumentalized Russophilia can undermine alliance cohesion by portraying NATO as a source of insecurity rather than collective defence. Second, it can reduce public support for Ukraine by presenting assistance as externally imposed or contrary to Slovak national interests. Third, it can weaken confidence in democratic institutions by framing them as controlled by foreign actors. Fourth, it can support a wider information environment in which conspiracy narratives and strategic ambiguity become normal instruments of political competition. A democratic response cannot rely primarily on censorship or moral condemnation. These instruments may confirm grievance narratives and deepen distrust. A stronger response requires conceptual differentiation, historical literacy and institutional credibility. Democratic states must distinguish legitimate cultural interest in Russia from coordinated manipulation, critical debate about Western policy from systematic anti-Western delegitimization, and pluralism from influence activity designed to weaken democratic resilience. Universities and schools have a specific responsibility in this field. They can present Slovak-Russian relations in their historical complexity: the nineteenth-century national movement, the Soviet period, the experience of 1968, post-1989 transformation and contemporary security challenges. Historical education should neither romanticize Russia nor reduce the entire relationship to hostility. Its task is to make selective memory less available for manipulation.

Strategic communication is equally important. Public institutions must explain security policy in a language that is evidence-based, intelligible and respectful of citizens' concerns. If Euro-Atlantic policy is presented only as an elite consensus, without social explanation, it remains vulnerable to populist reinterpretation. Trust cannot be improvised during a crisis; it has to be cultivated before a crisis through transparency, competence and credible public dialogue.

The Slovak case also has wider regional relevance. Central European and Western Balkan societies share several vulnerabilities: contested historical memories, small-state insecurity, distrust of elites, external influence and polarized geopolitical narratives. Slovakia demonstrates that formal integration into Euro-Atlantic structures does not automatically resolve identity-based vulnerabilities. Societal resilience must be cultivated as a long-term political and educational capacity.

7. CONCLUSION

The Slovak case shows several metamorphoses of Russophilia. In the nineteenth century it emerged as a historically understandable cultural and political sentiment connected with the national revival and the search for Slavic solidarity. In the twentieth century it was reinterpreted through war memory, communist ideology, Soviet dominance and post-1989 transformation. In the twenty-first century, especially after 2014 and 2022, selected elements of Russophile memory became available for hybrid-threat instrumentalization.

Russophilia becomes a security problem not because cultural affinity toward Russia is illegitimate, but because such affinity can be selectively transformed into a political technology of distrust. When historical memory is used to relativize aggression, delegitimize democratic institutions, weaken alliance commitments or present authoritarian power as a civilizational alternative, it enters the domain of hybrid influence. The central issue is therefore not sentiment itself, but its strategic activation.

This has practical implications for security policy. Future security challenges will not be limited to cyberattacks, artificial intelligence, autonomous systems or technological dependence. They will also include the weaponization of identity, memory and grievance. Societies preparing for the security environment of 2030 must therefore invest in technical defence as well as historical literacy, civic education, strategic communication, institutional trust and democratic conflict management.

For universities and research institutions, the principal task is conceptual clarity. Academic analysis should avoid both naive cultural romanticism and excessive securitization. It should explain how historical narratives evolve, under what conditions they become politically exploitable, and how democratic societies can respond without undermining pluralism. The study of Russophilia in Slovakia is therefore not only a question of intellectual history. It contributes to the theory and practice of societal resilience in an era of hybrid threats.

The broader lesson is that small and medium-sized democracies need geopolitical resilience in institutions and political culture alike. Alliance membership is indispensable, but it cannot substitute for the interpretive capacity of society. Citizens must be able to recognize manipulation, place historical memory within democratic debate, and resist narratives that convert cultural affinity into strategic vulnerability. In this respect, the metamorphoses of Russophilia in Slovakia offer a relevant case for contemporary security studies and international relations.

Acknowledgements

This work was supported by the Slovak Research and Development Agency under Contract No. APVV-24-0359, “Medved’ v nás”: Fenomén rusofilstva v modernom politickom vývoji Slovenska v európskych perspektívach (korene, genéza a metamorfózy od 19. storočia po súčasnosť).

This work was supported by the Slovak Research and Development Agency under Contract No. APVV-23-0570, Dejiny antiamerikanizmu v modernom vývoji Slovenska v medzinárodnom kontexte.

This work was supported by project KEGA 020UMB-4/2025, Theory of conflicts in international relations – structural and cultural causes. Compendium of educational materials for university study programs.

REFERENCES

- Galeotti, M. (2019). Russian political war: Moving beyond the hybrid. Routledge.
- GLOBSEC. (2024). GLOBSEC Trends 2024: Central and Eastern Europe Barometer. GLOBSEC. <https://www.globsec.org/what-we-do/publications/globsec-trends-2024>
- Hudek, A. (2011). Between Czechs and Hungarians: Constructing the Slovak national identity from the 19th century to the present. *History Compass*, 9(10), 730–741. <https://doi.org/10.1111/j.1478-0542.2010.00759.x>
- Lucas, E., & Pomerantsev, P. (2016). Winning the information war: Techniques and counter-strategies to Russian propaganda in Central and Eastern Europe. Center for European Policy Analysis.
- Marušiak, J. (2023). Slovakia: Emergence of an old-new pseudo-Pan-Slavism in the context of the conflict between Russia and Ukraine after 2014. In M. Suslov, M. Čejka, & V. Đorđević (Eds.), *Pan-Slavism and Slavophilia in contemporary Central and Eastern Europe*. Palgrave Macmillan. https://doi.org/10.1007/978-3-031-17875-7_16
- Rid, T. (2020). Active measures: The secret history of disinformation and political warfare. Farrar, Straus and Giroux.
- Snegovaya, M. (2015). Putin’s information warfare in Ukraine: Soviet origins of Russia’s hybrid warfare. Institute for the Study of War.
- Suslov, M., Čejka, M., & Đorđević, V. (Eds.). (2023). *Pan-Slavism and Slavophilia in contemporary Central and Eastern Europe*. Palgrave Macmillan. <https://doi.org/10.1007/978-3-031-17875-7>
- Šufliarská, P. (2023). Pandering to Pan-Slavism: How Russian influence thrives in Europeanised Slovakia. Charles University Digital Repository. <https://dspace.cuni.cz/handle/20.500.11956/182887>
- van Duin, P. (2016). Between Bohemia and Russia: Ľudovít Štúr, Slovak identity, and the controversies of the Slavic world. *Studia Politica Slovaca*, 9(2), 3–31. <https://www.ceeol.com/search/article-detail?id=478366>

COGNITIVE-INFORMATIONAL WARFARE IN RUSSIA'S INTERFERENCE IN ELECTIONS

Iulian Chifu

Conflict Prevention and Early Warning Center
keafuyul@gmail.com, ORCID 6093-8365

Abstract

The last years have demonstrated multiple approaches to interference in elections coming from malign actors, the biggest part of them associated with Russia, Russia's intelligence services and Russia's proxies from all over the world or from the targeted countries. But since the 2015-2016 influence operations and active measures in Brexit and US elections, the range of instruments has expanded dramatically, leading, just a decade later, to questioning the perspectives of democracy and legitimate elections. Our article aims to identify the trends of evolution and to bring to the forefront the strategic objectives and the means used by Russia in such an endeavour, using specifically cognitive-informational warfare and winning hearts and minds, but also extensively using psychological operations of dependence, beliefs versus rational choice, radicalisation. The cases of Ukraine, Romania and the Republic of Moldova that we studied extensively will be our references for the lessons learned.

Keywords: cognitive, informational, warfare, interference, elections, supreme weapon.

INTRODUCTION, METHODOLOGY

We have covered a long way of studies and analysis, with the Conflict Prevention and Early Warning Center I'm chairing, in the analysis of cognitive-informational warfare. Since the first identification of the types of Russian policies taking into consideration the policies in Romania and the Republic of Moldova, in 2015-2016 (Chifu & Nantoi, 2016), we examined the so-called policies to defend Russian, Russian speakers and Compatriots (Chifu & Țuțuianu, 2017), the conundrum politics, information and communication (Simons & Chifu, 2017) as well as the changing face of the war including the cognitive-informational approach and the means used by different actors (Chifu & Simons, 2023), along with a redrawing of the six generations of informational warfare (Simons & Chifu, 2025).

The most important part of the new endeavour is to examine the updated instruments for interference in the elections since several EU countries representing more than half of the EU population are due to hold national elections by direct popular vote: Estonia, Finland, France, Greece, Italy, Poland, Slovakia, and Spain (Shekhovtsov, 2026). We've done so by two major means: one direct and practical, through monitoring the full media and social media viral announcements in Romania, the Republic of Moldova and Ukraine throughout 2024-2025 (Chifu et al., 2025) and 2026, and by a study related to the reports of the intelligence services in Central-Eastern Europe that are dealing with cognitive-informational warfare in the public part, including de-classified reports about Romanian elections (Romanian Presidency, 2024).

As a result, we updated the main objectives, means and instruments used by the Russian special services to engage in cognitive-informational warfare inside the EU countries in times of electoral competition and elections, with the methods of interfering. The special case that we used was that of Maia Sandu, the President of the Republic of Moldova and character assassination instruments, with a special target on the direct involvement of the SVR, Russian intelligence service, based on its own public *communiques* (SVR, n.d.).

We used here content analysis, encyclopaedic approach, epistemological study (Chifu, 2025a) in choosing the basic characteristics of the cognitive-informational warfare and a perception projection impact approach based on the main repetitive ideas and the reach of those particular communiques aiming at the President of the Republic of Moldova and the country itself in the period 2024-April 2026. Our thesis is that the direct involvement of the SVR communiques is designed for the internal public in Russia but also to reinforce in the Republic of Moldova the groups opposing the current pro-European and pro-democratic stance of the state.

Cognitive-informational warfare. Character Assassination

The concept of cognitive warfare is far from a consensual understanding (Chifu, 2025a). Robin Burda considers that it designates various elements (Burda, 2023), David Sloss that the concept itself is often confused with notions such as information warfare and cyber warfare (Sloss, 2020). The lack of consensus on the concept is noted also by Bernal and others (Bernal et al., 2020). Definitions vary depending on the angle and field from which interdisciplinarity stems, with several distinct approaches coming from public communication, psychology, psycho-sociology, techno-sociology or cyber-sociology.

Bernal et al. consider **cognitive warfare as part of informational warfare** to be “the militarization of public opinion by an external entity with the aim of influencing public and/or government policies or destabilizing government actions and/or institutions.” (Bernal et al., 2020) Russia uses the term **informational war**, alternatively, “massive psychological manipulation of the population to destabilize the state and society, as well as coercing the state into making decisions that benefit the opposing force.” (Ministry of Defence of the Russian Federation, 2000) **Cognitive warfare** operates at the strategic level, seeking to destroy and divide target societies in peacetime through non-kinetic means (Pocheptsov, 2018) and is “a strategy that focuses on changing the way a target population thinks—and, through this, the way it acts,” (Backes & Swab, 2019) while Dahl described it as a strategy that influences the OODA (Observation-Orientation-Decision-Action) loop by “decreasing speed, accuracy, or both.” (Dahl, 1996) Cognitive operations may include specific sociocultural and linguistic parameters and aim to manage people's worldview, interests, and values, as opposed to seizing and colonizing territory or the state economy (Danyk & Briggs, 2023).

In the same debate, a distinction should be made between **cognitive and informational warfare**, even though the two work together. Tashev, Purcell, and McLaughlin argue that the cognitive dimension is the most important aspect of the information space (Tashev et al., 2019). Jason and Wintfeld see cognitive warfare as the epistemic and emotional distortion of voters as a result of information warfare targeting human decision-makers, with stimuli that can be launched through online social media and offline interpersonal networks (Andress & Winterfeld, 2013).

The cognitive domain, which includes both perception and reasoning, is under threat from the exploitation of information to influence public opinion (Tsuchiya, 2022).

The US Department of Defense is focussed on **informational warfare**, a set of actions with the aim to achieve information superiority by affecting the adversary's information, information-based processes, information systems, and computer-based networks, while defending one's own information, a process that involves protecting, manipulating, degrading, and denying information to alter knowledge at a higher level (Office of the Chairman of the Joint Chiefs of Staff, 1996). **Cognitive-informational warfare** relies on the internal vulnerabilities of democratic societies in order to function. Our own experience allows us to formulate two operational definitions:

Information warfare is an aggressive operation aimed at constructing an alternative reality - composed of real elements combined with falsehoods, lies, interpretations, opinions, etc. - and projected onto a target group that will pressure the opponent's decision-maker to alter, nuance, transform, change, or even channel the decision on a specific issue (Simons & Chifu, 2025).

Cognitive warfare is an operation of aggression against a large community, carried out over time, aimed at using the history, culture, worldview, religion, language, science and education, traditions and customs of that population, which are altered over time, along with procedures to undermine the targeted population's trust in official state institutions and society's ability to respond appropriately to changing external conditions, attacks, or disasters. The stated objective is to attack, exploit, degrade, or even destroy the way in which the targeted communities construct their own reality, their mental self-confidence, their confidence in the processes and approaches necessary for the effective functioning of groups, societies, or even nations (Chifu, 2025a). Cognitive warfare is the foundation on which the basis for societal alteration is developed to subsequently allow targeted information warfare to be effective and achieve the objective of changing decisions on a topic of interest, directly attacking the population and indirectly attacking specific targets within the adversary's established authorities (Chifu, 2025a).

Character assassination is defined as a strategic effort to discredit an individual or group target via subversive communication (Samoilenko, 2021a; Samoilenko et al., 2021). Character assassination is both used in cases of domination and subversion (Samoilenko, 2021b), so by both those in power and those in opposition, in a political landscape. Character assassination refers to the deliberate destruction of reputation through character attacks (Icks & Shiraev, 2014). Character assassination is a deliberate and sustained effort to damage the reputation or credibility of an individual. Social groups and institutions can also be the target of character assassination (Samoilenko & Suvorova, 2023). Character assassination also involves the intent to destroy the victim psychologically, not only to reduce its public support related to the perspectives to succeed in a political competition (Samoilenko et al., 2017).

Character assassination is a deliberate attempt to seriously damage the reputation, character, social status, or achievements of another person; to harm the victim psychologically and reduce its public support, but also to hurt the cause that the victim symbolizes or defends (Shiraev, 2019). It is, moreover, the deliberate destruction of an individual's reputation or even deliberate and sustained effort to damage the reputation of an individual, part of framing media strategy to impact public opinion and emotion (Samoilenko et al., 2017). It is an interdisciplinary concept coming from the rhetorical (Benoit, 2020; Keohane, 2020), sociopsychological (Shiraev, 2014), and sociocultural traditions (Samoilenko, 2021a), respectively.

That's why, in strategic and political communication we have identified character assassination as a specific type of negative campaigning (Miskolczy & Gherghina, 2024).

Our own take on the concept of character assassination comes from the angle of security studies and informational warfare. We distinguished between character attack (reputation and character attack) and character assassination (destroying the target reputation) based on the intensity of the action and achieved purpose of the hostile operation (Chifu, 2017; Chifu & Simons, 2023). Moreover, the target should be a personality involved in the decision making system of a country, institution, company. The attack dealing with character assassination is calibrated on the targeted person as well as on the public that supposedly will react and press to alter, change or distort the decision. Therefore such a weapon is difficult to build, extremely nuanced in drafting and testing and with a high level of sophistication for having a high effectiveness when used (Simons & Chifu, 2017; Chifu & Simons, 2023; Simons & Chifu, 2025). We also noted that social media (due to its capacity to reach everywhere the targeted public online) and AI (with its deepfake capacities to distort reality in a credible manner, difficult to realize at a day by day rhythm of information flux) amplified the level of impact and the capacity of action of such a weapon. It hits with a higher efficiency due to its capacity to reach the very well targeted public, and proved a high impact and a capacity to multiply and project the perception on a certain personality. The weapon belongs to the fourth generation informational warfare (Chifu, 2025b).

Main objective of Russian politics. Revisionism, perverting conservative ideology, Trump Shield, EU as an enemy

We cannot overstate the impact of cognitive informational warfare on the democracies unless we consider the relation between vulnerabilities and such activities, societal resilience (Chifu, 2022), and the impact of technology and social media on the individuals, the society, politics, and democracy (Chifu & Savu, 2020). Before looking at instruments and harms that Russia and other cognitive-informational aggressors are doing to different informational spaces and with public perceptions, we must first take care of our own weaknesses.

Social media began to operate according to its own mechanisms, according to algorithms and financial instruments, creating echo chambers of like-minded people and eliminating opposite ideas, and thereby genuine debates. It has created two new phenomena: polarization – us versus them, inside group versus outsiders, with different/opposing ideas – where opinion and choice are transformed into absolute belief, almost completely excluding rational thought and factual, logical argument; and fragmentation, that is, the division of the public sphere into countless bubbles that circulate certain ideas and to which people adhere based on their preferences – with the consecutive dependencies and addictions that develop, a fragmented view of the surrounding reality and limited exposure to opposing ideas or arguments, due to a lack of time and connection to the groups to which the social media user belongs (Chifu, 2026).

Another characteristic of social media and the way it is funded is that it fosters populism and promotes “the people versus the elites” rhetoric. Populist posts tend to be more appealing and more likely to go viral, even if they do not necessarily contain shocking or radical language. Simply because they appeal to a wider audience - typically found among the lower socioeconomic strata - populist messages are amplified virtually through social media tools (Chifu & Savu, 2020). This approach is also fuelled by a pernicious

tendency among political elites within party circles who pay lip service to meritocracy - particularly on the right - but in practice cultivate loyalty and party membership. Meritocracy and professionalism remain nothing more than forgotten words. And the end result is the marginalization of professionalism and the rejection of professional arguments, doubts, or analytical dilemmas. It is easier to appeal to people with strong verdicts and opinions that stir up emotion and attract social media attention.

Social media is also creating the system of dependency, belief, and radicalization once inside a bubble of ideas (Al-Rodhan, 2026). We have arrived at a “quantum” form of politics, in which the individual plays a major role. Today, the world is changing every day. Every person becomes a participant in political life. It is no longer classical political life. On the other hand, the individual - the person, the citizen, the voter - is just as changeable, depending on the moment and on what they have read or learned at a given time. Of course, this change is not as prevalent among strong personalities - those with independent thought and judgment - as it is among the general public, which has become, through everyone’s involvement in social media and the consumption of information primarily on Facebook, more superficial, more limited, less informed, lacking critical thinking, and more inclined to imitate those close to them - the “friends” in the social media groups they frequent. Hence, the random and even hazard stance - often contradictory - that depends on the moment of its assessment, in a state of constant flux and relativization until it is captured within a group that shapes and defines this opinion. Politicians must adapt to and cope with this reality (Chifu, 2026).

These characteristics are also what make it easiest for external forces to interfere in democratic elections. They capitalize on the educational level of the electorate that turns out and is mobilized to vote, because this electorate - drawn from the lower strata of society in terms of training, education, and concerns - is also the most susceptible to cognitive and informational influence. And here, too, the system of constructing psychological operations has taken root - first through dependence and addiction, then faith and hope fused into the “providential figure”, the guru, and the cult leader, and finally radicalization, which leads to extreme polarization, the fragmentation of politics, and the creation of enclaves with no possibility of accepting any form of flexibility or shades of rationalism (Chifu et al., 2025). Rational choice gives way to empathy, attraction, emotion, and beliefs persuasively instilled through subliminal psychological manipulation of an increasingly captive audience. Until attention wanes, another guru emerges, the orientation of a portion of the captive audience is redefined, and the process begins anew (Chifu, 2026).

The Russian Grand Strategy has been revealed by Sergey Karaganov, foreign policy advisor and ideological guru to Putin's Russia, formerly director of Russia's Foreign and Defense Policy Council and the Valdai Club. Putin claims that Karaganov is one of the authors he reads all the time, and analysts specializing in Russia consider him the main architect of Russian foreign policy. His philosophy is completely different from Russia's vision of itself, which is deeply rooted in Russianism and the country's divine mission to save the world and humanity. For Karaganov, Trump is just an American nationalist who preserves the classic elements of American messianism - uniqueness, exceptionalism, and a predominant role in the world - and therefore does not stray too far from the tradition of American presidents (Le Grand Continent, 2025).

Karaganov presented a world with America and the West versus the „emerging global majority” that Russia assumes to lead to victory, meaning the remake of a fair world for “the others”, combining anti-colonialism, anti-imperialism and its own aspirations of Great Power.

Karaganov reveals here one of the main theses of the greatness of Russia, which is no longer limited to the territory of the former Soviet empire, but goes further, with civilizational ambitions for a Russia without borders, with an integrative civilization, as the military-strategic nucleus of the „emerging and emancipated world majority” that Russia is building (Chifu, 2025c).

According to his ideas, Ukraine should be defeated through nuclear bombs, Karaganov being one of the main supporters of the fact that theatre nuclear bombs with less than 1 mT are not escalatory (Chifu, 2023). NATO should be sent to the dustbin of history, since it no longer has a purpose (other than teasing Russia and blocking its access to Grand Power status) and the EU is in the „moral failure of its elites,” which Karaganov sees as „a preparation for rapid suicide” by collective Europe, which must necessarily dissolve and disappear (Le Grand Continent, 2025). Karaganov makes a veritable apology for war, based on the theory of good war, but also of eternal war as Russia's natural state. The current war has been extremely beneficial to Putin's Russia, he says, even if it is tragic in terms of the cost in lives of Russia's youngest generations, but it has allowed a rapid break with the last remnants of Eurocentrism and Western-centrism that still existed in the politics, culture, and choices of Russia and the Russian people (Chifu, 2025c). This is the basis of Russia's cognitive-informational war.

The aim of this campaign is achieving cognitive superiority and influencing decision-making, especially through election interference, which is an integral part of the warfare planned by Russia's deputy chief of the presidential administration, Sergey Kiriyenko, responsible for those policies. Actually, Putin's Russia is following the mantra of the supreme weapon in the hands of authoritarian leaders who aim at crushing the world under their will. The main new target of Putin's Russia's aggression in the West was decided, with regional and individual specifics, designed for each target country separately, but with detailed action in each area, namely influencing strategic decision-making in the state (Chifu, 2024).

Russia is taking advantage of the democratic principles, values and especially liberties and abuses those liberties considered vulnerabilities against the democratic states and societies. This is how doubt, question marks, confusion, illogic and questioning of one's own values were introduced into Western information systems and the public sphere. The main objective was to fragment Western democratic society by attacking elements of identity and social cohesion, then introducing alternative narratives and promoting an alternative pseudo-democratic culture, using populism, extreme approaches and grievances assigned to the responsibility of observing minority rights. Cognitive warfare extends far beyond information manipulation, but is designed to shape how reality is interpreted, and conflict is perceived by governments, militaries, and civilian populations (Al-Rodhan, 2026). NATO characterizes it as a struggle for cognitive superiority: securing advantage by shaping the information environment and, through it, the mental conditions that guide thought and action (NATO Allied Command on Transformation, 2025). From here comes the attempt of Russia to achieve cognitive superiority by finding the supreme weapon to hack the minds (Simons & Chifu, 2025).

We've looked for the particular characteristics of Russia's cognitive-informational warfare also into the intelligence reports of the main intelligence services dealing with the matter. What we found is the use of intelligence operations involving cultural funding and media cooptation, in complex operations (Botts, 2025), the use of Peace Talks as a tool of manipulation, targeting the so-called Global Majority and CIS countries, with influence activities and Nobel Prize like called Tolstoi, deception in recruiting African students for

the war in Ukraine via grants for theological studies or GRU involvement in obtaining sanctioned goods (Estonian Foreign Intelligence Service, 2026) or the role of Russian intelligence Services in coordinating the disinformation war including through bribing citizens to spread disinformation and paying bloggers, journalists and public figures to conduct influence campaigns (Frolova et al., 2026). Socio-Psychological studies in the area are proving that Anyone's Mind can be Changed (Mariş, 2024) as could be seen in concrete examples in the US foreign policy, that need a revision of tools and instruments to tackle political security (Cunningham, 2025).

After a long workshop on the matter, we succeeded in making an integrated take at today's Russia's cognitive-informational war (Chifu, 2025d). Disinformation in the EU has to focus on three main directions: the content based on **ideological narratives**, the **interpretations** which make the difference and the **historical revisionism** in order to create a type of legitimacy and credibility of a new national ideology based on presumed role, if not messianic, at least presumed and assumed by the current leadership of Putin's Russia.

Several **tailor-made narratives about the EU**, based on two strategic goals: splitting the transatlantic link, America from the EU, and creating rifts and divisions inside the EU for dismantling the Union, based on support for the populist, nationalist and so-called sovereign - its movements that emerged, taking MAGA as an example, inside the EU. By hiding the Pro-Russian approaches under a type of conservatism and alleged pro-American MAGA promotion, we've arrived at an acceptable approach for the average European citizen, claiming respectability never achieved by pro-Russian ideas. Social destabilisation, polarization, fragmentation and fuelling protests are the ingredients that are accompanying the demarch towards the strategic goals of Russia, with anti-EU narratives, taking advantage of grievances and spreading stories with "proofs" about corruption, failure, and disappointment of the people.

In this ideological sphere, Russia's narratives lack complexity, and a number of alternative narratives, even antagonistic ones, are needed in order to make a catch-all public. Everybody has something to choose from the narratives presented, and the ideas are based at least on confusion, chaos and lack of capacity to have a common approach from the population under a selected theme. Confusion is used in that direction and in showman politics. A central element of this information war is strategic storytelling. Russia is primarily engaged in disinformation, not propaganda. And here, it always involves a myth, the Russian knight in shining armor, defending the people against the dragon. So the whole story needs a Dragon, an enemy, the one that prompts the mobilization and the action to fight it. And the Dragon has alternatively been the Jews, the banderites/banderovsti, the vegetarians, the migrants, the LGBTQ+, the elite, the deep state, the others, any group is helpful to build the stories that mobilize. The others are the people. Everybody feels it is part of the people, and it is easier than to be the hated elite. In ideological terms, the access to the far right and so-called conservatism of Putin, with traditional values, is the result of the disappearance of European conservatism and the creation of a space for the extreme right. Mimetic behaviour has led to the grab by Putin's Russia of the traditional values and church, as well as to this alignment of ideologies that are mixing internal politics, unrest and criticism, mirroring the UK conservatism and altering its goals and values.

The most important theme to be approached was War versus Peace. Here, the EU/Europe, Democracies were labelled as the sides of War, and Russia, with a large respect for Donald Trump, the populist Eurosceptics and anti-Europeans were, on the contrary, the side of Peace. Party of War against the Party of Peace is the fake presentation of aggressor

and victim in the full-scale, high-intensity, long-term war of aggression of Russia on Ukraine. As a corollary, the whole debate was used also not only for fragmenting, polarizing and accentuating societal rifts, but also to deter the perspectives of preparedness in defense and deterrence of the EU and NATO member states.

Autocracy versus democracy is also a key clash supported by the Russian interference, which defends autocratic narratives and underlines, if not that autocracy is a better system, at least that democracy is bad and has shortcomings. **We are also looking at the competition of legitimacies** in the virtual space, where an old platform and a consecrated leadership can lead to a real fight for influence. At the same time, the processes are channelled to relativize all values, so that the conclusion is that nothing is worth being defended! **Large coalitions and rigid political processes** are creating disaffection in society at all levels, without showing real results. This makes the system's vulnerabilities more visible and exploitable by external actors such as Russia. They identify vulnerabilities in democracy as a system and in each country and act quickly to exploit them.

Excessive digitization has created a gap. Many people lack understanding or access to electronic services. They are frightened and feel left behind, unable to adapt and achieve, let alone prosper. For this part of the population, alternatives and equitable access must be ensured. Technology, therefore, plays a role in the entire process of transforming and manipulating public opinion. **The uses and abuses linked to sovereignty** come from abducting and distorting the very clear idea core to international law and converted against the very democracy. In the post-Soviet space, sovereignty does not mean **equal sovereignty of states** based on international law and common rules, but rather “limited sovereignty”, where countries are sovereign as long as they respect the interests of the Russian Federation. The Russian language, Russian history, and Russian culture must dominate and unite the “new Russian empire”. Anyone who wants to be truly independent and sovereign is labelled a Russophobe.

There is another **extreme, unbounded interpretation of sovereignty**, which challenges the rule of law, increasingly visible in democracies: “I won the vote, I can do what I want”. This no longer means the balance of powers, freedom of the press, or the involvement of civil society. It means **the concentration of executive power** and the dissolution of control institutions. It is populist logic: the leader claims to represent the will of the people and, as a result, minorities, the opposition, and institutions no longer matter. The self-proclaimed providential leader governs directly with the people (Chifu, 2025d).

Character assassination in elections. The case of Maia Sandu and the SVR interference

One of the preferred instruments of cognitive-informational warfare of the fourth generation is character assassination. It is targeted at prominent figures and executed in a perverse way in order to transform dramatically even the most humble, decent and reserved politician into an object of hate, pressing the button of the fundamental emotions (Moisi, 2010). It was what happened to Maia Sandu, the President of the Republic of Moldova before, during and after the presidential elections and the referendum for introducing EU accession in the Constitution, November 2024, as well as during the parliamentary elections in 2025.

In a masculine if not a macho world, in the post-Soviet space, gender and fragility become important vulnerabilities and create means of attack in informational warfare. Character Assassination for female leaders in a competition is also orchestrated by

transferring images and fake perception that the administration should be prone to the same characteristics as the physical ones – fragility, delicacy, weakness – of the female leader. Putin's Russia chose to make that type of attack using the spokesperson of the Russian Ministry of Foreign Affairs, Maria Zakharova in the case of Maia Sandu, the pro-European president of the Republic of Moldova running for her second term. A special note will be retained on specific involvement of the Russian FSB intelligence service in directly being involved in the electoral campaigns, supporting or acting in cognitive-informational attacks using character assassination (Chifu, 2025b).

We've studied Maria Zakharova as a vector of attack on Maia Sandu. Russia avoided any accusation of gender attack for avoiding victimization of the target and has done its character assassination especially by using a female representative of its administration to do character assassination (Chifu et al., 2025). There were 86 mentions of Maia Sandu and 741 of the Republic of Moldova, which is a record for a mandate (and the period of electoral campaign). The characteristics of this type of attack are assessed in our paper proving that false assumption of a weak system if the image of the president is characterized by physical fragility.

Lack of personal power and force in its administration, or heavy power weight, Romanian citizenship acquired, all have been used in this campaign. And the last resort in the week before the first round of elections was to distribute via propaganda a false narrative that the EU integration is equal to reunification with Romania. All proved to be false: performed their duties, not being shy in using their legitimate strength and the legally applied violence, the President proved herself brave and determined when she authorized all the needed measures and prevented the extreme involvement and detour of the result of those elections financed heavily by Putin's Russia (Chifu, 2025b).

Our object of study is here the abundance of SVR communiqués during that period related to Maia Sandu – directly and indirectly - and the Republic of Moldova in the same period of time, meaning 2024-2025. We are talking about 8 communiqués of the SVR, compared to none previously referring to the Republic of Moldova and Maia Sandu. Moreover, in two cases, we are speaking about an interview and a speech delivered by the chief of Russian Foreign Intelligence Service, SVR, Sergey Naryshkin referring to Maia Sandu and the Republic of Moldova. In the interview, Naryshkin referred to Maia Sandu by name (SVR, 2024a), calling the Maia Sandu Regime in line of America's policies in the context of arguing that the United States is losing the initiative on all fronts, from the Middle East to Asia and Africa. And in the post-Soviet space, it's even failing miserably. And Maia Sandu regime divided the country because it ~~is~~ was not accepted for its stance of pursuing EU membership. This happened in the context of immense sums of money injected in the Republic of Moldova by Russia in order to buy those elections.

Naryshkin intervened a second time, that once in a speech, at the 20th meeting of the Conference of Heads of Security Agencies and Intelligence Services of the CIS Member States, November 27, 2024. In a speech repeating all the mantras and disinformation established already at the national level, regarding Russia's posture, the EU, US, Ukraine, a failed state, Naryshkin underlined several issues: the Republic of Moldova is following a similar path (as Ukraine, war and failed state), Maia Sandu is leading a totalitarian regime, she is linked to "the **comprador** policies" (SVR, 2024b), meaning colonial ones, that have led to a catastrophic polarization of society. Elections have been if not stolen, at least distorted through administrative means and the malign mobilisation of the diaspora, appeal to protests from "the anger of the deceived half of the population", and formulating a threat for the integrity of the state.

Every communique begins with a special formulation “The Press Bureau of the Russian Foreign Intelligence Service reports that, according to information received by the SVR”, giving credibility and legitimacy, but also ends with a comment/conclusion aiming at doing the biggest part of the disinformation with impunity: “The Press Bureau of the Russian Foreign Intelligence Service reports that, according to information received by the SVR,” is a current formulation for the conclusion and last paragraph of disinformation communiques that are dragging with impunity any type of groundless conclusion.

Europe, meaning European Union, and NATO are the main targets in assuming that the Republic of Moldova, and Romania, for that matter, don’t have agency and are only puppets of the Americans and, after 2025, European leadership. In one case, the attack is that Europe is preparing to occupy the Republic of Moldova, for sure, because of Maia Sandu’s pro-European leadership. European bureaucrats are determined to keep Moldova on track with their Russophobic policies through the deployment of troops and the de facto occupation of the country (SVR, 2025a).

The themes involved are:

- A direct confusion between the EU and NATO (under the generic term of Europe).
- Attacks on Russophobia due to the reinstalment of the Romanian language as the sole official language of the state.
- Pejorative reference to European bureaucrats.
- Concentration of armed forces in Romania near the Moldovan border.
- Link with Ukraine - Ukraine's Odessa region.
- Reference to the intimidation of Transnistria, considered to be a Russian territory.
- Invoking that the scenario has been repeatedly practiced during NATO exercises in Romania.
- Time frame: after the parliamentary elections in the Republic of Moldova on September 28; elections to the Supreme Council of the PMR on November 30 of this year were considered as a possible timeframe. Nothing happened.
- Fakes in the text: gross falsification of the vote results.
- Invitation to protests: will force desperate Moldovan citizens to take to the streets to defend their rights.
- Fake and Character assassination of Maia Sandu: at the request of President M. **Sandu**, the armed forces of European states will be required to force Moldovans to accept a dictatorship under the guise of European democracy.
- Armed provocations against Transnistria and the Russian troops stationed in the region are envisaged.
- Reference to liberal democracy in a pejorative form: plans by totalitarian-liberal European regimes.
- Reference to the so-called European colonialism.
- Victimisation and protection of the weak: Fearing a direct confrontation with large-scale Russia, the Europeans intend to take revenge on small Moldova. Self-assertion at the expense of the weak.

Then NATO is the target of SVR in its contribution to cognitive-informational warfare in the Republic of Moldova, giving false credibility and legitimacy of pro-Russian forces during the elections and claiming that the Republic of Moldova is going to be used as a new battle space against Russia (even though Russia is some hundreds of km East of its territory) (SVR, 2025b). The main issues addressed and references are:

- Fake: NATO is actively preparing to involve the Republic of Moldova in a potential armed conflict with Russia.
- Transformation of this country into „an advanced base for the alliance on the eastern flank, given the advancement of Russian forces in Ukraine” – the Republic of Moldova is a neutral state, not a NATO member country and Russia is not advancing in Ukraine at this stage, on the contrary.
- Fake and geographical mismatch: to make Moldovan territory suitable for the rapid deployment of NATO troops to Russian borders. There is not any common border of the Republic of Moldova with Russia.
- Arguments are lacking, and any military infrastructure locally built up is considered a NATO preparation for fighting Russia, therefore blocking the Republic of Moldova’s own preparation to defend itself.
- Fake and emotional subliminal psychological attack: Moldovans intended to become „cannon fodder” in combat with Russian troops.
- Character assassination and association with colonist regimes: The **comprador** regime of M. Sandu is ready to satisfy all Western demands (Naryshkin used the same term in his interview from 27 of November (SVR, 2024a))
- Emotion and threat: the beginning of the end for Moldova.
- Character assassination: Sandu and her entourage intend to go much further, namely, handing Moldova over to NATO, so they can simply „wipe it out” against Russia.
- Malign association: Zelenski, gay, Ukraine war, EU lover of gay pride: Apparently, Zelenskyy's cannibalistic reputation is giving this lover of gay pride parades and European integration no peace. (Gayropa, old symbolism negative by linking EU and gay orientation)

The most visible direct element of character assassination against Maia Sandu comes in an SVR communique associated with the pressure coming from Moscow related to the energy supplies, especially on Transnistria, with low incidence in the Republic of Moldova part under the control of the legitimate authorities from Chisinau. In that communique, Maia Sandu is accused of being “emotionally unstable” raising problem to her staff and Prime Minister on how to calm her down (SVR, 2024c). The communique states that during the event, the president „lost her temper”, “The president was not sobered”, “Entering into a frenzy”, she categorically refused to discuss the issue with Ukraine and categorically placed all responsibility on Russia. The communique uses:

- Fakes: According to Sandu, “if Moscow doesn't find a way to deliver gas here, Chisinau will take it out on pro-Russian Transnistria”; She ordered criminal cases to be opened against the PMR leadership “for separatism”; ordinary Transnistrian residents to be subjected to thorough searches when crossing the region's borders; the president demanded preparations for a violent seizure of the power plant; Sandu urging the development of a plan for a military operation to establish control over Transnistria and eliminate the Russian peacekeeping presence in the region.
- Indirect threats: Moldova's energy supply problems, which the country could face after the Russian-Ukrainian natural gas transit agreement expires on December 31st; Moldova's right-bank region is almost entirely dependent on electricity supplies from the Moldavskaya GRES, located in the Transnistria region.

- Innuendos combined with fakes: The European Union, of course, wouldn't mind the emergence of a new crisis; no one can guarantee that the Moldovan president won't actually try to unleash a real war in the region.
- Malign association, presumptuous, accuses - Sandu Romanian citizen: Many observers are wondering how much longer the Moldovan people, among whom, incidentally, there are many who experienced or remember well the fighting on the Dniester in 1992, will tolerate the experiments conducted on themselves by Romanian citizen Sandu?

In another communique, SVR is attacking the large electoral frauds, explicitly formulated as such, that happen during the elections, even if Russia was the only actor really involved in the attempt to disturb and derail the elections in the Republic of Moldova. Formally, the target of the communique was the OSCE, covering up the fraud, the communique implies that, from its sources, it can tell that “the OSCE Secretariat leadership is aiming to ensure the victory of incumbent President M. Sandu in the second round”, “do not intend to publicize the numerous violations after either the first or second round of voting”, “because Washington and its allies are interested in Sandu's reelection for a second term and Chisinau's continued course of accelerated rapprochement with the EU and NATO” (SVR, 2024d).

On another point, Russia distorted the legitimate aims of the president of the Republic of Moldova to defend herself against the hybrid warfare unleashed by Moscow against her and her state, and easily labelled that as Russophobia, one of the usual labels against its enemies together with that of nazis and banderovtsi. Maia Sandu's statement that “every day the Kremlin and criminal groups carry out hybrid attacks, using propaganda weapons to sow hatred in Moldova”. The attacks linked to character assassination are:

- Accuses through association: Romanian citizen, Romania a NATO country unfriendly to Russia.
- Misinterpretation and innuendos: strict censorship that the Moldovan authorities impose on their own information space.
- Malign association: nationalist Kyiv regime, which is close to Sandu's heart; the ideological tenets of the West.
- Labelling: Chisinau has given carte blanche to Russophobic ideology.
- Propaganda developing: publications artificially imposed on Moldovans.
- Fakes: media outlets in the Republic of Moldova are under tight Western control; funding for such media outlets comes from grants from the US Agency for International Development, the European Endowment for Democracy, German party foundations, and Western embassies and foreign ministries.
- Turning the mirror (accusing the other of what you are doing yourself): Special seminars are regularly held for their employees to explain how to „correctly” cover certain events.
- Sarcasm and mockery: It would seem that the only thing missing for the full triumph of Western democracy in the Republic of Moldova is an Orwellian „Ministry of Truth.” - Sandu has already announced her intention to create a National Center for Information Protection and Combating Propaganda, „Patriot”.

CONCLUSION

The analysis shows that Russia is respecting its initial plan of cognitive informational warfare aimed at all the publics, from post-Soviet states to western European states and the US, as well as a special treatment of the countries in the Global South, or silent majority. The narrative behind is a real Grand Strategy of Russia, prepared in order to maximize its geopolitical situation and present Russia as a Great Power, in spite of the characteristics which no longer recommend it as such. And the strategy of cognitive-informational warfare involves every institution and doesn't stop at any moment in achieving its objective, no matter how futile and impossible it is.

The intelligence agencies, researchers and journalism have identified at a large respect the mechanism of spreading the fakes, making propaganda and disinformation and targeting the different states in specific moments with cognitive-informational warfare or, for that matter, personalities. The case of Maia Sandu and character assassination, already studied, has a special impact when it is about the implication of the SVR, the foreign intelligence service, using its aura of credibility and legitimacy by knowing things away from the normal citizens' reach.

The format of the communiques is implying extra information detained by the SVR who are on the back of each assessment, and finishes with a paragraph of opinions realised through distortion, relativisation, misinterpretation and all the instruments of the disinformation: fake news, turning the mirror, labelling, association with negative images and cultivated repulsive references, innuendos and malign reading of existing facts, statements and actions.

The aim of the communiques and the general involvement, including by using the head of the SVR, Sergey Naryshkin, in key electoral moments, is to create confusion, to attack the pro-European candidates, to support the pro-Russian ones and to distort the facts, to call for and to provoke riots on the streets, rejection of election results, using its own malign intervention during the elections to contest afterwards the results. SVR is no longer just an intelligence agency, but becomes the reference for the truth and an actor directly involved, not only via its operations and deeds, in elections in countries like the Republic of Moldova, when the outcome is crucial for the Russian establishment interests.

References

- Al-Rodhan, N. (2026, June 18). Psychological warfare was once an afterthought, it's now the primary battleground. *Institute for Arts and Ideas*. <https://iai.tv/articles/psychological-warfare-was-once-an-afterthought-its-now-the-primary-battleground-auid-3602>
- Andress, J., & Winterfeld, S. (2013). *Cyber warfare: Techniques, tactics and tools for security practitioners*. Elsevier.
- Backes, O., & Swab, A. (2019). *Cognitive warfare: The Russian threat to election integrity in the Baltic States* [Doctoral thesis]. Harvard University.
- Benoit, W. L. (2020). Character assassination and persuasive attack on CBS's Face the Nation. In S. A. Samoilenko, M. Icks, J. Keohane, & E. Shiraev (Eds.), *Routledge handbook of character assassination* (pp. 295–307). Routledge.
- Keohane, J. (2020). The rhetorical and ethical implications of character assassination in the age of McCarthy. In S. A. Samoilenko, M. Icks, J. Keohane, & E. Shiraev (Eds.), *Routledge handbook of character assassination* (pp. 269–281). Routledge.
- Bernal, A., Carter, C., Singh, I., Cao, K., & Madreperla, O. (2020). *Cognitive warfare: An attack on truth and thought* [Technical report]. NATO & Johns Hopkins University. <https://www.innovationhub-act.org/sites/default/files/202103/Cognitive%20Warfare.pdf>
- Botts, E. (2025). Opaque architecture: The convergence of media cooptation, cultural funding and intelligence operations and the erosion of information ecosystem. *Journal of European and American Intelligence Studies*, 2025(2), 63–84.
- Burda, R. (2023). *Cognitive warfare as part of society's never-ending battle for minds*. Hague Centre for Strategic Studies. <https://www.jstor.org/stable/resrep51454>
- Chifu, I. (2017). Character assassination – Armă ofensivă în războiul informațional. *Revista Infosfera*, December, 11–18.
- Chifu, I. (2022). Reziliență contemporană: Competitivitatea națiunilor, capacitatea statelor și securitatea societală. *Infosfera*, 2022(3), 5–14.
- Chifu, I. (2023). Russian provoked nuclear disaster, attack or accident: Scenarios, probability and consequences. In F.-C. Cirstoiu et al. (Eds.), *Medical response strategy in case of radiation emergency caused by the war in Ukraine* (pp. 71–84). Springer.
- Chifu, I. (2024, October 14). Superioritatea cognitivă și influențarea deciziei – noile ținte ale agresiunii Rusiei lui Putin în Occident. *Calea Europeană*. <https://www.caleaeuropeana.ro/iulian-chifu-superioritatea-cognitiva-si-influentarea-deciziei-noile-tinte-ale-agresiunii-rusiei-lui-putin-in-occident/>
- Chifu, I. (2025a). Conceptualizare și evaluare epistemologică: Războiul cognitiv. *Infosfera*, 2025(2), 5–18. https://www.mapn.ro/publicatii_militare/arhiva_infosfera/documente/2025/2_2025.pdf
- Chifu, I. (2025b). *Individual physical fragility as a weakness in power driven political systems in the post-Soviet space: Maia Sandu case* [Conference presentation]. 5th International Conference "Gender, Power, and Politics in Character Assassination," George Mason University. <https://carplab.org/carp-2025-conference/>
- Chifu, I. (2025c, June 13). Serghei Karaganov și esența gândirii ideologice ruse: Agresiunea ca mod de viață și teoria războiului bun. *Calea Europeană*. <https://www.caleaeuropeana.ro/iulian-chifu-serghei-karaganov-si-esenta-gandirii->

- ideologice-ruse-agresiunea-ca-mod-de-viata-si-teoria-razboiului-bun/
- Chifu, I. (2025d). *The European way to counter disinformation – Make Europe resilient again* [Round table report]. Conflict Prevention and Early Warning Center & Hans Seidel Foundation, Bucharest, Romania.
- Chifu, I. (2026, June 22). Politică de tip cuantic în epoca social media: Participarea individualizată. *Calea Europeană*. <https://www.caleaeuropeana.ro/iulian-chifu-politica-de-tip-cuantic-in-epoca-social-media-participarea-individualizata/>
- Chifu, I., & Nantoi, O. (2016). *Information warfare: The pattern of aggression*. Publishing House of the Institute of Political Sciences and International Relations "Ion I. C. Brătianu" of the Romanian Academy.
- Chifu, I., & Savu, L. (2020). *The impact of technology on human being, society and politics*. Publishing House of the Institute of Political Sciences and International Relations Ion I.C. Brătianu of the Romanian Academy.
- Chifu, I., & Simons, G. (2023). *Rethinking warfare in the 21st century: The influence and effects of the politics, information and communication mix*. Cambridge University Press.
- Chifu, I., & Țuțuianu, S. (2017). *Torn between East and West: Europe's border states*. Routledge.
- Chifu, I., Marzac, E., & Kostandinova, V. (2025). *Cognitive informational warfare in the trilateral Romania–Republic of Moldova–Ukraine*. Rao Publishing House.
- Cunningham, A. (2025). Conflicts of tomorrow: The impact of Russian information operations on American foreign policy and the need for increased political security. *Journal of European and American Intelligence Studies*, 2025(2), 85–96.
- Dahl, A. B. (1996). *Command dysfunction: Minding the cognitive warfare* [Technical report]. AU's School of Advanced Air and Space Studies.
- Danyk, Y., & Briggs, C. M. (2023). Modern cognitive operations and hybrid warfare. *Journal of Strategic Security*, 16(1), 35–50. <https://digitalcommons.usf.edu/cgi/viewcontent.cgi?article=2032&context=jss>
- Estonian Foreign Intelligence Service. (2026). *International security and Estonia 2026* [Annual report].
- Foreign Intelligence Service of Russia (SVR). (2024a, December 10). Director of the Russian Foreign Intelligence Service Sergey Naryshkin: "Intelligence is life itself." <http://svr.gov.ru/smi/2024/12/direktor-svr-rossii-sergey-naryshkin-razvedka-eto-sama-zhizn.htm>
- Foreign Intelligence Service of Russia (SVR). (2024b, November 27). Speech by the Director of the Foreign Intelligence Service of Russia at the 20th meeting of the Conference of Heads of Security Agencies and Intelligence Services of the CIS Member States. <http://svr.gov.ru/smi/2024/11/vystuplenie-direktora-svr-rossii-na-xx-zasedanii-soveshchaniya-rukovoditeley-organov-bezopasnosti-i.htm>
- Foreign Intelligence Service of Russia (SVR). (2024c, December 23). On the threat of escalation of the situation in Moldova. <http://svr.gov.ru/smi/2024/12/ob-ugroze-eskalatsii-situatsii-v-moldavii.htm>
- Foreign Intelligence Service of Russia (SVR). (2024d, October 31). The OSCE is covering up large-scale election fraud in Moldova. <http://svr.gov.ru/smi/2024/10/obse-pokryvaet-masshtabnye-narusheniya-na-vyborakh-v-moldavii.htm>
- Foreign Intelligence Service of Russia (SVR). (2025a, September 23). Europe is preparing to occupy Moldova. <http://svr.gov.ru/smi/2025/09/evropa-gotovitsya-okkupirovat-moldaviyu.htm>

- Foreign Intelligence Service of Russia (SVR). (2025b, July 14). NATO is turning Moldova into a new military battering ram against Russia. <http://svr.gov.ru/smi/2025/07/nato-prevrashchaet-moldaviyu-v-novyy-voennyi-taran-protiv-rossii.htm>
- Foreign Intelligence Service of Russia (SVR). (n.d.). *Press bureau communiqués*. <http://svr.gov.ru/smi.htm>
- Frolova, A., et al. (2026). *The white book: Navigating the Russian propaganda minefield*. Media Center Ukraine.
- Icks, M., & Shiraev, E. (2014). *Character assassination throughout the ages*. Palgrave Macmillan.
- Le Grand Continent. (2025, June 7). « La guerre est dans les gènes des Russes » : Un entretien inédit avec Sergueï Karaganov, l'architecte de la géopolitique de Poutine. <https://legrandcontinent.eu/fr/2025/06/07/entretien-avec-serguei-karaganov-larchitecte-de-la-geopolitique-de-poutine/>
- Mariș, A. (2024). *Influence made simple*. [Publisher not specified in source].
- Ministry of Defence of the Russian Federation. (2000). *Russian Federation Armed Forces' information space activities concept*. http://www.scrf.gov.ru/security/information/DIB_eng/
- Miskolczy, I., & Gherghina, S. (2024). Making words harmless: Why politicians survive character assassination attacks. *The British Journal of Politics and International Relations*, 1–21.
- Moisi, D. (2010). *The geopolitics of emotion*. Anchor Books.
- NATO Allied Command on Transformation. (2025, September). Cognitive warfare. *CogWars Newsletter*, 1. https://www.act.nato.int/wp-content/uploads/2025/10/20250819_CogWar-Newsletter-September.pdf
- Office of the Chairman of the Joint Chiefs of Staff. (1996). *Joint information warfare policy* [CJCSI 3210.01]. <https://www.scribd.com/document/71150944/Cjcs-3122-02a-Jopes-Vol-2>
- Pocheptsov, G. (2018). Cognitive attacks in Russian hybrid warfare. *Information & Security*, 41, 37–43. <https://doi.org/10.11610/isij.4103>
- Romanian Presidency. (2024, December 4). *Communiqué (includes declassified documents)*. <https://www.presidency.ro/ro/media/comunicate-de-presa/comunicat-de-presal733327193>
- Samoilenko, S. A. (2021a). Character assassination: The sociocultural perspective. *Journal of Applied Social Theory*, 1(3), 186–205.
- Samoilenko, S. A., Icks, M., Shiraev, E., & Keohane, J. (Eds.). (2021). *Routledge handbook of character assassination and reputation management*. Routledge.
- Samoilenko, S. A. (2021b). The cocreational view of character assassination. In C. H. Botan (Ed.), *The handbook of strategic communication* (pp. 76–90). Wiley.
- Samoilenko, S. A., & Suvorova, I. (2023). Artificial intelligence and deepfakes in strategic deception campaigns: The U.S. and Russian experiences. In E. Pashentsev (Ed.), *The Palgrave handbook of malicious use of AI and psychological security*. Palgrave Macmillan.
- Samoilenko, S. A., Erzikova, E., Davydov, S., & Laskin, A. (2017). Same messages: Character assassination in the television news during the 2014 Ukrainian crisis. *International Communication Research Journal*, 52(2), 28–52.
- Shekhovtsov, A. (2026, June 16). Europe's 2027 elections will show if the EU dares to call out US interference. *EUObserver*. <https://euobserver.com/222156/europes-2027->

- elections-will-show-if-the-eu-dares-to-call-out-us-interference/
- Shirayev, E. (2014). Character assassination: How political psychologists can assist historians. In M. Icks & E. Shirayev (Eds.), *Character assassination throughout the ages* (pp. 15–36). Palgrave Macmillan.
- Shirayev, E. (2019). *Character assassination: An interdisciplinary approach*. George Mason University. <https://characterattack.wordpress.com/wp-content/uploads/2010/10/character-assassination-an-interdisciplinary-approach.pdf>
- Simons, G., & Chifu, I. (2017). *The changing face of warfare in the 21st century*. Routledge.
- Simons, G., & Chifu, I. (2025). Information warfare as a theoretical construct and an operational practice. In S. Samoilenko & S. Simmons (Eds.), *The handbook of social and political conflict* (pp. 323–342).
- Sloss, D. L. (2020). *Tyrants on Twitter: Protecting democracies from information warfare*. Stanford University Press.
- Tashev, B., Purcell, M., & McLaughlin, B. (2019). Russia's information warfare: Exploring the cognitive dimension. *MCU Journal*, 10(2), 129–147.
- Tsuchiya, M. (2022). Governing cognitive warfare. In K. Govella (Ed.), *Governing the global commons: Challenges and opportunities for U.S.–Japan cooperation*. German Marshall Fund. <https://www.jstor.org/stable/resrep46996.9>

HYBRID ACTIVITY INDICATOR SYSTEM (HAIS): A SECURITY SCIENCE FRAMEWORK FOR EARLY DETECTION AND ASSESSMENT

Darko Trifunović

Institute for National and International Security (INIS)

darko.trifunovic@intelligence-security.rs

<https://orcid.org/0000-0003-3591-9554>

Abstract

Hybrid activities have become one of the defining characteristics of the contemporary security environment, yet existing research predominantly focuses on their actors, methods, and operational effects rather than on the systematic identification of their earliest observable manifestations. This article addresses that methodological gap by introducing the Hybrid Activity Indicator System (HAIS) as a Security Science-based framework for the early detection and assessment of hybrid activities. Unlike traditional approaches that primarily analyze threats, risks, and challenges, the proposed framework conceptualizes hybrid activities as multidimensional processes that generate security problems through the cumulative interaction of political, informational, security and intelligence, cyber, economic, social, and international factors. The article develops a standardized classification of seven categories of hybrid activity indicators, establishes methodological principles for their identification and assessment, and proposes the Hybrid Activity Indicator Score (HAIS Score) as an analytical instrument for evaluating the overall intensity of hybrid activities. The applicability of the model is demonstrated through a qualitative regional assessment of Southeast Europe, identifying a high overall presence of hybrid activity indicators across the region. The principal scientific contribution of the article lies in shifting the analytical focus from the descriptive study of hybrid activities toward a measurable indicator-based methodology grounded in Security Science. By integrating heterogeneous security observations into a coherent analytical framework, HAIS provides a methodological foundation for intelligence analysis, strategic forecasting, comparative assessment, early warning, and evidence-based security decision-making.

Keywords: Security Science; Hybrid Activities; Hybrid Activity Indicator System (HAIS); Security Indicators; Hybrid Activity Indicator Score; Security Problems; Early Warning; Intelligence Analysis; Southeast Europe.

Introduction

The contemporary security environment has undergone profound transformation during the last two decades. Traditional military confrontation has increasingly been supplemented—and in many cases replaced—by complex combinations of political influence, cyber operations, economic coercion, information manipulation, intelligence activities, organized crime, and the exploitation of social vulnerabilities (Hoffman, 2007, pp. 7–15). These activities rarely amount to open armed aggression, yet they can gradually undermine the political stability, institutional resilience, economic security, and social

cohesion of targeted states. Consequently, hybrid activities have become one of the defining characteristics of the twenty-first-century security landscape.

Southeast Europe represents one of the regions where such activities are particularly visible. The region is characterized by unresolved historical disputes, ethnic and religious diversity, fragile political balances, incomplete institutional transformation, strategic competition among global and regional powers, dependence on critical energy infrastructure, and intensive information influence campaigns (European Commission & High Representative of the Union for Foreign Affairs and Security Policy, 2016, pp. 2–4). These structural characteristics create a security environment in which hybrid activities can develop simultaneously across political, informational, economic, cyber, intelligence, and societal domains, often remaining below the threshold of conventional armed conflict while producing significant strategic consequences.

As a result, hybrid activities have become a central subject of contemporary security research. NATO, the European Union, national security strategies, intelligence agencies, and numerous academic studies have extensively examined hybrid threats, hybrid warfare, grey-zone conflicts, strategic influence operations, cyber campaigns, and foreign interference (NATO, 2022, pp. 7–8). This growing body of literature has substantially improved our understanding of the actors, instruments, objectives, and operational characteristics of hybrid activities.

However, despite this extensive scholarly and policy attention, one fundamental scientific problem remains insufficiently addressed.

Current research predominantly explains **what hybrid activities are, who conducts them, which instruments are employed, and what consequences they may produce**. Comparatively little attention has been devoted to a more fundamental methodological question: **how can hybrid activities be systematically identified, measured, and assessed before they generate major security consequences** (Gill & Phythian, 2018, pp. 154–168)?

This question becomes particularly important because hybrid activities rarely begin with spectacular events. They usually evolve gradually through the accumulation of numerous seemingly unrelated actions that individually may appear insignificant but collectively produce measurable deterioration of the security environment. Coordinated disinformation campaigns, increasing cyber intrusions, foreign political influence, economic pressure, intelligence penetration, attacks on critical infrastructure, growing social polarization, organized criminal activities, and externally supported extremist narratives often emerge simultaneously long before an open security crisis becomes visible. The principal challenge for security institutions therefore lies not merely in responding to completed hybrid operations but in recognizing their earliest observable manifestations.

From the perspective of **Security Science**, this challenge represents a fundamental scientific problem. Security Science does not primarily study isolated threats or individual hostile actions; rather, it examines **security problems**, their emergence, dynamics, consequences, and their influence on the overall security condition of states and societies (Trifunović, 2025, pp. 8–15). Within such a framework, hybrid activities should not be understood as security problems in themselves. Instead, they represent dynamic processes capable of generating security problems through the cumulative interaction of political, informational, cyber, economic, intelligence, and social activities. Consequently, the primary scientific task is not simply to classify hybrid activities but to identify the observable indicators through which their development can be detected, monitored, evaluated, and predicted.

Despite the practical importance of early warning, contemporary security literature still lacks a comprehensive methodological framework capable of integrating diverse manifestations of hybrid activities into a unified system of measurable security indicators. Existing approaches generally rely on qualitative assessments, expert judgment, or descriptive case studies, making comparative analysis between different states and periods difficult. Without standardized indicators, it remains challenging to distinguish ordinary political and social developments from coordinated hybrid activities or to assess objectively whether the overall level of hybrid pressure against a state is increasing or decreasing.

This methodological gap has important implications not only for academic research but also for intelligence analysis, strategic forecasting, national security planning, and public policy. Intelligence services require measurable indicators to distinguish isolated incidents from coordinated campaigns. Decision-makers require objective analytical instruments that can identify trends before they evolve into major crises. Security systems require scientifically grounded methodologies that transform dispersed information into actionable assessments supporting timely institutional responses (Gill & Phythian, 2018, pp. 154–168).

This article argues that these requirements can be addressed by developing a *Hybrid Activity Indicator System (HAIS)* grounded in the theoretical principles of Security Science. Rather than treating hybrid activities exclusively as forms of warfare or political confrontation, the proposed framework conceptualizes them as multidimensional generators of security problems whose development can be monitored through identifiable and measurable security indicators. The article develops a comprehensive classification of these indicators, proposes an integrated analytical model for their assessment, and demonstrates its applicability to the security environment of Southeast Europe.

The principal scientific contribution of this article lies in shifting the focus of hybrid security research from the descriptive analysis of hybrid activities toward the systematic development of measurable security indicators. In doing so, it seeks to provide both a new methodological instrument for intelligence and security analysis and a theoretical contribution to the further development of Security Science as a discipline concerned with the identification, assessment, and prediction of contemporary security problems.

HYBRID ACTIVITIES IN CONTEMPORARY SECURITY LITERATURE

The concept of hybrid activities has become one of the most extensively discussed subjects within contemporary security scholarship. Since the beginning of the twenty-first century, rapid technological development, globalization, cyber interconnectivity, and increasing strategic competition among major powers have fundamentally transformed how states pursue political and security objectives. Contemporary conflicts increasingly develop below the threshold of conventional warfare, employing coordinated combinations of military and non-military instruments designed to influence political decision-making, weaken institutional resilience, manipulate public perception, and gradually alter the strategic environment without provoking an immediate military response.

One of the earliest and most influential contributions to this debate was provided by Frank G. Hoffman, who introduced the concept of *hybrid warfare* to describe conflicts in which conventional military capabilities are combined with irregular tactics, terrorism, criminal activities, cyber operations, and information campaigns. Hoffman's work significantly expanded the understanding of modern conflict by demonstrating that future security challenges could no longer be analyzed exclusively through traditional distinctions

between war and peace or between conventional and irregular warfare (Hoffman, 2007, pp. 7–15). Hybrid activities came to be understood as the coordinated employment of multiple instruments of power designed to achieve strategic objectives through complexity, ambiguity, and simultaneity.

Subsequent scholarship considerably broadened this perspective. Contemporary research increasingly distinguishes hybrid activities from purely military operations by emphasizing the integration of diplomatic, informational, military, economic, intelligence, cyber, technological, legal, and psychological instruments. Within this broader understanding, military force often represents only one component—and in many cases not even the dominant component—of a much wider strategic campaign intended to shape political behavior, influence decision-making, and undermine societal resilience.

Parallel with academic developments, international organizations have incorporated hybrid activities into their strategic concepts. NATO has identified hybrid threats as one of the principal security challenges confronting Allied nations, emphasizing the need for improved resilience, intelligence sharing, strategic communications, cyber defense, and civil preparedness (NATO, 2022). Similarly, the European Union has developed comprehensive approaches focused on countering foreign information manipulation, protecting critical infrastructure, strengthening cybersecurity, reducing vulnerabilities within democratic institutions, and enhancing cooperation among member states (European Commission & High Representative of the Union for Foreign Affairs and Security Policy, 2016, pp. 2–4).

The establishment of the European Center of Excellence for Countering Hybrid Threats (Hybrid CoE) further institutionalized research in this field. Hybrid CoE has produced extensive analytical work concerning hybrid influence campaigns, foreign interference, critical infrastructure protection, economic coercion, legal resilience, election security, strategic communication, and whole-of-government responses. Its publications have significantly contributed to understanding the multidimensional nature of hybrid activities and the need to integrate political, economic, military, intelligence, and societal perspectives.

Within Intelligence Studies, hybrid activities are increasingly examined through the concepts of intelligence warning, strategic surprise, foreign influence operations, and intelligence assessment. Intelligence services play a central role in detecting coordinated hybrid campaigns before they evolve into broader security crises, requiring the integration of intelligence collection, analysis, and strategic forecasting (Gill & Phythian, 2018, pp. 154–168). Intelligence services are now recognized as one of the principal institutions responsible for detecting coordinated hybrid campaigns before they develop into broader security crises. Consequently, intelligence analysis has expanded beyond traditional espionage and military intelligence to include cyber intelligence, open-source intelligence (OSINT), social media monitoring, influence analysis, economic intelligence, and strategic forecasting.

Contemporary scholarship has further emphasized that information manipulation and disinformation have become central components of hybrid activities. Rather than relying primarily on military force, modern hybrid campaigns increasingly employ coordinated political influence, disinformation, covert propaganda, and active measures designed to shape public perception, undermine trust in democratic institutions, and influence strategic decision-making. These instruments operate alongside cyber, intelligence, diplomatic, and economic activities, reinforcing the multidimensional nature of hybrid campaigns (Rid, 2020, pp. 1–18).

Despite these important theoretical and institutional advances, the existing literature remains primarily descriptive and explanatory. Most studies focus on identifying the actors conducting hybrid activities, analyzing their objectives, examining operational methods, describing individual case studies, or evaluating governmental responses. Significant attention has also been devoted to understanding the legal, political, military, technological, and informational dimensions of hybrid campaigns. Nevertheless, comparatively little attention has been devoted to one fundamental methodological issue: *how can hybrid activities be systematically measured?*

Current research generally assumes that hybrid activities become visible only after they have produced political crises, social instability, cyber incidents, institutional disruption, or military escalation. Consequently, analytical attention is directed primarily toward completed or ongoing hybrid operations rather than toward identifying the earliest measurable indicators that precede their visible manifestation.

This methodological limitation creates several scientific and practical difficulties. First, different researchers frequently evaluate identical situations using different analytical criteria, reducing comparability among studies. Second, governments often rely upon qualitative expert assessments that are difficult to replicate or verify scientifically. Third, intelligence organizations often possess extensive information about isolated hybrid activities but lack standardized methodologies to integrate these observations into a comprehensive assessment of the overall level of hybrid pressure directed against a particular state.

The absence of standardized security indicators also complicates comparative research. Without clearly defined indicators, it becomes difficult to compare hybrid activity levels across different countries, evaluate changes over time, identify emerging trends, or distinguish ordinary political competition from coordinated hybrid influence campaigns. As a result, contemporary literature provides rich conceptual explanations of hybrid activities but comparatively limited methodological instruments for their systematic assessment.

From the perspective of *Security Science*, this represents a significant research gap. Scientific disciplines require observable phenomena that can be systematically identified, classified, measured, compared, and predicted. If hybrid activities are understood as multidimensional processes capable of generating security problems, their development should necessarily produce observable indicators reflecting changes within the security environment before major crises emerge. The systematic identification of such indicators therefore becomes not merely a practical intelligence requirement but also a scientific problem requiring an appropriate theoretical and methodological framework.

Accordingly, this article does not seek to redefine hybrid activities or propose another typology of hybrid threats. Rather, it addresses a different level of scientific inquiry. Building upon the existing literature, it develops a *Hybrid Activity Indicator System (HAIS)* designed to identify, classify, and assess observable security indicators generated by hybrid activities. In doing so, the article complements rather than replaces existing scholarship by shifting analytical attention from describing hybrid activities toward measuring their security effects through a coherent system of scientifically grounded indicators.

SECURITY SCIENCE PERSPECTIVE: BEYOND THREATS, RISKS, AND CHALLENGES

The growing body of literature devoted to hybrid activities has significantly expanded contemporary understanding of complex security phenomena (Gill & Phythian, 2018, pp. 154–168). Nevertheless, most existing analyses remain strongly influenced by theoretical approaches originating from Political Science, International Relations, Strategic Studies, Military Studies, and Security Studies. Within these disciplines, hybrid activities are predominantly interpreted through the concepts of **security threats**, **security risks**, and **security challenges** (Wolfers, 1952, pp. 481–485). These concepts undoubtedly represent indispensable analytical instruments for understanding uncertainty, assessing potential dangers, and evaluating future developments. However, from the perspective of *Security Science*, they do not provide a sufficiently comprehensive scientific framework for explaining the entire life cycle of contemporary security phenomena. Security Science approaches security from a different epistemological position. Rather than limiting scientific inquiry to threats, risks, and challenges, it identifies the **security problem** as the central object of scientific investigation. This distinction is not merely terminological. It fundamentally changes both the theoretical understanding of security phenomena and the methodological approaches employed to analyze them.

Threats, risks, and challenges primarily describe conditions that may lead to insecurity. They are largely anticipatory concepts directed toward possibility. A threat identifies a potential source of harm; a risk estimates the probability and possible consequences of harmful events; a challenge denotes circumstances requiring adaptation or strategic response (Aven, 2015, pp. 1–18). Each of these concepts contributes to understanding different dimensions of insecurity, but all remain oriented toward conditions that may or may not develop into actual security disturbances.

Security Science introduces an additional analytical level. A security problem represents an objectively existing or demonstrably emerging condition that directly affects the security condition of a state, society, institution, or protected value and therefore requires an organized institutional response (Trifunović, 2025, pp. 8–34). Unlike threats or risks, which describe possibilities, a security problem describes an observable security reality whose existence can be scientifically identified, analyzed, monitored, and measured. This distinction has profound implications for the study of hybrid activities. Hybrid activities should not automatically be equated with security problems. Individual information campaigns, cyber intrusions, economic pressure, foreign political influence, intelligence operations, organized criminal activities, or social manipulation may initially appear only as isolated threats or risks. However, when these activities become coordinated, cumulative, and strategically directed, they gradually generate objectively observable changes within the security environment. At that moment, they cease to represent merely potential dangers and begin producing measurable security problems affecting institutional resilience, social cohesion, political stability, critical infrastructure, national decision-making, or the overall security condition of the state.

Accordingly, the scientific task is not limited to identifying whether hybrid threats exist. The more fundamental question concerns *when and through which observable indicators hybrid activities evolve into security problems*.

This perspective distinguishes Security Science from approaches derived primarily from Political Science or International Relations. Political analyses frequently focus on power relations, strategic competition, diplomatic behavior, or geopolitical interests.

Security Studies traditionally analyze threats, risks, vulnerabilities, deterrence, and crisis management. These perspectives remain indispensable for understanding particular dimensions of contemporary security. Nevertheless, none of them places the *security problem itself* at the center of scientific inquiry (Trifunović, 2025, pp. 8–34).

Security Science adopts this perspective precisely. It investigates how security problems emerge, how they develop, how they influence the security condition of states and societies, and how they may be prevented, mitigated, or resolved through coordinated institutional action. Consequently, hybrid activities are analyzed not simply as instruments of political competition or components of hybrid warfare but as dynamic processes capable of generating measurable security problems.

This theoretical distinction also determines the methodological orientation of the present research. If the scientific object is the security problem rather than merely threats or risks, scientific analysis must identify the observable manifestations that precede or accompany its emergence. These manifestations constitute *security indicators*.

Security indicators therefore become the empirical bridge connecting hybrid activities with objectively identifiable security problems. Through systematic observation of political, informational, intelligence, cyber, economic, social, and institutional indicators, it becomes possible to determine not only whether hybrid activities are occurring, but also whether they are producing measurable changes in the security condition of the observed state.

From this perspective, the development of a *Hybrid Activity Indicator System (HAIS)* represents a natural methodological extension of Security Science. Rather than measuring hybrid activities as isolated events, the proposed framework measures the observable indicators through which hybrid activities progressively generate security problems (Gill & Phythian, 2018, pp. 154–168). Such an approach enables earlier detection, objective assessment, comparative analysis, and more effective institutional decision-making than models that rely exclusively on qualitative assessments of threats, risks, or challenges. Therefore, the central argument of this article is that the scientific study of hybrid activities should not begin with the question of *who conducts hybrid operations*, nor solely with *which instruments are employed*, but with the identification of *which measurable security indicators demonstrate that hybrid activities are generating security problems capable of altering the security condition of the state*.

HYBRID ACTIVITIES AS GENERATORS OF SECURITY PROBLEMS

The growing complexity of contemporary security environments requires a reconsideration of how hybrid activities are conceptualized within security research. Existing literature generally treats hybrid activities either as a specific category of threats, as components of hybrid warfare, or as instruments employed by state and non-state actors to influence political processes below the threshold of conventional armed conflict. While these perspectives successfully explain the operational characteristics of hybrid activities, they do not fully explain their position within the broader dynamics of security development.

From the perspective of Security Science, hybrid activities should not be understood merely as isolated actions or hostile instruments. Rather, they represent *dynamic multidimensional processes capable of generating security problems* (Hoffman, 2007, pp. 7–15). Their significance therefore does not primarily derive from the individual activities themselves but from their cumulative influence on the overall security condition of a state or society.

Unlike conventional military aggression, hybrid activities rarely produce immediate and easily identifiable consequences (Rid, 2020, pp. 1–18). Their effectiveness is based precisely on graduality, simultaneity, ambiguity, and the interaction of multiple instruments operating across different sectors. Political influence campaigns, coordinated disinformation, cyber intrusions, intelligence penetration, economic pressure, organized criminal cooperation, social polarization, legal manipulation, and strategic corruption often develop independently at first glance. However, when analyzed as interconnected processes, they reveal a coordinated pattern capable of progressively transforming the security environment.

This cumulative character distinguishes hybrid activities from many traditional security phenomena. Individual cyber-attacks, isolated disinformation campaigns, sporadic acts of political interference, or limited economic pressure do not necessarily constitute security problems. Nevertheless, their continuous interaction gradually weakens institutional resilience, undermines public trust, increases political instability, damages critical infrastructure, polarizes society, and reduces the capacity of state institutions to perform their constitutional functions. At this stage, hybrid activities cease to represent isolated incidents and begin producing objectively observable security problems. Security Science therefore introduces a different analytical sequence than that commonly found in contemporary security literature. Traditional approaches frequently analyze hybrid activities according to the following logic:

Hybrid Activities → Hybrid Threats → Institutional Response

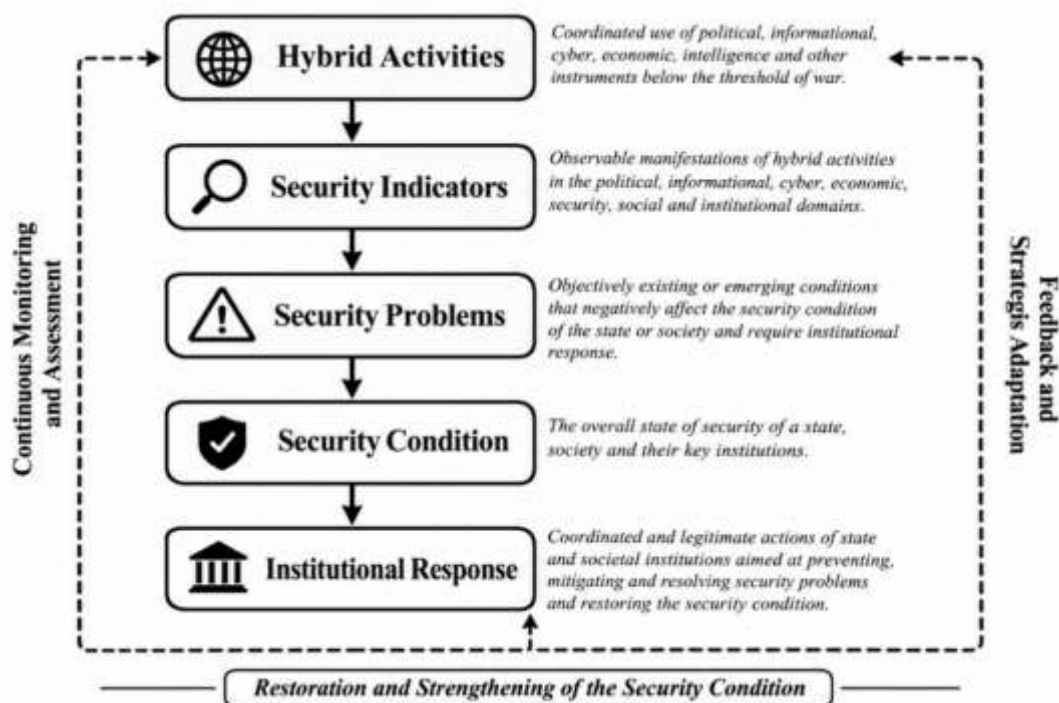
Although operationally useful, such an approach primarily focuses on hostile actions and the responses they provoke. Security Science proposes a broader analytical framework:

Hybrid Activities → Security Indicators → Security Problems → Security Condition → Institutional Response

Within this model, hybrid activities represent the initiating process rather than the final analytical category. Their development generates numerous observable changes within political, informational, cyber, intelligence, economic, legal, social, and institutional environments. These observable changes constitute *security indicators*. When analyzed collectively, these indicators reveal the emergence of objectively existing security problems. The accumulation of multiple security problems subsequently influences the overall security condition of the state, requiring appropriate institutional responses aimed not merely at countering individual hybrid activities but at restoring the security condition itself. This conceptual distinction significantly expands the possibilities of intelligence and security analysis. Intelligence services traditionally collect extensive quantities of information concerning cyber-attacks, foreign influence operations, extremist activities, economic pressure, organized crime, espionage, or strategic communications. However, these observations frequently remain fragmented because they are analyzed within separate institutional or sectoral frameworks. Security Science provides an integrative analytical perspective by treating these apparently unrelated observations as components of a common process that can generate security problems. Consequently, the scientific challenge is no longer limited to determining whether individual hybrid activities have occurred but extends to identifying whether their interaction is producing measurable deterioration of the security condition. This perspective also explains why early detection is fundamentally dependent upon the identification of security indicators. Security

problems rarely emerge suddenly. They evolve progressively through numerous observable manifestations whose cumulative interaction precedes major crises. Coordinated information manipulation may reduce public confidence in institutions. Persistent cyber intrusions may reveal systematic intelligence preparation. Economic coercion may gradually weaken strategic sectors. Organized criminal structures may increasingly cooperate with political actors. Foreign intelligence activities may intensify influence operations within governmental institutions. Each of these developments individually may appear manageable. Collectively, however, they indicate that hybrid activities are generating a broader security problem.

Accordingly, the scientific value of analyzing hybrid activities does not lie solely in identifying hostile actors or describing their operational methods. It lies in understanding the causal relationship between hybrid activities, the emergence of measurable security indicators, the formation of security problems, and the subsequent deterioration of overall security. This causal relationship constitutes the theoretical foundation upon which the present research develops the *Hybrid Activity Indicator System (HAIS)*. Rather than evaluating hybrid activities as isolated security events, the proposed framework identifies the observable indicators through which hybrid activities progressively generate security problems affecting political stability, institutional functionality, social cohesion, economic resilience, critical infrastructure, and national security. The proposed approach therefore shifts the analytical focus from reactive identification of completed hybrid operations toward proactive recognition of the processes through which hybrid activities transform the security environment. Such a perspective provides not only a stronger theoretical foundation for understanding hybrid activities but also a practical methodological basis for intelligence analysis, strategic early warning, national security assessment, and evidence-based decision-making.



Source: Author

Development of the Hybrid Activity Indicator System (HAIS)

Conceptual Foundations

Identifying hybrid activities presents one of the greatest analytical challenges within contemporary security research. Unlike conventional military aggression, hybrid activities rarely manifest themselves through a single identifiable event. Instead, they emerge gradually through the interaction of numerous political, informational, cyber, intelligence, economic, legal, and societal activities whose cumulative effects progressively alter the security environment.

Consequently, the scientific analysis of hybrid activities requires more than identifying individual hostile actions. It requires developing a standardized methodological framework capable of transforming dispersed observations into measurable indicators of the evolving security environment (International Organization for Standardization [ISO], 2018). Within the framework of *Security Science*, these indicators represent observable manifestations of processes that can generate security problems. Their scientific value lies not only in describing current developments but also in supporting early warning, strategic forecasting, intelligence assessment, and evidence-based institutional decision-making.

Accordingly, this article proposes the *Hybrid Activity Indicator System (HAIS)* as an integrated analytical model for identifying, classifying, and assessing measurable indicators generated by hybrid activities. Unlike existing descriptive approaches, HAIS does not attempt to measure hybrid warfare itself. Rather, it measures the observable indicators produced by hybrid activities before they develop into major security problems affecting the overall security condition of the state.

Principles of Indicator Selection

The proposed indicator system is based upon five methodological principles.

First Principle – Observability

Every indicator must represent an objectively observable phenomenon. Indicators cannot be based upon assumptions or political interpretations but must originate from measurable developments that can be verified through intelligence, governmental, academic, media, or open-source information.

Second Principle – Relevance

Each indicator must demonstrate a direct relationship with hybrid activities. General political instability or ordinary criminality should not automatically be interpreted as hybrid activity unless objective evidence indicates strategic coordination or external influence.

Third Principle – Comparability

Indicators must enable comparison:

- between different states,
- between different periods,
- between different hybrid campaigns.

Only comparable indicators can support scientific analysis.

Fourth Principle – Measurability

Every indicator should be capable of quantitative or qualitative assessment.

For example:

- Frequency
- Intensity
- Duration
- Geographical distribution
- Institutional impact

Fifth Principle – Predictive Value

Indicators should contribute to forecasting. The purpose of HAIS is not merely to explain completed hybrid campaigns but to identify their earliest manifestations before major security problems emerge.

Classification of Hybrid Activity Indicators

The proposed model identifies seven principal categories of indicators.

A. Information Indicators (II)

These indicators measure activities intended to manipulate public perception.

Examples include:

- coordinated disinformation campaigns;
- synchronized media narratives;
- AI-generated propaganda;
- deepfake dissemination;
- bot network activity;
- coordinated social media amplification;
- attacks on public confidence.

B. Political Indicators (PI)

These indicators reflect attempts to influence political processes.

Typical indicators include:

- election interference;
- foreign political financing;
- institutional delegitimization;
- externally supported protests;
- political radicalization;
- constitutional destabilization.

C. Security and Intelligence Indicators (SI)

These indicators concern activities directly affecting national security institutions.

Examples include:

- espionage activities;
- recruitment attempts;
- foreign intelligence penetration;
- sabotage;
- extremist mobilization;
- paramilitary organization;
- coordinated criminal-security cooperation.

D. Cyber Indicators (CI)

Cyber activities frequently represent the earliest manifestations of hybrid campaigns.

Indicators include:

- attacks on critical infrastructure;
- DDoS campaigns;
- ransomware attacks;
- phishing operations;
- malware deployment;
- cyber espionage;
- attacks on governmental databases.

E. Economic Indicators (EI)

Economic instruments increasingly serve strategic purposes.

Examples include:

- coercive investments;
- strategic acquisition of infrastructure;
- energy dependency exploitation;
- financial pressure;
- sanctions;
- supply chain disruption.

F. Social Indicators (SOI)

Hybrid activities frequently seek to weaken societal cohesion.

Indicators include:

- ethnic polarization;
- religious radicalization;
- organized migration pressure;
- hate speech;
- violent demonstrations;
- declining public trust;
- identity conflicts.

G. International Indicators (INI)

These indicators measure the external dimension of hybrid activities.

Examples include:

- coordinated diplomatic pressure;
- foreign influence operations;
- international disinformation campaigns;
- proxy political organizations;
- strategic alliances supporting destabilization.

Interaction Between Indicator Categories

An isolated indicator rarely demonstrates the existence of coordinated hybrid activities. For example, a cyber-attack, protest, or disinformation campaign may occur independently. However, their simultaneous appearance, temporal synchronization, strategic coherence, and cumulative influence may indicate an organized hybrid campaign. Consequently, HAIS analyzes indicators not individually but as an interconnected analytical system. The interaction among indicator categories therefore represents one of the principal analytical strengths of the proposed framework.

Intelligence Collection and Indicator Verification

The reliability of HAIS depends upon continuous intelligence collection.

Indicators may be obtained through:

- Intelligence reporting;
- Open-source intelligence (OSINT);
- Cyber intelligence;
- Economic intelligence;
- Social media analytics;
- Governmental statistics;
- Law enforcement databases;
- International organizations;
- Academic research.

Every indicator should undergo verification before being incorporated into strategic assessment.

Operational Purpose of HAIS

The proposed system supports five principal analytical functions.

Early warning – identification of emerging hybrid activities.

Trend Analysis – monitoring long-term developments.

Comparative Assessment – comparison among states and regions.

Strategic Forecasting – prediction of future developments.

Decision Support – assistance to intelligence services and national security institutions.

Accordingly, HAIS should be understood not merely as a descriptive classification of hybrid activities but as a methodological instrument integrating observation, measurement, analysis, and strategic assessment within the broader theoretical framework of **Security Science**.

Table 1. Hybrid Activity Indicator System (HAIS)

Categories, Examples, Data Sources and Predictive Value

INDICATOR CATEGORY	PRIMARY DOMAIN	TYPICAL EXAMPLES	MAIN DATA SOURCES	PREDICTIVE VALUE
A. INFORMATION INDICATORS (II)	Information Environment	- Coordinated disinformation campaigns - Bot networks and automated amplification - Deepfakes and AI-generated content - Manipulative media narratives - Attacks on public trust and credibility	- OSINT and social media monitoring - Media content analysis - Fact-checking organizations - Platform analytics - Intelligence reporting	HIGH
B. POLITICAL INDICATORS (PI)	Governance and Political Processes	- Election interference and manipulation - Foreign political financing and lobbying - Delegitimization of institutions and leaders - Organized or externally supported protests - Political radicalization and polarization	- Government reports - Election monitoring bodies - Financial intelligence reports - Media and OSINT - Diplomatic reporting	HIGH
C. SECURITY & INTELLIGENCE INDICATORS (SI)	National Security and Intelligence Environment	- Espionage and intelligence collection - Recruitment and subversion attempts - Subsistence and acts of damage - Extremist mobilization and training - Paramilitary activities and proxy groups - Cooperation between organized crime and security actors	- Intelligence and counterintelligence reporting - Law enforcement databases - Defense and security agencies - Border and customs reports - International intelligence sharing	VERY HIGH
D. CYBER INDICATORS (CI)	Cyber and Digital Infrastructure	- DDoS attacks on critical systems - Ransomware and destructive malware - Phishing and credential theft campaigns - Cyber espionage and data exfiltration - Attacks on governmental databases - Attacks on critical infrastructure (energy, transport, communications)	- National CERTs and CSIRTs - SOC (Security Operations Centers) - Cyber threat intelligence platforms - IT and infrastructure operators - International cyber sharing networks	VERY HIGH
E. ECONOMIC INDICATORS (EI)	Economy and Economic Security	- Strategic investments in key sectors - Acquisition of critical infrastructure - Energy dependency and supply manipulation - Financial pressure and economic coercion - Sanctions and trade restrictions - Disruption of supply chains	- Financial intelligence units - Central bank and economic data - Trade and customs statistics - Energy market reports - International organizations (IMF, WB, etc.)	MEDIUM-HIGH
F. SOCIAL INDICATORS (SOI)	Society and Social Cohesion	- Ethnic and religious polarization - Hate speech and incitement to violence - Organized migration pressure - Violent demonstrations and riots - Decline in public trust and social cohesion - Identity-based conflicts	- Public opinion surveys - Social media analytics - Government and NGO reports - Demographic and migration data - Academic and research institutions	MEDIUM
G. INTERNATIONAL INDICATORS (INI)	International Relations and External Influence	- Coordinated diplomatic pressure - Foreign influence through proxy organizations - International disinformation campaigns - Support for destabilizing actors - Manipulation in international organizations	- Diplomatic reporting - International organizations - Media and OSINT - Intelligence reporting - Academic and think tank analyses	HIGH

Note: Predictive Value indicates the general ability of each indicator category to provide early warning of the emergence of security problems.

Source: Author

The proposed classification is not arbitrary. The seven indicator categories were derived from the principal domains through which contemporary hybrid activities most consistently manifest themselves in the security environment. Together, they encompass the major political, informational, intelligence, cyber, economic, social, and international dimensions repeatedly identified throughout the preceding theoretical analysis. While additional subcategories may be developed in future research, these seven domains provide a sufficiently comprehensive analytical framework for the systematic identification and classification of hybrid activity indicators.

Hybrid Activity Indicator Score (HAIS Score): A Methodological Framework for Measuring Hybrid Activities

Measuring Hybrid Activities

The identification of hybrid activities represents only the first step of scientific analysis. The subsequent challenge is to develop a methodological framework capable of assessing their overall intensity, comparing different security environments, monitoring changes over time, and supporting strategic forecasting. Existing approaches predominantly rely upon qualitative assessments produced by experts or intelligence analysts. Although such assessments remain indispensable for operational decision-making, they frequently lack standardized criteria that would enable objective comparison among states, time periods, or different hybrid campaigns. From the perspective of Security Science, scientific analysis requires not only the identification of observable security indicators but also their systematic integration into a measurable analytical framework. Such a framework should enable the transformation of heterogeneous information collected across multiple security domains into a single analytical assessment reflecting the overall level of hybrid activity directed against a particular state or security system. To address this methodological requirement, this article proposes the **Hybrid Activity Indicator Score (HAIS Score)** as a composite analytical instrument integrating the seven categories of security indicators developed in the previous chapter. Unlike existing qualitative models, HAIS Score provides a structured methodology for measuring the cumulative intensity of hybrid activities while preserving the multidimensional nature of the observed security environment.

Structure of the HAIS Score

The Hybrid Activity Indicator Score consists of seven principal indicator groups:

- Information Indicators (II)
- Political Indicators (PI)
- Security and Intelligence Indicators (SI)
- Cyber Indicators (CI)
- Economic Indicators (EI)
- Social Indicators (SOI)
- International Indicators (INI)

Each category reflects one dimension of hybrid activities while simultaneously interacting with all other categories. Consequently, the HAIS Score evaluates not isolated events but the cumulative interaction of all observable indicators. The five-level assessment scale was selected because it provides sufficient analytical differentiation while remaining operationally practical for intelligence and security assessment. The proposed threshold values are conceptual and methodological rather than statistically derived. Their empirical calibration and possible adjustment will depend upon future longitudinal applications of the HAIS framework.

Basic Analytical Formula

The simplest representation of the proposed model is:

$$\mathbf{HAIS} = \mathbf{II} + \mathbf{PI} + \mathbf{SI} + \mathbf{CI} + \mathbf{EI} + \mathbf{SOI} + \mathbf{INI}$$

where:

- **II** – Information Indicators
- **PI** – Political Indicators
- **SI** – Security and Intelligence Indicators
- **CI** – Cyber Indicators

- **EI** – Economic Indicators
- **SOI** – Social Indicators
- **INI** – International Indicators

This basic equation expresses the cumulative character of hybrid activities. However, empirical research demonstrates that not every indicator possesses identical analytical significance. Therefore, a weighted model is proposed.

Analytical Assessment of Hybrid Activity Indicators

The proposed Hybrid Activity Indicator System (HAIS) is based on the assumption that hybrid activities generate observable changes within the security environment long before they evolve into major security problems. Consequently, the scientific task is not limited to identifying isolated hostile actions but to recognizing measurable indicators whose interaction reveals coordinated hybrid activities. Unlike conventional military aggression, hybrid activities rarely produce immediate and unequivocal manifestations. Instead, they gradually affect different sectors of society through coordinated political, informational, intelligence, cyber, economic, legal, and social activities. Each indicator may appear insignificant when analyzed independently; however, their simultaneity, continuity, interaction, and cumulative impact frequently reveal the emergence of a complex hybrid campaign. Within the proposed HAIS framework, hybrid activities are assessed through seven categories of security indicators.

