

Sovereignty in the Information Environment

A Policy Study
2026



Sovereignty in the Information Environment

A Policy Study
2026

Vienna, 2026

Disclaimer

The views, thoughts, and opinions expressed in this study are those of the individual authors alone and do not necessarily reflect the official policy or position of the Austrian Federal Ministry of Defence.

Imprint:

Media owner, publisher and producer:
Austrian Federal Ministry of Defence,
Roßauer Lände 1, 1090 Vienna

bmlv.gv.at

Editing: Daniel Hikes-Wurm, Milena Sagawa-Krasny

Photo credit/Cover image: iStock/kontekbrothers

Print: Heeresdruckzentrum, 1030 Vienna

Vienna, 2026

ISBN: 978-3-902275-69-1

Table of contents

Preface	5
1	
Christian Stadler: On Notion and Nature of ‘Sovereignty’ between Mimesis and Might	8
1.1 Polemological Prologue.....	10
1.2 On the History of the Concept of Sovereignty.....	12
1.2.1 On the Political Concept of Sovereignty (1500–1750).....	12
1.2.2 On the Ethical Concept of Sovereignty (1750–1850).....	16
1.2.3 On the Constitutional Concept of Sovereignty (1850–1950).....	19
1.3 On Notion and Nature of ‘Digital Sovereignty’.....	22
1.4 On René Girard and his Mimetic Concept of Sovereignty.....	28
1.5 Polemological Epilogue.....	31
2	
Matthias Postl: The Information Environment	32
2.1 Introduction	34
2.1.1 Distinction from cyberspace.....	35
2.2 Method.....	36
2.3 What is information?.....	37
2.3.1 Historical origins and interdisciplinary perspectives.....	37
2.3.2 Conclusion and classification of the concept of information within the information space.....	40
2.4 The EU information space – a classification.....	40
2.4.1 How the information space became a field of action.....	40

2.4.2 Protecting the democratic public sphere as a starting point.....	40
2.4.3 Platform architecture as a normative structure.....	42
2.4.4 Hybrid logics of influence in the information space: Foreign Information Manipulation and Interference (FIMI).....	44
2.4.5 Resilience as an organisational principle.....	46
2.4.6 Digital sovereignty in a fragmented order.....	47
2.4.7 Terminological consolidation: concepts of space in the European context.....	48
2.4.8 EU information space.....	49
2.5 The NATO Information Space – Operational Logic, Perception Management and Transatlantic Interdependence.....	50
2.5.1 From communication space to operational space.....	50
2.5.2 The NATO ‘Information Environment’: Structure and Definition.....	51
2.5.3 The information space as a space of power and discourse.....	52
2.5.4 Perception management: framing and agenda setting as operational mechanisms.....	53
2.5.5 Narratives as a strategic structuring of meaning.....	54
2.5.6 Operationalisation: The information environment as a strategic arena.....	55
2.5.7 Comparison with the EU: information space and information environment.....	56
2.5.8 Reference definition: The NATO information environment.....	56
2.6 The US information environment: From information control to structured meaning-making.....	58
2.6.1 Information as strategic infrastructure.....	58

2.6.2 The cognitive opening.....	59
2.6.3 The Information Environment as a Structured Framework of Meaning.....	59
2.6.4 The current definition of the 'information environment' in 2022.....	60
2.7 The Chinese information space as a cognitive space of order.....	61
2.7.1 From sovereignty to the question of order.....	61
2.7.2 Institutional integration and conceptual transformation.....	62
2.7.3 Cognitive security and warfare.....	63
2.8 The information space in the strategic thinking of the Russian Federation.....	64
2.9 Classification and comparison of concepts in the information space.....	68
2.9.1 Psychological Operations (PSYOPS).....	68
2.9.2 Propaganda and disinformation.....	71
2.10 Comparative Matrix of the Information Environment.....	73
3	
Alexandra Duca: Cognitive Warfare.....	76
3.1 Relevant Obligations under International Law.....	79
3.1.1 State sovereignty.....	80
3.1.2 Prohibition of Intervention.....	81
3.1.3 Due Diligence.....	83
3.1.4 Prohibition of the Threat or Use of Force.....	84
3.1.5 International Humanitarian Law.....	86
3.2 Possible Measures in Response to an Internationally Wrongful Act.....	87
3.2.1 State Responsibility and Countermeasures.....	87