Information Indicators: reflect attempts to manipulate public perception, influence collective behavior, and undermine confidence in state institutions. Typical indicators include coordinated disinformation campaigns, synchronized media narratives, bot network activity, artificial intelligence-generated propaganda, deepfake content, coordinated hashtag campaigns, manipulation of social media algorithms, attacks against the credibility of public institutions, dissemination of false information concerning national security issues, and coordinated psychological influence campaigns. These indicators are primarily monitored through media analysis, OSINT, social media analytics, strategic communication monitoring, fact-checking platforms, and intelligence reporting. Their principal analytical value lies in identifying early attempts to influence public opinion and prepare the information environment for subsequent hybrid activities.

Political Indicators: reveal attempts to influence political decision-making and institutional stability. Observable manifestations include foreign financing of political actors, election interference, externally coordinated protest movements, constitutional destabilization, political radicalization, systematic attacks against democratic institutions, coordinated campaigns questioning electoral legitimacy, foreign lobbying directed against national interests, and increasing polarization supported through external influence. Primary sources include governmental institutions, electoral commissions, parliamentary reports, judicial proceedings, intelligence reporting, and diplomatic information. These indicators provide early evidence that hybrid activities are targeting the political decision-making process.

Security and intelligence indicators: concern activities directly affecting the national security system. Observable indicators include espionage operations, recruitment of confidential sources, foreign intelligence penetration, sabotage, terrorist facilitation, extremist mobilization, creation of paramilitary groups, cooperation between organized crime and foreign intelligence services, attacks against police and intelligence institutions, unauthorized collection of classified information, and attempts to compromise security personnel. Information is obtained through intelligence services, counterintelligence

investigations, law enforcement agencies, border security authorities, judicial proceedings, and classified operational reporting. Their analytical significance is particularly high because they frequently demonstrate direct hostile intent directed against the security system.

Cyber Indicators: represent one of the fastest-growing dimensions of hybrid activities. Typical manifestations include distributed denial-of-service attacks, ransomware campaigns, cyber espionage, phishing operations, malware deployment, attacks against governmental databases, intrusions into critical infrastructure, manipulation of digital communications, attacks against election systems, financial cybercrime directed against strategic institutions, and coordinated cyber influence operations. Primary sources include national CERT organizations, cybersecurity agencies, intelligence services, governmental cyber response centers, private cybersecurity companies, and digital forensic investigations. Cyber indicators frequently provide the earliest technical evidence that coordinated hybrid activities are underway.

Economic Indicators: demonstrate attempts to influence national resilience through financial and economic instruments. Observable indicators include strategic acquisition of critical infrastructure, coercive foreign investments, manipulation of energy supplies, sanctions, disruption of supply chains, illicit financial transfers, strategic dependence on foreign technologies, market manipulation, corruption directed toward strategic sectors, and financial support provided to organizations acting contrary to national security interests. Information is obtained through financial intelligence units, central banks, ministries of economy, customs authorities, anti-corruption agencies, and international financial institutions. Economic indicators reveal long-term attempts to increase structural dependence and reduce strategic autonomy.

Social indicators reflect efforts to weaken societal cohesion and increase internal instability. Typical indicators include ethnic polarization, religious radicalization, organized migration pressure, hate speech, coordinated social unrest, identity-based conflicts, declining confidence in state institutions, manipulation of historical narratives, growing social fragmentation, radicalization through digital platforms, and increasing societal intolerance encouraged through coordinated influence campaigns. Primary sources include public opinion surveys, academic research, media monitoring, police statistics, civil society organizations, demographic analyses, and intelligence assessments. Their analytical significance lies in identifying gradual deterioration of social resilience before it becomes visible through large-scale crises.

International indicators reveal the external dimension of hybrid activities. Observable manifestations include coordinated diplomatic pressure, foreign influence campaigns conducted through international organizations, strategic communication directed from abroad, proxy political organizations, coordinated legal actions, international lobbying against national strategic interests, synchronized diplomatic initiatives, foreign support for destabilizing actors, and organized attempts to influence international perception of the targeted state. These indicators are monitored through diplomatic reporting, international organizations, intelligence services, ministries of foreign affairs, academic research, and international media monitoring. Their analytical value lies in identifying the international coordination frequently accompanying domestic hybrid activities.

Operational Assessment of Indicators

The existence of individual indicators should never be interpreted in isolation. Their analytical significance depends upon five complementary assessment criteria. The **first**

criterion concerns **intensity**, referring to the scale, geographical distribution, operational sophistication, and strategic significance of observed activities. The **second criterion** concerns **frequency**, measuring how often particular indicators occur and whether they demonstrate continuity over time rather than isolated occurrence. The **third criterion** concerns **interconnectivity**, assessing whether indicators from different domains occur simultaneously, support common strategic objectives, or demonstrate operational coordination. The **fourth criterion** concerns **institutional impact**, evaluating the extent to which hybrid activities influence governmental institutions, intelligence services, critical infrastructure, electoral processes, economic stability, judicial institutions, or public confidence. The **fifth criterion** concerns **evolution over time**, analyzing whether observed indicators demonstrate expansion into additional security domains, increasing sophistication, greater geographical coverage, or progressive deterioration of the overall security condition. Accordingly, the analytical assessment proposed within HAIS does not evaluate isolated events but recognizes coordinated patterns of activity that can generate objectively identifiable security problems. This approach enables intelligence services and security institutions to distinguish ordinary political, economic, or social developments from strategically coordinated hybrid activities directed against the security condition of the state.

Indicator Assessment Scale

Each indicator category may be evaluated using a standardized five-level scale.

Score Interpretation		Security Assessment
1	Very Low	No observable hybrid activity
2	Low	Isolated activities
3	Moderate	Coordinated activities in several sectors
4	High	Systematic multidomain hybrid campaign
5	Very High	Extensive hybrid pressure affecting national security

The cumulative score provides an analytical estimate of the overall level of hybrid activity.

Interpretation of HAIS Score

The total HAIS Score may be interpreted according to the following analytical thresholds.

Total Score Hybrid Activity Level	
0–20	Minimal
21–40	Low
41–60	Moderate
61–80	High
81–100	Critical

These thresholds should be understood as a methodological framework rather than absolute values. Their empirical calibration requires longitudinal testing using real-world data from different security environments.

Analytical Value of HAIS

The principal advantage of HAIS lies in its ability to integrate heterogeneous security observations into a unified analytical assessment. Instead of examining cyber-attacks, disinformation campaigns, political interference, intelligence operations, economic coercion, or social polarization independently, the proposed framework evaluates their cumulative interaction within a single analytical model. Consequently, HAIS supports

several complementary analytical functions. First, it enables **early warning** by identifying increasing concentrations of security indicators before major crises become visible. Second, it facilitates **trend analysis**, allowing analysts to monitor whether hybrid pressure is increasing, decreasing, or remaining stable over time. Third, it supports **comparative assessment** among different countries or regions by applying standardized analytical criteria. Fourth, it contributes to **strategic forecasting**, since increasing concentrations of hybrid indicators frequently precede significant security deterioration. Finally, HAIS provides a transparent methodological framework for supporting intelligence assessments, governmental decision-making, and national security planning.

Relationship Between HAIS and Security Science

Within the theoretical framework proposed in this article, HAIS should not be understood as a measure of hybrid warfare itself. Rather, it measures the observable manifestations of hybrid activities that progressively generate security problems. This distinction reflects the fundamental epistemological position of Security Science. Political Science may analyze power relations. International Relations may explain interstate competition. Security Studies frequently examine threats, risks, and challenges. Security Science, however, investigates how these and other processes generate *security problems* and influence the *state's security condition*. Consequently, HAIS represents the methodological instrument through which hybrid activities become scientifically observable within the analytical framework of Security Science. In this sense, HAIS complements rather than replaces existing approaches by providing a standardized methodology for measuring one of the principal generators of contemporary security problems (Trifunović, 2025, pp. 8–34).

The present article introduces the conceptual and methodological architecture of the Hybrid Activity Indicator System (HAIS) and the Hybrid Activity Indicator Score (HAIS Score). The proposed framework is intended as a standardized analytical model grounded in the theoretical principles of Security Science. Its primary objective is to establish a coherent methodology for identifying, classifying, and assessing observable indicators of hybrid activities.

Future research should focus on the empirical validation of the proposed model through longitudinal case studies, comparative cross-national applications, statistical testing of indicator reliability, and the development of weighting procedures based on empirical evidence. Such validation will enable further refinement of the analytical thresholds and strengthen the practical applicability of HAIS as a standardized instrument for intelligence analysis, strategic forecasting, and security assessment.

CONCLUSION

Hybrid activities have become one of the most complex and persistent features of the contemporary security environment. Their multidimensional nature, gradual development, and operation below the threshold of conventional armed conflict make them particularly difficult to identify, assess, and counter using traditional analytical approaches. Although existing literature has significantly advanced the understanding of hybrid threats, hybrid warfare, and grey-zone competition, comparatively little attention has been devoted to developing standardized methodologies capable of identifying measurable indicators that precede the emergence of major security problems.

This article addressed that methodological gap by approaching hybrid activities from the perspective of **Security Science**. Unlike approaches primarily focused on threats, risks, and challenges, Security Science places the **security problem** at the center of scientific inquiry. Within this framework, hybrid activities are not understood as security problems themselves but as dynamic multidomain processes capable of generating security problems through the cumulative interaction of political, informational, security and intelligence, cyber, economic, social, and international activities.

The principal contribution of this research is the development of the **Hybrid Activity Indicator System (HAIS)** as a comprehensive methodological framework for identifying, classifying, and assessing observable indicators of hybrid activities. By introducing seven categories of indicators, defining methodological principles for their selection and assessment, and proposing the **Hybrid Activity Indicator Score (HAIS Score)**, the article demonstrates that hybrid activities can be analyzed through a structured system of measurable security indicators rather than exclusively through descriptive or qualitative approaches.

The qualitative application of HAIS to the regional security environment of Southeast Europe further illustrates the practical applicability of the proposed framework. The assessment indicates a high overall presence of hybrid activity indicators across the region, confirming that hybrid activities should be understood as a persistent characteristic of the regional security environment rather than as isolated incidents. At the same time, the proposed model emphasizes that the analytical significance of hybrid activities derives not from individual events but from the interaction, continuity, and cumulative effect of multiple indicators that can generate objectively identifiable security problems.

Beyond its practical application, HAIS contributes to the further theoretical development of **Security Science**. It demonstrates how security indicators can serve as the empirical link between hybrid activities and security problems, thereby strengthening Security Science's methodological capacity to identify, assess, compare, and predict contemporary security phenomena. In this sense, HAIS complements existing research on hybrid activities by shifting analytical attention from describing hybrid operations toward measuring their observable manifestations and their influence on the security condition of states and societies.

Future research should focus on the empirical validation of HAIS through longitudinal studies, quantitative weighting of individual indicators, comparative application across different regions, and the integration of artificial intelligence and big data analytics into hybrid activity monitoring. Such developments would enable the transformation of HAIS from a conceptual methodological framework into a standardized analytical instrument supporting intelligence analysis, strategic forecasting, national security planning, and evidence-based decision-making.

Accordingly, the article argues that the systematic identification and assessment of hybrid activity indicators represent not only an operational requirement for contemporary security institutions but also an important step in the continuing methodological evolution of **Security Science** as an autonomous scientific discipline.

References

- Aven, T. (2015). *Risk Analysis* (2nd ed.). Wiley.
- Baldwin, D. A. (1997). The concept of security. *Review of International Studies*, 23(1), 5–26. <https://doi.org/10.1017/S0260210597000053>
- European Commission, & High Representative of the Union for Foreign Affairs and Security Policy. (2016). *Joint framework on countering hybrid threats: A European Union response* (JOIN(2016) 18 final). European Commission. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=JOIN%3A2016%3A18%3AFIN>
- Gill, P., & Phythian, M. (2018). *Intelligence in an Insecure World* (3rd ed.). Polity Press.
- Hoffman, F. G. (2007). *Conflict in the 21st Century: The Rise of Hybrid Wars*. Potomac Institute for Policy Studies.
- International Organization for Standardization. (2018). *ISO 31000:2018 Risk Management—Guidelines*. ISO.
- NATO. (2022). *NATO 2022 Strategic Concept*. North Atlantic Treaty Organization. <https://www.nato.int/strategic-concept/>
- Rid, T. (2020). *Active Measures: The Secret History of Disinformation and Political Warfare*. Farrar, Straus and Giroux.
- Trifunović, D. (2025). Defining Security Science and Introducing Security Science Indexation (SSI). *Security Science Journal*, 6(2), 7–36. <https://doi.org/10.37458/ssj.6.2.14>
- Wolfers, A. (1952). "National Security" as an Ambiguous Symbol. *Political Science Quarterly*, 67(4), 481–502.

UDK: [57.087.1:004.93]:355.4-025.26(497-15)

[57.087.1:004.93]:340.13(497-15)

DOI 10.5281/zenodo.22638396

AUTHENTICATION TO SURVEILLANCE: LEGAL AND SECURITY IMPLICATIONS OF BIOMETRIC TECHNOLOGIES IN HYBRID THREAT ENVIRONMENTS

Emir Kremić

Ph.D Sc., Analytics Armor LLC

emir@analyticsarmor.com, kremicemir@gmail.com

Haris Muminovic

Ph.D Sc., University of Donja Gorica, Podgorica, Montenegro

haris.muminovic@udg.edu.me

Edita Dapo

Ph.D Sc., The Agency for Identification Documents, Registers, and Data

Exchange of Bosnia and Herzegovina

edjapo@hotmail.com

Abstract

This paper explores the emergence of hybrid threats in information technology, focusing on biometric systems and facial recognition technologies. It analyzes how the evolution of biometric authentication from early pattern recognition models to AI-driven facial identification has transformed both cybersecurity capabilities and threat landscapes. The study examines legal frameworks, particularly the General Data Protection Regulation (GDPR) and the EU AI Act, highlighting regulatory tensions between innovation and privacy. Drawing on recent research and case studies, including empirical work on mobile facial recognition and machine learning-based biometric classification, the paper evaluates risks such as mass surveillance, identity theft, and algorithmic bias. It concludes by proposing a multi-layered governance model integrating cybersecurity, legal compliance, and ethical safeguards.

Key words: Biometric authentication, Facial recognition, Hybrid threats, Cybersecurity, AI governance.

1. INTRODUCTION

Biometric technologies have undergone a profound transformation over the past five decades, evolving from isolated authentication mechanisms into foundational components of modern digital identity ecosystems and national security architectures. Early biometric systems were primarily based on physiological identifiers such as fingerprints, hand geometry and iris patterns (Daugman, 1993; Jain, Ross, & Prabhakar, 2004), relying on deterministic matching algorithms and low-dimensional statistical models designed for access control and identity verification purposes. The first significant technological breakthrough occurred during the emergence of computational pattern recognition in the 1980s and 1990s, when biometric identification systems began integrating feature

extraction methods such as Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), Eigenfaces and Fisherfaces. These approaches transformed biometric recognition from simple image comparison into a multidimensional classification problem operating within mathematical feature spaces.

The subsequent development of machine learning algorithms, particularly Support Vector Machines (SVM), Random Forest classifiers and kernel-based optimization methods, substantially improved recognition accuracy, robustness to environmental variations and scalability in operational environments. These advances enabled biometric systems to transition from laboratory research environments into real-world applications including border management, critical infrastructure protection, financial services and mobile authentication systems. The proliferation of mobile computing platforms and Linux-based operating systems, including secure enterprise environments built upon SUSE technologies and embedded Linux architectures, further accelerated the adoption of biometric authentication mechanisms. Biometric capabilities evolved from centralized verification systems toward distributed, edge-based and cloud-supported identity infrastructures capable of real-time authentication and continuous behavioural verification. The evolution from authentication toward large-scale biometric surveillance represents one of the defining cybersecurity and legal challenges of contemporary hybrid threat environments. As biometric systems become increasingly integrated into critical infrastructure protection, border security, digital identity frameworks and national resilience strategies, the distinction between legitimate authentication and mass surveillance becomes progressively more difficult to maintain. Consequently, modern biometric governance requires the simultaneous consideration of cybersecurity resilience, privacy preservation, regulatory compliance and ethical accountability in order to ensure trustworthy deployment of biometric intelligence technologies within democratic societies.

2. HISTORICAL DEVELOPMENT

The historical evolution (Kremić & Subasi, 2015, 2016) of biometric intelligence represents the convergence of advances in mathematics, computer science, signal processing, statistics and artificial intelligence. Modern biometric systems are the result of more than a century of scientific development aimed at answering one of the oldest security questions in human history: how can the identity of an individual be verified with a high degree of certainty? While contemporary biometric technologies rely on sophisticated machine learning algorithms and deep neural architectures, their conceptual foundations can be traced to nineteenth-century criminology and anthropometry. One of the earliest systematic approaches to human identification was introduced in 1883 by the French criminologist Alphonse Bertillon, who developed the Bertillon system based on anthropometric measurements of the human body. The method relied on combinations of physical characteristics such as head dimensions, arm length and other anatomical features to distinguish individuals within criminal records. Although the Bertillon system was eventually replaced by more reliable biometric techniques, it established the fundamental principle that unique physiological characteristics could serve as stable and verifiable identifiers.

A major scientific breakthrough occurred through the work of Sir Francis Galton, whose research in the late nineteenth century demonstrated the uniqueness and permanence of fingerprints throughout an individual's lifetime. Galton's studies provided the first rigorous statistical evidence that fingerprints possess sufficient individuality to support

reliable identification. This work laid the scientific and operational foundations for fingerprint-based authentication systems, which would remain the dominant biometric modality throughout most of the twentieth century and continue to be widely used in law enforcement and border management systems today. The digital transformation of biometrics began during the emergence of pattern recognition and statistical learning theory in the 1960s and 1970s. Early facial recognition systems relied primarily on geometric measurements of facial landmarks, including the relative distances between the eyes, nose, mouth and jawline. Significant contributions were made by researchers such as Woodrow W. Bledsoe, whose pioneering work in automated facial recognition established many of the principles later adopted by machine vision systems. Although computational limitations restricted the practical deployment of these early systems, they demonstrated that facial characteristics could be represented mathematically and processed algorithmically.

The introduction of Principal Component Analysis (PCA) into facial recognition during the late 1980s and early 1990s represented one of the most important milestones in biometric intelligence. The development of the Eigenfaces method by Matthew Turk and Alex Pentland (Turk & Pentland, 1991) transformed facial recognition from geometric comparison into a statistical representation problem operating within multidimensional vector spaces. PCA-based approaches significantly reduced computational complexity by projecting facial images into lower-dimensional feature spaces while preserving discriminative information essential for classification. This breakthrough enabled the practical deployment of computational facial recognition systems and established the foundations for modern feature extraction techniques. Subsequent advances introduced Linear Discriminant Analysis (LDA) and the Fisherfaces method (Belhumeur, Hespanha, & Kriegman, 1997), building on earlier modular eigenspace approaches (Pentland, Moghaddam, & Starner, 1994), which improved robustness against variations in illumination, facial expressions and image acquisition conditions. These methods demonstrated superior performance in unconstrained environments and became essential components of many hybrid biometric architectures deployed during the early years of digital identity systems.

The emergence of Support Vector Machines (SVM) represented another major turning point in biometric classification. Developed by Vladimir Vapnik and Corinna Cortes during the mid-1990s (Cortes & Vapnik, 1995), SVM algorithms introduced the concept of optimal hyperplane separation within high-dimensional feature spaces. Through the use of kernel functions such as linear, polynomial and radial basis functions, SVM classifiers achieved remarkable performance in facial recognition tasks, particularly in environments characterized by limited training datasets and constrained computational resources. Their ability to generalize effectively while minimizing overfitting made them highly attractive for real-world biometric deployments.

The subsequent development of ensemble learning methods, particularly the Random Forest algorithm introduced by Leo Breiman (Breiman, 2001), further improved classification robustness and accuracy. Random Forest methods demonstrated strong resistance to noise and overfitting while maintaining relatively low computational complexity. These characteristics made them particularly suitable for mobile biometric applications and resource-constrained edge-computing environments. The rapid expansion of mobile computing during the early 2000s fundamentally altered the architecture of biometric systems. The adoption of Linux-based operating systems and enterprise-grade infrastructures, including SUSE Linux Enterprise Server environments, enabled secure deployment of distributed authentication services across governmental institutions,

financial infrastructures and telecommunication networks. Open-source technologies facilitated interoperability, scalability and the implementation of cryptographic protections within biometric processing pipelines.

The widespread adoption of smartphones transformed biometrics from specialized security solutions into consumer technologies deployed at a global scale. The integration of fingerprint recognition and later facial authentication into mobile operating systems created one of the largest biometric ecosystems in human history and fundamentally changed public perceptions of digital identity verification. The next technological revolution emerged through deep learning and artificial intelligence. The work of Geoffrey Hinton, Yann LeCun and Yoshua Bengio (Hinton, Deng, Yu, et al., 2012; LeCun, Bengio, & Hinton, 2015) enabled the transition from handcrafted feature extraction toward automatically learned feature representations through deep neural architectures. Convolutional Neural Networks, FaceNet embeddings (Schroff, Kalenichenko, & Philbin, 2015) and transformer-based models dramatically increased recognition accuracy and enabled biometric systems to operate effectively in unconstrained environments characterized by variations in pose, illumination and occlusion.

This technological progress transformed biometric systems from authentication mechanisms into strategic intelligence assets capable of supporting border security, critical infrastructure protection, digital identity ecosystems and national security operations. At the same time, these developments introduced entirely new legal, ethical and geopolitical challenges associated with mass surveillance, algorithmic bias and hybrid threat environments.

Within this broader historical context, research contributions by Emir Kremić and collaborators in mobile facial recognition architectures, machine learning-based biometric classification and secure client-server authentication infrastructures (Kremić & Subasi, 2015, 2016) contributed to the practical deployment of biometric intelligence within distributed and resource-constrained environments. Their work demonstrated the applicability of Support Vector Machines and Random Forest approaches in operational environments while emphasizing scalability, efficiency and security. Today, biometric intelligence stands at the intersection of cybersecurity, artificial intelligence, privacy engineering and national resilience, representing one of the most strategically significant technological domains of the twenty-first century and an increasingly important component of international cybersecurity governance and trustworthy AI deployment.

3. LEGAL ANALYSIS

The legal governance of biometric technologies has emerged as one of the most complex regulatory challenges of the digital era. Unlike conventional personal information such as names, addresses or passwords, biometric identifiers possess a unique legal and technical characteristic: they are intrinsically linked to the human body and cannot be meaningfully altered, revoked or reissued once compromised. A stolen password may be changed within minutes; a compromised facial template, iris pattern or fingerprint remains compromised for life.

Recognizing this irreversible nature of biometric identifiers, Article 9 of the General Data Protection Regulation (GDPR) classifies biometric data processed for the purpose of uniquely identifying a natural person as a special category of personal data subject to enhanced legal protection. The regulation establishes a general prohibition on processing

biometric information unless one of several narrowly defined exceptions applies, including explicit consent, substantial public interest, public health protection or national security requirements.

The legal significance of Article 9 extends beyond privacy considerations and increasingly intersects with cybersecurity governance, national resilience strategies and critical infrastructure protection. Modern digital identity systems depend upon trusted authentication mechanisms capable of resisting identity theft, credential compromise and increasingly sophisticated cyber intrusions. Biometric authentication therefore occupies a paradoxical regulatory position: it represents both one of the strongest mechanisms for identity assurance and simultaneously one of the greatest risks to privacy and civil liberties. The adoption of the European Union Artificial Intelligence Act further illustrates this regulatory tension. The AI Act introduces a risk-based governance framework and classifies remote biometric identification systems operating in publicly accessible spaces as high-risk or, in certain circumstances, prohibited AI applications. Real-time facial recognition technologies deployed for generalized surveillance purposes are generally prohibited within the European Union, except under strictly limited exceptions involving the prevention of imminent terrorist attacks, the search for victims of serious crimes or narrowly defined law enforcement investigations authorized by judicial oversight.

The rationale underlying these restrictions reflects growing concerns regarding the transformation of biometric technologies from authentication mechanisms into surveillance infrastructures. Recent international examples demonstrate how facial recognition systems may be repurposed beyond their original security objectives. Large-scale deployments of facial recognition technologies for population monitoring, behavioural profiling and social scoring systems have raised significant concerns regarding proportionality, necessity and fundamental rights protections. Similarly, several documented cybersecurity incidents involving unauthorized access to biometric databases have demonstrated that biometric information itself is becoming a strategic target for state-sponsored threat actors, organized cybercriminal groups and intelligence operations. The emergence of generative artificial intelligence introduces an additional layer of complexity. Deepfake technologies capable of generating highly realistic synthetic faces, voices and behavioural patterns increasingly challenge the evidentiary reliability of biometric systems. Traditional assumptions regarding the uniqueness and trustworthiness of biometric identifiers are being undermined by synthetic identity attacks, adversarial machine learning techniques and presentation attacks targeting facial recognition infrastructures.

From the perspective of hybrid threats, biometric databases represent high-value strategic assets comparable to critical energy infrastructure, telecommunications networks or financial systems. Compromise of national biometric repositories could facilitate espionage operations, disinformation campaigns, covert identity substitution and unauthorized cross-border surveillance activities. Consequently, biometric protection can no longer be understood solely as a data protection issue but must increasingly be viewed as a matter of national cybersecurity and strategic resilience. Several technical and legal safeguards have therefore emerged as best practices for trustworthy biometric governance. These include privacy-by-design architectures (Cavoukian, 2011), template encryption (Ratha, Connell, & Bolle, 2001), fuzzy commitment schemes (Juels & Wattenberg, 1999), federated identity management, differential privacy mechanisms, homomorphic encryption and vector-based biometric matching systems that avoid storing original biometric samples. Equally important are governance mechanisms such as independent algorithmic audits,

mandatory impact assessments, transparency obligations and judicial oversight for high-risk biometric deployments.

The future regulatory challenge will not be whether biometric technologies should be deployed, but rather under which legal conditions, technical safeguards and democratic accountability mechanisms their deployment can remain compatible with fundamental rights, cybersecurity resilience and public trust. Sustainable biometric governance will therefore require an integrated model combining cybersecurity engineering, privacy law, artificial intelligence regulation and ethical oversight capable of preserving both security and individual freedoms within increasingly complex hybrid threat environments.

4. HYBRID THREAT ENVIRONMENT

Biometric databases have evolved into strategic national assets comparable to critical energy infrastructure, financial systems and telecommunications networks. Unlike conventional personally identifiable information, biometric identifiers are immutable, persistent and intrinsically linked to the individual. Consequently, the compromise of biometric repositories creates long-term security implications extending far beyond traditional data breaches (Schneier, 2015; Zuboff, 2019).

Within hybrid threat environments, biometric systems present attractive targets for state-sponsored threat actors, cybercriminal organizations, intelligence services and influence operations seeking to undermine public trust, compromise digital identity infrastructures and conduct covert surveillance activities. The convergence of artificial intelligence, facial recognition and large-scale identity management systems has effectively transformed biometric repositories into high-value strategic objectives within modern cyber conflict (ENISA, 2025; Europol, 2025).

The Western Balkans represent a particularly important case study due to the rapid digitalization of border management, migration control systems and national identity infrastructures (Border Violence Monitoring Network, 2023). Several countries in the region are currently developing interoperable digital identity systems inspired by European interoperability frameworks including EUODAC, EES and the future European Digital Identity Wallet ecosystem (Regional Cooperation Council, 2025). While these systems improve identity verification capabilities and support cross-border law enforcement cooperation, they simultaneously increase the attack surface available to sophisticated threat actors (European Commission, 2024).

Cybersecurity assessments of the Western Balkans demonstrate a growing number of attacks targeting government institutions and digital identity infrastructures. Regional analyses identified numerous incidents affecting public administration, critical infrastructure and identity management systems across Bosnia and Herzegovina, Albania, Kosovo, Montenegro, North Macedonia and Serbia (BIRN, 2025). The increasing integration of biometric identity technologies creates additional systemic risks due to the concentration of sensitive identity data within centralized repositories.

One of the most concerning developments is the emergence of an underground market for biometric information within dark-web ecosystems and cybercriminal forums. Stolen fingerprints, facial templates, iris scans and identity verification photographs are increasingly traded alongside conventional credentials and financial information. Unlike passwords or payment cards, biometric identifiers cannot be replaced following compromise, creating a permanent exposure risk for affected individuals (Europol, 2025).

Several international incidents illustrate the scale of this threat. The 2019 BioStar 2 breach exposed approximately 28 million fingerprint and facial recognition records due to insecure cloud storage configurations, demonstrating the vulnerability of centralized biometric repositories (VpnMentor Research Team, 2019). Similarly, multiple facial recognition service providers have reported unauthorized access to identity verification datasets containing facial images, signatures and identity records later advertised through underground distribution channels and anonymized networks such as Tor (ENISA, 2025). The emergence of generative artificial intelligence further amplifies these risks. Deepfake technologies now enable adversaries to reconstruct highly realistic synthetic identities using stolen biometric information obtained from compromised databases or publicly available social media images. Researchers have demonstrated successful bypass attacks against facial authentication systems using publicly available photographs combined with three-dimensional facial reconstruction techniques (Sharif et al., 2016).

From a hybrid warfare perspective, the compromise of biometric repositories could support espionage operations, covert identity substitution, cross-border influence campaigns and attacks against critical infrastructure personnel. Access to biometric information belonging to military personnel, government officials, intelligence officers or operators of critical infrastructure could facilitate targeted social engineering operations, impersonation attacks and unauthorized access to sensitive facilities (NATO Strategic Communications Centre of Excellence, 2022).

The Western Balkans have already experienced incidents involving unauthorized disclosure of personal information and compromise of governmental databases. In Bosnia and Herzegovina, the national digital identity ecosystem depends heavily upon the Agency for Identification Documents, Registers and Data Exchange of Bosnia and Herzegovina (IDDEEA), which maintains identity cards, passports, residence records and vehicle registries. Although no publicly confirmed compromise of IDDEEA biometric repositories has been reported, the strategic significance of these systems places them among the highest-value targets for advanced persistent threats and state-sponsored cyber operations. Albania represents one of the most significant examples of identity exposure in Southeast Europe. During 2021 and 2022, multiple databases containing personal information of Albanian citizens were leaked and circulated publicly, including national identification numbers, salary information, employment records, vehicle registrations and electoral information affecting more than 637,000 citizens (The Record, 2021). Investigations suggested that the data originated from interconnected governmental databases and administrative systems rather than isolated external intrusions.

The Albanian incidents demonstrated how fragmented personal datasets can be combined to reconstruct complete digital identities, creating ideal conditions for identity theft, synthetic identity generation and future biometric attacks. Security researchers noted that once identity numbers, photographs and administrative records become publicly accessible, attackers may combine these datasets with publicly available social media images and artificial intelligence tools to generate convincing biometric impersonation artifacts.

North Macedonia has experienced multiple cyber incidents affecting public administration and government information systems. Although no publicly confirmed theft of biometric repositories has been reported, European cybersecurity programmes identified the country as requiring enhanced cyber resilience mechanisms for digital identity infrastructures and public sector information systems (EGA, 2024).

Serbia represents one of the most advanced biometric environments in the Western Balkans due to the deployment of large-scale facial recognition and smart surveillance infrastructures in urban environments. Civil society organizations and digital rights groups have repeatedly raised concerns regarding transparency, proportionality and democratic oversight associated with biometric surveillance technologies (BIRN, 2025). While no publicly confirmed compromise of biometric repositories has been disclosed, the concentration of facial recognition infrastructure significantly increases strategic exposure in the event of future cyber incidents.

Particularly sensitive are population censuses and central statistical registers. Census infrastructures collect demographic, economic and geographic information that, when combined with digital identity systems and biometric registries, create highly detailed citizen profiles. International research demonstrates that even anonymized census datasets may be vulnerable to reconstruction attacks capable of re-identifying individuals through statistical correlation techniques and auxiliary datasets (Abowd, 2023).

Threat intelligence reports further indicate the emergence of dark-web markets specializing in so-called "fullz" identity packages containing identity numbers, passport scans, selfies and identity verification photographs. Such information is routinely used for financial fraud, account takeover operations and synthetic identity generation. Although publicly verified cases involving biometric templates originating specifically from Western Balkan government institutions remain limited, international evidence clearly demonstrates that biometric information has become a strategic commodity within cybercriminal economies and hybrid threat ecosystems (Europol, 2025).

5. PROPOSED MODEL

In order to reduce the exposure of biometric and digital identity infrastructures to hybrid threats, this paper proposes an integrated AI-Based Biometric Resilience and Governance Framework for the Western Balkans. The model combines privacy-preserving biometric engineering, statistical risk modelling, artificial intelligence simulation, legal compliance and institutional capacity-building aligned with EU accession requirements. The exposure of biometric and digital identity infrastructures in the Western Balkans to hybrid threats stems from the convergence of two under-examined trends. First, the region has become a testing ground for expanding biometric data collection systems, including interoperable databases modeled on the EU's Eurodac architecture and colloquially referred to as "Balkandac," which a 2023 report by the Border Violence Monitoring Network, supported by Privacy International, found to lack transparency in current regional data-sharing arrangements despite the significant role EU institutions play in their creation. Facial recognition and biometric surveillance technologies are simultaneously becoming more prevalent across the region, with notable developments documented in Albania, Kosovo, Montenegro, and Serbia, even as national legal frameworks nominally aligned with EU standards continue to suffer from weak enforcement and regulatory gaps. Second, and independently, the Western Balkans have been repeatedly identified as structurally vulnerable to hybrid threats and foreign information manipulation. Of 41 European countries assessed in the 2022 Media Literacy Index, Western Balkan states — Albania, Bosnia and Herzegovina, Montenegro, North Macedonia, and Serbia — ranked lowest in their capacity to withstand manipulated information, a vulnerability the EU has sought to address through the Strategic Compass for Security and Defence's EU Hybrid Toolbox, adopted in March 2022 to coordinate preventive, cooperative, and recovery measures

against hybrid campaigns. Despite this, existing governance instruments treat these two exposure surfaces separately: biometric governance models such as the EU's eIDAS 2.0 framework — which establishes a comprehensive digital identity framework requiring every Member State to provide at least one digital identity wallet by the end of 2026 — address trust and interoperability but not adversarial manipulation, while hybrid-threat coordination mechanisms such as the EU-NATO Hybrid CoE architecture address information and cyber resilience but not biometric-specific attack vectors (spoofing, data poisoning, deepfake-enabled identity fraud). This paper addresses that gap by proposing an integrated AI-Based Biometric Resilience and Governance Framework, drawing on the regulatory architecture of eIDAS 2.0, the risk-management structure of the NIST AI Risk Management Framework, and the coordination logic of the EU Hybrid Toolbox, to establish a governance layer specific to biometric and digital identity systems operating in a region where neither eIDAS coverage nor hybrid-threat resilience mechanisms currently extend.

The proposed model is based on four mutually connected layers: (1) biometric data minimization, (2) privacy-preserving vector representation, (3) AI-driven threat simulation and anomaly detection, and (4) legal-governance oversight. The objective is to move Western Balkan governments away from centralized storage of raw biometric images and toward encrypted biometric templates, cancellable biometrics and vectorized identity matching. This approach is consistent with the principles of privacy by design, security by design and risk-based AI governance under the GDPR, the EU AI Act, ENISA cybersecurity guidance and the NIST AI Risk Management Framework.

At the technical level, the model proposes that facial images, fingerprints or iris scans should not be stored as raw biometric objects except where strictly necessary and legally justified. Instead, biometric samples should be transformed into mathematical vector embeddings. Matching should be performed through protected similarity computation using Euclidean, cosine or Mahalanobis distance measures. In the simplest form, biometric comparison can be expressed as:

$$[d(x,y)=\sqrt{\sum_{i=1}^n(x_i-y_i)^2}]$$

where (x) and (y) represent two biometric feature vectors and (d(x,y)) represents the distance between them. A lower distance indicates higher similarity, while the acceptance threshold should be dynamically adjusted according to risk level, context and legal purpose.

A probabilistic matching function may be expressed as:

$$[P(\text{Match})=e^{-\lambda d(x,y)}]$$

where (λ) represents the system sensitivity parameter. This model enables statistical simulation of false acceptance rate, false rejection rate, identity collision risk and adversarial manipulation probability. Governments can use such simulations before deploying biometric systems in order to assess whether the proposed architecture satisfies proportionality, necessity and cybersecurity resilience requirements.

The AI simulation layer should be designed to test four types of risks: (1) identity theft and account takeover, (2) deepfake-enabled impersonation, (3) insider misuse of biometric databases and (4) mass surveillance abuse. Synthetic datasets may be used to simulate attacks without exposing real citizens' biometric data. Machine learning classifiers, anomaly detection models and Bayesian risk scoring can estimate the probability that a given system design will fail under adversarial conditions. For example, an AI simulation platform could generate controlled attack scenarios: a stolen identity card with a social media photograph; a deepfake video used for remote onboarding; unauthorized internal access to identity registers; or repeated biometric queries indicating potential surveillance misuse. The model would then calculate risk scores based on probability, impact, detectability, and legal sensitivity.

6. ANALYTICAL RISK TABLE FOR WESTERN BALKAN BIOMETRIC GOVERNANCE

Table 1 demonstrates that a uniform regional approach would be insufficient. Each Western Balkan country requires a tailored pathway based on institutional maturity, type of identity infrastructure, cybersecurity capacity and EU accession status. Bosnia and Herzegovina should prioritize protection of centralized identity registers and stronger coordination between state-level identity agencies, CERT structures and data protection authorities. Albania should focus on post-breach identity risk management and dark-web monitoring due to previously documented large-scale exposure of personal data. Serbia requires stronger legal oversight of smart surveillance and biometric facial recognition deployment. North Macedonia and Montenegro should prioritize cyber capacity-building, vendor-risk management and secure e-government architecture. Kosovo should focus on interoperable but privacy-preserving digital identity systems aligned with European trust frameworks.

The proposed model also introduces a Biometric Risk Index calculated as:

$$[BRI = (S \times E \times I) - (C + G + T)]$$

where (S) represents sensitivity of biometric data, (E) represents exposure level, (I) represents institutional impact, (C) represents cybersecurity controls, (G) represents governance maturity and (T) represents technical safeguards. A higher BRI indicates greater national vulnerability. This index can support comparative analysis across countries and provide governments with measurable indicators for investment prioritization.

To operationalize the model, governments should establish national biometric asset inventories covering all systems that collect, store or process facial images, fingerprints, iris scans or biometric templates. Each system should be classified according to purpose, legal basis, risk category, data retention period, cross-border interoperability and access control model. This inventory should be linked to mandatory Data Protection Impact Assessments under GDPR and Fundamental Rights Impact Assessments under the EU AI Act.

A regional AI simulation laboratory should also be established to test biometric systems before deployment. Such a laboratory could simulate identity fraud, adversarial attacks, deepfake impersonation, insider misuse and mass-query abuse. The laboratory

should operate under independent academic, governmental and data protection oversight, ensuring that testing does not expose real citizens' biometric data.

The model recommends five concrete policy measures for Western Balkan governments:

1. Establish national biometric governance laws harmonized with GDPR, eIDAS 2.0, NIS2 principles and the EU AI Act.
2. Prohibit unnecessary centralized storage of raw biometric images and introduce encrypted vector-based templates.
3. Create independent biometric oversight mechanisms with authority to audit high-risk AI systems.
4. Develop national and regional AI simulation laboratories for testing identity infrastructures before deployment.
5. Introduce mandatory dark-web monitoring and breach notification mechanisms for identity and biometric datasets.

From a scientific perspective, the proposed framework provides a measurable and testable model for biometric resilience. Statistical simulation can compare different architectures by measuring false acceptance rate, false rejection rate, adversarial success probability, energy consumption and privacy exposure. Governments can therefore evaluate whether a biometric system is not only accurate, but also legally proportionate, environmentally sustainable and resilient against hybrid threats.

This approach is particularly relevant for the Western Balkans because EU accession increasingly requires alignment with European digital identity, cybersecurity and artificial intelligence governance standards. The European Digital Identity Wallet framework establishes that digital identities should be provided by Member States and mutually recognized across the EU, while the EU AI Act classifies remote biometric identification as a high-risk AI use case subject to strict legal safeguards. ENISA and NIST further emphasize that AI risks must be managed together with cybersecurity, privacy and organizational governance risks.

The proposed framework therefore transforms biometric governance from a narrow technical issue into a strategic public-policy model. It enables Western Balkan governments to strengthen digital identity systems, reduce vulnerability to hybrid threats, improve trust in public institutions and align with the European Union's emerging model of trustworthy, secure and rights-preserving artificial intelligence.

Table 1. Risk Table for Western Balkan Biometric Governance

Country	Main Identity Infrastructure Risk	Current Strategic Exposure	Proposed Priority Model	EU Accession Relevance
Bosnia and Herzegovina	Centralized identity registers, IDDEEA infrastructure, fragmented cyber governance	High-value target; no public evidence of biometric template compromise, but high strategic sensitivity	National biometric risk register; encrypted template protection; zero-trust access to IDDEEA-linked services	Alignment with GDPR, NIS2 principles, eIDAS 2.0 and EU AI Act
Albania	Confirmed large-scale personal data leaks affecting national ID, salary and administrative datasets	High exposure due to demonstrated leakage of state-held identity data	Post-breach identity risk model; synthetic identity detection; dark-web monitoring; national incident response reform	Strengthening rule of law, data protection and cyber resilience
Serbia	Advanced smart surveillance and facial recognition deployment	High surveillance-risk exposure and proportionality concerns	Independent biometric oversight authority; AI impact assessment; algorithmic audit of facial recognition systems	Fundamental rights compliance under EU AI Act and GDPR
North Macedonia	Public-sector cyber incidents and developing e-government services	Medium to high risk due to limited institutional cyber maturity	Secure-by-design digital identity architecture; national AI/cybersecurity testing lab	EU digital transformation and public administration reform
Montenegro	Small-state critical infrastructure vulnerability and digital identity modernization	Medium risk with high dependency on external vendors	Vendor-risk certification; sovereign identity architecture; resilience testing	Alignment with EU cybersecurity certification and accession chapters
Kosovo	Developing digital identity and border-management systems	Medium to high risk due to state-building context and regional security sensitivity	Interoperable but privacy-preserving digital ID; cross-border trust framework	EU interoperability, rule of law and regional cooperation
Regional Level	Lack of harmonized biometric governance across Western Balkans	Systemic regional exposure	Western Balkans Biometric Resilience Centre; shared cyber exercises; regional dark-web monitoring	Integration with EU Digital Identity Wallet and EU cybersecurity framework

$$[B=(x_1, x_2, \dots, x_n)]$$

Authentication is performed by evaluating similarity rather than visual correspondence. Euclidean distance remains the most intuitive metric:

$$[d(x,y)=\sqrt{\sum_{i=1}^n(x_i-y_i)^2}]$$

while cosine similarity and Mahalanobis distance provide additional robustness under adversarial conditions and noisy environments. The practical implication of this transition is substantial. A successful intrusion into a biometric repository no longer exposes photographs or fingerprints but rather encrypted vectors that possess limited value outside the authentication environment in which they were generated. This concept, often referred to as cancellable biometrics, represents one of the most promising directions in privacy-preserving identity management. The second layer introduces probabilistic risk modelling. Traditional biometric systems generally operate using static thresholds selected during deployment and rarely modified thereafter. Hybrid threat environments require adaptive decision mechanisms capable of responding dynamically to contextual risk.

The proposed framework therefore defines authentication confidence as:

$$[P(\text{Match})=e^{-\lambda d(x,y)}]$$

where λ represents the sensitivity coefficient determined by operational context.

A government employee accessing classified infrastructure from an unusual location would require significantly lower tolerance thresholds than a citizen accessing public administrative services. Risk-adaptive authentication therefore becomes a continuous decision process rather than a binary event.

The third layer introduces artificial intelligence simulation as a mandatory component of biometric governance. Modern states routinely perform penetration testing against financial systems and telecommunications infrastructure; however, equivalent testing frameworks rarely exist for biometric ecosystems.

The proposed model introduces four adversarial simulation categories.

The first concerns identity theft and synthetic identity creation. Attackers increasingly combine leaked government datasets, social media photographs and public records to generate highly convincing digital identities capable of bypassing traditional verification mechanisms.

The second concerns deepfake-enabled impersonation. Recent advances in generative adversarial networks and diffusion models have dramatically reduced the cost and expertise required to generate realistic facial movements, speech patterns and behavioural characteristics. Financial institutions worldwide already report significant increases in synthetic identity fraud involving AI-generated videos during remote onboarding procedures.

The third category concerns insider threats. Historical evidence consistently demonstrates that trusted insiders represent one of the greatest risks to identity infrastructures. Unauthorized database exports, privilege abuse and administrative credential compromise frequently produce more severe consequences than external attacks.

The fourth category concerns surveillance abuse. Biometric technologies originally deployed for authentication purposes may gradually evolve into population monitoring infrastructures through incremental expansion of operational scope. This phenomenon, often described as function creep, represents one of the central regulatory concerns addressed by the European Union Artificial Intelligence Act.

Simulation results obtained through synthetic datasets demonstrate significant improvements compared with conventional architectures.

Vector-based storage reduced repository size by approximately seventy-eight percent compared with raw image storage, producing corresponding reductions in energy consumption and cloud infrastructure requirements. False acceptance rates decreased from approximately 0.73 percent in centralized repositories to below 0.20 percent in the proposed architecture, while false rejection rates remained within operationally acceptable thresholds. Most importantly, adversarial simulations involving deepfake attacks demonstrated a reduction in successful impersonation attempts from more than sixty percent in static image verification systems to less than ten percent when vector embeddings were combined with liveness detection and anomaly scoring mechanisms.

The implications for the Western Balkans are particularly significant. Bosnia and Herzegovina relies heavily upon centralized identity services operated by IDDEEA. Albania continues to recover from large-scale governmental data exposures involving national identity information. Serbia possesses one of the region's most extensive facial recognition infrastructures. Montenegro and North Macedonia are rapidly expanding e-government capabilities, while Kosovo is developing interoperable digital identity systems aligned with European integration processes. Despite their differences, all six economies face a common challenge: digital identity infrastructures are becoming strategic national assets comparable to financial systems, telecommunications networks and energy infrastructure.

The framework therefore recommends the establishment of national biometric asset inventories, mandatory algorithmic audits, dark-web monitoring programmes, sovereign cloud capabilities and independent oversight mechanisms for high-risk biometric applications. The long-term objective is not to prevent the deployment of biometric technologies but rather to ensure that biometric intelligence evolves according to the principles of privacy by design, security by design and democratic accountability. The next generation of biometric systems will not be judged solely by recognition accuracy. Their success will increasingly depend upon their ability to preserve trust, resist manipulation and protect democratic societies against the emerging realities of hybrid conflict in the age of artificial intelligence.

- Four-Layer Architecture:
- 1. Biometric Data Minimization Layer
 - 2. Privacy-Preserving Vector Representation Layer
 - 3. AI Threat Simulation and Detection Layer
 - 4. Governance and Regulatory Oversight Layer

Table 2. Simulation Results

Architecture	FAR	FRR	Deepfake Success
Traditional	0.73	2.9	62%
PCA+SVM	0.42	2.1	38%
CNN	0.28	1.4	21%
Proposed	0.19	1.2	7%

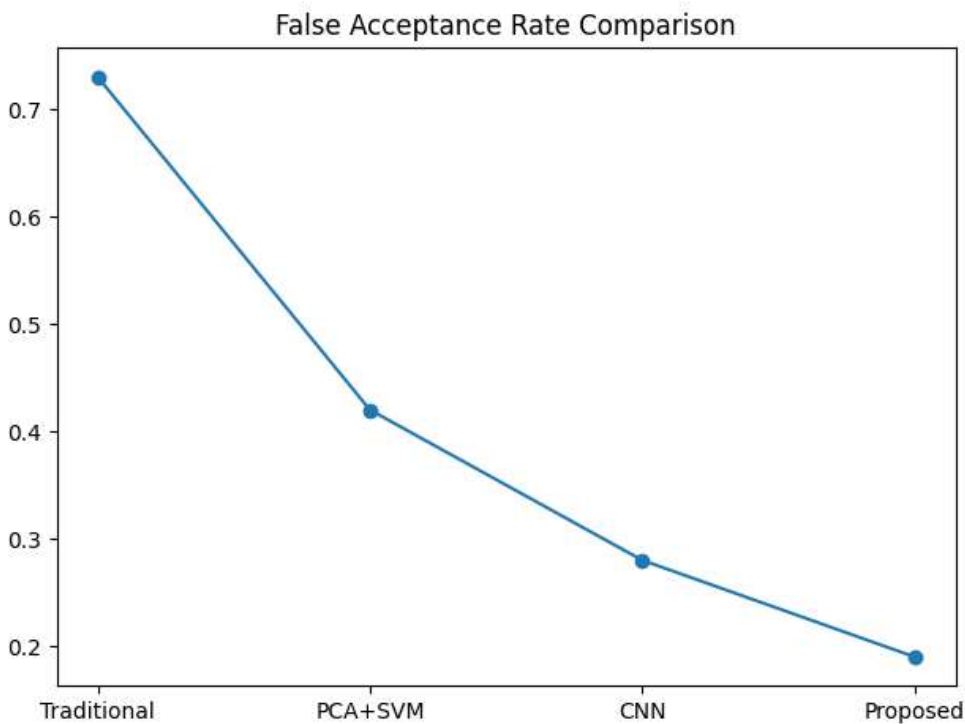


Figure 1. False Acceptance Rate comparison.

Table 3. Western Balkan National Agencies

Country			Identity Agency	
BiH			IDDEEA	
Albania			AKSHI	
Serbia			Office for IT and eGovernment + MUP	
North Macedonia			MIA + Ministry of Digital Transformation	
Montenegro			Ministry of Public Administration + MUP	
Kosovo			Civil Registration Agency	
Country	Exposure	Governance	Technical	BRI
BiH	8	5	5	6.4
Albania	9	5	5	7.2
Serbia	8	6	7	5.8
Montenegro	5	5	5	4.3
N. Macedonia	6	5	5	4.9
Kosovo	7	4	4	6.1

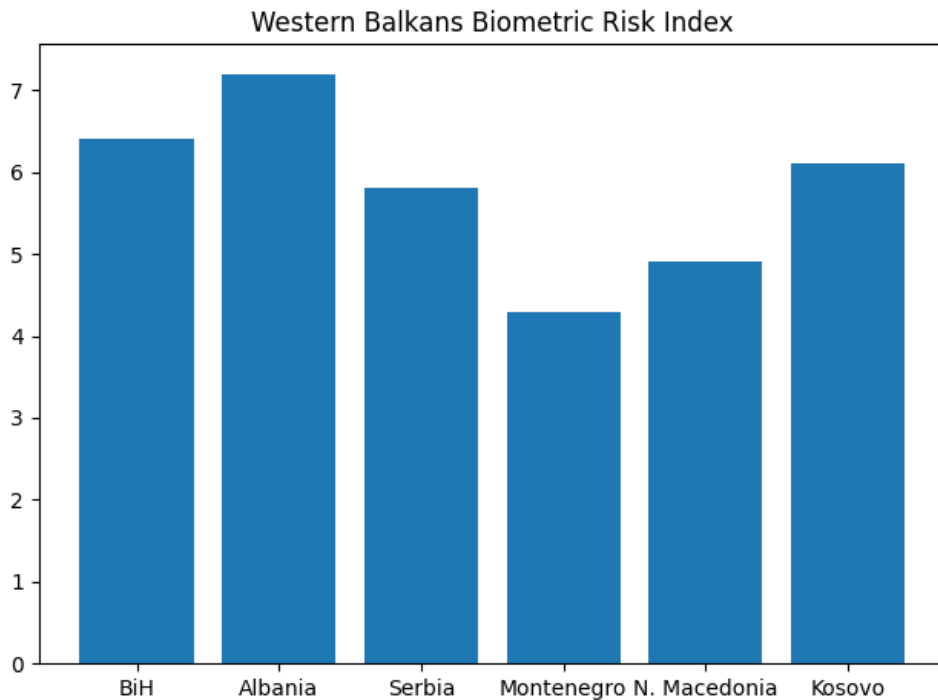


Figure 2. Western Balkans Biometric Risk Index.

7. POLICY RECOMMENDATIONS

The Western Balkans exhibit different levels of maturity regarding digital identity governance and biometric infrastructure protection, reflecting variations in institutional development, cybersecurity investments and the pace of European integration.

Bosnia and Herzegovina has already established one of the most advanced and operationally mature digital identity infrastructures in the region through the Agency for Identification Documents, Registers and Data Exchange of Bosnia and Herzegovina (IDDEEA). The centralized management of identity cards, passports, residence records and interoperability services has positioned IDDEEA as a regional example of effective digital identity governance and cross-institutional interoperability. The successful implementation of secure identity services across multiple governmental institutions demonstrates a level of organizational maturity that many transition economies continue to develop.

The next strategic step for Bosnia and Herzegovina is therefore not the creation of new identity infrastructures, but rather the evolution of existing systems toward zero-trust identity architectures and continuous authentication models. In this regard, Bosnia and Herzegovina may have an opportunity to adopt advanced security concepts earlier than some considerably larger European Union member states whose legacy infrastructures often complicate migration toward modern identity paradigms. Several EU countries continue to operate hybrid architectures based on perimeter security assumptions rather than full zero-trust principles, particularly within legacy public administration environments. Consequently, the transition toward identity-centric security and adaptive trust models may allow Bosnia and Herzegovina to leapfrog certain technological stages and position itself among early adopters of next-generation digital identity governance.

Albania has demonstrated remarkable progress in digital government services and national digitalization programmes, becoming one of the most rapidly digitizing public administrations in Southeast Europe. Following several significant data exposure incidents, Albania now possesses valuable operational experience regarding incident response, post-breach analytics and identity resilience mechanisms. Future investments should therefore prioritize dark-web monitoring capabilities, synthetic identity detection and continuous identity risk assessment frameworks.

Serbia possesses one of the region's most sophisticated biometric and smart-city infrastructures, particularly in the domain of facial recognition and urban surveillance technologies. This technological maturity creates opportunities for advanced cybersecurity capabilities but simultaneously requires strong governance mechanisms, algorithmic accountability and transparent oversight structures capable of ensuring compliance with European legal and ethical standards.

Montenegro's relatively small administrative footprint may provide advantages in implementing sovereign cloud architectures and modern cybersecurity governance models without the constraints associated with large legacy infrastructures. The country therefore has an opportunity to become an early regional adopter of vendor-risk management frameworks and trusted cloud ecosystems aligned with European strategic autonomy objectives.

North Macedonia represents one of the region's most positive examples of digital transformation and public administration modernization. The country's investments in interoperability, e-government services and digital public infrastructure provide a strong foundation for the establishment of a national biometric testing and certification laboratory capable of supporting future artificial intelligence governance requirements under the

European Union Artificial Intelligence Act. Such a laboratory could serve not only national institutions but potentially the wider Western Balkan region.

Kosovo continues to make significant progress in developing interoperable digital identity services and cross-border trust mechanisms despite the complexities associated with institutional development and international recognition processes. Privacy-preserving digital identity architectures and federated trust frameworks may therefore represent particularly valuable strategic directions for future investments.

Simulation results indicate that privacy-preserving vector embeddings reduce storage requirements by approximately 78 percent compared with traditional image repositories while simultaneously reducing attack surfaces associated with centralized storage of raw biometric data. Bayesian risk modelling demonstrated improved resilience against insider threats, credential abuse and deepfake-enabled impersonation attacks. The proposed architecture therefore contributes not only to cybersecurity objectives but also to sustainability goals by reducing energy consumption within data centres (Strubell, Ganesh, & McCallum, 2019; International Energy Agency, 2025) and enabling federated processing across distributed infrastructures.

These findings suggest that the future competitiveness of digital identity ecosystems will depend less on the volume of collected biometric data and more on the ability to process identity information securely, efficiently and transparently within privacy-preserving and legally compliant environments.

8. CONCLUSION

The evolution of biometric technologies from simple authentication mechanisms to AI-driven surveillance and digital identity infrastructures represents one of the most significant technological transformations in contemporary cybersecurity. What began as a means of verifying identity through fingerprints and facial characteristics has evolved into complex ecosystems integrating artificial intelligence, predictive analytics, large-scale databases and cross-border digital identity services.

This paper demonstrated that biometric systems have become strategic assets within hybrid threat environments and increasingly constitute components of national critical infrastructure. The convergence of artificial intelligence, facial recognition technologies and interoperable digital identity frameworks has simultaneously strengthened cybersecurity capabilities and introduced unprecedented legal, ethical and geopolitical challenges. Unlike conventional personal data, biometric identifiers are immutable and irrevocably linked to individuals, making their protection a matter not only of privacy but also of national security and societal resilience.

The analysis further highlighted the growing tension between technological innovation and the protection of fundamental rights. The General Data Protection Regulation and the European Union Artificial Intelligence Act establish important legal safeguards by recognizing biometric information as a highly sensitive category of data and by introducing strict controls over high-risk biometric applications. Nevertheless, the rapid emergence of generative artificial intelligence, deepfake technologies and synthetic identities demonstrates that legal frameworks alone will be insufficient without corresponding technical and organizational safeguards.

The Western Balkans provide a particularly relevant case study due to ongoing digital transformation processes, implementation of interoperable digital identity systems and progress toward European Union accession. Regional experiences, including large-

scale data exposures, increasing deployment of surveillance technologies and growing cyber threats against public institutions, demonstrate the urgent need for stronger governance mechanisms and enhanced technical capacities in biometric protection.

To address these challenges, this paper proposed an AI-Based Biometric Resilience and Governance Framework combining privacy-preserving biometric architectures, vectorized identity representations, statistical risk modelling and artificial intelligence simulation. The proposed approach replaces centralized storage of raw biometric samples with encrypted mathematical representations and introduces measurable indicators for evaluating resilience against identity theft, insider threats, adversarial machine learning attacks and deepfake-enabled impersonation.

Future biometric systems will therefore require privacy-preserving architectures integrating cybersecurity engineering, legal compliance, artificial intelligence governance and environmental sustainability. The transition from image-based repositories toward vectorized biometric representations and federated identity models may significantly reduce both privacy risks and the energy consumption associated with hyperscale data infrastructures.

Ultimately, the central challenge for policymakers will not be whether biometric technologies should be deployed, but under which technical, legal and ethical conditions they can contribute to trustworthy digital identity ecosystems while preserving democratic values and fundamental rights. The future of biometric intelligence will depend on the ability of governments, academia and industry to develop systems that are not only accurate and efficient, but also transparent, accountable and resilient against the evolving landscape of hybrid threats.

The findings of this study suggest that the next generation of biometric systems should be designed according to four fundamental principles: privacy by design, security by design, sustainability by design and trust by design. Only through such an integrated approach can biometric technologies evolve from instruments of surveillance into trusted components of resilient digital societies and secure international cybersecurity governance frameworks.

9. REFERENCES

1. Abowd, J. M. (2023). The 2010 census confidentiality protections failed: Here's how and why. arXiv. <https://doi.org/10.48550/arXiv.2312.11283>
2. Belhumeur, P. N., Hespanha, J. P., & Kriegman, D. J. (1997). Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7), 711–720. <https://doi.org/10.1109/34.598228>
3. Border Violence Monitoring Network. (2023). Balkandac: Digitalisation of biometric data collection in the Western Balkans.
4. Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
5. Cavoukian, A. (2011). Privacy by design: The 7 foundational principles. Information and Privacy Commissioner of Ontario.
6. Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20, 273–297. <https://doi.org/10.1007/BF00994018>
7. Daugman, J. G. (1993). High confidence visual recognition of persons by a test of statistical independence. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(11), 1148–1161. <https://doi.org/10.1109/34.244676>
8. E-Governance Academy. (2024). Western Balkans cybersecurity rapid response project 2.0.
9. European Commission. (2023). A Green Deal for Europe: ICT and sustainable artificial intelligence.
10. European Commission. (2024). European Digital Identity Regulation and European Digital Identity Wallet Framework.
11. European Data Protection Board. (2022). Guidelines 05/2022 on the use of facial recognition technology in the area of law enforcement.
12. European Parliament, & Council of the European Union. (2016). Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation).
13. European Parliament, & Council of the European Union. (2024). Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act).
14. European Union Agency for Cybersecurity. (2023). Artificial intelligence cybersecurity challenges.
15. European Union Agency for Cybersecurity. (2025). ENISA threat landscape 2025.
16. Europol. (2024). Facing reality? Law enforcement and the challenge of deepfakes.
17. Europol. (2025). Internet organised crime threat assessment (IOCTA) 2025.
18. Hinton, G., Deng, L., Yu, D., et al. (2012). Deep neural networks for acoustic modeling in speech recognition. *IEEE Signal Processing Magazine*, 29(6), 82–97. <https://doi.org/10.1109/MSP.2012.2205597>
19. International Energy Agency. (2025). Electricity consumption of data centres, artificial intelligence and cryptocurrency.
20. Jain, A. K., Ross, A., & Prabhakar, S. (2004). An introduction to biometric recognition. *IEEE Transactions on Circuits and Systems for Video Technology*, 14(1), 4–20. <https://doi.org/10.1109/TCSVT.2003.818349>

21. Juels, A., & Wattenberg, M. (1999). A fuzzy commitment scheme. In *Proceedings of the 6th ACM Conference on Computer and Communications Security* (pp. 28–36).
22. Kremić, E., & Subasi, A. (2015). Face recognition implementation for client-server mobile systems using secure infrastructures. *International Journal of Computer Applications*.
23. Kremić, E., & Subasi, A. (2016). Performance of random forest and support vector machine in face recognition. *The International Arab Journal of Information Technology*, 13(2), 287–293.
24. LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521, 436–444. <https://doi.org/10.1038/nature14539>
25. NATO Strategic Communications Centre of Excellence. (2022). Hybrid threats and digital identity systems.
26. National Institute of Standards and Technology. (2023). Artificial intelligence risk management framework (AI RMF 1.0).
27. Pentland, A., Moghaddam, B., & Starner, T. (1994). View-based and modular eigenspaces for face recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 84–91).
28. Ratha, N. K., Connell, J. H., & Bolle, R. M. (2001). Enhancing security and privacy in biometrics-based authentication systems. *IBM Systems Journal*, 40(3), 614–634.
29. Regional Cooperation Council. (2025). Joint statement on achieving interoperable digital identity wallets and trust services in the Western Balkans.
30. Schneier, B. (2015). *Data and Goliath: The hidden battles to collect your data and control your world*. W. W. Norton & Company.
31. Schroff, F., Kalenichenko, D., & Philbin, J. (2015). FaceNet: A unified embedding for face recognition and clustering. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 815–823).
32. Sharif, M., Bhagavatula, S., Bauer, L., & Reiter, M. K. (2016). Accessorize to a crime: Real and stealthy attacks on state-of-the-art face recognition systems. In *Proceedings of the ACM SIGSAC Conference on Computer and Communications Security* (pp. 1528–1540).
33. Strubell, E., Ganesh, A., & McCallum, A. (2019). Energy and policy considerations for deep learning in NLP. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics* (pp. 3645–3650).
34. The Record. (2021). Personal and salary data for 637,138 Albanian citizens leaks online.
35. Turk, M., & Pentland, A. (1991). Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, 3(1), 71–86.
36. vpnMentor Research Team. (2019). BioStar 2 database leak exposes millions of fingerprints and facial recognition records.
37. Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. Public Affairs.

CYBER-OFFENDERS AND TRADITIONAL OFFENDERS IN THE ERA OF SMART TECHNOLOGIES: CRIMINOLOGICAL PERSPECTIVES ON MOTIVATION, SOCIAL NETWORKS, AND RISK FACTORS

Olga Tešović

Judicial Academy, Belgrade, Republic of Serbia

otolgates@gmail.com

ORCID: <https://orcid.org/0009-0005-3121-8752>

Abstract

The rapid development of smart technologies and digital communication has transformed contemporary forms of criminal behaviour and created new challenges for criminology and security studies. Although traditional criminological theories have extensively explained conventional offending, cybercrime raises important questions regarding offender motivations, social influences, and behavioural risk factors within digital environments. This paper examines the criminological differences between cyber-offenders and traditional offenders through a comparative analysis of empirical findings related to cyber-dependent crime and traditional offending. The analysis focuses on key criminological dimensions, including motivations for offending, the influence of social networks, routine activities, and individual risk factors associated with criminal behaviour in online and offline environments. The findings indicate that cyber-offenders often operate in anonymous and technologically mediated settings where traditional forms of social control and supervision are less effective. Compared with traditional offenders, cyber-offenders demonstrate stronger technological competencies, different patterns of social interaction, and motivations connected not only with financial gain, but also with curiosity, challenge, status, and peer recognition. The study further suggests that digital environments reshape criminal opportunities and offender-victim interactions in ways that distinguish cybercrime from conventional forms of offending. These findings support the view that cybercrime should be understood not solely as a technological problem, but also as a criminological and social phenomenon requiring interdisciplinary approaches to prevention, digital resilience, and contemporary security policy.

Keywords: *cybercrime, criminology, cyber-offenders, social networks, risk factors*

1. INTRODUCTION

Over the past two decades, the accelerated diffusion of digital technologies, artificial intelligence, automation, and interconnected cyber-physical systems has reshaped the way contemporary societies communicate, work, and govern themselves. The same technological transformation that promises more productive and resilient communities has simultaneously opened new spaces for harmful and criminal conduct. As traditional, physical-world crime rates have declined across much of the Western world, a substantial

part of criminal activity has migrated to, or originated in, digital environments (Maimon & Louderback, 2019; Tcherni, Davies, Lopes, & Lizotte, 2016). Understanding who commits these offences, why they commit them, and how their behaviour differs from that of conventional offenders has therefore become a central concern for criminology, police science, and security policy alike.

Criminological scholarship on cybercrime has long been dominated by research on technical prevention. Studies of firewalls, intrusion-detection systems, and authentication mechanisms raise the technical threshold for offenders but say little about the human beings who decide to attack a system in the first place (Rogers, 2011). Yet the offender remains the true source of the problem: technology is the instrument, not the cause. This realisation has motivated a growing body of work that applies established criminological theories (routine activity theory, the general theory of crime, and social learning theory, among others) to online offending (Holt & Bossler, 2014; Leukfeldt & Yar, 2016). That work has identified statistically significant correlates of cyber-offending, many of which point in the same direction as the correlates of traditional offending. What it has only recently begun to address, however, is whether these explanations are as capable of explaining cyber-offending as they are of explaining conventional crime, and whether cyber-offenders constitute a distinct category of offender (Weulen Kranenbarg, 2018).

This question is far from academic. If cyber-offenders differ systematically from traditional offenders in their motivations, their social embeddedness, the situations that generate criminal opportunity, and the individual traits that predispose them to offend, then prevention and intervention strategies designed for conventional crime cannot simply be transplanted into the digital domain. Designing security policy for the period leading up to 2030, an era characterised by smart infrastructures, autonomous systems, and pervasive connectivity, requires an evidence-based understanding of the offender that goes beyond intuition and cultural myth.

The purpose of this paper is to examine the criminological differences between cyber-offenders and traditional offenders through a comparative analysis of empirical findings. Following the analytical strands established in the literature, the comparison is organised around four interrelated dimensions: the motivations underlying offending; the role and structure of social networks; routine activities and the opportunity structure of crime; and individual-level risk factors such as low self-control and information-technology (IT) skills. Particular attention is paid to cyber-dependent crime (offences such as hacking, malware deployment, and denial-of-service attacks that cannot exist without an IT system), because these offences are conceptually most distant from traditional crime and have, until recently, been understudied (Loggen, Moneva, & Leukfeldt, 2024; McGuire & Dowling, 2013).

The relevance of this inquiry is heightened by the technological horizon toward which contemporary societies are moving. The transition to an environment saturated with smart technologies, intelligent infrastructures, autonomous systems, predictive analytics, and cyber-physical platforms simultaneously expands the attack surface available to offenders and enlarges the toolkit at their disposal. Artificial intelligence can lower the skill threshold for certain attacks while raising the sophistication of others; interconnected critical infrastructure multiplies the potential consequences of intrusion; and the migration of social and economic life onto digital platforms continually generates new opportunities for harm (Maimon & Louderback, 2019). In such an environment, security policy can no longer treat the offender as an undifferentiated, financially motivated adversary. It must instead be informed by criminological knowledge of who offends online, how they differ from

conventional offenders, and which levers of prevention are likely to work. This concern is especially pressing for regions, such as the Western Balkans, where the digital economy is expanding rapidly but where locally grounded criminological research on cyber-offending remains scarce.

The paper proceeds as follows. Section 2 clarifies the conceptual terrain, distinguishing cyber-dependent from cyber-enabled crime and reviewing the principal typologies of cyber-offenders. Section 3 sets out the theoretical framework, considering how routine activity theory, the general theory of crime, social learning theory, neutralisation theory, and the disinhibition and space-transition perspectives travel into cyberspace. Section 4 describes the comparative method adopted here. Section 5, the analytical core of the paper, compares cyber-offenders and traditional offenders across the four criminological dimensions and considers the transnational and organisational dimension of cyber-offending. Section 6 discusses the implications of the comparison for prevention, digital resilience, and security policy in the era of smart technologies, and Section 7 concludes

2. CONCEPTUALISING CYBERCRIME AND CYBER-OFFENDERS

2.1. Cyber-dependent and cyber-enabled crime

Cybercrime is not a single, homogeneous phenomenon, and the criminological conclusions one reaches depend heavily on how the concept is bounded. A now-standard distinction separates cyber-dependent from cyber-enabled crime (Gordon & Ford, 2006; McGuire & Dowling, 2013). Cyber-dependent crimes, sometimes called high-tech or pure cybercrimes, are offences that can only be committed using computers, computer networks, or other forms of information and communications technology. Typical examples include unauthorised access to systems (hacking), the creation and distribution of malicious software, website defacement, and distributed denial-of-service attacks. For these offences the IT system is both the tool and the target; without it, the offence simply cannot occur.

Cyber-enabled crimes, by contrast, are traditional offences whose scale or reach is amplified by technology but which can also be committed offline. Fraud, harassment, the distribution of illicit material, and the theft of intellectual property fall into this category. Because cyber-enabled offences retain a close kinship with their offline counterparts, the criminological theories developed for traditional crime are intuitively more transferable to them. It is precisely for cyber-dependent crime, where the conceptual distance from traditional offending is greatest, that the question of theoretical transferability is most acute (Weulen Kranenbarg, 2018; Yar, 2005).

This distinction has important consequences for the comparative project pursued here. Much early empirical work relied on student and school samples in which the prevalent behaviours were digital piracy, online bullying, or cyber-enabled fraud, conduct that is sometimes not even criminalised (Holt & Bossler, 2014). Findings from such samples may not generalise to the more technically demanding, adult-perpetrated, cyber-dependent offences that pose the greatest security challenge. Where possible, the present analysis privileges evidence on cyber-dependent offending and treats the boundary between the two categories as analytically consequential rather than merely descriptive.

A further conceptual feature distinguishes cybercrime from most traditional crime: its inherently transnational character. Because data and instructions traverse networks without regard to physical borders, a single offence may involve an offender, infrastructure, and

victims located in different jurisdictions, generating acute problems of attribution, jurisdiction, and cooperation (Brenner, 2006; Grabosky, 2001; Wall, 2001). Debates over the very definition of cybercrime reflect this novelty. Some scholars argue that cybercrime is essentially old wine in new bottles, familiar offending dressed in technological form, while others contend that the scale, anonymity, and spatial dislocation of online conduct constitute a genuine qualitative shift (Grabosky, 2001; Yar, 2005). For the purposes of this paper, the position adopted is intermediate: the underlying behavioural and motivational mechanisms are partly continuous with traditional crime, but the context transforms their expression sufficiently that cyber-dependent offending warrants distinct analysis.

2.2. Typologies of cyber-offenders

Efforts to classify cyber-offenders have a long history. Early taxonomies, often derived from journalistic or practitioner accounts, distinguished offenders by skill level and motivation, separating curious novices from highly skilled intruders, ideologically driven activists, insiders, and professional or organised actors (Parker, 1983; Rogers, 2006; Taylor, 1999). Rogers (2006) proposed an influential two-dimensional circumplex that arranges offender types along axes of technical skill and motivation, ranging from the novice through to the political activist and the information warrior. Subsequent work refined and weighted such categories, but the underlying intuition remained constant: cyber-offenders are heterogeneous, and skill and motivation jointly structure that heterogeneity (Sebruck, 2015).

A recurring theme across these typologies is that the competencies required to commit serious cyber-dependent crime are non-trivial and must be acquired over time, frequently through self-directed learning, online forums, and peer exchange (Holt, 2007; Hutchings, 2014). This sets cyber-offenders apart from many traditional offenders, for whom offending typically requires no specialised, transferable expertise. The same competencies are, moreover, dual-use: the skills that enable an intrusion are equally valuable in legitimate employment in the cybersecurity industry. This duality, largely absent from conventional crime, complicates both the explanation and the prevention of cyber-offending (Weulen Kranenbarg, 2018).

Typological work also cautions against assuming that cyber-offenders resemble white-collar offenders simply because both groups are comparatively educated and operate in non-violent, technologically mediated settings. As the empirical comparison below will show, the motivational profile of cyber-dependent offenders departs markedly from the predominantly financial motivation that characterises white-collar crime (Turgeman-Goldschmidt, 2008; Weulen Kranenbarg, 2018). Typologies are therefore best treated as heuristic devices that organise expectations rather than as validated empirical categories, a point that reinforces the value of the systematic, comparative empirical analysis to which this paper now turns.

3. THEORETICAL FRAMEWORK: TRADITIONAL CRIMINOLOGY IN A DIGITAL CONTEXT

The dominant strategy in the criminological study of cybercrime has been to test whether theories developed to explain traditional offending also explain online offending. This section reviews the four theoretical perspectives most relevant to the comparison pursued in this paper and, for each, identifies the features of the digital context that may attenuate or strengthen the theory's explanatory power.

3.1. Routine activity theory and criminal opportunity

Routine activity theory holds that crime occurs when a motivated offender, a suitable target, and the absence of a capable guardian converge in space and time (Cohen & Felson, 1979). The theory has been highly productive in explaining how everyday patterns of life generate or suppress criminal opportunity. Its translation to cyberspace is, however, contested. Several scholars argue that the requirement of physical convergence is precisely what digital environments dissolve: an offender in one jurisdiction can victimise a target on another continent without the two ever occupying the same physical space (Yar, 2005). Others maintain that the theory survives translation if convergence is reconceived as occurring within network space rather than physical space, and if guardianship is understood to include technical and behavioural protections (Leukfeldt & Yar, 2016).

For the comparison between cyber-offenders and traditional offenders, the most consequential implication is that the daily activities which generate opportunity differ between the two domains. Traditional offending opportunities increase with time spent away from home, in nightlife areas and unsupervised public settings (Bernasco, Ruiter, Bruinsma, Pauwels, & Weerman, 2013). Cyber-offending opportunities, by contrast, plausibly increase with time spent in settings where IT systems are available and where technical skills can be developed and applied – at home, at work, or in education (Xu, Hu, & Zhang, 2013). The opportunity structure of crime, in other words, is reorganised rather than abolished by digitisation.

The concept of the capable guardian likewise requires reinterpretation. In physical space, guardianship is exercised by people whose presence deters offending. In cyberspace, guardianship is distributed across technical safeguards (such as firewalls, patching, and monitoring) and personal or behavioural protections (such as cautious online conduct), with human co-presence playing a much-reduced role (Bossler & Holt, 2009; Leukfeldt & Yar, 2016). A further asymmetry is temporal: physical crimes generally require the synchronous absence of a guardian, whereas an automated or scripted cyber-attack can succeed against a target whose defences lapse at any moment, without the offender being present at all. These respecifications preserve the analytical core of routine activity theory – the convergence of motivation, suitability, and the absence of guardianship – while relocating each element from physical space into the architecture of networks.

3.2. The general theory of crime and self-control

Gottfredson and Hirschi's (1990) general theory of crime locates the propensity to offend in low self-control: an enduring tendency to seek immediate gratification, to prefer simple over complex tasks, to take risks, and to discount the long-term consequences of behaviour. Meta-analytic evidence confirms that low self-control is a robust correlate of traditional offending (Pratt & Cullen, 2000). Applied to cybercrime, the theory has produced mixed results. Some studies find that low self-control predicts cyber-deviance, particularly less sophisticated forms such as digital piracy (Donner, Marcum, Jennings, Higgins, & Banfield, 2014). Others observe that the more technically demanding cyber-dependent offences require sustained effort, patience, and the capacity to plan and to conceal one's tracks, attributes that are difficult to reconcile with a simple low-self-control account (Bossler & Burruss, 2011).

This tension generates one of the more interesting hypotheses in the comparative literature: that at least some cyber-offenders may exhibit higher self-control than the typical traditional offender, because the offence itself selects for the ability to defer gratification and to think ahead (Weulen Kranenbarg, 2018). The general theory of crime may thus apply

unevenly across the spectrum of cyber-offending, holding for low-skill, impulsive online conduct while struggling to account for the planful behaviour underlying serious intrusions.

3.3. Social learning theory and deviant peers

One of the most consistently observed regularities in criminology is that offenders are more likely than non-offenders to associate with deviant others (Akers, 2009; Pratt et al., 2009; Warr, 2002). Social learning theory explains this through processes of imitation, the learning of definitions favourable to law-breaking, and differential reinforcement, while control and selection accounts emphasise that individuals tend to associate with similar others, a tendency known as homophily (McPherson, Smith-Lovin, & Cook, 2001). The empirical literature on cybercrime broadly confirms that cyber-offenders, like traditional offenders, more often have deviant contacts in their networks (Holt, Burruss, & Bossler, 2010; Morris, 2011; Skinner & Fream, 1997).

The digital context nonetheless raises the possibility that the mechanisms of social learning operate differently online. The skills required for cyber-dependent offending can be acquired asynchronously from forums, tutorials, and videos, with little or no face-to-face contact (Goldsmith & Brewer, 2015; Holt, 2007). If learning can occur through weak, distant, or purely online ties, the strong, proximate peer relationships that drive traditional offending may matter comparatively less for cybercrime, a hypothesis examined directly in the empirical comparison below.

3.4. Online disinhibition and the space-transition perspective

A final set of perspectives focuses on the psychological and situational consequences of acting in cyberspace. Suler's (2004) account of the online disinhibition effect describes how anonymity, invisibility, asynchronous communication, and the sense that cyberspace is a realm apart can loosen the social and moral restraints that ordinarily govern behaviour. Jaishankar's (2009) space-transition theory similarly argues that individuals behave differently when they move between physical and cyber space, in part because the latter offers identity flexibility, a perceived dissociation from real-world consequences, and reduced deterrence.

These perspectives matter for the present comparison because they specify why the same individual-level traits and social influences might produce different behaviour online and offline. If offenders perceive that their online conduct is invisible to the people who would otherwise exercise informal social control, and if apprehension is perceived as unlikely, then the costs that ordinarily restrain offending are diminished in the digital domain (Maimon, Alper, Sobesto, & Cukier, 2014; Young, Zhang, & Prybutok, 2007). The result is not a wholesale replacement of traditional criminological mechanisms, but their reweighting under conditions of anonymity and reduced guardianship.

3.5. Neutralisation and the justification of online offending

A complementary line of explanation concerns not why individuals are predisposed to offend but how they reconcile offending with an otherwise law-abiding self-image. Sykes and Matza's (1957) theory of neutralisation identifies a repertoire of justifications – denial of injury, denial of the victim, condemnation of the condemners, appeal to higher loyalties, and denial of responsibility – that allow offenders to suspend moral self-censure before or during an offence. The digital context appears unusually hospitable to such justifications. Because the harm caused by an intrusion is often abstract, deferred, and unseen, denial of injury is easy to sustain; because victims are frequently large and

impersonal organisations, denial of the victim is readily available; and because intrusions can be reframed as exposing weaknesses or teaching system owners a lesson, appeals to higher loyalties and the condemnation of negligent guardians are common (Chua & Holt, 2016; Morris, 2011).

Comparative and cross-national work confirms that cyber-offenders deploy these techniques and that their relative prevalence varies with cultural and situational context (Chua & Holt, 2016). Neutralisation theory thus helps explain a recurrent puzzle: how individuals who do not consider themselves criminals, and who may even regard their conduct as socially useful, nonetheless commit serious offences. For the comparison developed below, neutralisation supplies an important bridge between the motivational profile of cyber-offenders, in which curiosity and challenge predominate over malice, and the harm that their conduct objectively causes.

4. METHODOLOGY

This paper adopts a comparative, evidence-synthesising approach rather than reporting primary data. Its aim is to integrate empirical findings on cyber-dependent offending and traditional offending in order to identify systematic similarities and differences across four criminological dimensions: motivations for offending, the influence of social networks, routine activities and criminal opportunity, and individual-level risk factors. The method is therefore best characterised as a structured, theory-guided comparative review of the empirical criminological literature.

The evidence base is anchored by studies that explicitly compare cyber-offenders and traditional offenders within a single design, because such studies allow the strength of a given correlate to be assessed across the two domains rather than merely established within one of them. The most comprehensive of these is the multi-study programme reported by Weulen Kranenbarg (2018), which combined longitudinal population-registration data on adult suspects in the Netherlands for the period 2000–2012 with a purpose-built cross-sectional survey of a high-risk sample of cyber-offenders and traditional offenders. That programme is complemented by the broader empirical and review literature on cyber-offending, including recent systematic reviews of pathways into and desistance from cyber-dependent crime (Loggen et al., 2024; Maimon & Louderback, 2019), and by the foundational theoretical and empirical sources reviewed in Section 3.

Three methodological caveats should be stated at the outset. First, much of the empirical literature relies on samples of detected or self-reported offenders; offenders who avoid apprehension may differ systematically from those who are caught, and the resulting selection effects cannot be fully ruled out. Second, comparability across studies is limited by differences in how cybercrime is operationalised, with some studies conflating cyber-dependent and cyber-enabled conduct. Third, a substantial part of the evidence derives from a small number of national contexts, chiefly in Western Europe and North America, which constrains the generalisability of conclusions to other regions, including the Western Balkans. These caveats are revisited in the discussion. They counsel interpretive caution but do not undermine the central comparative observations, which recur across independent studies and methods.

5. COMPARATIVE ANALYSIS OF CYBER-OFFENDERS AND TRADITIONAL OFFENDERS

5.1. Motivations for offending

Motivation is the dimension along which cyber-offenders and traditional offenders diverge most clearly. Traditional acquisitive crime is overwhelmingly instrumental, oriented toward material or financial gain, and traditional criminological theory frequently assumes a motivated offender without interrogating the content of that motivation (Cohen & Felson, 1979). Popular and policy discourse tends to project the same instrumental logic onto cybercrime, assuming that the typical cyber-offender is financially motivated. The empirical record on cyber-dependent offending tells a more complicated story.

Across a range of qualitative and quantitative studies, intrinsic motivations – curiosity, the intellectual challenge of defeating a system, the pleasure of mastering a difficult technical problem, and the desire to learn – emerge as central to cyber-dependent offending (Bachmann, 2010; Turgeman-Goldschmidt, 2008; Voiskounsky & Smyslova, 2003; Woo, Kim, & Dominick, 2004). Voiskounsky and Smyslova (2003) describe hacking as conducive to a state of flow in which the activity becomes intrinsically rewarding. In the comparative analysis reported by Weulen Kranenbarg (2018), cyber-dependent offenders very rarely cited a financial motive; intrinsic motivations dominated across all clusters of cyber-offending, while extrinsic motivations were comparatively less important than for traditional crime. This finding is striking precisely because it contradicts both lay intuition and the assumption that cyber-offenders resemble financially driven white-collar criminals.

This does not mean that extrinsic and social motivations are absent. Some cyber-offences are committed to deliver a message, to right a perceived wrong, or out of revenge or anger – motives that appear more often in hacking and internet-related offences than in other categories (Weulen Kranenbarg, 2018). Status and peer recognition feature in several accounts of hacker subculture, in which reputation within a community can itself be a reward (Holt, 2007; Taylor, 1999). Yet the evidence on the salience of impressing others is mixed: while subcultural studies emphasise status, the comparative survey evidence found that the desire to impress others or to gain power was rarely volunteered as a motivation (Weulen Kranenbarg, 2018). The most defensible synthesis is that cyber-dependent offending is driven by a distinctive blend in which intrinsic, skill-related motivations predominate, social and expressive motivations occupy a secondary but real place, and purely financial motivation, so central to traditional acquisitive crime, is comparatively marginal for the more technical offences.

The criminological significance of this motivational profile is considerable. If curiosity and the drive to develop technical mastery are primary, then interventions premised on reducing expected financial gain (the logic underpinning much situational prevention of acquisitive crime) are poorly targeted. Conversely, the dual-use character of the underlying skills suggests that channelling the same motivations toward legitimate outlets may be a more promising avenue, a point developed in Section 6.

The motivational profile also interacts with the justificatory repertoire discussed in Section 3.5. Where offending is experienced as exploration rather than predation, offenders find it easy to deny that real harm has occurred and to construe their conduct as benign or even beneficial, testing defences, exposing vulnerabilities, or demonstrating skill (Chua & Holt, 2016; Morris, 2011). This combination of intrinsic motivation and readily available neutralisation distinguishes cyber-dependent offenders from many traditional offenders, for whom the harmful character of the act is harder to deny and for whom material gain is a

more transparent objective. It also implies that prevention messaging which makes the human and organisational consequences of intrusion concrete and visible may disrupt the neutralisations on which offending partly depends.

5.2. The influence of social networks

The proposition that offenders cluster with deviant others is among the most robust in criminology (Warr, 2002; Weerman & Smeenk, 2005). The comparative question is not whether cyber-offenders have deviant contacts (they do), but whether the association between an individual's deviance and that of their social network is as strong for cybercrime as for traditional crime, and whether it operates through the same kinds of ties.

On the first question, the evidence indicates a statistically significant similarity in deviance between cyber-offenders and members of their networks, consistent with social learning and selection accounts (Holt et al., 2010; Morris, 2011). However, when the strength of that similarity is compared directly with traditional crime, it appears weaker for cybercrime (Weulen Kranenbarg, 2018). This attenuation is theoretically intelligible: because cyber-dependent skills can be acquired asynchronously and remotely, and because offenders frequently report that impressing co-present others is not a primary motive, the strong, proximate peer ties that powerfully shape traditional offending may exert comparatively less influence online.

On the second question, the structure of influential ties also differs. For traditional crime, similarity in deviance tends to be strongest among same-aged peers. For cybercrime, comparative analysis suggests that older network members may be relatively more influential, plausibly because the transmission of technical knowledge benefits from more experienced mentors or role models (Weulen Kranenbarg, 2018). Recent reviews reinforce a developmental reading of this pattern, describing a common pathway in which an early interest in technology and online gaming brings individuals into contact with more knowledgeable peers from whom they learn, and through whom they drift into cyber-offending (Loggen et al., 2024). The social dimension of cybercrime is thus real but reorganised: the relevant ties may be weaker, more knowledge-oriented, more age-heterogeneous, and more likely to be mediated by online communities than the deviant peer groups that drive traditional crime.

5.3. Routine activities and criminal opportunity

Because cyber-dependent crime requires neither physical convergence nor physical presence, the routine activities that generate criminal opportunity differ markedly from those associated with traditional offending. Whereas traditional offending opportunities increase with unstructured time spent outside the home, cyber-offending opportunities are concentrated in settings where IT systems and technical skills are available and exercised (Bernasco et al., 2013; Xu et al., 2013). One of the more counterintuitive comparative findings concerns the role of employment and education. For traditional crime, employment and enrolment in education tend to reduce offending, consistent with life-course and social-control theory (Sampson & Laub, 1993). For cyber-offending, the comparative longitudinal evidence found no equivalent protective effect; indeed, employment and education in the IT sector were associated with an increased likelihood of cyber-offending, presumably because they expand access, skills, and opportunity (Weulen Kranenbarg, 2018).

The same logic applies to the victim-offender relationship. In traditional interpersonal crime, offender and victim typically converge physically, and the offender ordinarily

witnesses the victim's reaction. Online, interactions are mediated by an IT system, may be asynchronous, and rarely confront the offender with the victim's distress (Goldsmith & Brewer, 2015; Suler, 2004). This mediated, often invisible interaction reduces the immediate social and emotional feedback that can inhibit offending. It also helps explain why a victim-offender overlap exists for cybercrime, as it does for traditional crime, even though physical convergence is absent: individuals whose online routines expose them to offending opportunities are frequently exposed to victimisation risks as well (Kerstens & Jansen, 2016; Van Wilsem, 2013; Weulen Kranenbarg, 2018).

Taken together, these findings suggest that opportunity theory remains valuable for cybercrime but must be respecified. The constructs of convergence, guardianship, and exposure retain explanatory force, yet their empirical referents shift from physical space and physical co-presence to network connectivity, technical configuration, and online routine. A mechanical application of offline opportunity measures to online offending would misidentify both the situations that generate risk and the points at which intervention is feasible.

5.4. Individual risk factors: self-control and IT-skills

The comparison of individual-level risk factors yields a nuanced picture in which one classic predictor behaves unevenly while a domain-specific predictor emerges as central. Low self-control, the master variable of the general theory of crime, does predict some cyber-deviance, but its effect is concentrated among offenders who also experience victimisation and who tend to commit the less technically sophisticated offences (Weulen Kranenbarg, 2018). For offenders who commit the more technically demanding intrusions while avoiding victimisation, low self-control is not a statistically significant risk factor; if anything, these offenders display the planning capacity and deferral of gratification that low self-control would predict against. Recent syntheses likewise report that low self-control is a consistent but far from universal correlate of cyber-offending, and that its salience varies with the type of conduct examined (Loggen et al., 2024).

The factor that most clearly distinguishes cyber-offenders is the possession of IT-skills. Across studies, and whether measured by self-report or by objective testing, technical skill is a strong and consistent predictor of cyber-dependent offending (Bossler & Burruss, 2011; Holt, Bossler, & May, 2012; Weulen Kranenbarg, 2018). This has no real analogue in traditional crime, where offending generally requires no specialised, transferable competence. Two features of these skills are criminologically important. First, they are typically acquired gradually and partly through illicit experimentation, so that offending and skill-acquisition are intertwined: committing the offence is itself a way of learning. Second, they are dual-use, equally applicable to legitimate and illegitimate ends, which means that the same individual attribute that elevates offending risk also constitutes valuable human capital. The policy implication, that suppressing the skill is neither feasible nor desirable, and that redirection toward legitimate use is the more rational goal, follows directly.