3.2.2 The Threshold of Armed Attack and the Right of Self-Defence.....	88
3.3 Human Rights Implications.....	89
3.3.1 Right to Freedom of Expression.....	90
3.3.2 Right to Privacy and Data Protection.....	91
3.3.3 Right to Integrity.....	91
3.3.4 Right to Participate in Political and Public Affairs.....	92
3.4 Conclusion.....	92
4	
Philip Schäfer: Protection and Civil Liberties in the Digital Information Environment.....	94
4.1 Introduction.....	96
4.2 First key question.....	97
4.3 Policy recommendations regarding the first guiding question.....	98
4.4 Second key question.....	99
4.5 Policy recommendations on the second key question.....	99
4.6 Key findings on the balance between protection and freedom in the digital information space.....	100
4.7 The central role of the question of balance in democratic security architecture.....	101
4.8 Research.....	103
4.9 Definition of securitisation.....	103
4.10 Methodological overview.....	104
4.11 Systematic literature review.....	105
4.12 Thematic search strands and key findings.....	107
4.13 Development of the analytical framework and compilation of the strategic document corpus.....	110

4.14 Analysis of security policy strategy documents in the European context.....	112
4.15 A typology of democratic security architectures in the digital information space.....	115
4.16 An overview of ideal-type security architectures.....	118
4.17 Policy recommendations.....	119
4.18 Risks of military intervention and strategies against securitisation.....	121
4.19 Key findings on addressing securitisation in the digital information space.....	121
4.20 The strategic significance of the securitisation issue for democratic security architecture.....	122
4.21 Study.....	123
4.22 Methodological overview.....	123
4.23 Systematic literature review.....	126
4.24 Thematic search strands and key findings.....	128
4.25 An examination of the militarisation of strategic security documents.....	133
4.26 Operationalisation of securitisation in the digital information space.....	134
4.27 Indicators of the securitisation of the information space.....	135
4.28 The securitisation of the digital information space in European strategy documents.....	136
4.29 Indicator-based criteria of securitisation.....	139
4.30 How security strategies reinterpret the information space.....	141
4.31 Implications for democratic theory.....	143
4.32 Policy Recommendations.....	143
4.33 Conclusion.....	146

5

Mikael Weissmann and James Pamment: Governance of the Information Environment. Concrete Measures for Cognitive Security and Cognitive Warfare.....	148
5.1 Introduction & Scope.....	151
5.2 Delimitations.....	156
5.3 Holistic governance, hybrid threats and the role of a Ministry of Defence.....	157
5.4 The bigger picture – Holistic governance and whole-of-society defence as a prerequisite for success.....	159
5.5 The complexity of the security environment: peace, crisis, war and the grey-zone.....	161
5.6 The complexity of hybrid threats and cognitive warfare.....	163
5.7 MoD as an engine in the whole-of-society defence system (The Hourglass Model).....	166
5.8 Role of the MoD.....	172
5.9 MoD's contribution to whole-of-society defence and cognitive security.....	173
5.10 MoD Instruments: Capacity, Instruments and Tasks.....	174
5.11 What a Ministry of Defence Can Contribute to Holistic Governance and Security in the Information Sphere.....	175
5.12 Ethical-Legal Decision Framework (“Could/Should/Carry Out”).....	176
5.13 Capacity-instruments-task model for MoD.....	179
5.14 Target audience.....	181