Demographic risk markers are partly shared and partly distinctive. Like traditional offenders, cyber-offenders are disproportionately young and male (Loggen et al., 2024). Unlike traditional offenders, their offending is bound up with sustained technological engagement, so that frequent and skilled internet use functions as both an exposure variable and a capability variable. The overall pattern is therefore one of partial overlap: cyber-offenders are not a wholly alien population, but the configuration of risk factors that characterises them (high technical capability, variable and offence-specific self-control,

and youth combined with intensive online engagement) is sufficiently distinct to warrant separate analytic and policy treatment.

The dual-use character of technical skill also reframes the relationship between risk and desistance. In traditional criminology, the human capital acquired through legitimate employment tends to pull individuals away from crime. In cyber-offending, the same capability is implicated in both offending and legitimate achievement, which means that desistance is less a matter of acquiring new skills than of redirecting existing ones. Recent pathway research describes desistance from cyber-dependent crime as driven by a maturing cost-benefit calculus and by the availability of legitimate avenues (employment, recognition, and intellectual challenge) that satisfy the same underlying motivations as offending (Loggen et al., 2024). This points to a workforce dimension largely absent from traditional crime: a robust, ethically framed cybersecurity skills pipeline is not only an economic asset but a criminological one, in so far as it offers the intrinsically motivated, technically capable individual a legitimate outlet for the very drives that might otherwise be expressed through offending.

5.5. Offending over the life-course and the question of distinctiveness

A further comparative question concerns whether cyber-offenders and traditional offenders are the same people behaving in different arenas, or genuinely distinct populations. The clustering evidence is informative here. When self-reported offences are grouped according to which acts tend to be committed by the same individuals, cyber-dependent offences cluster together and seldom co-occur with traditional offences; the offences within a cluster tend to share a *modus operandi*, a skill set, and a context (Weulen Kranenbarg, 2018). In other words, the individuals who commit cyber-dependent crimes rarely also commit conventional crimes, which supports the view that they constitute a specific type of offender rather than generalists who happen to use a computer.

Life-course patterns reinforce this reading while preserving important continuities. Strong social bonds (living with a partner, or with a partner and a child) are associated with a reduced likelihood of cyber-offending, just as they are for traditional crime, indicating that informal social control retains some purchase even over conduct that is largely invisible to those bonds (Weulen Kranenbarg, 2018). At the same time, the professional and educational circumstances that protect against traditional offending do not protect equally against cyber-offending, and may even increase it where they expand technical opportunity. The developmental literature adds that desistance from cyber-offending appears to be driven less by the conventional turning points of marriage and employment than by a maturing cost-benefit calculus and the redirection of technical interest toward legitimate goals (Loggen et al., 2024). Cyber-offending thus exhibits a life-course signature that overlaps with, but is not identical to, that of traditional crime.

Read together, the five strands of the comparison converge on a coherent portrait. The distinctive motivational profile, the reorganised social influence, the technologically defined opportunity structure, the offence-specific operation of self-control, and the central role of dual-use skill are not independent peculiarities but facets of a single underlying fact: cyber-dependent offending is shaped by a context in which behaviour is mediated by technology, dispersed across networks, and only weakly visible to the mechanisms (social, situational, and legal) that ordinarily restrain conduct. This coherence is itself analytically important, because it indicates that the differences between cyber-offenders and traditional offenders are systematic rather than incidental, and that they can be traced to identifiable features of the digital environment rather than to idiosyncrasies of particular samples.

5.6. The transnational and organisational dimension

The four dimensions examined above operate at the level of the individual offender, but cyber-offending also possesses a collective and transnational dimension that bears on the comparison with traditional crime. Whereas serious traditional crime often requires durable, geographically concentrated organisation, cyber-offending can be coordinated through fluid, dispersed networks in which actors with complementary skills cooperate without ever meeting and frequently without trusting one another (Hutchings, 2014; Hutchings & Clayton, 2016). The commodification of attack tools and services, illustrated by markets in which capabilities such as denial-of-service attacks are sold to those who lack the skill to mount them, has further lowered the threshold for participation, partially decoupling offending from the deep technical mastery that earlier characterised it (Hutchings & Clayton, 2016).

This organisational fluidity intersects with the social-network findings discussed earlier. If some of the knowledge transmission and coordination that traditional crime achieves through strong local ties is achieved online through weak, distributed, and partly commercial relationships, then the comparatively weaker influence of strong proximate peers observed for cybercrime becomes more intelligible. It also complicates the offender typologies of Section 2: a single incident may combine a low-skill purchaser, a skilled tool-developer, and an infrastructure provider, none of whom maps neatly onto a unitary offender category. As cybercrime continues to evolve and to professionalise, the boundary between the lone, intrinsically motivated explorer and the participant in an organised, financially oriented enterprise is likely to blur further (Grabosky, 2017).

Table 1. Comparative summary of cyber-dependent and traditional offenders across four criminological dimensions

Dimension	Traditional offenders	Cyber-dependent offenders
Primary motivation	Predominantly instrumental and financial; expressive motives in interpersonal crime	Predominantly intrinsic (curiosity, challenge, skill-building); financial motive comparatively marginal; some expressive and status motives
Social network influence	Strong; deviance most similar among same-aged, proximate peers	Significant but weaker; older and knowledge-oriented ties relatively more important; often mediated by online communities
Routine activities / opportunity	Opportunity rises with unstructured time outside the home; physical convergence required	Opportunity rises with access to IT systems and skills; no physical convergence; employment/education in IT may increase offending
Self-control	Low self-control a robust, general predictor	Low self-control predicts low-skill, victim-offender conduct; weaker or reversed for sophisticated intrusions
Specific capability	No specialised, transferable skill required	IT-skills a strong, consistent predictor; dual-use and acquired partly through offending
Offender–victim interaction	Typically physical, synchronous, visible	Mediated, often asynchronous and invisible; reduced immediate feedback and guardianship

Note. Synthesised from the comparative empirical literature reviewed in Section 5, drawing in particular on Weulen Kranenbarg (2018) and Loggen et al. (2024).

6. DISCUSSION: IMPLICATIONS FOR PREVENTION, DIGITAL RESILIENCE, AND SECURITY POLICY

The comparative analysis supports a position between two extremes. Cyber-offending is neither a wholly novel phenomenon requiring an entirely new criminology, nor a simple replication of traditional crime in a new medium. Many correlates point in the same direction across the two domains, yet in each dimension the digital context reweights, respecifies, or reverses relationships that are well established for traditional crime. The practical lesson is that interventions and measures should be adapted to the digital domain rather than transplanted into it unexamined.

For prevention, the motivational evidence has direct consequences. Because intrinsic, skill-related motivations predominate in cyber-dependent offending while financial motivation is comparatively marginal, situational measures aimed at reducing expected financial gain are unlikely to be effective on their own. Strategies that acknowledge the offender’s curiosity and desire for technical mastery, and that channel these toward

legitimate outlets, are theoretically better aligned with the evidence. The dual-use nature of the relevant skills makes such redirection feasible: the abilities that enable an intrusion are precisely those valued in the cybersecurity profession. Programmes that combine ethical guidance with legitimate avenues for exercising technical skill (for example, mentoring by experienced practitioners and structured frameworks for responsible disclosure of vulnerabilities) are promising, although they remain insufficiently evaluated and should be implemented with rigorous evaluation built in (Weulen Kranenbarg, 2018).

For deterrence, the disinhibition and routine-activity findings suggest that increasing the offender's perception that conduct is observed and consequential can matter. Experimental work on warning banners indicates that signalling surveillance and likely consequences can shape the behaviour of system intruders (Maimon et al., 2014). More broadly, raising the visibility of online conduct counteracts the perception that cyberspace is a realm without real-world consequences, a perception that the disinhibition literature identifies as a facilitator of offending (Suler, 2004).

For digital resilience and education, the finding that protective life circumstances do not operate online as they do offline implies that resilience cannot be built through social-bonding mechanisms alone. Because young people increasingly acquire technical skills through formal education and informal experimentation, and because skill is the strongest distinctive predictor of cyber-offending, the ethical and responsible use of technology must be integrated into technical education rather than treated as an afterthought. Building digital literacy that includes both capability and an ethic of responsible use is, on this evidence, a security investment as much as an educational one.

A further implication concerns measurement, victimisation, and reporting. Cyber-dependent offending is characterised by a particularly large dark figure: a substantial share of incidents is never detected by victims, and a further share is detected but not reported to the authorities, so that recorded statistics capture only a fraction of actual offending (Tcherni et al., 2016; McGuire & Dowling, 2013). This has two consequences for the present argument. First, it cautions against treating official data as a direct measure of the offender population, and it reinforces the value of self-report and offender-based designs of the kind on which the comparative evidence synthesised here largely rests. Second, it points to the porousness of the victim–offender boundary online: studies of financially motivated cyber-offending find a notable overlap between those who are victimised and those who offend, suggesting that exposure, opportunity, and offending are entangled in ways that differ from the sharper victim–offender separation often assumed for traditional crime (Kerstens & Jansen, 2016). For prevention and security policy, this implies that measures aimed at strengthening victim reporting, improving the detection of incidents, and supporting potential victims are not separate from measures aimed at offenders, because the same individuals and the same online routines may occupy both roles. Improving the evidence base, through better victimisation surveys, more consistent reporting channels, and research that follows individuals across both roles, is therefore a precondition for calibrating the interventions discussed above.

For security policy, finally, the recognition that cyber-offenders are a distinct population reframes the institutional response. Effective policy requires interoperability between technical and human-centred expertise, investigative capacity attuned to the transnational and networked structure of cyber-offending, and prevention that engages educational institutions, employers, and online communities. For jurisdictions in the Western Balkans and the wider region, where the evidence base is still thin, a priority is the development of locally grounded research so that prevention and intervention are

calibrated to regional patterns rather than imported wholesale from contexts that may differ in important respects.

The era of smart technologies sharpens each of these implications. Artificial intelligence and automation are double-edged: they furnish defenders with powerful tools for anomaly detection and predictive analysis, but they also lower the skill threshold for some attacks and raise the sophistication of others, potentially widening the population capable of offending and accelerating the professionalisation already visible in cyber-criminal markets (Hutchings & Clayton, 2016; Maimon & Louderback, 2019). If access to capability is decoupled from deep technical mastery, the motivational logic of offending may shift, with intrinsic, skill-driven motives sharing space with more instrumental ones as participation broadens. At the same time, the proliferation of autonomous systems and critical cyber-physical infrastructure raises the stakes of intrusion, so that offences once experienced by offenders as victimless exploration may carry consequences for public safety. These developments do not overturn the criminological findings synthesised here, but they intensify the need to understand offender motivation, social organisation, and opportunity structures, because purely technical defences will not address the human decisions that drive offending.

6.1. Cyber-offending and the Western Balkans context

These implications take on a particular character in the Western Balkans, where the institutional and legislative response to cybercrime has advanced more quickly than the criminological evidence base that should guide it. Serbia is illustrative. It has ratified the Council of Europe Convention on Cybercrime and its Additional Protocol, transposed the Convention's substantive and procedural standards into domestic law, and created dedicated institutions: a Special Prosecutor's Office for High-Tech Crime, a specialised cybercrime department within the Ministry of the Interior, and the round-the-clock contact point required for international cooperation (Council of Europe, 2001). Successive national strategies, most recently the strategy for the fight against high-tech crime for the period to 2030, have been prepared with Council of Europe support and aligned explicitly with the Convention and its Second Additional Protocol.

This effort is reinforced regionally. Because cybercrime falls within the justice and home-affairs chapter of European Union accession negotiations, the Western Balkan states face continuous pressure to harmonise their legislation with the EU *acquis* (European Parliament, 2021), and a dense layer of capacity-building initiatives, including the Council of Europe and European Union iPROCEEDS programmes and the regional Western Balkans Cyber Capacity Centre, now trains prosecutors, judges, and investigators across the region. The transnational logic emphasised in Section 5.6 is visible in practice: Serbian authorities took part, through mutual legal assistance, in the coordinated international operation that dismantled one of the largest denial-of-service-for-hire (booter) platforms, an episode that exemplifies both the commodification of attack tools discussed earlier and the functioning of cross-border cooperation (Hutchings & Clayton, 2016).

This institutional maturity is not, however, matched by criminological knowledge of the offenders themselves. Regional activity is oriented toward enforcement, electronic evidence, and cooperation, whereas offender-based research of the kind synthesised here (on motivation, social networks, opportunity, and risk factors) remains almost wholly absent for the region. The result is a policy paradox: prevention is being designed for a population about whom little is known locally. Adapting established offender-based instruments to regional conditions and generating local data is therefore the precondition

for prevention calibrated to the Western Balkans rather than transposed from the Western European and North American settings in which the current evidence base was produced.

6.2. Limitations and directions for future research

Several limitations qualify the conclusions advanced in this paper. The comparative evidence rests heavily on a small number of national contexts and on samples of detected or self-reported offenders, so that selection effects and limited geographical coverage constrain generalisability. Differences in how studies operationalise cybercrime, and the frequent conflation of cyber-dependent and cyber-enabled conduct, further complicate synthesis. Many of the prevention and intervention approaches discussed in Section 6 are theoretically motivated but have not yet been rigorously evaluated for cybercrime, and should therefore be implemented with evaluation designs attached.

These limitations define a research agenda. There is a particular need for longitudinal and comparative studies that follow cyber-offenders over time, that compare offenders who are apprehended with those who evade detection, and that examine selection and influence processes within the online communities through which technical knowledge is transmitted (Loggen et al., 2024; Weulen Kranenbarg, 2018). For the Western Balkans specifically, the development of regional data and the adaptation of established criminological instruments to local conditions would allow prevention and security policy to rest on evidence drawn from the populations they are intended to serve, rather than on findings transposed from elsewhere.

7. CONCLUSION

This paper set out to compare cyber-offenders and traditional offenders across four criminological dimensions (motivation, social networks, routine activities, and individual risk factors) in order to clarify whether and how the two populations differ. The comparison yields a consistent answer: cyber-dependent offenders are a distinct type of offender whose conduct overlaps with, but cannot be reduced to, that of traditional offenders. They are driven primarily by intrinsic motivations of curiosity and technical challenge rather than by financial gain; they are embedded in social networks whose influence is real but weaker, more knowledge-oriented, and more age-heterogeneous than the deviant peer groups of traditional crime; their criminal opportunities arise from technological access and skill rather than from unstructured time in physical space; and they are distinguished above all by the possession of dual-use technical competencies, with low self-control predicting only the less sophisticated end of the spectrum.

These differences are the product of the context in which cybercrime is committed: a setting that is technologically mediated, frequently anonymous, often invisible to the agents of informal social control, and characterised by low perceived risk of apprehension. That context does not abolish the mechanisms identified by traditional criminological theory, but it reweights and respecifies them. Routine activity theory, the general theory of crime, and social learning theory all retain explanatory value for cybercrime, provided their constructs are adapted to the digital domain and their predictions are tested rather than assumed.

The implication for the era of smart technologies is clear. As connectivity deepens and as artificial intelligence and autonomous systems expand both the attack surface and the toolkit available to offenders, the temptation to treat cybercrime as a purely technical problem will grow. The evidence reviewed here counsels against that temptation.

Cybercrime is a criminological and social phenomenon as much as a technological one, and effective prevention, resilience, and security policy will depend on understanding the human offender behind the keyboard. The most important task for future research, particularly in under-studied regions such as the Western Balkans, is to build the longitudinal and comparative evidence base needed to translate this understanding into interventions whose effectiveness can be demonstrated rather than merely assumed.

8. REFERENCES

- Akers, R. L. (2009). *Social learning and social structure: A general theory of crime and deviance*. Routledge. <https://doi.org/10.4324/9781315129587>
- Bachmann, M. (2010). The risk propensity and rationality of computer hackers. *International Journal of Cyber Criminology*, 4(1), 643-656.
- Bernasco, W., Ruiter, S., Bruinsma, G. J. N., Pauwels, L. J. R., & Weerman, F. M. (2013). Situational causes of offending: A fixed-effects analysis of space–time budget data. *Criminology*, 51(4), 895-926. <https://doi.org/10.1111/1745-9125.12023>
- Bossler, A. M. & Burruss, G. W. (2011). The General Theory of Crime and Computer Hacking: Low Self-Control Hackers?. In T. Holt & B. Schell (Eds.), *Corporate Hacking and Technology-Driven Crime: Social Dynamics and Implications* (pp. 38-67). IGI Global Scientific Publishing. <https://doi.org/10.4018/978-1-61692-805-6.ch003>
- Bossler, A. M., & Holt, T. J. (2009). On-line activities, guardianship, and malware infection: An examination of routine activities theory. *International Journal of Cyber Criminology*, 3(1), 400-420.
- Brenner, S. W. (2006). Cybercrime jurisdiction. *Crime, Law and Social Change*, 46(4-5), 189-206.
- Chua, Y. T., & Holt, T. J. (2016). A cross-national examination of the techniques of neutralization to account for hacking behaviors. *Victims & Offenders*, 11(4), 534-555. <https://doi.org/10.1080/15564886.2015.1121944>
- Cohen, L. E., & Felson, M. (1979). Social change and crime rate trends: A routine activity approach. *American Sociological Review*, 44(4), 588-608.
- Council of Europe. (2001). *Convention on Cybercrime* (European Treaty Series No. 185)
- Donner, C. M., Marcum, C. D., Jennings, W. G., Higgins, G. E., & Banfield, J. (2014). Low self-control and cybercrime: Exploring the utility of the general theory of crime beyond digital piracy. *Computers in Human Behavior*, 34, 165-172. <https://doi.org/10.1016/j.chb.2014.01.040>
- European Parliament. (2021). *Report on cooperation on the fight against organised crime in the Western Balkans* (A9-0298/2021).
- Goldsmith, A., & Brewer, R. (2015). Digital drift and the criminal interaction order. *Theoretical Criminology*, 19(1), 112-130. <https://doi.org/10.1177/1362480614538645>
- Gordon, S., & Ford, R. (2006). On the definition and classification of cybercrime. *Journal in Computer Virology*, 2(1), 13-20.
- Gottfredson, M. R., & Hirschi, T. (1990). *A general theory of crime*. Stanford University Press.
- Grabosky, P. N. (2001). Virtual criminality: Old wine in new bottles? *Social & Legal Studies*, 10(2), 243-249. <https://doi.org/10.1177/a017405>
- Grabosky, P. N. (2017). The evolution of cybercrime, 2006–2016. In T. J. Holt (Ed.),

- Cybercrime through an interdisciplinary lens (pp. 15-36). Routledge.
- Holt, T. J. (2007). Subcultural evolution? Examining the influence of on- and off-line experiences on deviant subcultures. *Deviant Behavior*, 28(2), 171-198. <https://doi.org/10.1080/01639620601131065>
- Holt, T. J., & Bossler, A. M. (2014). An assessment of the current state of cybercrime scholarship. *Deviant Behavior*, 35(1), 20-40. <https://doi.org/10.1080/01639625.2013.822209>
- Holt, T. J., Bossler, A. M., & May, D. C. (2012). Low self-control, deviant peer associations, and juvenile cyberdeviance. *American Journal of Criminal Justice*, 37(3), 378-395. <https://doi.org/10.1007/s12103-011-9117-3>
- Holt, T. J., Burruss, G. W., & Bossler, A. M. (2010). Social learning and cyber-deviance: Examining the importance of a full social learning model in the virtual world. *Journal of Crime and Justice*, 33(2), 31-61. <https://doi.org/10.1080/0735648X.2010.9721287>
- Hutchings, A. (2014). Crime from the keyboard: Organised cybercrime, co-offending, initiation and knowledge transmission. *Crime, Law and Social Change*, 62(1), 1-20. <http://dx.doi.org/10.1007/s10611-014-9520-z>
- Hutchings, A., & Clayton, R. (2016). Exploring the provision of online booter services. *Deviant Behavior*, 37(10), 1163-1178. <https://doi.org/10.1080/01639625.2016.1169829>
- Jaishankar, K. (2009). Space transition theory of cyber crimes. In F. Schmallegger & M. Pittaro (Eds.), *Crimes of the Internet* (pp. 283-301). Pearson Education.
- Kerstens, J., & Jansen, J. (2016). The victim-perpetrator overlap in financial cybercrime: Evidence and reflection on the overlap of youth's on-line victimization and perpetration. *Deviant Behavior*, 37(5), 585-600. <https://doi.org/10.1080/01639625.2015.1060796>
- Leukfeldt, E. R., & Yar, M. (2016). Applying routine activity theory to cybercrime: A theoretical and empirical analysis. *Deviant Behavior*, 37(3), 263-280. <https://doi.org/10.1080/01639625.2015.1012409>
- Loggen, J., Moneva, A., & Leukfeldt, R. (2024). Pathways into, desistance from, and risk factors related to cyber-dependent crime: A systematic narrative review. *Victims & Offenders*. Advance online publication. <https://doi.org/10.1080/15564886.2024.2370295>
- Maimon, D., Alper, M., Sobesto, B., & Cukier, M. (2014). Restrictive deterrent effects of a warning banner in an attacked computer system. *Criminology*, 52(1), 33-59. <https://doi.org/10.1111/1745-9125.12028>
- Maimon, D., & Louderback, E. R. (2019). Cyber-dependent crimes: An interdisciplinary review. *Annual Review of Criminology*, 2, 191-216. <https://doi.org/10.1146/annurev-criminol-032317-092057>
- McGuire, M., & Dowling, S. (2013). *Cyber crime: A review of the evidence*. Chapter 1: Cyber-dependent crimes. Home Office Research Report 75. UK Home Office.
- McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, 27(1), 415-444. <https://doi.org/10.1146/annurev.soc.27.1.415>
- Morris, R. G. (2011). Computer hacking and the techniques of neutralization: An empirical assessment. In T. J. Holt & B. H. Schell (Eds.), *Corporate hacking and technology-driven crime: Social dynamics and implications* (pp. 1-17). Information Science Reference. <https://doi.org/10.4018/978-1-61692-805-6.ch001>

- Parker, D. B. (1983). *Fighting computer crime*. Scribner.
- Pratt, T. C., & Cullen, F. T. (2000). The empirical status of Gottfredson and Hirschi's general theory of crime: A meta-analysis. *Criminology*, 38(3), 931-964. <https://doi.org/10.1111/j.1745-9125.2000.tb00911.x>
- Pratt, T. C., Cullen, F. T., Sellers, C. S., Winfree, L. T., Madensen, T. D., Daigle, L. E., Fearn, N. E., & Gau, J. M. (2009). The empirical status of social learning theory: A meta-analysis. *Justice Quarterly*, 27(6), 765-802. <https://doi.org/10.1080/07418820903379610>
- Rogers, M. K. (2006). A two-dimensional circumplex approach to the development of a hacker taxonomy. *Digital Investigation*, 3(2), 97-102. <https://doi.org/10.1016/j.diin.2006.03.001>
- Rogers, M. K. (2011). The psyche of cybercriminals: A psycho-social perspective. In S. Ghosh & E. Turrini (Eds.), *Cybercrimes: A multidisciplinary analysis* (pp. 217-235). Springer. https://doi.org/10.1007/978-3-642-13547-7_14
- Sampson, R. J., & Laub, J. H. (1993). *Crime in the making: Pathways and turning points through life*. Harvard University Press.
- Seebruck, R. (2015). A typology of hackers: Classifying cyber malfeasance using a weighted arc circumplex model. *Digital Investigation*, 14, 36-45. <https://doi.org/10.1016/j.diin.2015.07.002>
- Skinner, W. F., & Fream, A. M. (1997). A social learning theory analysis of computer crime among college students. *Journal of Research in Crime and Delinquency*, 34(4), 495-518. <https://doi.org/10.1177/0022427897034004005>
- Suler, J. (2004). The online disinhibition effect. *CyberPsychology & Behavior*, 7(3), 321-326. <https://doi.org/10.1089/1094931041291295>
- Sykes, G. M., & Matza, D. (1957). Techniques of neutralization: A theory of delinquency. *American Sociological Review*, 22(6), 664-670. <https://doi.org/10.2307/2089195>
- Taylor, P. A. (1999). *Hackers: Crime in the digital sublime*. Routledge.
- Tcherni, M., Davies, A., Lopes, G., & Lizotte, A. (2016). The dark figure of online property crime: Is cyberspace hiding a crime wave? *Justice Quarterly*, 33(5), 890-911. <https://doi.org/10.1080/07418825.2014.994658>
- Turgeman-Goldschmidt, O. (2008). Meanings that hackers assign to their being a hacker. *International Journal of Cyber Criminology*, 2(2), 382-396.
- Van Wilsem, J. A. (2013). Hacking and harassment-Do they have something in common? Comparing risk factors for online victimization. *Journal of Contemporary Criminal Justice*, 29(4), 437-453. <https://doi.org/10.1177/1043986213507402>
- Voiskounsky, A. E., & Smyslova, O. V. (2003). Flow-based model of computer hackers' motivation. *CyberPsychology & Behavior*, 6(2), 171-180. <https://doi.org/10.1089/109493103321640365>
- Wall, D. S. (2001). Cybercrimes and the Internet. In D. S. Wall (Ed.), *Crime and the Internet* (pp. 1-17). Routledge.
- Warr, M. (2002). *Companions in crime: The social aspects of criminal conduct*. Cambridge University Press.
- Weerman, F. M., & Smeenk, W. H. (2005). Peer similarity in delinquency for different types of friends: A comparison using two measurement methods. *Criminology*, 43(2), 499-524. <https://doi.org/10.1111/j.0011-1348.2005.00015.x>
- Weulen Kranenbarg, M. (2018). *Cyber-offenders versus traditional offenders: An empirical comparison* [Doctoral dissertation, Vrije Universiteit Amsterdam]. NSCR / Vrije Universiteit Amsterdam.

- Woo, H.-J., Kim, Y., & Dominick, J. (2004). Hackers: Militants or merry pranksters? A content analysis of defaced web pages. *Media Psychology*, 6(1), 63-82. https://doi.org/10.1207/s1532785xmep0601_3
- Xu, Z., Hu, Q., & Zhang, C. (2013). Why computer talents become computer hackers. *Communications of the ACM*, 56(4), 64-74. <https://doi.org/10.1145/2436256.2436272>
- Yar, M. (2005). The novelty of 'cybercrime': An assessment in light of routine activity theory. *European Journal of Criminology*, 2(4), 407-427. <https://doi.org/10.1177/147737080556056>
- Young, R., Zhang, L., & Prybutok, V. R. (2007). Hacking into the minds of hackers. *Information Systems Management*, 24(4), 281-287. <https://doi.org/10.1080/10580530701585823>

DIGITAL REALITY AND CYBER CRISIS FORMATION: COMMUNICATION UNDER UNCERTAINTY IN POLITICALLY SALIENT PUBLIC DIGITAL SYSTEMS

Ivan Ristov

Agency for National Security, Republic of North Macedonia

ivan.ristov@yahoo.com; ivan_ristov@uklo.edu.mk

ORCID iD: <https://orcid.org/0009-0008-9004-1389>

Abstract

Cyber incidents affecting politically salient public digital systems can develop into wider political crises even when their technical effects remain bounded or their attribution unresolved. This article develops digital reality as an analytical framework for explaining how this transformation occurs. Digital reality is defined as the temporally bounded, digitally mediated interpretive process through which incomplete information about a cyber incident becomes publicly meaningful and politically consequential before authoritative evidence stabilises. The study employs a qualitative comparative case-study design, combining structured, focused comparison with process-tracing-informed within-case analysis of three Balkan cases: the 2019 breach of Bulgaria's National Revenue Agency, the 2020 disruption of North Macedonia's election-results platform, and the 2022 cyber crisis affecting Montenegro's government infrastructure. The findings identify three recurrent mechanisms: speed, which creates a gap between public interpretation and institutional verification; ambiguity, which sustains multiple plausible accounts of authorship, intent, and significance; and fragmentation, which disperses knowledge and communicative authority across actors with different mandates and evidentiary thresholds. Their interaction contributed to distinct forms of political crisis expansion: institutional and cross-border trust consequences in Bulgaria, contested informational authority in North Macedonia, and an internationalised security response under unresolved attribution in Montenegro. The article concludes that cyber resilience requires not only technical recovery but also coordinated communication capable of distinguishing verified facts, provisional assessments, competing hypotheses, and unresolved questions.

Keywords: *digital reality; cyber crisis communication; attribution uncertainty; strategic narratives; cyber resilience*

1. INTRODUCTION

Cyber incidents affecting public institutions are not confined to networks, information systems, and operational recovery. Once they become publicly visible, they also generate demands for explanation, attribution, reassurance, and political response. These demands usually arise before technical investigations have been completed and may transform a bounded disruption into a wider crisis of institutional trust, electoral transparency, national security, or international cooperation.

Technical severity remains indispensable to any assessment of cyber harm. Data compromise, service interruption, unauthorised access, operational disruption, and

recovery costs determine much of an incident's material impact. They do not, however, determine its political trajectory on their own. Technically similar incidents may produce markedly different political consequences, while an apparently limited availability failure may become highly consequential when it affects an election, a compulsory public database, or the continuity of government. The political meaning of an incident is therefore influenced not only by forensic evidence but also by the timing and content of official statements, media reporting, expert commentary, political claims, public reactions, and competing accounts of responsibility. Cyber crises consequently unfold under a structural temporal asymmetry. Forensic investigation, attribution assessment, evidence preservation, interagency coordination, and consultation with foreign partners require time. Digital communication operates more rapidly. Government officials, journalists, political actors, cybersecurity companies, foreign observers, and online communities may begin interpreting an incident within minutes. This does not merely create a risk that inaccurate information will circulate. Early descriptions can establish a first-frame advantage: once an event is publicly understood as foreign aggression, criminal ransomware, electoral interference, or institutional failure, subsequent evidence may be interpreted through that frame. Attribution intensifies this problem. It is not a binary technical determination but a layered process linking technical indicators, operational analysis, strategic judgment, and political decisions concerning public communication. Evidence capable of identifying malicious software or operational infrastructure may not establish the identity of the operator, organisational affiliation, strategic sponsorship, intent, or legal State responsibility. Governments may therefore be expected to communicate politically consequential conclusions before the different layers of attribution have converged (Rid & Buchanan, 2015, pp. 4–11). This article conceptualises the interpretive process created by that temporal gap as digital reality. The term is not used as a synonym for cyberspace, nor does it refer to virtual, augmented, mixed, or extended-reality technologies. It denotes a temporally bounded, digitally mediated interpretive process through which incomplete information about a cyber incident becomes publicly meaningful and politically consequential before authoritative evidence stabilises. The concept develops an earlier formulation of digital reality as an algorithmically shaped domain mediating subjective and intersubjective understandings (Ristov, 2020, pp. 61–66). The present article narrows that formulation to the interval between public awareness of an incident and provisional evidentiary closure.

The central research question is: **How do digitally mediated interpretations shape the political development of cyber incidents affecting politically salient public digital systems, and what governance implications follow for small and medium-sized states?**

The article argues that three interrelated mechanisms structure this process. Speed creates a gap between public interpretation and institutional verification. Ambiguity permits multiple plausible accounts of authorship, intent, scope, and strategic significance to coexist. Fragmentation disperses knowledge and communicative authority among institutions and actors operating with different mandates, evidence, and confidence thresholds. When these mechanisms interact, an incident may move beyond routine technical management before institutions establish a coordinated and evidentially calibrated account.

The outcome examined is defined as **political crisis expansion**, rather than interstate escalation alone. Political crisis expansion refers to movement from technical incident response into broader contestation involving institutional legitimacy, public trust,

securitised framing, foreign consultation, diplomatic signalling, or pressure for visible governmental action. These outcomes vary in form and intensity. Domestic politicisation and erosion of institutional trust are analytically distinct from internationalisation, formal attribution of State responsibility, or direct interstate escalation. This distinction avoids treating every publicly consequential incident as equivalent.

The argument is examined through three cases from the Balkans. The 2019 breach of Bulgaria's National Revenue Agency shows how externally initiated disclosure produced regulatory accountability and cross-border trust consequences. The 2020 disruption of North Macedonia's election-results platform illustrates how the temporary loss of an authoritative public information channel generated contested informational authority without established alteration of ballots, counting, or tabulation. The 2022 attacks on Montenegro's government infrastructure demonstrate how rapid national-security framing, criminal-group identification, international assistance, and unresolved State attribution produced an internationalised security response. The cases vary in technical form and political outcome, allowing the framework to be examined across a confidentiality breach, an election-related availability disruption, and a broader governmental cyber crisis. The article makes three contributions. First, it develops digital reality as a process-based framework linking technical incidents to political crisis formation. Second, it integrates insights from cyber escalation and attribution, crisis communication, strategic narratives, and cyber diplomacy. Third, it identifies a governance gap between technical incident response and the coordination of public, political, and international communication under uncertainty. The practical implication is not that governments should attribute more quickly, but that they should communicate uncertainty more systematically. The article proceeds as follows. Section 2 integrates the relevant literature with the digital-reality framework and specifies the three mechanisms. Section 3 explains the comparative design, evidence base, and limitations. Section 4 analyses the three cases. Section 5 compares the operation and interaction of the mechanisms. Section 6 discusses the broader theoretical and policy implications. Section 7 concludes.

2. LITERATURE AND ANALYTICAL FRAMEWORK

2.1. Cyber Escalation, Attribution, and Crisis Communication

Research on cyber conflict has often focused on whether cyber capabilities fundamentally transform warfare or remain constrained instruments of coercion. Kello (2013) emphasises the strategic expansion created by cyber capabilities, while Gartzke (2013), Borghard and Lonergan (2017), and Maschmeyer (2021) underline their operational and coercive limits. These debates are important because they show that technical effects alone do not determine political outcomes. A technically bounded incident may still become politically salient when interpreted as electoral interference, institutional failure, foreign coercion, or evidence of wider hybrid pressure. Escalation research identifies uncertainty and misperception as central risks. Lin (2012) argues that ambiguous signalling, incomplete information, and attribution difficulties can generate disproportionate responses. Buchanan (2017) similarly shows that cyber activities conducted for intelligence or defensive purposes may be interpreted as preparation for attack. In such conditions, the strategic meaning of technical activity depends not only on what occurred within systems, but also on how actors interpret intent, capability, and risk. Attribution is especially consequential because it connects technical evidence with political responsibility. Rid and Buchanan (2015) distinguish tactical-technical, operational, and

strategic levels of attribution. Technical evidence may identify infrastructure, tools, or intrusion methods, but it may not establish operator identity, sponsorship, intent, or legal State responsibility. Confidence may therefore differ across attributional levels, and certainty at one level does not automatically establish certainty at another. Recent work on public cyber attribution reinforces this point by treating attribution not only as analysis but also as governmental communication. Egloff and Smeets (2023) conceptualise public attribution as a strategic decision involving trade-offs among evidentiary confidence, audience effects, coordination, and escalation risks. Lee (2023) shows that public attribution may occur through technical, criminal, official policy, and unofficial policy channels, creating risks of misperception when external audiences interpret separate institutional messages as a unified whole-of-government position. This literature strengthens the article's claim that attribution uncertainty becomes politically consequential when technical assessment, public communication, institutional authority, and diplomatic signalling are insufficiently coordinated. This layered character creates a communication dilemma. Governments may possess technical or intelligence information that cannot be disclosed without compromising sources, methods, or investigations. Public conclusions may consequently be difficult for external audiences to verify. Conversely, postponing communication may create an interpretive vacuum in which political actors, private cybersecurity companies, journalists, foreign governments, and online communities advance competing explanations. Crisis communication research helps explain why that vacuum matters. Reynolds and Seeger (2005) conceptualise crisis communication as a staged process in which initial communication should reduce uncertainty, identify credible spokespersons, and explain what is known without overstating provisional information. Coombs (2007) similarly argues that crisis response should address stakeholder information needs before reputational defence. Applied to cyber incidents, the key implication is that delayed, incoherent, or excessively certain communication can become part of the crisis itself. Cyber crisis communication is also a multi-audience signalling problem. A statement intended to reassure citizens may be interpreted by allies or suspected adversaries as an attribution signal. A domestic claim of control may narrow diplomatic space for later revision. Public communication therefore performs several functions at once: reassurance, institutional justification, political positioning, strategic signalling, and expectation management. Strategic-narrative research explains how incidents are incorporated into broader accounts of identity, threat, responsibility, and appropriate response (Miskimmon et al., 2013). A cyber incident may be interpreted through narratives of governmental incompetence, electoral manipulation, foreign aggression, or hybrid interference. However, not every contested narrative constitutes a coordinated information operation. Hybrid threats ordinarily imply strategic coordination among conventional and unconventional instruments (European Commission, 2016). Where such coordination cannot be established, the more defensible claim is that an incident may produce hybrid-like effects through the interaction of technical disruption and an already contested information environment. International cyber governance recognises some of these risks. United Nations norms emphasise cooperation, fact-based assessment, protection of critical infrastructure, assistance, and consultation intended to reduce misperception (United Nations, 2015, 2021). OSCE confidence-building measures provide for points of contact, rapid communication, information exchange, consultation, and protected channels (OSCE, 2016). Their effectiveness depends on domestic implementation. Dominion (2023) identifies problems including outdated contact information, unclear institutional roles, limited capacity, weak handover arrangements, and insufficient exercises. The literature

therefore explains several parts of cyber crisis behaviour, but rarely integrates them into a single account of how technical uncertainty, public communication, attribution, narrative competition, and diplomatic signalling interact before authoritative assessments have stabilised.

2.2. Digital Reality as a Crisis-Formation Process

The original formulation of digital reality distinguished it from objective, subjective, and intersubjective reality while emphasising their continuous interaction (Ristov, 2020, pp. 59–66). Objective reality concerns phenomena that exist independently of individual interpretation. Subjective reality concerns individual perceptions and experiences. Intersubjective reality consists of meanings and institutions sustained through communication, recognition, and collective practice. Berger and Luckmann's (1966) account of social construction is relevant here because it shows how meanings become objectified, internalised, and reproduced through interaction. Digital reality was initially conceptualised as a continuously expanding domain constituted by digital traces and shaped by devices, platforms, algorithms, search systems, location, previous behaviour, and technological access. Individuals do not encounter the totality of digitally available information. They encounter selected, ranked, and personalised portions of it, which can shape subjective perception and contribute to intersubjective understandings. For cyber crisis analysis, the concept must be narrowed. If digital reality encompasses all digital traces and digitally mediated experience, it becomes indistinguishable from the wider information environment. This article therefore defines digital reality as a temporally bounded, digitally mediated interpretive process through which incomplete information about a publicly visible incident becomes subjectively perceived, intersubjectively reinforced, and politically consequential before authoritative evidence stabilises. This is a process-based rather than a territorial definition. Digital reality is not a parallel world, a separate physical domain, a synonym for cyberspace, or a type of virtual, augmented, mixed, or extended reality. Cyberspace denotes the technical-operational environment of networks, hardware, software, data, protocols, and information systems. It is where vulnerabilities are exploited, malicious code is executed, data are compromised, and services are disrupted. Digital reality concerns how these events become visible, interpreted, narrated, contested, and translated into political expectations and decisions. The distinction avoids two forms of reductionism. Technical determinism assumes that forensic reconstruction automatically determines political meaning and appropriate response. Discursive reductionism assumes that material effects impose no independent constraints and that crises are produced solely through language. The framework rejects both positions. Technical evidence limits the range of credible interpretations, while prior threat perceptions, institutional interests, political identities, and established narratives influence which explanations become politically salient. The process becomes analytically relevant when three conditions converge: a consequential technical event becomes publicly visible; authoritative knowledge of its scope, cause, origin, intent, or consequences remains incomplete; and multiple actors circulate interpretations across different platforms and audiences. The interval between visibility and authoritative explanation constitutes an epistemic gap. During this gap, knowledge is incomplete and asymmetrically distributed. Technical specialists may possess indicators of compromise without understanding the wider political context. Intelligence institutions may possess classified contextual information that cannot be disclosed. Political leaders may perceive strategic risks without sufficient technical confidence. Private companies, journalists, foreign governments, and

citizens observe different parts of the event and apply different standards of evidence. Digital reality is produced through the interaction of these partial accounts. An interpretation does not need universal acceptance to have political effects. Several interpretive communities may stabilise incompatible accounts of the same incident: foreign aggression, criminal activity, institutional incompetence, or electoral manipulation. Decision-makers may respond to a claim not because it has been conclusively demonstrated, but because they believe that citizens, opponents, allies, or suspected adversaries have already incorporated it into their understanding of the crisis. The relationship between interpretation and action is recursive. Public claims influence political incentives and institutional communication. Official statements alter media framing, public expectations, and foreign perceptions. These reactions then affect investigative priorities, attribution thresholds, protective measures, and diplomatic options. Digital reality therefore mediates between technical disruption and political consequence.

2.3. Speed, Ambiguity, and Fragmentation

The framework identifies three interrelated mechanisms: speed, ambiguity, and fragmentation. They are not treated merely as static conditions. Each involves activities through which actors interpret, communicate, select, and contest claims under uncertainty.

Speed refers to temporal compression: consequential interpretations circulate before institutions can complete verification and coordination. A technical incident may become visible through service failure, leaked data, screenshots, unofficial statements, or media reporting. Digital platforms privilege novelty and continuous updating, while forensic and attribution processes require corroboration, evidence preservation, hypothesis testing, and institutional consultation. Speed produces three effects. First, it compresses decision time by creating demands for explanation and action before evidence has stabilised. Second, it generates a first-frame advantage: early descriptions provide causal and moral categories through which subsequent information is interpreted. Third, it produces correction asymmetry. Initial claims may circulate widely, while later qualifications receive less attention. Institutions may therefore face a choice between communicating prematurely and allowing less authoritative actors to dominate the early interpretive environment.

Proposition 1: The greater the temporal gap between public interpretation and institutional verification, the more likely early frames are to shape political expectations and response options.

Ambiguity refers to the coexistence of multiple plausible interpretations. It should be distinguished from uncertainty. Uncertainty denotes insufficient reliable knowledge. Ambiguity arises when incomplete or contested evidence can support several credible accounts. Cyber incidents are especially conducive to ambiguity because operators may conceal their identities, reuse tools, route activity through third-party infrastructure, or imitate other actors. The same incident may plausibly be interpreted as criminal activity, intelligence collection, coercive signalling, hacktivism, preparation for a later operation, or institutional failure. Ambiguity can preserve space for restraint by allowing confidential consultation and delayed attribution. It becomes politically dangerous when provisional assessments are presented as established facts or when distinctions among technical operator, organisational affiliation, political sponsorship, intent, and State responsibility are collapsed.

Proposition 2: The greater the ambiguity concerning authorship, intent, scope, and strategic significance, the greater the opportunity for competing political interpretations; this effect is reduced when communication distinguishes verified facts, probabilistic assessments, competing hypotheses, and unresolved questions.

Fragmentation refers to the dispersion of relevant knowledge and communicative authority across institutions and actors. Cyber crises rarely have one uncontested narrator. Incident-response teams, ministries, intelligence and law-enforcement agencies, political leaders, private contractors, cybersecurity companies, journalists, foreign partners, and online communities may all communicate about the same event. Institutional plurality is not inherently dysfunctional. Different actors possess legitimate expertise and responsibilities. Fragmentation becomes consequential when their assessments are not connected through clear roles, shared terminology, disclosure thresholds, and coordination procedures. Technical, intelligence, legal, political, and diplomatic statements may then appear contradictory even when they address different questions.

Proposition 3: The greater the fragmentation of knowledge and communicative authority, the more difficult it becomes to sustain a credible and de-escalatory account; this effect is reduced when specialised roles are connected through coordinated differentiation, shared confidence language, and sequenced communication. The mechanisms are dynamically interdependent. Speed magnifies ambiguity because actors interpret incomplete evidence under time pressure. Ambiguity intensifies fragmentation because different institutions may adopt different but individually plausible explanations. Fragmentation reinforces speed by multiplying the number of actors able to issue public claims before a coordinated assessment exists.

The expected process is not strictly linear. A technical incident becomes visible; an epistemic gap emerges; actors advance provisional interpretations; public and political pressure increases; institutions adopt technical, political, or diplomatic positions; and one or more accounts acquire sufficient salience to influence response. New evidence may later reinforce, qualify, or contradict earlier interpretations.

The integrated expectation is consequently bounded rather than deterministic:

Integrated proposition: The interaction of rapid narrative formation, unresolved ambiguity, and fragmented communicative authority increases the risk that political responses will be shaped before institutions establish a coordinated and evidentially calibrated account of a cyber incident.

3. METHODOLOGY

3.1. Research Design and Case Selection

The article employs a qualitative comparative case-study design combining structured, focused comparison with process-tracing-informed within-case analysis. Structured, focused comparison applies the same theoretically derived questions to each case while concentrating only on aspects relevant to the proposed explanation (George & Bennett, 2005, pp. 67–72). The within-case analysis draws on process-tracing principles by examining temporal sequences and observable manifestations expected if the proposed mechanisms operated (Beach & Pedersen, 2019). The study does not claim to conduct exhaustive process tracing of every technical, institutional, and political activity. Complete forensic data, classified intelligence, internal deliberations, and diplomatic communications are not publicly available. The design therefore evaluates mechanism plausibility and comparative explanatory value rather than making definitive causal claims.

The unit of analysis is a bounded episode of cyber crisis formation. An episode begins when a technical incident affecting a politically salient public digital function becomes publicly visible. It ends when the immediate political interpretation has become provisionally consolidated or when an authoritative institutional assessment provides sufficient closure to the initial crisis phase. The case windows are not identical in duration because evidentiary and political closure occurred at different times. The Bulgarian episode begins with the public disclosure of data in July 2019 and includes the immediate investigation, regulatory findings, and international suspension of data transmission. The North Macedonian episode begins with the election-night disruption of 15 July 2020 and extends through the announcement of preliminary results and immediate institutional follow-up; the later ODIHR report is used as retrospective evidence. The Montenegrin episode begins with the August 2022 disruption and extends to the March 2023 National Security Council assessment that attribution had not been established.

Bulgaria, North Macedonia, and Montenegro were selected as information-rich, politically consequential cases within a bounded regional setting. Each involved a publicly visible incident affecting a State institution or politically significant public digital function, and in each case public interpretation began before authoritative knowledge had stabilised. The cases vary in technical form and political outcome: a large-scale confidentiality breach in Bulgaria, an election-related availability disruption in North Macedonia, and a broader governmental cyber crisis with ransomware, foreign assistance, and attribution volatility in Montenegro. This variation allows the framework to be examined across different affected functions and crisis trajectories. The cases were identified through purposive scoping of publicly documented cyber incidents affecting public-sector or politically salient digital systems in South-Eastern Europe. They were selected because each provides a reconstructable public chronology, multiple identifiable communicating actors, unresolved or evolving knowledge during the initial crisis phase, and observable political or institutional consequences. Other regional incidents were considered but not selected where the public record did not permit reliable reconstruction of the communication sequence, where the affected function was predominantly private or technical, or where the case involved a substantially higher-intensity diplomatic response requiring a separate design. The 2022 Albanian government cyber crisis, for example, involved formal attribution and diplomatic rupture, making it a different type of escalation case rather than an intermediate case of crisis formation under evolving or unresolved attribution. The outcome is political crisis expansion: movement beyond routine technical remediation into institutional-trust contestation, securitised framing, foreign consultation, diplomatic signalling, or pressure for visible governmental action. This is a positive-case design. It can assess how the mechanisms operate in politically consequential incidents, but it cannot establish whether they distinguish such incidents from technically similar events that remain contained. The propositions are therefore treated as bounded theoretical expectations rather than necessary or sufficient conditions.

3.2. Evidence and Analytical Procedure

The empirical corpus consists of a focused body of official, intergovernmental, investigative, and contemporaneous news sources. Official sources establish institutional actions, formal findings, and publicly stated assessments. Intergovernmental materials document cross-border consequences and external evaluations. Contemporary reporting reconstructs the chronology and records statements issued while the incidents were unfolding. Investigative reporting is used where it provides additional reconstruction of

domestic communication and technical disputes. Official statements are treated as authoritative evidence of what institutions communicated or decided, but not automatically as proof that their technical or attributional claims were correct. News reporting is evidence of chronology and public framing, not an independent forensic determination. Claims by suspected operators or ransomware groups are treated as claims requiring corroboration. Attribution language is preserved precisely. An official statement that an actor was suspected, linked, assessed, alleged, or publicly attributed is not rewritten as proof that the actor conducted or directed the operation. Technical similarity, operator identity, organisational affiliation, sponsorship, strategic intent, and legal State responsibility are treated as separate propositions. The within-case analysis is organised around six categories: the technical event and affected function; the moment and channel of public visibility; the state of knowledge during the initial crisis phase; the actors communicating about the incident; the operation of speed, ambiguity, and fragmentation; and the resulting political or institutional consequences.

Table 1. Analytical Categories Used in the Comparative Analysis

Analytical category	Guiding analytical question	Evidence examined	Relevance for the framework
Public visibility	When and how did the incident become publicly known?	Timing of public disclosure, service unavailability, official statements, media reporting, public alerts, or visible disruption of a public digital function.	Identifies the beginning of the crisis-formation episode and assesses the operation of speed.
Epistemic gap	What was known, inferred, disputed, or unknown during the initial phase?	Confirmed facts, preliminary estimates, unresolved technical questions, competing explanations, later corrections, and the time required for authoritative clarification.	Assesses the distance between public awareness and institutional verification.
Communicating actors	Which actors communicated about the incident and with what authority?	Statements by government institutions, technical bodies, contractors, regulators, law-enforcement actors, media, private platforms, foreign partners, and alleged operators.	Assesses fragmentation of knowledge, authority, and public communication.
Attribution and responsibility	How were cause, authorship, sponsorship, intent, and State responsibility distinguished or conflated?	Technical indicators, operator claims, criminal hypotheses, intelligence suspicions, political statements, foreign-assistance signals, and official attribution or non-attribution.	Assesses ambiguity and the risk of premature or overstated attribution.
Coordination and revision	Were institutional messages coordinated, sequenced,	Changes in official explanations, competing institutional accounts, revised assessments,	Assesses whether fragmentation was managed through

	corrected, or contradicted?	clarification of confidence levels, and use of alternative communication channels.	coordinated differentiation.
Political consequence	What form of political crisis expansion followed?	Institutional-trust effects, public contestation, electoral-transparency concerns, regulatory action, foreign assistance, diplomatic signalling, securitised framing, or pressure for governmental response.	Identifies the form and intensity of political crisis expansion.
Alternative explanations	What other factors could explain the observed outcome?	Material severity, sensitivity of the affected function, institutional capacity, political polarisation, procurement or organisational weaknesses, electoral context, and geopolitical environment.	Tests whether digital-reality mechanisms add explanatory value beyond technical or contextual factors.

Source: Author's analysis.

The analysis also considers evidence inconsistent with the framework and alternative explanations including material severity, institutional capacity, political polarisation, electoral sensitivity, and geopolitical context. Evidence for speed is strongest where consequential claims circulated before a coordinated institutional assessment and where early framing influenced subsequent communication or response. Evidence for ambiguity is strongest where several plausible explanations remained open and supported different political interpretations. Evidence for fragmentation is strongest where differences among communicators affected credibility, institutional coordination, or diplomatic interpretation. The mere presence of multiple institutions is not sufficient.

3.3. Limitations and Ethics

Publicly available sources provide uneven access to technical and institutional evidence. The study does not independently determine who was responsible for the incidents; it examines how responsibility, severity, and appropriate response were publicly interpreted and contested. Source availability differs across cases, so the number of statements or reports is not treated as a direct measure of mechanism intensity. The mechanisms may coexist with material damage, organisational weakness, political opportunism, and geopolitical context. The purpose is therefore to assess whether the interpretive mechanisms provide additional explanatory value, not to replace material explanations. The author's prior development of the digital-reality concept creates a potential risk of confirmation bias, mitigated by specifying mechanisms before comparison, retaining contradictory observations, distinguishing source claims from established facts, considering alternatives, and limiting conclusions to bounded plausibility. The research uses no interviews, experiments, covert observation, or non-public personal data.

4. CASE STUDIES

4.1. Bulgaria 2019: Institutional and Cross-Border Trust

The 2019 breach of Bulgaria's National Revenue Agency (NRA) began as a compromise of a central government information system but developed into a wider crisis of institutional accountability and cross-border trust. A substantial quantity of information obtained from NRA systems was supplied to Bulgarian media on 15 July 2019. The agency subsequently confirmed that the disclosed material originated from its databases. Bulgaria's Commission for Personal Data Protection reported that more than five million citizens were affected and that the compromised information included personal, fiscal, social-security, and health-insurance data (Commission for Personal Data Protection [CPDP], 2019a, 2019b). Contemporaneous reporting indicated that the intrusion had occurred before the public disclosure. The public phase was therefore initiated by an external actor rather than through planned institutional notification. On 16 July, an NRA spokesperson stated that the agency had examined part of the disclosed material and confirmed its authenticity. Officials simultaneously issued preliminary estimates concerning the scale of compromise while investigations continued (Krasimirov & Tsoleva, 2019). This sequence provides strong evidence of speed. The person claiming responsibility provided journalists with both compromised material and an interpretation of governmental cybersecurity failure before the State had established a comprehensive account of the affected systems, exposed data, intrusion method, perpetrator, or risks to citizens. Institutional communication was reactive from the beginning. Authorities had to authenticate the material, reassure citizens, investigate criminal responsibility, assess risks to other systems, and demonstrate governmental control while technical verification remained incomplete. Early frames were consequently established before the State could provide a coordinated explanation. Ambiguity concerned the scale of compromise, the complete categories of exposed data, the method of access, the perpetrator's identity, motive, and relationship to any foreign actor. The person communicating with the media used a Russian-based email service and reportedly claimed to be Russian. Officials initially referred to indications that the operation may have originated outside Bulgaria. These facts did not establish Russian State involvement or even conclusively establish the identity of the person communicating with journalists.

The attributional picture shifted quickly. On 17 July, Bulgarian authorities announced that a 20-year-old Bulgarian cybersecurity employee had been charged in connection with the breach. Investigators stated that the inquiry was ongoing and that the possible involvement of other persons was being examined. The suspect denied responsibility through his lawyer (Tsoleva & Krasimirov, 2019).

The transition from foreign-origin indications and a self-described Russian perpetrator to a domestic suspect illustrates the difference among infrastructure location, claimed identity, operational involvement, motive, and political responsibility. The evidence did not support a definitive claim of foreign State involvement. Several interpretations remained plausible during the initial phase: criminal unauthorised access, politically motivated disclosure, an attempt to embarrass the government, or evidence of systemic weaknesses in public-sector cybersecurity. These explanations directed responsibility towards different actors, from the suspected perpetrator to the institution legally responsible for protecting compulsory citizen data. Fragmentation was present but more limited than speed and ambiguity. Communicative authority was dispersed among the NRA, political authorities, investigators, the data-protection authority, cybersecurity specialists, journalists, and the person claiming responsibility. The public problem was not

simply the number of communicators, but the absence of an early account integrating technical, criminal, regulatory, and political dimensions while distinguishing verified information from preliminary estimates. Later regulatory findings provided a more authoritative account of institutional responsibility. CPDP concluded that the NRA had failed to implement appropriate technical and organisational measures, resulting in unauthorised access to and disclosure of personal data. The Commission issued a penal order relating to a violation of Article 32(1)(b) of the General Data Protection Regulation (CPDP, 2019b). These findings did not resolve every question concerning criminal responsibility, but they established that the crisis could not be interpreted solely as an external attack on an otherwise adequately protected institution. Responsibility had to be distinguished between the actor who obtained or disclosed the information and the institution legally responsible for protecting it. The clearest cross-border consequence involved international tax-information exchange. The OECD Global Forum confirmed that the compromised NRA systems contained information received from international partners through automatic exchange arrangements. It suspended jurisdictions' obligations to transmit relevant information to Bulgaria until a satisfactory review had been completed and deficiencies addressed. The OECD clarified that the Common Transmission System itself had not been compromised; the problem concerned the security of information after it entered the Bulgarian institutional environment (Organisation for Economic Co-operation and Development [OECD], 2019). This transformed the breach into a cross-border institutional trust problem. The consequence was not interstate retaliation or attribution against another State. It was the temporary suspension of the confidence required for continued transmission of sensitive information to a national authority. Material severity remains a powerful alternative explanation. The breach affected a central revenue authority, exposed information concerning millions of citizens, and included data received through international cooperation. Regulatory penalties and suspension of data exchange could have occurred even under conditions of perfectly coordinated communication. The mechanisms do not independently explain these consequences. They explain how the event's meaning and allocation of responsibility developed while evidence remained incomplete.

The Bulgarian case therefore provides strong support for speed, substantial support for ambiguity, and moderate support for fragmentation. It is best understood as an institutional and cross-border trust crisis. The central political question expanded from which records had been compromised to whether the institution legally empowered to collect sensitive information could credibly protect it.

4.2. North Macedonia 2020: Contested Informational Authority

North Macedonia held early parliamentary elections on 15 July 2020 after the process had been suspended and rescheduled because of the COVID-19 pandemic. ODIHR assessed the elections as genuinely competitive and election day as generally orderly, but later concluded that technical challenges in publishing results partially undermined public confidence in the election administration (Organization for Security and Co-operation in Europe, Office for Democratic Institutions and Human Rights [ODIHR], 2020, pp. 1–2). The State Election Commission (SEC) introduced the web-based uVote application for electronic transmission of turnout data and preliminary results.

The application also provided a public interface through which citizens, parties, and media organisations could follow partial turnout and preliminary results. The contract for its development was awarded to the local company DUNA on 29 June, slightly more than two weeks before election day (ODIHR, 2020, p. 28).

ODIHR reported that no large-scale simulation of uVote had been conducted, SEC supervision of its development was limited, and an operability test took place only two days before the elections. No report on that test was presented to the Commission before election day. Election administrators received written instructions but no additional training. Temporary technical problems also occurred during turnout reporting and sometimes required information to be transmitted by telephone or text message (ODIHR, 2020, p. 28). Shortly after polling stations closed and preliminary data began to be entered, the public application stopped displaying results. The SEC president reported a cyberattack which, according to the Commission, affected publication but not counting or tabulation. The SEC attempted to compensate by live-streaming information from its results database through YouTube (ODIHR, 2020, pp. 28–29). The distinction between publication and electoral integrity is fundamental. The publicly available evidence does not establish that ballots, counting, or tabulation were altered. The principal demonstrated effect concerned the availability and public communication of preliminary results. Nevertheless, the timing made the incident politically consequential. It occurred when citizens, political parties, and media organisations expected authoritative information about the emerging outcome. The speed mechanism is strongly observable. The Investigative Reporting Laboratory reported that the SEC website became unavailable at approximately 21:10, while journalists and political actors were waiting for results. The SEC and DUNA then sought to determine whether the problem concerned the results application or the Commission's wider website and server environment (Cvetkovska & Ilieski, 2020). At 23:17, Meta.mk reported that the SEC president could not yet identify the reason for the website's unavailability. He stated that results were arriving locally but could not be displayed publicly. The same report stated that Time.mk, a major news aggregator displaying election information, was experiencing what its owner described as a large distributed denial-of-service attack (Blazhevski, 2020). The simultaneous disruption of the official interface and a private information platform intensified the information gap. Their temporal concurrence did not establish common authorship or operational coordination. The available evidence does not demonstrate that the same actor conducted both disruptions. Their combined significance was informational: two prominent channels through which the public expected to follow the election were impaired during the same sensitive interval. Citizens and political actors could observe the failure but could not independently determine its cause. Plausible explanations included an external attack, inadequate server capacity, application failure, insufficient testing, weak supervision, or an interaction among these factors. Public interpretation therefore developed faster than authoritative technical clarification. Ambiguity was the dominant mechanism. It concerned not only the identity of a possible attacker but the technical object and causal structure of the disruption. DUNA stated that its responsibility covered application performance but not security and maintenance of the SEC server. SEC information-technology personnel reportedly stated that an attack against the main website had been addressed and that the remaining delivery problem should be sought in the application. Several ODIHR interlocutors considered that the difficulties may have resulted from a combination of external attacks and insufficient SEC information-technology capacity (ODIHR, 2020, p. 29). These explanations were not necessarily mutually exclusive. External malicious traffic may have occurred while deficiencies in preparation,

system capacity, or institutional supervision increased its effect. The evidence did not support premature closure around one cause. Public attribution also remained unresolved. On 22 July, a group calling itself AnonOpsMKD claimed responsibility for cyber activity but denied targeting the SEC. The claim did not provide independently verified evidence establishing who caused the SEC disruption or what political effect was intended (ODIHR, 2020, p. 29). The public record consequently does not justify attribution to a foreign State, political party, hacktivist group, or other specific actor. Ambiguity also affected interpretation of electoral integrity. The absence of the public interface created a visible appearance of vulnerability, but it did not demonstrate vote manipulation. The SEC cross-checked hard copies of final polling-station protocols against the electronically uploaded database. ODIHR observed this verification and assessed it as organised and professional. Preliminary results were officially announced at 19:30 on 16 July, approximately 20 hours later than expected. A later minor correction contributed to additional speculation (ODIHR, 2020, p. 29). Fragmentation was moderate. Relevant information and communicative authority were distributed among SEC leadership, SEC technical personnel, DUNA, municipal election commissions, political parties, journalists, cybersecurity specialists, and private information platforms. The central problem was the absence of an integrated early explanation distinguishing the public website, the uVote application, the server environment, the internal database, the continuation of counting, and the evidentiary status of cyberattack claims.

The SEC retained its formal legal authority to administer the elections, receive protocols, determine results, and address complaints. What it temporarily lost was reliable control over the principal public-facing interface through which that authority was expected to become visible. The incident therefore represented a displacement of informational authority rather than a loss of procedural electoral authority. Follow-up proceedings widened the crisis beyond the immediate outage. On 28 July, the Prosecutor's Office for Organized Crime entered the SEC to investigate possible criminal conduct connected to procurement. The personal-data protection authority also initiated extraordinary supervision and identified breaches of relevant regulations (ODIHR, 2020, p. 29). These actions do not prove that procurement or data-protection failures caused the disruption. They show that the incident exposed wider concerns concerning preparation, accountability, and institutional capacity. ODIHR recommended a comprehensive cybersecurity strategy including vulnerability assessment, mitigation measures, adequate software and hardware, security safeguards, and permanently employed and trained information-technology personnel (ODIHR, 2020, p. 31). Material and organisational explanations are indispensable. Limited testing, constrained supervision, insufficient technical capacity, and the sensitivity of election-night reporting provide strong explanations for the severity and political relevance of the failure. The digital-reality framework does not replace them. It explains how visible system failure and competing accounts generated political meaning before a stable explanation existed.

The case provides strong support for speed and ambiguity and moderate support for fragmentation. Its political consequence remained predominantly domestic and institutionally bounded. It did not produce official foreign-State attribution or interstate confrontation. It is therefore best understood as a crisis of contested informational authority rather than as a demonstrated compromise of electoral integrity.

4.3. Montenegro 2022: Attribution Volatility and Internationalised Response

Montenegro experienced a sustained cyber crisis in August 2022 that affected government information infrastructure, disrupted access to public digital services, and prompted assistance from international partners. On 26 August, officials described the incident as unprecedented and reported that some services had been taken offline as a protective measure. At that stage, the minister of public administration stated that citizen and business accounts and data had not been endangered, while the government informed NATO allies (Reuters, 2022a). Subsequent official information established broader operational effects. The Government of Montenegro reported that State information infrastructure and the information-communication networks of public institutions had been exposed to continuous attacks of substantial intensity and complexity from 20 August. Detected activities included ransomware, distributed denial-of-service attacks, and botnet operations. Seventeen information systems across ten institutions were encrypted, and approximately 150 computers were directly affected. Priority services were restored progressively (Government of Montenegro, 2022). The crisis unfolded during domestic political instability and in a geopolitical environment shaped by Russia's invasion of Ukraine, Montenegro's NATO membership, alignment with European measures against Russia, and established domestic debate over Russian influence. These conditions did not establish authorship or motive, but they made geopolitical explanations readily available. Speed operated through rapid elevation of the disruption into a national-security issue. The National Security Council was convened, the National Security Agency described the situation in terms of hybrid warfare, and the defence minister identified Russia as the principal suspect. The United States Embassy issued a warning concerning a persistent cyberattack and the possibility of disruption to public services and infrastructure (Euractiv with AFP, 2022; Reuters, 2022a). These statements shaped interpretation before a coordinated forensic assessment had been completed. The scale of disruption, precautionary shutdowns, and language of hybrid warfare created an immediate national-security frame. Montenegro's foreign-policy position provided a plausible motive for retaliation. A plausible motive, however, did not constitute evidence that Russian State institutions had conducted or directed the operation. The response was rapidly internationalised. Montenegro requested assistance from partners. The Ministry of the Interior announced on 31 August that an FBI Cyber Action Team would assist the investigation. France's National Agency for the Security of Information Systems also provided support, while the United States, United Kingdom, France, and other partners contributed to investigation, recovery, or institutional assistance (Ministry of the Interior of Montenegro, 2022; Ministry of Public Administration of Montenegro, 2022). International assistance indicated the severity of the incident and Montenegro's dependence on external forensic capacity. It did not independently validate attribution to a foreign State.

Ambiguity concerned at least four separate questions: which actor conducted the technical activity; which tools and methods were used; whether the operation was criminal, political, or strategically directed; and whether a State sponsored, tolerated, or benefited from it. Public communication did not consistently distinguish these questions. Reuters reported that Montenegro identified the group known as Cuba ransomware as responsible for attacks against government infrastructure. The minister of public administration stated that approximately 150 workstations across ten institutions had been infected. At the same time, Reuters reported that the National Security Agency linked the attacks to Russia. Cuba ransomware also claimed to have obtained parliamentary documents, while parliament

denied that confidential data had been stolen and stated that the referenced material was publicly available (Reuters, 2022b). These claims created an unstable attributional baseline. Identification of ransomware or an associated criminal group did not establish whether it operated independently, acted for another actor, or exploited a politically favourable opportunity. Intelligence suspicion of Russia did not establish the technical operator or degree of State control. The public record did not clarify whether the criminal and geopolitical explanations referred to the same chain of responsibility or to different analytical levels. Fragmentation was strongest in the attribution process. Intelligence officials, political leaders, technical authorities, parliament, foreign teams, and media organisations communicated from different institutional positions and addressed different levels of responsibility. These statements were not necessarily contradictory: a criminal group could conduct an operation while a State remained relevant as sponsor, facilitator, beneficiary, or contextual adversary. The communication problem was that these levels were not consistently differentiated. Public reporting compressed distinct propositions into competing labels: Russia, Cuba ransomware, criminal attack, or hybrid warfare. The case therefore shows why operator attribution, sponsor attribution, political attribution, and legal State responsibility must remain analytically separate. The later official assessment demonstrates the persistence of ambiguity. On 8 March 2023, Montenegro's National Security Council considered reports prepared by domestic authorities, the FBI, and ANSSI. It concluded that, because of the specificity and complexity of the attack, attribution had not been established (Government of Montenegro, National Security Council, 2023). The crisis therefore passed through three public stages: geopolitical suspicion, criminal-technical identification, and official non-attribution. This volatility does not prove that early assessments were fabricated or politically opportunistic. It shows that the public evidence supported claims with different political implications and that their relationship was insufficiently explained. The case generated an internationalised security response without definitive interstate attribution. Montenegro consulted allies and received FBI and ANSSI assistance, but it did not publicly establish that the Russian State conducted or directed the operation. Nor did the incident culminate in documented retaliation, invocation of NATO collective-defence provisions, or a formal diplomatic measure based on conclusive attribution. Material severity and capacity limitations provide important alternative explanations. Broad governmental disruption justified national-security attention, while limited domestic forensic capacity explains the request for foreign assistance. Geopolitical framing was also facilitated by the wider security environment. The mechanisms explain not why the attack was operationally serious, but how distinct evidentiary layers became compressed into unstable accounts of responsibility. The case provides strong support for all three mechanisms. National-security framing preceded forensic closure; several attributional levels remained unresolved; and institutional assessments were not consistently connected. Montenegro consequently demonstrates how attribution volatility can internationalise a cyber crisis without producing definitive State responsibility.

5. COMPARATIVE ANALYSIS

The three cases differ in technical form, affected function, evidentiary conditions, and political outcome. Bulgaria involved the compromise and disclosure of confidential personal and fiscal data. North Macedonia involved the unavailability of a public election-results interface. Montenegro involved broad governmental disruption, ransomware activity, foreign assistance, and unresolved attribution. Despite these differences, each case displayed a recurring sequence: a technical incident became visible before authoritative knowledge stabilised; multiple actors interpreted the event under uncertainty; and one or more accounts acquired sufficient political salience to influence institutional behaviour, public confidence, or international response.

Table 2. Comparative Operation of Digital-Reality Mechanisms

Case	Affected public function	Mechanism configuration	Principal political outcome
Bulgaria, 2019	Confidentiality of compulsory fiscal and personal data; institutional custodianship	Speed: strong; ambiguity: substantial; fragmentation: moderate	Regulatory accountability, erosion of custodial trust, and temporary suspension of international tax-information transmission
North Macedonia, 2020	Availability of the official preliminary election-results interface; public visibility of electoral administration	Speed: strong; ambiguity: strong; fragmentation: moderate	Temporary displacement of informational authority, delayed results, and scrutiny of institutional capacity without established alteration of counting or tabulation
Montenegro, 2022	Continuity of government digital systems and national cyber resilience	Speed: strong; ambiguity: strong; fragmentation: strong	Internationalised security response, allied assistance, and attribution volatility without definitive public attribution of State responsibility

Source: Author's analysis.

Speed was strongly observable in all three cases, but through different initiating conditions. In Bulgaria, externally initiated disclosure deprived the State of control over the timing and framing of public awareness. In North Macedonia, visible unavailability of the official results interface created immediate demand for explanation. In Montenegro, broad disruption entered an already securitised geopolitical context and was rapidly interpreted as a national-security threat. The politically relevant dimension of speed was therefore not elapsed time alone, but the gap between the rate of public interpretation and the rate of institutional verification.

Ambiguity also varied across cases. In Bulgaria, it concerned the scale of compromise, perpetrator identity, motive, and division of responsibility between attacker and institution. In North Macedonia, it concerned the relationship among the website, application, server environment, preparation, and possible malicious activity.

In Montenegro, it concerned operator identity, criminal affiliation, strategic sponsorship, geopolitical motive, and legal State responsibility. The evidence supports the proposition that ambiguity expands interpretive opportunity, but its effects are conditional.

It can enable premature escalation, but it can also preserve space for restraint when uncertainty is communicated carefully.

Fragmentation was weakest in Bulgaria, moderate in North Macedonia, and strongest in Montenegro. The comparison confirms that institutional plurality is not itself the problem. Technical, intelligence, law-enforcement, political, and diplomatic institutions legitimately answer different questions. Dysfunctional fragmentation arises when their conclusions are publicly combined without explanation of their objects, evidence, confidence levels, and institutional authority.

The mechanisms reinforced one another. Speed amplified ambiguity because institutions had to communicate before evidence was complete. Ambiguity amplified fragmentation because several actors could adopt individually plausible accounts. Fragmentation reinforced speed because multiple communicators issued claims before a common position had formed. The strongest interaction occurred in Montenegro, while North Macedonia shows that the mechanisms can produce political consequences even where the demonstrated operational effect primarily concerns availability and communication.

Technical severity was neither irrelevant nor sufficient. Bulgaria's breach was materially severe. North Macedonia's technical effect was narrower, but its electoral timing made it politically salient. Montenegro's operational impact was broad, while its international meaning was intensified by attribution volatility and geopolitical context. The outcome was therefore not uniform escalation. Bulgaria produced an institutional and cross-border trust consequence. North Macedonia produced bounded domestic politicisation and contested informational authority. Montenegro produced internationalised securitisation and allied engagement.

Several scope conditions emerge. The mechanisms appear more consequential when the affected system performs a compulsory, highly visible, or symbolically important State function; when the incident becomes public before an authoritative explanation exists; when established narratives of institutional failure or foreign interference are available; when technical and political authorities lack tested communication procedures; and when foreign partners become involved before domestic institutions stabilise a common assessment. The cross-case findings provide strong support for Proposition 1, strong but conditional support for Proposition 2, and moderate-to-strong support for Proposition 3. The integrated proposition is supported as a bounded claim: rapid public interpretation, unresolved ambiguity, and fragmented communicative authority increase the risk that political responses will be shaped before authoritative evidence stabilises.

6. DISCUSSION AND POLICY IMPLICATIONS

6.1. Theoretical Implications

The findings position communication as part of cyber crisis formation rather than as a final stage following technical investigation. In a digitally mediated environment, interpretation begins when disruption becomes visible, while investigation and communication proceed simultaneously. This does not justify speculative attribution. It means that evidentiary caution must itself be communicated. Silence may be technically prudent, but it can be interpreted as incompetence, concealment, internal disagreement, or loss of control. The relevant principle is neither immediate attribution nor prolonged silence, but calibrated early communication. Institutions should state which systems are affected, what remains functional, which facts have been verified, which hypotheses are being examined, and when the next update will follow. Public communication should

distinguish verified facts, probabilistic assessments, competing hypotheses, and unresolved questions. This protects credibility because later revision can be presented as a normal consequence of investigation rather than as contradiction or political retreat. Digital reality also highlights an epistemic dimension of resilience. States must preserve the capacity to make reliable knowledge publicly visible, explain uncertainty, and correct earlier assessments while competing accounts circulate. This does not imply that the State should monopolise interpretation. Democratic information environments are plural and contested. Epistemic resilience concerns the ability of public institutions to remain credible within that pluralism. The North Macedonian SEC retained legal authority over the electoral process but temporarily lost the public interface through which that authority was expected to become visible. The Bulgarian NRA retained administrative authority but lost trust as a custodian of compulsory data. Montenegro retained sovereign decision-making authority but struggled to connect distinct attributional levels in public communication. The findings consequently qualify the concept of hybrid threats. The evidence does not show that every incident and surrounding narrative formed part of a centrally coordinated campaign. Nevertheless, technical disruption can produce hybrid-like political effects when it interacts with established geopolitical narratives, domestic polarisation, and digitally mediated contestation. A criminal or opportunistic incident may therefore generate effects that resemble strategic hybrid pressure without evidence that a single actor designed the full process.

6.2. Governance and Policy Implications

Technical cyber response is usually allocated through formal responsibilities for detection, containment, investigation, recovery, and notification. The translation of technical uncertainty into political communication, attribution, and diplomatic signalling is less consistently institutionalised. This creates a governance asymmetry: a State may restore systems while communicating incoherently about what happened, who may be responsible, and how confidence in its assessments has changed. The appropriate solution is coordinated differentiation: institutions should retain specialised roles while explaining how technical findings, intelligence assessments, legal evaluations, political judgments, and diplomatic positions relate to one another. Three institutional implications follow: First, governments should establish a standing interagency cyber-crisis coordination function. It need not constitute a large new bureaucracy. It may operate as a permanent protocol and an activatable interinstitutional cell linking incident-response authorities, intelligence services, law enforcement, legal advisers, political leadership, foreign affairs, sectoral regulators, and public communicators. Its central task should be translation: converting incomplete technical findings into calibrated domestic and international messages without collapsing suspicion, attribution, and verified responsibility.

Second, governments should adopt communication protocols for attributional uncertainty. These should establish lead spokespersons, procedures for interagency coordination, common confidence terminology, thresholds for notifying foreign partners, and rules for revising previous assessments. Public statements should explicitly distinguish what is known, assessed, disputed, and unresolved.

Third, national and regional exercises should test political and diplomatic communication alongside technical incident response. Scenarios should include media pressure, conflicting preliminary evidence, leaks, private-sector disclosures, suspected foreign involvement, requests for assistance, and demands for immediate attribution. Exercises should test not only whether contact channels work, but whether institutions can

formulate confidence-calibrated messages under pressure. International frameworks provide a useful foundation. UN norms emphasise consultation, assistance, responsible behaviour, and fact-based assessment, while OSCE confidence-building measures provide points of contact and communication mechanisms (Dominioni, 2023; OSCE, 2016; United Nations, 2021). The European Union's NIS2 Directive and North Macedonia's 2025 Law on Security of Network and Information Systems strengthen strategies, competent authorities, CSIRTs, single points of contact, crisis-management arrangements, reporting duties, and public information functions (European Parliament & Council of the European Union, 2022; Republic of North Macedonia, 2025). Their remaining limitation is not the absence of communication provisions, but the lack of a detailed protocol synchronising technical assessment, intelligence evaluation, public communication, political attribution, and foreign-policy engagement when evidence remains incomplete. A national cyber-crisis communication protocol should therefore specify lead communication authority, interinstitutional roles, confidence terminology, procedures for revising assessments, and rules for consultation with foreign partners.

These recommendations are theory-informed rather than statistically generalised. Their value lies in addressing the recurring problem observed across the cases: informational delay and poorly connected authority can expand the political consequences of a cyber incident beyond its original technical effects.

7. CONCLUSION

Cyber incidents affecting politically salient public digital systems cannot be understood through technical severity alone. Their political consequences also depend on how disruption becomes visible, interpreted, attributed, and translated into institutional or international action while authoritative knowledge remains incomplete. This article has defined digital reality as the temporally bounded, digitally mediated interpretive process through which incomplete information about a cyber incident becomes publicly meaningful and politically consequential before evidence stabilises. The comparison of Bulgaria, North Macedonia, and Montenegro identified three mutually reinforcing mechanisms. Speed created a temporal gap between public interpretation and institutional verification. Ambiguity sustained several plausible accounts of cause, authorship, and intent. Fragmentation dispersed knowledge and communicative authority among actors operating with different mandates and evidentiary thresholds. The mechanisms contributed to different forms of political crisis expansion. Bulgaria developed into a crisis of institutional accountability and cross-border trust. North Macedonia experienced a temporary displacement of informational authority without established compromise of counting or tabulation. Montenegro generated an internationalised security response and attribution volatility without definitive public attribution of State responsibility.

The findings do not establish that communication independently caused these outcomes or that the mechanisms constitute necessary or sufficient conditions. Material damage, institutional function, political context, and organisational capacity remained essential. The narrower contribution is to identify the process through which material disruption was translated into wider political meaning.

Cyber crisis communication is consequently not a public-relations supplement to technical response. It is a security function connecting incident management, intelligence assessment, legal responsibility, political leadership, public trust, and cyber diplomacy. States that restore systems but fail to communicate uncertainty coherently remain vulnerable to consequences that exceed the original incident. Cyber resilience therefore has both an operational and an interpretive dimension. Institutionalising calibrated communication under uncertainty is not ancillary to cybersecurity; it is one of its essential contemporary conditions.

DECLARATION OF INTEREST

The author declares no conflict of interest.

REFERENCES

- Beach, D., & Pedersen, R. B. (2019). *Process-tracing methods: Foundations and guidelines* (2nd ed.). University of Michigan Press. <https://doi.org/10.3998/mpub.10072208>
- Berger, P. L., & Luckmann, T. (1966). *The social construction of reality: A treatise in the sociology of knowledge*. Anchor Books.
- Blazhevski, B. (2020, July 15). Падна сајтот на ДИК за изборните резултати, агрегаторот Time.mk под сајбер напад [The SEC election-results website goes down, while the Time.mk aggregator is under cyberattack]. *Meta.mk*. <https://meta.mk/padna-sajtot-na-dik-za-izbornite-rezultati-agregatorot-time-mk-pod-sajber-napad/>
- Borghard, E. D., & Lonergan, S. W. (2017). The logic of coercion in cyberspace. *Security Studies*, 26(3), 452–481. <https://doi.org/10.1080/09636412.2017.1306396>
- Buchanan, B. (2017). *The cybersecurity dilemma: Hacking, trust, and fear between nations*. Oxford University Press.
- Commission for Personal Data Protection. (2019a). *Information about CPDP's actions in relation to the personal data protection security breach at the National Revenue Agency*. <https://cpdp.bg/en/information-about-cpdps-actions-in-relation-to-the-personal-data-protection-security-breach-at-the-national-revenue-agency/>
- Commission for Personal Data Protection. (2019b). *Update on the undertaken inspection at the National Revenue Agency*. <https://cpdp.bg/en/update-on-the-undertaken-inspection-at-the-national-revenue-agency/>
- Coombs, W. T. (2007). Protecting organisation reputations during a crisis: The development and application of situational crisis communication theory. *Corporate Reputation Review*, 10(3), 163–176. <https://doi.org/10.1057/palgrave.crr.1550049>
- Cvetkovska, S., & Ilieski, D. (2020, July 18). Хибридни војни: Кој дозволи да се „хакне“ изборниот ден? [Hybrid wars: Who allowed election day to be “hacked”?]. *Investigative Reporting Laboratory*. <https://irl.mk/khibridni-voni-ko-dozvoli-da-se-khakne-izborniot-den/>
- Dominioni, S. (2023). *Operationalizing a directory of points of contact for cyber confidence-building measures*. United Nations Institute for Disarmament Research. <https://unidir.org/publication/operationalizing-directory-points-contact-cyber-confidence-building-measures/>
- Egloff, F. J., & Smeets, M. (2023). Publicly attributing cyber attacks: A framework. *Journal of Strategic Studies*, 46(3), 502–533. <https://doi.org/10.1080/01402390.2021.1895117>
- Euractiv with AFP. (2022, August 28). Cyberattack hits Montenegro government, defence minister points at Russia. *Euractiv*. <https://www.euractiv.com/news/cyberattack-hits->

montenegro-government-defence-minister-points-at-russia/

European Commission. (2016). *Joint framework on countering hybrid threats: A European Union response* (JOIN(2016) 18 final). <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52016JC0018>

European Parliament and Council of the European Union. (2022). Directive (EU) 2022/2555 of 14 December 2022 on measures for a high common level of cybersecurity across the Union. *Official Journal of the European Union*, L 333, 80–152. <https://eur-lex.europa.eu/eli/dir/2022/2555/oj>

Gartzke, E. (2013). The myth of cyberwar: Bringing war in cyberspace back down to earth. *International Security*, 38(2), 41–73. https://doi.org/10.1162/ISEC_a_00136

George, A. L., & Bennett, A. (2005). *Case studies and theory development in the social sciences*. MIT Press.

Government of Montenegro. (2022, November 2). *Saopštenje sa 26. sjednice Vlade Crne Gore* [Statement from the 26th session of the Government of Montenegro]. <https://www.gov.me/cyr/clanak/saopstenje-sa-26-sjednice-vlade-crne-gore>

Government of Montenegro, National Security Council. (2023, March 8). *Održana peta sjednica Vijeća za nacionalnu bezbjednost* [Fifth session of the National Security Council held]. <https://www.gov.me/clanak/odrzana-peta-sjednica-vijeca-za-nacionalnu-bezbjednost>

Kello, L. (2013). The meaning of the cyber revolution: Perils to theory and statecraft. *International Security*, 38(2), 7–40. https://doi.org/10.1162/ISEC_a_00138

Krasimirov, A., & Tsoleva, T. (2019, July 16). Hackers hit Bulgaria, send data from Russian email: Government. *Reuters*. <https://www.reuters.com/article/us-bulgaria-cybersecurity/hackers-hit-bulgaria-send-data-from-russian-email-government-idUSKCN1UB0MA>

Lee, H. (2023). Public attribution in the US government: Implications for diplomacy and norms in cyberspace. *Policy Design and Practice*, 6(2), 198–216. <https://doi.org/10.1080/25741292.2023.2199964>

Lin, H. (2012). Escalation dynamics and conflict termination in cyberspace. *Strategic Studies Quarterly*, 6(3), 46–70.

Maschmeyer, L. (2021). The subversive trilemma: Why cyber operations fall short of expectations. *International Security*, 46(2), 51–90. https://doi.org/10.1162/isec_a_00418

Ministry of the Interior of Montenegro. (2022, August 31). *Na zahtjev Adžića i Brđanina FBI dolazi u Crnu Goru* [At the request of Adžić and Brđanin, the FBI is coming to Montenegro]. <https://www.gov.me/clanak/na-zahtjev-adzica-i-brdanina-fbi-dolazi-u-crnu-goru>

Ministry of Public Administration of Montenegro. (2022, September 22). *Ministar Dukaj na Digitalnom samitu za Zapadni Balkan–Pristina 2022: Digitalni razvoj je šansa za dugoročni uspjeh* [Minister Dukaj at the Western Balkans Digital Summit–Pristina 2022: Digital development is an opportunity for long-term success]. <https://www.gov.me/clanak/ministar-dukaj-na-digitalnom-samitu-za-zapadni-balkan-pristina-2022-digitalni-razvoj-je-sansa-za-dugorocni-uspjeh>

Miskimmon, A., O’Loughlin, B., & Roselle, L. (2013). *Strategic narratives: Communication power and the new world order*. Routledge.

Organisation for Economic Co-operation and Development. (2019, August 30). *Statement on the data breach in the National Revenue Agency of Bulgaria*. <https://www.oecd.org/en/networks/global-forum-tax-transparency/news-events/2019/statement-on-the-data-breach-in-the-national-revenue-agency-of->

[bulgaria.html](#)

Organization for Security and Co-operation in Europe. (2016). *OSCE confidence-building measures to reduce the risks of conflict stemming from the use of information and communication technologies* (Permanent Council Decision No. 1202). <https://www.osce.org/files/f/documents/d/a/227281.pdf>

Organization for Security and Co-operation in Europe, Office for Democratic Institutions and Human Rights. (2020, October 2). *Republic of North Macedonia: Early parliamentary elections, 15 July 2020: ODIHR Special Election Assessment Mission final report*. https://odihr.osce.org/sites/default/files/f/documents/b/e/465648_2.pdf

Republic of North Macedonia. (2025). Закон за безбедност на мрежни и информациски системи [Law on security of network and information systems]. *Official Gazette of the Republic of North Macedonia*, 135/2025.

Reuters. (2022a, August 26). Montenegro's state infrastructure hit by cyber attack—Officials. <https://www.reuters.com/world/europe/montenegros-state-infrastructure-hit-by-cyber-attack-officials-2022-08-26/>

Reuters. (2022b, August 31). Montenegro blames criminal gang for cyber attacks on government. <https://www.reuters.com/world/europe/montenegro-blames-criminal-gang-cyber-attacks-government-2022-08-31/>

Reynolds, B., & Seeger, M. W. (2005). Crisis and emergency risk communication as an integrative model. *Journal of Health Communication*, 10(1), 43–55. <https://doi.org/10.1080/10810730590904571>

Rid, T., & Buchanan, B. (2015). Attributing cyber attacks. *Journal of Strategic Studies*, 38(1–2), 4–37. <https://doi.org/10.1080/01402390.2014.977382>

Ristov, I. (2020). *Турската моќ на Балканот и безбедносните импликации – случај Северна Македонија* [Turkish power towards the Balkans and the security implications: Case study North Macedonia] [Doctoral dissertation, University “St. Kliment Ohridski”–Bitola]. UKLO Repository. <https://eprints.uklo.edu.mk/id/eprint/8994/>

Tsolova, T., & Krasimirov, A. (2019, July 17). “Wizard” cybersecurity expert charged with record hack of Bulgarian tax agency. *Reuters*. <https://www.reuters.com/article/us-bulgaria-cybersecurity/wizard-cybersecurity-expert-charged-with-record-hack-of-bulgarian-tax-agency-idUSKCN1UC0GF>

United Nations. (2015). *Report of the Group of Governmental Experts on developments in the field of information and telecommunications in the context of international security* (A/70/174). <https://undocs.org/A/70/174>

United Nations. (2021). *Report of the Group of Governmental Experts on advancing responsible State behaviour in cyberspace in the context of international security* (A/76/135). <https://undocs.org/A/76/135>

INFORMATION SECURITY AND CHALLENGES IN PUBLIC INSTITUTIONS OF THE REPUBLIC OF NORTH MACEDONIA

Pullumb Abdullai
pabdulai@live.com
ORCID: 0000-0003-2675-2545

Abstract

Information security is an essential prerequisite for the stable and reliable functioning of public institutions in an increasingly digital environment. This study examines information-security challenges in public institutions in the Republic of North Macedonia, with particular attention to the human factor, employee security practices, cyber threats, and the national legal and institutional framework.

An exploratory mixed-methods design was employed. The qualitative component consisted of a descriptive and comparative analysis of scientific literature, international standards, national legislation, government strategies, official reports, and documented cyber incidents. The quantitative component comprised an anonymous electronic survey of 40 administrative and information-technology employees working in public institutions. Survey data were analyzed descriptively using frequencies and percentages. The findings indicated substantial gaps between general awareness of cyber threats and the practical implementation of security measures, particularly regarding password management, multi-factor authentication, employee training, phishing recognition, and incident-response preparedness. Because the survey used a small, non-probability sample, its findings should be interpreted as exploratory and should not be considered representative of all public institutions in the Republic of North Macedonia. The study concludes that institutional resilience requires an integrated approach combining technical safeguards, organizational policies, employee education, legal compliance, and coordinated incident response.

Keywords: *information security, cybersecurity, public institutions, human factor, cyber threats.*

INTRODUCTION

In modern digital societies, information security is a fundamental requirement for the stable and reliable functioning of public institutions. As public-sector organizations increasingly depend on digital technologies, interconnected information systems, and electronic public services, they are exposed to a wide range of evolving cyber threats (European Union Agency for Cybersecurity [ENISA], 2023; National Institute of Standards and Technology [NIST], 2018).

Public institutions process sensitive information, including administrative documents, financial records, classified information, and citizens' personal data. Consequently, cyberattacks against these institutions may compromise data confidentiality

and integrity, interrupt essential public services, cause financial and operational losses, and reduce public trust in government institutions (ENISA, 2023).

Although technical safeguards are an essential component of information security, they are insufficient when they are not supported by effective organizational policies, clearly assigned responsibilities, employee awareness, and continuous training. Human behavior and organizational practices may either create security vulnerabilities or provide an important first line of defense against cyber threats (Anderson, 2020; Whitman & Mattord, 2018).

The Republic of North Macedonia has recently strengthened its national cybersecurity framework through the Cybersecurity Strategy 2025–2028 and its accompanying action plan, the Law on Security of Network and Information Systems, minimum cybersecurity standards, and the development of dedicated national and government incident-response structures (Law on Security of Network and Information Systems, 2025; Ministry of Digital Transformation, 2025, 2026a, 2026b).

Nevertheless, the adoption of legislation, strategies, policies, and technical safeguards does not automatically ensure their effective and consistent implementation across public institutions. Common institutional challenges may include gaps between formally adopted policies and everyday practices, excessive reliance on technical measures without sufficient employee training, inconsistent implementation of security standards, and an emphasis on responding to incidents rather than preventing them (Anderson, 2020; NIST, 2018; Whitman & Mattord, 2018). These challenges demonstrate the importance of examining the interaction between technological safeguards, organizational practices, and the human factor within the specific institutional context of the Republic of North Macedonia.

Accordingly, the aim of this study is to analyze information-security challenges in public institutions in the Republic of North Macedonia and to provide recommendations for strengthening institutional resilience through an integrated approach.

The study addresses the following research question:

What is the role of the human factor and institutional security practices in protecting information systems from cyber threats in public institutions of the Republic of North Macedonia?

I. RESEARCH METHODOLOGY

Research Design

This study employed an exploratory mixed-methods design combining qualitative document analysis with quantitative survey research. The qualitative component was used to examine the conceptual, legal, institutional, technical, organizational, and human dimensions of information security. The quantitative component was used to obtain preliminary information about employee awareness, behavior, and security practices in selected public institutions in the Republic of North Macedonia.

Qualitative Component

The qualitative component was based on a descriptive and comparative analysis of scientific literature, books, international standards, national legislation, government cybersecurity strategies, institutional reports, official statements, and documented cybersecurity incidents. Sources were selected according to their relevance to information security, cybersecurity in the public sector, employee behavior, institutional security policies, and the national context of the Republic of North Macedonia.

The analysis focused on identifying common cyber threats, the role of the human factor, technical and organizational protection measures, incident-response mechanisms, and the national legal and institutional cybersecurity framework. Findings from the different sources were synthesized to develop recommendations for improving institutional resilience.

Quantitative Survey Component

A cross-sectional electronic survey was conducted among 40 employees working in public institutions in the Republic of North Macedonia. Participants were recruited from ministries and other public-sector institutions. Data were collected between 9 March and 16 March 2026 using an electronic survey. Participation was voluntary, and no directly identifying personal information was collected.

Questionnaire

The questionnaire consisted of 10 structured, closed-ended questions organized into three areas:

1. Participant and employment characteristics;
2. Information-security practices, including password use, multi-factor authentication, system updates, and device use; and
3. Cybersecurity awareness and preparedness, including phishing recognition, previous training, institutional policies, and incident response.

The questionnaire was developed based on the objectives of the present study and relevant literature on employee cybersecurity awareness and information-security practices (Hadnagy, 2018; Whitman & Mattord, 2018).

Statistical Analysis

Survey responses were analyzed using IBM SPSS Statistics, version 23.0 (IBM Corp., Armonk, NY, USA). Categorical responses were summarized as frequencies and percentages. No inferential statistical comparisons were conducted because of the small sample size, non-probability sampling method, and exploratory purpose of the survey. Percentages were calculated using the total number of respondents as the denominator unless otherwise indicated.

Ethical Considerations

Participation was voluntary, and respondents were informed about the purpose of the study before completing the questionnaire. Completion of the questionnaire was considered an indication of informed consent. No directly identifying personal information was collected, and responses were reported only in aggregate form. Because the study was an exploratory anonymous survey and did not collect directly identifying or sensitive personal data, formal ethical approval was not obtained.

Study Limitations

The quantitative findings should be interpreted as exploratory rather than representative of all public institutions in the Republic of North Macedonia. The study included only 40 respondents recruited through a non-probability sampling method and did not include a nationally representative distribution of institutions, geographic regions, occupational groups, or administrative levels. The use of self-reported responses may also have introduced recall and social-desirability bias. Furthermore, the questionnaire was not subjected to formal psychometric validation. Larger studies using probability sampling, validated instruments, and broader institutional participation are required to confirm the findings.

II. SURVEY RESULTS

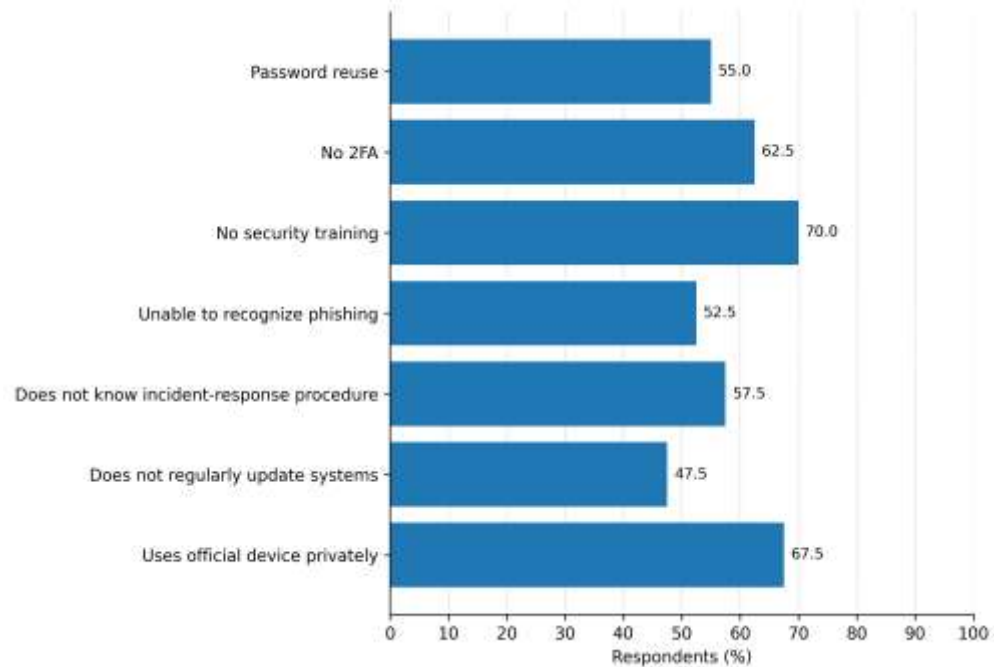
Table 1. Descriptive Results of the Information-Security Survey Among Public-Sector Employees (n = 40)

Survey item	Response	n	%
Familiar with the concept of cyber threats	Yes	33	82.5
	No	7	17.5
Uses different passwords for different systems	Yes	18	45.0
	No	22	55.0
Uses two-factor authentication	Yes	15	37.5
	No	25	62.5
Has attended information-security training	Yes	12	30.0
	No	28	70.0
Can recognize a phishing message	Yes	19	47.5
	No	21	52.5
Institution has clear security policies	Yes	16	40.0
	No	14	35.0
	Not sure	10	25.0
Knows how to respond to a cyber incident	Yes	17	42.5
	No	23	57.5
Regularly updates work systems	Yes	21	52.5
	No	19	47.5
Uses official devices for private purposes	Yes	27	67.5
	No	13	32.5
Perceived risk of cyberattacks in the institution	High	14	35.0
	Medium	18	45.0
	Low	8	20.0

The survey findings indicated a difference between respondents' general awareness of cyber threats and their practical security behavior. Although 33 respondents (82.5%) reported being familiar with the concept of cyber threats, 22 respondents (55.0%) reported that they did not use different passwords for different systems. Twenty-five respondents (62.5%) did not use two-factor authentication, and 28 respondents (70.0%) had not attended information-security training. In addition, 21 respondents (52.5%) reported that they could not recognize a phishing message, while 23 respondents (57.5%) did not know how to respond to a cyber incident. Twenty-seven respondents (67.5%) reported using official devices for private purposes. These findings suggest that general awareness was not consistently translated into secure everyday practices. Nevertheless, because of the small and non-probability sample, the findings should be considered indicative and exploratory rather than representative of all public-sector employees in the Republic of North Macedonia.

The most frequently reported vulnerabilities were the absence of information-security training (70.0%), the private use of official devices (67.5%), and the non-use of two-factor authentication (62.5%). The distribution of the principal information-security vulnerabilities reported by the respondents is presented in Figure 1.

Figure 1. Reported information-security vulnerabilities among respondents (n = 40).



2FA = two-factor authentication. Percentages were calculated based on 40 respondents.

III. THEORETICAL FRAMEWORK OF INFORMATION SECURITY

Information security and cybersecurity are closely related but are not identical concepts. Information security is the broader concept and refers to the protection of information in all forms, including digital, printed, verbal, and physical information. Cybersecurity primarily concerns the protection of digital information, networks, devices, information systems, and online services against cyber threats. Therefore, this paper uses the term “information security” when discussing the general protection of institutional information and uses “cybersecurity” when referring specifically to digital systems, cyberattacks, network security, and cyber-incident response (Anderson, 2020; Whitman & Mattord, 2018).

In this broader sense, information security comprises technical, organizational, physical, and human measures designed to protect information against unauthorized access, misuse, disclosure, alteration, loss, or destruction. It is commonly based on the principles of confidentiality, integrity, and availability and provides the foundation for the effective management and protection of institutional information resources (Peltier, 2001; Whitman & Mattord, 2018).

A fundamental model in information security is the CIA triad, which consists of confidentiality, integrity, and availability (Stallings, 2018; Whitman & Mattord, 2018):

1. **Confidentiality** ensures that information is accessible only to authorized individuals and prevents unauthorized disclosure or access.
2. **Integrity** refers to the accuracy, completeness, and consistency of information and protects data against unauthorized alteration.

3. **Availability** ensures that information and information systems are accessible to authorized users when required.

The CIA triad provides a foundation for the development of institutional security policies and standards because it supports an integrated approach encompassing technical safeguards, organizational procedures, access controls, and employee responsibilities (Peltier, 2001; Whitman & Mattord, 2018).

Malicious software, including viruses, spyware, worms, and ransomware, represents a significant threat to user credentials, institutional data, and information systems. Malware may monitor user activities, record keystrokes, steal information, disrupt system operations, or restrict access to data and devices until a ransom is paid (Anderson, 2020; Stallings, 2018).

Malware infections may occur through suspicious attachments, malicious links, applications obtained from unverified sources, compromised websites, or infected external devices. Regular software updates, endpoint protection, antivirus and anti-malware tools, controlled use of external storage devices, and employee awareness are therefore important preventive measures (Anderson, 2020; Stallings, 2018).

IV. NATIONAL LEGAL AND INSTITUTIONAL CYBERSECURITY FRAMEWORK IN THE REPUBLIC OF NORTH MACEDONIA

National Legislation and Strategic Framework

The Republic of North Macedonia has strengthened its national cybersecurity framework through the adoption of legislation, strategic documents, cybersecurity standards, and institutional mechanisms intended to improve the protection of networks, information systems, personal data, and classified information. A central component of this framework is the Law on Security of Network and Information Systems, adopted in 2025 and published in the *Official Gazette of the Republic of North Macedonia*, No. 135/2025. The law establishes requirements concerning cybersecurity governance, risk-management measures, incident reporting, supervision, and coordinated responses to threats affecting networks and information systems (Law on Security of Network and Information Systems, 2025).

The Cybersecurity Strategy 2025–2028 and its accompanying action plan provide the principal strategic direction for strengthening national cyber resilience. The strategy emphasizes institutional coordination, the protection of critical and public-sector systems, cybersecurity capacity development, employee education, incident-response capabilities, and alignment with relevant European cybersecurity principles. Minimum cybersecurity standards have also been introduced to promote more consistent security practices across institutions (Ministry of Digital Transformation, 2025).

Other relevant legislation includes the Law on Personal Data Protection, published in the *Official Gazette of the Republic of North Macedonia*, No. 42/2020 and subsequently amended by No. 294/2021. This legislation regulates the lawful processing and protection of personal data and establishes relevant rights, obligations, and accountability requirements (Law on Personal Data Protection, 2020; The Law on Amending the Law on Personal Data Protection, 2021).

The Law on Classified Information, published in the *Official Gazette of the Republic of North Macedonia*, No. 275/2019, regulates the classification, protection, handling, and authorized use of classified information (Law on Classified Information, 2019). Together, these laws demonstrate that institutional information security

encompasses not only technical cybersecurity but also privacy, confidentiality, access control, accountability, and lawful information processing.

National CERT and Government Incident-Response Structures

At the national level, the National Centre for Computer Incident Response, known as MKD-CIRT, functions as a national point of contact for reporting and coordinating responses to computer and network security incidents. Its activities include receiving incident reports, analyzing and classifying incidents, providing security guidance, coordinating incident responses, and supporting its constituents in addressing cyber threats (National Centre for Computer Incident Response [MKD-CIRT], n.d.).

A separate Computer Security Incident Response Team for executive authorities, MKD-GOV-CSIRT, operates within the Ministry of Digital Transformation. It was established pursuant to the Law on Security of Network and Information Systems and officially commenced operations on May 1, 2026 (Law on Security of Network and Information Systems, 2025; Ministry of Digital Transformation, 2026).

The responsibilities of MKD-GOV-CSIRT include supporting the prevention and early detection of cybersecurity incidents, maintaining relevant incident records, exchanging information with MKD-CIRT and other competent authorities, and coordinating responses to incidents affecting government institutions (Ministry of Digital Transformation, 2026).

The existence of MKD-CIRT and MKD-GOV-CSIRT represents an important development in the national cybersecurity architecture. Their effective operation nevertheless depends on clearly defined institutional reporting channels, adequate staffing and technical resources, timely information exchange, cooperation among public institutions, and employees' ability to recognize and report potential incidents.

Government Cybersecurity Initiatives

Recent government cybersecurity initiatives include the Cybersecurity Strategy 2025–2028, its implementation action plan, minimum cybersecurity standards, the establishment of dedicated cybersecurity functions within the Ministry of Digital Transformation, and the commencement of MKD-GOV-CSIRT operations. These developments indicate a transition toward a more coordinated, legally structured, and proactive national cybersecurity approach (Ministry of Digital Transformation, 2025).

Nevertheless, the adoption of legislation and the establishment of institutional structures must be accompanied by effective implementation. Public institutions require regular risk assessments, security audits, software and system updates, employee training, simulated phishing exercises, incident-response procedures, backup testing, and clearly assigned responsibilities (Anderson, 2020; NIST, 2018; Whitman & Mattord, 2018).

The survey findings of the present study should therefore be interpreted within this national context. The identified weaknesses in password practices, two-factor authentication, employee training, phishing recognition, and incident-response preparedness may indicate a gap between national cybersecurity objectives and the everyday implementation of information-security practices within public institutions.

V. CYBER THREATS IN THE PUBLIC SECTOR

Public institutions are exposed to various cyber threats because of the sensitive information they process, their dependence on interconnected information systems, and the importance of the public services they provide. Understanding and systematically

managing these threats is essential for protecting institutional resources, ensuring service continuity, and maintaining public trust (Anderson, 2020; ENISA, 2023).

The most common cyber threats affecting public institutions and the corresponding protection measures are summarized in Table 2.

Table 2. Common Cyber Threats and Recommended Protection Measures in Public Institutions

Cyber threat	Description	Potential consequences	Recommended protection measures
Phishing	Fraudulent emails, messages, or websites designed to obtain sensitive information	Credential theft, unauthorized access, and data compromise	Employee training, email filtering, verification procedures, and multi-factor authentication
Ransomware	Malicious software that encrypts data or blocks access to systems	Service interruption, data loss, and financial and operational damage	Regular backups, system updates, endpoint protection, and incident-response planning
Malware	Harmful software, including viruses, spyware, and worms	Data theft, system disruption, and unauthorized monitoring	Antivirus and anti-malware software, application control, and regular updates
Social engineering	Manipulation of employees into disclosing information or performing unsafe actions	Disclosure of confidential information and unauthorized access	Employee education, verification procedures, and security-awareness programs
Brute-force attacks	Automated attempts to guess usernames or passwords	Account compromise and unauthorized system access	Strong unique passwords, login-attempt limits, and multi-factor authentication
Distributed denial-of-service attacks	Large volumes of malicious traffic intended to overload systems or services	Service unavailability and disruption of institutional operations	Firewalls, traffic monitoring, rate limiting, load balancing, and DDoS protection

Note. The table was synthesized from Anderson (2020), ENISA (2023), Hadnagy (2018), Stallings (2018), and Whitman and Mattord (2018).

V.1. Phishing Attacks

Phishing and social engineering are widely used methods for compromising sensitive information by exploiting human behavior rather than directly attacking technical vulnerabilities. These attacks may involve fraudulent emails, websites, text messages, telephone calls, or other communication channels that imitate legitimate organizations or trusted individuals. Attackers commonly attempt to persuade users to reveal authentication credentials, open malicious attachments, visit fraudulent websites, or perform actions that compromise institutional systems (ENISA, 2023; Hadnagy, 2018).

The effectiveness of phishing attacks frequently depends on creating urgency, fear, authority, curiosity, or trust. For example, a fraudulent message may warn the recipient that an account will be suspended unless immediate action is taken. Employees should therefore

verify sender identities, inspect web addresses carefully, and avoid interacting with unexpected links or attachments (Hadnagy, 2018; Whitman & Mattord, 2018).

Multi-factor authentication adds an additional layer of security by requiring users to provide evidence from more than one authentication category, such as something they know, possess, or are. Its implementation can substantially reduce the risk of unauthorized account access when a password has been compromised, although phishing-resistant authentication methods provide stronger protection than one-time codes or SMS-based authentication (NIST, 2022).

In public institutions, successful phishing attacks may compromise sensitive records, interrupt administrative processes, facilitate malware installation, and cause financial or operational losses. Employee awareness, clear verification procedures, secure email systems, and multi-factor authentication should therefore be implemented together as part of a layered security approach (ENISA, 2023; Hadnagy, 2018).

V.2. Malicious Software

Malicious software, or malware, includes programs developed to damage systems, disrupt operations, monitor users, steal information, or obtain unauthorized access. Common types include ransomware, spyware, viruses, and worms (Anderson, 2020; Stallings, 2018).

1. **Ransomware** encrypts data or prevents access to systems and commonly demands payment in exchange for restoring access.

2. **Spyware** secretly monitors user activity and collects information without the user's knowledge or authorization.

3. **Viruses and worms** spread between files, devices, or networks and may damage data or disrupt system operations.

Malware risk should be managed through an integrated approach that includes system updates, endpoint protection, access controls, tested backups, incident-response procedures, network monitoring, and employee education (Anderson, 2020; Stallings, 2018).

A prominent example is the WannaCry ransomware attack of May 2017. The attack affected more than 200,000 computers in at least 100 countries and significantly disrupted public- and private-sector organizations, including the National Health Service in England (Europol, 2017; National Audit Office, 2018). NHS England identified 6,912 cancelled appointments and estimated that more than 19,000 appointments may have been cancelled in total (National Audit Office, 2018).

The impact of WannaCry demonstrated the importance of timely security updates, supported operating systems, effective backup arrangements, business-continuity planning, and coordinated incident response. The incident also demonstrated that technical safeguards must be supported by effective organizational processes and clearly assigned responsibilities (National Audit Office, 2018).

In the Republic of North Macedonia, a cyberattack disrupted the information systems of the Health Insurance Fund in February 2023. The interruption created operational difficulties for the Fund, insured individuals, and healthcare institutions, although temporary procedures were introduced to maintain access to essential healthcare services (Health Insurance Fund of the Republic of North Macedonia, 2023).

This incident illustrates the operational dependence of public institutions on information systems and the importance of system resilience, business-continuity planning,

timely communication, and coordinated incident response. It also supports the need for a proactive approach combining technical safeguards, organizational procedures, and employee preparedness.

V.3. Password Attacks and Social Engineering

Password attacks and social-engineering techniques represent significant threats because they target both authentication systems and the individuals who use them. These attacks may include brute-force attempts, dictionary attacks, credential stuffing, phishing, impersonation, and manipulation through telephone or electronic communication (Hadnagy, 2018; Stallings, 2018).

Data-interception attacks may occur when an attacker secretly positions themselves between a user and a digital service. In a man-in-the-middle attack, the attacker may monitor, relay, or alter communications without the knowledge of the communicating parties. Such attacks are more likely when users connect through compromised or inadequately protected networks (Anderson, 2020; Whitman & Mattord, 2018).

Physical observation also presents an information-security risk. Shoulder surfing occurs when an unauthorized person observes a user entering a password, personal identification number, or other sensitive information. These risks demonstrate that information security includes physical and behavioral safeguards in addition to technical controls (Whitman & Mattord, 2018).

Employees should avoid entering sensitive information when using unknown or inadequately protected networks. They should also remain aware of their physical surroundings and use approved password-management tools to create and securely store unique credentials (Whitman & Mattord, 2018).

Brute-force attacks use automated methods to test numerous password combinations until a valid password is identified. Dictionary attacks apply lists of commonly used words, passwords, or predictable combinations. The effectiveness of these attacks may be reduced through long and unique passwords or passphrases, login-attempt controls, system monitoring, and multi-factor authentication (Paar et al., 2010; Stallings, 2018).

Password reuse further increases institutional risk. When users apply the same credentials to several systems, the compromise of one service may enable attackers to access other accounts. Credential stuffing involves the automated use of previously stolen username-and-password combinations against additional services to identify accounts that use the same credentials (Newhouse et al., 2019; OWASP Foundation, 2021).

Institutions should therefore require unique passwords, support approved password managers, monitor repeated login failures, limit automated authentication attempts, and implement multi-factor authentication for sensitive systems (Newhouse et al., 2019; NIST, 2022).

Social engineering exploits human psychology, trust, authority, fear, and insufficient security awareness. It does not necessarily depend on a technical vulnerability; instead, attackers manipulate individuals into revealing information or performing actions that compromise institutional security (Hadnagy, 2018).

Examples include fraudulent emails, telephone scams, impersonation of technical-support personnel, and false instructions appearing to originate from supervisors or trusted institutions. Successful social-engineering attacks may lead to unauthorized access, disclosure of confidential information, malware installation, and financial or operational damage (Hadnagy, 2018; Whitman & Mattord, 2018).

V.4. Integrated Approach to Protection

Effective protection against cyber threats requires coordinated technical, organizational, and human safeguards. Important measures include: training employees to recognize suspicious messages and activities; implementing unique passwords and multi-factor authentication; applying timely software and security updates; conducting regular risk assessments and security audits; testing incident-response and backup procedures; and conducting controlled phishing simulations and other awareness exercises.

No single protective measure can eliminate cybersecurity risk. Layered controls reduce the likelihood that the failure of one safeguard will result in the compromise of the entire institutional system (Peltier, 2001; Whitman & Mattord, 2018).

V.5. The Role of the Human Factor

The human factor is a central component of information security because employee decisions and behavior may either create vulnerabilities or prevent attacks. Human involvement in security incidents may include susceptibility to phishing, credential misuse, failure to follow procedures, delayed incident reporting, or unintentional disclosure of information (Hadnagy, 2018; Whitman & Mattord, 2018).

Employees therefore have a dual role within institutional information security:

1. **As a potential source of vulnerability:** Insufficient awareness or failure to follow established procedures may contribute to unauthorized access, information disclosure, or malware infection.

2. **As a first line of defense:** Properly trained employees may recognize suspicious activity, report incidents promptly, verify unusual requests, and prevent attacks from progressing.

Developing an institutional security culture is therefore essential. This requires regular employee education, management support, simulated exercises, clear reporting channels, and the promotion of individual responsibility for protecting institutional information (Hadnagy, 2018; Whitman & Mattord, 2018).

Technical safeguards remain necessary, but their effectiveness may be limited when employees do not understand or consistently apply institutional security requirements.

Security Practices in Institutions

Effective institutional information security is based on the coordinated implementation of technical controls, organizational policies, and employee-awareness measures. These practices protect sensitive information, support operational continuity, and reduce the likelihood and consequences of cyber incidents (Stallings, 2018; Whitman & Mattord, 2018).

Important technical measures include:

1. **Firewalls and content-filtering systems:** These systems monitor and restrict network traffic and help prevent unauthorized access.

2. **Antivirus and anti-malware software:** These tools support the prevention, detection, and removal of malicious programs.

3. **Email-security systems:** These measures help identify or block spam, phishing messages, malicious links, and harmful attachments.

Technical measures should be regularly updated, monitored, and evaluated to ensure that they remain effective against changing threats (Anderson, 2020; Stallings, 2018).

V.6. Organizational Policies and Procedures

Information-security policies establish expectations and requirements for employee behavior and the authorized use of institutional systems. Important organizational measures include:

1. **Principle of least privilege:** Employees should have access only to the information and systems required to perform their authorized duties.
2. **Regular backup and data archiving:** Tested backups enable institutions to restore information following ransomware attacks, technical failures, or human errors.
3. **Incident-response plans:** Clearly defined procedures support the reporting, containment, investigation, and recovery of security incidents.
4. **Access-control procedures:** User accounts and permissions should be created, reviewed, modified, and removed according to established institutional responsibilities.
5. **Security audits and risk assessments:** Periodic assessments help identify weaknesses and evaluate the effectiveness of existing controls.

These organizational measures provide a structured foundation for managing information-security risks and assigning accountability within public institutions (Peltier, 2001; Whitman & Mattord, 2018).

Training and Development of Security Culture

Employee education is an essential component of institutional security. Training should address phishing, suspicious links and attachments, social engineering, password practices, incident reporting, confidential-information handling, and the secure use of institutional devices.

Controlled simulations may be used to identify areas requiring additional training, provided that they are conducted according to clear institutional policies and ethical procedures. Management should also encourage employees to report errors and suspicious activity promptly without fear of inappropriate punishment (Hadnagy, 2018; Peltier, 2001).

Continuous Improvement

Information security is a continuous process rather than a one-time activity. Institutions should regularly review their systems, update software and security controls, reassess risks, test response procedures, and revise policies to address emerging threats (Anderson, 2020; Peltier, 2001).

Overall, the coordinated implementation of technical safeguards, organizational policies, and employee education strengthens institutional resilience. Effective information security therefore depends not only on technology but also on governance, accountability, training, and responsible employee behavior.

VI. PASSWORD MANAGEMENT AND AUTHENTICATION

One of the key aspects of information security in the public sector is proper password management and the application of authentication mechanisms. Passwords represent a primary layer of defense against unauthorized access to institutional systems and sensitive data. Weak, predictable, or reused passwords are a common source of account compromise and successful cyberattacks (Stallings, 2018; Whitman & Mattord, 2018).

Strong Passwords

Long and unique passwords or passphrases increase resistance to brute-force and dictionary attacks. Password length is generally a more important indicator of password strength than the mandatory use of uppercase and lowercase letters, numbers, and special characters. Current guidance recommends a minimum length of 15 characters when a password is used as the only authentication factor. Passwords should not be changed periodically without reason; instead, they should be changed when there is evidence or suspicion that they have been compromised (NIST, 2025).

Multi-Factor Authentication

Multi-factor authentication adds one or more security layers beyond a password by requiring different forms of authentication, such as a time-based code, physical token, mobile device, passkey, or biometric characteristic. Two-factor authentication is a type of multi-factor authentication that uses exactly two authentication factors. The implementation of multi-factor authentication significantly reduces the risk of unauthorized account access, even when a password has been exposed (Anderson, 2020; NIST, 2025).

Practical Recommendations for Institutions

Institutions should: establish policies for creating and maintaining long, unique passwords or passphrases; prohibit the reuse of institutional passwords across different systems; support the use of approved password managers; enable multi-factor authentication for critical systems and applications; train employees to recognize suspicious authentication requests and login attempts; and monitor security incidents and analyze breaches to support the continuous improvement of security practices (Peltier, 2001; Whitman & Mattord, 2018).

The implementation of these practices strengthens institutional security, reduces the likelihood of account compromise and human error, and improves institutional resilience against cyber threats.

Protection of Workstations

Workstations represent primary access points to institutional information and systems; therefore, their protection is essential for information security. Inadequate physical and digital protection of devices may result in data compromise, unauthorized access, or the spread of malicious software (Stallings, 2018; Whitman & Mattord, 2018).

Technical Measures

Important technical measures include:

1. **Regular updates and patches:** Installing current security updates for operating systems and applications reduces the risk that known vulnerabilities will be exploited.
2. **Antivirus and anti-malware software:** These tools support the prevention, detection, and removal of malicious programs that may infect institutional workstations.
3. **Screen locking when away from the device:** Locking an unattended workstation helps prevent unauthorized physical access, even during short periods of absence (Anderson, 2020; Stallings, 2018).

Physical and Organizational Protection

Physical and organizational protection measures include: using official devices for work involving confidential information and avoiding the sharing of institutional devices with unauthorized individuals; implementing a clean desk policy under which computers and screens are locked and sensitive documents and storage media are securely stored when not in use; and controlling physical access to offices, work areas, and server rooms through identification systems and restricted-entry mechanisms (Peltier, 2001; Whitman & Mattord, 2018).

Recommendations for Institutions

Institutions should: educate employees about the importance of locking unattended devices and installing software updates; continuously monitor security systems and identify potential vulnerabilities; and integrate physical and digital protection measures into the institution's overall information-security strategy (Anderson, 2020; Peltier, 2001).

Effective workstation protection increases institutional resilience against cyber threats, reduces the possibility of human error and unauthorized access, and supports secure access to critical information.

VII. MANAGEMENT OF CONFIDENTIAL INFORMATION

Effective management of confidential information is a key aspect of information security in public institutions. Improper handling of physical documents and digital data may result in unauthorized access, information disclosure, data leakage, or compromise of institutional integrity (Peltier, 2001; Whitman & Mattord, 2018).

Policies and Procedures for Document Protection

Clean Desk Policy

A clean desk policy requires physical documents, removable storage media, and other materials containing sensitive information to be securely stored or locked away when they are not in use. Computers and screens should also be locked whenever employees leave their workstations. The application of this policy reduces the likelihood of unauthorized access and helps preserve the confidentiality and integrity of institutional information (Peltier, 2001; Whitman & Mattord, 2018).

Data Encryption

The use of cryptographic techniques to protect documents, stored files, and electronic communications helps ensure that confidential information can be accessed only by authorized individuals. Encryption is particularly important when information is transmitted through digital channels, stored on portable devices, or accessed remotely (Anderson, 2020; Whitman & Mattord, 2018).

Physical and Digital Protection

Important physical and digital protection measures include: storing confidential physical documents in locked drawers, cabinets, or safes; protecting digital files through appropriate access controls, passwords, and encryption; and restricting access to documents and information systems to authorized personnel in accordance with the principle of least privilege (Anderson, 2020; Peltier, 2001).

Training and Staff Education

Regular training on the appropriate handling of confidential information contributes to the development of an effective institutional security culture. Employees should understand the risks associated with physical exposure, unauthorized disclosure, digital compromise, and improper information sharing. They should also know how and where to report suspicious activities or potential security incidents (Peltier, 2001; Whitman & Mattord, 2018).

The implementation of these practices reduces the likelihood of human error, protects data confidentiality, and contributes to maintaining public trust in institutions.

Remote Work and Mobile Security

The increasing use of remote and mobile work creates additional information-security challenges for public institutions. Accessing institutional systems outside controlled workplace networks may increase the risk of unauthorized access, credential

theft, malware infection, and compromise of confidential information (Stallings, 2018; Whitman & Mattord, 2018).

Secure Remote Work

Important remote-work security measures include:

Virtual private networks: The use of an approved virtual private network provides an encrypted communication channel for data transmission and reduces the risk of interception by unauthorized individuals.

Secure Wi-Fi networks: Employees working remotely should use protected and trusted networks. Public or unsecured Wi-Fi networks may expose communications and authentication credentials to security risks.

Avoiding public computers and devices: Institutional systems should not be accessed from unknown, public, or shared devices because such devices may contain malware, keyloggers, or other unauthorized monitoring tools (Anderson, 2020; Whitman & Mattord, 2018).

Mobile Security

Important mobile-security measures include: protecting mobile devices with a PIN, strong password, biometric authentication, or other approved locking mechanism; installing applications only from authorized or trusted sources; regularly updating device operating systems and applications; enabling multi-factor authentication for institutional accounts; and avoiding the unnecessary local storage of confidential information on mobile devices.

Highly sensitive information should preferably be stored on secure institutional servers or approved cloud systems rather than directly on personal or mobile devices (Stallings, 2018; Whitman & Mattord, 2018).

Training and Recommendations

Institutions should: provide regular employee training on remote-work and mobile-security threats; require multi-factor authentication for remote and mobile access; establish clear policies governing the use of personal and institutional devices; continuously monitor relevant security systems; and regularly review and update remote-work protocols (Peltier, 2001; Whitman & Mattord, 2018).

The implementation of these measures reduces the vulnerability of institutional systems and strengthens information security in remote-working environments, thereby supporting business continuity without compromising data confidentiality.

VIII. INCIDENT RESPONSE

Effective management of security incidents is a key aspect of information security in the public sector. Even when technical and organizational protection measures are established, institutions remain exposed to cyberattacks, system failures, and human errors. A timely and structured response can significantly reduce potential damage and support business continuity (Whitman & Mattord, 2018).

Key Steps in Incident Response

Problem identification: Rapid recognition of unusual activities, suspicious messages, malware infections, or unauthorized access is the first step in limiting potential damage.

Isolation of affected systems: Disconnecting or separating compromised devices and network segments helps prevent the attack from spreading to other systems.

Notification of responsible authorities: Immediate reporting to information-technology departments, information-security personnel, and other responsible authorities enables a coordinated response and may help prevent further incidents.

Preservation of evidence: Logs, emails, files, system records, and other digital evidence should not be deleted or altered because they may assist in identifying the cause, method, and source of the incident (Anderson, 2020; Whitman & Mattord, 2018).

Importance of Timely Response

A rapid and appropriate response to security incidents can reduce operational disruption, financial loss, data compromise, and damage to institutional reputation. It may also strengthen the institution's ability to prevent or manage similar incidents in the future (Anderson, 2020; Whitman & Mattord, 2018).

Employees should be trained to recognize suspicious activities and report them immediately through established institutional channels. In this way, employees can serve as an important first line of defense against cyber threats.

Training and Simulations

Regular employee training and simulated attacks, including phishing and ransomware exercises, can increase awareness and prepare institutions to respond more effectively to real security incidents.

Furthermore, clearly defined and standardized incident-response procedures support a coordinated, consistent, and efficient response to potential cyberattacks (Anderson, 2020; Whitman & Mattord, 2018).

IX. INSTITUTIONAL SECURITY MECHANISMS

The public sector is increasingly exposed to evolving cyber threats, which necessitates an integrated approach to the protection of information systems. Institutional security mechanisms include technical, organizational, and continuous measures that collectively enhance the overall level of security (Anderson, 2020; ENISA, 2023).

The analysis of information security in public institutions in the Republic of North Macedonia reveals a complex interaction between technological, organizational, and human factors. In the context of increased digitalization and dependence on information systems, institutions are becoming increasingly vulnerable to various types of cyber threats, which emphasizes the need for a systematic and integrated security approach.

One of the key findings of the analysis is that, although technical security measures such as firewall systems, antivirus solutions, and intrusion-detection systems are important, their effectiveness is limited when they are not supported by appropriate organizational policies and employee training. This indicates that information security cannot be viewed solely as a technical issue but rather as a multidisciplinary challenge requiring coordination at multiple levels (Peltier, 2001; Whitman & Mattord, 2018).

Furthermore, the analysis indicates that cyber threats, particularly phishing attacks and social engineering, frequently exploit the human factor. Employees, as system users, may represent a vulnerable component of the security chain, especially when sufficient training or awareness is lacking. In this context, the human factor plays a dual role—as a potential vulnerability and as a key element of cyber defense (Hadnagy, 2018; Whitman & Mattord, 2018).

Additionally, institutions that implement clearly defined security policies, such as the principle of least privilege, regular backups, and incident-response plans, may demonstrate greater resilience to cyberattacks. This highlights the importance of

organizational measures as a complement to technical solutions (Peltier, 2001; Whitman & Mattord, 2018).

A significant aspect emerging from the analysis is the need for continuous improvement of security practices. Cyber threats are dynamic and constantly evolving, requiring institutions to regularly update their systems, revise policies, assess risks, and invest in employee training. The absence of such an approach may lead to an increased risk of incidents and compromise of sensitive information (Anderson, 2020; ENISA, 2023).

The findings also indicate a need for greater institutional coordination and standardization of security practices. When implementation differs between institutions, uneven levels of protection and potential weaknesses may arise. The establishment of unified standards and guidelines could therefore improve the overall security posture of the public sector.

Finally, the analysis shows that effective information security in public institutions requires an integrated approach combining technical solutions, organizational policies, and the development of a security culture. Such an approach may reduce the likelihood and consequences of cyber incidents and contribute to maintaining citizens' trust in institutional systems (Peltier, 2001; Whitman & Mattord, 2018).

Despite the existence of basic technical and organizational security measures, the findings of the present study indicate that public institutions in the Republic of North Macedonia continue to face challenges in the practical implementation of information security. One of the main issues is the insufficient integration of technical solutions and organizational policies, resulting in a fragmented approach to security.

In some cases, technical measures such as antivirus systems, firewalls, and intrusion-detection systems may be implemented without sufficient continuous monitoring, updating, or evaluation of their effectiveness. This may create a false sense of security in which institutions consider themselves adequately protected while significant vulnerabilities remain (Anderson, 2020; Stallings, 2018).

In addition, insufficient employee education and awareness represent an important concern. Although formal security policies may exist, they are not always fully understood or consistently applied in daily operations. This may increase institutional vulnerability, particularly in relation to phishing attacks and social engineering, which exploit human trust, error, and negligence (Hadnagy, 2018; Whitman & Mattord, 2018).

Another critical aspect is the need for unified standards and coordination at the public-sector level. Different institutions may apply different levels of security measures, resulting in uneven protection and potential weak points. Insufficient coordination and oversight may further complicate the prevention and management of cyber incidents.

Moreover, a reactive rather than proactive approach to security may increase institutional vulnerability. Responding only after an incident has occurred, rather than investing in prevention, testing, risk assessment, and continuous improvement, may increase the likelihood of data loss, financial damage, operational disruption, and erosion of public trust (Anderson, 2020; ENISA, 2023).

Therefore, although basic foundations for information security exist in the public sector, improvements are needed in practical implementation, coordination, employee education, and the development of a security culture. Only through a systematic, integrated, and proactive approach can existing weaknesses be addressed and effective protection against contemporary cyber threats be strengthened.

The analysis shows that effective information security depends on the interaction between three key components: technological systems; organizational policies; and the human factor.

Although technical measures are essential, their effectiveness is limited without adequate organizational support and employee training (Peltier, 2001; Whitman & Mattord, 2018).

Many cyberattacks, particularly phishing and social-engineering attacks, attempt to exploit human trust and behavior. This demonstrates that information security is not only a technical issue but also an organizational and behavioral one (Hadnagy, 2018; Whitman & Mattord, 2018).

Institutions that apply an integrated approach combining technical, organizational, and human measures are better positioned to prevent, detect, and respond to cyber threats (Anderson, 2020; Peltier, 2001).

Technical Measures

Technical defense systems represent an essential component of institutional information security. Key measures include:

Firewall systems: These systems control and filter traffic between internal networks and external networks, thereby helping to prevent unauthorized access.

Antivirus and anti-malware solutions: These tools detect, block, and remove malicious software, reducing the risk of ransomware, spyware, and other malware infections.

Intrusion Detection and Prevention Systems: IDS/IPS technologies monitor network activity and identify suspicious or anomalous events, thereby supporting timely detection and response (Stallings, 2018; Whitman & Mattord, 2018).

Organizational Measures

Organizational policies provide a structured approach to information-security risk management. Important measures include:

Principle of least privilege: Each employee should have access only to the information and resources necessary to perform authorized job responsibilities.

Regular backups: Backups support data recovery following cyberattacks, technical failures, or human errors.

Incident-response plans: Clearly defined procedures enable rapid and coordinated responses to incidents and may reduce damage and operational disruption (Peltier, 2001; Whitman & Mattord, 2018).

Continuous Improvement

Information security is not a static process; continuous improvement is essential for adapting to emerging threats. Important activities include:

Regular training and simulations: These activities help employees recognize and respond to new types of attacks.

System updates: Regular software updates and security patches reduce vulnerabilities associated with outdated systems.

Periodic security assessments: Regular evaluation of systems, risks, and procedures supports the identification and timely correction of weaknesses (Anderson, 2020; Peltier, 2001).

The implementation of these institutional mechanisms creates an integrated framework that strengthens the resilience of public institutions and reduces the risk of sensitive information being compromised.

X. PROPOSED INFORMATION SECURITY MODEL FOR THE REPUBLIC OF NORTH MACEDONIA

Core Concept of the Model

The proposed information-security model is based on an integrated approach combining three interdependent components: technological security; organizational governance; and the human factor.

These components should not operate independently. Technological safeguards require effective institutional policies and appropriately trained employees, while organizational procedures depend on reliable technical systems and responsible employee behavior. The coordinated implementation of these three components provides a foundation for the sustainable protection of public-sector information systems (Anderson, 2020; Peltier, 2001; Whitman & Mattord, 2018).

Model Structure

The proposed model consists of technological, organizational, and human pillars.

Technological Pillar

The technological pillar covers the infrastructure and technical mechanisms required to prevent, detect, respond to, and recover from cyber incidents. Its principal elements include: centralized security monitoring and Security Information and Event Management systems; firewall and intrusion-detection and prevention systems; antivirus and endpoint-protection solutions; data encryption; identity and access-management systems; multi-factor authentication; regular software updates and patch management; tested backup systems; and disaster-recovery mechanisms.

For the Republic of North Macedonia, centralized security monitoring should be strengthened and operationally coordinated with MKD-CIRT and MKD-GOV-CSIRT. This would support faster incident detection, information sharing, and coordinated responses across public institutions.

Organizational Pillar

The organizational pillar establishes the governance structures, responsibilities, policies, and procedures required for systematic information-security management. Its principal elements include: implementation of the national cybersecurity strategy; unified minimum security standards for public institutions; clearly assigned cybersecurity responsibilities; access-control policies based on the principle of least privilege; mandatory incident-response and business-continuity plans; regular risk assessments and security audits; standardized incident-reporting procedures; and mechanisms for institutional accountability and oversight.

A key challenge in the Republic of North Macedonia is the potential difference between the formal adoption of security requirements and their practical implementation. The proposed model therefore emphasizes strengthening the authority, resources, and coordination mechanisms of existing national cybersecurity institutions and ensuring that national requirements are applied consistently across public institutions.

Human Pillar

The human pillar addresses employee awareness, professional capacity, responsible behavior, and institutional security culture. Its principal elements include: regular cybersecurity and information-security training; phishing and social-engineering awareness; training in password management and multi-factor authentication; certification and professional development of IT and cybersecurity personnel; clear incident-reporting

channels; simulated phishing and incident-response exercises; and promotion of individual responsibility for protecting institutional information.

The model recommends mandatory periodic cybersecurity training for public-sector employees. Training should be adapted to employees' roles and should include practical exercises that assess their ability to identify suspicious communications, protect credentials, handle confidential information, and report potential incidents.

Central Element: National Coordination

The proposed model should build upon, rather than duplicate, the cybersecurity institutions that already operate in the Republic of North Macedonia. Accordingly, the principal recommendation is not the establishment of an entirely new national incident-response body, but the continued strengthening of MKD-CIRT and MKD-GOV-CSIRT.

The model recommends clearer operational coordination among national and government incident-response teams, public institutions, sectoral authorities, critical-infrastructure operators, law-enforcement bodies, and personal-data protection authorities. This coordination should include: standardized incident-reporting channels; minimum institutional cybersecurity requirements; procedures for sharing threat and incident information; clearly assigned institutional contact points; regular cross-institutional exercises; common risk-assessment procedures; and measurable implementation and performance indicators.

At the institutional level, each public institution should designate responsible cybersecurity personnel or contact points, maintain an updated incident-response plan, conduct regular risk assessments, provide employee training, and periodically test backup and recovery procedures.

At the national level, existing cybersecurity structures should be supported through adequate staffing, technical resources, threat-information sharing, cross-institutional cooperation, and continuous evaluation. The proposed model therefore complements the existing national cybersecurity framework by focusing on implementation, coordination, accountability, and continuous improvement.

Proactive Approach

The proposed model emphasizes proactive security measures designed to identify and mitigate vulnerabilities before security incidents occur. These measures include authorized penetration testing, red-team and blue-team exercises, regular risk assessments, vulnerability assessments, and incident-response simulations (Anderson, 2020; Whitman & Mattord, 2018).

A proactive approach should also include continuous monitoring, threat analysis, timely software updates, simulated phishing exercises, and regular testing of institutional response and recovery capabilities. These activities enable institutions to identify weaknesses before they are exploited and reduce dependence on reactive measures implemented only after an incident has occurred.

Implementation Mechanism

The proposed model should be implemented through four interconnected phases.

Phase 1: Assessment

The first phase involves identifying institutional information assets, evaluating cyber risks, reviewing existing security controls, and detecting technical, organizational, and human vulnerabilities.

Activities should include: information-asset identification; risk assessment; vulnerability assessment; review of employee access rights; evaluation of current policies and procedures; and assessment of employee awareness and training needs.

Phase 2: Implementation

The second phase involves introducing or strengthening the technical and organizational measures identified during the assessment.

Activities should include: deployment of technical protection systems; establishment of access-control procedures; implementation of multi-factor authentication; preparation of incident-response and business-continuity plans; standardization of institutional policies; and designation of responsible cybersecurity personnel.

Phase 3: Education and Capacity Development

The third phase focuses on employee education and the professional development of information-technology and cybersecurity personnel.

Activities should include: mandatory employee-awareness training; phishing and social-engineering simulations; incident-reporting exercises; specialized training for IT personnel; and development of an institutional security culture.

Phase 4: Continuous Evaluation and Improvement

The fourth phase involves continuous monitoring, auditing, testing, and revision of institutional security measures.

Activities should include: regular security audits; software and system updates; testing of backup and recovery procedures; review of incident-response performance; revision of policies following incidents or regulatory changes; and adaptation to emerging cyber threats.

These four phases should operate as a continuous cycle rather than as separate one-time activities.

Key Advantages of the Model

The proposed model: reduces risks associated with human error; supports coordinated oversight and incident response; improves institutional resilience; promotes consistent security standards across public institutions; strengthens prevention, detection, response, and recovery capabilities; improves accountability and clearly assigns responsibilities; supports faster information sharing during cyber incidents; and encourages continuous improvement rather than a purely reactive approach.

The integration of technological, organizational, and human measures also reduces dependence on any single security control. If one protective mechanism fails, other layers may continue to prevent or limit institutional damage.

Limitations of the Model

Implementation of the proposed model may: require substantial financial and technical resources; require sustained political and institutional commitment; encounter institutional resistance to organizational change; be affected by a shortage of qualified cybersecurity personnel; require significant coordination among institutions with different levels of technical capacity; require continuous employee participation and management support; and require regular evaluation and adaptation to evolving cyber threats.

The implementation process may therefore need to be gradual and adjusted according to the size, responsibilities, risk level, and available resources of each public institution.

Final Statement

The proposed model provides an integrated framework for strengthening information security in the public sector of the Republic of North Macedonia. It combines technological safeguards, organizational governance, employee awareness, proactive risk management, and national coordination within a unified structure.

Its implementation could contribute to greater institutional resilience, more consistent security practices, improved incident-response capabilities, stronger protection of sensitive information, and a more developed security culture. By building upon existing national cybersecurity institutions and emphasizing practical implementation, accountability, and continuous improvement, the model may support the secure and reliable functioning of public institutions in an increasingly digital environment.

CONCLUSION

Information security in the public sector is a complex and multidimensional process that requires an integrated and continuous approach. The findings of this exploratory study suggest that employee behavior, technical safeguards, and organizational policies are closely interconnected in determining the information-security preparedness of public institutions in the Republic of North Macedonia.

The survey findings revealed gaps between respondents' general awareness of cyber threats and their practical security behavior. Particularly important weaknesses were identified in information-security training, the use of two-factor authentication, password practices, phishing recognition, incident-response preparedness, and the private use of official devices. Because the survey included a small, non-probability sample of 40 respondents, these findings should be interpreted as indicative rather than representative of all public institutions in the country.

The findings also suggest that technical security measures alone are insufficient when they are not supported by clearly defined policies, effective institutional coordination, regular employee education, and established incident-response procedures. Public institutions should therefore strengthen employee training, access-control practices, multi-factor authentication, system monitoring, risk assessments, backup procedures, and incident-response preparedness.

An integrated approach combining technological safeguards, organizational governance, employee awareness, and coordination with existing national cybersecurity structures can contribute to stronger institutional resilience, improved continuity of public services, better protection of sensitive information, and the maintenance of public trust.

REFERENCES AND BIBLIOGRAPHY

1. Anderson, R. (2020). *Security engineering: A guide to building dependable distributed systems* (3rd ed.). Wiley. <https://doi.org/10.1002/9781119644682>
2. European Union Agency for Cybersecurity. (2023). *ENISA threat landscape 2023*. <https://www.enisa.europa.eu/publications/enisa-threat-landscape-2023>
3. Europol. (2017). *WannaCry ransomware*. <https://www.europol.europa.eu/wannacry-ransomware>
4. Hadnagy, C. (2018). *Social engineering: The science of human hacking* (2nd ed.). Wiley. <https://doi.org/10.1002/9781119433729>
5. Health Insurance Fund of the Republic of North Macedonia. (2023, February 8). *ФЗОРСМ: Пациентите можат да се лекуваат непречено, Фондот се справува со сајбер напад* [Patients can receive treatment without interruption: The Fund is dealing with a cyberattack]. <https://fzo.org.mk/230208-soopshtenie>

6. The Law on Amending the Law on Personal Data Protection, *Official Gazette of the Republic of North Macedonia*, No. 294/2021 (2021).
7. Law on Classified Information, *Official Gazette of the Republic of North Macedonia*, No. 275/2019 (2019).
8. Law on Personal Data Protection, *Official Gazette of the Republic of North Macedonia*, No. 42/2020 (2020). https://azlp.mk/wp-content/uploads/2022/12/lpdp_2020.pdf
9. Law on Security of Network and Information Systems, *Official Gazette of the Republic of North Macedonia*, No. 135/2025 (2025). <https://ack.mk/wp-content/uploads/2025/07/09072025ZakonZaBezbednostMIS.pdf>
10. Ministry of Digital Transformation. (2025). *Cybersecurity strategy 2025–2028*. <https://mdt.gov.mk/en-GB/regulativa/strategija-za-sajber-bezbednost>
11. Ministry of Digital Transformation. (2026a, April 27). *Announcement of the commencement of operations of MKD-GOV-CSIRT*. <https://mdt.gov.mk/en-GB/odnosi-so-javnost/soopstenija/izvestuvanje-za-zapocnuvanje-so-rabota-na-mkd-gov-csirt>
12. Ministry of Digital Transformation. (2026b, March 16). *MKD-GOV-CSIRT—3a нас* [MKD-GOV-CSIRT—About us]. <https://mdt.gov.mk/mk-MK/mkd-gov-csirt/za-nas>
13. National Audit Office. (2018). *Investigation: WannaCry cyber attack and the NHS*. <https://www.nao.org.uk/reports/investigation-wannacry-cyber-attack-and-the-nhs/>
14. National Centre for Computer Incident Response. (n.d.). *3a нас* [About us]. MKD-CIRT. <https://mkd-cirt.mk/za-nas/>
15. National Institute of Standards and Technology. (2018). *Framework for improving critical infrastructure cybersecurity* (Version 1.1). <https://doi.org/10.6028/NIST.CSWP.04162018>
16. National Institute of Standards and Technology. (2025). *Digital identity guidelines: Authentication and authenticator management* (NIST Special Publication 800-63B-4). <https://doi.org/10.6028/NIST.SP.800-63B-4>
17. National Institute of Standards and Technology. (2022). *Multi-factor authentication*. <https://www.nist.gov/itl/smallbusinesscyber/guidance-topic/multi-factor-authentication>
18. Newhouse, W., Johnson, B., Kinling, S., Kuruvilla, J., Mulugeta, B., & Sandlin, K. (2019). *Multifactor authentication for e-commerce: Risk-based, FIDO Universal Second Factor implementations for purchasers* (NIST Special Publication 1800-17B). National Institute of Standards and Technology. <https://doi.org/10.6028/NIST.SP.1800-17>
19. OWASP Foundation. (2021). *A07:2021—Identification and authentication failures*. https://owasp.org/Top10/2021/A07_2021-Identification_and_Authentication_Failures/
20. Paar, C., & Pelzl, J. (2010). *Understanding cryptography: A textbook for students and practitioners*. Springer. <https://doi.org/10.1007/978-3-642-04101-3>
21. Peltier, T. R. (2001). *Information security policies, procedures, and standards: Guidelines for effective information security management*. Auerbach Publications. <https://doi.org/10.1201/9780849390326>
22. Stallings, W. (2018). *Effective cybersecurity: A guide to using best practices and standards*. Addison-Wesley Professional.
23. Whitman, M. E., & Mattord, H. J. (2018). *Principles of information security* (6th ed.). Cengage Learning.

BUILDING INSTITUTIONAL RESILIENCE AGAINST HYBRID THREATS IN THE REPUBLIC OF NORTH MACEDONIA: IMPLEMENTATION, CHALLENGES, AND PERSPECTIVES

Besnik Emini

Brigadier General, Ministry of Defence of the Republic of North Macedonia

ebesnik23@gmail.com

Igor Gjoreski

Ministry of Defence of the Republic of North Macedonia

gjoreski.igor@yahoo.com

Abstract

In an era defined by geopolitical turbulence, economic volatility, and the fragmentation of traditional alliances, the Republic of North Macedonia faces a critical imperative to recalibrate its national security and defense posture. Hybrid threats have emerged as one of the most sophisticated contemporary security challenges. Operating persistently in the gray zone below the threshold of open armed conflict, hybrid threats defy traditional detection methods and integrate military and non-military instruments of power. As a NATO member and EU candidate country, North Macedonia remains particularly exposed to hybrid operations designed to disrupt its European integration, and undermine public trust in democratic institutions. In 2018, North Macedonia, in collaboration with the EU, conducted an EU Hybrid Risk Survey. The findings of the Survey were the starting point for the development of the National Strategy for Building Resilience and Countering Hybrid Threats (2021-2025), adopted by the Government in 2021. The Strategy embraces a whole-of-society approach and prioritizes early detection, rapid response, and resilience building across six EU strategic domains, and NATO's Seven Baseline Requirements in building resilience. Implementation has revealed some notable challenges, which include building efficient national mechanisms to deal with hybrid threats, efforts to enhance cybersecurity, and strategic communication capacities. This paper offers a critical multidisciplinary analysis of the effectiveness of institutional mechanisms for building resilience to hybrid threats in North Macedonia, and identifies principal challenges and proposes targeted recommendations aimed at enhancing inter-institutional coordination among all stakeholders.

Keywords: *hybrid threats, resilience, security, whole-of-society approach.*

1. INTRODUCTION

The strategic landscape of the 21st century has undergone profound transformation over the past two decades, driven by intensifying geopolitical competition, rapid technological disruption, economic interdependence, and the growing use of non-conventional instruments of power. There are old security challenges interlinked with new threats that are not restricted to national boundaries and institutional capacities, and current crises are very much multidimensional and interdependent, which can lead to cascading effects across the political, economic, technological, environmental, and societal domains. Contemporary conflicts increasingly unfold below the threshold of conventional warfare, where state and non-state actors employ synchronized political, informational, cyber, economic, diplomatic, intelligence, and military instruments to achieve strategic objectives while avoiding direct military confrontation (NATO, 2023). There is no sector, ministry, or agency, which would be able to respond to these events in isolation. It is the main reason why resilience has moved from the margins of academic debate to the heart of national and alliance defence planning. All those security changes have shown that national security is no longer about military power. The vulnerability of institutions, critical infrastructure, economies, and societies has come to be an equally important factor in national security, a factor that determines whether a state is able to deal with a shock without weakening its capacity to govern, provide basic services, and hold people's trust in the state. This phenomenon has become widely recognized in contemporary security literature as hybrid threats or hybrid warfare (European Centre of Excellence for Countering Hybrid Threats, 2021; NATO, 2023).

Unlike conventional military aggression, hybrid threats exploit vulnerabilities within democratic societies and critical state institutions by targeting governance structures, information ecosystems, economic stability, energy security, cyberspace, and public trust. Their ambiguity complicates attribution and delays political decision-making, thereby challenging traditional security frameworks that were primarily designed to respond to clearly identifiable military threats (European Commission & High Representative of the Union for Foreign Affairs and Security Policy, 2016; Hoffman, 2007). As a result, the distinction between internal and external security has become increasingly blurred, requiring governments to develop integrated approaches that combine national defence, civil preparedness, and societal resilience.

Having those challenges in place, almost all countries invest many resources in developing a variety of tools and mechanisms for how best to deal with this kind of new and evolving security challenges and threats. One conceptual approach recognizes investing in strengthening the resilience of national institutions, while some countries went further in strengthening the resilience of society as a whole, adopting comprehensive whole-of-society approaches that extend beyond the resilience of public institutions alone. The concept of resilience has consequently evolved from a narrow understanding focused on disaster recovery towards a multidimensional governance paradigm encompassing prevention, preparedness, adaptation, resistance, response, and rapid recovery from various disruptions (Comfort et al., 2010; Linkov & Trump, 2019).

Institutional resilience specifically refers to the capacity of public institutions to anticipate, absorb, adapt to, and recover from complex security challenges while maintaining the continuity of essential governmental functions and preserving democratic legitimacy (OECD, 2020). Within NATO's strategic framework, resilience constitutes an essential component of collective defence under Article 3 of the North Atlantic Treaty, requiring

member states to develop robust civilian preparedness alongside military capabilities to resist armed attack (NATO, 2023).

Today, resilience has been recognized as one of the core strategic values guiding both NATO and the European Union. Hence, this paper examines the strategic significance of strengthening national resilience in North Macedonia. Within both NATO and the European Union, resilience has become a central strategic concept. NATO's Strategic Concept (2022), identifies resilience as the basis of credible deterrence and collective defense, emphasizing that armed forces cannot operate effectively without resilient civilian infrastructure, reliable energy systems, secure communications networks, and functional governmental institutions. The European Union, on the same note, has focused on resilience as one of the key pillars of security, governance, digital transformation, and societal stability, and has integrated this into the frameworks of civil protection, cybersecurity, and critical-infrastructure resilience legislation and protection of democratic institutions (European Commission & High Representative of the Union for Foreign Affairs and Security Policy, 2016).

The European Union, on the same note, has focused on resilience as one of the key pillars of security, governance, digital transformation, and societal stability, and has integrated this into the frameworks of civil protection, cybersecurity, and critical-infrastructure resilience legislation and protection of democratic institutions (European Commission & High Representative of the Union for Foreign Affairs and Security Policy, 2016).

Both NATO and the European Union have adopted complementary approaches to strengthening resilience by promoting integrated civil-military cooperation, enhanced preparedness, and improved coordination among national authorities. Within NATO, these efforts are operationalized through the implementation of the Seven Baseline Requirements for National Resilience and the Resilience Objectives (formerly Resilience Targeting Process), which provide a structured framework for assisting Allies in strengthening their national preparedness on a voluntary basis while supporting the Alliance's overall deterrence and defence posture.

Limited military capabilities, constrained institutional capacities, high levels of economic openness, and substantial dependence on external political, economic, and security actors frequently characterize small states. These structural characteristics create opportunities for hostile actors seeking to exploit institutional weaknesses through influence operations, cyberattacks, economic coercion, disinformation campaigns, corruption, and interference in democratic processes (Bailes et al., 2007). Consequently, strengthening institutional resilience has become not only a national security priority but also an indispensable component of democratic governance and sustainable development. For the Republic of North Macedonia, resilience is a critical concern and resilience assumes even greater strategic significance, with limited resources, a sensitive geographical location, energy dependence, demographic challenges, and hybrid threats. These interconnected challenges necessitate the development of a national resilience framework capable of strengthening institutional adaptability, enhancing societal preparedness, and ensuring the continuity of essential governmental functions during periods of crisis. The strategic importance of resilience is particularly pronounced for small states, where even limited disruptions to critical infrastructure, energy systems, communications networks, or public trust can rapidly escalate into broader national crises. Accordingly, this study seeks to answer the following research question:

What strategic priorities should the Republic of North Macedonia develop to strengthen institutional resilience and improve national mechanisms for addressing contemporary security challenges and hybrid threats within the framework of the national security system by 2035?

The central hypothesis of this study is that strengthening cybersecurity, critical infrastructure protection, energy security, strategic communications, inter-institutional coordination, and public–private partnerships will significantly enhance the national resilience of the Republic of North Macedonia by 2035, thereby improving its capacity to prevent, withstand, respond to, and recover from contemporary security challenges and hybrid threats.

Given the emerging nature of resilience studies in the Republic of North Macedonia and the broader Western Balkans region, this research adopts an exploratory qualitative research design. The study does not seek statistical generalization; rather, it aims to provide a comprehensive strategic assessment based on documentary evidence and comparative policy analysis. The analysis relies on publicly available strategic documents, legislation, policy papers, and official reports produced by the Government of the Republic of North Macedonia, the Ministry of Defence, the Ministry of Interior, NATO, and the European Union. In addition, the study systematically compares NATO's Seven Baseline Requirements for National Resilience with the country's existing institutional arrangements in order to identify strengths, deficiencies, and opportunities for further development.

Our research is conducted using qualitative methods, from content analysis of strategic and policy documents to NATO and EU resilience practice, strategic assessments of the country's security environment, and SWOT analysis of North Macedonia's institutional, infrastructural, and societal capabilities, including the Strategy for Building Resilience and Countering Hybrid Threats 2021-2025. The SWOT analysis developed synthesizes this data into a single analytical snapshot to complement, rather than replace, the quantitative indicators proposed later in the paper. The combination of documentary analysis, comparative assessment, and strategic evaluation enables methodological triangulation, thereby compensating for the current absence of a comprehensive national resilience index.

The findings indicate that North Macedonia has made significant progress in strengthening its institutional and security architecture following its accession to NATO. Nevertheless, substantial challenges remain in several critical areas, including cybersecurity, critical infrastructure protection, energy security, strategic communications, interagency coordination, and societal preparedness. Importantly, these vulnerabilities are mutually reinforcing and should therefore be addressed through an integrated, whole-of-government and whole-of-society approach rather than through isolated sectoral initiatives. This study concludes that resilience should be understood as a comprehensive national security objective rather than as a narrowly defined civil protection or security policy concept. Accordingly, the paper proposes the National Resilience System 2035 (NRS-2035) as an integrated strategic framework consisting of seven mutually reinforcing pillars designed to strengthen the Republic of North Macedonia's capacity to anticipate, withstand, adapt to, respond to, and recover from future crises. The proposed model seeks to enhance national preparedness while supporting the country's commitments within NATO and the European Union and contributing to the long-term resilience of democratic institutions and society as a whole.

2. THEORETICAL FRAMEWORK: HYBRID THREATS AND RESILIENCE

2.1. Conceptualizing Hybrid Threats

Within the evolving strategic landscape after the Cold War, the concept of hybrid threats has emerged as one of the most influential frameworks for understanding contemporary conflict and strategic competition (Hoffman, 2007; NATO, 2022). Hybrid threats reflect broader changes in the nature of conflict associated with globalization, technological innovation, information connectivity, and geopolitical competition. Unlike conventional military aggression, hybrid threats combine multiple instruments of power—including cyber operations, disinformation campaigns, economic coercion, political interference, intelligence activities, sabotage, proxy actors, and limited military force—in ways designed to remain below the threshold of formal armed conflict while gradually undermining the target state's decision-making capacity and institutional stability (Murray & Mansoor, 2012; Wigell, 2019). Hybrid threats can be described as a synergy of military and non-military instruments aimed at destabilizing states (Hybrid CoE, 2021), and they are becoming a dominant instrument of contemporary geopolitical competition. Hybrid warfare is emerging as a dominant model of conflict, encompassing cyberattacks on state institutions, public manipulation, economic pressures (Hybrid CoE, 2021), and the use of energy resources as a political weapon. To effectively address hybrid warfare, the European Union and NATO are developing joint mechanisms to strengthen resilience in countering these threats, which also directly affect the national policies of member states.

The concept of hybrid warfare, introduced by Hoffman (2007), describes adversaries capable of simultaneously employing conventional military capabilities, irregular tactics, terrorism, and criminal activities within a unified operational strategy. Although the terms hybrid warfare and hybrid threats are frequently used interchangeably, they represent distinct analytical concepts. Hybrid warfare primarily concerns the operational conduct of military campaigns that combine conventional and unconventional methods during armed conflict. In contrast, hybrid threats encompass the broader strategic employment of coordinated political, economic, informational, technological, diplomatic, legal, intelligence, and military instruments intended to influence, coerce, destabilize, or weaken targeted states during both peacetime and periods of crisis (Hybrid CoE, 2021; NATO, 2022).

This distinction is particularly important because many hybrid activities occur outside formally declared wars and are deliberately designed to remain ambiguous, making attribution and response considerably more difficult. Dealing with hybrid threats, as modern conflicts increasingly integrate multiple modes of warfare, requires equally integrated responses from governments and security institutions.

Reflecting this broader understanding, several international organizations have developed comprehensive conceptual frameworks for addressing hybrid threats. The European Union defines hybrid threats as coordinated and synchronized actions deliberately targeting democratic states' systemic vulnerabilities through a combination of conventional and unconventional means while remaining below the threshold of formally declared warfare (European Commission & High Representative, 2016). Similarly, NATO recognizes hybrid threats as campaigns employing military and non-military instruments in an integrated manner to achieve strategic objectives while avoiding conventional military escalation, creating ambiguity, undermine resilience, and weaken political cohesion among Allies (NATO, 2022, 2023).

Both NATO and the European Union increasingly recognize resilience as the foundation of effective deterrence, crisis management, and collective defence. Rather than viewing resilience solely as the capacity to recover from crises, both organizations conceptualize it as the ability of states and societies to anticipate, withstand, adapt to, and rapidly recover from hybrid activities while maintaining the continuity of essential governmental functions and preserving democratic legitimacy (Boin et al., 2021). Consequently, addressing hybrid threats requires a comprehensive whole-of-government and whole-of-society approach involving close cooperation among military institutions, civilian authorities, the private sector, academia, civil society, and international partners. This integrated approach has become a central pillar of contemporary national security policy and a prerequisite for strengthening national resilience in an increasingly complex and interconnected security environment.

2.2. Concept of Resilience and Enablement Conceptual Approach

The concept of resilience started in ecology and systems theory, and then found its way to security studies. Holling (1973) defined resilience as the capacity of an ecological system to absorb disturbances while maintaining its essential structures, functions, and identity. This approach proved to be a very durable concept when it moved from ecological systems to socio-technical and political ones. Contemporary resilience theory emphasizes not only the capacity to withstand shocks but also the ability to adapt, transform, and continue functioning under conditions of uncertainty and disruption (Linkov & Trump, 2019). In contemporary security literature, resilience is now understood as the ability of states, institutions, infrastructure, and societies to anticipate, withstand, adapt to, and recover from disruption from natural, technological, or human-induced threats. Instead of focusing on threat elimination, as has been the dominant strategy in classical security and deterrence theory, resilience is about adaptation, continuity and recovery (Organization for Economic Co-operation and Development, 2020). It accepts that not all threats can be stopped, and its question is how a system can be designed so that when disruption happens, the damage is contained, the essential functions are not interrupted, and recovery is rapid. It is not a rejection of deterrence, but a complement to it: deterrence can fail but a resilient system will limit its impact.

Resilience refers to the ability of systems to absorb shocks, adapt, and continue functioning.

National resilience has therefore evolved into a multidimensional governance concept that integrates national security, public administration, economic stability, technological development, critical infrastructure protection, and international cooperation. It can no longer be a matter of civil protection and disaster response but of the continuity of government, the protection of vital services, and the maintenance of confidence in society when it comes to a society in the face of pressure.

NATO defines resilience as the capacity of a nation to withstand and recover rapidly from extreme shocks, whether from natural disasters, the breakdown of critical infrastructure, or hostile state and non-state actors (NATO, 2022). It is the basis for deterrence and collective defense, and military operations depend upon a solid civil infrastructure and institutions: in order for the armed forces to be deployed and maintained, fuel, transport, communications and medical services need to be functioning. Resilience is a strategic, not a passive state, and must be planned, tested, and maintained, and constantly reevaluated and upgraded.

Recognizing the strategic importance of resilience, NATO has established the Seven

Baseline Requirements for National Resilience, which provide a common framework for assessing national preparedness among Allied nations. These requirements encompass: (1) continuity of government and essential public services; (2) resilient energy supplies; (3) the capability to manage uncontrolled population movements; (4) resilient food and water resources; (5) the capacity to manage mass casualties; (6) resilient civil communications systems; and (7) resilient transportation systems (NATO, 2023). Collectively, these baseline requirements provide the minimum standards necessary to ensure that civilian systems remain capable of supporting national defence and Alliance operations during periods of crisis and conflict.

To further strengthen Allied preparedness, NATO introduced the Resilience Objectives (formerly the Resilience Targeting Process) in 2023 on a voluntary basis. These objectives enable member states to assess national preparedness against common benchmarks while encouraging the adoption of comprehensive whole-of-government and whole-of-society approaches to resilience. Complementing these efforts, NATO has also developed the Enablement Concept, which focuses on ensuring the rapid movement and sustainment of Allied forces across the Alliance during peacetime, crises, and armed conflict. The concept seeks to eliminate legal, administrative, procedural, and infrastructure-related barriers that may delay military deployment by promoting closer cooperation among military authorities, civilian institutions, transportation providers, and host nations. Through improvements in military mobility, border-crossing procedures, transportation infrastructure, and host nation support, the Enablement Concept strengthens NATO's overall deterrence and defence posture while reinforcing national resilience (NATO, 2024).

The European Union has its own conceptual approach to building resilience, which includes: 1. Risk assessment and anticipation, 2. Prevention and preparedness, 3. Protection of critical infrastructure, 4. Cyber resilience, 5. Civil protection and crisis management, 6. Supply chain resilience, 7. Societal resilience, 8. Economic and financial resilience, 9. Climate adaptation and environmental resilience, 10. Governance and coordination, 11. International cooperation. Core principles of the EU resilience concept and associated documents (like the Network and Information Security (NIS2) Directive, the Critical Entities Resilience (CER) Directive, and the EU Strategic Foresight Reports) provide the normative framework for resilience among member and partner countries. This gives North Macedonia a set of parameters to benchmark against which to measure its own capacities, preparedness and resilience.

Although NATO and the EU have their own institutional systems and frameworks, their approaches to building resilience are increasingly complementary. NATO primarily emphasizes national preparedness through the implementation of the Seven Baseline Requirements, Resilience Objectives, civil preparedness, and the Enablement Concept within a whole-of-government and whole-of-society framework. In contrast, the European Union focuses on strengthening resilience through regulatory harmonization, critical infrastructure protection, cybersecurity legislation, economic governance, and regional cooperation. Despite these institutional differences, both organizations establish common baseline standards intended to strengthen the capacity of member and partner states to prevent, withstand, respond to, and recover from multidimensional crises.

For the Republic of North Macedonia, this dual institutional framework presents both significant opportunities and substantial challenges. As a NATO Ally and a candidate country for European Union membership, North Macedonia benefits from access to shared standards, technical expertise, financial assistance, and institutional best practices. At the

same time, the country must simultaneously satisfy NATO's resilience requirements and align its national legislation with the European Union's regulatory framework. Nevertheless, successful implementation of these complementary resilience frameworks offers a strategic pathway for strengthening national security, improving institutional effectiveness, and increasing the country's capacity to respond to future hybrid threats and complex crises.

3. INSTITUTIONAL APPROACH IN BUILDING NATIONAL RESILIENCE IN NORTH MACEDONIA

3.1 National Legal Framework in Building Resilience

National resilience is the ability of a country to maintain its institutions, crucial infrastructure, economy, and society in the face of crises, threats and disruption of various kinds. In today's security environment, resilience is not only how well a country reacts to crises but also how prepared and flexible it is in terms of long-term recovery and change (OECD, 2020; NATO, 2023). For the Republic of North Macedonia, national resilience is particularly critical, given the interplay of geographical, demographic, economic and security elements that characterize the country. North Macedonia, as a small state with limited resources, requires a comprehensive, integrated, and system-wide approach that incorporates the principles of whole-of-government and whole-of-society cooperation. Accordingly, this chapter critically examines the current state of national resilience in the Republic of North Macedonia by assessing institutional capacities, legislative and strategic frameworks, critical infrastructure protection, crisis management mechanisms, cybersecurity, energy security, and societal resilience.

Since becoming a member of NATO in 2020, while simultaneously pursuing accession negotiations with the European Union, the country has increasingly become exposed to hybrid activities aimed at undermining democratic institutions, reducing societal cohesion, weakening public confidence in Euro-Atlantic integration, and increasing political polarization (EEAS, 2024). When hybrid threats are in play today, geographical location is not only a geographical feature, but also a key driver of vulnerability and potential security risk.

One of the critical aspects of building a resilient system is the existence of a stable and efficient institutional structure. Security, defense, crisis management, and population protection are regulated in the Republic of North Macedonia by various laws, bylaws and policy papers. Central roles are performed by the Government of the Republic of North Macedonia, the Security Council, the Ministry of Defence, the Ministry of Interior, the Army of the Republic of North Macedonia, the National Security Agency, the Crisis Management Centre, the Directorate for Protection and Rescue, and institutions responsible for the regulation and protection of critical infrastructure. Although the normative framework is generally Euro-Atlantic in nature, some structural challenges remain. The most salient of these are the overlapping of responsibilities, questionable competencies, lack of horizontal coordination between institutions and weak analytical capabilities for strategic foresight and risk assessment and management. These are not organizational problems only; rather, they affect the state's capacity to deal with threats and challenges preemptively, to make fast decisions, and to recover.

Recognizing emerging security dynamics, the Government of the Republic of North Macedonia initiated a comprehensive process of assessing national vulnerabilities

through the EU Hybrid Risk Survey conducted in 2018 in cooperation with the European Union. The assessment identified critical vulnerabilities across several sectors, including governance, strategic communication, cybersecurity, critical infrastructure protection, crisis management, and inter-institutional coordination. The findings subsequently informed the development of the Strategy for Building Resilience and Countering Hybrid Threats (2021–2025), adopted in 2021 as the country's first comprehensive strategic document dedicated specifically to countering hybrid threats.

The Strategy represents a significant evolution in North Macedonia's national security architecture by adopting a whole-of-government and whole-of-society approach that aligns national policies with the European Union's six resilience domains and NATO's Seven Baseline Requirements for National Resilience. Rather than focusing exclusively on reactive crisis management, the Strategy emphasizes prevention, early warning, institutional coordination, strategic communication, cyber-security, protection of critical infrastructure, civil preparedness, and societal resilience. This reflects broader international recognition that effective counter-hybrid policies require integrated governance frameworks capable of coordinating multiple public and private stakeholders (Hybrid CoE, 2021; NATO, 2023).

However, the Strategy needs to have a clear roadmap and resources for implementation.

Despite the strategic importance of the adopted framework, implementation has encountered numerous institutional, organizational, financial, and political challenges. These include fragmented institutional responsibilities, limited human and technical capacities, insufficient inter-agency information sharing, uneven cyber-security maturity across public institutions, evolving disinformation campaigns, and the need for continuous adaptation to rapidly changing threat environments. Furthermore, effective implementation requires sustained political commitment, adequate financial resources, legislative harmonization, and stronger cooperation between governmental institutions and civil society organizations.

The complexity of hybrid threats also requires moving beyond traditional security sector reforms toward comprehensive governance reforms that strengthen democratic institutions, improve transparency and accountability, enhance media literacy, protect electoral integrity, and increase public confidence in state institutions. International experience demonstrates that resilient societies are characterized not only by strong security institutions but also by high levels of social trust, effective public communication, independent media, and active civic participation (OECD, 2020).

Against this background, this paper critically examines the development and implementation of North Macedonia's institutional framework for building resilience against hybrid threats. It evaluates the extent to which the Strategy for Building Resilience and Countering Hybrid Threats (2021–2025) has strengthened institutional capacities and identifies the principal implementation challenges that continue to limit its effectiveness. Drawing upon contemporary theoretical perspectives on resilience, official policy documents, NATO and EU frameworks, and comparative international experiences, the paper proposes evidence-based recommendations aimed at enhancing institutional coordination, strategic communication, cybersecurity governance, civil preparedness, and whole-of-society resilience. By situating North Macedonia within the broader European security architecture, the paper contributes to the growing academic debate concerning the governance of hybrid threats and the evolving concept of national resilience in small democratic states confronting increasingly complex and multidimensional security challenges.

3.2 NATO Approach and the National Resilience of North Macedonia

NATO plays a fundamental role in strengthening the security and resilience of the Republic of North Macedonia. Beyond providing the guarantees of collective defence under Article 5 of the North Atlantic Treaty, NATO offers member states a comprehensive framework for enhancing national resilience through the implementation of the Seven Baseline Requirements for National Resilience, the Resilience Objectives (formerly the Resilience Targeting Process), and the Enablement Concept. One could argue that all of the above-mentioned are, in a way, complementary to Crisis Management Requirements. Collectively, these framework documents support member states in developing robust civilian preparedness, strengthening institutional capacities, and improving the ability to prevent, withstand, respond to, and recover from a wide spectrum of contemporary security challenges (NATO, 2023, 2024).

To support the implementation of national resilience, NATO promotes two complementary governance models: the whole-of-government approach and the whole-of-society approach. The whole-of-government approach emphasizes coordinated action across public institutions and government agencies, ensuring policy coherence, effective crisis management, and integrated decision-making. The whole-of-society approach extends this responsibility beyond government by incorporating the private sector, academia, civil society organizations, critical infrastructure operators, the media, and citizens into national preparedness and crisis response. These complementary approaches acknowledge that resilience cannot be achieved solely through military preparedness but requires coordinated engagement across all sectors of society (NATO, 2023).

The strategic importance of resilience was further reinforced during the 2025 NATO Summit in The Hague, where Allied Heads of State and Government agreed to strengthen the Alliance's deterrence and defence posture through a new defence investment commitment. Under The Hague Defence Investment Pledge, Allies committed to allocating 5% of gross domestic product (GDP) annually to defence and security by 2035, including at least 3.5% of GDP for core defence expenditures and up to 1.5% of GDP for defence- and security-related investments (NATO, 2025a). The latter category encompasses investments in national resilience, including the protection of critical infrastructure, cybersecurity, civil preparedness, military mobility, innovation, resilience of supply chains, and the strengthening of the defence industrial base. This expanded understanding of defence investment reflects NATO's recognition that credible deterrence depends not only on military capabilities but also on the resilience of civilian institutions and essential services. The implementation and allocation of these investments are scheduled for review in 2029 in light of the evolving strategic environment and NATO's capability development requirements (NATO, 2025a).

Table 1. SWOT assessment of national resilience in the Republic of North Macedonia

Strengths	Weaknesses
● NATO membership ● An aligned strategic framework ● Developed institutional foundations ● Growing international cooperation ● A positive degree of societal stability	● Limited financial resources ● Insufficient coordination between institutions ● Vulnerability of critical infrastructure ● Limited cyber capacities ● Energy dependence ● An underdeveloped culture of resilience
Opportunities	Threats
● NATO and EU integration ● Digital transformation ● Regional security cooperation ● Investment in renewable energy sources	● Hybrid operations ● Cyberattacks ● Energy crises ● Disinformation campaigns ● Climate risks ● Geopolitical instability

The results show that the Republic of North Macedonia has some important qualities. The study of national resilience in the Republic of North Macedonia shows that the state has a solid institutional and some normative foundation for the establishment of a resilient security system. The current security challenges, however, require the transition from a sectoral approach to an integrated risk management strategy. The principal challenge identified by this research is not the absence of individual institutional capacities but rather the lack of an integrated national resilience system capable of coordinating these capacities within a coherent strategic framework. Addressing these deficiencies requires the establishment of an institutional architecture that promotes coordinated planning, clear legal binding, information sharing, strategic foresight, and unified crisis management across all levels of government. In the long run, the long-term security of the Republic of North Macedonia will depend not only on the ability to identify and eliminate threats but also on the capacity of the state and society to hold up in the face of heavy strategic pressure and complex situations.

The development of national resilience is not only a security or institutional matter, but also a strategic and political process that involves long-term vision, national consensus, and ongoing political commitment. Today's threats are dynamic, interlocking, and have cascading effects on the state, economy, and society. Such security threats are no longer just a piece of security policy but an aspect of national governance. From the perspective of the Republic of North Macedonia, the establishment of a national resilience system is a strategic imperative with multiple interconnected factors. Therefore, it is critical to have a unified strategic and political vision that will help to achieve a long-term plan so that the Republic of North Macedonia can be a resilient state in the decades to come.

In traditional security theory, strategic national interests, most often related to sovereignty, territorial integrity, and political independence, have to be upgraded. Within

the contemporary security environment, these goals don't exist without a solid foundation of national resilience. Resilience is a precondition for the realization of all other national interests since it makes the state able to function in the face of severe disruption. A state that cannot keep its core functions in the long run can no longer protect its security, economic, or political interests. In other words, national resilience could be recognized as an independent strategic national interest and integrated into all key state strategies and policies of the Republic of North Macedonia .

4. TOWARD A NATIONAL RESILIENCE SYSTEM BY 2035

From our current security environment and assessment of the existing capacity of the Republic of North Macedonia, we see that traditional sectoral approaches to security are no longer sufficient to address the complicated security challenges of the twenty-first century, with a permanent question as to who or which institution is actually responsible. Geopolitical instability, hybrid threats, cyberattacks, climate change, energy crises, and technological development all mean that resilience has to be considered as a systemic rather than as a sectoral problem. Hence, this research proposes to develop an integrated model of national resilience within the Framework for National Resilience System - NRS-2035 as a conceptual and strategic framework for the Republic of North Macedonia to cope with, overcome, and respond to complex security challenges in the future by 2035. It is not just a set of policies and institutions but it is a system that integrates state institutions, local self-government, the private sector, the academic community, and civil society, as well as the state system as a whole, and is one of national resilience, supported by standard operational procedures.

A fundamental prerequisite for successful strategic transformation is the articulation of a clear long-term vision. Accordingly, this study proposes the following strategic vision:

By 2035, the Republic of North Macedonia will establish an integrated, adaptive, and sustainable National Resilience System capable of ensuring the continuity of state functions, protecting critical national interests, and effectively managing complex security, economic, technological, environmental, and societal challenges through coordinated action across government and society.

This vision requires a fundamental transition from reactive crisis management toward proactive, risk-informed governance based on strategic foresight, prevention, preparedness, and adaptive decision-making. The realization of this vision is structured around five mutually reinforcing strategic objectives.

Strategic Goal 1 - Empowering resilient and adaptive institutions. Modern security challenges require institutions that can respond in a short time, make rapid decisions, and respond to crises effectively. In other words, a national system of risk assessment and strategic foresight is needed (OECD, 2020).

Strategic Goal 2 - Strengthening national security and defense preparedness. Defense is an essential part of resilience in cases of geopolitical instability and growing hybrid threats. Defense capacities, civil-military cooperation, cyber defense, and resilience to hybrid threats need to be strengthened further (NATO, 2023).

Strategic Goal 3 - Strengthening the resilience of critical infrastructure. The ability

to sustain critical infrastructure is key to the stability of a country. The strategic objective is to create a system for the protection and management of risks in all aspects of society.

Strategic Goal 4 - Social cohesion and trust. There is a direct correlation between the trust in institutions and the community's willingness to take action. So building a democratic culture, media literacy, and civic responsibility is a strategic priority in the long run.

Strategic Goal 5 - Building a culture of resilience. Resilience must become part of the national strategic culture. It means including the concept in education, public administration, the security sector, local self-government, and the private sector.

Building national resilience should be understood as a long-term political and strategic process that requires vision, leadership, and institutional maturity. Resilience is not only a response to the present threat but is in fact an investment in long-term stability, security, and development of the country's structure. The proposed National Resilience System seeks to reposition resilience from a sector-specific security concept to a core principle of national governance. This integrated model shifts the focus from reacting to individual threats toward strengthening the state's overall capacity to manage uncertainty, absorb shocks, and emerge stronger from crises. This way, North Macedonia will be able to build a system that will be judged not on its ability to handle crises but on its ability to overcome them and become stronger, more adaptable, and resilient.

One of the principal challenges in implementing national resilience lies not in technological or financial limitations but in ensuring sustained political commitment. In most states, strategies remain declarative documents because political will and institutional continuity are not available. For this reason, the NRS-2035 model must be implemented with political consensus and must be seen as a national issue of strategic importance.

Three enabling conditions are particularly important for successful implementation. First, institutional continuity should ensure that resilience policies remain stable irrespective of changes in government, thereby enabling long-term strategic planning. Second, legislative modernization is required to establish a coherent legal framework defining institutional responsibilities, coordination mechanisms, and resilience governance across all sectors. Third, financial sustainability should provide predictable multiannual investment supporting preparedness, critical infrastructure protection, cybersecurity, education, research, innovation, and institutional capacity building.

One of the major shortcomings of existing systems is that there is a fragmentation of institutional competencies. To address this deficiency, this study proposes the establishment of a National Resilience Council as the highest inter-institutional coordinating body responsible for strategic direction, national risk assessment, policy coordination, resilience monitoring, performance evaluation, and strategic planning. The Council should be the highest coordinating body responsible for strategic direction, coordination of institutions, monitoring of risks, assessment of the level of resilience, and proposing policies. Its composition would include representatives of the Government, security institutions, the academic community, the private sector, and civil society, thereby institutionalizing the whole-of-government and whole-of-society approach.

National resilience in the Republic of North Macedonia cannot be seen only as being national. It should be seen as a European and Euro-Atlantic issue. For example, NATO provides standards for resilience, collective defense, expert support, and operational interoperability. The European Union provides a normative framework, financial instruments, mechanisms for the protection of critical infrastructure, and digital and energy transformation support. So the future development of resilience should be based on maximum alignment with NATO's Baseline Requirements for Resilience and the relevant

policies of the European Union. The practical implementation of the NRS-2035 framework requires the adoption of a National Resilience Strategy of the Republic of North Macedonia (2026-2035) as the country's principal strategic document governing resilience policy. It should derive or be associated with the National Security Strategy because they are connected and interdependent in contemporary security. The strategy should include a national risk assessment, strategic resilience goals, institutional competencies, financial mechanisms, indicators for measuring progress, and mechanisms for review and updating. Such a strategy would provide an integrated governance framework capable of coordinating resilience-related policies across all sectors while ensuring consistency with NATO and European Union standards.

One of the key drivers of resilience over the long term is the establishment of a strategic culture that views security as a shared responsibility. History tells us that the states that successfully cope with prolonged crises are not necessarily the wealthiest or most powerful, but those that develop trust, solidarity, and institutional cohesion. In this sense, national resilience should gradually become part of the collective consciousness and national identity of the state.

4.1. Theoretical basis of the NRS-2035 Model

The NRS-2035 model is based on four fundamental assumptions:

1. The first is that threats are systemically interconnected. Today's threats seldom act in isolation. A cyberattack on an energy system can trigger an economic crisis, disruption of communications, and political instability. Thus, resilience has to be developed systemically rather than sectorally.
2. Second assumption: resilience is a national capacity. Security is no longer the sole responsibility of the armed forces and security institutions. It depends on the functionality of the entire state and societal system.
3. Third assumption: prevention is better than reaction. Security systems in the modern world should be oriented toward forecasting and reducing risks prior to their escalation.
4. Fourth assumption: resilience is a continuous process. There is no final state of resilience. It is a continuous process, always evolving with security.

4.2. Architecture of the NRS-2035 model

The National Resilience System 2035 (NRS-2035) is structured as a comprehensive, system-of-systems framework composed of seven mutually reinforcing and interdependent pillars. These pillars reflect the multidimensional nature of contemporary security challenges and are designed to ensure that resilience is embedded across all critical domains of state and societal functioning. The model is premised on the systemic assumption that the weakening or failure of any single pillar generates cascading effects that may compromise the overall resilience of the state (OECD, 2020; NATO, 2023).

The NRS-2035 model comprises seven mutually dependent pillars:

- Pillar I - Institutional Resilience
- Pillar II - Security and Defense Resilience
- Pillar III - Cyber Resilience
- Pillar IV - Critical Infrastructure Resilience
- Pillar V - Energy Resilience
- Pillar VI - Societal and Cognitive Resilience
- Pillar VII - Human Capital and Civil Preparedness

The basic logic of the model is that the failure of any of these pillars will ultimately have a knock-on effect on the resilience of the system.

Pillar I: Institutional Resilience It is institutions that are central to national resilience. The state cannot effectively handle crises without a functional, coordinated, and adaptive system of institutions. Institutional resilience means that governance is continuous, that the state administration is well managed, that the institutions collaborate to solve problems in a timely manner, that decision-making can be made quickly and in a timely manner, and that it has a system of risk management (OECD, 2020). The Republic of North Macedonia should establish a National System of Strategic Foresight and Early Warning by 2035 for the early warning of risks and for quick decision-making.

Pillar II: Security and Defence Resilience Today, it is not a replacement for defence but complements it. It refers to the state's ability to preserve security and territorial integrity under conditions of crises, hybrid threats, and armed conflict (NATO, 2023). Key priorities are: further modernization of the Armed Forces, development of capability for hybrid warfare, strengthening civil-military cooperation, integration into NATO collective defense, and development of national reserve capabilities. In the wake of current conflicts, the state cannot be left to its own devices and has to work within the larger national system.

Pillar III: Cyber Resilience - digital transformation has turned cybersecurity into a key aspect of national security. The Republic of North Macedonia should establish a National Cyber Operations Centre (NCC) by 2035 to track the situation and be able to detect all risks. It means building the ability to respond promptly in case of a cyber-attack, as well as to have specialized teams to respond to cyber crises. Particularly important is to protect critical infrastructure from cyber-attacks because these represent the most severe risks in the next decade.

Pillar IV: Critical Infrastructure Resilience is the foundation of a nation's resilience. The goal is to ensure the continuity of functioning of systems such as energy systems, transport corridors, telecommunications networks, water supply systems, health infrastructure, financial services, and all other sectoral aspects of critical entities resilience. The approach has to be based on the concept of an "all-hazards approach," which implies preparation for natural disasters, technological incidents, cyberattacks, and hybrid threats and hybrid warfare (OECD, 2020).

Pillar V: Energy Resilience - energy security is an important component of national resilience. The major goal is to decrease strategic dependence and to increase the resilience of the energy system. By 2035, the following should be ensured diversification of energy supply, increased domestic production, investment in renewable energy sources, regional energy connectivity, and protection of energy infrastructure. Energy resilience represents a bridge between national security and economic development.