Preservation and Promotion of Democratic Resilience.....	184
6.1 Executive Summary and Key Takeaways.....	186
6.2 Core Findings.....	188
6.3 Implications for Strategic Communication.....	191
6.4 Introduction.....	194
6.5 Definition of Core Constructs.....	195
6.6 The Complex Framework for Studying Misinformation Dynamics.....	197
6.7 The PRORESIL Project.....	199
6.8 Study 1 – Understanding Susceptibility to Misinformation.....	200
6.8.1 Database and Screening Procedure.....	201
6.8.2 Scholarly Research on the Susceptibility to Misinformation (RQ1).....	202
6.8.3 Results.....	204
6.8.4 Most Prevalent Topics in the Academic Literature (RQ2).....	208
6.8.5 Results.....	209
6.8.6 Summary of Core Findings Study 1.....	213
6.8.7 Limitations and Pathways for Future Research.....	214
6.9 Study 2 – Tackling Misinformation.....	215
6.9.1 Why study interventions against misinformation?.....	216
6.9.2 How scholars discuss interventions against misinformation (RQ3).....	218
6.9.3 Results.....	219

6.9.4 How do state responses to misinformation during recent international conflicts compare to the interventions emphasised in the scholarly literature (RQ4)?.....	224
6.9.5 Results.....	225
6.9.6 Summary of Core Findings Study 2.....	230
6.10 Study 3 – Democratic Resilience Against Misinformation.....	232
6.10.1 Results: Austrians’ general preferences (RQ5).....	238
6.10.2 Results: Austrians’ Preferences in Times of Crises (RQ6).....	242
6.10.3 Citizens’ susceptibility to misinformation and their preferences for different information environments (RQ7).....	247
6.10.4 Identifying Susceptible Target Groups.....	254
6.10.5 Three Clusters: Institutional Mainstream, Networked Apart, and Right-Wing Authoritarian.....	254
6.10.6 How are citizens’ preferences related to susceptibility towards misinformation? (RQ7).....	258
6.10.7 Summary of Core Findings Study 3.....	261
6.11 General Discussion of Limitations and Directions for Future Research.....	262

Conclusion.....	264
------------------------	------------

Appendix: André K. Rodarte and Lena Frischlich:

Preservation and Promotion of Democratic Resilience –

A Scientific Basis for Policy Recommendations.....	273
---	------------

Bibliography.....	292
--------------------------	------------

1 Christian Stadler: On Notion and Nature of ‘Sovereignty’ between Mimesis and Might – A Polemological Reflection on Digital Sovereignty in the Information Space.....	292
2 Matthias Postl: The Information Environment – A Comparative Analysis and Contextualisation of Existing Concepts.....	294
3 Alexandra Duca: Cognitive Warfare – A Legal Assessment.....	304
4 Philip Schäfer: Protection and Civil Liberties in the Digital Information Environment.....	306
5 Mikael Weissmann and James Pamment: Governance of the Information Environment – Concrete Measures for Cognitive Security and Cognitive Warfare: The Role of a Ministry of Defence (MoD).....	316
6 André K. Rodarte und Lena Frischlich: Preservation and Promotion of Democratic Resilience – A Scientific Basis for Policy Recommendations.....	318

Preface

The information environment has emerged as a pivotal arena of the conflict landscape in the 21st century. Hybrid threats and the digital transformation are increasingly blurring the boundaries between the civilian and military spheres. Hybrid conflicts challenge the classical dichotomy between war and peace that has underpinned international law and traditional strategic thinking. State and non-state actors alike are employing a broad spectrum of tactics below the threshold of war, ranging from infrastructure sabotage and economic coercion to coordinated information operations on social media. These methods are designed to destabilise societies from within, undermine their cohesion, and erode public trust.

While human perception, its underlying processes, and potential avenues for psychological manipulation have long been of relevance to security policy strategies, modern technologies have reshaped the dynamics in this domain. James Giordano captured this heightened significance succinctly, stating, “The brain will be the battlefield of the 21st century.”¹ Ultimately, it is the perception of reality that forms the foundation for the formation of political will as well as for the capacity for action and defence, both for individuals and for communities.