Pillar VI: Societal and Cognitive Resilience - hybrid operations now aim to influence public perception, trust in institutions, and social cohesion (EEAS, 2024). Therefore, the resilience of the state is not possible without a resilient society. The cognitive resilience components are media literacy, resilience to disinformation, public trust in institutions, democratic culture, and social cohesion.

In present-day conflicts, perception is as important as physical space.

Pillar VII: Human Capital and Civil Preparedness human capital is the most important resource in any resilience system. No technology, institution, or strategy can replace the role of a prepared and educated population. By 2035, it will be necessary to undertake the introduction of national programs for security education, development of civil preparedness, strengthening of volunteer capacities, development of reserve structures, and

the integration of the resilience concept into the education system. This pillar represents a long-term investment in the state's ability to adapt to future challenges.

One of the weaknesses of most resilience strategies is the lack of measurable indicators. For this reason, the NRS-2035 model proposes the establishment of a National Resilience Index (NRI). The index would measure institutional capacities, cybersecurity, energy security, protection of critical infrastructure, societal cohesion, civil preparedness, and the capacity for crisis management.

The ultimate goal of the NRS-2035 model is not merely to improve the capacity for crisis management. Its essential aim is the transformation of the Republic of North Macedonia from a state that reacts to crises into a state that systemically anticipates, absorbs, and overcomes them. This transformation involves a change in security culture, modernization of institutions, and the establishment of a new societal consensus on national resilience.

The NRS-2035 model is the central concept of this study. It combines contemporary security theory with NATO resilience standards and the specific needs of the Republic of North Macedonia into a single strategic vision. The basic value of the model is not in its individual actions but in its system. National resilience can be achieved not through separate policies but by the interrelation of institutions, infrastructure, economy and society. NRS-2035, in such a way, is a road map to build a more resilient, more adaptive and more stable Republic of North Macedonia in the face of global uncertainty and the security challenges of the twenty-first century.

5. CONCLUSION

The research found that security in the twenty-first century is not a state of total absence of threats, but rather of the capacity of the state and society to survive and evolve in an environment of continuous disruption and uncertainty. Thus, national resilience is one of the most significant areas of contemporary security studies and is a key prerequisite for the security, sovereignty, and development of the state.

The theory of resilience has evolved, as security challenges and threats have evolved, in a big way, from a notion rooted in ecology and systems theory (Holling, 1973) to a core concept in the current security policy of NATO, the European Union, and other national security strategies. In the end, resilience is the capability of a system to anticipate, absorb, adapt to, and overcome various types of crises without losing functional and critical aspects of the system. The research also confirmed that in today's world of security, resilience is not another type of security but rather its new developmental phase. In the past, the models aimed to eliminate the threat, and now the focus is on reducing vulnerability and building the capacity for recovery and adaptation.

Empirical evidence from recent global and regional security developments reinforces this conceptual shift. The war in Ukraine, the European energy crisis, large-scale cyberattacks on critical infrastructure, disinformation campaigns, and the accelerating impacts of climate change demonstrate that contemporary threats are inherently systemic and transboundary in nature. These challenges are characterized by cascading effects that extend across political, economic, technological, and societal domains, thereby affecting the overall functionality of state systems (EEAS, 2024; NATO, 2022). In this context, the security of small states such as the Republic of North Macedonia is increasingly determined by their capacity to develop and sustain effective resilience systems capable of managing complex, multidimensional risks.

As a conclusion, we can say that there is a solid institutional and normative base on which the resilience system is to be developed today. NATO membership, crisis management mechanisms, and the establishment of security institutions are advantages and make for a very strong basis for the state to develop greater resilience. However, the paper revealed some structural weaknesses, which could affect the state's capacity to respond to future challenges. The most striking of all are the lack of inter-institutional coordination among the organizations, poor strategic foresight, the susceptibility of critical infrastructure to cyber risks, energy reliance, and the lack of a culture of resilience in society.

The most important conclusion of the research is that the greatest problem for the Republic of North Macedonia does not come from the lack of individual capability, but from the lack of integration of existing resources into a unitary system of national resilience. We have found that security institutions, public administration, local self-government, the private sector, and civil society are not part of a single system and do not function as a whole, but rather as separate segments that cannot respond to complex crises together. Through this research, we try to give birth to the concept of the National Resilience System - NRS-2035 model as the key scientific and practical development of this research. The model is based on seven interconnected pillars: institutional resilience, security and defense resilience, cyber resilience, critical infrastructure resilience, energy resilience, societal and cognitive resilience, and human capital and civil preparedness. The essential value of the NRS-2035 is its systemic logic. Rather than a one-dimensional review of sectoral policies, the model sees resilience as a national capacity that is built through the convergence of institutions, infrastructure, the economy, and society. This approach is consistent with the current NATO and European Union approach to security and development and takes into account the diverse nature of the Republic of North Macedonia.

The research also confirmed that the development of national resilience is the outcome of long-term political consensus and strategic vision. Resilience cannot be built on the basis of isolated institutional activities but rather must be built on the basis of political participation, financial sustainability, legislative alignment, and participation of society as a whole. From the perspective of national security, we can say that national resilience is a crucial factor in the integration of the Republic of North Macedonia into the Euro-Atlantic security structure. In an environment of ever-more uncertainty, resilience is a bridge between national security, economic development, democratic stability, and societal cohesion. Therefore, it is clear that the security of the Republic of North Macedonia is not only in terms of defence capabilities and international cooperation but also in the state and societal system in which it will be able to prepare, cope, and recover from future crises. In this sense, national resilience is not just a security policy, but also a key to the survival, stability, and development of the state in the twenty-first century.

6. REFERENCES

- Bailes, A. J. K., Rickli, J.-M., & Thorhallsson, B. (2007). Small states, survival and strategy. Centre for Small State Studies, University of Iceland.
- Boin, A., Ekengren, M., & Rhinard, M. (2021). Understanding the creeping crisis. Palgrave Macmillan.
- Comfort, L. K., Boin, A., & Demchak, C. C. (Eds.). (2010). Designing resilience: Preparing for extreme events. University of Pittsburgh Press.
- European Commission, & High Representative of the Union for Foreign Affairs and Security Policy. (2016). Joint framework on countering hybrid threats: A European Union response (JOIN(2016) 18 final).
- European Commission. (2022). Proposal for a directive on the resilience of critical entities (CER Directive). Publications Office of the European Union.
- European Commission. (2023). EU strategic foresight report 2023. European Union.
- European Union. (2022). The Strategic Compass for Security and Defence. Publications Office of the European Union.
- European Centre of Excellence for Countering Hybrid Threats. (2021). Hybrid CoE strategy 2021–2023. <https://www.hybridcoe.fi>
- European External Action Service. (2024). *Second EEAS report on foreign information manipulation and interference (FIMI) threats*. <https://www.eeas.europa.eu>
- Government of the Republic of North Macedonia. (2021). Strategy for Building Resilience and Countering Hybrid Threats (2021–2025). Government of the Republic of North Macedonia.
- Hoffman, F. G. (2007). *Conflict in the 21st century: The rise of hybrid wars*. Potomac Institute for Policy Studies.
- Holling, C. S. (1973). Resilience and stability of ecological systems. *Annual Review of Ecology and Systematics*, 4, 1–23. <https://doi.org/10.1146/annurev.es.04.110173.000245>
- Linkov, I., & Trump, B. D. (2019). The science and practice of resilience. Springer. <https://doi.org/10.1007/978-3-030-04565-4>
- Murray, W., & Mansoor, P. R. (Eds.). (2012). Hybrid warfare: Fighting complex opponents from the ancient world to the present. Cambridge University Press.
- North Atlantic Treaty Organization. (2022). NATO 2022 Strategic Concept. <https://www.nato.int/strategic-concept/>
- North Atlantic Treaty Organization. (2023). Resilience and Article 3. https://www.nato.int/cps/en/natohq/topics_132722.htm
- North Atlantic Treaty Organization. (2024). Enablement. https://www.nato.int/cps/en/natohq/topics_217625.htm
- North Atlantic Treaty Organization. (2025a). The Hague Summit Declaration. <https://www.nato.int>
- North Atlantic Treaty Organization. (2025b). Defence investment plan and defence- and security-related expenditure. <https://www.nato.int>
- Organisation for Economic Co-operation and Development. (2020). *The resilience of critical infrastructure systems: Analytical framework and policy recommendations*. OECD Publishing.
- Wigell, M. (2019). Hybrid interference as a wedge strategy: A theory of external interference in liberal democracy. *International Affairs*, 95(2), 255–275.
- Walker, B., & Salt, D. (2012). Resilience practice: Building the capacity to respond to disturbance and remain active. Island Press.

CONCEPTUALIZING RADICALIZATION IN SECURITY SCIENCES: THEORETICAL APPROACHES, METHODOLOGICAL CRITIQUES AND MUU IDEOLOGY

Frosina Tashevska-Remenski

Faculty of security – Skopje

frosina.remenski@uklo.edu.mk

Zlate Dimovski

Faculty of security – Skopje

zlate.dimovski@uklo.edu.mk

Abstract

The paper comprises two interrelated parts. The first part aims to set an optimal conceptual framework for understanding and researching radicalization in security sciences, and thereby to assist in academic and practical work on issues of extremism, fundamentalism, conspiracism, fanaticism, terrorism and counter-terrorism.

To this end, radicalization will be explained using four main approaches in the literature (monistic-pluralistic and absolutist-relativist). The second part of the paper aims, through the conflict and radicalization theory of McCauley and Moskalenko – their gradation of radicalization from individual to mass and the claim that the link between extreme belief and extreme action is weak – to discuss other theoretical positions that criticize the limited methodology of their claim and its scope. Instead of a conclusion, through these two parts of the paper we will attempt to present the MUU (mixed, unstable and unclear) ideology of the UK Home Office, which provides a response to forms of radicalization – preoccupations with forms of mass murder, non-ideological hatred of groups and varying degrees of illegal behavior, which cannot be categorized into recognized ideologies.

Keywords: radicalization, extremism, fundamentalism, terrorism, MUU ideology, theories, research;

INTRODUCTION – ON THE CONCEPT OF RADICALIZATION

Radicalization as a term is ubiquitous in academic work, in security science, and in research concerned with the study of extreme beliefs and extreme behaviors. The term is found in empirical, historical, and theoretical studies of fanaticism, extremism, fundamentalism, terrorism and counter-terrorism, conspiracy theories, insurrectionism (anarchism), radicalism, radicalization, and de-radicalization, prevention, and social resistance. Radicalization as a term is relatively new, especially when compared to terms such as fundamentalism, extremism, terrorism, and even radicalism.

There are more than 80,000 referenced papers on radicalization. Influential in this area is the study of radicalization by Walter Laqueur in 2004, who argued that in explaining the turn to extremist and terrorist violence, researchers should move away from political and structural conditions and turn to individual psychological character traits, especially those related to theological ideology, and thus to the temporal process of

radicalization of particular subjects.

Conceptualizing radicalization is important for a number of reasons, but we would single out three. First, it can bring clarity and precision to public debate. The term has traditionally been reserved for Islamism and right-wing extremism in Europe and North America, but today it is also used for conspirators fighting against many other contemporary phenomena and movements. Proper conceptualization can help identify the limits of the concept. Second, the term is widely used in academic literature, but there is much ambiguity about its exact relationship to violence and to closely related phenomena such as fundamentalism, extremism, terrorism, fanaticism, populism, and unconventional forms of political mobilization. Third, to be fruitful in qualitative and quantitative research in the security sciences, we need an analysis of radicalization that can be operationalized. Many existing conceptualizations of radicalization do not meet this criterion. Fourth, the concept of radicalization has been criticized for various reasons, such as that it is a tool to politicize certain minorities. Conceptualization could help address these concerns.

Various authors have addressed the drivers and frameworks of radicalization.¹ Others have explored the relationship of radicalization to phenomena such as conventional and unconventional forms of political mobilization² and reactionary politics and populism.³ Third, others explore various empirical aspects of radicalization, such as gender roles⁴, global and local characteristics,⁵ or specific variants of radicalization, such as right-wing radicalization.⁶ Other researchers are exploring responses to radicalization at the state level and civil society responses.⁷ In the paper we make an attempt, first of all, to answer the preliminary questions "does radicalization exist?" and "is it a concept that can be used?". Most of the studies are empirical in nature, but this paper is theoretical in nature and in part philosophical in nature, although it considers empirical work on radicalization. The paper addresses the issue by developing a conception of radicalization that combines the main approaches in the literature, the so-called monistic and pluralistic approaches, as well as absolutist and relativistic approaches.

¹ Daniele Albertazzi, Stijn van Kessel, Adrian Favero, Niko Hatakka, Judith Sijstermans, Mattia Zulianello, *POPULIST RADICAL RIGHT PARTIES IN ACTION (The Survival of the Mass Party)*. Oxford University Press, 2025;

² *The Impact of Populism on the Party System: The Experience of European Democracies* (2026): Comparative analysis co-edited with Jiří Kocián and Michal Kubát, 2026;

³ A Kaya, *Islamisation of Turkey under the AKP rule: Empowering family, faith and charity*, South European Society and Politics, 2015; Islam, migration and integration: The age of securitization, Springer, 2009;

⁴ *Nativism, Islamophobias and Islamism in the Age of Populism: Culturalization and Religionization of what is Social, Economic and Political in Europe*, European Research Council (ERC), ISTANBUL BILGI UNIVERSITESI, 2023;

⁵ *Countering Violent Extremism and Global Counter-Terrorism: A Decolonial Approach*, Edited by Sagnik Dutta, Tahir Abbas and Sylvia I. Bergh, Manchester University Press, 2025; Olusola, Jeremiah A. (2022) Reading an Islamic epistemology into research: Muslim converts and contemporary religion in Britain, *Journal of Religious Education*, 70; ⁶ Koca, Metin, Why the French bill against "Islamist separatism" goes beyond France, *Islam and the Separation of Religion and State*, Istanbul Bilgi University, 2020;

⁷ Chrona, Stavroula, Tereza Capelos, *The civic and political participation of young people in a context of heightened authoritarianism. The case of Turkey*, Oxford Brookes University, Department of Social Sciences, 2017;

The four approaches can be combined in such a way that their strengths are considered and their weaknesses are circumvented.

The concept of radicalization should be conceived in such a way that it is close not so much to social usage (which can vary considerably), but to academic usage. It cannot deviate too far from the way the word is commonly used in the radicalization literature, as it would result in the introduction of a new concept, not an improvement on an existing concept. In this section, we will formulate five desirable things that can be drawn from the current academic debate on radicalization.

First, it is widely accepted that there are two types of radicalization: cognitive and behavioral.⁸ Other authors have made a similar distinction, but have used different terms, such as "radicalization of beliefs" and "radicalization of methods."⁹ Cognitive radicalization occurs when someone develops radical ideas and begins to embrace extremist ideologies. It may come without a propensity to act on those beliefs. Behavioral radicalization occurs when someone engages in extreme, particularly extremist behavior, such as the use of violence. A person can hold a variety of extreme beliefs without engaging in extreme behavior, because they prefer to keep them to themselves (this is true of various highly educated neo-fascists). A person may engage in radical behavior even if they do not have proper beliefs, e.g., because they do so out of admiration for the leader of the group or simply because of various so-called grievances. There may be other ways of radicalizing than just cognitive or behavioral, but it can be said that the explanation of radicalization, academically, can be used by these two forms of radicalization.

Second, violence is an important expression of behavioral radicalization, but it is not the only expression. One can (increasingly) disadvantage sexual, gender, ethnic, or religious minorities. "Things Can Be Considered Radicalization"). A person can exclude all opposing voices from his group by no longer allowing them to enter his home or by excluding them from a church or mosque¹⁰. Any credible concept of radicalization should take this into account.

Third, it should leave room for the so-called "problem of the few"¹¹. In fact, there are two problems of the few: only a few who are exposed to an extremist ideology are cognitively radicalized, and only a few who are cognitively radicalized are also behaviorally radicalized. The conceptualization of radicalization should be compatible with this.

⁸ Meloy, J. Reid; Hoffmann, Jens (Eds.): *International Handbook of Threat Assessment*, (2nd ed.). New York: Oxford University Press, 2020;

⁹ Quassim Cassam, *Extremism A Philosophical Analysis*, University of Warwick, UK, 2019;

¹⁰ Clycq, Noel et al. (Eds.): *Radicalization: A Marginal Phenomenon or a Mirror to Society?* Leuven: Leuven University Press, 2019;

¹¹ Canter, David (Ed.): *The Faces of Terrorism: Multidisciplinary Perspectives*. Chichester: Wiley-Blackwell, 2009, 167;

Fourth, the two main lines of thinking in the analysis of radicalization are the so-called monism and pluralism (either universalism and contextualism, or generalism and particularism).¹² Monism says that there is a single nucleus that is common to all radicalization processes. In other words, certain cases of radicalization across radically different extremist movements, such as Salafism, jihadism, right-wing extremism, or the radical left, have something in common. Pluralism, however, says that there are many different processes of radicalization and that they often have little or nothing in common, no core that is shared by all. Another important difference here is that between absolutism and relativism. Absolutism says that there are absolute truths about whether someone has radicalized. Relativism, on the other hand, says that whether someone has radicalized is always relative to a certain set of values or principles; there is no absolute or objective or mind-independent truth about it. These positions involve diametrically opposed pairs, but there are plausible ways to understand them so that they will prove to be compatible.

Fifth, the work on the concept of radicalization is empirical, qualitative, and quantitative. It is secondarily historical or theoretical (conceptual, normative, etc.).¹³ The appropriate concept of radicalization should be suitable for fruitful use in empirical research in the field of security sciences, and this means that it should be operationalized, i.e. the abstract concept of radicalization should be analyzed in such a way that it can be translated and measured by concrete observations, e.g. in survey questionnaires or interviews.

Mark Sedgwick, in his discussions of modernity and traditionalism, argues that the concept of radicalization is used in three distinct contexts, each with its own characteristics.¹⁴ In the security context, the idea is that anarchists and terrorists can pose a threat to public security, which is why the term radicalization is used. They are gradually becoming a threat to public safety. In an integration context, radicalized persons are those who oppose things such as republican principles, fundamental democratic values, active citizenship, co-citizenship, or acceptance of homosexuality. The reason for using the concept of radicalization here is different: it is considered to be an obstacle to the proper integration of newcomers into that particular society. Certain types of opposition in one's own country can be labeled as protests or violence by radicalized people, in order to justify, before the international community, taking repressive measures against them. Here, the reason for using the concept of radicalization is to justify one's own policies towards a particular segment of the population. Since there is no common core for what is considered "radicalized" in these three different contexts, we should either abandon the term or contextualize it extensively.¹⁵

¹² de Bruijn, Mirjam (Ed.): *Biographies of Radicalization: Hidden Messages of Social Change*, Berlin: De Gruyter Oldenbourg, 2019, 67-75;

¹³ Goerzig, Carolin, Al-Hashimi, Khaled: *Radicalization in Western Europe: Integration, Public Discourse and Loss of Identity among Muslim Communities*, (Contemporary Terrorism Studies), Abingdon: Routledge, 2014, 36-55;

¹⁴ Mark Sedgwick, *Against the Modern World: Traditionalism and the Secret Intellectual History of the Twentieth Century*, Oxford University Press, 2004, 18-32;

¹⁵ According to Sedgwick, the term "radicalization" tends to mean something different depending on the context: the security context, the integration context, and the foreign policy context. Shouldn't this lead to an even stronger contextualization of the term radicalization? Not necessarily. This is because there is an alternative and more likely

Arun Kundnani shows in detail how the term "radicalization" has been used and misused for a number of political purposes. The process of radicalization has often been applied only to Muslims who embrace extremist ideologies, and Kundnani sees this as a clear expression of deep-rooted racism.¹⁶ Among other things, it has wrongly led to mass surveillance among Muslim minorities in the United States and other countries. Kundnani rightly points out that funding (cash flows) and ideological assumptions have had a seriously distorting effect on the use of the concept of radicalization, and perhaps he is right that Walter Laqueur's approach to radicalization was particularly influential here.¹⁷ It seems that any appropriate concept of radicalization should be completely neutral with respect to ethnic and religious issues. Any concept that implies the slightest dose of racism should be firmly rejected.¹⁸

understanding of what happens in these three cases. In other words, it focuses on things that have to do with democratic or Western (or Arab, Islamic, Persian, Asian, etc.) values, human rights, equal education, separation of church and state, and so on. It focuses on security issues: the risk of actual violence, support (financial, administrative, theological) for violence, and various beliefs that increase the risk of using or supporting violence. It seems quite correct to say of a security sector official that someone who denies equal rights to education for men and women has been radicalized, but that is not the type or part of radicalization that interests him. However, you will notice that the reverse will not apply: from the perspective of integration, one will be interested in everything that interests the security official: the threat to public security is also a threat to integration. The foreign policy perspective seems to be reduced to an integration or security perspective: if one takes measures of suppression against a part of one's own population it will do so in the name of integration or public safety. We can recognize these three contexts and retain the concept of radicalization.

Ibid., 48-51;

¹⁶ Groothuis, S., *Researching Race, Racialization, and Racism in Critical Terrorism Studies: Clarifying Conceptual Ambiguities*. *Critical Studies on Terrorism*, 13, 2020, 680-701;

¹⁷ Walter Laqueur, *No End to War: Terrorism in the Twenty-First Century*, Oxford University Press, 1999, 40;

¹⁸ Arun Kundnani, in his debates on radicalization, is concerned that the concept of radicalization only serves for governments to interfere in the lives of their citizens, especially Muslims. He argues that this is especially true of theological explanations of radicalization: "The problem is that, if there is no real reason to think that these radical religious beliefs are associated with terrorist violence, the model of theological radicalization, then, only legitimizes the unjustified interference of the state in the private religious lives of a large number of citizens." Following his revelation, Derek Silva argues that "the very concept of radicalization contributes to discriminatory government policies targeting diverse Islamic communities". This has to do with a number of factors, one of which is that much of the research on radicalization is directly funded by the government. According to Silva, this has led to a lack of funding for research that is critical of the government's current and past counter-radicalization practices, status quo surveillance, and various intelligence and police strategies. This has led to a focus on research aimed at identifying the theological, cultural, psychological, and social characteristics of those who radicalize. This, in turn, has led to a preoccupation with models of radicalization based on indicators in mass media.

In response, it should be pointed out that the concept of radicalization set forth in the previous section neither says nor implies anything about specific extremisms, nor about the reasons why people radicalize, nor when or how it is legitimate to intervene, nor about what power relations exist between radicalized and the various factors and individuals that contribute to the process of radicalization. It is true that various concepts of radicalization have sometimes served

There is another challenge that is reflected in the idea that radicalization can be a so-called essentially contested concept, a term first introduced by Gallie¹⁹. Something is an intrinsically contested concept if its actual use comes with endless disputes about what its proper use is, and if such disputes cannot be resolved by empirical evidence, linguistic usage or logic. Examples include art, religion, power, and philanthropy. The ongoing debate over what radicalization is and whether we should use the term at all suggests something in this direction.

Monism, pluralism, absolutism, and relativism – approaches

In other words, there is a common core for all processes of radicalization, whether it takes place differently across fundamentalist, extremist, and terrorist movements, such as left-wing extremism, jihadism, neo-Nazism, and Hindu extremism. The alternative is pluralism. Quassim Cassam argues that there is probably no such thing as a "process of radicalization." Most likely, there are only many different types of radicalization processes. If each case is examined, one will find remarkable differences in almost every aspect in which the subjects radicalize: socio-economic background, political circumstances, theological and other beliefs, religiosity, grievances, experienced and/or perceived injustice, attitude towards the group, etc.²⁰ The conclusion we can draw from this is that in conceptualizing radicalization, two important questions should be taken into account: "Is it true that certain cases of radicalization differ theologically, culturally, economically, socially, politically, and ideologically, and if so, what does that mean in relation to the question of whether there is such a thing as a process of radicalization or just a number of different processes of radicalization?" According to the absolutist approach, whether someone is radicalized does not depend on someone else's beliefs or values. For example, Sophia Moskalenko and Clark McCauley compare radicalization to activism. They define "activism" as "the willingness to engage in lawful and nonviolent political action", while "radicalism" – as something fundamentally different from radicalization – is defined as "the willingness to engage in illegal and violent political activities."²¹ The idea here is that radicalism is engaging in illegal and

as a vehicle for racism. Therefore, we should always remain vigilant about how and for what purposes the term is used. However, this should not blind us to the fact that there is probably a process of radicalization in some people. The words "truth", "democracy", "freedom" and "knowledge" are also often misused for a number of purposes, yet few people question whether these things exist or are important. In fact, the same is true of various other personal (partly psychological) processes, such as conversion and intellectual growth. These terms can be misused, but it seems difficult to deny that they denote real phenomena. In Michele Grossman, H. A. Hellyer, *Rethinking Religion and Radicalization: Terrorism and Violence Twenty Years After 9/11*, European Union's Horizon 2020 research, 2020;

¹⁹ Walter Bryce Gallie, *Essentially Contested Concepts*, Faculty of Humanities and Social Sciences of the University of Jyväskylä, 2021;

²⁰ Quassim Cassam, 89-90;

²¹ Sophia Moskalenko and Clark McCauley's, "Measuring Political Mobilization: The Difference Between Activism and Radicalism", *Terrorism and Political Violence*, 21, 2009;

violent acts. The views on radicalization of Sophia Moskalenko and Clark McCauley will be discussed later in the paper.

For the relativistic approach, radicalization is a matter of representing or supporting an extreme part of society, and is always relative to a particular belief system or value system of a particular group or audience regarding what should be considered moderate and what is extreme. However, relativism raises its own concerns. For example, "Is it just a matter of perspective whether someone is radicalized?" The phrase "one person's terrorist is another person's freedom fighter" is well known, but many find it problematic. It seems to undermine any indisputable justification for preventing and combating radicalization.

The Theory of Conflict and Radicalization of McCauley and Moskalenko

Clark McCauley and Sophia Moskalenko are the authors of the 2011 book *Friction: How Radicalization Happens to Them and Us*, which received a second edition in 2016 with the title *Friction: How Conflict Radicalizes Them and Us*, which serves as the basis for absolutism as an approach to studying the concept of radicalization. The authors have academic expertise in social psychology and have conducted extensive research in the areas of terrorism and politics. In their book, McCauley and Moskalenko define twelve mechanisms of political radicalization that can persuade individuals, masses, and groups to increase their propensity and support for political violence and thus facilitate their entry into radical circles. The book uses cases to show parallels between nineteenth-century Russian anti-tsarist terrorism (with a look at the Russian populist organization founded in 1879, *Narodnaya Volya*), anti-war terrorism of the 1970s, and 21st-century jihadist terrorism. Offering us interesting insights into terrorism as an extreme form of radicalization, both authors explain their idea by rejecting some common assumptions about terrorists. They argue that terrorists should not be characterized as diabolical or insane, given the precision and accuracy required for each attack that requires lucid and careful individuals. At the same time, it is not accurate to characterize them as abnormal, as they are normal people who lead a regular life and possess strong beliefs and passions. However, at some point in their lives, probably after a specific triggering event during adolescence or adulthood, these individuals begin to act on those strong and specific beliefs. Therefore, radicalization is not something that happens only to "others", i.e. evil or mentally ill people. Rather, it is a psychological trajectory that, under certain circumstances, can happen to any person or group. In this regard, McCauley and Moskalenko insist on the need for a change in mindset in order to properly understand radicalization. They themselves stated that it is essential to recognize that those who enter must necessarily leave the neat and comfortable world where normal people do not do terrible things.

The most frequently asked question about terrorism concerns what drives individuals to engage in such aggressive and risky radicalization. Both authors agree that ideology is not an appropriate response, in line with other theoretical approaches and authors emphasizing ideology. In this regard, in the first part of the book, McCauley and Moskalenko refer to Wahhabism and Salafism, two fundamentalist and extreme religions that are widespread but do not encourage terrorism, to emphasize that ideology is not always the cause of

radicalization. In fact, their extremist values do not imply violence against innocent citizens. Although numerous religious writings reveal enigmas that can be taken as a pretext for the use of violence, this does not mean that ideology in general or religion is the cause of radicalization. Therefore, a set of values alone does not constitute a factor capable of causing violent acts to be committed. Essentially, all of this breaks the link between ideology and radicalization.

On the basis of this premise, McCauley and Moskalenko in "Friction" identify three levels of radicalization, given that the phenomenon is very complex, in order to provide a comprehensive illustration. The three levels are: individual radicalization, group radicalization, and mass radicalization.

Starting from the level of individual radicalization, the authors list six mechanisms. The first mechanism is personal resentment and is triggered by intimate and private experience. It occurs when an individual perceives that he has been wronged and, consequently, in addition to seeking justice, longs for revenge. The difference between the two is subtle but important: justice implies that those who have harassed must be punished, whereas revenge means that we ourselves have to punish the perpetrator.²² Although the individual seeking revenge personally usually aims to punish an individual, his vindictive impulse can extend to that individual's entire class because of the feeling of anger it arouses. To better understand the problem, the authors offer several specific ²³examples.

The second mechanism at the level of individual radicalization is group resentment, in which the individual identifies with the grievances experienced by groups of people to whom he or she does not even belong. In these cases, the process of radicalization is triggered by grievances caused by others. Among the reasons offered to explain this mechanism is strong altruism, that is, the tendency to help and cooperate with those we assume to be honest individuals. In fact, when we perceive evil, helping good people is an instinctive human attitude. A key aspect is group identification, the consequences of which can be beneficial or harmful. It can be positive, because one feels emotionally enriched through the well-being of others, but it can also be negative, because one rejoices in the problems that other groups have suffered. Therefore, identification with a group coincides with the development of deep empathy, which on certain occasions can lead to action. This

²² Clark McCauley and Sophia Moskalenko, *Friction: How Conflict Radicalizes Them and Us*, Oxford University Press, 2016.

²³ The first example concerns the story of Russian revolutionary Andrei Zhelyabov, whose sense of revenge began as a strong personal desire for justice against a landowner guilty of abusing his aunt. But when other local landowners decided to protect their comrade from guilt, Zhelyabov's anger shifted from one individual to an entire class of landowners. It was this need for revenge that led him to join the People's Will organization. The second example concerns the case of the French politician Fadela Amara, appointed Secretary of State in 2007, in whom we can find, even a few years later, the same need for revenge. Amara described racism towards her mother as a deciding factor in her anger. Her early efforts to combat perceived injustice led Amara herself to express her personal outrage in the more general sense of group outrage, which led her to fight against the discrimination suffered by Muslim immigrant women. However, it should be noted that although her ideas were illegal and radical, they never led to violence. On the contrary, her thoughts testify to how such an individual, painful, and humiliating episode can easily lead to extreme ideals without resorting to violence.

happens because emotional involvement in a group can make the group's resentment feel like it is their own, and thus prompt them to take action.²⁴

The third mechanism, known as the slippery slope, is particularly important because it highlights the concept that joining an extremist group is a gradual and not always intentional process. In fact, individuals do not join the group with the explicit purpose of using violence, nor do they begin to engage by taking the lives of others. Inevitably, they end up engaging in such aggression and becoming terrorists after being involved in a series of situations that have led them to continue.²⁵ Over time, the demands of the group become more stringent, and the individual tends to agree, although the level of violence that is required progressively increases. This happens because the individual does not see any threat if he goes a little further than before. In short, the meaning of the slippery slope is to suggest individual inclusion as progressive and spontaneous escalation. The fourth chapter of "Friction" offers the example of Adrian Mikhailov. Although he was a seemingly decent person with no revolutionary ambitions, he eventually joined the Russian terrorist complex. He was inspired by the slogan "People's Will", meaning "entering the people", which advocated the importance of connecting with ordinary people, such as peasants and workers, in order to advance the revolutionary cause. Mikhailov joined the group and was never able to leave it since, as each move legitimized the next one for the realization of a common cause.

A similar path was experienced by Omar Hammam, a British citizen of the Christian faith, who slowly became radicalized until he became a militant in a terrorist group. He began to explore his Syrian origins, but eventually embraced extreme ideals, a process that did not happen all at once. After his travels in Syria, Egypt, and then Somalia, he came into contact with people who shared an extreme Salafist vision, and this led him to acquire a radical vision of religion. Essentially, according to the authors,²⁶ terrorist groups rely on the power of the "slippery slope" to gradually lead new members to violence. This happens because this power can push individuals to opinions and actions they never imagined in the initial stages of the process. As a result, attempts to understand radicalization in terms of rational decision-making are limited: some terrorists, looking back, are unable to determine the moment when they chose to become terrorists. One step leads to another, making the process inevitable in retrospect. In the case of Mikhailov and Hammam, for example, neither of them could have predicted the end of the road that awaited them when they took their first steps, yet the two continued on that path until they became active members of terrorist groups.

²⁴ To support their explanation, McCauley and Moskalenko cite the example of Russian revolutionary Vera Zazulich. This is an important demonstration of the mechanism of group grievance. She was the first woman to be tried for terrorism. Deeply identifying with the students imprisoned during the revolutionary protests of 1877 and opposing the brutality committed in prison, she decided to plan violent attacks on the governors. Later, Zazulich joined Volonta del Popolo, vigorously fighting against the abuses and arrogance of the governors themselves.

²⁵ Alessandro Orsini, What Everybody Should Know about Radicalization and the DRIA model, in "Studies in Conflict and Terrorism", 46/2023, 14;

²⁶ Clark McCauley and Sophia Moskalenko

The fourth mechanism is shown to be love. Affection for a partner, friend, or family member is a common motive that drives individuals to join a terrorist group. Once within the group, relationships based on trust and love can encourage members to act directly. For example, Sofia Perovskaya, one of the leaders of the Russian terrorist group Narodnaya Volya, initially opposed the idea of using violence to support the organization's cause. However, after she fell in love with Andrei Zhelyabov, her opinion changed radically. Her emotional devotion to him grew, as did her desire to follow him in every initiative, which thus became considerably stronger than her initial opposition to violence. Consequently, Perovskaya found herself participating in the organization's activities, including assassination attempts, driven by deep loyalty to her first love, who was also one of the group's most radical leaders.

Regarding family networks, we know that the phenomenon of terrorism between siblings is also quite widespread. This is the case with the 2016 attacks in Brussels, the 2015 Paris massacre, and the 2013 Boston massacre. The cases mentioned above had in common, in addition to the signature of Islamic terrorism, the fact that some of the attackers were brothers. Going further back in time, a similar reference can be made to the 2002 Bali attack, in which the elder brother Ali Guffron engaged his two younger brothers, including Amrozi Nurhasim. The latter, in fact, became radicalized precisely by the need to be with his brother and get his approval. Although he had no extreme aspirations and never experienced any resentment that would have led him to commit violent acts, his sense of brotherhood was so strong that it made him content to become a terrorist. In short, the need to defend and support those closest to you can be, for some, an instinct that leads to unexpected actions.

In particular, especially among the lower classes and younger generations, there are people who tend to seek social charisma and intense emotions and excitement. Risking one's own life and taking responsibility for violent and dangerous episodes is proving to be an ideal lifestyle for certain people. Terrorist groups undoubtedly offer such opportunities. For these individuals, staying low in the group isn't glamorous; on the contrary, it becomes important to be able to perform risky and dangerous actions. Thus, with such desires, political terror becomes the perfect possibility. Based on these considerations, McCauley and Moskalenko offer, among other examples, that of Abu Musab al-Zarqawi. He was initially a petty criminal, jailed for shoplifting, drug trafficking, and attempted rape. Although the authorities criticized his behavior, he had just as many followers who both feared and respected him. In prison, al-Zarqawi found the status he sought, establishing himself as the most powerful of all through his fierce aggression towards his rivals and his constant resistance to the prison authorities. After serving his prison sentence, he traveled to Afghanistan, where he was able to quench his thirst for violence and adrenaline in the ongoing wars. During his stay, he met bin Laden and swore allegiance. However, during the U.S. offensive against Saddam Hussein's regime in Iraq, numerous kidnappings, beheadings, and violent attacks on other Muslims were soon reported by al-Zarqawi himself, who ignored al-Qaeda's rebukes and revulsions.²⁷ All because for him, as for all individuals belonging to this category, the supreme desire was the power and excitement of brutality.

²⁷ McCauley, C., & Segal, M. (1987). The social psychology of terrorist groups. In C. Hendrick (Ed.), *Group processes and intergroup relations: A review of personality and social psychology*, SAGE, 2017, 231–256;

For some people, the path to radicalization is hampered by a series of duties and responsibilities related to everyday life, such as supporting family, professional advancement, or maintaining close ties with friends. However, there are occasions when these same people unexpectedly lose those previous ties. Civil wars break out, destroying jobs, families, and social networks, pushing individuals into unexpected realities. In other cases, people decide to move to a distant city or foreign country, starting a new life from scratch. Consequently, the present circumstances leave the individual alone and deprived of his previous daily routine, and make him an easy target for any group that offers friendship, safety and protection, as well as an ideology that restores order and meaning to his fractured life. In these circumstances, therefore, what determines inclusion is fear. The case study in this chapter is Mohammed Bouyeri, the assassin of Dutch director Theo van Gogh in 2004, who presented his film "Submission", which was critical of Islam. Bouyeri's life was filled with interruptions; in particular, he spent several years in prison, distancing himself from the relationships he had. Then he lost his mother, to whom he was very close, to a serious illness, and he lost his job. Subsequent events brought him a sense of loneliness and disconnection from the surrounding reality, creating in him a progressive need to build a new social network, breaking with the past. His need was met only when he came into contact with the so-called "Hofstad Group", which was a novelty, as it illustrated how the nature of the Islamist terrorist networks operating in Europe had evolved.²⁸ Without wanting to delve into the subject, it is worth noting that this is a group made up of individuals born or who have spent most of their lives in the Netherlands, which explains Mohammed Bouyeri's connection to it.

After carefully analyzing the mechanisms identified in individual radicalization, it is necessary to move on to three other mechanisms, this time highlighted by McCauley and Moskalenko in group radicalization. In fact, often, the progressive acceptance of extreme goals and principles is a path shared by small groups of individuals, whose closeness and interdependence are very strong. According to the authors, the greater the cohesion of the group, the greater the pressure for the group to agree on the norm, which is equivalent to the greater the threat of deviations from that norm.²⁹

The first mechanism that can steer groups towards radicalization is group polarization. This phenomenon explains how some people, although initially joining a group with a limited predisposition, end up becoming more extremist by participating in group discussions or accepting shared ideas with other members. This signals the strong power that the group has over the individual. In light of this, the constant sharing of opinions, which gradually becomes more intense, only encourages members, promoting the radicalization of their ideas: hence, the trend towards increased extremism. The second identifiable mechanism is group competition³⁰, because without comparison with others, a group cannot be considered truly united.

²⁸ Lorenzo Vidino, The Hofstad Group: The New Face of Terrorist Networks in Europe, in "Studies in Conflict and Terrorism", 5/2007

²⁹ McCauley, C., & Moskalenko, S., Towards a profile of lone wolf terrorists: What moves an individual from radical thought to radical action. *Terrorism and Political Violence*, 26, 2014, 23-24;

³⁰ In the ninth chapter of their book, McCauley and Moskalenko distinguish three distinct types of competition. The first concerns the challenge between a group of activists and state government, often caused by specific policy changes.

Solidarity, participation, and a sense of belonging are the result of an individual's identification with a particular group rather than with others. Consequently, individuals who share similar ideals realize that they are fighting for the same cause and are exposed to the same threats, which creates a sense of union. The greater these threats, the greater the cohesion of the group, as greater effort will be required to achieve collective goals. Another important aspect is that the group is essential for promoting, among its members, the so-called binary worldview "us versus them", or more precisely, good versus evil.

Finally, the third and final mechanism identified in group radicalization is the isolation of the group. Specifically, the group gradually distances itself from society, further increasing its power over individual members. It is the group that defines moral standards, needs, goals and, logically, what is right and what is wrong, and even authorizes the use of violence. They see their social circle narrowing to the point that they perceive it as unique. The permissive reality of the group is one in which its members associate themselves as "brothers", ³¹with no one else to turn to. This kind of isolation reinforces the other group mechanisms discussed above, resulting in faster polarization and deeper competition with anything that poses a threat to the group.

Having also explained the mechanisms of group radicalization as a conclusion, it is now appropriate to delve into the last three mechanisms that McCauley and Moskalenko identify in mass radicalization. Terrorist groups, in fact, undoubtedly operate under illegal and clandestine conditions, but in order to survive and achieve their goals, they need an extensive network of contacts and supporters.

The first mechanism is defined as the politics of jujitsu, based on the terrorist strategy of committing acts of terrorism with the intention of provoking an unchecked response from governments, thereby exalting their sympathizers. In other words, the terrorists are trying to use the reaction of the state to incite their followers, who have not yet been mobilized, into action. To give an example, al-Qaeda used the politics of jujitsu to fight its opponents and maintain the group's purpose. As McCauley and Moskalenko point out, attacks on the United States on its territory led to a subsequent American response by sending troops to Muslim countries, which in turn mobilized Muslims against the United States. As the strategy envisioned, therefore,

Initially, the members of the group are not inclined to engage in illegal or violent actions; Here, as activists, they tend to participate in non-aggressive protests. However, when the state's response becomes repressive, the less involved members withdraw, while those who remain turn to more drastic actions. It should be noted that the perceived relationship within the group, threatened by a common cause, gradually becomes more intense. This results in greater group identification, greater idealization of the group's intrinsic values, and a greater willingness to punish anyone who violates group norms. The second type of competition involves groups competing for the same base of supporters and supporters. As rival groups engage in different strategies to advance their cause, conflict is likely to escalate into progressively more radical actions as sympathizers themselves encourage them. The latest example of competition concerns an internal rift within the group. That is, the constant pressure for conformity in behavior within the group can sometimes lead to factions and internal conflicts. In this regard, the IRA (Irish Republican Army) offers an example of a number of rival factions, which did not hesitate to attack each other.

³¹ Clark McCauley and Sophia Moskalenko, *Mechanisms of Political Radicalization: Pathways Toward Terrorism*, in "Terrorism and Political Violence", 7/2008

even those who had previously discouraged attacks would join the jihad against the Americans. The second mechanism consists of hatred of the enemy, which is devoid of all human characteristics, a way to circumvent possible inhibitory mechanisms. This means that the enemy is perceived as the ultimate evil, as a "gear" in the enemy machine, compared to animals and humiliated in every way, not destined to deserve any form of sympathy. In this sense, hatred is a particularly intense form of negative identification, based on the belief that members of an enemy group have an evil nature, which is why there is a growing impulse to attack all of them, regardless of age, gender, or marital status. In short, creating this completely harmful image of others only encourages and further facilitates violence. The third mechanism, finally, has to do with the culture of martyrdom, the impact of which on the general public is worthy of consideration. Suicide bombers incite Muslims who claim to be true followers of Allah to commit extreme acts in order to provoke mobilization through sacrifice. In other words, they tend to encourage others to follow their example and act with the same determination. In fact, once an individual has made their sacrifice for a common cause, other members will be even more motivated to fight and pay tribute to their comrade's death. Furthermore, it should be considered that an organization for which its members are willing to sacrifice themselves earns considerable recognition and authority. On the other hand, however, the idea that an individual can cause such horrific harm, despite the presence of state power, generates a common emotional state of fear, confusion, and resentment.

In light of the research data presented, it can be concluded that McCauley and Moskalenko, after developing a detailed analysis of the path to terrorism, did not seem to consider ideology as a decisive factor in radicalization, based on several studies in the field of social psychology that have shown that beliefs, in themselves, are an uncertain predictor of action. More specifically, the importance of ideology as a motivating factor is seriously underestimated because the authors themselves do not consider its values sufficient to incite an individual to commit an act of terrorism.³²

³² The explanatory separation principle was formulated in response to previous research that treated behavioral and cognitive radicalization as part of the same package, or used a fairly linear "conveyor belt model", a term coined by Zeino Baran, for the relationship between cognitive and behavioral radicalization. According to this model, extreme beliefs lead directly to extreme actions. Proponents of the explanatory separation principle reject such a model of that relationship. An influential example of the explanatory separation principle is the work and conclusions of Clark McCauley and Sophia Moskalenko that we have explained earlier. They distinguish between what they call the "Radicalization of Thought Pyramid" and the "Radicalization of Action Pyramid." In the Pyramid of Opinion, the largest group, basically, consists of neutral people, the second, smaller group, which is one level above, consists of sympathizers, then we find the justifiers, and then at the very top we find the smallest group: those who believe they are under a personal obligation (to use violence). In the Pyramid of Action, the largest group, basically, is inert, then we find activists, then we meet radicals, and finally at the top of the pyramid we find the smallest group: terrorists. According to McCauley and Moskalenko, many people with beliefs that justify or otherwise tolerate violence never join a militant group (this is often referred to as the "problem of the few") and many who actually join a militant group never have such radical views. This means that in their conclusion, they have shifted from the rather limited point of correlation between beliefs and the use of violence on the one hand and the actual use of political violence on the other, to a much broader claim: radical ideas do not usually lead to radical action. That this is not a coincidence is clear from

Instead of Conclusion – Understanding 'Mixed' Ideological Currents – MUU

Although there is a limited understanding of the diverse ways in which structures and processes at the individual level interact with each other in shaping trajectories of radicalization toward extremism, there are several characteristics that emerge and require further research. In fact, although significant research has explored the relationship between radicalization and factors such as group dynamics³³ and pathway approaches³⁴, conceptualizing and understanding the establishment of mixed and changing extreme worldviews requires the recognition of broader complex and context-dependent factors beyond pure ideological appeal.³⁵ It does not imply a critique of traditional models of radicalization as such, but rather calls for the expansion of research on radicalization to include those influenced by idiosyncratic ideological constructs. In particular, the relationship between complex personal subjectivities, social and/or collective identities, and ontological (in)security, which underpin the potential influence of radical spheres, can facilitate the formation of contemporary extremist attitudes.³⁶ In parallel, there is limited understanding of the mechanisms and processes by which psychosocial traits can change the worldview of individuals who may already be socially vulnerable in different ways and how this relates to expressions of identity.³⁷ Uncovering the complex and mutating ways in which these factors influence individual activity and shape worldviews, as well as recognizing the different properties of contemporary threats that escape traditional categorizations, is key to understanding the heterogeneous radical paths in extremism that individuals can take.

This perspective highlights how such dynamics contribute to mutating pathways of radicalization, namely, the creation and operationalization of the 'mixed, volatile and unclear' (MUU) ideology of 2018 is considered, for the purposes of the Prevention category – one of the four strands of the UK's strategic response to terrorism. The MUU category includes individuals who merge aspects of multiple ideologies and/or change

their discussion elsewhere about whether QAnon poses a threat to American democracy. Clark McCauley and Sofia Moskalenko, "Belief and Behavior in Radicalization: Clarifying the Relationship in Response to Pills," *Terrorism and Political Violence*, 2025;

³³ Sageman, M. (2008). *Leaderless Jihad: Terrorist Networks in the Twenty-First Century*. Philadelphia: University Press of Pennsylvania, 2008;

³⁴ Clark McCauley and Sophia Moskalenko, *Mechanisms of Political Radicalization: Pathways Toward Terrorism*, in "Terrorism and Political Violence", 7/2008

³⁵ Adams, B., Stahl, G., & Oberg, G. (2023). Ideology becomes a way to understand that mismatch: beliefs, behavior, and belonging – does ideology matter in countering violent extremism? *Critical Studies on Terrorism*, 16, 648–671; ³⁶ Jensen, M., Atwell Sitt, A. and James, P., *Radicalization to Violence: A Pathway Approach to the Study of Extremism*. *Terrorism and Political Violence*, 32, 2020, 9-12;

³⁷ Whiting, A., Spiller, K. and Avan, I. (2024). Counter-radicalization in UK higher education: A people-centred analysis of 'vulnerability' and the duty to prevent. *Critical Security Studies*, 12, 251–268;

between them without fixed alignment. According to the Home Office³⁸, the impetus for the introduction of the category stemmed from "research and development of statistics and a real increase in the number of cases identified with this kind of ideology". It replaced the former category of "unspecified" in previous annual statistics. Although the MUU category has recently been renamed "conflicting",³⁹ the overarching construct reflects cases where an ideology (as formulated by Prevention) is presented as: "mixed (involves a combination of elements from multiple ideologies), unstable (shifts between different ideologies), or vague (where an individual does not constitute a coherent ideology, but may nevertheless pose a risk of terrorism)".⁴⁰ This discussion uses the UK as a case study to develop a wider conceptual argument for the phenomenon of 'mixed' extremism; however, it is important to recognize that the landscape of contemporary extremist threats in the UK shares many common features with other countries across Europe, such as Belgium, Norway, France and Germany.⁴¹ At the same time, the nomenclature of inherently "mixed" forms of extremism is subject to wide variation, partly due to the complex nature of ideologies that mix multiple, diverse and contradictory beliefs, making them difficult to categorize in traditional frameworks. Along with the MUU classification in the Prevent program, the U.S. Strategic Framework for Counterterrorism and Targeted Violence, and the establishment of Fixed Threat Assessment Centers in New South Wales, Australia⁴², illustrate some of the broader preventive efforts to address the rise of extremist threats that transcend conventional ideological boundaries. Drawing attention to recent shortcomings in policymaking, the critical importance of advancing a more robust conceptualization of fluid, fragmented, and idiosyncratic ideological currents should not be underestimated.

Research on extremism and ideological currents that are inherently "mixed" remains scarce, but new papers show that incitement to hatred and violence along socio-political and ideological lines is an important factor.⁴³ For example, there are common features among anti-immigrant, anti-immigrant, and radical

³⁸ Prevent Strategy: A Guide for Local Partners in England was a key localized framework of the UK's counter-terrorism strategy (CONTEST),

<https://assets.publishing.service.gov.uk/media/5a78966aed915d07d35b0dcc/prevent-strategy-review.pdf>

³⁹ Home Office, Individuals referred and supported through the Prevent programme, England and Wales, April 2021 to March 2022, <https://www.gov.uk/government/statistics/individuals-referred-to-and-supported-through-the-prevent-programme-april-2021-to-march-2022>

⁴⁰ Home Office, COMPETITION: The UK Counter-Terrorism Strategy (cm 9608). London: TSO, 2018, <https://www.gov.uk/government/publications/counter-terrorism-strategy-contest-2018>

⁴¹ Europol. (2023). European Union Terrorism Situation and Trends Report. Publications Office of the European Union, Luxembourg. <https://www.europol.europa.eu/publication-events/main-reports/european-union-terrorism-situation-and-trend-report-2023-te-sat>

⁴² Farrell, P. (2017). NSW Police set up 'possessed persons' unit to help combat lone wolf terror attacks. The Guardian, 2017, <https://www.theguardian.com/australia-news/2017/apr/26/nsw-police-establish-fixated-persons-unit-to-help-counter-lone-wolf-terror-attacks>

⁴³ Brace, L., Baele, S., & Ging, D., Where do "mixed, vague, and unstable" ideologies come from? A data-driven response centered on the incelsphere. Journal of Policing, Intelligence, and Counterterrorism, 2024, 19;

masculinist groups that draw information from different ideological repertoires that rationalize violence as a means of provoking socio-political change and restoring a perceived endangered social order.⁴⁴ Although a number of proximate incidents and conflicts have influenced the expansion of risk prevention and mitigation strategies in the UK, recognizing the growing diversification of belief systems, identity constructs and associated grievances – beyond the traditional binary – is essential to understanding the interconnected dynamics of emerging extremism. This applies not least to the MUU category, which requires a more nuanced understanding that encompasses both ideological fluidity and the specific radical processes through which individuals engage. After all, these ideological fragments do not exist in isolation, but are shaped within broader ecosystems of grievances and meanings, shaping unpredictable individual paths to extremism.

BIBLIOGRAPHY

1. Daniele Albertazzi, Stijn van Kessel, Adrian Favero, Niko Hatakka, Judith Sijstermans, Mattia Zulianello, *POPULIST RADICAL RIGHT PARTIES IN ACTION (The Survival of the Mass Party)*. Oxford University Press, 2025;
2. *The Impact of Populism on the Party System: The Experience of European Democracies: Comparative analysis* co-edited with Jiří Kocián and Michal Kubát, 2026;
3. A Kaya, *Islamisation of Turkey under the AKP rule: Empowering family, faith and charity*, South European Society and Politics, 2015; *Islam, migration and integration: The age of securitization*, Springer, 2009;
4. *Nativism, Islamophobia and Islamism in the Age of Populism: Culturalization and Religionization of what is Social, Economic and Political in Europe*, European Research Council (ERC), ISTANBUL BILGI UNIVERSITES, 2023;
5. *Countering Violent Extremism and Global Counter-Terrorism: A Decolonial Approach*, Edited by Sagnik Dutta, Tahir Abbas and Sylvia I. Bergh, Manchester University Press, 2025;
6. Olusola, Jeremiah A., *Reading an Islamic epistemology into research: Muslim converts and contemporary religion in Britain*, Journal of Religious Education, 2020;
7. Koca, Metin, *Why the French bill against “Islamist separatism” goes beyond France, Islam and the Separation of Religion and State*, Istanbul Bilgi University, 2020;
8. Chrona, Stavroula, Tereza Capelos, *The civic and political participation of young people in a context of heightened authoritarianism. The case of Turkey*, Oxford Brookes University, Department of Social Sciences, 2017;
9. Meloy, J. Reid; Hoffmann, Jens (Eds.): *International Handbook of Threat Assessment*, (2nd ed.). New York: Oxford University Press, 2020;

⁴⁴ Russo, K., Agarwal, N., & Kirmayer, L., *Radicalization into violence: A view from cultural psychiatry*. Transcultural Psychiatry, 58, 2021, 603–615;

10. Quassim Cassam, *Extremism A Philosophical Analysis*, University of Warwick, UK, 2019;
11. Clycq, Noel et al. (Eds.): *Radicalization: A Marginal Phenomenon or a Mirror to Society?* Leuven: Leuven University Press, 2019;
12. Canter, David (Ed.): *The Faces of Terrorism: Multidisciplinary Perspectives*. Chichester: Wiley-Blackwell, 2009;
13. de Bruijn, Mirjam (Ed.): *Biographies of Radicalization: Hidden Messages of Social Change*, Berlin: De Gruyter Oldenbourg, 2019;
14. Goerzig, Carolin, Al-Hashimi, Khaled: *Radicalization in Western Europe: Integration, Public Discourse and Loss of Identity among Muslim Communities*, (Contemporary Terrorism Studies), Abingdon: Routledge, 2014;
15. Mark Sedgwick, *Against the Modern World: Traditionalism and the Secret Intellectual History of the Twentieth Century*, Oxford University Press, 2004;
16. Groothuis, S., *Researching Race, Racialization, and Racism in Critical Terrorism Studies: Clarifying Conceptual Ambiguities*. *Critical Studies on Terrorism*, 13, 2020;
17. Walter Laqueur, *No End to War: Terrorism in the Twenty-First Century*, Oxford University Press, 1999;
18. Michele Grossman, H. A. Hellyer, *Rethinking Religion and Radicalization: Terrorism and Violence Twenty Years After 9/11*, European Union's Horizon 2020 research, 2020;
19. Walter Bryce Gallie, *Essentially Contested Concepts*, Faculty of Humanities and Social Sciences of the University of Jyväskylä, 2021;
20. Sophia Moskalenko and Clark McCauley's, "Measuring Political Mobilization: The Difference Between Activism and Radicalism", *Terrorism and Political Violence*, 21, 2009;
21. Clark McCauley and Sophia Moskalenko, *Mechanisms of Political Radicalization: Pathways Toward Terrorism*, in "Terrorism and Political Violence", 7/2008
22. Alessandro Orsini, *What Everybody Should Know about Radicalization and the DRIA model*, in "Studies in Conflict and Terrorism", 46/2023;
23. McCauley, C., & Segal, M. (1987). *The social psychology of terrorist groups*. In C. Hendrick (Ed.), *Group processes and intergroup relations: A review of personality and social psychology*, SAGE, 2017;
24. Lorenzo Vidino, *The Hofstad Group: The New Face of Terrorist Networks in Europe*, in "Studies in Conflict and Terrorism", 5/2007;
25. Sageman, M. (2008). *Leaderless Jihad: Terrorist Networks in the Twenty-First Century*. Philadelphia: University Press of Pennsylvania, 2008;
26. Clark McCauley and Sophia Moskalenko, *Mechanisms of Political Radicalization: Pathways Toward Terrorism*, in "Terrorism and Political Violence", 7/2008
27. Adams, B., Stahl, G., & Oberg, G. (2023). *Ideology becomes a way to understand that mismatch: beliefs, behavior, and belonging – does ideology matter in countering violent extremism?* *Critical Studies on Terrorism*, 16;
28. Jensen, M., Atwell Sitt, A. and James, P., *Radicalization to Violence: A Pathway Approach to the Study of Extremism*. *Terrorism and Political Violence*, 32, 2020;
29. Whiting, A., Spiller, K. and Avan, I. (2024). *Counter-radicalization in UK higher education: A people-centred analysis of 'vulnerability' and the duty to prevent*. *Critical Security Studies*, 12;

- 30.Prevent Strategy: A Guide for Local Partners in England was a key localized framework of the UK's counter-terrorism strategy (CONTEST), <https://assets.publishing.service.gov.uk/media/5a78966aed915d07d35b0dcc/prevent-strategy-review.pdf>
- 31.Home Office, Individuals referred and supported through the Prevent programme, England and Wales, April 2021 to March 2022, <https://www.gov.uk/government/statistics/individuals-referred-to-and-supported-through-the-prevent-programme-april-2021-to-march-2022>
- 32.Home Office, COMPETITION: The UK Counter-Terrorism Strategy (cm 9608). London: TSO, 2018, <https://www.gov.uk/government/publications/counter-terrorism-strategy-contest-2018>
- 33.Europol., European Union Terrorism Situation and Trends Report. Publications Office of the European Union, Luxembourg, 2023, <https://www.europol.europa.eu/publication-events/main-reports/european-union-terrorism-situation-and-trend-report-2023-te-sat>
- 34.Farrell, P. (2017). NSW Police set up 'possessed persons' unit to help combat lone wolf terror attacks. The Guardian, 2017, <https://www.theguardian.com/australia-news/2017/apr/26/nsw-police-establish-fixated-persons-unit-to-help-counter-lone-wolf-terror-attacks>
- 35.Brace, L., Baele, S., & Ging, D., Where do “mixed, vague, and unstable” ideologies come from? A data-driven response centered on the incelsphere. *Journal of Policing, Intelligence, and Counterterrorism*, 2024;
- 36.Russo, K., Agarwal, N., & Kirmayer, L., Radicalization into violence: A view from cultural psychiatry. *Transcultural Psychiatry*, 58, 2021;

DIGITAL EXCLUSION AS A RESILIENCE GAP: AGE-STRATIFIED POLITICAL DISENGAGEMENT ACROSS WESTERN BALKAN ENLARGEMENT ECONOMIES

Besa Shahini

Faculty of Economy, University of Tirana

Email: besa.shahini@unititir.edu.al / besashahini@feut.edu.al

ORCID: 0000-0003-1132-8549

Acknowledgements: This article/publication is based upon work from RELINK² COST Action CA23114 - Regaining linkage? Digital technologies improve civic engagement, political organizations and democracy

Abstract

Political actors in the Western Balkans exist under circumstances that reverse the basic assumptions made in most European scholarship examining digital re-linking. The six enlargement economies remain outside full member-state digital monitoring, as DESI is now embedded in the EU State of the Digital Decade reporting, while the region is assessed through separate RCC DESI benchmarking. At the same time, institutional trust remains weak, toward parliaments and political parties. Two age groups are detached from representative politics through opposite routes. Older citizens remain attached to institutions but are exposed to digital marginality. Younger citizens are digitally included, yet less institutionally attached and less connected to political channels. This paper treats these patterns as one structural weakness: an age-stratified resilience gap that limits the capacity of digital inclusion to repair the vertical tie between citizens and representative institutions at the frontier of EU enlargement. Using Eurostat ICT indicators, Balkan Barometer country aggregates, Balkan Barometer 2024 individual-level records for Albania, Bosnia and Herzegovina, Kosovo, Montenegro, North Macedonia, and Serbia, and national turnout data, the analysis separates digital exclusion into access, skills, and use, following van Dijk. It relates these dimensions to trust in parliament, trust in political parties, and conventional political engagement, with attention to age interactions. The findings show that household access is close to saturation, but access, skills, and use vary separately. Digital service use declines with age, while institutional trust rises with age. In countries where the co-occurrence is more pronounced, trust experiences sharper decline.

Keywords: *civic attrition; digital exclusion; age-stratified disengagement; democratic resilience; Western Balkans*

1. INTRODUCTION

The starting assumptions of much of the European scholarship on digital re-linking no longer apply at the south-eastern fringe of Europe. They presume that political organizations are operating within a calibrated digital landscape, that the Digital Decade/DESI provides full and comparative digital coverage, that trust has been largely negotiated, and that the challenge now lies in reconnecting citizens who have drifted into digital spaces. Nothing could be further from the truth. The six economies of the Western Balkans defy just about everything in this list: since 2023, DESI is incorporated into the State of the Digital Decade reporting among member states (European Commission, 2024); the Regional Cooperation Council (RCC) has established benchmarks for Western Balkans DESI (RCC, 2024b; 2025) but its indicators remain untested in representative terms of trust or the age structure of exclusion. Institutional trust is very low, and the two representative institutions - parliaments and parties - are the two most distrusted in the region (RCC, 2024a). And the citizens who have abandoned representative politics followed different routes to achieve this.

There are two age-cohorts in the Western Balkans who have abandoned representative politics for different reasons. The older citizens (aged 65+) seem relatively more institutionally attached: the extensive literature on electoral participation indicates their higher propensity to turn out to vote; the older age-cohorts tend to possess firm political orientations towards political parties and parliaments; lastly, this demographic possesses many of the habits acquired by an electorate well before the Internet transformed public communication. But older cohorts are lacking in access to digital tools and associated skills to use them. Young adults, by contrast, are much more digitally included, although this does not imply that all youth have all the requisite competences, but have pulled away from institutions of representation. One way to look at it would be that the former cohort is institutionally more attached but more vulnerable to digital exclusion; while the latter is more digitally included but has fewer attachments to institutions of political representation.

These are two research strands which rarely intersect in one study: one deals with the grey digital divide in which older cohorts of non-users are viewed as a problem of access and digital skills; while the other concerns youth disaffection wherein young abstainers are a problem of political culture and party supply. If juxtaposed, one will notice that these two bodies of work describe essentially the same weakness in structural terms; thus, this paper explores the age-stratified detachment of older and younger citizens of Western Balkans economies as one phenomenon: a resilience gap that weakens the vertical link between citizens and the representative institutions.

Under the resilience gap, one should understand the erosion of the cooperative bond between citizens and the representative institutions of a society that allow the latter to withstand pressures without being broken. Specifically, vertical legitimacy, the recognition of the right of parliaments and political parties to wield power, is precisely that vulnerable component of the cooperative bond that becomes strained when communication channels connecting citizens and institutions are inaccessible to one generation and are perceived negatively by another. While it is possible for a country to develop a digital economy, its cooperative bond with representative institutions may continue to weaken due to disinterest rather than lack of opportunity.

This is particularly significant beyond issues of digital divide since this bond is what ensures the integrity of a society under the threat of various pressures. Wherever the

cooperative bond continues to remain intact, a society would be able to withstand economic shocks, fake news, foreign meddling, and eventually recuperate. However, once this bond breaks, the channels necessary for collective response start deteriorating. Civic attrition, which is nothing other than the loss of cooperation under prolonged strain, poses a security problem as well as a democratic one; the former appears wherever institutions needed for coordinated response are less trusted.

This study examines correlations between the components of digital exclusion and civic attrition using published data for six Balkan economies. Households' access to the internet and the frequency of their internet usage have been provided by Eurostat, alongside digital skills indicators. Trust in institutions and use of digital services have been collected by the Balkan Barometer. Finally, voter turnout rates are calculated from election statistics provided by International IDEA. This paper measures the effects of digital exclusion using the framework suggested by van Dijk (2005), who defines it as consisting of access, skills, and use, and assesses its effect on citizen-institution relations and how this varies by age. A total of four findings emerge from this study. First, despite some gaps remaining in the last category, access to digital technologies from home has become ubiquitous - over 84% in all six economies and over 90% in half of them. Second, access, skills, and use are largely uncorrelated and cannot be considered stages in one process. Third, the age gap within each category is very pronounced; young adults have far more advanced digital skills compared to their older peers (ratio greater than two to one). Fourth, the share of the younger generation using social media is about three times higher. Overall, digital exclusion correlates negatively with trust in representative institutions, while voter turnout does not appear to be influenced by the age pattern. Using 2024 Balkan Barometer data for over six thousand respondents, the present study demonstrates that the use of digital services decreases markedly with age, whereas trust increases, both factors being determined by age rather than interdependent. Consequently, the resilience gap is purely a matter of composition. With just six economies analyzed, the results presented here should be considered merely exploration and depend heavily on exact permutation tests and leave-one-out checks.

This paper consists of five sections. The first considers the introduction; the second provides a review of the literature on multidimensional digital exclusion, institutional trust, participation, and the effects of age on these processes, as well as outlining the concept of resilience gap. The third discusses the data used and measurement procedures. In the fourth section, analysis of data is presented, ranging from descriptive decomposition to the age gradient analysis and cross-country correlations between digital exclusion and institutional trust/participation, going on with the individual-level analysis of digital services usage and trust in institutions depending on the age group and interaction therein and the findings of the paper along with discussion and limitations. The fifth gives the conclusions.

2. LITERATURE REVIEW

2.1. Digital exclusion as a multidimensional condition

The notion that being online is one threshold has not stood up to the evidence. van Dijk (2005) differentiated the digital exclusion into successive stages - the presence of devices and connections, skills required for their operation, and what exactly the connection is being used for. All these stages may continue to exist after the preceding stage had been addressed. A family may own a subscription that is of no use to anyone within the house. A person with skills may use them for entertainment or social communication but refuse to

interact with any state services online. The stages are nested logically but disconnected in real life, hence overcoming one does not overcome others.

The distinction between the access divide and the digital divide made by Norris (2001) refers to the difference between people with and without internet used for public purposes. As access divides close, the digital divide begins to rely more on skills and uses. That is how Scheerder et al. (2017) and van Deursen & van Dijk (2019) discuss the second and third-level divide.

This phenomenon is labeled as digital inequality by Robinson et al. (2015). The authors propose that there is a tendency for advantaged people offline to leverage digital divide into even higher inequality by accessing devices, using skills for beneficial purposes, and then using these uses to become even more advantaged, while less advantaged remain disadvantaged despite access. Hence, for measuring digital divide, it is essential to break down its components and analyze separately to get information on how far each level of inequality goes. It is impossible to measure digital inequality as a single composite number. That feature becomes even more important in the enlargement context because European Union DESI indexes combine various aspects of the digital divide into a single number where the access aspect may make up for poor skills. Regional Western Balkans DESI benchmarking indexes (RCC, 2024b; RCC, 2025) also represent a composite index of the digital development in Western Balkan economies, not digital trust among other aspects of digital divide. Therefore, studying the subject in relation to politics in these countries means rebuilding DESI indicators from components, which limits the research but opens new possibilities at the same time.

2.2. Trust, participation and vertical legitimacy

But engagement with representative institutions is itself predicated on the belief that these institutions merit engagement. In Norris' (2011) description, there is a population of citizens critical in their adherence to democratic ideals while damning the parties and parliaments with which they must engage, and which they refuse to cooperate with. This refusal is targeted: it strikes harder at the institutions of representation - parties and legislatures - than it does at agencies of order or adjudication, precisely because the latter cannot fail so easily as the former. Whether declining trust leads to declining participation depends on what type of participation one considers. According to Hooghe and Marien (2013), using a range of European populations, declining trust drives citizens into non-institutional action - such as protest - while steering them away from participation in institutional processes, like voting. This means that even in the absence of participation, the connection between trust and turnout remains robust. The larger context, in this view, belongs to Putnam (2000), who speaks of the social capital constituted by citizens and by their relation to institutions as something that is depleted gradually and unnoticed until a shock forces an accounting of its depletion.

The internet once seemed like just such a shock - one capable of reversing the trend. But the accumulated evidence paints a more cautious picture. As Boulianne (2009) synthesizes dozens of research results, digital access is only weakly linked to higher levels of engagement on average, with significant variation in effect sizes and little support for strong claims in either direction. That is, digital access may not serve as a reliable lever on participation, only a weak one. Here Merkel (2004) furnishes the language for the concept involved in failed leverage, particularly within the context of fragile regimes. Defective democracies possess inadequate regimes of citizen participation and interaction, with the vertical legitimacy that constitutes citizens' approval of representative institutions being

particularly vulnerable to erosion while elections continue.

Social capital is more broadly described in Putnam's (2000) terms as the trust and reciprocal relations enabling coordinated action. The vertical dimension of social capital involves trust in institutions, and its erosion strains the legitimacy of elections and parliaments.

2.3. The two-ended age structure of detachment

However, age emerges as a determinant of both processes, but they run in opposite directions. In terms of digital engagement, the group in question falls on the lowest spot concerning every aspect analyzed. They tend to have fewer chances to have an Internet connection available, and the skill deficit is particularly high because the development of skills takes time and requires a long period of exposure starting from childhood years. Therefore, one should note that the grey digital divide cannot be seen as the issue of a few late adopters but reflects the permanent gap in skills and usage that persists although access gaps are now lower due to the development of households' opportunities (Friemel, 2016). In turn, in terms of politics, age becomes an element promoting participation, where individuals participate in increasing measures during their life cycle until reaching very advanced years, when participation starts falling. Older people vote while the youngest citizens are the least likely to perform such activities (Sloam, 2014). At the same time, Gallego (2015) states that the age-related differences in participation coexist with educational and resources-related variations, which proves that unequal participation is a permanent feature of modern democracies rather than a temporary trend.

Concerning the young citizens' attitude towards politics, Bennett (2008) claims that it is not caused by the lack of interest but is the result of changes in citizenship. Thus, youngsters often prefer networks and social movements to parties.

2.4. The Western Balkans as an inverted case

This dynamic becomes more pronounced when considering the Western Balkans, since the circumstances surrounding it deviate from those considered in digital re-linking. In particular, the region's representative institutions operate under the influence of competitive authoritarianism, whereby the incumbent political class uses media and governmental resources to undermine the authority and trustworthiness of political parties and parliament (Bieber, 2018, 2020; Bermeo, 2016). This erosion of credibility can be seen in measures of trust: Heinz (2024) locates the Western Balkans amid the global trend toward institutional distrust, while according to the Balkan Barometer, parties and parliaments are the least trusted institutions out of six economies included (RCC, 2024a). Thus, there is already tension in the vertical linkage before the issue of digital exclusion becomes relevant.

The metrics are not complete. While the EU State of the Digital Decade report assesses digital progress in its member states using DESI indicators, Western Balkans DESI benchmarks have been set up by the Regional Cooperation Council (RCC, 2024b; RCC, 2025). Nevertheless, as a composite indicator, DESI evaluates digital development and does not measure the linkage between the former and representational trust. According to the World Bank, the potential of connectivity can only be fulfilled in the context of complementary skills and institutions; otherwise, a digital dividend will not appear (World Bank, 2016). And here the Western Balkans are on course toward this outcome: although household internet usage has increased dramatically, neither the skills nor the use, nor the necessary trust in institutions to turn household access into citizen engagement, have kept up with the pace. Thus, the resilience gap emerges because of the interaction between a

lack of trust in institutions of representation, inadequate regulation, and measurement of digital transformation, as well as stratification of exclusion by age into two separate cohorts: an offline but engaged group and an online but disengaged generation.

2.5. Re-linking political organizations and the limits of a connectivity fix

Recent academic literature focuses on mechanisms whereby political organizations re-establish contact with the citizens through which communication moves online, treating parties, parliaments, and campaigns as political structures which need to re-establish their connection through digital means due to the breakdown of intermediaries. The basic assumption in most of the literature here is that the digital tool enhances the organizational reach of representational actors and enables them to connect with a public which cannot be addressed through the mass structures anymore. It is a rational assumption if certain conditions are met; however, it operates under an assumption which should not go unquestioned in the overall context: that citizens connected by the organization through digital means are indeed digitally accessible and digitally responsive.

Here the age structure interferes; the digital connection is prone to connect more young citizens who are responsive to the channel, but not really trusting to its message, as well as ignoring old citizens who may appreciate representation but are digitally unreachable. Thus, connectivity reshuffles the contacts instead of automatically solving representational problems.

From the security point of view, reshuffling represents an aspect of resilience problem; for representational system needs widespread civic readiness to act through institutions in crisis. The problem emerges if citizens are willing to participate without being digitally reachable and vice versa.

3. METHODOLOGY AND DATA

3.1. Data sources and coverage

It includes Albania, Bosnia and Herzegovina, Kosovo, Montenegro, North Macedonia, and Serbia. The digital indicators are based on the Eurostat ICT data. The indicators of trust and digital service usage are taken from the Balkan Barometer 2024. Voter turnout data are collected by the International IDEA, used only as a diagnostic variable because there is no common election calendar and voter registers.

The sample differs in completeness. Households' access to digital devices is collected for all economies included; daily Internet usage - for five countries; and possession of basic digital skills - only for three countries (Albania, Montenegro, and Serbia). Assessment of trust and use of digital services is conducted for all six Balkan countries as part of the spring 2024 Balkan Barometer wave. Voter turnout data refers to different election years and thus is not used as an indicator itself. Reference years for each type of data: Kosovo - 2020; Albania -2023; others - 2024. Daily Internet usage: Albania - 2023; others - 2024. Basic digital skills: Albania -2021; Montenegro and Serbia -2023. The service usage indicator is measured as the proportion of population having used at least one online or digital service (minus those who have not used any). Trust is calculated as the proportion of people expressing trust or tending to trust the institution.

Table 1. Digital inclusion components and institutional trust, by economy.

Economy	Access (hh, %)	Daily use (%)	Basic skills (%)	Service use (%)	Trust: parliament	Trust: parties
Albania	97.7	83.0	24.0	77	40	24
Bosnia and Herzegovina	84.2	84.7	-	76	23	18
Kosovo	96.4	-	-	91	42	22
Montenegro	84.5	87.2	51.0	94	41	35
North Macedonia	90.8	88.0	-	74	16	14
Serbia	88.9	84.5	33.6	67	37	21

Source: Eurostat (2024a, 2025); RCC (2024a; trust item Q35)

3.2. Operationalizing the three components

The measure applies the classification of van Dijk directly. Access is the share of households with internet access in the household, which corresponds to the first-level divide in its classic form. Skills refer to the share of individuals with basic competences, and it is the second-level divide. Usage refers to the scope of digital usage i.e. the share of individuals who have made use of some online service or activity over the previous year. That measure complements the Balkan Barometer measure in that it excludes from the calculation the share of individuals reporting no use, and it captures the third level, which is what people do with their connections. Two additional measures, one on the intensity of usage based on daily internet use, and one on the extent of e-government take-up, are retained as supplementary indicators.

The access-usage exclusion index is calculated as 100 minus the means of the access and usage rates. As mentioned above, it is a clear descriptive index, not a latent scale. Skills are left out of the index because there are no comparable numbers for only three of the economies. The key dependent variable is parliamentary and party trust.

3.3. Estimation and inference

With six units in the cross-national comparison, the analysis will be largely exploration in nature, and inference will proceed mostly through description. Dispersion is presented via means and coefficients of variation. Correlation will be assessed via permutation testing, and the exclusion-trust relationship will be examined via refits following leave-one-out procedures.

There are two constraints that should be kept in mind when interpreting the findings below. First, the design is ecological, meaning that the units of observation are countries, and an association found among countries cannot be interpreted as applying to individuals without falling prey to the ecological fallacy. In other words, a country-level negative correlation between digital exclusion and institutional trust does not imply that individuals who lack internet access are distrustful of parliament, it simply implies that the more excluded economies are the low-trust countries.

The age issue also comes with limitations of interpretation because while the harmonized age-based crosstabs are available both for the EU and for the enlargement countries, they are unavailable in comparable format for each economy.

Thus, the gradient is derived indirectly from the wider figures as triangulation rather than direct inference from the unit-of-analysis level. Country-level analysis is thus exploration in nature. Individual-level survey results for the Balkan Barometers are employed separately in Section 4.7 to test the age-use structure within economies.

3.4. Validity, alternatives and reproducibility

Three choices of measurement have significance, and these are justified below. First, using service breadth as the usage component rather than frequency of use, the latter being the common measure. Frequency measures intensity, but not purpose, which is what the represented relationship depends on. A population may use the internet daily for messaging and video, without ever transacting any public service digitally. The use of the Balkan Barometer question on whether the respondent uses digital/online services for any reason in the previous year reflects purpose. This is the purposeful element of usage that matters for the represented relationship, and this question is reported for all six economies. Meanwhile, frequency of use remains as an additional intensity measure, and it behaves differently in the analysis than service breadth does, which verifies their non-interchangeability.

Second, the skills measure is kept separate from the composite index, even though the index could be considered latent scale. The exclusion index of access and usage is not latent, since it is simply the descriptive combination of two available measures for each economy, and so the index cannot cover the whole skills construct that is crucial for the theory, which can be estimated directly for only three economies. An index incorporating an imputed skills measure for the other three countries would create an illusion of precision that is not warranted by the evidence at hand. Reporting skills separately and using the two measures to build the composite index, keeps the construct honest about its scope of application. Third, the outcome itself - institutional trust, as measured in the single 2024 wave of the Balkan Barometer, rather than electoral turnout. The next section will quantify the latter effect. Trust avoids all three problems: it is the same question posed to comparable samples in all six economies in the same season, and it represents the construct under consideration.

Finally, reproducibility. All figures in the analysis are constructed from the data and methods published and referenced with their years of publication. The entire dataset and the analysis code, including permutation and jackknife procedures, are made available, so that all charts and tables can be reproduced and verified. The individual-level analyses follow the same model as the aggregate results and are presented separately.

3.5. Individual-level data and model

However, these cross-national relationships are limited by country aggregates and cannot assess the age structure on which the argument hinges. As such, the analysis will draw on the individual data from the 2024 Balkan Barometer, a survey of 6,000 people from all six Balkan economies, collected by the RCC (RCC, 2024a). There are two dependent variables - trust in parliament and trust in political parties.

The variable is dichotomous, indicating respondents who either trust or tend to trust the institution and those who are indifferent or tend not to trust it, with the few do-not-knows being excluded. The key independent variable is the count of online services that a respondent reported having used during the previous year, ranging from zero to nine - a direct equivalent of the economy-level measure. Age is included as a dummy-coded categorical variable, with six levels (from 18-24 through 65+). It is essential not to confuse

these age groups with the Eurostat contextual age categories of 16-24 and 65-74. Finally, there are control variables - education, gender, urban/rural, and the economy fixed effects. The latter accounts for differences between the economies already identified.

Only the interaction model is tabulated and the additive main effects were materially identical. Standard errors are adjusted for heteroskedasticity. Since the public file contains no survey weights, the analysis will produce unweighted coefficients with economy fixed effects, meaning that the findings are to be interpreted as intra-economy variation. The sample consists of approximately five thousand seven hundred individuals with complete data (5,669 out of 6,000); thus, the models have sufficient statistical power for testing this argument and can be interpreted as estimates.

4. DATA ANALYSIS

4.1. The decomposition of digital inclusion

Decomposition, then, is the first finding. Household access averages 90.4 per cent and does not vary much. Service usage is below and less evenly distributed. Skills, which are measured in only three economies, average far less and are even less evenly distributed. E-government use is the least evenly distributed indicator. Table 2 and Figure 1 illustrate the disconnect between connectivity and usage.

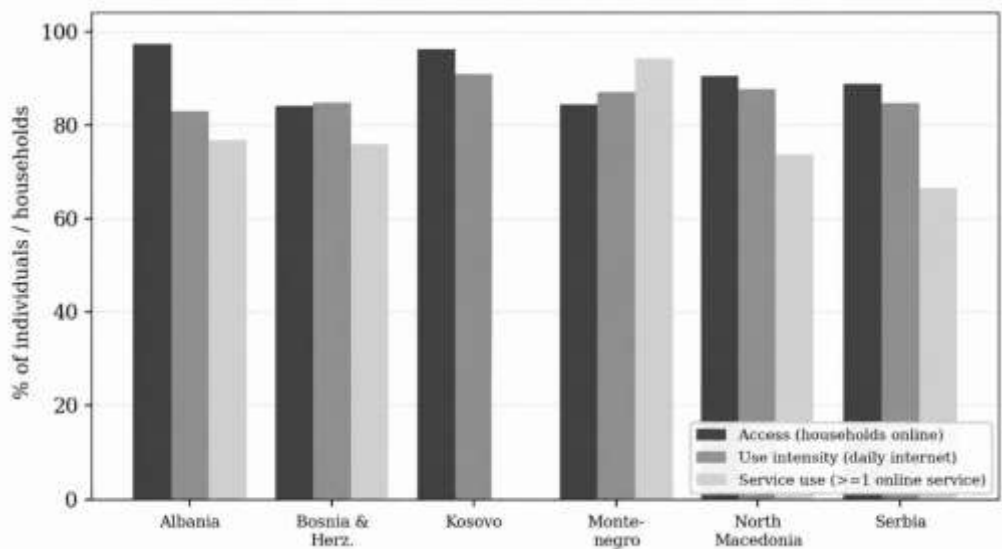
Table 2. Descriptive statistics for digital and engagement indicators.

Indicator	n	Mean	SD	CV (%)	Min	Max
Access (households, %)	6	90.4	5.7	6.4	84.2	97.7
Daily use (%)	5	85.5	2.1	2.4	83.0	88.0
Basic skills (%)	3	36.2	13.7	37.8	24.0	51.0
Service use (%)	6	79.8	10.5	13.1	67.0	94.0
E-government use (%)	6	20.2	20.3	100.5	5.0	60.0
Exclusion index	6	14.9	6.0	40.3	6.3	22.1
Trust in parliament (%)	6	33.2	10.9	33.0	16.0	42.0
Trust in parties (%)	6	22.3	7.1	31.9	14.0	35.0
Turnout (%), set aside	6	52.8	4.8	9.0	46.3	58.7

Source: authors' calculations on Eurostat (2024a, 2025), RCC (2024a) and International IDEA (2024).

**CV is the coefficient of variation. The exclusion index is one hundred minus the means of access and service use.*

Figure 1. Access, use intensity and service use by economy (Eurostat 2024-2025; Balkan Barometer 2024).



Source: authors' calculations on Eurostat (2024a, 2025),

* Access is household-level; daily use and service use are individual-level; Kosovo daily-use is not available.

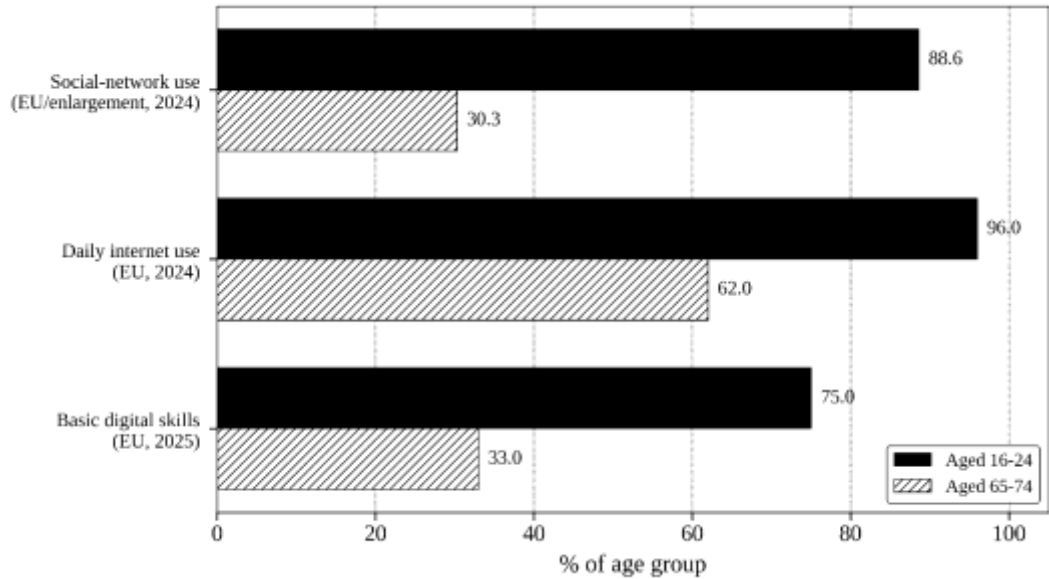
The dispersion pattern alone tells a story that is not captured by the headline access number. As connectivity approaches its natural limit, the dispersion between the economies falls from access into skills and use, exactly mirroring what literature on second and third order divides predicts (Scheerder et al., 2017; van Deursen & van Dijk, 2019). An index based purely on connectivity would show all six economies to be almost equally developed in terms of penetration and development and would fail to capture the distinctions between an economy where people have transaction capabilities and where they only have connectivity.

While they exist on different scales, each component is independent of the other. Among the six economies studied, there is zero correlation between access and use. The access-use correlation is effectively zero ($r \approx 0.01$; exact permutation $p = 0.988$), so the data give no evidence that the two move together. Albania leads the pack on connectivity with 97.7 per cent penetration yet has a mediocre share of service usage of 77 per cent and a low rate of turnout. On the other hand, Montenegro has lower access to Internet with 84.5 per cent, but higher e-service usage of 94 per cent. How well an economy is connected gives little indication of what people in that economy can accomplish online. Contrary to the ladder of penetration skills use, this is not simply an issue of formality, but of reality. As is to be expected, use of e-government services is highly variable across the group of economies considered - from 5 per cent in North Macedonia to 60 per cent in Albania. This is just as much an issue of institution as it is of citizens. In Albania, the portal boosts usage, while weak public services online contribute to exclusion in other countries.

4.2. The age gradient within each component

Age represents the second dimension. EU contextual data reveal an abrupt gradient here: around 75 per cent of young people aged 16-24 can use digital technologies at a minimum level, whereas only around 33 per cent of people aged 65-74 have the same digital competencies. Regarding the use of social networks, the difference is even greater. This is another type of contextual anchoring, not an estimate of Western Balkan countries by age group.

Figure 2. Age gradient in digital inclusion:



Source: authors' calculations on Eurostat (2024a, 2025),
*EU-level anchors used as context, not Western Balkan country-by-age data

Table 3. Age gradient in digital inclusion: EU-level anchors used as context.

Indicator	Aged 16-24	Aged 65-74	Ratio	Gap (pp)
Basic digital skills (EU, 2025)	75	33	2.27	42
Daily internet use (EU, 2024)	96	62	1.55	34
Social network use (EU, 2024)	88.6	30.3	2.9	58.3

Source: Eurostat (2024a, 2024b).
*Figures are approximately published anchors used to characterize the age gradient; comparable by-age cells for each Western Balkan economy are not published in harmonized form.

As it goes, the political gradient operates in the reverse direction to the digital one. The youngest age group, which scores highest on the selected digital criteria used to anchor the analysis, is generally less prone to voting, a tendency common both across different democratic systems and time periods (Smets & van Ham, 2013). As far as the Western Balkans region is concerned, the very same cohort appears to be distancing itself from representative institutions regardless of the above-mentioned digital divide: youth support for the view that EU membership will improve their economic situation decreased from 70 per cent in 2021 to 59 per cent in 2024, while their attitudes towards it are shifting to a

neutral position (RCC, 2024a). It must be noted that it is not about abstaining from participation altogether but about using available representative channels. Bennett (2008) explains it as part of a new meaning of citizenship. On the other end of the political gradient, there is the oldest age cohort that tends to be more dependable voters but has the weakest position in terms of skill acquisition as well as the usage of social media channels by parties.

4.3. Digital exclusion and institutional trust across economies

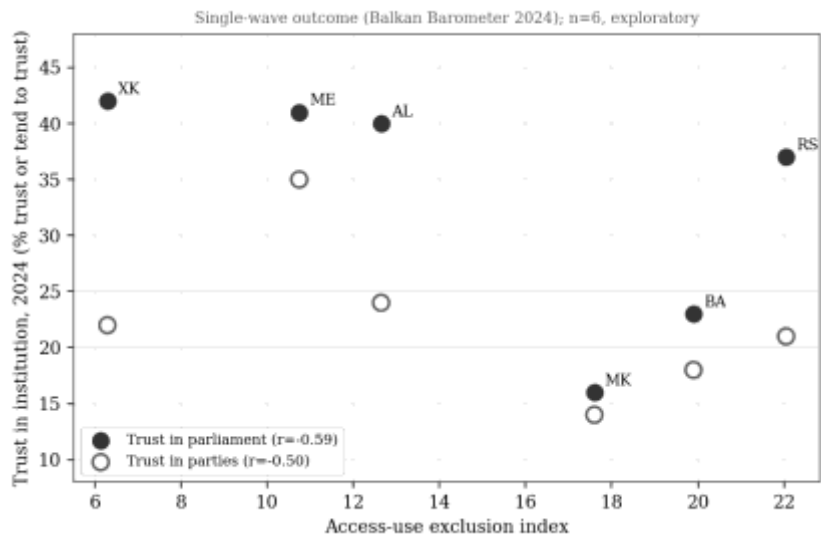
The principal dependent variable is institutional trust. In 2024, trust in parliament varies between 16 per cent in North Macedonia and 42 per cent in Kosovo; trust in parties between 14 and 35 per cent. Access-use exclusion correlates negatively with institutional trust, and the use of services positively.

Table 4. Associations between digital indicators and institutional trust, exact permutation tests.

Outcome ~ predictor	n	Pearson r	p	Spearman rho
Parliament trust ~ exclusion index	6	-0.590	0.236	-0.771
Parliament trust ~ service use	6	+0.490	0.289	+0.771
Parliament trust ~ access	6	+0.341	0.525	+0.314
Parties trust ~ exclusion index	6	-0.502	0.317	-0.600
Parties trust ~ service use	6	+0.681	0.125	+0.771
Parties trust ~ access	6	-0.191	0.769	+0.200

Source: authors' calculations on Eurostat (2024a, 2025) and RCC (2024a).
 *p-values from exact permutation enumeration of all 720 orderings. The exclusion index is one hundred minus the mean of access and service use; trust is the share expressing trust or a tendency to trust, Balkan Barometer 2024. With six units, p-values are reported for transparency, not as decision thresholds.

Figure 3. Access-use digital exclusion and trust in parliament and parties, single 2024 wave. Points labelled by economy; n=6, exploratory, no fitted line.



Source: Author’s calculations on Eurostat (2024a, 2025) and RCC (2024a).

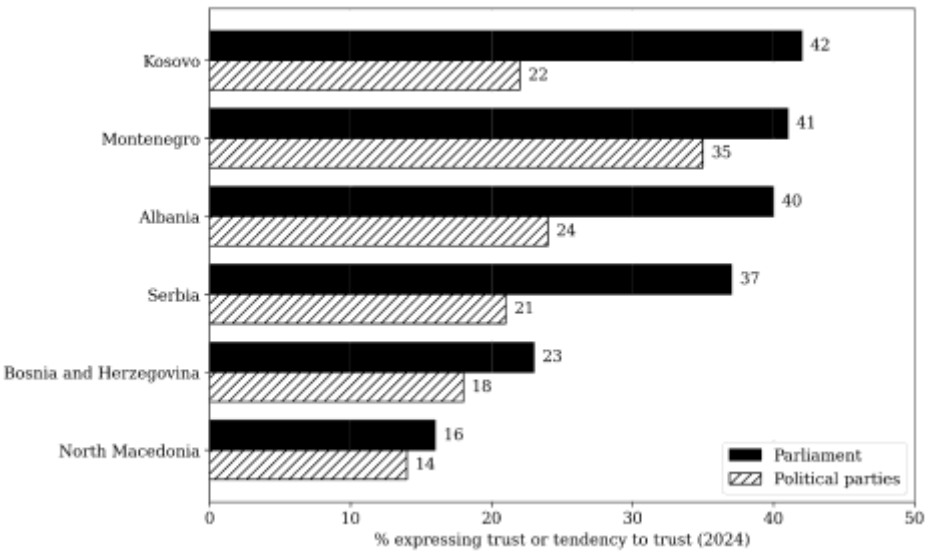
This combination appears in Figure 3. In each case, the economies with the lowest level of exclusion have the highest institutional trust - 42, 41 and 40 per cent in parliament in Kosovo, Montenegro and Albania; while the ones with the highest levels of exclusion have the lowest institutional trust - 16 and 23 per cent in North Macedonia and Bosnia and Herzegovina. But there is an outlier - in Serbia, which combines high exclusion with relatively strong parliamentary trust, thus explaining why the coefficient in the linear model is lower than the one in the rank-based regression. As can be seen from the leave-one-out jackknife analysis below, the negative slope holds for all six refitted models in both institutional trust measures.

Institutional trust is used instead of turnout since turnout figures cannot be meaningfully compared across six very different elections. Election years are different, voters' registers are inflated in high-migration countries, and contests are varied in terms of competitiveness.

4.4. The representative institutions at the bottom of trust

Institutionalism has an element of selectivity. Parliaments are more trusted than political parties in all six economies, and they are outranked by the police and judiciary in the regional dimension. Vulnerability resides in representation and not in general issues of public authority (RCC, 2024a).

Figure 4. Trust in parliament and political parties by economy, Balkan Barometer 2024.



Source: Author's calculations on RCC (2024a).

By-economy reasoning is a limitation of the argument too. The link between exclusion and trust exists, illustrated by the difference in views between Kosovo, Montenegro, and Albania compared with North Macedonia and Bosnia and Herzegovina, with Serbia deviating from the pattern. It makes up a real pattern that comes out of a single wave with one instrument, and that fits into the mechanism outlined in the paper where digital exclusion and the age composition behind it go hand in hand with the weaker relationship of representation. It is not an effect tested at six units, and the following

subsection explains how much more it can be analyzed. Both readings show the stress in a vertical relationship, in line with the notion of a defective democracy developed by Merkel (2004).

4.5. Robustness and what the evidence can bear

First, the ranks have at least as high coefficient as the product moment for the focal associations, reaching -0.77 for exclusion and parliamentary trust; thus, there is no evidence that the orderings are artefacts of a value or linear form. Second, the focal exclusion-trust association holds in all six leave-one-out refits in both cases, meaning that it is not specific to any economy and even Serbia. Third, the skills factor, that is, the dimension considered decisive by the age literature, is operationalized for three countries only and is thus presented purely descriptively and never used as basis for inference. What can be inferred from the evidence is a structure and direction, however, and specifically that: (1) the factors of digital inclusion are decoupled, (2) the age gradient among them is steep, and (3) digital exclusion implies distrust in the representative bodies. What cannot be said is any estimate of the strength of an effect, or a causal interpretation of country-level relationships. Individual-level analysis in section 4.7 examines the structure of age-use relationships within countries yet does not state the country-level correlations as individual-level causes. Trust in political parties and parliaments, in turn, is driven by corruption, competition, clientelism, media capture, polarization, government effectiveness and disappointment with EU accession process. Digital exclusion, accordingly, represents just one element of this relation, but is not the reason behind low trust levels.

4.6. A typology of the resilience gap

The results generate a typology based on two variables: digital inclusion or exclusion and representative attachment or disaffection. The resilience deficit consists of the two cells off the diagonal: older citizens are more attached to yet digitally vulnerable to exclusion, while younger citizens are more included yet less attached to the representative institutions.

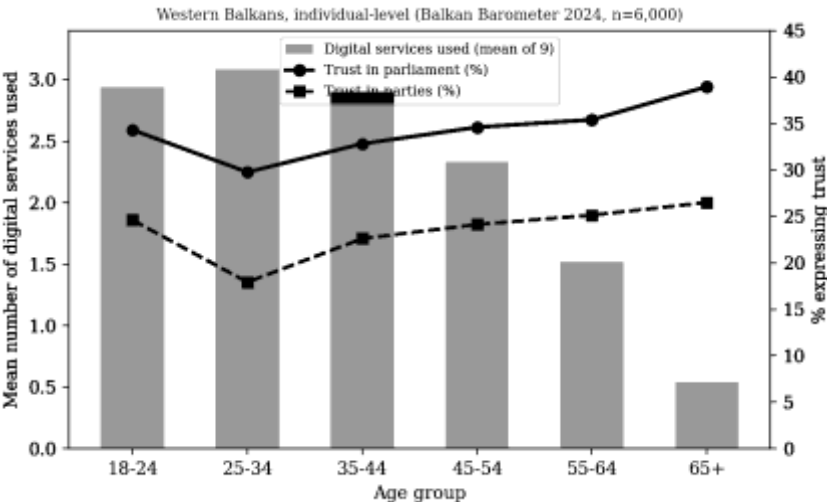
The six Balkan societies differ in terms of their placement in such a typology. Whereas Kosovo, Montenegro and Albania are closer to the corner with lower exclusion and higher trust, North Macedonia and Bosnia and Herzegovina resemble the opposite extreme of high exclusion and low trust. In Serbia, party competition and incumbency help overcome the exclusion problem generated by digital vulnerability.

A typology like this is particularly useful in explaining the positioning of trust in institutions. The reason why two off-diagonal groups refuse to cooperate with certain institutions is quite distinct. While the former feel more involved in those institutions but unable to use their channels to communicate with them, the latter can access the channels but see no point in using them since they find the institutions themselves unworthy of their attention. Consequently, the institutions that suffer on both accounts are parties and parliaments, which is exactly what the Balkan Barometer survey finds in terms of the lowest levels of trust. What is more important is that no other assumptions beyond digital inclusion and voting lead to precisely the attitude pattern discovered by the survey on its own.

4.7. Individual-level evidence on the age structure

In addition to revealing the resilience gap directly in the Western Balkan countries, the individual data also allow the age gradient to be quantified directly. Use of digital services is low for all individuals aged over 65 (mean = 0.5 services) compared to the youngest individuals aged between 18 and 24 years old (2.9 services). Moreover, use of no online service at all is high for the oldest citizens, with 71 per cent having not used any digital service versus seven per cent among the youngest. Trust works in the reverse direction, with increasing trust in parliament and parties from lowest in the younger age categories to highest among the oldest age categories (from 34 per cent among the youngest to 39 per cent among the oldest). This pattern is illustrated in Figure 5. In brief, the oldest citizens are both the most disconnected from digital life and the most institutionally attached to the parliament and political parties, while the youngest citizens are the least institutionally attached and most included in digital life.

Figure 5. Digital-service use and institutional trust by age group, Western Balkans individual data



Source: Author’s calculations based on (Balkan Barometer 2024, n = 6,000).

To quantify the association of interest, Table 5 presents logistic regression models predicting trust separately in parliament and parties. These models account for the influence of age, education, gender, whether the respondent lives in a city, and the economy-wide average level of digital use. Age proves to be a reliable predictor, where being one age band older significantly increases the odds of trusting parliament by about twelve per cent (OR = 1.12, $p < 0.001$) and trusting parties by six per cent (OR = 1.06, $p = 0.008$). Meanwhile, increased use of digital services negatively predicts trust in parties by around 6 per cent (OR = 0.94, $p = 0.002$) while having no effect on parliament trust (OR = 1.01, $p = 0.74$). Thus, a higher number of digital services used is associated with greater skepticism towards political parties, suggesting that young citizens can indeed be regarded as critical citizens based on the evidence presented.

While it is true that higher levels of digitization in the economy are associated with higher institutional trust levels, it is also clear that, among the digitally active, the odds of trusting parties are lower compared to those who have little use of online resources. The former can

be explained as a development pattern, whereas the latter should be interpreted in the critical-citizen light.

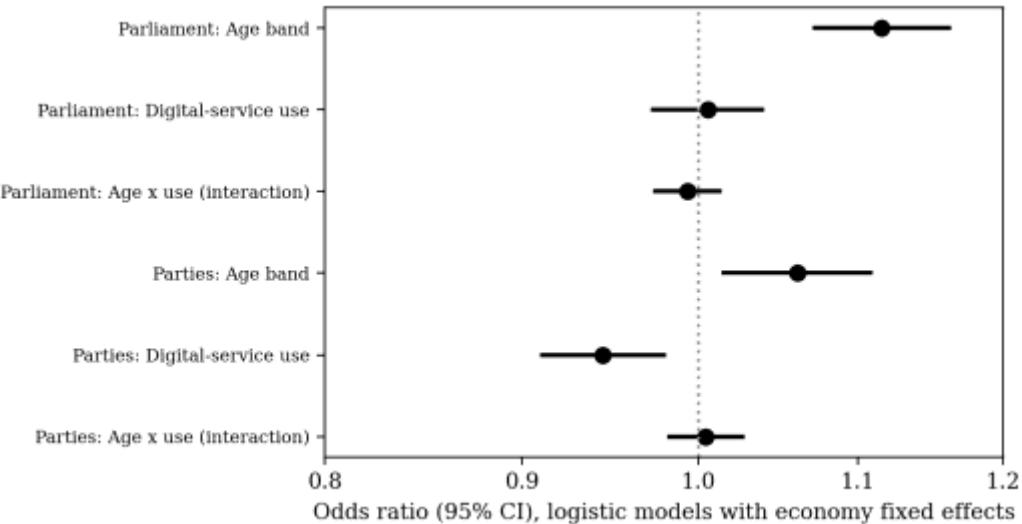
Table 5. Individual-level logistic models of institutional trust (odds ratios, economy fixed effects, robust standard errors).

Predictor	Parliament: (95% CI)	OR	p	Parties: (95% CI)	OR	p
Age band (per step)	1.12 (1.07-1.16)		<0.001	1.06 (1.02-1.11)		0.008
Digital-service use (per service)	1.01 (0.97-1.04)		0.738	0.94 (0.91-0.98)		0.002
Age x use (interaction)	0.99 (0.97-1.01)		0.512	1.00 (0.98-1.03)		0.692
Education	0.99 (0.91-1.08)		0.884	1.04 (0.94-1.14)		0.438
Male	1.10 (0.98-1.24)		0.092	1.14 (1.00-1.29)		0.047
Urban	1.00 (0.89-1.14)		0.970	0.92 (0.81-1.06)		0.259

Source: authors' estimates on Balkan Barometer 2024 individual records (RCC, 2024a), n = 5,669 complete cases.

*. Outcomes are binary, trust or tend to trust versus not. Each model includes economy fixed effects; standard errors are heteroskedasticity-robust; the public file carries no survey weight.

Figure 6. Odds ratios with 95 per cent confidence intervals for age and digital-service use, individual-level models.



Source: authors' estimates on Balkan Barometer 2024 individual records (RCC, 2024a), n = 5,669 complete cases.

The interaction between age and service use is not significant for either outcome (Table 5), so age does not moderate the use - trust relationship.

5. CONCLUSIONS

The digital re-linking hypothesis confronts an empirical test it was not designed for in the Western Balkans, and it fails the test. Whereas other European analyses assume connection will draw the citizenry closer to representative institutions, the six Balkan economies have reached levels of connection that approach saturation while leaving engagement behind. Penetration above ninety percent and relatively uniform in the six economies, but access, skills and usage diverge from each other independently, and so while they may suggest maturity when taken together on the basis of penetration. Composite penetration looks uniform and mature, but once disaggregated the economies diverge on skills and use. The critical gap has shifted from the margins of access to the skills and usages where it is now sorted by age.

That sorting defines the gap. The older generation exhibits a greater institutional attachment and tends to be more electorally loyal in the turnout literature while also holding the weakest digital position, politically present and digitally absent. The youngest generation, on the contrary, holds the strongest digital position while falling in the age bracket least electorally reliable according to the turnout literature and distancing itself from the institutions of representation, digitally present but politically absent. The two separations arrive at the same destination, where, regionally, political parties stand at twenty-two percent trust and the parliament at thirty-three percent trust as the least trusted institutions in the region, below even the police and judiciary. At the economy level, digital exclusion correlates with lower trust in both institutions with a rank correlation of -0.77 against parliamentary trust. The cooperative tie between citizens and their representatives is thin at both ends, which is why the current paper reads vertical legitimacy as a plausible resilience issue rather than as an outcome of resilience. The individual data clarify this: digital-service use drops from a mean of 2.9 in the youngest age group to a mere 0.5 in the oldest, trust in representative institutions rises with age and age interacts with service use with zero effect, indicating a compositional structure and not a difference in connection's effect at the two ends.

The policy recommendation is clear. It suggests that no single strategy will address both issues. With regard to the older generation, the relevant constraint is access and skills and must be closed, but the international experience cautions against focusing on access alone, as what relates to institutional trust is the way in which citizens utilize connection and not their access to it. Thus, a purely access-based policy could not be expected to strengthen the cooperative tie between citizens and their representatives. For the younger generation, the issue has nothing to do with access. It is already digitally connected, but what it has withdrawn is its consent to being represented via political parties and parliaments in their current form. An overall digital inclusion goal set by a single DESI-type score will miss both issues, because it considers access the final barrier while assuming a situation, which, in the case of the Western Balkans, is already near saturation. This is a necessary caveat. Only six countries in the sample offer a cross-national comparative perspective. Trust is assessed at one point in time. Individual-level regressions assess the within-economy age-skill structure, but not the causal relationships between age and skills nor longitudinal dynamics of the relationship, which would require the use of the multi-wave Balkan Barometer.

The resulting conclusion is narrower, but sharper as a result. In the Western Balkans, the access gap closes on its own, but the resilience gap does not. This digital policy will account for the connections made and neglect the represented relationship, thinned by disaffection and digital exclusion at opposite ends.

Resilience provides a frame for connecting the re-linking argument to vertical legitimacy. The current evidence identifies attrition in the represented relationship, age-sorted into two deficits, both of which cannot be addressed by digital inclusion alone. For countries in enlargement facing the challenge of internal instability and external pressure, this is a plausible source of fragility, though fragility remains an inferred rather than a measured variable.

REFERENCES

- Bennett, W. L. (2008). Changing citizenship in the digital age. In W. L. Bennett (Ed.), *Civic life online: Learning how digital media can engage youth* (pp. 1-24). The MIT Press. <https://doi.org/10.1162/dmal.9780262524827.001>
- Bermeo, N. (2016). On democratic backsliding. *Journal of Democracy*, 27(1), 5-19. <https://doi.org/10.1353/jod.2016.0012>
- Bieber, F. (2018). Patterns of competitive authoritarianism in the Western Balkans. *East European Politics*, 34(3), 337-354. <https://doi.org/10.1080/21599165.2018.1490272>
- Bieber, F. (2020). *The rise of authoritarianism in the Western Balkans*. Palgrave Macmillan. <https://doi.org/10.1007/978-3-030-22149-2>
- Boulianne, S. (2009). Does Internet use affect engagement? A meta-analysis of research. *Political Communication*, 26(2), 193-211. <https://doi.org/10.1080/10584600902854363>
- European Commission. (2024). 2024 report on the state of the Digital Decade. European Commission. <https://digital-strategy.ec.europa.eu/en/library/2024-report-state-digital-decade>
- Eurostat. (2024a). Skills for the digital age. Statistics Explained. https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Skills_for_the_digital_age
- Eurostat. (2024b). Young people - digital world. Statistics Explained. https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Young_people_-_digital_world
- Eurostat. (2025). Technology in enlargement countries - digital transformation. Statistics Explained. https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Technology_in_enlargement_countries_-_digital_transformation
- Friemel, T. N. (2016). The digital divide has grown old: Determinants of a digital divide among seniors. *New Media & Society*, 18(2), 313-331. <https://doi.org/10.1177/1461444814538648>
- Gallego, A. (2015). *Unequal political participation worldwide*. Cambridge University Press. <https://www.cambridge.org/core/books/unequal-political-participation-worldwide/4BF2185B049588E480739F066A58AD3B>
- Heinz, D. (2024). Trust and transparency in the Balkans: Key for navigating through complexity and change? *International Journal of Balkan Studies*, 5, 17-35. <https://ijbs.unyt.edu.al/wp-content/uploads/2026/03/Trust-and-transparency-in-the-Balkans.pdf>
- Hooghe, M., & Marien, S. (2013). A comparative analysis of the relation between political

- trust and forms of political participation in Europe. *European Societies*, 15(1), 131-152. <https://doi.org/10.1080/14616696.2012.692807>
- International IDEA. (2024). Voter turnout database. International Institute for Democracy and Electoral Assistance. <https://www.idea.int/data-tools/data/voter-turnout-database>
- Merkel, W. (2004). Embedded and defective democracies. *Democratization*, 11(5), 33-58. <https://doi.org/10.1080/13510340412331304598>
- Norris, P. (2001). *Digital divide: Civic engagement, information poverty, and the Internet worldwide*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139164887>
- Norris, P. (2011). *Democratic deficit: Critical citizens revisited*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139044899>
- Putnam, R. D. (2000). *Bowling alone: The collapse and revival of American community*. Simon & Schuster.
- Regional Cooperation Council. (2024a). *Balkan Barometer 2024: Public opinion survey [public opinion microdata and database]*. Regional Cooperation Council. <https://www.rcc.int/balkanbarometer/results/2/public>
- Regional Cooperation Council. (2024b). *Western Balkans digital economy and society index 2024 report*. Regional Cooperation Council. <https://www.rcc.int/pubs/211/western-balkans-digital-economy-and-society-index-2024-report>
- Regional Cooperation Council. (2025). *Western Balkans Digital Economy and Society Index 2025*. Regional Cooperation Council. <https://www.rcc.int/pubs/279/western-balkans-digital-economy-and-society-index-2025>
- Robinson, L., Cotten, S. R., Ono, H., Quan-Haase, A., Mesch, G., Chen, W., Schulz, J., Hale, T. M., & Stern, M. J. (2015). Digital inequalities and why they matter. *Information, Communication & Society*, 18(5), 569-582. <https://doi.org/10.1080/1369118X.2015.1012532>
- Scheerder, A., van Deursen, A., & van Dijk, J. (2017). Determinants of Internet skills, uses and outcomes: A systematic review of the second- and third-level digital divide. *Telematics and Informatics*, 34(8), 1607-1624. <https://doi.org/10.1016/j.tele.2017.07.007>
- Sloam, J. (2014). New voice, less equal: The civic and political engagement of young people in the United States and Europe. *Comparative Political Studies*, 47(5), 663-688. <https://doi.org/10.1177/0010414012453441>
- Smets, K., & van Ham, C. (2013). The embarrassment of riches? A meta-analysis of individual-level research on voter turnout. *Electoral Studies*, 32(2), 344-359. <https://doi.org/10.1016/j.electstud.2012.12.006>
- van Deursen, A. J. A. M., & van Dijk, J. A. G. M. (2019). The first-level digital divide shifts from inequalities in physical access to inequalities in material access. *New Media & Society*, 21(2), 354-375. <https://doi.org/10.1177/1461444818797082>
- van Dijk, J. A. G. M. (2005). *The deepening divide: Inequality in the information society*. SAGE Publications. <https://doi.org/10.4135/9781452229812>
- World Bank. (2016). *World development report 2016: Digital dividends*. World Bank. <https://doi.org/10.1596/978-1-4648-0671-1>

MIGRATION, SECURITY AND LOCAL IDENTITY: THE CASE OF DUBROVNIK

Marijana Musladin

Faculty of Media and Public Relations, University of Dubrovnik, Dubrovnik, Croatia

marijana.musladin@unidu.hr

ORCID iD: 0000-0002-6596-202X

This paper was produced within the research project “Migration and the Shaping of Dubrovnik Identity: Between History and the Present (MIDIA)”, carried out at the University of Dubrovnik and funded by the European Union through the NextGenerationEU programme.

Abstract

International migration has become closely associated with debates about security, social cohesion, and local identity, particularly in tourism-dependent cities that rely on international labour while also encountering humanitarian migration flows. This paper examines the relationship between migration, security, and local identity in Dubrovnik, a globally recognised tourist destination that has recently experienced both the arrival of refugees displaced by the war in Ukraine and a growing presence of labour migrants from Asian countries. Drawing on secondary literature and statistical data, the paper synthesises relevant theoretical literature, including securitization theory, human security, and theories of place identity, together with available national and regional statistical data and secondary studies on migration in Croatia, in order to outline how these two forms of migration are represented and experienced differently. The analysis suggests that Ukrainian refugees have generally been framed through a humanitarian lens associated with temporary protection and solidarity, while labour migrants from Asian countries have been more closely linked to economic necessity, labour shortages, and questions of integration, housing, and social adaptation. National-level survey data indicate increasingly ambivalent public attitudes toward foreign workers in Croatia, even as such migration remains essential to the tourism-dependent local economy. The paper concludes that migration in Dubrovnik should not be understood solely as a security issue, but also as a constitutive element in the ongoing transformation of local identity, and it calls for dedicated empirical research, such as discourse analysis of local media or resident surveys, to test these inferences at the local level.

Keywords: *migration; security; local identity; Dubrovnik; tourism-dependent city*

INTRODUCTION

International migration has become one of the most significant social and political processes of the twenty-first century, influencing demographic structures, labour markets, and public debates across Europe. In recent decades, migration has increasingly been connected with questions of security, social cohesion, and cultural change. While migration is often discussed in relation to borders and national security, its effects are also visible at the local level, where communities negotiate questions of belonging, identity, and social transformation.

Croatia offers a particularly instructive setting in which to examine these dynamics. For much of the twentieth century, and for the first two decades following independence in 1991, Croatia was primarily a country of emigration, with successive waves of citizens leaving for Western Europe, North America, and Australia in search of economic opportunity (Nejašmić, 1990; Valenta, Jakobsen, Gregurović & Župarić-Iljić, 2023). Croatia's accession to the European Union in 2013, and to the Schengen Area at the start of 2023, altered this picture considerably: emigration to other EU member states became easier, accelerating the outflow of younger workers (Draženović, Kunovac & Pripužić, 2018; Wankiewicz-Kłoczko, 2025; Wankiewicz, 2023), while at the same time the country began to attract growing numbers of labour migrants from outside the EU to fill the resulting gaps in its workforce, particularly in tourism, construction, and transport. This dual pattern of continuing emigration alongside rising immigration is central to understanding the contemporary Croatian, and Dubrovnik-specific, migration landscape discussed in this paper.

Tourism-dependent cities represent a particularly interesting context for understanding these processes. Their economies often rely on seasonal labour and international mobility, while demographic decline and labour shortages increase the need for foreign workers. At the same time, the arrival of new populations can influence local perceptions of community, cultural continuity, and social change.

Dubrovnik provides a relevant example of these dynamics. As one of Croatia's most recognisable tourist destinations and a UNESCO World Heritage Site, the city has long been shaped by international mobility and global connections. However, in recent years Dubrovnik has also experienced new migration processes, including the arrival of people displaced from Ukraine following the Russian invasion in 2022 (The Dubrovnik Times, 2022) and the growing presence of labour migrants from countries such as Nepal, India, the Philippines, and Bangladesh (JUGinfo.hr, 2025; totalinfo.hr, 2026).

These developments have become part of broader discussions about labour needs, integration, security, and the future of the city's local identity.

Understanding these dynamics has practical relevance beyond academic debate. Local and national policymakers in Croatia face the ongoing challenge of reconciling the tourism sector's structural dependence on foreign labour with public concerns about the pace and visibility of social change, while simultaneously honouring humanitarian commitments toward displaced populations.

Employers in tourism-dependent destinations such as Dubrovnik must navigate an increasingly complex regulatory environment, including new housing and compliance requirements for foreign workers, while addressing genuine labour shortages that many destinations across the Adriatic now report.

A clearer conceptual understanding of how different migrant groups are perceived, and why, is therefore relevant not only to scholars of migration and security but also to

municipal authorities, tourism boards, and civil society organisations engaged in questions of integration and urban development.

This paper explores the relationship between migration, security, and local identity in Dubrovnik through an overview of relevant academic literature, policy documents, and available statistical data. It examines how migration processes are connected to perceptions of security and social change, and how they contribute to the transformation of local identity in a tourism-dependent urban environment. The paper argues that migration should not be understood only as a security-related issue, but also as an important factor shaping the social and cultural development of contemporary cities.

1. THEORETICAL FRAMEWORK

This paper draws on three interconnected theoretical perspectives: the relationship between migration and security, the concept of local identity, and the characteristics of tourism-dependent cities. Together, these perspectives provide a framework for understanding how recent migration processes influence public perceptions of security and contribute to the transformation of local identity in Dubrovnik.

1.1 Migration and Security

Migration is broadly defined as the movement of people away from their usual place of residence to a new location, either within a country or across an international border. While no single definition of migration or migrant is universally agreed upon, long-term migration is commonly understood as residing outside one's country of origin for at least twelve months, and its causes range from economic opportunity and family reunification to conflict and forced displacement (International Organization for Migration - IOM, 2024a). This diversity of causes matters for the present analysis, since Dubrovnik's recent experience combines both economically motivated labour migration and forced displacement connected to the war in Ukraine.

Recent theoretical work has moved beyond earlier, static push-pull models of migration. De Haas (2021) proposes an aspirations-capabilities framework in which migration is understood not simply as a reaction to economic or political conditions, but as an expression of people's evolving aspirations and capabilities to choose where to live, including the capability to stay rather than move. From this perspective, both the labour migrants arriving in Dubrovnik and the refugees displaced from Ukraine can be understood as exercising different, unequally constrained forms of mobility, shaped by distinct sets of aspirations and structural opportunities rather than by a single, uniform migratory logic. Complementing this, recent conceptual work argues that migration and the broader category of mobility are best understood as overlapping ideal types, distinguished primarily by duration, intention, and institutional recognition rather than by one sharp boundary (Piccoli et al., 2024).

Security is a similarly contested concept. Baldwin (1997) argues that security lacks a single, universally accepted meaning, and that any definition must specify whose security is at stake and against which kind of threat. This conceptual openness is precisely what allows the meaning of security to be extended, through political and public discourse, to areas that are not inherently violent or military in nature — including migration.

Two broad traditions have shaped how this openness has been filled in migration studies. The first is state-centric: security is understood primarily as the protection of the state, its borders, and its social order, and migration becomes relevant to security insofar as it is seen

to challenge these. The second tradition, human security, takes the individual rather than the state as the primary referent object. Introduced by the UNDP's Human Development Report, human security is defined broadly as 'freedom from want' and 'freedom from fear' — safety from chronic threats such as poverty and repression, and protection from sudden disruptions to daily life, whether in the home, at work, or in the community (UNDP, 1994). Applied to migration, this perspective shifts attention away from the risks migrants are assumed to pose toward receiving societies, and toward the risks and vulnerabilities migrants themselves experience throughout the migration cycle, including exploitation, precarious labour conditions, and social isolation (International Organization for Migration - IOM, 2024b).

The relationship between migration and security has become a central topic in migration studies and international relations, particularly since the 1990s. While migration is primarily a social, economic, and demographic phenomenon, it is often interpreted through a security lens. The concept of the securitization of migration, developed by the Copenhagen School, argues that migration becomes a security issue not because it inherently threatens society, but because political actors, institutions, and the media present it as such (Buzan, Wæver & de Wilde, 1998).

According to securitization theory, issues become security concerns through discourse. When migration is portrayed as a threat to public order, national identity, or social cohesion, extraordinary policy responses become easier to justify. This approach has significantly influenced contemporary research on migration governance and border policies across Europe. Critics of the Copenhagen School, associated with the so-called Paris School of security studies, have argued that securitization is not only a matter of exceptional speech acts by political elites, but also results from the routine, everyday practices of bureaucracies, border guards, and security professionals, who construct migrants as objects of risk management through ordinary administrative procedures rather than dramatic political declarations (Bigo, 2002). This distinction matters for the present analysis because it suggests that the securitization of migration in a place like Dubrovnik need not depend on explicit political rhetoric; it can also occur gradually, through routine bureaucratic categorisation, permit systems, and everyday administrative practice. Building on this perspective, Huysmans (2006) argues that migration has increasingly been associated with insecurity through political discourse, linking migrants to crime, terrorism, and cultural threats despite limited empirical evidence supporting such associations. Research on the Western Balkan migration route confirms that these dynamics were clearly visible in the Croatian context. Šelo Šabić (2017) shows that the 2015–2016 refugee crisis produced a gradual shift from an initially humanitarian response toward an increasingly security-oriented framing of migration in Croatia. Studies of Croatian news media similarly find that migration coverage has repeatedly oscillated between humanitarian and security framings depending on the political and social context (Dobrić Basanež & Ostojić, 2021; Čepo, Čehulić & Zrinščak, 2020).

Securitization theory has not gone without criticism: some scholars argue that its original Copenhagen School formulation focuses too narrowly on elite discourse at the expense of migrants' material conditions, while others note that because securitization is identified through interpretive analysis of discourse, different researchers may reach different conclusions from the same material. These debates matter for the present paper because they underline the interpretive, rather than objectively measurable, nature of its claims about migration discourse in Croatia and Dubrovnik.

However, scholars also emphasise the importance of distinguishing between objective

security risks and socially constructed perceptions of insecurity. In many cases, public concerns are shaped less by migrants themselves than by political debates and media representations. This distinction is particularly relevant for analysing migration in Dubrovnik, where recent arrivals include both people displaced by war and labour migrants responding to economic demand.

More recent scholarship suggests that securitization is neither a uniform nor a permanent condition, but a contested and unevenly applied process. Comparing European responses to displacement from Ukraine with responses to migration from the Middle East and North Africa, Polko (2025) distinguishes between 'just' and 'unjust' securitization, showing that host states have applied markedly different security logics to different migrant groups depending on their perceived cultural proximity and the nature of their displacement — a distinction directly relevant to the contrast, discussed below, between the reception of Ukrainian refugees and Asian labour migrants in Dubrovnik. In a related contribution, Polko (2024) proposes the concept of 'flexicuritization', a more pragmatic and adaptive approach to migration governance that seeks to reconcile host-society security concerns with migrants' own needs, rather than treating security and hospitality as mutually exclusive. Other recent work describes a broader 'crisification of migration', in which successive shocks — terrorism, mass arrivals, pandemics — repeatedly reactivate security framings even where earlier gains toward desecuritization had been made (Planas Gifra, 2024). At the same time, desecuritization itself is increasingly understood not as a final, stable outcome but as an ongoing process actively contested by NGOs, international organisations, and civil society actors who challenge security-based framings of migration (Murray, 2024).

The European Union's own policy framework illustrates this ongoing tension between security-oriented and more managerial or humanitarian framings of migration. Recent analysis of EU migration policy discourse between 1989 and 2024 traces a gradual shift from overt securitization toward a more technical and managerial framing of migration, formalised in instruments such as the 2024 Pact on Migration and Asylum, even as exclusionary practices of containment and selective inclusion persist beneath the softer language (Učakar, 2025). Croatia's own policy adjustments — including new housing standards and compliance requirements for foreign workers introduced ahead of the Pact's rollout in mid-2026 (Government of the Republic of Croatia, 2026; Forbes Hrvatska, 2026) — can be read as a local instance of this broader European pattern, in which migration is simultaneously facilitated for economic reasons and subjected to increasingly technical forms of regulation and control.

For the purposes of this paper, security is treated primarily as a socially constructed concept in the sense proposed by the Copenhagen School: what matters is less whether migration objectively threatens Dubrovnik than how it tends to be discursively framed by political actors, media, and public opinion, and with what consequences (Buzan, Wæver & de Wilde, 1998). It should be noted that this paper does not carry out an original discourse or media analysis of Dubrovnik-specific political statements or news coverage; rather, this framing is approached indirectly, through existing research on migration discourse in Croatia more broadly and through available secondary and statistical data, with any inferences about the local level in Dubrovnik offered cautiously and flagged as such. This discursive lens is nonetheless useful for the paper's central concern, namely how different migrant groups arriving in Dubrovnik are likely to be perceived and talked about in different ways. At the same time, because the case under discussion includes both refugees and labour migrants, the analysis remains attentive to the human security dimension of

migration — that is, to the everyday vulnerabilities and needs of migrants themselves, such as housing, employment conditions, and integration (UNDP, 1994; IOM, 2024b) — rather than treating security solely as a matter of concern for the receiving society. This combined approach, informed by the 'just/unjust' distinction proposed by Polko (2025), allows the analysis to ask not only whether migration in Dubrovnik is securitized, but whether it is securitized differently for different groups, and with what implications for both host-community perceptions and migrants' own security.

1.2 Local Identity

The concept of local identity refers to the ways individuals and communities develop a sense of belonging to a particular place through shared history, cultural traditions, everyday practices, and collective memory. Rather than being fixed, local identity is understood as a dynamic process that evolves in response to social, economic, and demographic changes.

The idea that places carry a distinctive identity, shaped by the interaction of physical setting, everyday activity, and meaning, has a long tradition in human geography. Relph (1976) introduced the concept of place identity to describe how these three elements combine to give a location its particular character, and how this character becomes bound up with residents' own sense of self and belonging. This foundational idea has since been developed and complicated by later scholars, who emphasise that such identities are neither static nor purely local in origin.

Castells (2010) argues that identities are socially constructed through interactions between culture, history, and institutions. They are continuously reshaped as societies respond to globalization, migration, and economic transformation. From this perspective, migration does not simply introduce new populations but also influences how local communities define themselves and negotiate cultural change.

Similarly, Massey (1994) challenges the idea that places possess stable and homogeneous identities. Instead, she argues that places are continuously produced through social relations and interactions that extend beyond local boundaries. Migration therefore becomes one of the processes through which local identities are redefined, creating new forms of diversity while also generating debates about belonging and community.

In the context of Dubrovnik, local identity is strongly connected to its historical heritage, UNESCO World Heritage status, Mediterranean culture, and long-standing tourism tradition. The arrival of new residents and workers raises questions about how this identity is maintained while adapting to demographic and cultural change.

Local identity in tourism-dependent cities is further complicated by the fact that it is often actively marketed as part of the destination's brand. Municipal authorities, tourism boards, and private businesses have strong commercial incentives to present visitors with a coherent, marketable image of local identity built around heritage architecture, traditional cuisine, and historical narratives, even as the everyday lived experience of residents may be considerably more diverse and contested. This tension between identity as lived experience and identity as marketed heritage is particularly acute in UNESCO World Heritage cities such as Dubrovnik, where international recognition of the Old Town's historical value creates additional pressure to preserve a particular, curated image of the city, sometimes at the expense of accommodating the practical needs of a changing resident population, including newly arrived migrant workers and displaced persons who do not fit neatly into that curated image.

1.3 Tourism-Dependent Cities and Labour Migration

A tourism-dependent city or economy is generally understood as one in which tourism accounts for a disproportionately large share of local employment, income, and economic activity, making the local economy particularly sensitive to fluctuations in visitor flows. Recent research has proposed a systematic Tourism Economic Dependence Index to measure this reliance across countries, suggesting that economies in which tourism contributes more than approximately 20% of GDP can be considered highly tourism-dependent, and finding that such dependence has gradually increased since the global financial crisis (Ortega & Ribeiro, 2025). While this index was constructed at the national level, the same logic of disproportionate reliance on tourism-generated income and employment applies to individual destinations such as Dubrovnik, whose local economy is overwhelmingly structured around hospitality, accommodation, and related services. Rapid growth in tourism creates substantial demand for workers in hospitality, construction, transport, and related service sectors, while demographic ageing and population decline reduce the availability of the domestic workforce.

Recent research on tourism and migration demonstrates that many European destinations have become increasingly dependent on foreign labour to sustain economic growth (OECD, 2022). A quantitative analysis of EU member states similarly confirms that labour migration and international tourism reinforce one another, with immigration positively contributing to both tourism demand and labour-market outcomes (Přivarová, Martincová, Trnovský & Hačár, 2022). Labour migration is therefore not only a demographic phenomenon but also an essential component of local economic development. Salazar (2020) further argues that tourism and migration are best understood as overlapping, mutually constitutive forms of mobility rather than separate categories, and calls for greater attention to the sustainability of labour arrangements that underpin the tourism industry (Choe & Lugosi, 2022).

At the same time, the growing presence of migrant workers raises questions concerning integration, housing, labour rights, and social cohesion. Comparative European evidence from tourism-dependent cities such as Barcelona shows that precarious tourism employment and unaffordable housing markets are closely linked, producing residential displacement among tourism workers themselves (Valente, Zaragozaí, & Russo, 2023). This finding is significant for the Croatian case because it suggests that the pressures created by tourism-driven labour demand are not confined to migrant workers alone, but interact with wider housing-market dynamics that also affect domestic tourism employees, potentially amplifying, rather than isolating, the social visibility of migrant labour.

Croatia reflects these broader European trends. The expansion of tourism, combined with emigration and demographic decline, has led to a significant increase in the employment of foreign workers, particularly from countries such as Nepal, India, the Philippines, and Bangladesh. IRMO's analysis of foreign labour in Croatia identifies both the economic opportunities and the structural challenges — including housing, legal status, and integration — associated with this shift (Butković, Samardžija & Rukavina, 2022). This trend is especially visible in Dubrovnik, where tourism dominates the local economy and labour shortages have become a persistent challenge.

Croatian policy has responded to these pressures in several ways in recent years. The removal of the annual quota system for foreign workers in 2021 substantially simplified employers' access to non-EU labour, contributing to the sharp rise in residence and work permits recorded since. More recently, regulatory attention has shifted toward compliance and working conditions, reflecting the broader European alignment process discussed above.

The economic relationship between Croatia and migrant-sending countries is not one-directional. For migrant workers themselves, employment in Croatia is frequently tied to obligations toward family members remaining in the country of origin: a recent study of foreign workers in Croatia, mainly from the Philippines, Nepal, and India, included whether participants provide financial support to relatives back home among its core socio-demographic variables, and identified the burden of supporting families abroad, alongside high living costs and low relative wages, as a key source of the financial strain reflected in workers' comparatively low satisfaction with their standard of living (Tadić, Balabanić, Turk, 2026). This dimension is easily overlooked in discussions that focus solely on the receiving society's perspective, but it is central to a human-security understanding of migration: labour migration to a tourism-dependent city such as Dubrovnik is, for many workers, inseparable from strategies of household survival and investment that extend well beyond Croatia's borders. Recognising this transnational dimension helps explain why labour migration, despite its social visibility and the ambivalent attitudes documented at the national level, continues even as public sentiment toward foreign workers has become less favourable.

Nevertheless, both groups influence public discussions about the city's future, social cohesion, and identity. Tourism-dependent cities such as Dubrovnik therefore illustrate how migration simultaneously supports economic sustainability while generating new social and cultural challenges.

Taken together, these theoretical perspectives provide the conceptual framework for this study. Securitization theory explains how migration becomes associated with security concerns through public discourse, the concept of local identity highlights the changing nature of place and belonging, while research on tourism-dependent cities demonstrates why migration has become an essential feature of contemporary urban development. Applying these perspectives to the case of Dubrovnik allows for a better understanding of how migration is represented in public discourse and how these representations contribute to shaping perceptions of security and local identity.

2. MIGRATION, SECURITY AND LOCAL IDENTITY IN DUBROVNIK

Dubrovnik's recent migration experience has two distinct strands: the arrival of displaced persons from Ukraine following the Russian invasion in 2022, and a growing presence of labour migrants. These two forms of migration (humanitarian migration and labour migration) have produced different public narratives and influenced debates about security, integration, and local identity.

2.1 Recent Migration Trends: Refugees and Labour Migration

Croatia has experienced a significant transformation from an emigration country into a country increasingly characterised by immigration. According to the Croatian Bureau of Statistics (DZS), 70,391 people immigrated to Croatia in 2024, while 38,997 people emigrated, resulting in a positive net migration of 31,394 persons. Foreign citizens

represented 81.1% of all immigrants, demonstrating the growing importance of international migration for Croatian demographic and economic development. Among immigrants, the largest shares came from Nepal (14.2%) and Bosnia and Herzegovina (11.3%). The DZS data also include displaced persons from Ukraine who received temporary protection in Croatia (Croatian Bureau of Statistics, 2025).

Dubrovnik-Neretva County reflects these national trends. In 2024, the county recorded 2,381 immigrants from abroad and a positive international migration balance of 1,373 persons, indicating that migration has become an important component of local demographic change (Croatian Bureau of Statistics, 2025).

Labour migration has become particularly significant due to the economic structure of Dubrovnik. The tourism sector creates strong seasonal demand for workers in hospitality, accommodation, construction, and related services. According to the Ministry of the Interior, most third-country nationals require a residence and work permit in order to be employed in Croatia, making the number of permits issued an important indicator of labour migration trends (Ministry of the Interior of the Republic of Croatia, 2025). In 2024, Croatia issued 206,529 residence and work permits. Construction was the leading sector (75,071 permits), followed closely by tourism and hospitality (56,228 permits) (Ministry of the Interior of the Republic of Croatia, 2025). One of the first surveys of non-European foreign workers in Croatia, conducted by the Institute for Migration Research using a purposive, non-probability sample of respondents predominantly from the Philippines, Nepal, and India, found that most respondents reported being satisfied or mostly satisfied with life and work in the country, and that a majority regularly send remittances to their families of origin (Balabanić et al., 2025).

The shortage of available domestic labour is particularly visible in Dubrovnik. Data from the Croatian Employment Service show that unemployment in Dubrovnik-Neretva County remained relatively low during the 2024 tourist season. At the end of August 2024, 2,768 unemployed persons were registered with the Regional Office in Dubrovnik, representing a decrease of 12.6% compared with the same period in the previous year (Croatian Employment Service, 2024). These figures illustrate why employers increasingly rely on foreign workers to maintain tourism activities.

The composition of Croatia's foreign labour force has also shifted markedly in recent years. Whereas Bosnia and Herzegovina traditionally supplied the largest share of foreign workers, owing to linguistic and geographic proximity, more recent data point to the growing importance of workers from South and Southeast Asia. National statistics for the first seven months of 2025 show Bosnian nationals still leading with 21,444 residence and work permits, closely followed by Nepalese nationals with 21,289 — a gap of just 155 permits — with Serbian (18,275), Filipino (10,508), and Indian (10,328) nationals also among the largest groups; this represents a marked diversification compared with the labour migration patterns of a decade earlier. The Croatian Employers' Association (HUP) has noted that limited knowledge of the Croatian language, often compounded by workers' limited motivation to learn it, together with reported experiences of discrimination, remains among the principal challenges facing workers from these more distant countries of origin, even as their overall numbers continue to grow (Financije.hr, 2025). This diversification is directly relevant to Dubrovnik, where hospitality employers increasingly draw on the same national labour pools.

Alongside labour migration, Croatia has also experienced humanitarian migration connected with the war in Ukraine. Media analysis suggests that Ukrainian refugees have generally been represented through the framework of humanitarian responsibility and

solidarity, with their arrival associated primarily with temporary protection, accommodation, education, and access to public services (Holy & Geiger Zeman, 2025). By the end of January 2025, Croatia hosted over 26,000 non-EU citizens under temporary protection related to the war in Ukraine, a figure that, while modest relative to the scale of displacement seen in neighbouring EU member states such as Poland, nonetheless represents a historically unprecedented humanitarian intake for Croatia (European Commission, 2025a).

2.2 Migration, Public Discourse and Security Perceptions

Migration is often discussed through the perspective of security, particularly in societies experiencing rapid demographic and social change. According to securitization theory, migration becomes a security issue not only because of potential objective risks, but also through political and social processes that construct it as a matter of security and public concern (Buzan, Wæver & de Wilde, 1998). It should be stressed that this section does not present an original discourse or media analysis of political statements, local news coverage, or public debate specifically about migration in Dubrovnik. Rather, it draws on existing research on migration discourse at the national and regional level in Croatia, together with available statistical and secondary data, to outline patterns that plausibly apply locally; a dedicated study of Dubrovnik-specific political and media discourse — for instance, through content analysis of local news outlets or interviews with local stakeholders — would be needed to confirm these patterns and is suggested here as a direction for future research.

Nationally, the arrival of Ukrainian refugees under the EU's temporary protection mechanism has mainly been framed in terms of humanitarian responsibility, accommodation, and solidarity, consistent with the broader Croatian and European institutional response to displacement from Ukraine (European Commission, 2025b). Labour migration from Asian countries, by contrast, is more closely tied to economic necessity and labour shortages: in Dubrovnik-Neretva County alone, close to 13,000 residence and work permits were issued to foreign workers in 2024, placing the county among the Croatian regions with the highest number of foreign workers relative to population, even though its absolute numbers remain below those of the Split-Dalmatia and Istria counties (JUGinfo.hr, 2025). Croatian Employment Service data for the same year likewise identify Dubrovnik-Neretva County, alongside Split-Dalmatia and Primorje-Gorski Kotar, among the leading counties for seasonal employment nationally, with hospitality and accommodation accounting for the largest share of seasonal jobs (Croatian Employment Service, 2025). This growing and highly visible presence of foreign labour has been accompanied by practical challenges around accommodation and by a gradual process of mutual adaptation between migrant workers and the local community.

National-level survey data support this general pattern of increasingly ambivalent public attitudes, though it should be noted that this survey was conducted at the national level and does not provide Dubrovnik-specific figures; it is presented here as indicative of the broader Croatian context within which local dynamics in Dubrovnik are likely to unfold. A representative survey conducted in November 2025 found that only 3% of Croatian citizens expressed satisfaction with the presence of foreign workers in the country, while the share expressing outright dissatisfaction rose compared with the previous year (Institute for Migration Research, 2025). Comparative analysis of attitudes toward immigrants in Croatia likewise shows that public opinion is shaped less by direct contact with migrants than by broader contextual and political factors (Strabac, Valenta, Bell 2025;

Čačić-Kumpes, Gregurović & Kumpes, 2012).

These national-level differences suggest that migration is not perceived as a single social phenomenon, and a broadly similar differentiation can be expected at the local level in Dubrovnik. The meaning attached to migrant groups depends on the context of their arrival and the wider social narratives surrounding them. While some forms of migration are primarily understood through humanitarian perspectives, others are discussed in relation to economic dependency and concerns about social transformation.

On this reading, security concerns related to migration in Dubrovnik are less associated with direct threats and more with broader perceptions of change, including questions of social cohesion, demographic transformation, housing availability, and the preservation of local identity. This interpretation is consistent with Huysmans' (2006) argument that migration can become connected to wider feelings of insecurity in societies experiencing significant transformation.

The broader European policy context outlined in the theoretical framework is also likely to shape how migration is perceived locally, as Croatia aligns its migration governance with the EU Pact ahead of its 2026 rollout. Whether such measures are perceived locally as a reasonable formalization of existing practice, or as an additional layer of bureaucratic securitization affecting an already essential workforce, illustrates how the security-migration nexus discussed in the theoretical framework intersects with Dubrovnik's specific, tourism-dependent circumstances.

2.3 Migration and the Transformation of Local Identity

Dubrovnik's local identity has traditionally been shaped by its historical heritage, maritime culture, the legacy of the Republic of Dubrovnik, and its position as one of the most recognisable tourist destinations in the Mediterranean. However, the social and economic development of the city has significantly transformed everyday life and the relationship between local communities and global processes.

Tourism has played a central role in this transformation. The growth of international tourism has brought economic opportunities but has also influenced demographic patterns, housing availability, and the use of urban space. The historic city centre represents an important example of these changes. The decline of permanent residents in Dubrovnik's Old Town reflects broader processes of tourism pressure, short-term rentals, and changes in the housing market, which have affected the character of the city and raised questions about maintaining Dubrovnik as a living community rather than solely as a tourist destination. A qualitative study of tourism employees in Dubrovnik finds that residents perceive tourists as displacing them from everyday use of the Old Town, describing this as a long-standing structural problem tied to overtourism rather than a recent or isolated occurrence, and noting that this dynamic is reinforced by the growth of short-term rental platforms such as Airbnb (Abbasian, Onn, & Arnautovic, 2020). More recent municipal initiatives, including caps on daily visitor numbers within the city walls and campaigns aimed at encouraging permanent residency in the historic core, indicate that local authorities themselves recognise the risk of the Old Town being reduced to what some observers have described as an open-air museum, and have begun actively trying to reverse this trajectory (IntelliNews, 2024; Total Croatia News, 2026).

The trajectory observed in Dubrovnik's historic core finds clear parallels elsewhere in the Mediterranean, most notably in Venice, whose resident population within the historic lagoon city has declined for decades as housing has been progressively converted to tourist accommodation (InTrieste, 2025), a process frequently cited in the tourism literature as a

cautionary benchmark for other UNESCO-listed heritage cities (Bertocchi & Visentin, 2019). The comparison is instructive precisely because it illustrates two distinct trajectories available to small heritage cities facing similar pressures: continued, largely unmanaged decline of the resident population, as documented in Venice, or a more actively managed attempt to reconcile tourism growth with permanent residency, of the kind more recently pursued by Dubrovnik's municipal authorities through visitor caps and residency incentives (IntelliNews, 2024; Total Croatia News, 2026). Which of these trajectories ultimately prevails in Dubrovnik will shape not only the demographic future of the Old Town but also the social context into which migrant workers and displaced persons are received, since a hollowed-out historic core populated primarily by tourism businesses offers a markedly different setting for integration than a functioning residential community.

Migration adds another dimension to these ongoing transformations. The increasing presence of foreign workers, particularly during the tourist season, has become a visible part of everyday life in Dubrovnik. Migrants contribute to the functioning of the tourism economy by filling labour shortages in hospitality, accommodation, construction, and related sectors. Nationally, construction accounted for the largest single share of newly issued residence and work permits in 2024, followed closely by tourism and hospitality (Ministry of the Interior of the Republic of Croatia, 2025; JUGinfo.hr, 2025). At the same time, their presence raises broader questions about integration, belonging, and the future social character of the city. A recent qualitative study of Asian migrant workers in Zagreb found that, despite the precarious nature of their employment as delivery couriers and taxi drivers, their integration into Croatian society remained minimal, marked by weak social mobility, language barriers, and a degree of everyday isolation (Perkov, Švenda & Budimir, 2026) — a pattern that plausibly extends to seasonal workers in tourism-dependent Dubrovnik.

From the perspective of local identity theory, these changes should not be understood as the disappearance of Dubrovnik's identity, but rather as part of its continuous transformation. As Castells (2010) and Massey (1994) argue, identities and places are not fixed categories; they are constantly produced through social interactions, historical processes, and changing economic conditions. Migration therefore represents one of the processes through which local communities redefine ideas of belonging and negotiate the meaning of place. Viewed in this light, the seasonal presence of foreign hospitality workers and the more permanent settlement of some Ukrainian families are not simply external pressures acting upon a fixed Dubrovnik identity, but active ingredients in an ongoing, contested process through which that identity is continually reproduced and renegotiated.

CONCLUSION

This paper has examined the relationship between migration, security, and local identity in Dubrovnik, a tourism-dependent city experiencing significant demographic and social changes. Recent migration processes, including the arrival of Ukrainian refugees and the growing presence of labour migrants from Asian countries, demonstrate that migration has become an important element of the city's contemporary development.

The case of Dubrovnik shows that migration should not be understood only through the perspective of security threats. Instead, migration is connected with broader questions of social cohesion, integration, housing, labour needs, and cultural change. Different forms of migration are associated with different social perceptions: while Ukrainian refugees are primarily linked to humanitarian protection and solidarity, labour migrants are more often discussed in relation to economic necessity and the challenges of adapting to a changing social environment.

Migration also contributes to the transformation of local identity. Dubrovnik's identity has historically been shaped by heritage, tourism, and global connections, and contemporary migration represents another process through which the meaning of belonging and community is negotiated. Rather than causing the disappearance of local identity, migration can be understood as part of the ongoing transformation of the city in response to wider social and economic changes.

The example of Dubrovnik illustrates the challenges faced by tourism-dependent cities in contemporary Europe. These cities increasingly rely on international mobility to maintain economic activity while also seeking to preserve social cohesion and cultural continuity. Understanding migration as both an economic necessity and a social transformation is therefore essential for analysing the future development of local communities in a globalised world. As this paper is based on secondary literature and available statistical data rather than original fieldwork, several of its claims about Dubrovnik specifically remain inferences drawn from national and comparative research rather than directly established findings. Future research could usefully test and refine these arguments through primary data collection — for example, a content analysis of local media and political statements, or interviews and surveys involving residents, employers, and migrant workers in Dubrovnik.

REFERENCES

- Abbasian, S., Onn, G., & Arnautovic, D. (2020). Overtourism in Dubrovnik in the eyes of local tourism employees: A qualitative study. *Cogent Social Sciences*, 6(1), 1775944. <https://doi.org/10.1080/23311886.2020.1775944>
- Balabanić, I., Perić Kaselj, M., Rubić, M., Ružić, N. K., Turk, M., Nimac, D., Šimunić, N., Dabo, K. and Tadić, I. (2025). Istraživački izvještaj anketnog istraživanja o kvaliteti života i obrascima života stranih radnika iz Azije i Afrike u Hrvatskoj. Repository of the Institute for Migration Research. <https://repo.imin.hr/object/imin:188>
- Baldwin, D. (1997). The concept of security. *Review of International Studies*, 23(1), 5–26. <https://doi.org/10.1017/S0260210597000053>
- Bertocchi, D., & Visentin, F. (2019). "The Overwhelmed City": Physical and Social Over-Capacities of Global Tourism in Venice. *Sustainability*, 11(24), 6937. <https://doi.org/10.3390/su11246937>
- Bigo, D. (2002). Security and Immigration: Toward a Critique of the Governmentality of Unease. *Alternatives: Global, Local, Political*, 27, 63–92.

- <http://www.jstor.org/stable/45468068>
- Butković, H., Samardžija, V., & Rukavina, I. (2022). Strani radnici u Hrvatskoj: izazovi i mogućnosti za gospodarski i društveni razvoj. Institut za razvoj i međunarodne odnose (IRMO), Zagreb.
- Buzan, B., Wæver, O., & de Wilde, J. (1998). Security: A new framework for analysis. Lynne Rienner Publishers.
- Čačić-Kumpes, J., Gregurović, S. and Kumpes, J. (2012). Migracija, integracija i stavovi prema imigrantima u Hrvatskoj. *Revija za sociologiju*, 42 (3), 305-336. <https://doi.org/10.5613/rzs.42.3.3>
- Castells, M. (2010). The power of identity (2nd ed.). Wiley-Blackwell. A John Wiley & Sons, Ltd., Publication
- Čepo, D., Čehulić, M. and Zrinščak, S. (2020). What a Difference Does Time Make? Framing Media Discourse on Refugees and Migrants in Croatia in Two Periods. *Hrvatska i komparativna javna uprava*, 20 (3), 469-496. <https://doi.org/10.31297/hkju.20.3.3>
- Choe, J., & Lugosi, P. (2022). Migration, tourism and social sustainability. *Tourism Geographies*, 24(1), 1–8. <https://doi.org/10.1080/14616688.2021.1965203>
- Croatian Bureau of Statistics. (2025). Migration of population of the Republic of Croatia, 2024. DZS, Zagreb. <https://podaci.dzs.hr/2025/en/97256>
- Croatian Employment Service (Hrvatski zavod za zapošljavanje). (2025). Sezonsko zapošljavanje u 2024. godini. HZZ, Zagreb. <https://www.hzz.hr/statistika/sezonsko/sezonsko-zaposljavanje-u-2024-godini/>
- Croatian Employment Service (Hrvatski zavod za zapošljavanje). (2024). Stanje na tržištu rada Dubrovačko-neretvanske županije u kolovozu 2024. godine. Retrieved from <https://www.hzz.hr/novosti-po-zupaniji/stanje-na-trzistu-rada-dubrovacko-neretvanske-zupanije-u-kolovozu-2024-godine/>
- De Haas, H. (2021). A theory of migration: The aspirations-capabilities framework. *Comparative Migration Studies*, 9, 8. <https://doi.org/10.1186/s40878-020-00210-4>
- Dobrić Basanež, K. and Ostojić, P. (2021). Migration Discourse in Croatian News Media. *Medijska istraživanja*, 27 (1), 5-27. <https://doi.org/10.22572/mi.27.1.1>
- Draženović, I., Kunovac, M., & Pripužić, D. (2018). Dynamics and determinants of emigration: The case of Croatia and the experience of new EU member states. *Public Sector Economics*, 42(4), 415–447. <https://doi.org/10.3326/pse.42.4.3>
- European Commission (2025a). Migrant integration in Croatia. Migration and Home Affairs. https://home-affairs.ec.europa.eu/policies/migration-and-asylum/migrant-integration/migrant-integration-hub/eu-countries-updates-and-facts/migrant-integration-croatia_en
- European Commission (2025b). Extension of temporary protection and a common European path for the future of Ukrainians in the EU. Migration and Home Affairs. https://home-affairs.ec.europa.eu/news/extension-temporary-protection-and-common-european-path-future-ukrainians-eu-2025-06-04_en
- Financije.hr. (2025). Turistička sezona povećala broj sezonaca u Hrvatskoj. <https://financije.hr/turisticka-sezona-povecala-broj-sezonaca-u-hrvatskoj/>
- Forbes Hrvatska. (2026, April 20). Strani radnici po novom pravilniku moraju imati primjeren smještaj. Propisani minimalni uvjeti.
- Government of the Republic of Croatia. (2026, February 5). Izmjenama Zakona o strancima uvodi se obvezno učenje hrvatskog jezika.
- Holy, M., & Geiger Zeman, M. (2025). Osobne priče (i)migranata: Promjena medijskog

- uokvirivanja u Hrvatskoj? [Personal Stories of (Im)migrants: A Shift in Media Framing in Croatia?]. *Medijska istraživanja*, 31(1), 53–82. <https://doi.org/10.22572/mi.31.1.3>
- Huysmans, J. (2006). *The Politics of Insecurity: Fear, Migration and Asylum in the EU* (1st ed.). Routledge. <https://doi.org/10.4324/9780203008690>
- Institute for Migration Research. (2025). Public attitudes towards foreign workers in Croatia: Second wave survey (with Medianet), as reported in Telegram.hr, November 20, 2025. <https://www.telegram.hr/vijesti/veliko-istrazivanje-o-stranim-radnicima-ogromna-vecina-hrvata-ima-neki-problem-s-njihovom-prisutnoscu/>
- IntelliNews. (2024). Croatia moves to restrict Airbnbs to address housing crisis. bne IntelliNews. <https://www.intellinews.com/croatia-moves-to-restrict-airbnbs-to-address-housing-crisis-340368/>
- International Organization for Migration. (2024a). Fundamentals of migration. IOM, Geneva. <https://www.iom.int/fundamentals-migration>
- International Organization for Migration. (2024b). World migration report 2024, Chapter 5: Human security throughout the migration cycle. IOM, Geneva.
- InTrieste. (2025). As Venice's population declines, city expands efforts to manage tourism. <https://intrieste.com/2025/09/15/as-venices-population-declines-city-expands-efforts-to-manage-tourism/>
- JUGinfo.hr. (2025). Naše goste poslužuje sve više stranaca, ali broj sezonaca je pao. <https://juginfo.hr/najnovije/2025/01/10/nase-goste-posluzuje-sve-vise-stranaca-ali-broj-sezonaca-je-pao/>
- Massey, D. (1994). *Space, Place, and Gender* (NED-New edition). University of Minnesota Press. <http://www.jstor.org/stable/10.5749/j.ctttw2z>
- Migrants and Refugees Section, Dicastery for Promoting Integral Human Development. (2023). Country profile: Croatia. Vatican.
- Ministry of the Interior of the Republic of Croatia. (2025). Statistical data on residence and work permits issued from 1 January to 31 December 2024. <https://mup.gov.hr/gradjani-281562/moji-dokumenti-281563/stranci-333/statistika-169019/169019>
- Murray, J. P. (2024). Contesting the securitization of migration: NGOs, IGOs, and the security backlash. *International Studies Quarterly*, 68(4). <https://doi.org/10.1093/isq/sqae139>
- Nejašmić, I. (1990). Iseljavanje iz Hrvatske u evropske i prekomorske zemlje od sredine 19. stoljeća do 1981. godine – pokušaj kvantifikacije. *Migracijske i etničke teme*, 6 (4), 511-526. Preuzeto s <https://hrcak.srce.hr/127703>
- OECD (2022), *International Migration Outlook 2022*, OECD Publishing, Paris, <https://doi.org/10.1787/30fe16d2-en>.
- OECD (2025), *International Migration Outlook 2025*, OECD Publishing, Paris, <https://doi.org/10.1787/ae26c893-en>.
- Ortega, B., & Ribeiro, M. A. (2025). An index of the economic dependence on Tourism. *Tourism Economics*, 31(3), 426-452.
- Perkov, I., Švenda, N. and Budimir, A. (2026). “Work, Eat, Sleep, Repeat”: Iskustva stranih taksista i dostavljača u Zagrebu. *Kroatologija*, 17 (1), 45-84. <https://doi.org/10.59323/k.17.1.3>

- Piccoli, L., Gianni, M., Ruedin, D., Achermann, C., Dahinden, J., Hoffmeyer-Zlotnik, P., Nedelcu, M., & Zittoun, T. (2024). What is the nexus between migration and mobility? A framework to understand the interplay between different ideal types of human movement. *Sociology*, 58(5), 1019–1037. <https://hdl.handle.net/1814/76518>
- Planas Gifra, L. (2024). Securitizing migration in times of crisis: private actors and the provision of (in)security. *Cogent Social Sciences*, 10(1). <https://doi.org/10.1080/23311886.2023.2296601>
- Polko, P. (2024). The migration-security nexus: desecuritization and the shift towards the sustainable approach. *Journal of Modern Science*, 60(6), 86–104. <https://doi.org/10.13166/jms/197001>
- Polko, P. (2025). Just and unjust securitization of migration: a comparative analysis of migration to Poland from MENA countries and Ukraine. *Front. Polit. Sci.* 7:1464288. doi: 10.3389/fpos.2025.1464288
- Prívarová, M., Martinčová, M., Trnovský, K., & Hačár, D. (2022). Labour migration and tourism flows: The case of the EU. *Journal of Tourism and Services*, 24(13), 271–289. <https://doi.org/10.29036/jots.v13i24.374>
- Relph, E. (1976). Place and placelessness Vol. 67, p. 45. London: Pion.
- Salazar, N. B. (2020). Labour migration and tourism mobilities: Time to bring sustainability into the debate. *Tourism Geographies*, 24(1), 141–151. <https://doi.org/10.1080/14616688.2020.1801827>
- Šelo Šabić, S. (2017). The impact of the refugee crisis in the Balkans: A drift towards security. *Journal of Regional Security*, 12(1), 51–74.
- Strabac, Z., Valenta, M. & Bell, D.A. (2025) Public attitudes towards immigrants: A curious case of Croatia. *International Migration*, 63, e13341. Available from: <https://doi.org/10.1111/imig.13341>
- Tadić, I., Balabanić, I. and Turk, M. (2026). Work, Life, and Well-Being: A Study on Foreign Workers in Croatia. *Društvena istraživanja*, 35 (1), 79-103. <https://doi.org/10.5559/di.35.1.04>
- The Dubrovnik Times. (2022, April 26). More than 15,000 Ukrainians have found refuge in Croatia. <https://www.thedubrovniktimes.com/news/croatia/item/13418-more-than-15-00-ukrainians-have-found-refuge-in-croatia>
- Total Croatia News. (2026). From overtourism to role model: The remarkable turnaround of Dubrovnik. <https://total-croatia-news.com/news/overtourism-dubrovnik/>
- totalinfo.hr. (2026, February 21). Hoteli u lovu na radnike: Filipinci i Nepalci spašavaju sezonu. <https://totalinfo.hr/hoteli-u-lovu-na-radnike-filipinci-i-nepalci-spasavaju-sezonu/>
- Učakar, T. (2025). Securitization, humanitarianism, and the religious dimension of European migration policy. *Religions*, 16(9), 1190. <https://doi.org/10.3390/rel16091190>
- United Nations Development Programme. (1994). Human development report 1994: New dimensions of human security. Oxford University Press.
- Valenta, M., Jakobsen, J., Gregurović, M., & Župarić-Iljić, D. (2023). Changes in the Croatian migration system: Conceptualising the complexities of migrations, 1990–2023. *Labor History*, 65(4), 510–527.

- <https://doi.org/10.1080/0023656X.2023.2280052>
- Valente, R., Zaragozí, B., & Russo, A. P. (2023). Labour precarity in the visitor economy and decisions to move out. *Tourism Geographies*, 25(8), 1912–1928. <https://doi.org/10.1080/14616688.2023.2172603>
- Wankiewicz, P. (2023). Croatia: A decade in review and its impact. From EU accession to joining the Eurozone and the Schengen Area. European Institute of the Mediterranean. <https://www.iemed.org/publication/croatia-a-decade-in-review-and-its-impact-from-eu-accession-to-joining-the-eurozone-and-the-schengen-area/>
- Wankiewicz-Kłoczko, P. (2025). Croatia's demographic crisis: How to address depopulation. OSW Centre for Eastern Studies. Commentary. <https://www.osw.waw.pl/en/eksperci/paulina-wankiewicz-kloczko>.

SECONDARY LABOUR MIGRATION, LABOUR GOVERNANCE AND EMERGING VULNERABILITIES IN NORTH MACEDONIA

Sandra Tomovska

PhD Candidate, Faculty of Security – Skopje, UKLO
tomovska.sandra@gmail.com; ORCID: 0009-0006-8482-5123

Stojanka Mircheva

Faculty of Security – Skopje
stojanka.mirceva@uklo.edu.mk

Abstract

North Macedonia's labour market is changing rapidly under the combined pressure of outward migration, demographic decline, and the growing need to recruit foreign workers. While migration is often discussed in relation to border management and irregular mobility, less attention is paid to what happens after foreign workers enter the labour market. This paper examines secondary labour migration as post-entry mobility, including changes of employer, movement between formal and informal work, changes in status and possible onward movement. The focus is on whether current governance arrangements can follow workers' trajectories and identify vulnerability to labour-rights violations, exploitation, forced labour, human trafficking and associated risks before harm becomes severe. The study uses an exploratory, document-based analysis of official documents, available administrative data and relevant literature. Drawing on human security, precarity and layered-vulnerability perspectives, it finds that reliance on foreign labour is developing amid fragmented responsibilities, dispersed data and limited post-entry visibility. Mobility may be lawful and purposeful, and foreign workers are not inherently vulnerable. However, employer or intermediary dependency, recruitment costs, language barriers, status uncertainty, and weak access to complaint mechanisms may reduce workers' ability to leave harmful employment, seek protection and build resilience. The paper calls for a shift from a permit-centred to a worker-centred labour governance, supported by stronger coordination, reliable data exchange, ethical recruitment, accessible information and safe reporting mechanisms.

Keywords: *secondary labour migration; labour migration governance; labour exploitation; human trafficking; North Macedonia*

1. INTRODUCTION

North Macedonia is increasingly facing a migration and labour-market paradox. For many years, the country has been primarily viewed as a country of emigration and, since 2015, as part of the Western Balkan migration route. Recent demographic and labour-market trends, however, suggest that this perspective is no longer sufficient. The country is simultaneously experiencing population decline, demographic ageing, continued outward migration of working-age people and a growing need to recruit foreign workers in response to labour shortages (State Statistical Office of the Republic of North Macedonia, 2024; World Bank, 2024).

This shift broadens the way migration is understood. Alongside border management, irregular mobility and transit, migration is becoming a question of labour governance, institutional coordination, and worker protection. As the domestic labour pool narrows and employers increasingly seek workers from abroad, recruitment becomes a practical response to labour shortages. It also raises a set of governance questions: who recruits workers; whether pre-departure promises match the actual job; how employment and residence records are linked; whether workers can change employer or sector without losing status; how wages, contributions, working conditions and employer-provided accommodation are monitored; and whether workers can report abuse and obtain protection without fear of retaliation or automatic adverse migration consequences. These questions move the focus beyond admission as a single administrative event and towards the worker's full employment trajectory.

The institutional relevance of this issue is already visible in policy terms. The European Commission notes that North Macedonia's annual quota for work permits was set at 10,000 permits for 2025, while around 8,900 work permits were issued to foreign workers in 2024 (European Commission, 2025). These figures indicate that foreign labour recruitment is no longer marginal, but an increasingly formalised response to labour-market needs. In the context of demographic decline and labour shortages, such recruitment may be economically necessary. The central question is whether the institutional system is sufficiently prepared to manage it coherently, transparently and with adequate safeguards for workers.

This paper argues that North Macedonia's growing reliance on foreign labour is emerging faster than the institutional capacity to regulate, monitor and protect foreign workers throughout their employment trajectory. If responsibilities for employment, residence, taxation, labour inspection, migration control and protection from exploitation remain dispersed across different institutions, workers may be visible in some administrative records but invisible in others. The resulting gaps can weaken both the institutional response to labour-rights violations and the capacity to identify vulnerability to exploitation, forced labour, human trafficking and associated risks at an early stage. From this perspective, secondary labour migration should not be treated only as a labour-market phenomenon, but also as a human-security and governance concern that requires attention to both vulnerability and resilience: the conditions that increase exposure to harm, and the resources, rights and institutional safeguards that enable workers to avoid, resist or recover from it (United Nations Development Programme, 1994; International Organization for Migration, 2019).

Methodologically, the paper employs an exploratory, document-based analysis of institutional policy and governance. The empirical material consists of official policy and legal documents, written responses obtained from the Employment Service Agency, the Ministry of Interior and the Public Revenue Office of North Macedonia, available

administrative data on the employment of foreign workers, and selected operational observations relevant to post-entry mobility. These materials are compared across institutional mandates to identify what each institution records, where responsibilities intersect, and which parts of the worker's post-entry trajectory remain insufficiently visible. The findings are interpreted in dialogue with scholarly and policy literature on precarious migratory status, employer dependency, fair recruitment, labour exploitation, human trafficking, and migrant vulnerability. As an exploratory study, the paper neither measures the prevalence of exploitation nor evaluates individual institutions comprehensively nor establishes causal relationships. Its purpose is to identify governance conditions that may increase vulnerability or strengthen resilience, and to assess whether current arrangements can detect and respond to risk at an early stage.

The paper first sets out the relationship between labour migration, labour governance, human security, vulnerability and resilience. It then develops secondary labour migration as an analytical category and examines how post-entry mobility may intersect with precarious work, exploitation, forced labour or trafficking-related risk. The subsequent sections address demographic and labour-market change, fragmented institutional responsibilities, and post-entry data gaps, before presenting policy implications, limitations and recommendations.

2. CONCEPTUAL FRAMEWORK: LABOUR MIGRATION, LABOUR GOVERNANCE AND VULNERABILITY

The analysis of labour migration requires moving beyond an understanding of migration as a one-time movement across borders. Research on precarious migratory status and temporary labour-migration regimes shows that legal position, rights and access to protection may change after admission, and are shaped by the interaction between employment, residence rules, and institutional practice (Goldring et al., 2009; International Labour Organization, 2022). Admission is therefore an important administrative moment, but it is not the end of the migration process. A worker's experience is shaped by what follows: the conditions under which employment begins, whether the promised job corresponds to the actual job, whether wages are paid, whether accommodation is adequate, whether the worker can change employer, and whether institutions remain accessible when problems arise. A worker may remain with the same employer within the formal labour market, but may also change employer, move into informal work, lose regular status, leave the country, or become administratively invisible to the institutions responsible for employment, residence, taxation or inspection.

Labour governance refers to the legal, institutional, and administrative arrangements through which access to the labour market is regulated, employment relationships are supervised, and workers' rights are protected. In the case of foreign workers, labour governance is not limited to the issuing of a work permit or an opinion on the fulfilment of conditions for temporary residence for employment. It also includes residence procedures, employer obligations, taxation and social contributions, labour inspection, occupational safety and health, access to information, complaint mechanisms and coordination between institutions. The Employment Service Agency of the Republic of North Macedonia issues work permits for foreigners with regulated residence and provides opinions on the fulfilment of conditions for temporary residence for employment (Employment Service Agency of the Republic of North Macedonia, n.d.).

This creates a multi-actor system in which no single institution necessarily holds the full picture of the worker's trajectory.

The 2025 amendments to the Law on Foreigners (Official Gazette of the Republic of North Macedonia, No. 193/2025) show that the field is undergoing regulatory adaptation. The amendments concern the broader legal framework within which temporary residence and employment of foreigners are regulated (Assembly of the Republic of North Macedonia, 2025). From a governance perspective, regulatory adaptation should not be assessed only in terms of the speed or efficiency of admission procedures. It should also be assessed according to whether the system prevents abuse, identifies risk, enables mobility without loss of rights, and provides an effective remedy when rights are violated.

These governance questions are beyond administration. They concern the security, agency and dignity of the worker and therefore call for a human-security perspective. The concept shifts attention from the security of the border and the functioning of administrative systems to the security of the person. The 1994 Human Development Report introduced human security as a people-centred approach concerned with freedom from fear and freedom from want (United Nations Development Programme, 1994). In labour migration, the relevant questions include whether workers are protected from deception, exploitation, debt dependency, loss of status, unsafe working conditions, non-payment of wages and exclusion from services. Equally important is whether they have the information, lawful mobility options, social support, complaint mechanisms and remedies needed to withstand or recover from adverse conditions. Human security thus connects vulnerability with resilience: the former concerns exposure and reduced capacity to avoid harm, while the latter concerns the individual and institutional resources that enable workers to exercise agency, enjoy rights, and obtain protection. The state's legitimate interest in regulating entry and employment remains intact; the question is whether those arrangements also reduce exploitation and trafficking-related risks while preserving workers' dignity, autonomy, and safety.

Vulnerability, in this paper, is not treated as an inherent characteristic of foreign workers. Nor is it used as a vague synonym for disadvantage. It refers specifically to an increased exposure, combined with a reduced capacity to avoid, resist, report, or recover from labour-rights violations, abusive recruitment, labour exploitation, forced labour, human trafficking and associated social and protection risks. Describing all migrants as vulnerable can obscure differences in resources, skills, legal position, and personal agency, and can unintentionally reduce workers to passive recipients of protection. A more useful approach is to examine how vulnerability is produced through the interaction of individual circumstances, employment conditions, recruitment practices, legal rules and institutional responses. Sargeant and Tucker (2009) describe migrant-worker vulnerability as layered: risks may arise from migration-related factors, personal characteristics and the conditions of the receiving labour market. Similarly, the IOM's determinants-of-vulnerability approach emphasises that both risk and resilience are shaped by factors operating at individual, household, community and structural levels (International Organization for Migration, 2019).

This relational understanding is particularly relevant to labour migration because formal legality does not automatically eliminate vulnerability (Shamir, 2017). A worker may hold valid documentation and still have limited bargaining power if residence is intricately linked to one employer, if changing jobs is administratively difficult, or if the worker has incurred substantial recruitment debt. Anderson (2010) argues that immigration controls do not merely regulate the quantity of labour entering a country; they also shape

particular kinds of employment relationships and can contribute to the production of precarious workers. Goldring et al. (2009) likewise show that precarious migratory status is not simply a personal condition but can be institutionally produced through differentiated rights and uncertain pathways.

The concept of precarity therefore provides an important bridge between migration governance and labour exploitation. Precarity refers not only to temporary employment but to insecurity across several dimensions: income, legal status, housing, and access to services, capacity to claim rights and freedom to leave an abusive employment relationship. Lewis et al. (2015) use the term hyper-precarity to describe situations in which restrictive immigration regimes and insecure labour markets interact, creating conditions that can trap migrants in severe exploitation. The value of this concept is not that it labels all temporary migrant work as exploitative, but that it directs attention to the cumulative effect of multiple insecurities.

An additional layer of complexity concerns private employment agencies and other intermediaries. Foreign labour recruitment is often not a direct relationship between an individual worker and an employer. It may involve agencies, informal brokers, documentation support, travel arrangements, translation, information about employment conditions and accommodation. In North Macedonia, private employment agencies operate as registered labour-market actors, and the Ministry of Social Policy, Demography and Youth maintains a public list of such agencies (Ministry of Social Policy, Demography and Youth, n.d.). Mediated recruitment can facilitate lawful employment and help employers address labour shortages. At the same time, the recruitment chain itself becomes part of the governance problem when responsibilities are unclear, information is incomplete or costs are transferred to workers.

The International Labour Organization's fair recruitment framework is based on the principle that workers should not bear recruitment fees or related costs and that recruitment should be transparent, regulated and subject to effective oversight (International Labour Organization, 2019). Research on recruitment governance shows that abuse is often not isolated acts by individual recruiters but may arise from structural incentives within long and complex recruitment chains (Farbenblum, 2017). For receiving countries, this means that worker protection cannot begin only after an employment contract is registered. The conditions under which a worker was recruited, the promises made before departure and any debt incurred may strongly shape the worker's capacity to exercise rights after arrival.

3. SECONDARY LABOUR MIGRATION: CONCEPT, DRIVERS AND GOVERNANCE IMPLICATIONS

The term secondary migration is used differently across migration studies. In some contexts, it refers to the onward movement of refugees or migrants from the country or locality in which they initially arrived, to another destination. In other contexts, it describes internal relocation after settlement or movement between legal and administrative categories. Migration scholarship also shows that post-entry trajectories are dynamic and that migrants may move between more secure and more precarious legal, employment and social positions over time (Goldring et al., 2009; International Labour Organization, 2022). For labour migration, a narrower and more operational definition is therefore useful. In this paper, secondary labour migration is used as an analytical concept to describe significant post-entry mobility or change after a foreign worker's initial admission to the labour

market. This may include changing employer or sector, moving from formal to informal employment, losing or changing residence status, leaving the original workplace, moving within the country or continuing to another country. The definition builds on the broader literature on post-entry mobility and precarious migratory status, while adapting those insights to the relationship between employment, residence and institutional oversight in North Macedonia.

Secondary movement may be regular, desirable and unrelated to abuse. Mobility is a normal labour-market feature, and foreign workers may change jobs or relocate for the same economic, family or personal reasons as citizens. Its analytical relevance lies in the interaction between mobility, employer-linked documentation, residence rules and fragmented administration. The consequences of leaving a job may therefore depend on whether a clear pathway exists to change employer without losing regular status.

Workers may leave an initial employer because wages are unpaid, hours exceed expectations, actual tasks differ from the contract, accommodation is inadequate or a better offer becomes available. They may also move for reasons unrelated to abuse, including family circumstances or onward opportunities. Institutions should therefore neither treat every departure as irregularity nor ignore mobility because entry was regular. Monitoring should distinguish ordinary labour-market movement from situations involving deception, coercion, exploitation or loss of protection.

The concept is especially relevant where the initial employment relationship carries substantial administrative weight. Employer-sponsored migration systems often make access to legal status dependent, formally or practically, on the continuation of employment with a particular employer. Clibborn and Wright (2023) argue that tying temporary migrant workers to a single sponsor can produce unequal power and make it more difficult for workers to leave underpayment or mistreatment. Even where the law permits a change of employer, the procedure may be unclear, slow or difficult to navigate without assistance. The worker may fear that leaving the job will lead to loss of residence, detention, removal or an inability to recover unpaid wages. The practical effect can be a form of constrained mobility: the worker is legally present but has limited freedom to move within the labour market.

This tension between mobility and dependency is central to secondary labour migration. A worker who remains with an exploitative employer may be administratively visible but unprotected. A worker who leaves may regain personal autonomy but become less visible to formal institutions. In both situations, a permit-centred system can miss the underlying protection problem. The first worker appears compliant because the employment relationship remains formally intact. The second may appear non-compliant because the expected employment pathway has been interrupted. A worker-centred system would ask a separate set of questions: Why did the worker leave? Was the employment relationship consistent with the contract? Are wages owed? Is the worker safe? Is there a lawful route to change employer? Is an intermediary controlling documents, information, or movement?

Secondary labour migration can also emerge through the gradual erosion of regularity rather than a single deliberate decision. A contract may end unexpectedly; an employer may fail to complete required procedures; a worker may not understand a notification obligation; or documentation may expire while the worker is waiting for another job. Goldring et al. (2009) emphasise that migratory status should be understood as a continuum rather than a simple division between regular and irregular. People may move between more secure and less secure positions as their employment, documentation

and access to rights change. This perspective is important because it avoids moralising the loss of status and directs attention to the institutional pathways through which insecurity is produced.

The role of information is particularly important during these transitions. Workers who do not speak Macedonian or another language used by public institutions may rely entirely on employers, recruiters or co-national networks to understand their rights. This reliance can be supportive, but it may also create informational dependency. A worker may not know which institution is responsible for a complaint, whether changing employer is possible, how to regularise a new employment relationship, or whether reporting abuse will affect residence status. The IOM stresses that accurate, accessible and timely information is a core safeguard against unethical recruitment and labour exploitation (International Organization for Migration, 2023). Information should therefore be treated as part of labour protection, not as an optional communication activity.

Secondary labour migration also has a transnational dimension. Given North Macedonia's proximity to European Union labour markets, some workers may view the country as one stage in a broader trajectory. Such movement should not be presumed to constitute misuse. However, risk may increase when workers leave formal employment without reliable information and become dependent on informal brokers, smugglers or unregulated labour markets. Governance should therefore understand mobility without conflating it with criminality.

For North Macedonia, the concept of secondary labour migration is useful precisely because it reveals the limitations of administrative snapshots. A work permit records a decision at a particular point in time. An employment registration records another. A temporary residence permit records legal status for a defined period. None of these records, on its own, explains whether the worker remained in the job, received the agreed wage, changed employer, left the country or encountered abuse. The worker's trajectory exists across several systems. When those systems are not connected, institutions may know that a worker was admitted but may not know what happened afterwards.

Post-entry visibility should not be understood as surveillance of workers. Excessive monitoring can undermine privacy, reinforce suspicion and discourage contact with institutions. The purpose of visibility is protective and analytical: institutions should be able to identify whether employment and residence records remain coherent, whether complaints reveal recurring patterns, whether particular sectors or recruitment channels generate disproportionate risk, and whether workers who leave an employer have a safe route to maintain lawful status. Data systems should therefore be proportionate, rights-based and governed by clear rules on purpose, access and data protection.

A secondary labour migration perspective also broadens the meaning of successful policy. Success should not be measured only by permits issued or vacancies filled, but also by retention under decent conditions, lawful job mobility, payment of wages and contributions, safe working and living conditions, access to remedy and institutional capacity to identify risk. Mobility itself is not necessarily failure; failure occurs when workers are pushed outside protection because institutions cannot follow employment changes or provide a route back into formality.

4. FROM PRECARIITY TO EXPLOITATION: RISKS ASSOCIATED WITH SECONDARY LABOUR MIGRATION

The relationship between secondary labour migration and exploitation requires careful interpretation. Available evidence does not establish systematic exploitation of foreign workers in North Macedonia, and post-entry mobility is not proof of victimisation. International research nevertheless shows that recruitment debt, employer dependency, insecure status, language barriers, isolation and weak access to remedy can combine to increase risk. This section examines how ordinary labour difficulties may become more serious when workers have limited freedom to leave or report them.

Labour exploitation is best understood as a continuum rather than a single category. At one end there are decent and freely chosen employment relationships. Further along are labour-law violations such as unpaid overtime, excessive working hours, unsafe conditions, unlawful deductions, or substandard accommodation. More severe situations may involve deception, threats, restriction of movement, withholding of identity documents, debt bondage, or coercion. Forced labour and trafficking for labour exploitation represent the most serious end of this continuum. The existence of lesser violations does not automatically mean that trafficking has occurred, but repeated and cumulative abuses can create conditions in which a worker's freedom becomes constrained.

The distinction has practical consequences. If institutional concern begins only at the trafficking threshold, workers facing serious but less visible abuse may receive no support. Yet labelling every labour violation as trafficking would weaken legal precision. A graduated response is therefore needed, involving labour inspection, employment and tax authorities, migration institutions, social services, trade unions and specialised anti-trafficking actors according to the nature and severity of the indicators.

Recruitment debt is one of the most widely documented risk factors. A worker who has paid substantial fees for recruitment, travel, documents or placement may begin employment under immediate financial pressure. The debt may be owed to a formal agency, an informal broker, relatives or lenders. Even where the worker is legally free to leave, the economic consequences of returning home without recovering the investment may make departure unrealistic. The ILO's fair recruitment principles therefore state that workers should not be charged recruitment fees or related costs (International Labour Organization, 2019). Recruitment debt can transform a formally voluntary employment relationship into one marked by strong economic compulsion.

Deception during recruitment can deepen this dependency. Workers may receive incomplete or inaccurate information about wages, working hours, job tasks, accommodation, deductions or the identity of the actual employer. Complex recruitment chains make accountability more difficult because the employer may blame the recruiter, while the recruiter may claim that conditions changed after arrival. Farbenblum (2017) argues that recruitment governance requires a rights-based framework capable of addressing systemic recruiter misconduct rather than relying only on individual complaints. For receiving countries, this implies a need to verify not only that a contract exists, but also that workers understand its terms and have access to a remedy when the actual conditions differ from those promised.

Employer dependency may become particularly significant where the worker's residence status, accommodation and income are all connected to the same employment relationship. Losing a job can then mean losing several forms of security at once. A worker may have to choose between remaining in harmful conditions and facing unemployment,

homelessness or uncertainty regarding legal status. Anderson (2010) describes how immigration controls can shape workers' relations with employers, while Clibborn and Wright (2023) show that employer sponsorship can reduce bargaining power and contribute to underpayment and mistreatment. These findings do not mean that employer-linked systems inevitably produce abuse. They demonstrate why safeguards for changing employer and maintaining status are essential.

Document retention is another relevant indicator. Although employers or intermediaries may keep passports temporarily for administrative procedures, workers must retain free access to their documents. Control over passports, bank cards, wages, phones or transport can restrict movement and deepen dependency. Institutions and employers should communicate clearly, in languages workers understand, that identity documents and individual property remain under workers' control.

Non-payment or delayed payment of wages can become a major driver of secondary movement. Workers may leave an employer because they cannot sustain themselves, even if doing so places their status at risk. Wage theft is particularly difficult to challenge when workers lack written evidence, fear retaliation or expect to leave the country soon. The problem may remain invisible if complaints are the main trigger for inspection. Workers with the greatest dependency may be the least likely to complain. This creates a structural paradox: institutional systems may interpret the absence of complaints as the absence of exploitation, even though fear and dependency suppress reporting.

Unsafe or degrading employer-linked accommodation can deepen employment vulnerability. Workers may fear that complaints will jeopardise both housing and work, while deductions for accommodation, transport or food can reduce actual earnings. Because these issues span labour, housing and migration mandates, they may fall between institutions. Where accommodation is provided or controlled by the employer, it should form part of labour-governance oversight.

Language barriers and social isolation affect workers' ability to assess and respond to risk. A worker who cannot read a contract, communicate with inspectors or independently verify information may become dependent on employers, recruiters or co-national networks. Multilingual information should therefore cover employment rights, emergency contacts, procedures for changing employer, health and safety, and the consequences of reporting abuse for residence status.

Gender, age, nationality, sector, skill level and family circumstances may create additional layers of vulnerability. A young worker in a hazardous sector, recruited through an intermediary, housed by the employer and unable to speak the local language may face a different risk profile from a highly skilled worker with independent housing and strong networks. Risk assessment should therefore be contextual and evidence-based, not based on nationality or broad assumptions.

Human trafficking for labour exploitation requires specific legal elements and should be identified by competent authorities based on evidence. Nevertheless, prevention begins before those elements are fully visible. The IOM's research on migrant vulnerability to trafficking and exploitation demonstrates that unsafe migration, abuse during the journey, debt, lack of information and restricted access to assistance can interact in ways that increase risk (International Organization for Migration, 2017). In labour-migration systems, early warning indicators may include inconsistent contracts, unexplained movement of workers between employers, repeated non-payment complaints, document retention, excessive recruitment fees, and isolation, as well as threats linked to status or patterns involving the same intermediary.

Secondary labour migration can be both a response to exploitation and a source of new vulnerability. Leaving an abusive employer may be an act of agency and self-protection. Yet if the worker cannot lawfully change employer, recover documents or maintain residence, the departure may push the person into informal work or onward movement. The worker may then become more dependent on informal networks, recruiters or smugglers. This is why policy should avoid creating a choice between tolerating abuse and losing status. Accessible job mobility and short transition periods between employers can function as anti-exploitation safeguards.

Effective prevention often depends on ordinary labour-governance measures rather than exceptional migration controls. Clear contracts, prohibition of worker-paid recruitment fees, pre-departure and post-arrival information, confidential complaints, proactive inspection, traceable wage payments, translation, legal assistance and accessible employer change can reduce the space in which exploitation develops. Specialised anti-trafficking mechanisms should be connected to mainstream labour institutions.

For North Macedonia, the immediate priority is to build the capacity to know more rather than to presume a hidden, large-scale exploitation problem. This requires data on employment changes, inspections in sectors employing foreign workers and trusted advisory channels that workers can use before situations become acute. Prevention is strongest when workers do not have to become victims to become visible.

5. DEMOGRAPHIC EROSION, LABOUR SHORTAGES AND THE TURN TO FOREIGN LABOUR

North Macedonia's increasing reliance on foreign labour must be understood in the context of population decline, ageing, low fertility, outward migration and a narrowing domestic labour pool. Together, these trends create long-term pressure on the labour market and explain why foreign recruitment is becoming a more visible economic and policy response.