These complex developments fundamentally question the military’s traditional understanding of its role in society. On the one hand, maintaining conventional armed forces of adequate quality and quantity remains essential to ensure credible deterrence against classic military confrontations. On the other hand, the manifestations of hybrid conflicts demand new forms of resilience. In this context, a critical question arises: What role should a Ministry of Defence (MoD) and the military play?

1 NATO Allied Command Transformation Innovation Hub (2020) Cognitive Warfare. Norfolk, VA: NATO ACT Innovation Hub. Verfügbar unter: https://innovation-hub-act.org/wp-content/uploads/2023/12/20210122_CW-Final.pdf (Accessed: 12 May 2026).

These considerations served as the starting point for the multidisciplinary study at hand, “Sovereignty in the Information Environment“. It draws on philosophy, legal studies, political science, and psychology for a holistic analysis. Beyond its theoretical foundation, the study relies on empirical research, including the results of a representative survey of over 1,900 participants conducted across Austria.

The study examines two main areas of focus. First, it addresses a number of fundamental conceptual and legal questions:

- Is the concept of sovereignty viable in the digital realm, and what aspects does it encompass?
- How do we define “information“, and how do global actors such as the EU, the US, NATO, China, and Russia differ in their conceptions of the information space?
- How should cognitive warfare be assessed from a legal perspective?

The second area concerns governance in its various facets.

A singular focus on security measures always carries the risk of “securitisation“, wherein political problems are defined as existential threats to legitimise extraordinary interventions. This raises crucial questions:

- How can democratic states maintain a balance between security interests and civil liberties?
- What undesirable or unintended effects may arise from the deployment of military forces, how can these be counteracted, and how can this risk be managed?

To outline concrete options for action for the Ministry of Defence and the military, the study then turns to the following questions:

- How can holistic governance in the information space strengthen protection against influence operations, and what contribution can a Ministry of Defence make in this regard?
- What instruments are available to a Ministry of Defence, and which capabilities and tasks are required?
- Which target groups should the Ministry of Defence address?