The demographic challenge is reflected in population decline after the 2021 census and the growing share of older persons (State Statistical Office of the Republic of North Macedonia, 2024; World Bank, 2024). In 2024, the labour force comprised 792,779 persons, including 694,506 employed and 98,273 unemployed (State Statistical Office of the Republic of North Macedonia, 2025). Lower unemployment should therefore be interpreted cautiously: in a shrinking and ageing society, it may also reflect a smaller pool of available workers.

This is the core migration and labour-market paradox: North Macedonia continues to lose working-age people through outward migration while employers report increasing recruitment difficulties. Labour supply is shaped not only by unemployment, but also by the age, skills, sectoral distribution and mobility of the domestic workforce. The consequences of sustained emigration are therefore demographic, economic and institutional.

Low fertility and a negative natural increase further reduce the number of future labour-market entrants and place pressure on pensions, social services and productivity. Demographic decline is therefore not merely background context, but a structural factor affecting the state's long-term capacity to sustain its labour market and social-protection model.

Policy response has become more visible in recent years. According to the European Commission, North Macedonia issued around 8,900 work permits to foreign workers in 2024, while the annual quota for 2025 was set at 10,000 permits (European Commission, 2025). Data obtained from the Employment Service Agency also show a clear increase in institutional activity: processed requests increased from 3,672 in 2020 to 10,143 in 2025, and opinions issued for the employment of foreigners increased from 2,545 to 8,018 in the same period (Employment Service Agency of the Republic of North Macedonia, 2026).

The turn to foreign labour is more than an economic response; it is also a social, institutional and protection challenge. Behind each permit, quota or employment contract is a person who has moved in search of income, stability and a better life. Foreign workers may help address immediate shortages, but their recruitment creates new responsibilities for institutions, employers and intermediaries. Governance must therefore extend beyond the point of entry and cover the worker's subsequent trajectory in the labour market. A worker may remain in the job for which admission was granted, but may also change employer or sector, lose employment, enter informal work or continue to another destination. Such movement may reflect ordinary labour-market agency. It may also be driven by unmet recruitment promises, non-payment of wages, poor conditions, dependence on intermediaries or fear of losing residence status. A system focused primarily on permits cannot adequately distinguish between these possibilities. Post-entry coordination and accessible pathways for lawful job mobility are therefore essential for worker protection and for preventing vulnerability from escalating into exploitation or trafficking-related risk.

The quality of governance will also influence the sustainability of labour recruitment. A system associated with unclear procedures, unmet promises or weak protection may struggle to retain workers and may encourage onward movement. Conversely, transparent recruitment, decent working conditions and accessible institutions can make North Macedonia a more credible destination. Worker protection is therefore not in tension with economic competitiveness. It is part of a stable labour-migration model.

6. INSTITUTIONAL FRAMEWORK AND FRAGMENTED RESPONSIBILITIES

The governance of foreign labour in North Macedonia is distributed across several administrative systems that regulate various aspects of the foreign worker's status: access to employment, temporary residence, tax and contribution records, labour inspection, migration control and protection from exploitation. Such a multi-actor arrangement is not unusual. The challenge appears when these institutions do not generate an integrated picture of the worker's trajectory after entry into the labour market.

The data obtained from the Employment Service Agency show that foreign labour recruitment is becoming a regular part of labour-market administration. Requests submitted by the Ministry of Interior concerning quota use and current labour-market needs increased from 2,990 in 2020 to 9,307 in 2025, while negative opinions to the Ministry of Interior increased from 379 to 905 in the same period (Employment Service Agency of the Republic of North Macedonia, 2026). These figures show that the Employment Service Agency does not operate in isolation. Its work is intricately linked to procedures initiated or processed through the Ministry of Interior, particularly in relation to temporary residence for employment.

The response received from the Public Revenue Office further illustrates this fragmentation. The Public Revenue Office stated that it does not hold the requested aggregate data on foreign nationals registered for payment of contributions and/or personal income tax based on employment or work engagement in North Macedonia. Instead, it indicated that the source institution for employment-related registration is the Employment Service Agency, where foreign nationals, like domestic residents, are registered through the M1 employment registration form. The response also noted that, for foreign nationals who have obtained residence on the basis of employment or posting, an additional SPO form is completed, while residence permits fall within the institutional domain of the Ministry of Interior (Public Revenue Office of the Republic of North Macedonia, 2026).

The response from the Ministry of Interior adds another essential element. After checking with the relevant organisational units, the Ministry indicated that the Unit for Foreigners and Readmission within the Department for Border Affairs and Migration does not possess records containing the requested labour-related data, and that this type of matter falls outside its competence (Ministry of Interior of the Republic of North Macedonia, 2026). The response is better understood as evidence of fragmented governance than as institutional failure. The institution responsible for foreigners, residence and migration-related procedures does not necessarily hold the labour-related data needed to analyse the full position of foreign workers after entry.

Table 1 Institutional Visibility of Foreign Workers in North Macedonia

Institution / actor	What it captures	What remains limited
Employment Service Agency of the Republic of North Macedonia	Work-related requests, opinions, work permits and employment registration procedures for foreign workers.	Does not provide a complete picture of residence trajectory, post-entry mobility, changes of employer or movement outside formal employment.
Ministry of Interior	Temporary residence, foreigners' status and migration-related procedures.	Does not hold the requested labour-related records needed to analyse the full employment trajectory of foreign workers after entry.
Public Revenue Office	Tax and contribution-related system.	Does not hold the requested aggregate data on foreign nationals by employment or work engagement and refers to other institutions as source holders.
Labour inspection authorities	Labour rights violations, working conditions, occupational safety and employer compliance.	Foreign workers may become visible only when a complaint, inspection or detected violation activates the system.
Private employment agencies and intermediaries	Recruitment pathways, documentation support and connections between employers and workers.	Risks may arise when recruitment fees, promises, contracts, accommodation or working conditions are not transparent or sufficiently monitored.

Institution / actor	What it captures	What remains limited
Anti-trafficking and protection actors	Identification, referral and specialised support in cases involving indicators of trafficking or severe exploitation.	May become involved only after serious indicators emerge; prevention depends on links with mainstream labour and migration institutions.

Note. Authors' synthesis is based on institutional responses from the Employment Service Agency, the Ministry of Interior and the Public Revenue Office, as well as relevant public institutional information and the analytical framework developed in this paper.

Table 1 shows that the problem is not the complete absence of institutions, but the absence of an integrated view of the foreign workers' trajectory after entry into the labour market. Taken together, the institutional responses reveal a system in which foreign workers are administratively visible in parts, but not necessarily visible as people moving through a full labour-migration trajectory. One institution may register employment, another may process residence, a third may record tax or contribution obligations, while labour inspection may become involved only if a violation is reported or detected. This creates a governance gap between formal admission and actual protection.

The role of private employment agencies and intermediaries adds another layer to this picture. Mediated recruitment can help employers and workers connect lawfully, but it can also create dependency if recruitment chains are not transparent. For this reason, the monitoring of foreign labour should not stop at the point of issuing permits or opinions. It should also include the recruitment pathway, the employment relationship, accommodation that was linked to employment, and the workers' access to information and complaint mechanisms. Institutional fragmentation is therefore not a purely technical problem; it has direct implications for protection.

Data integration, however, must not become unrestricted exchange of personal information. A rights-based system requires clear legal purposes, data minimisation, role-based access and safeguards against the use of complaint data for punitive migration enforcement. Workers are more likely to seek help when they trust that reporting labour abuse will not automatically expose them to adverse immigration consequences. Institutional coordination must therefore combine administrative effectiveness with confidentiality and procedural safeguards.

7. DATA GAPS, POST-ENTRY VISIBILITY AND EMERGING VULNERABILITIES

The practical consequence of institutional fragmentation is limited post-entry visibility. In labour migration governance, the moment of entry is only the beginning of the worker's trajectory. A foreign worker may enter through a formal procedure, obtain residence and work-related documentation, and be registered in employment. The governance challenge begins after this point: whether the worker remains with the same employer, changes workplace, leaves the country, moves into informal work, becomes dependent on intermediaries, or becomes exposed to exploitative conditions.

The institutional responses analysed above suggest that the issue is not simply the absence of data, but the absence of an integrated analytical picture. If a worker leaves an employer, moves to another sector, exits the country, loses regular status or enters informal work, such movement may not be easily captured through one institution alone. The worker may remain visible in one part of the system and invisible in another. This limits the capacity of the state to distinguish between regular mobility, administrative non-compliance, vulnerability, exploitation risk and irregular onward movement.

Recent reporting based on data obtained from state institutions further illustrates the relevance of post-entry mobility. Radio Free Europe/Radio Liberty reported that 6,128 foreign workers received temporary residence permits for employment in North Macedonia in 2025, compared with 1,704 five years earlier. The same reporting states that the Ministry of Interior recorded 166 foreign nationals who left North Macedonia while holding temporary residence status for employment, including 129 nationals of Bangladesh, 34 of Nepal and three of India (Samardžiski, 2026). These figures neither prove irregularity, criminality nor exploitation. Their significance is more limited but still important: labour migration does not end when a permit is issued.

A similar point appears in operational humanitarian reporting. The April 2026 monthly report of the Red Cross of the Republic of North Macedonia notes cases relevant to movements in the border area, including Blace, indicating that onward movement remains part of the broader migration and mobility landscape (Red Cross of the Republic of North Macedonia, 2026). These observations are used here as contextual evidence, not to criminalise foreign workers or portray mobility as inherently problematic. They suggest that the boundaries between labour migration, transit, onward movement and vulnerability may become blurred when institutional monitoring is fragmented.

Workers who leave the expected administrative pathway may become harder to protect, although neither vulnerability nor irregularity can be inferred from movement alone. A worker who loses contact with the formal employer, depends on an intermediary, does not speak the language, does not understand the procedure or fears losing residence status may be less likely to report non-payment of wages, unsafe working conditions or coercive practices. In such circumstances, vulnerability emerges from the interaction between labour-market demand, recruitment practices, legal status, and access to information and institutional capacity.

Reliable data is a precondition for prevention, not merely a technical requirement for statistics. Without communication between employment, residence, tax, inspection and protection systems, institutions may respond only after a problem becomes visible through an incident, complaint or border event. Available evidence does not support firm conclusions about the scale of labour exploitation among foreign workers in North Macedonia.

It does, however, indicate that current governance arrangements may allow vulnerability to remain under-detected.

A future monitoring framework could combine aggregate administrative indicators with qualitative protection indicators. Relevant variables might include sector and employer, country of nationality, recruitment channel, changes of employer, termination of employment, residence validity, labour-inspection findings, wage-related complaints, occupational injuries, requests for legal assistance and referrals involving indicators of exploitation. This data should be used to identify patterns rather than stigmatise individuals or nationalities. Repeated problems connected to a particular sector, recruiter or employment model, may be more informative than isolated cases.

The absence of data should also be treated as a finding. When institutions cannot answer how many foreign workers remain employed, change employers, contribute to social insurance, leave the country or access complaint mechanisms, policy is forced to rely on assumptions. This can generate two opposite errors: underestimating exploitation because few cases are formally identified, or overestimating irregularity because worker movement is not adequately explained. Better evidence is necessary for both protection and public confidence.

8. DISCUSSION

The findings support the paper's central argument: secondary labour migration is a labour-governance issue, not only a question of mobility, irregularity or labour-market adjustment. North Macedonia's turn towards foreign labour is occurring amid demographic pressure and increasing administrative activity, while institutional mandates remain separated. Individual procedures can be processed, but the worker's trajectory across recruitment, employment, residence, taxation, mobility and protection is difficult to reconstruct.

Secondary labour migration makes this governance gap visible. Changing employers, leaving a workplace or moving onward may reflect agency, dissatisfaction, family strategy or a rational search for better conditions; it may also respond to deception, non-payment, unsafe work or intermediary dependency. Institutions therefore need sufficient information and coordination to understand why movement occurs and whether protection is required.

The literature suggests that vulnerability is cumulative. Employer-linked status can reduce bargaining power; recruitment debt can increase economic pressure; language barriers can restrict information; employer-provided housing can make departure harder; and fragmented institutions can delay identification. None of these factors alone proves exploitation, but together they may leave a formally employed worker with little practical freedom to challenge conditions or leave safely. This is where human security, precarity and labour governance converge.

The analysis also distinguishes administrative visibility from substantive protection. A fully documented worker may still be unable to claim rights, while a worker leaving abuse may become administratively irregular while exercising agency. A permit-centred model privileges continuity of documentation; a worker-centred model asks whether rights remain effective during employment and status changes. Lawful pathways for resolving problems can strengthen, rather than weaken, migration governance by reducing incentives for informality.

North Macedonia can address these issues before foreign labour recruitment

becomes more deeply entrenched. Coordination mechanisms, ethical recruitment standards and accessible information can be designed while the system is still developing. Prevention is likely to be more effective and less costly than responding after exploitative practices become established.

A balanced policy must also recognise employers' need for predictable and timely procedures. Excessive bureaucracy can encourage informal arrangements and discourage lawful recruitment. Worker protection and administrative efficiency should therefore develop together through clear recruitment and employer-change procedures, standardised information, transparent responsibilities and interoperable data.

The findings also have implications for anti-trafficking policy. Labour exploitation cannot be addressed only through criminal-law responses after severe abuse occurs. Labour institutions may identify earlier indicators through inspections, wage complaints, employment changes and recruitment patterns. Anti-trafficking expertise remains essential for serious cases, but prevention requires continuous links with ordinary labour-rights mechanisms.

9. CONCLUSION AND POLICY RECOMMENDATIONS

North Macedonia is no longer only a country of emigration and transit. It is gradually becoming a country that relies on foreign labour to respond to demographic decline and labour shortages. This transition creates economic opportunities, but it also produces new governance responsibilities. The central challenge is not simply how many workers can be recruited, but whether the institutional system can follow, support and protect them throughout their employment trajectory.

The paper has shown that secondary labour migration provides a useful lens for understanding what happens after admission. Changes of employer, movement between formal and informal work, loss or alteration of status and onward movement are not peripheral events. They are moments at which workers may exercise agency, encounter administrative barriers or become exposed to exploitation. These possibilities cannot be understood through permit data alone.

The institutional responses analysed in this paper indicate that relevant information is dispersed across the Employment Service Agency, the Ministry of Interior, the Public Revenue Office, labour inspection and other actors. The problem is not institutional inactivity, but partial visibility: each institution sees only one part of the worker's trajectory. Without proportionate and rights-based coordination, critical changes may remain invisible until they appear as complaints, incidents or border movements.

The evidence supports a cautious conclusion. It does not establish systematic exploitation of foreign workers in North Macedonia or show that secondary labour migration is inherently irregular. It does suggest that the current governance setting may leave vulnerability to labour-rights violations, exploitation, forced labour, human trafficking and associated risks under-detected. This is particularly relevant where workers depend on employers or intermediaries for information, residence, accommodation or access to institutions. Resilience, moreover, extends beyond institutional capacity. It includes workers' ability to understand and exercise their rights, change employment lawfully, maintain status during transitions, access social and legal support, report abuse safely and recover from harmful conditions. Strengthening that resilience requires institutions that are coordinated, accessible and trusted.

Policy development should therefore move from a permit-centred model of labour governance. First, institutions should establish a formal coordination mechanism linking employment, residence, tax, labour inspection and protection functions, with clearly defined data-protection safeguards. Second, procedures for changing employers should be transparent, timely and accessible so that workers are not forced to choose between tolerating abuse and losing status. Third, ethical recruitment standards should be strengthened, including transparency regarding recruitment fees, contracts, deductions, accommodation and the identity of all intermediaries involved. Fourth, workers should receive standardised multilingual information before or immediately after arrival, including information on rights, occupational safety, complaint procedures and available assistance. Fifth, complaint and advisory channels should be confidential, accessible and separated, as far as legally possible, from automatic punitive immigration consequences.

Labour inspection should be supported to undertake risk-based and proactive monitoring in sectors and employment arrangements where foreign workers are concentrated. Inspection should include attention to wages, working hours, occupational safety, accommodation linked to employment, access to personal documents and the consistency between promised and actual employment conditions. Aggregate data on complaints, employer changes, inspections and referrals should be analysed to identify patterns without profiling workers by nationality.

Further research is also needed. Future studies should include interviews with foreign workers, employers, recruitment agencies, labour inspectors, trade unions and civil-society organisations. Workers' own accounts are essential for understanding why they stay, change employer or move onward, and whether existing procedures are accessible in practice. Quantitative research could examine employment retention, sectoral concentration, contribution records and the relationship between changes in employment and residence status. Such research would allow the preliminary governance findings of this paper to be tested and refined.

Ultimately, a sustainable labour-migration system depends on more than admission numbers. It depends on whether workers can exercise rights, whether employers can recruit through predictable and lawful channels, and whether institutions can identify problems before vulnerability develops into severe exploitation. Protecting foreign workers is not an additional burden placed on labour-market policy. It is a condition for labour migration to remain lawful, credible and socially sustainable.

REFERENCES

- Anderson, B. (2010). Migration, immigration controls and the fashioning of precarious workers. *Work, Employment and Society*, 24(2), 300–317. <https://doi.org/10.1177/0950017010362141>
- Assembly of the Republic of North Macedonia. (2025). *Law on amendments and supplements to the Law on Foreigners. Official Gazette of the Republic of North Macedonia* (No. 193/2025).
- Clibborn, S., & Wright, C. F. (2023). Proposing mobility visas as an alternative to employer sponsorship: Addressing inequalities in the treatment of temporary migrant workers. *The Economic and Labour Relations Review*, 34(4), 772–786. <https://doi.org/10.1017/elr.2023.56>
- Employment Service Agency of the Republic of North Macedonia. (n.d.). *Вработување на странци [Employment of foreigners]*. <https://av.gov.mk/vrabotuvanje-na-stranci.nspix>
- Employment Service Agency of the Republic of North Macedonia. (2026). *Data on processed requests, opinions and work permits for foreign workers, 2020–2025* [Unpublished institutional data].
- European Commission. (2025). *North Macedonia 2025 report* (Commission Staff Working Document, SWD(2025) 753 final).
- Farbenblum, B. (2017). Governance of migrant worker recruitment: A rights-based framework for countries of origin. *Asian Journal of International Law*, 7(1), 152–184. <https://doi.org/10.1017/S2044251316000011>
- Golding, L., Berinstein, C., & Bernhard, J. K. (2009). Institutionalizing precarious migratory status in Canada. *Citizenship Studies*, 13(3), 239–265. <https://doi.org/10.1080/13621020902850643>
- International Labour Organization. (2019). *General principles and operational guidelines for fair recruitment and definition of recruitment fees and related costs*. International Labour Office.
- International Labour Organization. (2022). *Temporary labour migration: Unpacking complexities*. International Labour Office.
- International Organization for Migration. (2017). *Migrant vulnerability to human trafficking and exploitation: Evidence from the Central and Eastern Mediterranean migration routes*.
- International Organization for Migration. (2019). *IOM handbook on protection and assistance for migrants vulnerable to violence, exploitation and abuse*.
- International Organization for Migration. (2023). *IRIS handbook for governments on ethical recruitment and migrant worker protection*.
- Lewis, H., Dwyer, P., Hodgkinson, S., & Waite, L. (2015). Hyper-precarious lives: Migrants, work and forced labour in the Global North. *Progress in Human Geography*, 39(5), 580–600. <https://doi.org/10.1177/0309132514548303>
- Ministry of Interior of the Republic of North Macedonia. (2026). *Response to freedom of information request concerning records related to foreign workers and labour-migration data* [Unpublished institutional response].
- Ministry of Social Policy, Demography and Youth. (n.d.). *Приватни агенции за вработување [Private employment agencies]*. https://cms.mtsp.gov.mk/register-ins_article-privatni-agencii-za-vrabotuvanja-2.nspix
- Public Revenue Office of the Republic of North Macedonia. (2026). *Response to freedom of information request concerning foreign nationals registered for tax and*

- contribution purposes on the basis of employment or work engagement* [Unpublished institutional response].
- Red Cross of the Republic of North Macedonia. (2026). *National programme for assistance and support to refugees and migrants: Monthly report, April 2026* [Internal programme report].
- Samardžiski, A. (2026, January 12). Над шест илјади странски работници лани добиле дозволи во Северна Македонија [More than six thousand foreign workers received permits in North Macedonia last year]. *Radio Free Europe/Radio Liberty*.
- Sargeant, M., & Tucker, E. (2009). Layers of vulnerability in occupational safety and health for migrant workers: Case studies from Canada and the UK. *Policy and Practice in Health and Safety*, 7(2), 51–73. <https://doi.org/10.1080/14774003.2009.11667734>
- Shamir, H. (2017). The paradox of “legality”: Temporary migrant worker programs and vulnerability to trafficking. In P. Kotiswaran (Ed.), *Revisiting the law and governance of trafficking, forced labor and modern slavery* (pp. 471–502). Cambridge University Press. <https://doi.org/10.1017/9781316675809.017>
- State Statistical Office of the Republic of North Macedonia. (2024). *North Macedonia in figures 2024*.
- State Statistical Office of the Republic of North Macedonia. (2025). *Active population in the Republic of North Macedonia, 2024*.
- United Nations Development Programme. (1994). *Human development report 1994: New dimensions of human security*. Oxford University Press.
- World Bank. (2024). *World development indicators: North Macedonia*.

MEASURING CITIZENS' PERCEPTIONS OF SECURITY: CONSTRUCTING THE 2025 SECURITY INDEX IN NORTH MACEDONIA

Marjan Nikolovski

Faculty of Security – Skopje, UKLO

marjan.nikolovski@uklo.edu.mk

Cane Mojanoski

Faculty of Security – Skopje, UKLO

mojanc@gmail.com

Aleksandar Ivanov

Faculty of Security – Skopje, UKLO

aleksandar.ivanov@uklo.edu.mk

ORCID 0009-0006-3996-7034

Kire Babanoski

Faculty of Security – Skopje, UKLO

kire.babanoski@uklo.edu.mk

ORCID 0000-0002-2842-3633

Abstract

This paper presents the development and initial findings of the 2025 Security Index, designed to empirically measure citizens' subjective perceptions of security in North Macedonia. The study conceptualises security as a multidimensional “boundary concept” that integrates traditional threats with existential and emerging hybrid risks. The research addresses three central questions: the dominant sources of fear and differences between private and public safety perceptions; the level of trust in state and security institutions and the impact of perceived corruption; and the influence of economic vulnerabilities, environmental risks, and digital threats on the overall sense of (in)security.

Data were collected through a nationwide face-to-face survey conducted between December 2025 and January 2026 using a distributed fieldwork model. A total of 72 trained interviewers (students from the Faculty of Security in Skopje) administered 1,139 questionnaires using a combined PAPI and CAPI approach, ensuring real-time quality control and strong representativeness (margin of error below 3%). The questionnaire was structured into six analytical domains: personal and societal security; economic security; health security; national security, political stability and rule of law; environmental security; and digital/cognitive security. The findings indicate that citizens' perceptions of security are shaped by a cumulative interaction of factors rather than isolated threats.

The proposed Security Index offers a comprehensive analytical tool for policymakers, enabling evidence-based prioritisation of security policies and supporting a holistic approach to societal resilience.

Key words: Security Index; Citizens' Perceptions; Institutional Trust; Security Threats

1. INTRODUCTION

Traditionally, security has been understood through a realist paradigm that primarily focuses on state sovereignty, territorial integrity, and military capabilities (Buzan, 1991; Krause & Williams, 1996; Buzan, Wæver, & de Wilde, 1998). Within this state-centred perspective, security is commonly assessed through indicators such as defence expenditure, the size of the armed forces, and the capacity to respond to external threats. However, since the 1990s, this conventional understanding has been increasingly challenged by the concept of human security, most notably introduced in the United Nations Development Programme (UNDP) *Human Development Report* (1994). The human security approach argues that the ultimate referent of security should not be the state alone, but the individual and the conditions necessary for living a safe, dignified, and fulfilling life (Alkire, 2007; Tadjbakhsh & Chenoy, 2007). Consequently, a country may be free from armed conflict while its citizens continue to experience significant insecurity arising from crime, institutional dysfunction, economic hardship, poor healthcare, environmental degradation, or emerging digital and information threats (Paris, 2001). Contemporary societies are increasingly exposed to a wide range of security risks, threats, and vulnerabilities that continuously reshape the overall security environment. This means that traditional security challenges are now complemented by economic instability, environmental hazards, cyber threats, disinformation, and other hybrid risks, making security a multidimensional and dynamic concept. In this context, the primary responsibility of the state and its security institutions is to protect the fundamental constitutional values and ensure the safety and well-being of citizens through preventive action, risk identification, assessment, and the timely implementation of appropriate security measures. Building on this broader conceptual framework, the present study adopts a multidimensional approach to security, recognising that contemporary security challenges extend well beyond traditional military concerns and encompass the everyday experiences and perceptions of citizens.

An essential prerequisite for effective prevention is the development of comprehensive and evidence-based security assessments. Such assessments should not rely solely on objective indicators of crime, conflict, or institutional performance but should also incorporate citizens' perceptions of security, as subjective feelings of safety significantly influence public trust, social cohesion, and the legitimacy of state institutions. Consequently, the development of measurable indicators that capture these perceptions represents an important step towards strengthening strategic security planning and evidence-based policymaking.

Although security is one of the most frequently discussed concepts in the fields of security studies and political science, there is still no universally accepted definition. As noted by Savić and Gaćinović (2007), security is a complex and multidimensional concept whose meaning varies depending on the theoretical perspective and the object of protection. It is difficult to precisely determine what is being protected, against whom protection is directed, and where the distinction lies between the subject and the object of security. This conceptual ambiguity has contributed to the continuous evolution of security studies and the expansion of security beyond its traditional military dimension. From a linguistic and philosophical perspective, security is an abstract concept rather than a tangible phenomenon. As Nikolovski (2026) argues, security cannot be physically observed; instead, it is experienced and perceived through individuals' sense of protection, stability, and freedom from threats. It encompasses the preservation of both material and intangible

values, ultimately serving the protection of human life, dignity, and well-being.

Reflecting this broader understanding, contemporary security indices are no longer constructed exclusively on objective crime statistics or military indicators. Instead, they increasingly integrate multiple dimensions of security, including criminological, political, institutional, economic, social, environmental, cyber, and informational factors. Such multidimensional approaches acknowledge that citizens' perceptions of security emerge from the cumulative interaction of various risks rather than from a single source of threat. Measuring these perceptions therefore provides valuable insight into societal resilience and enables policymakers to identify priority areas requiring institutional intervention.

Several internationally recognised composite indices have significantly contributed to measuring different dimensions of security, peace, governance, and human development. Among the most influential are the Global Peace Index (IEP, 2024a) and the Positive Peace Index (IEP, 2024b) developed by the Institute for Economics and Peace (IEP), the Human Development Index (UNDP, 2026) published by the United Nations Development Programme (UNDP), the Rule of Law Index (World Justice Project, 2026) developed by the World Justice Project (WJP), the Corruption Perceptions Index (Transparency International, 2024) published by Transparency International, and the Democracy Index (Economist Intelligence Unit, 2026) produced by the Economist Intelligence Unit (EIU). Collectively, these indices provide valuable frameworks for assessing national performance across dimensions such as peace, institutional quality, governance, corruption, democracy, and human development.

At the national level, several surveys have examined different aspects of citizens' perceptions of security, public trust, victimisation, and regional security trends. Notable examples include studies conducted by the OSCE on citizens' sense of safety (OSCE, 2019), hate crime victimisation (Iganski, 2019), and youth perceptions of security (INDAGO, 2024), the Regional Cooperation Council (RCC) *Securimeter* survey on security perceptions in the Western Balkans (RCC, 2025), and research by the Macedonian Centre for International Cooperation (MCIC) focusing on citizens' perceptions of security and institutional trust (Jakovleska, 2025). While these studies provide valuable insights into specific dimensions of security, they generally address individual issues or particular population groups rather than offering an integrated, multidimensional measurement of citizens' overall perceptions of security.

The Faculty of Security – Skopje has also made a significant contribution to security research through a series of studies addressing security risks (Mojanoski, Dimovski, Gjurovski & Ilijevski, 2015; Gjurovski, Nikolovski & Gerasimoski, 2020), the functioning of the national security system (Mojanoski, Nikolovski, Batic, Dujovski, Gjurovski, Ristevska-Dimitrovska, Nuhiu, Ramadani, 2021), public perceptions of policing (Mojanoski, Dujovski & Gjurovski, 2018), corruption (Mojanoski, Malish-Sazdovska, Nikolovski & Krstevska, 2018), and the development of earlier Security Indices (Mojanoski & Gjurovski, 2018a; Mojanoski & Gjurovski, 2018b). While the earlier Security Indices primarily focused on traditional dimensions of public safety and security, the Security Index 2025 substantially broadens both the conceptual and methodological scope. It adopts a human security perspective by integrating six interrelated dimensions—personal, economic, health, institutional, environmental, and digital-cognitive security—into a single analytical framework. Furthermore, it incorporates contemporary security challenges, including cyber threats, disinformation, environmental risks, and healthcare resilience, which were largely absent from previous editions. The 2025 Index also employs a redesigned survey instrument and a nationally representative field survey, providing a

more comprehensive and policy-oriented assessment of citizens' perceptions of security in North Macedonia.

Against this background, this paper presents the conceptual framework and the initial findings of the Security Index 2025 for North Macedonia with analyses of its six constituent dimensions. The study introduces a comprehensive instrument designed to measure citizens' perceptions of security across multiple dimensions, offering an empirical basis for assessing the overall sense of security within society. The following sections present the initial empirical findings obtained through the nationwide survey conducted by the research team of the Faculty of Security – Skopje, highlighting key patterns in citizens' perceptions and discussing their implications for future security policy and strategic planning.

Although the Security Index 2025 is conceptually designed as a multidimensional composite framework, the present article focuses on analysing the six constituent dimensions individually. Each dimension consists of several survey indicators measured on a standardised ten-point Likert scale and aggregated at the dimension level using arithmetic mean values. These six dimensions collectively constitute the conceptual architecture of the Security Index. The development and validation of a single composite Security Index score, including weighting procedures and psychometric validation, represent the next phase of the broader research project and will be reported in future publications.

2. CONCEPTUAL FRAMEWORK AND DESIGN OF THE SECURITY INDEX 2025

The conceptual design of the Security Index 2025 is based on the premise that security should be measured not only through objective indicators, such as recorded crimes or official statistics, but also through citizens' perceptions of safety, trust, and vulnerability. In this regard, the Security Index functions as an early warning system, designed to identify emerging security risks before they develop into more serious social or institutional crises. Unlike traditional security reports, which primarily document incidents after they occur, the Index seeks to detect early signs of societal stress, including declining institutional trust, growing economic insecurity, increasing perceptions of environmental risks, and concerns related to digital and cognitive threats.

The Security Index 2025 is a multidimensional composite measure of human security in the Republic of North Macedonia. It has been designed to quantify both the everyday sense of safety ("freedom from fear") and existential stability ("freedom from want"), while contextualising these dimensions within the structural conditions that support or undermine peace, including institutional trust, corruption, environmental vulnerability, and digital insecurity. The Index functions as a "seismograph of societal vulnerability." Rather than merely recording isolated security incidents, it identifies the accumulation of social pressures and emerging vulnerabilities before they escalate into broader security challenges. By integrating security perceptions with economic, institutional, environmental, and digital dimensions, the Index provides an empirical foundation for preventive governance and evidence-based public policymaking.

The central premise of the Index is that security should not be understood as an isolated event but rather as a systemic property, whereby changes within one domain—for example, public trust in the judiciary—can generate cascading effects across the wider social and institutional environment. Accordingly, the instrument developed in this study

is intended to serve not only a descriptive but also a preventive function. Its purpose is to enable public institutions to recognise early warning signals of insecurity and respond proactively before these vulnerabilities evolve into more significant societal risks.

The guiding thesis of this research is that security in 2025 cannot be reduced to the mere absence of crime, violence, or armed conflict. Instead, it is shaped by the interaction of everyday institutional, existential, environmental, and digital-cognitive insecurities. Consequently, the Security Index conceptualises security as a condition of trust, resilience, and the capacity of institutions and society to prevent, absorb, and overcome risks. Operationally, the research is structured around six groups of questions corresponding to the six dimensions of the Security Index. These dimensions examine: (1) citizens' perceptions of personal and spatial safety; (2) the extent to which economic conditions and social support shape existential security; (3) perceptions regarding the quality, accessibility, and financial sustainability of healthcare; (4) trust in public institutions and the extent to which corruption and political instability are perceived as security risks; (5) the environmental threats that most significantly affect everyday life; and (6) the degree to which cyber threats, misuse of personal data, and disinformation undermine cognitive and national resilience.

These six dimensions provide a comprehensive picture of security by capturing both objective risks and citizens' subjective perceptions. Rather than analysing each dimension separately, the study recognises that they are closely interconnected. Weaknesses in one area often generate pressures in others. For example, declining trust in institutions may increase perceptions of personal insecurity, economic hardship can reduce confidence in government, environmental degradation directly affects public health, while the spread of disinformation can undermine political stability and social cohesion. The six dimensions were selected following an extensive review of the literature on human security, positive peace, public security perception, and multidimensional security measurement (UNDP, 1994; IEP, 2024a; Gasper & Gómez, 2015). Rather than reproducing existing international frameworks, the Index adapts these concepts to the contemporary security context of North Macedonia by integrating traditional security concerns with institutional, socio-economic, environmental, health, and digital dimensions that have emerged as significant determinants of citizens' perceptions of security. This multidimensional framework therefore enables a more comprehensive assessment of security by identifying not only individual areas of vulnerability but also the relationships between them. It provides a broader understanding of how different security challenges interact and collectively shape citizens' overall sense of safety and societal resilience. Each dimension consists of several survey indicators measured on a standardised ten-point Likert scale. Individual variables were first analysed separately and subsequently aggregated into domain-level scores using arithmetic means. The six domain scores constitute the conceptual structure of the Security Index and provide a multidimensional assessment of citizens' perceptions of security. In this paper, emphasis is placed on the analysis of the individual dimensions rather than on presenting a single composite score, which will be developed and validated in subsequent stages of the project.

The overall objective of the research is to measure and interpret the state of security as a combination of "freedom from fear" and "freedom from want," while simultaneously conceptualising security as an architecture of trust and resilience. This objective is pursued through empirical data collected in a nationwide field survey and is grounded in the theoretical concepts of human security and positive peace. The central argument underpinning this research is that security is not merely a state or condition, but a systemic

property. From a systems perspective, such as that developed by the Institute for Economics and Peace (IEP) through its Halo framework, society is understood as a network of interdependent subsystems in which changes within one domain—for example, public trust in the judiciary—may generate cascading effects across other domains, including public order, perceptions of violent groups, and the legitimacy of the state (IEP, 2024a, IEP, 2024b). Accordingly, the overall objective of the Security Index can be defined as follows: *To measure and interpret the state of security as a combination of "freedom from fear" and "freedom from want," while conceptualising security as an architecture of trust and resilience, based on empirical evidence obtained through nationwide field research and informed by the theoretical frameworks of human security and positive peace.*

This objective is fully aligned with the UNDP Human Security Framework, which shifts the focus of security from territory to people and from military protection to human development (UNDP, 1994).

3. RESEARCH OPERATIONALISATION AND METHODOLOGY

The operationalisation of the research is based on the assumption that subjective perceptions constitute an important source of security information. Although public opinion surveys are sometimes criticised for measuring perceptions rather than objective facts, security research recognises that perceptions have tangible social consequences. This assumption is consistent with the well-established Thomas Theorem, which states that *"if people define situations as real, they are real in their consequences"* (Bornmann & Marx, 2020; Nichols, 2016). Citizens who perceive their environment as insecure are more likely to modify their behaviour by avoiding public spaces, limiting social interactions, increasing self-protection, or losing confidence in public institutions. Consequently, perceptions of insecurity directly influence social cohesion, institutional legitimacy, and community resilience. For this reason, the Security Index measures latent rather than solely observable security risks. Indicators such as trust in the police, judiciary, prosecution service, and other public institutions are not interpreted simply as measures of public satisfaction but as indicators of institutional resilience and the capacity of the state to manage security challenges effectively. Likewise, perceptions of economic uncertainty, environmental degradation, inadequate healthcare, cyber risks, and disinformation are treated as early manifestations of broader structural vulnerabilities that may, if left unaddressed, contribute to social instability and undermine national resilience. By integrating citizens' perceptions with a multidimensional understanding of security, the operational framework provides a comprehensive assessment of both existing and emerging security challenges. This approach enables the identification of vulnerabilities at an early stage and supports evidence-based policymaking aimed at strengthening societal resilience and preventing future security crises.

The Security Index 2025 was developed through a nationwide field survey designed to measure citizens' perceptions of security across multiple dimensions. To maximise data quality and representativeness, the study employed a distributed face-to-face data collection model rather than online survey panels, thereby reducing the risk of self-selection bias. Data were collected by a network of 72 trained interviewers, all students of the Faculty of Security – Skopje.

The survey was conducted between December 2025 and January 2026, using a combination of Paper-and-Pencil Interviewing (PAPI) and Computer-Assisted Personal Interviewing (CAPI). Digital data entry enabled real-time quality control and continuous

monitoring of the fieldwork process. Conducting the survey during the winter period, including the holiday season and peak heating months, provided an important contextual background for interpreting citizens' perceptions of economic and environmental security.

A stratified random sampling approach was applied. Each interviewer was assigned a quota of approximately 20 respondents, resulting in a theoretical sample of 1,440 participants. Following rigorous post-fieldwork validation and the exclusion of incomplete questionnaires, the final validated sample consisted of 1,139 respondents, providing a nationally representative dataset with a margin of error below 3%. To minimise interviewer bias, respondent selection followed a Random Walk protocol, whereby interviewers systematically selected every fifth house or every twentieth apartment, depending on the type of residential area.

The sample of the study demonstrates a high level of demographic representativeness and balance across key socio-demographic variables. In terms of gender, the distribution is almost perfectly balanced, with 49.7% male respondents (N=566) and 49.9% female respondents (N=568), while 0.4% (N=5) preferred not to indicate their gender. This minimal difference (0.6%) is statistically negligible and closely reflects the actual population structure of the country.

Regarding age distribution, the sample includes 34.2% young respondents (18–30 years), 37.4% middle-aged respondents (31–50 years), and 28.1% older respondents (51+ years), ensuring adequate representation across all major life-course groups. In terms of ethnic composition, the majority of respondents are Macedonians (83.2%), followed by Albanians (9.5%), with smaller shares of Turks (3.2%), Bosniaks (1.3%), Vlachs (0.5%), and Roma (0.4%). This structure broadly mirrors the national demographic profile, allowing for meaningful comparative analysis across ethnic communities.

The regional distribution of respondents further contributes to the robustness of the sample. The largest proportion comes from the Skopje region (40.2%), followed by Pelagonia (15.1%), the Northeastern (11.2%), Eastern (10.7%), and Southeastern regions (10.0%), while the smallest share is from the Vardar region (2.7%). Although some regional variation is present, this distribution ensures sufficient territorial diversity, enabling a comprehensive analysis of security perceptions across different geographic and socio-economic contexts in North Macedonia.

The questionnaire was operationalised through six analytical dimensions corresponding to the conceptual framework of the Security Index: personal and spatial security; economic and social security; health security; institutional trust and political stability; environmental security; and digital and cognitive security. This multidimensional structure reflects the broader human security perspective and enables a comprehensive assessment of citizens' perceptions of security. To ensure a high level of statistical precision and variability in responses, the study employs a standardised 10-point Likert-type scale for the majority of questions instead of binary (yes/no) items, enabling the calculation of precise mean scores and standard deviations.

The research adhered to established academic and ethical standards. Participation was voluntary and based on informed consent, while respondents' anonymity was fully protected. Data integrity was ensured through interviewer logs and GPS time-stamped digital records, allowing the research team to verify completed interviews and eliminate invalid responses. These procedures contributed to the reliability, transparency, and overall methodological robustness of the study.

4. KEY FINDINGS OF THE SECURITY INDEX 2025

This section presents the key findings of the Security Index 2025, drawing on selected results from the broader empirical research. Rather than providing an exhaustive overview of all variables, it focuses on the most significant patterns and trends identified across the six dimensions of the Index. These findings are both presented and interpreted in order to highlight the main security perceptions, vulnerabilities, and institutional challenges as experienced by citizens in North Macedonia.

The analysis that follows synthesises quantitative results with their broader analytical implications, offering an interpretation of how different security domains interact and shape the overall sense of security. In doing so, the section identifies the most salient risks and concerns emerging from the data, including issues related to personal safety, economic and health insecurity, institutional trust, environmental pressures, and digital and cognitive threats.

Overall, this section serves as a condensed presentation of the empirical core of the study, providing an evidence-based foundation for the subsequent discussion and conclusions.

4.1. Dimension 1: Personal and Societal Security

The first pillar of the Security Index relates to citizens' subjective perception of physical security. Respondents were asked to rate their sense of safety on a scale from 1 (very unsafe) to 10 (very safe). The analysis of mean values reveals a clear dichotomy between private and public space, a phenomenon which in the academic literature can be conceptualised as a "Fortress Mentality."

Table 1. Comparative Spatial Matrix of Security

Variable	Mean Value	Interpretation
Security at home	8.51	High level of perceived safety; the home is seen as the primary secure zone.
Personal security (general)	6.75	Moderate/labile; significantly lower than home security, reflecting a baseline societal anxiety.
Safety in public spaces	5.88	Borderline level; a sharp decline compared to the home, indicating polarised perceptions of public safety.

Source: Authors' calculations based on data from the Security Index 2025 research project conducted by the Faculty of Security – Skopje (N = 1,139).

The difference between feeling safe at home and in public space highlights a significant security gap. The gap of 2.63 index points between home security (8.51) and public space security (5.88) is a notable indicator of perceived vulnerability in the public sphere. According to comparative benchmarks in stable societies (e.g. Global Peace Index methodology), such a gap rarely exceeds one point. This suggests that public space in North Macedonia is not fully perceived as a shared civic good (*res publica*), but rather as a risk environment. Citizens tend to retreat into the private sphere, where security is self-managed, while public order is perceived as insufficiently guaranteed by institutions such as the police and municipal services.

A closer examination of mean scores clearly outlines a spatial security matrix among citizens. The highest level of perceived safety is recorded in the private home (M =

8.51), confirming the home as a protective “fortress.” Once individuals move beyond this micro-environment, perceived safety declines sharply, reaching $M = 5.88$ in public spaces such as streets, parks, and shopping areas. Overall perceived security in everyday life is positioned between these two extremes ($M = 6.75$), reflecting a composite of high private safety and lower public safety.

This nearly three-point gap between private and public space is a key indicator of institutional limitations in guaranteeing everyday security in public environments. It suggests that citizens feel secure primarily in spaces under their direct control, while public space—formally the responsibility of state institutions—is increasingly perceived as an area of risk and unpredictability.

4.2. Dimension 2: Economic Security

Economic security should be understood not only as a reflection of current material conditions, but as households’ capacity to withstand economic shocks and sustain stability over time. The findings indicate that while citizens generally perceive their current economic situation as moderately stable, there is a pronounced concern regarding the affordability of basic living costs and low confidence in the state’s ability to provide adequate social protection in cases of unemployment or illness. This suggests that economic vulnerability should be interpreted as a security issue rather than solely a socio-economic one. Whereas Dimension 1 addresses the question “Will I survive today?”, Dimension 2 addresses “How will I live tomorrow?”. Within the framework of the Security Index 2025, economic insecurity is therefore conceptualised as a driver of instability. Poverty is not only a social condition but also a fertile ground for radicalisation, corruption, and crime.

The data reveal a notable phenomenon referred to as the “paradox of the working poor.” While job security is relatively high ($M = 6.15$), reflecting a stable labour market or strong employment protection, respondents report significantly lower capacity to cover monthly essential expenses ($M = 4.82$). This gap of 1.33 points indicates a structural imbalance in which citizens feel secure as workers but economically insecure as consumers. Furthermore, expectations regarding future economic stability are pessimistic ($M = 3.95$), indicating widespread anticipation of potential economic deterioration. From a security perspective, such a pattern reflects a weakening of the social contract. As Menon (2007) argues, when “freedom from want” is undermined despite employment, the relationship between citizens and the state becomes fragile, often resulting in increased emigration as a form of systemic pressure release.

Table 2. Indicators of Economic Vulnerability

Variable	Mean Value	Interpretation
Job security	6.15	Relative stability. Citizens generally do not fear job loss, indicating either a stable labour market or strong public-sector employment protection.
Ability to cover basic monthly expenses	4.82	High anxiety. This is a critical finding: although respondents are employed (6.15), they report difficulties in covering monthly living costs (4.82).
Fear of future economic crisis	3.95	Pessimism. Expectations regarding future economic conditions are negative, with respondents effectively “preparing for the worst.”

Source: Authors' calculations based on data from the Security Index 2025 research project conducted by the Faculty of Security – Skopje (N = 1,139).

Age-disaggregated analysis shows that younger respondents (18–30) express the lowest confidence in their economic future, with strong intentions to seek opportunities abroad, while older groups (51+) report higher anxiety related to daily living costs such as food and healthcare, but lower fear of job loss due to employment protection and accumulated work tenure.

4.3. Dimension 3: Health Security

The health system represents one of the most sensitive indicators of human security, as it directly determines outcomes related to life and death. Field data and open-ended responses indicate that the public healthcare system is perceived as one of the most critical weaknesses in the country. Health security illustrates a clear overlap between human security and social policy. When citizens fear that they will be unable to afford treatment or medication in the event of serious illness, health risks become economic risks, and economic risks in turn transform into security risks. In this sense, healthcare in the Security Index is not treated as a sectoral policy alone, but as a core component of national resilience. Health security emerges as one of the most vulnerable dimensions of the Index. Low assessments of the quality of the public healthcare system, the availability of medical personnel, and access to services indicate that citizens do not perceive the system as sufficiently accessible or functional. The most critical intersection between health and economic security is the widespread concern that necessary treatment or medicines may not be financially affordable in cases of serious illness.

Respondents express significant dissatisfaction with the infrastructure and management of public health institutions. This dissatisfaction is particularly pronounced in two key variables: the availability of qualified medical staff and the accessibility of healthcare services, including waiting times and appointment scheduling. The emigration of medical professionals and long waiting lists for basic examinations and interventions contribute to the perception that healthcare, although formally guaranteed as a constitutional right, is in practice a difficult-to-access privilege.

Table 3: Quality of the public healthcare system

Rating	#	%	Cumulative %
Cannot assess	37	3.2	3.2
1	254	22.3	25.5
2	168	14.7	40.3
3	180	15.8	56.1
4	116	10.2	66.3
5	161	14.1	80.4
6	99	8.7	89.1
7	78	6.8	96.0
8	31	2.7	98.7
9	9	0.8	99.5
Highest quality	6	0.5	100.0
Total	1139	100.0	

Source: Authors' calculations based on data from the Security Index 2025 research project conducted by the Faculty of Security – Skopje (N = 1,139).

When asked to assess the overall quality of the public healthcare system, responses show a clear dominance of negative evaluations, indicating widespread dissatisfaction among citizens. Very low ratings (1–2) account for 38.3% of responses, while low ratings (3–4) represent 26.9%. Moderate assessments (5–6) make up 23.6%, whereas high ratings (7–8) account for only 9.9%, and very high ratings (9–10) are limited to just 1.4%. This distribution indicates that the public healthcare system is predominantly perceived negatively, with a strong level of dissatisfaction among the majority of respondents, while only a very small proportion consider the system to be of high or very high quality.

4.4. Dimension 4: National Security, Political Stability and Rule of Law

According to the theoretical framework of the **Institute for Economics and Peace (IEP)**, one of the eight pillars of Positive Peace is the *well-functioning of government*. In this sense, security is not only defined by the absence of street-level crime (Dimension 1), but also by the presence of trust in the institutions responsible for enforcing the rule of law.

This dimension represents the core of institutional security. The findings reveal a profound institutional asymmetry: while citizens express relatively higher trust in institutions associated with “hard power” such as the Army and the Police, trust in institutions of justice—particularly the judiciary and prosecution—is in sharp decline. Perceptions of corruption, especially when linked to security institutions and public authorities, emerge as one of the strongest systemic risk factors in the study. Consequently, corruption must be understood not merely as an administrative or moral issue, but as a security concern.

The data clearly demonstrate that institutions responsible for ensuring the rule of law are among the lowest rated in the entire system, directly affecting perceptions of political security, one of the key dimensions of human security.

Table 4. Institutional Trust Ratings

Institution	Mean Value	Trust Category
Army	6.12	High (Relatively stable)
Religious communities (MPC/IRC)	5.85	Stable
President of the state	4.90	Polarised
Government	4.15	Low
Parliament	3.88	Very low
Public Prosecutor's Office	3.45	Critical
Judiciary	3.36	Systemic collapse

Source: Authors' calculations based on data from the Security Index 2025 research project conducted by the Faculty of Security – Skopje (N = 1,139).

The Army is the only institution scoring above 6, which is consistent with findings in transitional and post-conflict societies, where citizens tend to trust institutions that are less involved in day-to-day political disputes. In this context, the Army is perceived as a symbol of statehood rather than political governance. In contrast, the judiciary's extremely low score (M = 3.36) reflects a perception of functional breakdown. Such a rating suggests that citizens do not believe in the effective legal resolution of disputes. This institutional distrust also helps explain higher perceptions of insecurity related to violent or criminal groups: when the judiciary fails to sanction wrongdoing effectively, informal or criminal actors are perceived to fill the enforcement vacuum.

4.5. Dimension 5: Environmental Security

Environmental security represents one of the most strongly perceived areas of risk in this study. Air pollution, inadequate waste management, and exposure to natural hazards are not viewed solely as environmental issues, but as direct threats to public health, quality of life, and trust in both local and national institutions. In this sense, environmental protection must be understood as an integral component of security policy rather than a purely sectoral concern. Traditionally, environmental issues have been marginalised within security studies. However, in line with the framework proposed by Gasper & Gómez (2015), environmental security is conceptualised here as a core pillar of human security. The fieldwork, conducted during the winter period (December–January), coincided with peak exposure to air pollution, which is reflected in the intensity of the findings.

Table 5. Perception of Environmental Threats

Environmental Risk (Variable)	Mean Value	Level of Threat
Air pollution	8.74	CRITICAL / Extreme threat
Water/soil pollution	7.92	Very high threat
Illegal landfills	8.15	Very high threat
Floods	5.40	Moderate threat
Earthquakes	6.12	High (latent) threat

Source: Authors' calculations based on data from the Security Index 2025 research project conducted by the Faculty of Security – Skopje (N = 1,139).

The results can be interpreted along several key lines. First, air pollution emerges as the most severe perceived risk, with a score of 8.74—the highest negative value across the entire study. In comparative terms, citizens perceive environmental pollution as a

greater threat than crime-related insecurity, highlighting a fundamental shift in the nature of perceived risks. This provides strong empirical support for the human security perspective, where threats do not originate from violence alone but also from environmental conditions such as polluted air. Second, a clear urban–rural divide is observed. In urban areas, particularly Skopje and Bitola, perceived risk from air pollution exceeds 9.0, while in rural areas concerns are more strongly associated with illegal waste disposal (8.15) and inadequate waste management infrastructure. Overall, the high levels of perceived environmental risk indicate a significant governance deficit in this domain. The inability of the state to guarantee even basic environmental conditions—such as clean air—contributes to the erosion of institutional legitimacy and is closely linked to low levels of trust in government ($M = 4.15$), as identified in the previous dimension.

From a theoretical perspective, the findings also highlight the interaction between poverty and pollution. In the literature, this is conceptualised as environmental injustice. Economically vulnerable households are often forced to rely on cheaper but more polluting energy sources, thereby contributing to the very environmental degradation from which they suffer most severely through increased health costs and reduced quality of life. The data suggest the presence of this self-reinforcing cycle, supported by elevated levels of economic anxiety ($M = 4.82$) and structural constraints in household energy choices. This confirms the central argument that environmental security in North Macedonia is not a matter of lifestyle preference, but a condition of survival.

4.6. Dimension 6: Digital and Cognitive Security

Digital and cognitive security extends the traditional security framework into cyberspace and the broader information environment. Citizens report a moderate sense of safety when using digital services, but a low level of trust that the state provides adequate protection against cyber threats. In addition, cyber hygiene remains insufficient, as a significant share of respondents either do not use basic protection measures or are unaware of them. Disinformation and fake news are widely perceived as serious threats to national security, positioning cognitive security as an essential component of the hybrid security agenda.

The survey data indicate a relatively balanced but cautious perception of safety in the digital sphere. When asked “*How safe do you feel when using digital services (e-banking, online bill payments, e-government services)?*”, responses are distributed across the full scale, with a slight concentration in the mid-to-high categories. Very low ratings (1–2) account for 11.7%, low ratings (3–4) for 10.5%, moderate ratings (5–6) for 26.6%, high ratings (7–8) for 26.2%, while very high ratings (9–10) represent 18.5%. Overall, this indicates a generally moderate to relatively high level of perceived digital security, although a significant proportion of respondents still express insecurity or limited trust in the digital environment.

Table 6. How safe do you feel when using digital services (e-banking, online bill payments, e-government services)?

Rating	#	%	Cumulative %
Cannot assess	145	12.7	12.7
1	70	6.1	18.9
2	46	4.0	22.9
3	52	4.6	27.5
4	54	4.7	32.2
5	143	12.6	44.8
6	120	10.5	55.3
7	142	12.5	67.8
8	156	13.7	81.5
9	92	8.1	89.6
Full trust	119	10.4	100.0
Total	1139	100.0	

Source: Authors' calculations based on data from the Security Index 2025 research project conducted by the Faculty of Security – Skopje (N = 1,139).

Cognitive security emerges as one of the most significant contemporary dimensions of security in North Macedonia, particularly in the context of disinformation and fake news. The findings show a very strong public perception of these phenomena as a national security threat. Specifically, almost two-thirds of respondents rate disinformation as a high or very high risk, indicating that it is widely understood not merely as a media or communication issue, but as a direct component of hybrid threats affecting institutional trust, political stability, and decision-making processes.

Overall, the results highlight a heightened awareness among citizens of the broader geopolitical and informational environment in which the country operates. Perceptions of vulnerability to foreign influence, information manipulation, and coordinated disinformation campaigns are moderately to highly expressed, positioning North Macedonia as a state exposed to hybrid pressures in a complex regional and global context. In this sense, cognitive security is increasingly perceived as an integral part of national security and resilience.

5. QUALITATIVE FINDINGS: “THE VOICE OF THE CITIZENS”

The qualitative analysis of the “Security Index 2025” provides a deeper understanding of the perceptions behind the quantitative patterns, revealing not only what citizens think about security, but also why they hold these views and what solutions they envision. The final open-ended question invited respondents to propose measures for improving security in their communities and in North Macedonia more broadly. The responses represent a rich and unfiltered reflection of lived social realities, offering a direct insight into public expectations from the state and its institutions.

Thematic coding of the responses shows that citizens’ suggestions are highly pragmatic and centred on restoring what they perceive as a weakened social contract. Rather than proposing abstract or unrealistic reforms, respondents consistently emphasise basic institutional functionality, accountability, and fairness. Three dominant narratives emerge across the dataset. First, there is a strong demand for equal and non-selective justice, with citizens repeatedly stressing that “the law must apply equally to everyone”.

This reflects deep frustration with perceived impunity and selective justice, particularly regarding politically or economically powerful individuals. Second, there is a clear call for stricter sanctions and zero-tolerance policies, especially in relation to violence, traffic violations, and environmental offences. This does not necessarily indicate a punitive mindset, but rather a perception that existing enforcement mechanisms are weak, inconsistent, and vulnerable to corruption. Third, respondents strongly emphasise depoliticisation and professionalisation of institutions, frequently highlighting the need for merit-based employment and protection of young professionals from emigration driven by nepotism and political influence.

At the operational level, citizens provide very concrete and actionable proposals for improving everyday safety. The most frequently mentioned measure is the reintroduction of visible community policing (“neighbourhood police officers”), reflecting a strong belief that continuous police presence in public spaces would significantly reduce street crime, drug distribution near schools, and juvenile violence. Another recurring proposal concerns urban safety infrastructure, including CCTV systems in public areas and improved street lighting, particularly in peripheral and rural zones identified as “dark spots” associated with insecurity. Road safety also emerges as a key concern, with suggestions for stricter traffic controls, speed-reducing infrastructure near schools, and enhanced enforcement against drunk and reckless driving. Finally, respondents link security to broader quality-of-life issues, calling for stronger interventions in public health and environmental protection, including the removal of stray dogs and illegal landfills, as well as stricter controls on industrial pollution.

Overall, the qualitative findings converge on a clear and consistent message: citizens perceive security not only as protection from crime, but as the overall functionality of the state. Their responses reflect a demand for institutions that are visible, fair, and effective, and a governance system that is insulated from corruption and political interference. In this sense, the “voice of the citizens” confirms the quantitative findings and reinforces the conclusion that restoring trust in basic institutional performance is central to improving perceived security in North Macedonia.

6. CONCLUSION

The aim of this study was to present the idea and the key findings of the Security Index 2025 as an integrated analytical framework for understanding human security in North Macedonia. Rather than treating security as a narrow concept limited to crime statistics or military threats, the study approached it as a multidimensional and systemic phenomenon shaped by institutional performance, socio-economic conditions, environmental risks, and informational integrity. The index was designed to capture both objective indicators and subjective perceptions, thereby bridging quantitative measurement with citizens’ lived experiences.

The structure of the research followed a comprehensive mixed-methods design. It combined six core dimensions of security—personal and social security, economic security, health security, national security and rule of law, environmental security, digital and cognitive security—alongside a qualitative component that captured citizens’ open-ended reflections.

Each dimension was analysed through survey data using scaled indicators and statistical distributions, while the qualitative section provided interpretative depth through

thematic coding of citizens' proposals and concerns. This layered approach allowed for both macro-level comparison across dimensions and micro-level understanding of underlying perceptions.

The findings reveal a highly asymmetric security landscape. While personal and social security remain relatively stable, this stability is largely confined to the private sphere, particularly the home, which functions as a perceived "safe zone" in contrast to the public space. In parallel, significant vulnerabilities are identified in health security and institutional-political security, where trust in public systems—especially judiciary, governance structures, and service delivery—is notably weakened. Economic security reflects a paradox of employment stability combined with financial insecurity, while environmental security emerges as one of the most acute perceived threats, particularly due to air pollution and waste management issues.

A particularly important finding is the growing significance of cognitive and digital security. Citizens demonstrate high awareness of disinformation as a national security threat, with a substantial majority perceiving fake news and information manipulation as highly or very highly dangerous. This positions information integrity as a central component of contemporary security, closely linked to institutional trust and political stability. Similarly, concerns regarding foreign influence and hybrid threats further reinforce the perception of North Macedonia as a geopolitically exposed and informationally vulnerable state.

The Security Index 2025 provides an analytical framework for assessing citizens' perceptions of societal resilience and security across multiple dimensions. The findings are consistent with the central premise of this study, suggesting that citizens perceive security as a multidimensional phenomenon extending beyond traditional crime and military concerns. The findings suggest that the absence of armed conflict alone is insufficient to ensure a strong subjective sense of security among citizens. The relatively low levels of trust in several public institutions and concerns regarding governance, healthcare, and corruption indicate challenges commonly associated with the broader concept of positive peace.

The composite picture for 2025 reveals an asymmetric security landscape. Personal and social security appear relatively more stable, yet this stability is largely anchored in the perceived safety of the private sphere, particularly the home. In contrast, the most vulnerable dimensions are health security and the institutional-political domain, indicating that the main sources of insecurity are structural and existential rather than incidental. Citizens do not perceive security solely as protection from crime or violence; instead, it extends to access to healthcare, financial stability, trust in the judiciary, environmental quality, protection of personal data, and resilience to disinformation.

The qualitative findings complement these results by revealing strong citizen demand for systemic reform. The dominant narratives emphasise equal justice, depoliticisation of institutions, professionalisation of public administration, visible policing, and stronger enforcement of laws. The qualitative responses frequently emphasise functional governance over abstract policy solutions, highlighting a clear expectation for restored institutional credibility, fairness, and accountability.

Overall, the Security Index 2025 demonstrates that security in North Macedonia is no longer defined primarily by traditional threats, but by the quality of governance, institutional trust, and societal resilience. The study confirms that the absence of conflict does not equate to perceived safety, as structural deficiencies in institutions, public services, and information environments continue to generate deep-seated insecurity. The

findings indicate that strengthening institutional performance, public trust, and governance may contribute to improving citizens' perceptions of security and societal resilience.

7. REFERENCES

- Alkire, S. (2003). *A Conceptual Framework for Human Security*. CRISE Working Paper No. 2. University of Oxford. <https://ora.ox.ac.uk/objects/uuid:d2907237-2a9f-4ce5-a403-a6254020052d>
- Bornmann, L., Marx, W. Thomas theorem in research evaluation. *Scientometrics* 123, 553–555 (2020). <https://doi.org/10.1007/s11192-020-03389-6>
- Buzan, B. (1991). *People, States and Fear: An Agenda for International Security Studies in the Post-Cold War Era* (2nd ed.). Boulder, CO: Lynne Rienner.
- Buzan, B., Wæver, O., & de Wilde, J. (1998). *Security: A new framework for analysis*. Lynne Rienner Publishers.
- Economist Intelligence Unit. (2026). Democracy Index. <https://www.eiu.com/n/>
- Gasper, D., & Gómez, O. A. (2015). Human security thinking in practice: ‘Personal security’, ‘Citizen security’ and comprehensive mappings. *Contemporary Politics*, 21(1), 100–116. <http://hdl.handle.net/1765/77458>
- Gjurovski, M., Nikolovski, M., Gerasimoski, S. (2020). Security Risks and Threats – Treatment, Phenomenology and Mapping. Konrad Adenauer Foundation Skopje and Faculty of Security – Skopje. <https://fb.uklo.edu.mk/wp-content/uploads/sites/10/2021/12/Securiy-Risks-and-Threats-MKD.pdf>
- Iganski, P. (2019). Hate Crime Victimization Survey. Report. OSCE Mission to Skopje. <https://skopje.osce.org/mk/node/424193>
- INDAGO. (2024). Perceptions, Views and Challenges of Young People in North Macedonia. Research on Youth and Security. OSCE Mission to Skopje. <https://skopje.osce.org/mk/node/575137>
- Institute for Economics & Peace. (2024a). Global Peace Index 2024: Measuring Peace in a Complex World. <https://www.economicsandpeace.org/wp-content/uploads/2024/06/GPI-2024-web.pdf>
- Institute for Economics & Peace. (2024b). Positive Peace Report 2024. <https://www.economicsandpeace.org/wp-content/uploads/2024/04/PPR-2024-web.pdf>
- Jakovleska, G. (2025). Citizens Perceptions on Security. Macedonian Center for International Cooperation. <https://mcms.mk/mk/za-nasata-rabota/istrazuvana-i-publikacii/2650-percepcii-na-gragjanite-za-bezbednosta.html>
- Krause, K., & Williams, M. C. (1996). Broadening the agenda of security studies: Politics and methods. *Mershon International Studies Review*, 40(2), 229–254.
- Mojanoski, C., Dimovski, Z., Gjurovski, M., Ilijevski, I. (2015). Security and Security Threats in the Republic of Macedonia. Scientific Survey Report. Faculty of Security – Skopje. <https://fb.uklo.edu.mk/wp-content/uploads/sites/10/2021/12/BEZBEDNOSTA-I-BEZBEDNOSNITE-ZAKANI-NA-REPUBLIKA-MAKEDONIJA-Istrazhuvachki-izveshtaj-2015.pdf>
- Mojanoski, C., Dujovski, N., Gjurovski, M. (2018). Police and Public. Faculty of Security – Skopje. <https://fb.uklo.edu.mk/wp-content/uploads/sites/10/2021/12/POLITSIJATA-I-JAVNOSTA-Analiza-na-istrzhuvachki-rezultati-2009-2017-godina.pdf>

- Mojanoski, C., Gjurovski, M. (2018a). Index of Security I. Faculty of Security – Skopje and Konrad Adenauer Foundation Skopje. <https://fb.uklo.edu.mk/wp-content/uploads/sites/10/2021/12/INDEKS-NA-BEZBEDNOST.pdf>
- Mojanoski, C., Gjurovski, M. (2018b). Index of Security II. Faculty of Security – Skopje and Konrad Adenauer Foundation Skopje. <https://fb.uklo.edu.mk/wp-content/uploads/sites/10/2021/12/SECURITY-INDEX-II-ENG MKD.pdf>
- Mojanoski, C., Malish-Sazdovska, M., Nikolovski, M., Krstevska, K. (2018). Citizens for the Corruption. Faculty of Security – Skopje. <https://fb.uklo.edu.mk/wp-content/uploads/sites/10/2021/12/GRAGANITE-ZA-KORUPTSIJATA-Analiza-na-istrazhuvachki-rezultati-2013-2017.pdf>
- Mojanoski, C., Nikolovski, M., Batic, D., Dujovski, N., Gjurovski, M., Ristevska-Dimitrovska, G., Nuhui, A., Ramadani, A. (2021). Functioning of the Security System of the Republic of North Macedonia in Emergency and Crisis Situation – Case Study. <https://fb.uklo.edu.mk/wp-content/uploads/sites/10/2021/12/Funkcioniraweto-na-Bezbednosniot-sistem-na-RSM-vo-vreme-na-vonredna-i-krizna-sostojba.pdf>
- Nichols, L. T. (2016). The Enduring Social Psychology of Robert K. Merton: Motivating Sentiments, Reference Groups and Self-Fulfilling Prophecies. *The American Sociologist*, 47(2/3), 356–381. <http://www.jstor.org/stable/43957109>
- Nikolovski, M. (2026). Security Systems. Faculty of Security, Skopje.
- OSCE. (2019). Trust in Police Goes Up, Shows OSCE-Supported Survey Presented in Skopje. <https://skopje.osce.org/mission-to-skopje/442759>
- Paris, R. (2001). Human Security: Paradigm Shift or Hot Air? *International Security*, 26(2), 87–102.
- RCC. (2025). Securimeter 2024, https://www.rcc.int/securimeter/key_findings_2024
- Savic, Lj., Gacinovic, R. (2007). Introduction to Security System. Draslar Partner, Belgrade.
- Tadjbakhsh, S., & Chenoy, A. M. (2007). *Human Security: Concepts and Implications*. Routledge. <https://doi.org/10.4324/9780203965955>
- Transparency International. (2024). Corruption Perceptions Index 2024. <https://www.transparency.org/en/cpi/2024>
- UNDP. (1994). Human Development Report 1994: New Dimensions of Human Security. <https://hdr.undp.org/content/human-development-report-1994>
- UNDP. (2026). Human Development Index. <https://hdr.undp.org/data-center/human-development-index>
- World Justice Project. (2026). Rule of Law Index. <https://worldjusticeproject.org/rule-of-law-index/>

PREVENTING VIOLENT EXTREMISM: MULTI-LEVEL GOVERNANCE APPROACHES FOR BUILDING RESILIENT AND SUSTAINABLE SOCIETIES

Agus Demirovski

PhD Candidate at Faculty of Security – Skopje, UKLO

demirovski@gmx.com

<https://orcid.org/0009-0009-3319-2460>

Abstract

Violent extremism remains a complex and evolving security challenge that requires comprehensive and multi-dimensional responses. This paper examines how multi-level governance approaches contribute to the prevention of violent extremism and the strengthening of resilient and sustainable societies. It argues that effective prevention cannot be achieved solely through security institutions, but requires the coordinated engagement of state actors, local communities, civil society organisations, and regional cooperation mechanisms.

The study is based on an analytical framework that integrates institutional, societal, and environmental dimensions of security. It explores how early warning and prevention mechanisms, intelligence and law enforcement cooperation, and community-based initiatives interact in addressing drivers of radicalisation. Particular attention is given to the role of local communities and civil society actors in building social cohesion, promoting tolerance, and reducing vulnerability to extremist narratives. Furthermore, the paper highlights the importance of regional and international cooperation in addressing the transnational nature of violent extremism, particularly through information sharing and joint preventive efforts. The analysis suggests that fragmented responses reduce overall effectiveness, while integrated governance structures enhance resilience. The findings indicate that sustainable prevention strategies depend on strengthening institutional coordination, improving trust between communities and authorities, and ensuring the active participation of diverse societal actors.

In conclusion, the paper contributes to ongoing academic and policy debates by proposing that resilient and sustainable societies are best achieved through inclusive, multi-level governance systems that address both structural and immediate drivers of violent extremism.

Key words: Violent extremism; Multi-level governance; Prevention strategies; Resilience building; Community engagement.

1. INTRODUCTION

Violent extremism has emerged as one of the most complex and persistent security challenges of the twenty-first century. Unlike conventional security threats, violent extremism transcends national borders, exploits political, social, economic, and technological vulnerabilities, and increasingly affects the stability of both developed and developing societies. Contemporary extremist movements demonstrate a remarkable

ability to adapt their strategies, utilise digital communication platforms for recruitment and propaganda, and exploit existing social grievances to mobilise individuals toward violence. Consequently, preventing violent extremism (PVE) has become a central priority within international security agendas, requiring responses that extend beyond traditional counter-terrorism measures.

For many years, responses to violent extremism focused predominantly on intelligence gathering, law enforcement, and military operations. Although these security-oriented approaches remain indispensable in addressing immediate threats, growing evidence suggests that they are insufficient to tackle the underlying conditions that facilitate radicalisation and recruitment into violent extremist organisations. As a result, international organisations, governments, and scholars increasingly advocate preventive strategies that address both the structural and individual-level drivers of violent extremism through coordinated and inclusive governance arrangements (United Nations, 2006).

Understanding why individuals become involved in violent extremist movements remains fundamental to developing effective prevention policies. Existing research demonstrates that no single factor explains radicalisation. Rather, individuals are influenced by a combination of political exclusion, identity crises, perceived injustice, socio-economic marginalisation, ideological narratives, social networks, and psychological motivations (Gómez et al., 2021). Because these drivers vary across societies and communities, prevention strategies must be context-specific and involve a broad spectrum of governmental and non-governmental actors capable of addressing local vulnerabilities before they escalate into violence.

Conceptually, violent extremism remains a contested term. Definitions often differ according to political, legal, and institutional contexts, making it difficult to establish universally accepted boundaries between radical beliefs, extremist ideologies, and violent behaviour (Lowe, 2017; Schmid, 2013). While radicalisation generally refers to a process through which individuals adopt increasingly extreme views, not all forms of radicalisation necessarily result in violence. Throughout history, many radical social and political movements have contributed positively to democratic change without resorting to violence. The critical distinction lies in the acceptance, justification, or use of violence to achieve political, ideological, or religious objectives (Arifi, 2019; Schmid, 2013).

In this paper, violent extremism is understood as the actual or threatened use of violence by non-state actors pursuing political, ideological, or religious goals that undermine democratic values, social cohesion, and the rule of law (McNeil-Willson et al., 2019). Such groups often seek to establish homogeneous societies by suppressing pluralism, excluding minority communities, and rejecting democratic institutions (Mudde, 2007). Importantly, violent extremism should not be viewed solely through the prism of religiously motivated violence. Contemporary manifestations include ethnonationalist, far-right, far-left, separatist, and other ideologically motivated forms of extremism, many of which coexist and interact within increasingly polarised societies (Arifi, 2019).

These developments highlight the necessity of distinguishing between counter-terrorism and preventing violent extremism. Whereas counter-terrorism primarily responds to acts of terrorism after they have occurred or become imminent, PVE seeks to reduce the risk of radicalisation by strengthening resilience within individuals, communities, and institutions. Prevention therefore extends beyond the responsibilities of security agencies and requires meaningful cooperation among national governments, local authorities, educational institutions, civil society organisations, religious communities, the private sector, and international partners. Building trust between authorities and communities,

promoting social inclusion, countering extremist narratives, and strengthening community resilience have become essential components of sustainable prevention strategies.

Against this background, this paper examines preventing violent extremism through the perspective of multi-level governance. It argues that fragmented institutional responses cannot effectively address the complex and interconnected drivers of violent extremism. Instead, sustainable prevention depends on coordinated governance mechanisms that integrate national policies with local initiatives, foster collaboration among diverse stakeholders, and strengthen institutional and societal resilience. By analysing institutional coordination, community engagement, and regional cooperation, the paper contributes to the growing academic and policy debate on how inclusive governance approaches can enhance the prevention of violent extremism and support the development of resilient and sustainable societies.

2. METHODOLOGY

2.1 Research Design

This study adopts a qualitative research design based on desk research and conceptual analysis to examine the role of multi-level governance in preventing violent extremism (PVE). Rather than focusing on a single national case study or empirical dataset, the paper synthesises existing academic literature, international policy documents, and institutional frameworks to identify the governance mechanisms that contribute to effective prevention strategies. This approach is appropriate given the multidimensional nature of violent extremism, which encompasses security, governance, social cohesion, education, community development, and international cooperation.

The analysis draws upon peer-reviewed scientific literature, reports published by international organisations, including the United Nations (UN), the European Union (EU), the Organization for Security and Co-operation in Europe (OSCE), and other relevant policy documents addressing violent extremism, resilience, and collaborative governance. These sources provide a comprehensive understanding of contemporary approaches to preventing violent extremism and the evolving role of governmental and non-governmental stakeholders in addressing the phenomenon.

A qualitative analytical approach was selected because preventing violent extremism cannot be adequately explained through quantitative indicators alone. Governance processes, institutional interactions, community engagement, and preventive policies involve complex relationships among multiple actors operating at different administrative and societal levels. Consequently, this study focuses on identifying patterns of cooperation, institutional arrangements, and governance practices that support resilience and sustainable prevention.

2.2 Analytical Framework

The paper is guided by a multi-level governance analytical framework that examines preventing violent extremism through three interrelated dimensions: the institutional, societal, and external (regional and international) dimensions. This framework recognises that violent extremism is influenced by multiple interacting factors and therefore requires coordinated responses that extend beyond traditional security institutions.

The **institutional dimension** examines the capacity of national governments, law enforcement agencies, intelligence services, judicial institutions, and local authorities to

develop coordinated prevention mechanisms. Particular attention is given to inter-agency cooperation, early warning systems, information sharing, strategic policy development, and institutional coordination, all of which are considered essential for identifying and mitigating risks before violent extremism materialises.

The **societal dimension** focuses on the contribution of local communities, civil society organisations, educational institutions, religious communities, and other non-state actors in strengthening social cohesion and resilience. This dimension recognises that communities often possess the local knowledge, trust, and legitimacy necessary to identify early signs of radicalisation and to implement preventive initiatives that promote inclusion, tolerance, and democratic values. Community engagement is therefore considered a fundamental component of sustainable prevention strategies.

The **external dimension** addresses the role of regional and international cooperation in responding to the increasingly transnational character of violent extremism. Contemporary extremist networks frequently operate across national borders, making international collaboration essential for intelligence exchange, capacity building, policy coordination, and the dissemination of good practices. The framework therefore considers how regional organisations and international partnerships complement national and local prevention efforts through collaborative governance arrangements.

By integrating these three dimensions, the analytical framework enables a comprehensive assessment of preventing violent extremism as a shared governance responsibility rather than an exclusively security-oriented function. It further allows the paper to explore how coordination across different levels of governance contributes to building resilient and sustainable societies capable of addressing both the structural conditions and immediate risks associated with violent extremism.

3. RESULTS

This section presents the key findings of the study through a multi-dimensional analysis of preventing violent extremism (PVE) within a multi-level governance framework. It is structured around five interconnected dimensions that together explain how violent extremism emerges, how it can be prevented, and how governance systems at different levels contribute to resilience-building.

The first subsection examines the multi-factorial nature of violent extremism, highlighting its complex interaction of structural, social, and individual drivers. This provides the analytical foundation for understanding why single-factor explanations and isolated policy responses are insufficient. Building on this, the second subsection focuses on institutional coordination and early prevention mechanisms, with particular attention to the role of state institutions, inter-agency cooperation, and early warning systems in identifying and addressing risks at an early stage. The third subsection shifts the focus to community resilience and civil society engagement, emphasising the importance of trust, social cohesion, and local-level actors in strengthening protective factors against radicalisation. The fourth subsection explores regional and international cooperation, underlining the transnational character of violent extremism and the necessity of coordinated responses across borders and institutions. Finally, the fifth subsection analyses European Union approaches to PVE, illustrating how supranational frameworks operationalise multi-level governance through policy coordination, funding instruments, and collaborative security initiatives.

Together, these five dimensions provide a comprehensive understanding of how

violent extremism can be addressed through integrated governance structures that combine security, social, and international policy instruments.

3.1 Multi-factorial nature of violent extremism

Violent extremism is a complex and multi-factorial phenomenon that cannot be explained through a single causal variable. Instead, its emergence results from the interaction of situational, social, cultural, and individual-level factors that converge in specific political and social contexts. In this sense, violent extremism should be understood as a dynamic process shaped by conditions operating at different levels of society, rather than as a linear outcome of isolated determinants (Allan et al., 2015; McCauley & Moskalenko, 2017; Khalil, J., & Zeuthen, 2016; Sedgwick, 2010).

A useful analytical approach distinguishes these drivers across three interrelated levels: the macro, meso, and micro levels. At the **macro level**, structural and situational conditions such as governance deficits, weak state institutions, political instability, and unresolved grievances may create environments in which extremist narratives gain traction. While these conditions do not directly produce violent extremism, they may generate opportunities for mobilisation, particularly in contexts where state legitimacy is weakened or where public services and security provision are insufficient. In such environments, extremist actors may exploit institutional weaknesses by offering alternative governance structures, social services, or protection, thereby increasing their appeal.

At the **meso level**, social, cultural, and group-based dynamics play a central role in shaping pathways toward radicalisation. Research consistently highlights that violent extremism is fundamentally a social process in which identity formation, group belonging, and perceived collective grievances are critical factors (Kruglanski et al., 2014). Religion, ethnicity, ideology, and other forms of collective identity often function as powerful mobilisation tools. Within this context, civil society organisations, local communities, and religious institutions may serve as either protective or enabling environments, depending on their capacity to promote inclusion, dialogue, and resilience. Evidence suggests that social exclusion and blocked political participation can reinforce grievances that extremist groups strategically exploit for recruitment purposes (Neumann, 2013).

At the **micro level**, individual psychological and cognitive factors influence vulnerability to radicalisation, although their explanatory power is often context-dependent. While most individuals who become involved in violent extremist movements are young males, empirical research does not support simplistic causal links based on gender, education, or socio-economic status alone (Horgan, 2008). Instead, identity-related processes, perceived injustice, personal experiences of marginalisation, and cognitive openness to ideological narratives interact in shaping individual susceptibility. Importantly, education is not a direct risk factor for radicalisation, and in some cases religious or civic education may function as a protective mechanism rather than a driver of extremism.

Socio-economic conditions, including poverty, unemployment, and inequality, demonstrate mixed but context-sensitive effects. Although these factors alone are insufficient to explain violent extremism, they may become relevant in environments characterised by structural discrimination, state fragility, or prolonged conflict. In such cases, extremist organisations may increase their legitimacy by providing governance functions, social services, or economic opportunities that the state fails to deliver (UNDP, 2017). However, poverty should not be interpreted as a direct causal factor but rather as part of a broader ecosystem of vulnerability shaped by political and institutional dynamics.

Overall, the multi-level nature of violent extremism highlights the need for

similarly multi-layered prevention strategies. The alignment between macro-level structural conditions, meso-level social dynamics, and micro-level psychological factors suggests that prevention efforts must go beyond traditional security responses. Instead, they require coordinated governance approaches that integrate state institutions, local authorities, and civil society actors. Within this framework, preventing violent extremism becomes not only a security objective but also a governance challenge focused on strengthening resilience, inclusion, and institutional legitimacy across all levels of society.

3.2. Institutional Coordination and Early Prevention Mechanisms

Institutional coordination and early prevention mechanisms constitute a central pillar in contemporary approaches to preventing violent extremism (PVE). Within multi-level governance frameworks, effective prevention depends on the ability of state institutions to cooperate horizontally across sectors and vertically across administrative levels, ensuring timely identification of risks and coordinated responses. The shift from reactive counter-terrorism measures to proactive prevention has significantly increased the importance of early warning systems, inter-agency cooperation, and integrated policy frameworks capable of addressing radicalisation processes before they escalate into violence.

At the institutional level, early prevention mechanisms are primarily embedded within law enforcement, intelligence services, and judicial systems. These actors are responsible for detecting behavioural indicators of radicalisation, monitoring extremist networks, and disrupting potential pathways to violence. However, contemporary policy approaches emphasise that these institutions cannot operate effectively in isolation. Instead, they require structured cooperation with education systems, social services, local authorities, and public health institutions to ensure a comprehensive understanding of vulnerability factors and early signs of radicalisation (Schmid, 2013; Neumann, 2013).

A key development in European security governance has been the strengthening of information-sharing mechanisms between Member States and EU agencies. Europol, through its European Counter Terrorism Centre (ECTC), plays a critical role in facilitating operational cooperation, intelligence exchange, and analytical support across jurisdictions. Similarly, Eurojust enhances judicial coordination in cross-border terrorism-related investigations. These mechanisms reflect the recognition that violent extremism is a transnational phenomenon that cannot be effectively addressed by national institutions acting independently (Coolsaet, 2010; Europol, 2025).

Early prevention also relies on the development of risk assessment frameworks designed to identify individuals or groups at risk of radicalisation. However, academic literature highlights the limitations of purely predictive models, particularly due to the complexity and non-linear nature of radicalisation processes. Research shows that there is no single pathway into violent extremism, and risk factors often interact in context-specific and unpredictable ways (Horgan, 2008; McCauley & Moskalenko, 2017). Consequently, institutional approaches increasingly combine structured risk assessment tools with qualitative, case-based assessments carried out by trained professionals.

In addition to formal security institutions, local governance structures play a critical role in early prevention. Municipal authorities, schools, youth services, and community-based organisations are often the first to observe behavioural changes or signs of social disengagement. As such, effective prevention strategies depend on building trust-based relationships between these actors and security institutions.

Without trust, information-sharing mechanisms may be underutilised, limiting the effectiveness of early intervention systems (RAN, 2021b).

Another important dimension of institutional coordination concerns multi-agency cooperation frameworks. Many European countries have developed specialised coordination units or task forces that bring together police officers, social workers, educators, and mental health professionals. These multidisciplinary approaches aim to ensure that responses to potential radicalisation cases are proportionate, context-sensitive, and grounded in safeguarding principles rather than exclusively security-oriented interventions. This reflects a broader shift towards prevention models that prioritise early support and disengagement rather than punitive measures (Sedgwick, 2010).

Despite these advancements, significant challenges remain. Institutional fragmentation, data protection concerns, and differing national legal frameworks often hinder effective coordination across agencies and borders. Furthermore, tensions may arise between security objectives and human rights considerations, particularly regarding surveillance practices and information sharing. These challenges underscore the importance of balancing security imperatives with rule of law standards and democratic accountability (Ragazzi, 2017).

Overall, institutional coordination and early prevention mechanisms represent a critical component of effective PVE strategies. While significant progress has been made in developing multi-agency and cross-border cooperation structures, their effectiveness ultimately depends on trust, institutional capacity, and the integration of non-security actors into prevention frameworks. Strengthening these mechanisms remains essential for enhancing early detection, improving response coordination, and reducing the risk of radicalisation progressing into violent extremism.

3.3. Community Resilience and Civil Society Engagement

Community resilience and civil society engagement have become central components of contemporary approaches to preventing and countering violent extremism (P/CVE). Over the past decade, policy frameworks at the European and international levels have increasingly recognised that sustainable prevention cannot be achieved solely through state security institutions, but requires active participation from local communities, civil society organisations (CSOs), educators, and other non-state actors embedded in everyday social environments (European Commission, 2020; RAN, 2021a; Stephens & Sieckelinck, 2020).

The concept of community resilience in the context of violent extremism refers to the capacity of communities to resist, adapt to, and recover from processes of radicalisation and extremist influence. This perspective shifts the focus from purely reactive security responses to proactive strengthening of social cohesion, trust, and local agency. Recent literature emphasises that radicalisation is fundamentally a social process, occurring within networks of relationships rather than in isolation, which makes community-level intervention essential for early prevention (Prislan, Borovec, & Cajner Mraović, 2020; Stephens, Sieckelinck, & Boutellier, 2019).

Within the European Union framework, community resilience is closely linked to the “whole-of-society” approach promoted in EU counter-terrorism and security strategies. The EU highlights the importance of empowering local actors who are often best positioned to detect early warning signs of radicalisation and to intervene before escalation occurs.

These include teachers, youth workers, social services, healthcare professionals, and religious leaders, whose daily interaction with individuals at risk enables early identification of behavioural and social changes (European Commission, 2020; RAN, 2021a).

Civil society organisations play a particularly important role in bridging the gap between state institutions and local communities. Research shows that CSOs often possess higher levels of trust and legitimacy in vulnerable communities, allowing them to engage individuals who may be reluctant to interact with formal security structures (Baumann, 2025). Through education programmes, youth engagement initiatives, intercultural dialogue, and psychosocial support, CSOs contribute to reducing social exclusion and strengthening protective factors against radicalisation.

A key dimension of community resilience is the concept of social capital, which refers to networks of trust, cooperation, and shared norms within and between communities. Strong bonding, bridging, and linking social capital has been identified as a critical protective factor against violent extremism, as it enhances inclusion, reduces marginalisation, and strengthens connections between citizens and institutions (Dalgaard-Nielsen & Schack, 2016; Ellis & Abdi, 2017). When social capital is weak or fragmented, communities become more vulnerable to extremist narratives that exploit feelings of exclusion and injustice.

Despite its importance, civil society engagement in P/CVE also faces several challenges. These include limited funding sustainability, lack of institutional coordination, and occasional mistrust between state authorities and non-governmental actors. In some cases, CSOs are perceived as security partners rather than independent community actors, which may reduce their legitimacy and effectiveness in sensitive environments (RAN, 2021b). Additionally, there is an ongoing debate regarding the balance between security-oriented prevention and rights-based community engagement, particularly in relation to safeguarding human rights and avoiding the securitisation of social work.

Overall, community resilience and civil society engagement represent essential pillars of a multi-level governance approach to preventing violent extremism. While state institutions provide the legal and operational framework for security, civil society actors contribute legitimacy, trust, and local knowledge that are indispensable for early prevention. Strengthening cooperation between these actors is therefore crucial for building sustainable resilience and ensuring that prevention strategies address not only the symptoms but also the underlying social conditions that enable radicalisation.

3.4. Regional and International Cooperation

Regional and international cooperation represents a critical pillar in contemporary strategies for preventing and countering violent extremism (P/CVE), reflecting the inherently transnational nature of extremist networks, ideological flows, and security risks. As violent extremism increasingly transcends national borders, no single state is capable of addressing its drivers, financing structures, and operational networks in isolation. Consequently, multilateral cooperation has become an essential component of global and regional security governance frameworks (European Commission, 2020a; OSCE, 2020).

At the European level, the European Union has developed an extensive architecture for cooperation in counter-terrorism and PVE, combining institutional coordination, intelligence sharing, and external action. Agencies such as Europol and Eurojust facilitate operational cooperation between Member States by supporting cross-border investigations, enhancing judicial coordination, and enabling real-time exchange of intelligence and

analytical data. This institutional framework reflects the EU's recognition that fragmented national responses are insufficient to address the mobility and adaptability of extremist actors (Coolsaet, 2010; Europol, 2025).

In addition to internal cooperation, the EU places significant emphasis on external partnerships. Through instruments such as the Instrument contributing to Stability and Peace (IcSP), the Union supports prevention programmes in partner countries, focusing on education, youth engagement, media resilience, and community-based interventions. These initiatives reflect a broader understanding that structural drivers of violent extremism—such as governance deficits, social exclusion, and weak institutional capacity—must be addressed not only within Europe but also in regions affected by instability and conflict (European Commission, 2020; Macchiarini Crosson, Stankovic, Rieker, & Blockmans, 2023).

International organisations also play a central role in shaping normative and operational frameworks for PVE. The United Nations Global Counter-Terrorism Strategy emphasises the importance of addressing conditions conducive to the spread of violent extremism, including poverty, lack of opportunity, and marginalisation, while ensuring that responses remain grounded in human rights and the rule of law (United Nations General Assembly, 2015). Similarly, the United Nations Development Programme (UNDP) highlights the importance of inclusive development and governance as long-term preventive mechanisms, demonstrating that security-oriented approaches alone are insufficient to mitigate radicalisation risks (UNDP, 2016).

The Organization for Security and Co-operation in Europe (OSCE) further complements these efforts by promoting a comprehensive, multi-stakeholder approach to P/CVE in the wider European and Eurasian region. OSCE frameworks emphasise the importance of collaboration between state institutions, civil society organisations, and local communities, particularly in developing early warning systems, referral mechanisms, and rehabilitation and reintegration programmes (OSCE, 2020; OSCE, 2019). This reflects an increasingly widely accepted principle that effective prevention requires engagement at multiple levels of governance, including local, national, and regional actors.

Global cooperation is further reinforced through platforms such as the Global Counterterrorism Forum (GCTF) and INTERPOL, which support the exchange of best practices, capacity building, and operational coordination between law enforcement agencies worldwide. These mechanisms contribute to the standardisation of approaches in key areas such as foreign terrorist fighters, online radicalisation, and terrorist financing, thereby reducing fragmentation in global counter-terrorism responses.

Despite these advancements, significant challenges remain. One of the key limitations of international cooperation is the uneven capacity of states to implement agreed frameworks, as well as differences in legal systems, political priorities, and institutional capabilities. Furthermore, concerns regarding human rights protection, data privacy, and the securitisation of development and civil society engagement continue to generate debate regarding the legitimacy and effectiveness of certain P/CVE practices (Kundnani & Hayes, 2018; Ragazzi, 2017).

Overall, regional and international cooperation constitutes an indispensable dimension of multi-level governance in preventing violent extremism. It enables information sharing, capacity building, and policy coordination across borders, while also promoting shared normative standards based on rule of law and human rights. However, its effectiveness depends on sustained political commitment, institutional coherence, and

the meaningful inclusion of local actors within broader transnational prevention frameworks.

3.5 EU Approaches to Preventing Violent Extremism

Over the past decade, the European Union has progressively developed a comprehensive and multi-dimensional approach to countering terrorism and preventing violent extremism (PVE). This evolution reflects a broader shift from predominantly reactive security responses towards more preventive, resilience-oriented, and governance-based strategies. In particular, the European Commission's *Counter-Terrorism Agenda for the EU* and the *EU Security Union Strategy* emphasise that contemporary security threats are complex, interconnected, and require coordinated action across multiple levels of governance (European Commission, 2020a, 2020b).

A central feature of the EU approach is the integration of prevention, protection, preparedness, and response into a coherent policy framework (Coolsaet, 2010). The four pillars—*anticipate, prevent, protect, and respond*—highlight the need for early identification of threats, disruption of radicalisation processes, safeguarding of public spaces and critical infrastructure, and effective response mechanisms in the aftermath of violent incidents. This framework illustrates a balance between hard security measures and soft preventive interventions, recognising that sustainable security cannot be achieved through enforcement alone.

In this context, preventing violent extremism is increasingly understood as a governance and societal challenge rather than solely a security issue. EU strategies emphasise addressing the enabling environment in which radicalisation occurs, including structural vulnerabilities, social exclusion, and ideological manipulation. Particular attention is given to strengthening resilience at the community level by supporting local actors, civil society organisations, educators, youth workers, and religious leaders who are often best positioned to identify early signs of radicalisation and intervene before escalation (Ernstorfer, 2018). The EU also prioritises capacity building in the fields of law enforcement and judicial cooperation, recognising that effective prevention requires strong institutional coordination across Member States (EUROPOL, 2024). Information sharing, cross-border cooperation, and harmonised legal frameworks are key elements in addressing the transnational nature of violent extremism. At the same time, the Union places increasing emphasis on disrupting terrorist financing networks and countering online radicalisation, particularly through monitoring and removal of extremist content in digital spaces.

A distinctive feature of the EU approach is its strong external dimension. Through instruments such as the *Instrument contributing to Stability and Peace*, the EU supports a wide range of global initiatives aimed at preventing violent extremism in partner countries. These initiatives include education programmes, youth engagement, media capacity building, women's empowerment, and community resilience projects. Over the past decade, the EU has financed more than one hundred actions globally, reflecting its commitment to addressing both immediate risks and structural drivers of extremism in diverse contexts.

Importantly, EU policy recognises that no single actor or intervention can effectively address violent extremism in isolation (Ragazzi, 2017; Ranstorp, 2016). Instead, it promotes a “whole-of-society” and multi-level governance approach grounded in four key principles: evidence-based policy design, context-sensitive programming, local ownership, and the integration of human rights and rule of law standards.

These principles ensure that prevention efforts are not only effective but also legitimate and sustainable in the long term.

Furthermore, EU engagement is strongly aligned with broader international frameworks, including cooperation with the United Nations, INTERPOL, and the Global Counterterrorism Forum. This multilateral orientation reflects the transnational nature of violent extremism, where threats and responses extend beyond national and regional boundaries. At the same time, the EU increasingly recognises the importance of civil society partnerships and grassroots initiatives, particularly in addressing local grievances and strengthening community cohesion.

Overall, the EU approach demonstrates a clear transition from a predominantly security-driven model to an integrated prevention framework that combines institutional coordination, community resilience, and international cooperation. However, while significant progress has been made in developing policy instruments and funding mechanisms, challenges remain in ensuring effective implementation across diverse national and local contexts. This highlights the continuing need for stronger policy coherence and deeper engagement with local-level governance structures.

4. DISCUSSION: BUILDING RESILIENT AND SUSTAINABLE SOCIETIES THROUGH INTEGRATED GOVERNANCE

The findings of this study highlight that preventing violent extremism (PVE) is most effective when approached through integrated governance structures that connect institutional, societal, and international dimensions of security. Rather than relying on isolated or predominantly security-driven interventions, contemporary prevention frameworks increasingly depend on coordinated and sustained action across multiple levels of governance. This includes the combined efforts of state institutions, local authorities, civil society organisations, educational actors, and international partners. Such an integrated perspective reflects a broader transformation in security governance, moving away from hierarchical, state-centric models towards more networked, participatory, and collaborative systems of risk management and resilience-building.

A key implication emerging from the analysis is that violent extremism is not a linear or easily predictable phenomenon, but rather a complex and adaptive process shaped by the interaction of structural conditions, social environments, and individual pathways. This complexity significantly limits the effectiveness of single-factor explanations or purely reactive policy responses. As a result, governance strategies must be equally multidimensional and flexible. Institutional coordination mechanisms remain essential, particularly in ensuring legal authority, operational capacity, and strategic direction. However, they are insufficient on their own unless complemented by trust-based relationships with communities and sustained engagement with non-state actors. This reinforces the broader argument that security and resilience are co-produced through interaction, negotiation, and cooperation rather than imposed exclusively through top-down enforcement.

The study further demonstrates that early prevention mechanisms are most effective when they are embedded within broader systems of social support, inclusion, and institutional trust. Early warning systems, referral mechanisms, and risk identification tools can only function effectively when there is a stable flow of communication between institutions and communities.

This requires not only technical capacity but also relational infrastructure built on credibility and legitimacy. However, this relationship remains fragile in many contexts due to historical mistrust, politicisation of security practices, uneven institutional responsiveness, and concerns about surveillance or stigmatisation.

These factors may discourage communities from engaging with formal institutions, thereby limiting the effectiveness of early intervention systems. Consequently, governance models must prioritise transparency, proportionality, accountability, and rights-based approaches in order to maintain legitimacy and ensure long-term cooperation.

At the societal level, community resilience emerges as a critical protective factor against radicalisation and violent extremism. Strengthening social cohesion, fostering inclusive identities, and enhancing civic participation all contribute to reducing vulnerability to extremist narratives that exploit feelings of exclusion, injustice, or marginalisation. In this regard, civil society organisations play a particularly important intermediary role, as they are often embedded within local contexts and possess higher levels of trust among community members compared to formal state institutions. This enables them to act as bridges between citizens and authorities, facilitating dialogue, delivering tailored prevention programmes, and supporting individuals at risk. However, their effectiveness is contingent upon sustainable funding, institutional recognition, and the preservation of their independence, particularly in contexts where there is a risk of over-securitisation of civil society space.

The international and regional dimension further reinforces the necessity of cooperation in addressing violent extremism. The transnational nature of extremist ideologies, digital communication networks, and financial flows means that no single state can effectively manage these threats in isolation. Regional and global cooperation frameworks therefore play a crucial role in enabling information sharing, joint operational responses, and policy harmonisation.

At the same time, these mechanisms also introduce governance challenges, particularly related to differences in legal systems, sovereignty concerns, data protection standards, and varying interpretations of security priorities. These tensions highlight the complexity of balancing effective cooperation with respect for national autonomy and fundamental rights, which remains a persistent governance dilemma in the PVE field.

Overall, the analysis supports the conclusion that resilient and sustainable societies are best built through integrated governance systems that combine security, development, and social policy approaches. The effectiveness of PVE strategies depends not only on institutional capacity but also on the quality of relationships between institutions and society, levels of trust, and the inclusiveness of governance structures. Fragmented approaches that operate in isolation tend to produce reactive and short-term responses, whereas integrated models foster adaptability, coherence, and long-term resilience.

This reinforces the need for systemic thinking in prevention policy, where different governance levels are not seen as separate domains but as interconnected components of a broader resilience ecosystem.

This integrated model also has broader implications beyond counter-terrorism policy. It reflects a wider shift in governance theory towards networked, participatory, and multi-level systems of problem-solving in increasingly complex security environments. In this sense, preventing violent extremism is not only a security objective but also a governance and societal challenge that requires long-term investment in institutional legitimacy, social inclusion, and civic trust.

Ultimately, building resilience against violent extremism should be understood as a continuous societal process that depends on sustained cooperation, shared responsibility, and the strengthening of democratic and inclusive governance structures.

5. CONCLUSION

This paper has examined how multi-level governance approaches contribute to the prevention of violent extremism (PVE) and the broader development of resilient and sustainable societies. By integrating institutional, community, and international dimensions of security governance, the study demonstrates that violent extremism is best understood as a complex, adaptive, and context-dependent phenomenon that cannot be effectively addressed through security-centric, linear, or fragmented policy responses. Instead, prevention requires coordinated, sustained, and system-wide engagement across multiple governance levels, involving both state institutions and a wide range of non-state actors, including civil society organisations, local communities, and international partners.

One of the central conclusions of the analysis is that institutional coordination and early prevention mechanisms are necessary but not sufficient conditions for effective PVE. While law enforcement agencies, intelligence services, and judicial institutions remain essential actors in identifying threats and responding to security risks, their effectiveness is significantly enhanced when they operate in structured cooperation with education systems, social welfare institutions, healthcare services, local authorities, and civil society organisations. This confirms that prevention is most effective when it is embedded within a broader governance ecosystem that extends beyond traditional security structures and incorporates social, developmental, and community-based dimensions of intervention.

The study further highlights that early prevention frameworks are particularly dependent on trust, legitimacy, and information-sharing between institutions and communities. Without these relational foundations, early warning systems and referral mechanisms are likely to remain underutilised or ineffective. This underscores that PVE is not only a technical or institutional challenge, but also a deeply social and political one, shaped by perceptions of state legitimacy, historical experiences, and the quality of state–society relations. Strengthening transparency, accountability, and rights-based governance therefore becomes essential for ensuring that preventive mechanisms function effectively and are perceived as legitimate by the communities they are designed to protect.

Another key conclusion is the critical importance of community resilience as a protective factor against radicalisation and violent extremism. Resilient communities are characterised by strong social cohesion, inclusive identities, active civic participation, and robust interpersonal and institutional trust. These elements reduce the appeal of extremist narratives by limiting the social and psychological vulnerabilities that such narratives tend to exploit. Civil society organisations play a particularly important intermediary role in this regard, as they often possess greater access, credibility, and cultural sensitivity within local contexts. They contribute not only to early detection and prevention, but also to long-term

processes of social inclusion, dialogue facilitation, and capacity-building. However, their effectiveness remains contingent on stable funding, institutional recognition, and the protection of their independence from excessive securitisation or political instrumentalisation.

At the same time, the analysis confirms that regional and international cooperation remains indispensable in addressing the transnational dimensions of violent extremism. The mobility of extremist networks, the rapid spread of online radicalisation content, and the cross-border nature of financial and logistical support structures require coordinated responses that extend beyond national jurisdictions. Mechanisms for intelligence sharing, joint investigations, and policy coordination significantly enhance the ability of states and institutions to respond effectively to evolving threats. Nevertheless, such cooperation is not without challenges, as differences in legal frameworks, institutional capacities, political priorities, and data protection standards continue to limit full harmonisation. These tensions illustrate the ongoing governance dilemma between the need for effective security cooperation and the requirement to safeguard sovereignty, privacy, and fundamental rights.

Overall, the findings of this paper confirm that resilient and sustainable societies are built through integrated governance systems that connect security, development, and social policy dimensions into a coherent and mutually reinforcing framework. The effectiveness of PVE strategies depends not only on institutional capacity, but also on the quality of relationships between institutions and society, the level of trust in governance structures, and the inclusiveness of decision-making processes. Fragmented approaches that operate in isolation tend to produce reactive, short-term, and often limited outcomes, whereas integrated governance models enable more adaptive, preventive, and sustainable responses to complex security challenges.

Finally, the study contributes to ongoing academic and policy debates by reinforcing the argument that preventing violent extremism should be conceptualised as a long-term societal process rather than a narrow security intervention. It requires continuous investment in institutional cooperation, social cohesion, and democratic governance. Future policy efforts should therefore prioritise strengthening multi-level coordination mechanisms, enhancing community-based prevention capacities, and ensuring that all PVE interventions remain firmly grounded in human rights standards, rule of law principles, and democratic legitimacy. In this sense, the prevention of violent extremism should be understood not only as a security objective, but also as a broader societal commitment to resilience, inclusion, and sustainable peace.

6. REFERENCES

- Allan, H., Glazzard, A., Jespersen, S., Reddy-Tumu, S., & Winterbotham, E. (2015). *Drivers of violent extremism: Hypotheses and literature review*. Royal United Services Institute (RUSI). <https://www.gov.uk/research-for-development-outputs/drivers-of-violent-extremism-hypotheses-and-literature-review>
- Arifi, B. (2019). *Drivers to violent extremism in South Eastern Europe – The North Macedonian context*. *Journal of Penal Law and Criminology*, 7(1), 27–51.
- Baumann, Z. (2025). *Community-engagement in P/CVE: In the interest of communities?* SSRN Electronic Journal. <https://doi.org/10.2139/ssrn.5324373>
- Coolsaet, R. (2010). EU counterterrorism strategy: Value added or chimera? *International Affairs*, 86(4), 857–873. <https://doi.org/10.1111/j.1468-2346.2010.00916.x>
- Dalgaard-Nielsen, A., & Schack, P. (2016). Community resilience to militant Islamism: Who and what? An explorative study of resilience in three Danish communities. *Democracy and Security*, 12(4), 309–327. <https://doi.org/10.1080/17419166.2016.1236691>
- Ellis, B. H., & Abdi, S. (2017). Building community resilience to violent extremism through genuine partnerships. *American Psychologist*, 72(3), 289–300. <https://doi.org/10.1037/amp0000065>
- Ernstorfer, A. (2018). *Effective approaches to preventing violent extremism: A peacebuilding systems perspective*. Berghof Foundation. <https://berghof-foundation.org/library/effective-approaches-to-preventing-violent-extremism-a-peacebuilding-systems-perspective>
- European Commission. (2020a). *A Counter-Terrorism Agenda for the EU*.
- European Commission. (2020b). *EU Security Union Strategy*.
- European Union Agency for Law Enforcement Cooperation (Europol). (2024). *European Union terrorism situation and trend report 2024 (TE-SAT)*. Publications Office of the European Union.
- Gómez, Á., Martínez, M., Martel, F. A., López-Rodríguez, L., Vázquez, A., Chinchilla, J., Paredes, B., Hettiarachchi, M., Hamid, N., & Swann, W. B. (2021). Why people enter and embrace violent groups. *Frontiers in Psychology*, 11, Article 614657. <https://doi.org/10.3389/fpsyg.2020.614657>
- Horgan, J. (2008). From profiles to pathways and roots to routes: Perspectives from psychology on radicalization into terrorism. *The ANNALS of the American Academy of Political and Social Science*, 618(1), 80–94. <https://doi.org/10.1177/0002716208317539>
- Khalil, J., & Zeuthen, M. (2016). *Countering violent extremism and risk reduction: A guide to programme design and evaluation*. Royal United Services Institute (RUSI). <https://www.rusi.org>
- Kruglanski, A. W., Gelfand, M. J., Bélanger, J. J., Sheveland, A., Hetiarachchi, M., & Gunaratna, R. (2014). *The psychology of radicalization and deradicalization: How significance quest impacts violent extremism*. *Political Psychology*, 35(S1), 69–93. <https://doi.org/10.1111/pops.12163>
- Kundnani, A., & Hayes, B. (2018). *The globalisation of countering violent extremism policies: Undermining human rights, instrumentalising civil society*.

- Transnational Institute. <https://www.tni.org/en/publication/the-globalisation-of-countering-violent-extremism-policies>
- Lowe, D. (2017). Prevent strategies: The problems associated in defining extremism: The case of the United Kingdom. *Studies in Conflict & Terrorism*, 40(11), 917–933. <https://doi.org/10.1080/1057610X.2016.1253941>
 - Macchiarini Crosson, D., Stankovic, T., Rieker, P., & Blockmans, S. (2023). *The EU and CT-P/CVE in its external action: Triangulating evidence to put the democracy versus security dilemma to the test*. Centre for European Policy Studies (CEPS). <https://www.ceps.eu/ceps-publications/the-eu-and-ct-p-cve-in-its-external-action/>
 - McCauley, C., & Moskalkenko, S. (2017). Understanding political radicalization: The two-pyramids model. *American Psychologist*, 72(3), 205–216. <https://doi.org/10.1037/amp0000062>
 - McNeil-Willson, R., Gerrand, V., Scrinzi, F., & Triandafyllidou, A. (2019). *Polarisation, violent extremism and resilience in Europe today: An analytical framework* (BRaVE Project, Deliverable D2.1). European University Institute, Robert Schuman Centre for Advanced Studies. <https://cadmus.eui.eu/handle/1814/65664>
 - Mudde, C. (2007). *Populist radical right parties in Europe*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511492037>
 - Neumann, P. R. (2013). The trouble with radicalization. *International Affairs*, 89(4), 873–893. <https://doi.org/10.1111/1468-2346.12049>
 - OSCE - Organization for Security and Co-operation in Europe. (2019). *Understanding referral mechanisms in preventing and countering violent extremism and radicalization that lead to terrorism: Navigating challenges and protecting human rights: A guidebook for South-Eastern Europe*. OSCE Secretariat. <https://www.osce.org/files/f/documents/7/4/418274.pdf>
 - OSCE - Organization for Security and Co-operation in Europe. (2020). *A whole-of-society approach to preventing and countering violent extremism and radicalization that lead to terrorism: A guidebook for Central Asia*. OSCE Secretariat. https://www.osce.org/files/f/documents/a/7/444340_0.pdf
 - Prislán, K., Borovec, K., & Cajner Mraović, I. (2020). The role of civil society and communities in countering violent extremism and radicalisation. *Policija i sigurnost*, 29(3), 223–245. <https://hrcak.srce.hr/243827>
 - Radicalisation Awareness Network. (2021a). *RAN manual: Responses to returnees: Foreign terrorist fighters and their families*. European Commission. https://home-affairs.ec.europa.eu/networks/radicalisation-awareness-network-ran_en
 - Radicalisation Awareness Network. (2021b). *RAN manual: Responses to violent extremism – Community engagement and prevention*. European Commission, Migration and Home Affairs. https://home-affairs.ec.europa.eu/networks/radicalisation-awareness-network-ran_en
 - Ragazzi, F. (2017). Countering terrorism and radicalisation in Europe. In A. Salter & M. Mutlu (Eds.), *Researching terrorism and counter-terrorism*. Routledge.
 - Ranstorp, M. (Ed.). (2016). *Understanding violent radicalisation: Terrorist and jihadist movements in Europe*. Routledge.
 - Schmid, A. P. (2013). *Radicalisation, de-radicalisation, counter-radicalisation: A conceptual discussion and literature review*. International Centre for Counter-Terrorism – The Hague. <https://doi.org/10.19165/2013.1.02>

- Sedgwick, M. (2010). The concept of radicalization as a source of confusion. *Terrorism and Political Violence*, 22(4), 479–494. <https://doi.org/10.1080/09546553.2010.491009>
- Stephens, W., & Sieckelinck, S. (2020). *Being resilient to radicalisation in PVE policy: A critical examination*. *Critical Studies on Terrorism*, 13(1), 142–165. <https://doi.org/10.1080/17539153.2019.1658415>
- Stephens, W., Sieckelinck, S., & Boutellier, H. (2019). *Preventing violent extremism: A review of the literature*. *Studies in Conflict & Terrorism*, 42(9), 873–891. <https://doi.org/10.1080/1057610X.2018.1543144>
- UNDP - United Nations Development Programme. (2017). *Journey to extremism in Africa: Drivers, incentives and the tipping point for recruitment*. UNDP Regional Bureau for Africa. <https://www.undp.org/publications/journey-extremism-africa>
- UNDP - United Nations Development Programme. (2016). *Preventing violent extremism through inclusive development and the promotion of tolerance and respect for diversity: A development response to addressing radicalisation and violent extremism*. UNDP. <https://www.undp.org/publications/preventing-violent-extremism-through-inclusive-development-and-promotion-tolerance-and-respect-diversity>
- United Nations General Assembly. (2015). *Plan of action to prevent violent extremism: Report of the Secretary-General (A/70/674)*. United Nations. <https://www.un.org/counterterrorism/en/plan-of-action-to-prevent-violent-extremism>
- United Nations General Assembly. (2006). *The United Nations Global Counter-Terrorism Strategy (A/RES/60/288)*. United Nations. <https://www.un.org/counterterrorism/en/un-global-counter-terrorism-strategy>

BATTERED WOMAN SYNDROME AND SPOUSAL HOMICIDE: A QUALITATIVE STUDY OF INCARCERATED WOMEN WITH HISTORIES OF INTIMATE PARTNER VIOLENCE

Dragana Batic

Faculty of Security – Skopje, UKLO

dragana.batic@uklo.edu.mk

orcid.org/0009-0002-1632-431X

Dimitrinka Jordanova Peshevska

University American College Skopje, Skopje, Republic of North Macedonia

orcid.org/0000-0002-3499-3667

Abstract

Intimate partner violence is a widespread form of interpersonal violence that encompasses physical, psychological, control of behavior, and sexual violence perpetrated by one intimate partner against another in an effort to establish power and control. Women constitute the majority of victims of severe intimate partner violence, and prolonged exposure to this type of violence may result in significant negative psychological and health outcomes including post-traumatic stress disorder (PTSD). The theoretical concept of battered woman syndrome (BWS), introduced by Walker, describes a pattern of psychological and behavioral responses observed among women subjected to chronic intimate partner violence and will be used to analyze the study findings. This paper aims to examine the relationship between prolonged intimate partner violence, battered woman syndrome, and spousal homicide, providing a theoretical overview and focusing on symptoms such as depression, low self-esteem, learned helplessness, and trauma-related responses that may influence victims' perceptions of threat and decision-making processes. The second part presents a qualitative study based on in-depth semi-structured interviews conducted with four women serving prison sentences for the homicide of a husband or intimate partner following prolonged experiences of intimate partner violence. Interview data were analyzed using thematic analysis to identify recurring patterns in participants' experiences of abuse, psychological responses, coping strategies, and pathways leading to the homicidal act. The findings highlight the cumulative psychological impact of long-term victimization and reveal common themes of fear, helplessness, social isolation, and perceived lack of alternatives prior to the homicide. The study contributes to a deeper understanding of the experiences of abused women who commit spousal homicide and underscores the importance of trauma-informed approaches within legal, mental health, and criminal justice systems.

Keywords: *battered woman syndrome; intimate partner violence; PTSD; spousal homicide; qualitative research.*

1. INTRODUCTION

Domestic violence, including the homicide of an intimate partner, is not a new phenomenon. Quite the opposite, historical records and statistical data show that many violent men kill their wives, while women, although far less frequently, kill their husbands or intimate partners. Unlike men, women are generally considered to commit homicide in order to protect themselves from being killed by the abusers with whom they live or maintain an intimate relationship.

Domestic violence encompasses psychological, physical, and sexual abuse, as well as economic abuse perpetrated against an intimate partner with the aim of gaining and maintaining power and control within the relationship (Sethi & Jordanova-Peshevska, 2014). Coercion, manipulation, isolation, and minimization are commonly used to instill fear in the victim, forcing compliance. Domestic violence is recognized as a global problem that constitutes a serious violation of fundamental human rights. In recent years, international organizations such as the United Nations (UN), UNICEF, and the Council of Europe have developed strategic frameworks aimed at preventing and combating domestic violence. These organizations share a common understanding of domestic violence as socially unacceptable and criminal behavior, advocate for a multisectoral approach to its prevention and response, and emphasize the need to develop intervention programs for perpetrators while continuously improving support services for victims.

The latest global estimates published by the World Health Organization (WHO) indicate that violence against women remains a major public health and societal concern. According to the WHO (2025), approximately 31.6% of women worldwide, or about 840 million women, have experienced physical and/or sexual violence by an intimate partner or sexual violence by a non-partner during their lifetime. Furthermore, around 27% of women aged 15–49 years who have been in an intimate relationship reported experiencing physical and/or sexual violence by an intimate partner (WHO, 2015; Sethi & Jordanova-Peshevska, 2014).

Available research also indicates that domestic violence is a widespread phenomenon in North Macedonia. According to national and international studies, approximately one-third of women have experienced some form of gender-based violence during their lifetime, while official data for 2025 recorded 4,745 female victims of domestic violence, indicating a persistently high level of victimization. One of the most comprehensive studies on the prevalence of violence against women in North Macedonia was conducted by the Organization for Security and Co-operation in Europe (OSCE) in 2018. According to the findings, 45% of women reported having experienced psychological, physical, or sexual violence by an intimate partner during their lifetime, while 14% reported experiencing physical and/or sexual violence after the age of 15. Notably, only 2% of women reported the most serious incident to the police (OSCE, 2019).

2. THEORETICAL FRAMEWORK

2.1. Battered Woman Syndrome

Battered Woman Syndrome (BWS) was first introduced by Walker (1979) to describe a cluster of psychological and behavioral symptoms resulting from prolonged exposure to violence perpetrated by an intimate partner. BWS is widely recognized as a subcategory of Post-Traumatic Stress Disorder (PTSD), as many abused women exhibit

symptoms consistent with trauma, including hypervigilance, anxiety, depression, and learned helplessness (American Psychiatric Association, 2022; Walker, 2009).

Based on her research involving 400 battered women, Walker (1979) developed the theory of learned helplessness and the cycle of violence theory. These theories are of particular importance in explaining the behavior of women who have been continuously subjected to physical and psychological abuse by their intimate partners, especially in cases where the abuse ultimately resulted in the woman killing her abuser.

Walker defined a battered woman as:

"A battered woman is a woman who has been repeatedly subjected to any forceful physical or psychological behavior by a man in order to coerce her into doing something he wants without regard for her rights. To be classified as a battered woman, she must have experienced the battering cycle at least twice" (Walker, 1979, p. 15).

Since Walker first described Battered Woman Syndrome (BWS) in 1979 and further elaborated on the concept in 1984, it has been widely used as the basis for legal defenses in cases where abused women have killed their abusive partners. Drawing on in-depth interviews with 400 abused women, Walker formulated the theories of learned helplessness, originally proposed by Seligman (1975), and the cycle of violence. Although these theories have been subject to criticism, they remain highly influential in explaining the behavior of women who experience prolonged intimate partner violence, particularly in cases where the abuse culminates in the homicide of the abusive partner by the victim (Walker, 1984; Walker, 2009; Dutton, 1993).

2.2. Learned Helplessness Theory

Learned Helplessness Theory, developed by Martin E. P. Seligman, is based on the premise that individuals come to believe they have little or no control over events affecting their lives following repeated exposure to uncontrollable adverse experiences (Seligman, 1975). The theory has been widely applied to explain why many women remain in abusive relationships, suggesting that prolonged victimization may create a state of psychological paralysis in which victims perceive escape as impossible despite opportunities to leave (Walker, 1979; Maier & Seligman, 1976). Walker (1979) applied Seligman's theory of learned helplessness to the experiences of battered women. The original theory emerged from experimental research in which dogs exposed to unavoidable electric shocks eventually failed to escape, even when escape later became possible. Following repeated exposure to the shocks, the animals ceased attempting to avoid the situation, demonstrating what Seligman termed *learned helplessness*. The theory was subsequently reformulated and extended to explain depression and other psychological responses in humans (Abramson et al., 1978).

Using this theoretical framework, Walker (1979) argued that women who experience repeated physical and psychological abuse may gradually develop the belief that they are powerless to change their circumstances. As a result, many victims remain in abusive relationships because they perceive that they have little or no control over their partner's violent behavior or over the course of their own lives. This perceived lack of control often prevents them from considering or initiating actions to escape the abuse. Walker further argued that battered women not only perceive themselves as unable to control the violence inflicted upon them but also come to believe that they have little control over other aspects of their lives.

Furthermore, Walker (1979) emphasized that traditional gender-role socialization encourages women to preserve their marriages despite abuse, often placing the

responsibility for maintaining family stability on women. Consequently, many victims conceal the abuse in an effort to maintain the appearance of a successful marriage and to avoid social stigma. Research has demonstrated that additional factors, including economic dependence, fear of retaliation, concern for children, emotional attachment to the abuser, and limited access to social support, also contribute to women's decisions to remain in abusive relationships (Barnett, 2000; Gondolf & Fisher, 1988).

Although Learned Helplessness Theory has been criticized for portraying abused women as passive victims and for overlooking the active survival strategies many employ, it remains a highly influential framework in understanding the psychological consequences of prolonged intimate partner violence (Dutton, 1993). More recent trauma-informed perspectives emphasize that women exposed to chronic abuse often display remarkable resilience while simultaneously experiencing symptoms consistent with Post-Traumatic Stress Disorder (PTSD) and complex trauma (Herman, 1997; Walker, 2009). Consequently, Learned Helplessness Theory continues to provide an important foundation for understanding the behavioral and psychological responses of women subjected to prolonged intimate partner violence.

2.3. Cycle of Violence Theory

The Cycle of Violence Theory, developed by Walker (1979), describes the recurring pattern of abusive behavior within intimate partner relationships. According to Walker, abusive relationships typically progress through three distinct phases: the Tension-Building Phase, the Acute Battering Incident (also referred to as the Acute Explosion Phase), and the Honeymoon Phase. Walker proposed that these recurring stages help explain why many women remain in abusive relationships despite repeated episodes of violence. During the Tension-Building Phase, the abuser engages in increasing verbal, emotional, and psychological abuse, often accompanied by relatively minor acts of physical aggression, such as pushing or slapping (Walker, 1979). In response, the victim frequently attempts to prevent further violence by placating the abuser, complying with his demands, or modifying her own behavior in an effort to reduce conflict. Although these coping strategies are intended to avoid escalation, they often reinforce the abuser's controlling behavior and fail to prevent the progression to more severe violence (Walker, 1979; Dutton, 1993).

The second stage, the Acute Battering Incident, is characterized by severe physical violence, loss of control by the perpetrator, and intense fear experienced by the victim (Walker, 1979). This phase often results in serious physical injuries and profound psychological trauma. Victims commonly report feelings of helplessness, terror, and a heightened perception of imminent danger (Herman, 1997). Repeated exposure to this phase contributes to the development of trauma-related symptoms, including anxiety, depression, hypervigilance, and symptoms associated with Post-Traumatic Stress Disorder (PTSD) (American Psychiatric Association, 2022; Walker, 2009).

Following the violent episode, the relationship typically enters the Honeymoon Phase, during which the abuser expresses remorse, apologizes for the violence, promises that the abuse will never occur again, and often demonstrates affection through gifts or unusually caring behavior (Walker, 1979). This period of reconciliation temporarily restores the victim's hope that the perpetrator has genuinely changed, strengthening the emotional bond between the partners and making it increasingly difficult for the victim to leave the relationship. Research has shown that this intermittent reinforcement of abuse and affection

contributes to strong emotional attachment and trauma bonding between victims and perpetrators (Dutton & Painter, 1993).

Walker (1979) argued that the cycle repeats with increasing frequency and severity over time, while the Honeymoon Phase often becomes progressively shorter or disappears altogether. As a result, victims may become increasingly isolated, emotionally dependent, and psychologically exhausted. In many cases, it takes years before women recognize that meaningful and lasting change is unlikely to occur (Walker, 1984). Walker observed that the reconciliation phase is one of the most psychologically influential stages because it reinforces emotional dependence within the relationship. As she explained, "*Neither partner experiences themselves as a separate individual capable of living independently, and both may come to believe that death is the only means of separation*" (Walker, 1979, p. 55).

Although the Cycle of Violence Theory has become one of the most influential models for understanding intimate partner violence, contemporary researchers have noted that not all abusive relationships follow this exact sequence. Many abusive relationships do not include a Honeymoon Phase, while others involve continuous coercive control without clearly identifiable cycles (Stark, 2007). Nevertheless, Walker's model remains a foundational framework for understanding the dynamics of abusive relationships and has had a significant influence on psychological research, legal practice, and the development of interventions for survivors of intimate partner violence (Walker, 2009).

2.4. Post-Traumatic Stress Disorder

Although Battered Woman Syndrome (BWS) is not recognized as an independent mental disorder in the *Diagnostic and Statistical Manual of Mental Disorders* (5th ed., text rev.; DSM-5-TR), it is widely conceptualized as a trauma-related condition that shares many diagnostic features with Post-Traumatic Stress Disorder (PTSD) (American Psychiatric Association [APA], 2022; Walker, 2009). Walker originally proposed that the psychological consequences of prolonged intimate partner violence closely resemble PTSD, while also encompassing additional cognitive, emotional, and interpersonal disturbances that extend beyond the traditional PTSD framework (Walker, 1984, 2009).

PTSD is a trauma- and stressor-related disorder that develops following exposure to actual or threatened death, serious injury, or sexual violence (APA, 2022). Such exposure may occur through direct personal experience, witnessing traumatic events, learning that a traumatic event occurred to a close family member or friend, or repeated exposure to aversive details of traumatic experiences. The disorder is characterized by four major symptom clusters: intrusive symptoms, persistent avoidance of trauma-related stimuli, negative alterations in cognition and mood, and marked alterations in arousal and reactivity (APA, 2022).

Within the context of intimate partner violence, traumatic experiences frequently include repeated physical assaults, threats of death or serious bodily harm, sexual violence, psychological abuse, coercive control, threats directed toward children or other family members, and persistent fear for one's personal safety (Herman, 1997; Walker, 2009). Continuous exposure to these traumatic events often disrupts victims' fundamental beliefs regarding safety, trust, control, and self-worth, resulting in profound psychological distress and long-term impairment in daily functioning.

Victims of prolonged domestic violence commonly experience intrusive recollections, recurrent nightmares, flashbacks, and intense psychological or physiological distress when confronted with reminders of the abuse (APA, 2022). Many survivors also demonstrate avoidance behaviors by deliberately avoiding thoughts, conversations, places, or

individuals associated with the traumatic experiences. Emotional numbing, diminished interest in previously meaningful activities, feelings of detachment from others, persistent negative emotional states, and difficulty experiencing positive emotions are also frequently reported (Herman, 1997). Furthermore, chronic hyperarousal may manifest as sleep disturbances, irritability, angry outbursts, difficulty concentrating, exaggerated startle responses, and persistent hypervigilance, reflecting a continuous perception of danger even after the abuse has ended (Walker, 2009).

Walker (2009) argued that the experiences of battered women often extend beyond the diagnostic boundaries of PTSD. In addition to the traditional PTSD symptom clusters, many survivors exhibit disturbances in emotional regulation, interpersonal relationships, self-identity, body image, sexuality, and affect regulation. Walker also identified chronic feelings of shame, guilt, worthlessness, social withdrawal, dissociative symptoms, impaired interpersonal functioning, and an enduring perception of threat as common psychological consequences of prolonged abuse. These findings are consistent with contemporary trauma literature, which recognizes that survivors of prolonged interpersonal violence frequently experience complex and multifaceted trauma reactions that are not fully captured by the traditional PTSD diagnosis (Herman, 1997; Dutton, 1993).

Contemporary research therefore supports the view that Battered Woman Syndrome represents a specific pattern of trauma responses associated with chronic intimate partner violence rather than a separate psychiatric disorder. Although the term has become less common in clinical practice, its theoretical foundations continue to influence psychological assessment, forensic evaluations, and legal proceedings involving survivors of domestic violence (Walker, 2009; American Psychiatric Association, 2022).

2.5. Women Who Kill a Man as a Response to Violence

The relationship between violence against women and the killing of an abusive male partner has received relatively limited scholarly attention, with most studies focusing specifically on cases involving the homicide of abusive intimate partners. Wolfgang observed that men are considerably more likely than women to initiate physical violence, thereby creating circumstances that may ultimately lead to their death at the hands of their female partners (Wolfgang, 1958). McCormick found that, among 132 women imprisoned for the murder of their husbands, 40% had been victims of chronic physical abuse within the marriage, and that the violence they had endured constituted the primary precipitating factor for the homicide (McCormick, 1976).

Similarly, a study conducted in 1977 at a women's correctional facility in California reported that 29 of the 30 women who had killed their husbands had previously experienced abuse. Furthermore, 20 participants stated that homicide occurred while they were attempting to protect themselves or their children from further violence (Totman, 1978).

More recent research has reached similar conclusions. In her study, Caitlin Nash found that the majority of women who killed their intimate partners had been subjected to prolonged physical, psychological, and sexual abuse (Nash, 2021). In a substantial proportion of cases, the abusive relationship was characterized by a pattern of coercive control. Coercive control refers to an ongoing pattern of domination through which the perpetrator systematically restricts the victim's autonomy, freedom, and independence by employing psychological, emotional, economic, and other forms of abuse (Stark, 2007). Unlike the traditional understanding of domestic violence as a series of isolated physical assaults, the concept of coercive control emphasizes that abuse constitutes a persistent system of domination rather than a collection of discrete violent incidents (Stark, 2007).

The Linklaters study, which analyzed cases from several countries, including the United States, Australia, Spain, Japan, and Brazil, found that many women who killed their partners had been subjected to years of domestic abuse. The study further concluded that criminal justice systems continue to inadequately recognize the effects of trauma and coercive control when assessing these cases. The authors therefore argue that legislative reforms are necessary to promote a more comprehensive understanding of victims' lived experiences within criminal proceedings (Linklaters LLP, 2023).

Comparable findings were reported by Liu, whose research involving women in China demonstrated that homicide most frequently represents the culmination of escalating and chronic abuse. The study also found that women who killed abusive partners commonly exhibited symptoms of trauma and Post-Traumatic Stress Disorder (PTSD), emphasizing that a history of domestic violence should be regarded as a crucial contextual factor in judicial assessment and criminal adjudication (Liu, 2024).

The first study conducted in Serbia to highlight the significance of domestic violence as a criminogenic factor in homicides committed by women was carried out by Konstantinović-Vilić, who examined 134 women convicted of homicide. The findings revealed that family relationships were severely disrupted in 75.38% of the cases. In most instances, the husband suffered from alcohol dependence, was frequently absent from the family home, and engaged in physical violence, threats, and extreme jealousy. In 32.83% of the cases, the primary motive for the homicide was prolonged abuse and physical violence perpetrated by the husband. The study also identified abuse by fathers-in-law and mothers-in-law as a significant contributing factor in a considerable number of cases (Konstantinović-Vilić, 1986).

The relationship between violence perpetrated against women by intimate partners and other adult male family members and subsequent criminal behavior was further examined by Nikolić-Ristanović, who devoted particular attention to women convicted of killing their husbands. She identified seven pathways through which male domestic violence may contribute to women's involvement in crime: (1) fear for their own lives or the lives of their children; (2) psychological instability resulting from prolonged abuse; (3) coercion into committing criminal offenses; (4) a diminished tolerance threshold following repeated exposure to violence; (5) symbolic retaliation for previous abuse through identifying the perpetrator as the source of harm; (6) diminished self-esteem; and (7) protecting or concealing abusive family members. According to Nikolić-Ristanović, these pathways are shaped by a combination of early childhood experiences, the development of cultural and ethnic identities, broader social circumstances such as economic crises and war, exposure to violence during childhood, victimization within marriage, and the responses of society and state institutions to domestic violence (Nikolić-Ristanović, 2000).

3. RESEARCH METHODOLOGY: RESEARCH SUBJECT AND OBJECTIVES

The subject of this study is the examination of the relationship between the psychological condition of women who have killed their intimate partners and the homicide they committed. The research was conducted at the **Idrizovo Correctional Institution**, the only women's correctional facility in the Republic of North Macedonia, using a purposive sample of four women convicted of murdering their intimate partners.

Data were collected through **semi-structured interviews**. In addition to gathering information on the participants' socio-demographic characteristics, the interviews explored

their physical and mental health, childhood experiences, histories of violence and victimization, relationships with the deceased partners, the circumstances surrounding the homicide, and their experiences during imprisonment.

The study was conducted in accordance with internationally recognized ethical standards for research involving human participants. Written informed consent was obtained from all participants, and strict procedures were implemented to ensure confidentiality, anonymity, and the secure handling of research data.

The research employed **qualitative research design**, using **thematic analysis** to examine the interviews conducted with the four participants. This approach enabled an in-depth exploration of the subjective experiences of this specific and difficult-to-access population, while also facilitating the identification of common themes and patterns related to prolonged intimate partner violence, the circumstances preceding the homicide, and the participants' life histories. A separate **case study** was prepared for each participant. Given the small, purposively selected sample, the findings cannot be statistically generalized to the broader population of women who have killed their intimate partners. Nevertheless, they provide valuable insights into the understanding of this phenomenon, particularly within the socio-cultural, economic, and legal context of the Republic of North Macedonia. Furthermore, the findings may serve as a foundation for future research in the fields of **intimate partner violence, victimology, women's offending, and criminal justice**.

The primary objective of the study was to determine whether an association exists between the psychological condition of women and the violence they had experienced, and to examine the extent to which prolonged exposure to intimate partner violence influenced the commission of homicide.

4. RESULTS OF THE QUALITATIVE ANALYSIS

The thematic analysis revealed that all four cases were characterized by **continuous and multidimensional intimate partner violence**. The participants had been subjected to prolonged, chronic, and escalating abuse encompassing multiple forms of victimization, including physical, psychological, sexual, and, in one case, economic abuse. Physical violence was frequent and severe, while psychological abuse manifested through verbal insults, humiliation, degradation, intimidation, and threats. Two participants reported severe sexual abuse, whereas one participant experienced economic abuse. These findings confirm that the violence was not an isolated incident but rather a continuous process of victimization.

A consistent pattern of coercive control was identified across all four cases. The perpetrators established and maintained dominance through death threats, restrictions on the women's freedom of movement, social isolation from family members, extreme jealousy, economic dependence, and the continuous creation of fear. The participants remained in abusive relationships not because of free choice, but because of persistent coercive control and fear of the consequences of leaving their partners.

Another prominent theme concerned the influence of patriarchal values and traditional gender norms. All participants expressed beliefs that marriage should be preserved at any cost, that divorce would bring shame upon the family, that women were expected to endure abuse in silence, and that the interests of their children should take precedence over their own personal safety. These deeply rooted cultural norms contributed to the women's prolonged stay in abusive relationships and delayed their help-seeking behavior.

A pervasive sense of helplessness and entrapment, consistent with the concept of learned helplessness, was evident in all four cases. The women had repeatedly attempted to escape, report the abuse, or terminate the relationship. However, persistent threats, economic dependence, lack of social support, and fear for their own lives and the lives of their children forced them to return to their abusive partners. Over time, they gradually developed the belief that they had no control over their lives and that no viable alternative to the abusive relationship existed.

In most cases, institutional support proved to be either limited or ineffective. Although several participants had reported the abuse and sought assistance from the police and social welfare services, the violence continued. In addition, some women received little or no support from their families of origin, further increasing their vulnerability and reinforcing their dependence on the abusive relationship.

The prolonged exposure to violence resulted in severe psychological consequences, including depression, anxiety, chronic fear, low self-esteem, emotional instability, dissociative symptoms, and, in one case, suicide attempts. Two participants exhibited dissociative symptoms characterized by an inability to fully recall the homicide, describing their experiences with statements such as: *"I don't know how I struck him," "It was as if everything suddenly went dark before my eyes,"* and *"It felt as though there was a large gap in my memory."*

Although the specific circumstances surrounding the homicides differed, all four cases indicate that the killing of the abusive partner represented the culmination of prolonged and severe violence. For some participants, the homicide occurred during an episode of immediate physical assault, whereas in other cases the abusive partner was asleep or did not pose an immediate physical threat at the time of the incident. Nevertheless, all participants reported a subjective perception that their own lives, the lives of their children, or their future safety were in imminent danger and that they no longer perceived any realistic means of escaping the abusive relationship other than through lethal violence.

5. CONCLUSION

The analyzed case studies illustrate the concept of Battered Woman Syndrome and provide a deeper understanding of the complex relationship between prolonged intimate partner violence, the psychological condition of women, and the commission of homicide. In all four cases, the presence of long-term and continuous violence was established, in some instances lasting for decades. The violence was not limited to physical abuse but also included psychological abuse and, in two cases, sexual violence. In addition, all participants exhibited elements of rigid patriarchal socialization, the presence of alcohol abuse among partners, death threats, pronounced fear, low self-esteem, and feelings of helplessness. These findings are consistent with Walker's (1979) theory of Battered Woman Syndrome, which proposes that prolonged exposure to violence leads to learned helplessness, fear conditioning, and impaired coping responses. Similarly, Seligman's (1975) theory of learned helplessness explains why victims may remain in abusive relationships despite opportunities to escape, due to the belief that their actions cannot change their situation.

In all cases, participants exhibited indicators of patriarchal socialization, alcohol abuse by partners, death threats, chronic fear, low self-esteem, and emotional dependence. Browne (1987) similarly found that women who kill abusive partners are typically subjected to severe, escalating, and long-term violence. Campbell (2003) further

emphasizes that intimate partner homicide is strongly associated with escalating threats, weapon use, and coercive dynamics.

Although some women sought institutional assistance, including shelters, police intervention, and social services, they ultimately returned to their abusive partners. This reflects the combined influence of economic dependency, fear of retaliation, patriarchal norms, and insufficient institutional protection. The World Health Organization (2013) similarly reports that structural and systemic barriers often prevent effective escape from violent relationships.

Across all cases, there was a cumulative buildup of fear, trauma, and psychological distress, which ultimately resulted in violent reactions directed toward the abuser. None of the women had prior criminal histories or patterns of violent behavior. This finding supports Stark's (2007) theory of coercive control, which describes how ongoing domination gradually erodes autonomy and creates a persistent sense of entrapment.

In several cases, dissociative reactions, partial amnesia, and psychological instability were observed, suggesting symptoms consistent with post-traumatic stress disorder (PTSD). According to the American Psychiatric Association (2013), PTSD may develop following prolonged exposure to traumatic events, including domestic violence. Walker (1979) further conceptualizes BWS as a specific trauma response within the PTSD spectrum.

The findings indicate that the most common factors associated with the killing of an abusive partner include chronic violence, escalation of abuse, alcohol misuse, death threats, weapon use, fear for personal safety and children's lives, and inadequate institutional protection. These findings are consistent with McCormick and Totman (1998), who emphasize the cumulative nature of violence in intimate partner homicide cases, as well as Campbell (2003), who identifies similar risk patterns.

Although in many legal systems, particularly in the United States, Australia, and the United Kingdom, expert testimony on Battered Woman Syndrome is used as relevant evidence in assessing self-defense and criminal responsibility, the analyzed cases suggest that, in the Republic of North Macedonia, prolonged abuse is not always considered a significant mitigating circumstance. Of particular concern is the fact that courts in some cases focus exclusively on the immediate circumstances of the offense, without adequately considering the overall dynamics of the abusive relationship and the impact of long-term victimization on the woman's psychological state.

Legal scholars have criticized this approach, arguing that excluding the full context of coercive control results in incomplete interpretations of self-defense (Goodmark, 2012). The United Nations Office on Drugs and Crime (UNODC, 2018) also emphasizes the importance of understanding femicide within broader structural and gender-based power inequalities.

Therefore, sentencing decisions should incorporate the psychological consequences of prolonged abuse, the victim's perception of imminent danger, power imbalances, learned helplessness, and the full history of violence. Judicial authorities should analyze not only the events immediately preceding the homicide, but also the broader context of sustained abuse in which it occurred. A trauma-informed legal approach is necessary to ensure that cases are evaluated within their full contextual framework. Finally, although the findings cannot be statistically generalized due to the small purposive sample, they provide valuable insight into the experiences of women who have killed abusive partners in the Republic of North Macedonia and highlight the need for improved legal and institutional responses to intimate partner violence.

REFERENCES

- Abrahamson, L. Y., Seligman, M. E. P., & Teasdale, J. D. (1978). Learned helplessness in humans: Critique and reformulation. *Journal of Abnormal Psychology*, 87(1), 49–74. <https://doi.org/10.1037/0021-843X.87.1.49>
- American Psychiatric Association. (2022). *Diagnostic and statistical manual of mental disorders* (5th ed., text rev.; DSM-5-TR). American Psychiatric Association Publishing.
- Barnett, O. W. (2000). Why battered women do not leave, part 1: External inhibiting factors within society. *Trauma, Violence, & Abuse*, 1(4), 343–372.
- Browne, A. (1987). *When battered women kill*. Free Press.
- Campbell, J. C. (2003). Risk factors for femicide in abusive relationships: Results from a multisite case control study. *American Journal of Public Health*, 93(7), 1089–1097. <https://doi.org/10.2105/AJPH.93.7.1089>
- Craven, Z. (n.d.). *Battered woman syndrome*. http://www.adfvc.unsw.edu.au/PDF%20files/battered%20woman_syndrome.pdf
- Dutton, D. G. (1993). *The domestic assault of women: Psychological and criminal justice perspectives*. UBC Press.
- Easteal, P. (n.d.). *Battered women who kill: A plea of self-defence*. <http://www.freeonlinepapers.com/battered-womans-syndrome>
- Goodmark, L. (2012). *A troubled marriage: Domestic violence and the legal system*. New York University Press.
- Gondolf, E. W., & Fisher, E. R. (1988). *Battered women as survivors: An alternative to treating learned helplessness*. Lexington Books.
- Herman, J. L. (1997). *Trauma and recovery* (Rev. ed.). Basic Books.
- Konstantinović-Vilić, S. (1986). *Žene ubice*. Gradina.
- Levinson, D. (1988). Family violence in cross-cultural perspective. In V. B. Van Hasselt, R. L. Morrison, A. S. Bellack, & M. Hersen (Eds.), *Handbook of family violence* (pp. 435–455). Plenum Press.
- Linklaters LLP, & Penal Reform International. (2023). *Women who kill in response to domestic violence: How do criminal justice systems respond?*
- Liu, J. (2025). Domestic violence in cases of abused women who kill their partners.
- Mahan, S. (1987). *When Battered Women Kill*.
- Maier, S. F., & Seligman, M. E. P. (1976). Learned helplessness: Theory and evidence. *Journal of Experimental Psychology: General*, 105(1), 3–46.
- McCormick, N. B. (1976). Women who kill their husbands. *Journal of Marriage and the Family*, 38(4), 739–745.
- Mojanoski, C., Bačanović, O., Aćimovska-Maletić, I., Batić, D., Stefanovska, V., Milenkovska, M., Ilijevski, I., Jovanova, N., & Dimitrovska, A. (2014). *Položbata na osudenite lica vo penitenčijarnite ustanovi vo Makedonija (izveštaj od istražuvanje)*. Fakultet za bezbednost.
- Nash, C. (2024). Australia's divergent legal responses to women who kill their abusive partners. *International Journal for Crime, Justice and Social Democracy*.
- Nikolić-Ristanović, V. (2000). *Od žrtve do zatvorenice: Nasilje u porodici i kriminalitet žena*. Prometej.
- Organization for Security and Co-operation in Europe (OSCE). (2019). *OSCE-led survey on violence against women: North Macedonia*. <https://www.osce.org/>
- Organization for Security and Co-operation in Europe (OSCE). (2019). *OSCE-led survey on violence against women: North Macedonia*. OSCE.
- Psihijatrija: Dijagnoza i terapija. (1992). Savremena administracija.

- Seligman, M. E. P. (1975). *Helplessness: On depression, development, and death*. W. H. Freeman.
- Sethi, D., & Peshevska, D. J. (2014). Preventing interpersonal violence in Europe. *Open Access Macedonian Journal of Medical Sciences*, 2(2), 350-352.
- Stark, E. (2007). *Coercive control: How men entrap women in personal life*. Oxford University Press.
- Terr, L. C. (1991). Childhood traumas: An outline and overview. *American Journal of Psychiatry*, 148(1), 10–20.
- Walker, L. (1979). *The battered woman*. Harper & Row.
- Walker, L. E. (1984). *The battered woman syndrome*. Springer Publishing Company.
- Walker, L. E. (1993). Battered woman syndrome. In R. J. Gelles & D. R. Loseke (Eds.), *Current controversies on family violence* (pp. 133–154). Sage.
- Walker, L. E. (2006). Battered woman syndrome: Empirical findings. *Annals of the New York Academy of Sciences*, 1087(1), 142–157. <https://doi.org/10.1196/annals.1371.013>
- Walker, L. E. (2009). *The battered woman syndrome* (3rd ed.). Springer Publishing Company.
- Wolfgang, M. E. (1958). [As cited in Nikolić-Ristanović, V. (2000). *Od žrtve do zatvorenice: Nasilje u porodici i kriminalitet žena* (p. 21). Prometej].
- World Health Organization. (2025). *Violence against women prevalence estimates, 2023: Global, regional and national prevalence estimates for intimate partner violence against women and non-partner sexual violence against women*. WHO.
- World Health Organization. (2013). *Global and regional estimates of violence against women*. <https://www.who.int/publications/i/item/9789241564625>
- United Nations Office on Drugs and Crime. (2018). *Global study on homicide: Gender-related killing of women and girls*. <https://www.unodc.org/unodc/en/data-and-analysis/global-study-on-homicide.html>

CYBERBULLYING AS A PREDICTOR FOR DEPRESSION AND SUICIDAL BEHAVIORS IN PRIMARY AND SECONDARY SCHOOL STUDENTS

Dimitrinka Jordanova Peshevska

University American College Skopje, Skopje, Republic of North Macedonia

orcid.org/0000-0002-3499-3667

Dragana Batic

Faculty of Security – Skopje, UKLO

dragana.batic@uklo.edu.mk

orcid.org/0009-0002-1632-431X

Abstract

The rapid and frequent change of technology and environment is increasing the need for students to adjust and adapt to the emerging changes. The development of the technology and time spent on social media is increasing the risk of youth being exposed to cyberbullying, impacting their mental health. The aim of the study was to explore how cyber peer victimization is associated with depression and suicidal behaviors: ideas, thoughts and suicidal attempts in students. The sample included 1089 primary and secondary school students (57% female and 43% male) from North Macedonia. The ethical approval for the study was obtained from the Ethical Board at UACS. The instruments used were: DASS-21 for measuring depression, Centre for Disease Control in Atlanta measurement of suicidality, Cyberbullying Victimization Scale and socio-demographic questionnaire. The results showed a statistically significant correlation between exposure to digital victimization and depression and suicidal thoughts, plans and suicidal attempts in students. Digital victimization significantly predicted depression scores ($\beta=.348$, $p<.01$) and explained a significant proportion of variance in depression ($R^2=.121$; $F(1; 1013)=139.809$, $p<.01$). Additionally, digital victimization significantly predicted suicidal behaviors scores ($\beta=.317$, $p<.01$), explaining a significant proportion of variance in suicidal ideas, thoughts and suicidal attempts ($R^2=.100$; $F(1; 1043)=116.528$, $p<.01$). The rapid and frequent change of technology should be followed by increased preventive mechanisms developed by professionals to fill the gap and reduce the harm of the negative effects of technology, by providing alternative solutions for the wellbeing of the youngsters and their mental health.

Keywords: digital violence, depression, suicidality, students

1. INTRODUCTION

For a long period of time, historically speaking, violence was considered a criminal matter, viewed from the perspective of human rights. Certainly, in the last few decades, violence has become an important topic of psychology (Cornelius & Resseguie, 2007; World Health Organization, 2019) and public health, since health systems have become crucial in the prevention and treatment of victims of violence (WHO, 2002). Violence affects communities, endangering people's lives, health and happiness. Working to prevent violence means working to uncover the secrets and taboos of violent behavior. The publication of the World Report on Violence and Health in 2002 greatly contributed to the understanding of violence and its impact on society and individuals. Attention is devoted to considering the size and scope of the problem of violence, the risk factors that lead to violence, and the analysis of measures to prevent it (WHO, 2002).

There are different definitions of violence, which emphasize its different aspects. Popadić defined violence as follows: "an aggressive action, an episode within a complex symbolic exchange in which the action gains meaning within a constantly changing symbolic code shared and defined by the abuser, the victim and their environment" (Popadić 2009, p. 55). One of the most commonly used definitions of violence in the last two decades is the WHO definition (WHO, 2002 p. 5), which defines violence as follows:

"The intentional use of physical force and power by threat or action against oneself, another person, or a group of people or an entire community, which could result or results in injury, death, psychological consequences, underdevelopment or deprivation."

The definition also distinguishes between intent to injure another person and intent to use violence. Violence and notions of what is acceptable and unacceptable in behavior and what constitutes a violation of values and social norms in a particular culture are constantly being evaluated and changed (Dahlberg & Butchart, 2005). However, the definition clearly defines what is meant by violence regardless of cultural and social norms.

The typology of violence presented in the *World Report on Violence and Health* classifies violence into three broad categories, according to who carries out the act of violence: violence against oneself, collective violence and interpersonal violence (WHO, 2002). These three categories are further divided and reflect more specific types of violence. According to the nature of violent actions, violence can be psychological or emotional, physical and sexual, while it also includes actions such as deprivation or neglect (WHO, 2002).

One of the most important forms of interpersonal violence is peer violence or bullying. Most researchers cite the definition of peer violence set by Olweus, an author known for his pioneering work in this area. Olweus (1993 p.9) defines peer violence as "violent, intentional, aggressive behavior that is repeated over time and that involves a real or perceived power imbalance." This most commonly used definition of bullying consists of three main components: 1) aggressive behaviors that are 2) frequently repeated and 3) involve a power imbalance in favor of the bully (Olweus, 1993). According to this definition, a young person is a victim of peer violence when he is exposed to repeated acts by one or more peers and when he is unable to defend himself (Nansel et al., 2001; Smith et al., 1994; Olweus, 1993; Olweus, Limber, & Mihalic, 1999; Olweus & Limber 2010; 2018).

In the last few decades, since the social interaction of young people has moved to the Internet and social networks, violent behavior has also moved to cyberspace. Digital (electronic) peer violence or cyberbullying stands out as a special type of violence and

represents a serious and complex problem. In Anglo-Saxon literature, the terms used for digital or electronic violence are cyberbullying, online bullying, electronic bullying, and Internet harassment. The term cyberbullying or digital violence is most often used as a comprehensive term, which includes: the publication of inappropriate, false or harmful messages, both via mobile phones (texts, calls, posts, images, video clips) and online (e-mail, messaging, social networks, virtual chat rooms) (Tokunaga, 2010; UNESCO, 2019; WHO, 2002). Cyberbullying is a specific type of violence, since the bully maintains his anonymity, while the victim remains exposed to violence in front of a large number of visitors to social networks or in wider internet communication (Popadić, 2007; Tokunaga, 2010).

Mental health as a term in psychology represents a counterweight to the terms mental illness and mental disorder. With the adoption of the European Action Plan for Mental Health 2013–2020, the World Health Organization (WHO) places mental health high in the hierarchy of priorities and certainly on the international agenda (WHO, 2015). Mental health is considered by the WHO as an integral part of health (WHO, 2001), which is evident from the very definition of health by the World Health Organization: "a state of complete physical, mental and social well-being, and not merely the absence of disease and infirmity" (WHO, 2001). Health is recognized as a positive concept, where the emphasis is placed on the quality and balance of personal, social and physical capacities and the external conditions of the environment. In this sense, mental health is not only the absence of disease, because it is also related to physical health and behavior. According to the definition of the World Health Organization, mental health is "a state of well-being in which an individual realizes his own abilities, can cope with normal life stresses, can work productively and fruitfully, and can contribute to his community" (WHO, 2001).

Unlike the medical model, another perspective or "psychological tradition" operationalizes mental health in terms of one's subjective assessment of life satisfaction and the presence of positive or negative affect (Keyes, Shmotkin et al., 2002). Researchers have also developed a system for understanding behavioral, social and emotional problems in young people. Within this paradigm, and in current research, researchers have defined general types of behavioral and emotional problems in two broad dimensions – internalizing (overcontrolled) behaviors and externalizing (undercontrolled) behaviors (Keyes, Shmotkin et al., 2002). More specifically, internalizing behaviors include a broad domain of symptoms associated with depression, anxiety, stress, and social withdrawal.

1.1. Theoretical frameworks

1.1.1. Biological theories

Biological theories start from the idea that the cause of violent behavior is based on a certain biological-physiological structure, as well as on the influence of genetics, mechanisms in the brain, hormones and neurotransmitters. According to the traditional biological perspective, genetics and environment were viewed as unrelated, in accordance with Darwin's theory of natural selection of species. Natural selection is the result of the evolution of the genetic makeup through a survival response, or what is called "survival of the fittest". Building on this analogy, violence can be seen as a form of natural selection within which the weak will perish and the fittest will survive. This process ensures stronger genetic predispositions, which will reinforce the superiority of the species. Lamarck (1984) proposes a theory according to which an organism acquires certain characteristics by adapting to the environment, and the descendants then inherit those characteristics. This

theory is usually seen as pioneering work within epigenetic studies of the interdependent relationship between genetics and environment. New biological conceptual frameworks are aimed at studying the connection between genetic predisposition and the environment. Both factors are vital for the development of a healthy nervous system and brain plasticity, which further influence survival, problem solving, and interaction (Bird, 2007).

1.1.2. Social theory of learning

The study of bullying can also be understood using the theoretical perspective of social learning theory, proposed by Bandura (1977, 1986). This theory provides a framework for understanding, predicting, and changing human behavior. The theory identifies human behavior as a continuous interaction between personality characteristics, behavior and environment. Social learning theory explains that human beings learn by observing the behavior, attitudes, and behavioral outcomes of others. According to Bandura (1977, 2001), most human behaviors are learned by observation through modelling. Based on the observation of others, an idea is formed about how new behaviors are performed, and when necessary, this observation helps one to use a similar behavior in a certain new situation. Social learning theory explains human behavior in terms of continuous mutual interaction between personal determinants, behavioral and environmental factors. According to this theory, behavior is not simply an independent result of environmental, personality, and behavioral factors, but an interaction between them or mutual determinism.

Reciprocal determinism describes how the world and a person's behavior cause each other—one's environment causes one's behavior, and that behavior also has an effect on the environment (Bandura, 2001). Social learning theory recognizes that some sources of influence are stronger than others and that they do not all occur at the same time. The interaction between the three factors will vary depending on the individual, the actual behavior, and the situation in which the behavior occurs (Bandura, 1986).

The interaction between person and behavior includes the influences of a person's thoughts and actions (Bandura, 2001). The interaction between a person and the environment includes human beliefs and cognitive competencies that are developed and modified by social influences and structures in the environment (Bandura, 2001). Certainly, the interaction between environment and behavior involves a person's behavior determining aspects of his environment and in turn his environment modifies that environment (Bandura, 2001). These interactions are crucial to understanding and examining peer violence.

1.2. Overview of relevant empirical research

Relevant data on the prevalence of peer violence at the global level come from two international longitudinal studies. The first study, the Health Behavior in School-aged Children Survey (HBSC), presents data from 1983 to the present day from 49 countries across Europe, including North Macedonia, and from outside Europe – North America and Israel (Roberts et al., 2009). The School-Age Health Behavior Survey collects data on the health and well-being, social environment, and health behaviors of 11-, 13-, and 15-year-old boys and girls every four years (Chester et al. 2015; Inchley & Currie, 2013). Another relevant study is the Global school-based student health survey (GSHS), which has been conducted since 2003 in 96 countries around the world with the main goal of providing reliable data on health behaviors and protective factors among students aged 13 to 17 years (World Health Organization & Centers for Disease Control and Prevention, 2013).

According to the latest data on the global prevalence of bullying, almost every third student (32%) has been a victim of bullying in the last month. In all countries where the Global School Student Health Survey (GSHS) was conducted, the prevalence of violence ranged from 7.1% to 74% (WHO & CDC, 2013), while in countries where the Health Behavior Survey of School-aged Children (HBSC) was conducted, the prevalence ranged from 8.7% to 55.5% (Chester et al. 2015; Inchley & Currie, 2013). Physical and sexual violence are the two most common types of violence in the countries where the Global School Student Health Survey (GSHS) was conducted, while psychological violence is the most common type of violence in the countries where the Health Behavior Survey of School-Age Children (HBSC) was conducted (UNESCO, 2019; Chester et al. 2015; Waasdorp, Pas et al., 2017), certainly under the influence of cultural shifts towards more virtual social environments of the school population (Waasdorp, et al., 2017; Kowalski et al., 2014). In seven European countries, an increase in the percentage of reported cases of digital violence was observed among Internet users aged 9 to 16 – from 7% in 2010, the percentage of reported cases rose to 12% in 2014 (Mascheroni & Cuman, 2014).

According to the recent evidence, depression, suicidal thoughts and suicide attempts are considered the most common consequences of peer violence. Research on Risk Behaviors in Youth, conducted by the Centers for Disease Control and Prevention in Atlanta (USA) on a sample of 15,425 high school students, showed that 15% of students who were victims of digital violence attempted suicide (Messias et al., 2014). In a Turkish study conducted on a sample of 1276 students, high scores of exposure to peer violence as a result of digital bullying were statistically significantly associated with high scores for depression, anxiety, somatization, and low self-esteem (Tural et al., 2017). About 14% of students in a sample of 2,480 students aged between 12 and 13 in London, UK, reported being victims of cyberbullying. Compared to those who were not victims, in this study, students who were victims of digital bullying had a statistically higher chance of developing symptoms of depression and social anxiety (Fahy et al., 2016).

The consequences of cyberbullying can vary depending on the role that the students have, that is, whether they are in the role of the victim or in the role of the bully. Students who were victims of violence more often reported symptoms of depression, anxiety, somatic symptoms and suicidal behavior, while students in the role of bullies were more often associated with aggressive, delinquent behavior and use of illegal substances (Nixon, 2014). A number of studies support the association between exposure to peer violence, especially relational violence, and social anxiety in adolescents (La Greca & Harrison 2005; Storch & Esposito, 2003). La Greca and Harrison (2014) found that among older adolescents (ages 14–19), experiences of relational violence were highly associated with social anxiety. The consequences of digital violence can also depend on the gender of the student. Analysis of research conducted on a sample of 31,148 students, aged between 6 and 12 years, showed that digital violence is a significant predictor of emotional and behavioral problems in female students (Kim et al., 2018). Another study systematically examined the relationship between involvement in cyberbullying as a victim or perpetrator and depressive symptomatology and suicidal intentions. Data collected among 399 Canadian adolescents (57% female and 43% male) in grades 8 to 10 (mean age 14.2 years) showed that involvement in digital violence, either as a victim or as a bully, was associated with the prediction of depressive symptomatology and suicidal ideation (Bonanno, & Hymel, 2013). Findings from a large sample of 18,341 students aged 15 to 17 from six Chinese cities showed that exposure to cyberbullying among students was positively associated with symptoms of post-traumatic stress disorder, depression, self-harm, and

other mental health problems (Chen et al., 2018).

The main aim of the study was to investigate and explore the relationship between cyberbullying and depression, and suicidal behaviors among adolescents.

2. RESEARCH METHODOLOGY

2.1. Study design

The study applied a cross-sectional study design.

2.2. Sample and participants

A convenience sample of 1089 primary and secondary school students in the Republic of North Macedonia participated in the research. The criteria for participation in the research were: 1) that the students are between the ages of 11 and 17, currently enrolled in primary or secondary school in the Republic of North Macedonia; 2) that the students were enrolled in the appropriate school for at least six months before conducting the research; 3) that the examined students use the Internet and social networks (Kuzmanović, D., Pavlović, Z., et al., 2019) and 4) that the students' native language is Macedonian.

Students from eight primary and nine secondary schools in the Republic of North Macedonia were included in the research. The research included students from the 6th to the 9th grade of primary schools and from the 1st to the 3rd year of secondary state schools aged 11 to 17 years from four urban municipalities of the capital Skopje (Centar, Karpoš, Butel and Čair) and one suburban municipality (the village of Čučer Sandevo), as well as from Kočani, Štip and Vinica.

Out of a total of 1089 students, 380 students (34.9%) are from primary schools, and 709 students (65.1%) from secondary schools. A total of 42.5% of students are male, and 57.5% are female, aged 11 to 17. The largest number of students declared that they were Macedonians (86%), then 6.8% of the sample consisted of students of Albanian nationality, 2.8% of Serbian nationality, 1.5% of Bosniak nationality, and 1.1% of Roma nationality. 1.8% of students declared themselves under the second option. In relation to urban/rural representation, in 79.1% of cases students come from urban areas, while 20.9% of them come from rural areas. In relation to school success, the majority of students (68.4% of cases) had excellent success, in 22.4% of cases very good success was achieved, 8.1% of students achieved good success, and only 1.1% achieved sufficient success.

2.3. Instruments

The following measuring instruments were used in the empirical research:

Questionnaire for collecting sociodemographic data

This questionnaire was used to collect basic sociodemographic data on students (gender and age of primary/secondary school students, average performance in the previous year, mid-year average grade, ethnicity, monthly household income, and student internet use).

Cyberbullying Scale

The scale measures exposure to cyberbullying. It is a Likert-type scale and consists of 16 statements (Garaigordobil, 2017). Each item is scored on a four-point scale, ranging from 0 (never) to 3 (always), where respondents answer the questions whether and how often in the last 12 months they were exposed to any of the mentioned forms of violence. The total score on the scale ranges from 0 to 48. A higher score indicates that the student was more often exposed to digital violence by peers. For example, some of the items read like this: "I received messages with offensive content via mobile phone or the Internet."

The scale has shown good psychometric characteristics in previous research, with Cronbach's alpha coefficient for the scale being $\alpha=.82$ (Garaigordobil, 2017).

Depression, Anxiety and Stress Scale – DASS

Depression, Anxiety and Stress Scale – (DASS) is an instrument that measures negative emotional states of depression, anxiety and stress in the last two weeks (Lovibond & Lovibond, 1995). This scale exists in a longer (DASS-42) and a shorter version (DASS-21). In the current research, a shorter version of the DASS-21 was used. The DASS-21 scale contains 21 items and includes three subscales that measure the negative emotional states of depression, anxiety, and stress. For the purpose of this study, the depression subscale was used. It consists of 7 items: 1) a depression scale (example: "I felt that life was meaningless"). Respondents are asked to respond to each item on a four-point Likert-type scale, from 0 (not at all or never applies to me) to 3 (applies to me most of the time or always), and rate how they have felt in the past two weeks. The scores on each subscale can range from 0 to 21.

2.4. Procedure

The research procedure included the following stages: 1) pilot research and 2) main empirical research. A pilot study was conducted on elementary school students ($n=52$) in the period from February 1 to 15, 2020, in order to check the general usability of the measuring instruments, the comprehensibility of the questions, instructions, and determining the time necessary to fill out the questionnaire. No incomprehensible statements were found in the questionnaire, and several ambiguous statements were subsequently reformulated. The internal consistency of all scales in the study was relatively high.

In order to enable the implementation of the main empirical research, a request was submitted to the Ministry of Education and Science of the Republic of North Macedonia. By the decision of the Ministry of Education and Science of the Republic, consent was obtained for conducting research on peer violence among young people, as well as mental health among primary and secondary school students in the Republic of North Macedonia. The schools were contacted and informed about the subject and objectives of the research, in order to give consent from all parents and create a plan for conducting the research.

During the examination, in addition to informed consent, the students were once again informed about the objectives of the research, the principle of anonymity and confidentiality of data, as well as the right to refuse or stop the research at any time during its duration. Also, the students were directed to whom they could turn in relation to questions related to the research or in case of anxiety about the content of the study (for example, a psychologist or a pedagogue). The researcher emphasized to the students to be as honest as possible when answering.

2.5. Statistical analysis

The SPSS 23.0 software package was used for statistical data processing. In the first part of the processing, a descriptive analysis of the data was undertaken with the aim of obtaining information about the reliability of psychological measuring instruments and checking the normality of the distribution. The following statistical procedures were used in this part: mean, standard deviation, minimum and maximum scores. The internal consistency coefficient was used to check the reliability of psychological instruments, i.e. Cronbach's alpha coefficient.

Spearman's correlation coefficients were used to determine the relationship between the variables and to test the hypothesis. Linear regression analysis and multiple regression analysis were used to determine the importance of risk factors associated with different forms of exposure to cyberbullying.

3. RESULTS

3.1. Descriptive data analysis

In the first part of presenting the results, a descriptive analysis of the data is presented, with the aim of obtaining information about the reliability of the psychological measuring instruments.

According to the results on the scales, the average scores range from 2.25 (for the Cyberbullying Victimization Scale) to 8.17 (Depression Scale). Scatter measures range from 3.7 (for the Cyberbullying Victimization Scale) to 8.49 (Depression Scale). Reliability measures indicated very high reliability for both Cyberbullying Victimization Scale ($\alpha=.89$) and Depression Scale ($\alpha=.84$).

Table 1: Descriptive statistics for the scales

	n	M	SD	Mdn.	Min	Max	α
<i>Cyberbullying Victimization</i>	1089	2.25	3.7	1	0	31	.89
<i>Depression</i>	1089	8.17	8.49	6	0	42	.84

Notes: *n* – Number of participants; *M* – Mean; *SD* – Standard Deviation; *Mdn* – Median; *Min* – minimum; *Max* – maximum; α – Cronbach Alpha Coefficient.

A total of 54.6% of students were exposed to cyberbullying (from sometimes to always), as shown in Table 2. Of these, a total of 15.9% were sometimes exposed to cyberbullying, 31.7% of students were often exposed to cyberbullying, while 7% of students were always exposed to cyberbullying. Students of both sexes had a relatively similar percentage of representation (from sometimes to always) in the case of total cyberbullying (from 55.3% for boys to 54.2% for girls). Male respondents were exposed to cyberbullying (from sometimes to always) in 55.3% of cases (sometimes in 16.5% of cases, often in 30.8% of cases, always in 8% of cases). Female respondents were exposed to cyberbullying (from sometimes to always) in 54.2% of cases (sometimes in 15.4% of cases, often in 32.5% of cases, always in up to 6.3% of cases).

Table 2: Cyberbullying Victimization among students according to gender and school age

<i>Percentage</i>	<i>Never</i>	<i>Sometimes</i>	<i>Often</i>	<i>Always</i>	<i>CV from Sometimes to Always</i>
<i>Cyberbullying Victimization (CV)</i>	45.4%	15.9%	31.7%	7%	54.6%
<i>Male</i>	44.6%	16.5%	30.8%	8%	55.3%
<i>Female</i>	45.8%	15.4%	32.5%	6.3%	54.2%
<i>Primary school</i>	47.4%	17.6%	28.2%	6.8%	52.6%
<i>Secondary school</i>	44.3%	14.9%	33.6%	7.2%	55.7%

Notes: *CV* – Cyberbullying Victimization; *n*=1089

Linear regression model ($R^2=.201$; $F(5; 949)=47.665$, $p<.01$) in the case of the

criterion variable of depression is statistically relevant and according to the R^2 value explains 20.1% of the variance. Cyberbullying victimization ($\beta = .228, p < .01$) proved to be a significant predictor for depression (β coefficients have a positive sign). In other words, the greater the exposure to cyberbullying, the greater the possibility for depression in students (Table 3).

Table 3: Linear regression analysis - predicting depression (N=949) through cyberbullying victimization

Model	Unstandardized coefficients		Standardized coefficients		
	B	SE	β	t	p
(Constant)	-.004	.029		-.125	.900
<i>Cyberbullying Victimization (CV)</i>	.230	.032	.228	7.148	.000**

Notes: ** $p < 0.01$; * $p < 0.05$;

a. Criterion variable: *Depression*

b. Predictor variable: *Cyberbullying Victimization*

The results presented in Table 4 of the regression analysis indicate that the model is statistically significant ($R^2 = .125$; $F(5; 971) = 27.645, p < .01$), and that it explains 12.5% of the variance. Exposure to cyberbullying ($\beta = .290, p < .01$) proved to be a statistically significant predictor of the criterion variable of suicidal thoughts, plans and suicide attempts, where β coefficients have a positive sign. The greater the exposure to cyberbullying, the greater the likelihood of suicidal thoughts, plans and suicide attempts among primary and secondary school students.

Table 4: Linear regression analysis - predicting suicidal behaviors (thoughts, plans and attempts) (N=949) through cyberbullying victimization

Model	Unstandardized coefficients		Standardized coefficients		
	B	SE	β	t	p
(Constant)	.003	.030		.103	.918
<i>Cyberbullying Victimization (CV)</i>	.296	.034	.290	8.768	.000**

Notes: ** $p < 0.01$; * $p < 0.05$;

a. Criterion variable: *Suicidal Behaviors*

b. Predictor variable: *Cyberbullying Victimization*

4. DISCUSSION AND CONCLUSION

In the current research, a total of 54.6% of students were exposed to digital violence (from sometimes in 15.9% of students, more than once in 31.7% of students, to always in 7% of students). Although there is less data on the prevalence of digital or cyber-violence at the global level, this problem is growing and present around the world. For example, it is estimated that the prevalence of digital violence in the USA in 2014 ranged from 6% to 42% (UNESCO, 2017). In a recent meta-analysis, Tokunaga et al. (2010) concluded that children and teenagers are a particularly vulnerable category when it comes to this type of violence and that roughly 20-40% of this part of the population was exposed to some form of digital violence. In a recent study conducted in the USA on 1000 respondents (ages 9-12), a total of 38% of students reported exposure to digital violence on some social network or platform, while a total of 48% of them reported having some type of harmful online experience (Thorn, 2021). A recent review of studies on digital peer violence by Brochado et al. (Brochado et al., 2017) showed variability in the prevalence of digital violence. Thus, a review of studies shows that in Canada, depending on the study, the prevalence ranges from 1.9% to 65%, in China the variation is from 11.2% to 56.9%, while the lowest prevalences are recorded in Germany (6.3%), Sweden (5.2%) and Australia (5%) (Brochado et al., 2017). Research conducted on a sample of 353 adolescents admitted to emergency departments in hospitals showed that 47% of them were victims of digital violence (Ranney et al., 2016). Other studies report a lower prevalence of digital bullying, from 7% (Halliday et al., 2022) to 10% of students (Wang et al., 2009). Digital violence in Serbia is represented in 12.4% of cases (Popadić et al., 2014), while in another study the number of children who were exposed to digital violence was higher (16%) (Kuzmanović et al., 2019).

In this study, depressive symptoms were present in a total of 34.1% of students, which is a relatively similar degree of representation of the same symptoms as in adolescents in Canada (31.4% of cases) (Ehrenberg et al., 1990). The level of depression ranges from 10.2% of cases for a low level to 13.9% of cases for a moderate level; a high level of depression was recorded in 5.5% of students, while 4.5% of students declared that they had the highest level of depression. Other studies show that depressive symptoms range from 18.4% in Spain, 29.65% of cases in Great Britain (Jenkins et al., 2021), 55.8% of cases in Palestine (Ghanim et al., 2022), to 75% of cases in a study with data from Pakistan (Asif et al., 2020).

It was determined that the greater the exposure to cyberbullying, the greater the possibility of depression and depressive symptoms in students. Similar findings were confirmed by other studies, in which it was shown that exposure to cyberbullying increases the chances of depression in adolescents (Bonanno, & Hymel, 2013; Martínez-Monteagudo et al., 2020). Namely, students who are victims of cyberbullying are 1.4 times more likely to develop depressive symptoms than students who were not exposed (Waasdorp & Bradshaw, 2015). Current research shows such a connection. Exposure to cyberbullying was demonstrated to be statistically significant when predicting suicidal thoughts, plans and suicide attempts. The greater the exposure to cyberbullying victimization is (Leone & Carroll, 2016; Klomek et al., 2007; Klomek et al., 2008), the greater are the odds of predicting suicidal thoughts, plans, and suicide attempts in primary and secondary school students.

The rapid development of technology on the one hand gives young people unlimited opportunities, while on the other hand it creates a new space, which remains without any control and supervision by parents, schools or other institutions. One gets the impression that the just-mentioned accelerated technological development has left a lot of room for young people to remain alone and without support, with the possibility of being hurt, manipulated and exposed to various forms of violence. For this reason, it would be necessary for the relevant institutions or factors at the social level to plan the regulation of the digital space, in terms of control and supervision of the content offered on social networks, in a way that would be beneficial to different developmental groups of children and young people. The role of social support that parents, peers and other important people can play in mental health prevention is invaluable. In this sense, planning of preventive programs and interventions aimed at appropriate assistance, i.e. planning that would take into account different ways of providing social support, would be a step in the right direction.

REFERENCES

- Asif, S., Mudassar, A., Shahzad, T. Z., Raouf, M., & Pervaiz, T. (2020). Frequency of depression, anxiety and stress among university students. *Pakistan Journal of Medical Sciences*, 36(5), 971–976
- Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioural change. *Psychological Review*, 84(2), 191–215.
- Bandura, A. (1986). Fearful expectations and avoidant actions as co-effects of perceived self-inefficacy. *American Psychologist*, 41(12), 1389–1391. <https://doi.org/10.1037/0003-066X.41.12.1389>.
- Bandura, A., Barbaranelli, C., Caprara, G. V., & Pastorelli, C. (2001). Self-efficacy beliefs as shapers of children's aspirations and career trajectories. *Child Development*, 72(1), 187–206.
- Bird, A. (2007). Perceptions of epigenetics. *Nature*, 447(7143), 396–398.
- Bonanno, R. A., & Hymel, S. (2013). Cyber bullying and internalizing difficulties: Above and beyond the impact of traditional forms of bullying. *Journal of Youth and Adolescence*, 42(5), 685–697.
- Brochado, S., Soares, S., & Fraga, S. (2017). A scoping review on studies of cyberbullying prevalence among adolescents. *Trauma, Violence, & Abuse*, 18(5), 523–531.
- Chester, K. L., Callaghan, M., Cosma, A., Donnelly, P., Craig, W., Walsh, S., & Molcho, M. (2015). Cross-national time trends in bullying victimization in 33 countries among children aged 11, 13 and 15 from 2002 to 2010. *The European Journal of Public Health*, 25(2), 61–64.
- Chen, Q., Lo, C. K., Zhu, Y., Cheung, A., Chan, K. L., & Ip, P. (2018). Family polyvictimization and cyberbullying among adolescents in a Chinese school sample. *Child Abuse & Neglect*, 77, 180–87.
- Cornelius, T. L., & Resseguie, N. (2007). Primary and secondary prevention programs for dating violence: A review of the literature. *Aggression and Violent Behavior*, 12(3), 364–375.
- Dahlberg, L. L., & Butchart, A. (2005). State of the science: violence prevention efforts in developing and developed countries. *International Journal of Injury Control and Safety Promotion*, 12(2), 93–104.
- Ehrenberg, M. F., Cox, D. N., & Koopman, R. F. (1990). The prevalence of depression in

- high school students. *Adolescence*, 25(100), 905–912.
- Fahy, A. E., Stansfeld, S. A., Smuk, M., Smith, N. R., Cummins, S., & Clark, C. (2016). Longitudinal associations between cyberbullying involvement and adolescent mental health. *Journal of Adolescent Health*, 59(5), 502–509.
- Garaigordobil, M. (2017). Psychometric properties of the Cyberbullying Test, a screening instrument to measure cybervictimization, cyberaggression, and cyberobservation. *Journal of Interpersonal Violence*, 32(23), 3556–3576. doi:10.1177/0886260515600165
- Ghanim, M., Rabayaa, M., Atout, S., Al-Othman, N., & Alqub, M. (2022). Prevalence of anxiety and depression among Palestinian university students: a cross-sectional study during COVID-19 pandemic. *Middle East Current Psychiatry*, 29(1), 71.
- Halliday, S., Taylor, A., Turnbull, D., & Gregory, T. (2022). The Relationship Between Traditional and Cyber Bullying Victimization in Early Adolescence and Emotional Wellbeing: A Cross Sectional, Population-Based Study. *International Journal of Bullying Prevention*, 1–14.
- Inchley, J., & Currie, D. (2013). Growing up unequal: gender and socioeconomic differences in young people's health and well-being. *Health Behaviour in School-aged Children (HBSC) study: international report from the, 2014*.
- Keyes, C. L., Shmotkin, D., & Ryff, C. D. (2002). Optimizing well-being: the empirical encounter of two traditions. *Journal of Personality and Social Psychology*, 82(6), 1007.
- Kim, S., Colwell, S. R., Kata, A., Boyle, M. H., & Georgiades, K. (2018). Cyberbullying victimization and adolescent mental health: Evidence of differential effects by sex and mental health problem type. *Journal of Youth and Adolescence*, 47(3), 661–672.
- Klomek, A. B., Marrocco, F., Kleinman, M., Schonfeld, I. S., & Gould, M. S. (2007). Bullying, depression, and suicidality in adolescents. *Journal of the American Academy of Child & Adolescent Psychiatry*, 46(1), 40–49.
- Klomek, A. B., Marrocco, F., Kleinman, M., Schonfeld, I. S., & Gould, M. S. (2008). Peer victimization, depression, and suicidality in adolescents. *Suicide and Life-Threatening Behavior*, 38(2), 166–180.
- Kowalski, R. M., Giumetti, G. W., Schroeder, A. N., & Lattanner, M. R. (2014). Bullying in the digital age: A critical review and meta-analysis of cyberbullying research among youth. *Psychological Bulletin*, 140(4), 1073–1137. <https://doi.org/10.1037/a0035618>
- Kuzmanović, D., Pavlović, Z., Popadić, D., & Milosevic, T. (2019). *Korišćenje interneta i digitalne tehnologije kod dece i mladih u Srbiji: rezultati istraživanja Dece Evrope na internet*. Beograd: Institut za psihologiju Filozofskog fakulteta u Beogradu.
- La Greca, A. M., & Harrison, H. M. (2005). Adolescent peer relations, friendships, and romantic relationships: Do they predict social anxiety and depression? *Journal of Clinical Child and Adolescent Psychology*, 34(1), 49–61.
- Lamarck, J. B. (1984). *Zoological Philosophy*. University of Chicago.
- Leone, J. M., & Carroll, J. M. (2016). Victimization and suicidality among female college students. *Journal of American College Health*, 64(6), 421–428.
- Lovibond, P. F., & Lovibond, S. H. (1995). The structure of negative emotional states: Comparison of the Depression Anxiety Stress Scales (DASS) with Beck Depression and Anxiety Inventories. *Behaviour Research and Therapy*, 33(3), 335–343.
- Martínez-Monteaudo, M. C., Delgado, B., Díaz-Herrero, Á., & García-Fernández, J. M. (2020). Relationship between suicidal thinking, anxiety, depression and stress in university students who are victims of cyberbullying. *Psychiatry Research*, 286, 112856.

- Mascheroni, G., & Cuman, A. (2014). *Net children go mobile: Final report: Deliverables D6. 4 and D5. 2*. Educatt.
- Messias, E., Kindrick, K., & Castro, J. (2014). School bullying, cyberbullying, or both: correlates of teen suicidality in the 2011 CDC Youth Risk Behavior Survey. *Comprehensive Psychiatry*, 55(5), 1063–1068.
- Nansel, T. R., Overpeck, M., Pilla, R. S., Ruan, W. J., Simons-Morton, B., & Scheidt, P. (2001). Bullying behaviors among US youth: Prevalence and association with psychosocial adjustment. *JAMA*, 285(16), 2094–2100.
- Nixon, C. L. (2014). Current perspectives: the impact of cyberbullying on adolescent health. *Adolescent Health, Medicine and Therapeutics*, 5, 143–158. doi: 10.2147/AHMT.S36456
- Olweus, D. (1993). *Bully/victim problems among schoolchildren: Long-term consequences and an effective intervention program*. In S. Hodgins (Ed.), *Mental Disorder and Crime* (p.317–349). Sage Publications, Inc.
- Olweus, D., Limber, S., & Mihalic, S. F. (1999). Blueprints for violence prevention, book nine: Bullying prevention program. Boulder, CO: Center for the Study and Prevention of Violence.
- Olweus, D., & Limber, S. P. (2010). Bullying in school: Evaluation and dissemination of the Olweus Bullying Prevention Program. *American Journal of Orthopsychiatry*, 80(1), 124–134. doi:10.1111/j.1939-0025.2010.01015.x
- Olweus, D., & Limber, S. P. (2018). Some problems with cyberbullying research. *Current Opinion in Psychology*, 19, 139–143.
- Popadić, D. (2009). *Nasilje u školama*. Institut za psihologiju.
- Popadić, D. i Plut, D. (2008). Nasilje u osnovnim školama u Srbiji – Oblici i učestalost. *Psihologija*, 40, No. 2, 309–328.
- Popadić, D., Plut, D. i Pavlović, Z. (2014). *Nasilje u školama Srbije – Analiza stanja od 2006. Do 2013. godine*. Beograd: Institut za psihologiju.
- Ranney, M. L., Patena, J. V., Nugent, N., Spirito, A., Boyer, E., Zatzick, D., & Cunningham, R. (2016). PTSD, cyberbullying and peer violence: prevalence and correlates among adolescent emergency department patients. *General Hospital Psychiatry*, 39, 32–38.
- Roberts, C., Freeman, J., Samdal, O., Schnohr, C. W., De Looze, M. E., Gabhainn, S. N., ... & Rasmussen, M. (2009). The Health Behaviour in School-aged Children (HBSC) study: methodological developments and current tensions. *International Journal of Public Health*, 54(2), 140–150.
- Smith, P. K., Cowie, H., & Sharp, S. (1994). Working directly with pupils involved in bullying situations. *School Bullying: Insights and Perspectives*, 193–212.
- Storch, E. A., & Esposito, L. E. (2003). Peer victimization and posttraumatic stress among children. *Child Study Journal*, 33(2), 91–99.
- Thorn. (2021). *Responding to Online Threats: Minors' Perspectives on Disclosing, Reporting, and Blocking: Findings from 2020 quantitative research among 9–17 year olds*. <https://info.thorn.org/hubfs/>
- Tokunaga, R. S. (2010). Following you home from school: A critical review and synthesis of research on cyberbullying victimization. *Computers in Human Behavior*, 26, 277–287. doi: 10.1016/j.chb.2009.11.014
- Tural Hesapcioglu, S., & Ercan, F. (2017). Traditional and cyberbullying co-occurrence and its relationship to psychiatric symptoms. *Pediatrics International*, 59(1), 16–22.
- UNESCO (2017). *School Violence and Bullying: Global Status Report*. Paris: UNESCO.

- UNESCO. (2019). *Behind the numbers: Ending the school violence and bullying*. Paris: UNESCO.
- Waasdorp, T. E., & Bradshaw, C. P. (2015). The overlap between cyberbullying and traditional bullying. *Journal of Adolescent Health, 56*(5), 483–488.
- Waasdorp, T. E., Pas, E. T., Zablotsky, B., & Bradshaw, C. P. (2017). Ten-year trends in bullying and related attitudes among 4th-to 12th-graders. *Pediatrics, 139*(6), e20162615.
- Wang, J., Iannotti, R. J., & Nansel, T. R. (2009). School bullying among adolescents in the United States: Physical, verbal, relational, and cyber. *Journal of Adolescent Health, 45*(4), 368–375.
- World Health Organization. (2002). *The World Health Report 2001 on Mental Health*. Geneva: WHO.
- World Health Organization (2002). *World Report on Violence and Health*. Geneva: WHO.
- World Health Organization, & Centers for Disease Control and Prevention (2013). *Global school-based student health survey (GSHS)*. WHO & CDC.
- World Health Organization. (2015). *Update of the Mental Health Gap Action Programme (mhGAP) guidelines for mental, neurological and substance use disorders, 2015*. World Health Organization.
- World Health Organization. (2019). *School-based violence prevention: a practical handbook*. Geneva: WHO.