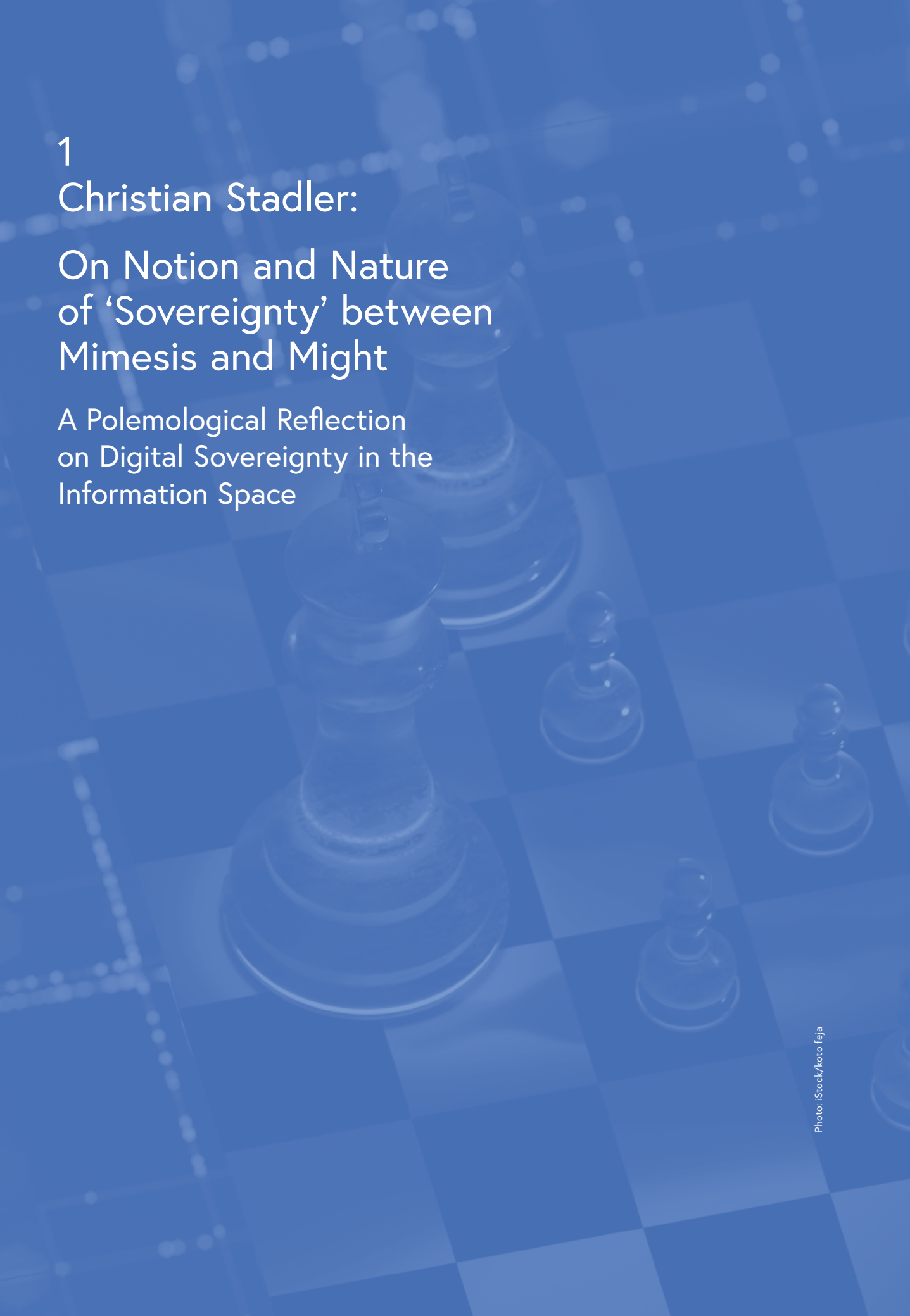
- Which target groups are vulnerable to influence, and why?
- What constitutes effective strategic communication?
- What are protective factors against manipulation in the information space?

The study highlights the necessity for comprehensive state and societal cooperation, both at the national and international levels. The information environment is a complex system that far exceeds the jurisdictional scope of any single ministry. It is not the task of the Ministry of Defence or the military to assume roles that belong to political leadership, civil authorities, civil society, or the private sector. However, due to their specific competencies and capabilities – including intelligence gathering, strategic analysis, military planning, and strategic communication – the Ministry of Defence and the military occupy a unique position for fostering cooperation with national and international partners.

It is our hope that readers will find this study both informative and thought-provoking.

Daniel Hikes-Wurm and Milena Sagawa-Krasny

Vienna, 27 May 2026



1

Christian Stadler:

On Notion and Nature of 'Sovereignty' between Mimesis and Might

A Polemological Reflection
on Digital Sovereignty in the
Information Space



On Notion and Nature of ‘Sovereignty’ between Mimesis and Might. A Polemological Reflection on Digital Sovereignty in the Information Space.

MMag. DDr Christian Stadler, Associate Professor of Legal Philosophy at the University of Vienna, has served as the Scientific Director of the Polemological and Legal-Ethical research collaboration between the University of Vienna and the National Defence Academy Vienna since 2013. His research focus areas include political philosophy, polemology, and the ethics of public security.

1.1 Polemological Prologue

Polemology is a modern science of the 20th century that has its roots in ancient thinking, such as Plato's political philosophy. Polemology is based on Plato's classical three-part theory of the soul and thus also of society, which distinguishes between Epithymeticon – Thymoeides – Logistikon, meaning: nourishing soul – defensive soul – teaching soul or nourishing class – defensive class – teaching class. Each of these three dimensions has a specific virtue: moderation, courage and wisdom. When these three dimensions are in harmony, Plato speaks of 'justice' – whether it be the justice of the individual or that of the community. This justice is what matters – first to Plato, then to Augustine, and finally to Schelling again. Moderation – courage – wisdom – justice are the four classical cardinal virtues. This psychological and social triad also forms the basis of the polemology concept developed at the National Defence Academy in Vienna (in cooperation with the University of Vienna): polemology has the task or function of providing a comprehensive picture of the situation for further military assessment. As the name suggests, this

involves the core of every military situation, i.e. gaining the most comprehensive possible understanding of what is currently happening as a conflict (*polemos*).

In order to enable such a comprehensive assessment of the situation, at least three dimensions must be considered and reflected upon: (1) the *Oikos* dimension (this concerns the foundations of (social) survival, the economy (real and fiscal), resources (human and natural), energy and its pathways, trade and its pathways, but also climate, topography, vegetation, etc.). This must be considered separately from the (2) *Polis* level, i.e. the level of political, normative and military conflicts. This concerns (geo)politics, (geo)strategy, power, institutions, rule, dominance, but also legal and constitutional culture and, ultimately, the military and war. Finally, the third dimension addresses the cultural dimension of a conflict, the (3) *Logos* level. This concerns language, religion, worldview, conception of humanity, civilisation, history and its narratives in literature and music and philosophy itself.

What can only be hinted at here is the so-called 'polemological formula': *Oikos* x *Logos* => *Polis*. This means that the respective socio-economic conditions for survival must be assessed in the light of the respective socio-cultural characteristics in order to obtain a prognosis or evaluation of the expected *Polis* dimension – i.e. ultimately war or peace.

Building on this basic structure of the polemological analysis model, it is now necessary to situate the present topic appropriately: the focus is on 'digital sovereignty in the information space'. This topic touches upon both the *Oikos* and the *Logos* dimensions:

The *Oikos* dimension concerns the question of hardware and software within the information space and thus also digital information flows; we shall refer to this as 'information infrastructure'. In this area, the issue – in short – is the extent to which one has control over this information infrastructure, including networks, servers, communication technologies, information storage systems, etc. This question is central, but concerns 'only' the *Oikos* sphere, whose strategic power potential lies in either destroying or withdrawing this infrastructure. Whilst in recent years – whether in the context of religiously or state-sponsored terrorism – we have placed greater emphasis on the issue of disrupting or destroying this infrastructure, it has become clear in recent months – since the Trump administration's overt power politics – that the

withdrawal or ‘shutting down’ of this communications infrastructure could also be used as an instrument of power. This dimension has been inconceivable in the North Atlantic region since the founding of NATO but has now become very apparent.

The *Logos* dimension, however, concerns the content – the texts, images and narratives – produced within the information space. This aspect of content, which has a far-reaching impact and can exert influence both directly and indirectly, will hereinafter be referred to as “information impressions”. These constitute the other substantive dimension of what takes place within the information space.

Logos relates to *Oikos*, and information impressions relate to the information infrastructure, just as new wine relates to old wineskins, we should be wary of overlooking the acidity of the new wine by focusing solely on the stability of the old wineskins.

Before we get to the heart of the matter, we need to clarify the key political concept – the concept of ‘sovereignty’, so that we can understand what has been the case up to now – at least since the beginning of the modern era.²

1.2 On the History of the Concept of Sovereignty

1.2.1 On the Political Concept of Sovereignty (1500–1750)

Sovereignty is a distinctly modern concept, which was described – still unnamed – by Niccolò Machiavelli in 1513 in his *‘Il Principe’*. The focus of his ideas is no longer the Christian *‘bonum’*, but the Pagan *‘potestum’*: it is no longer about the (medieval) ‘Good’ that can be found, but about the rational-strategic institutionalisation of ‘Power’. This paradigm shift is central to the first author

2 Just a thought experiment: European antiquity (Greece and Rome) lasted 2,000 years, the European Middle Ages lasted 1,000 years, and the European modern era lasted 500 years. Which era have we been living in since the turn of the millennium?

of the modern era to coin the term 'sovereignty'. But that's not all! In his main political work, *'Discorsi'*, Machiavelli developed a very modern concept of the Republic, taking up the basic ideas of the Roman Republic³ and transforming them for the Modern Age. This work also clearly describes the concept of sovereignty, even if the term itself has not yet been introduced. Machiavelli describes the republic as being supported by the people and shaped by the constant conflict between the elites and the people. Another factor already addressed by Machiavelli is the willingness and ability of citizens to defend themselves with weapons in hand. Machiavelli therefore understands sovereignty very pragmatically as a concept of empowerment of the people. What Machiavelli no longer addressed at all was the question of justice within the state and of the state itself. In this respect, he differs fundamentally from Augustine, who, 1,000 years before Machiavelli, had established precisely this connection as constitutive of a 'state'. Machiavelli was thus already a pioneer of the intellectual developments of the 18th century, albeit without an ethical foundation. This was to become indispensable for German Idealism (esp. J.G.Fichte).

Jean Bodin, about half a century later than Machiavelli, introduced the term "Sovereignty" in his major work *'Les six livres de la République'* in 1576: The ruler is sovereign because – according to Bodin – as God's governor on earth, he deserves absolute obedience: he has intrinsically unlimited legislative power. The only authorities to which he is subject are the will of God and divine natural law.⁴ Bodin therefore defines sovereignty: *'La souveraineté est la puissance absolue et perpétuelle d'une République.'* This power is to be understood both internally and externally. In addition: just like Machiavelli, Bodin was not a 'monarchist'; for him, the people (in a republic) can also be the sovereign. The central point is that sovereignty means undivided power and its exercise. This is also a criticism of the doctrine of monarchy, the right to resist and even tyrannicide. Nor should we forget the traumatic event of St Bartholomew's Day (1572), therefore – similar to later Thomas Hobbes – Bodin argued that a right to resist would cause constant uprisings and instability. His

3 Like its ancient Greek predecessors, the Roman Republic was also a pagan state that owed its stability largely to polytheism.

4 Although this concept may still sound 'medieval', it already lacks (unlike the famous Spanish 'Magna Charta Libertatum' of 1188) any institutional oversight of this conformity or lack of contradiction. The modern era is dawning and cannot (or will not) conceal its ancient pagan heritage.

central idea was that stability of the state is more important than the right to resist. Sovereignty consists – according to Bodin – of the following criteria:

1. Absolute (supreme power); Sovereignty is the highest authority in the state. There is no *secular authority above the sovereign*.
2. Permanent (perpetual); Sovereignty is not limited in time. The sovereign possesses power permanently.
3. Indivisible, that is, Sovereignty cannot be divided between several authorities – checks and balances do not fit into Bodin's concept.
4. Legislative power: According to Bodin, the most important concrete sign of sovereignty is the power to make laws without the consent of another.

Although this concept of sovereignty is quite abstract, it stands to reason that Bodin expected the monarchy to be the best embodiment of these criteria. However, in principle, a republic can also be the bearer of these principles of sovereignty. In this sense, some 200 years later, Jean-Jacques Rousseau argued for what was, in effect, absolute republican *popular sovereignty*, which in his radical interpretation led to the French Revolution.

Thomas Hobbes also presented a pioneering concept of sovereignty in his major work '*Leviathan*' in 1651. Like Bodin, it is central to Hobbes that sovereignty is undivided and unified, that it is absolute and therefore tolerates no contradiction. The main difference is that Hobbes no longer legitimises sovereignty theologically ('representative of God'), but purely politically – in the form of a 'social contract': the people conclude a social contract in the name of security and peace in order to transfer all necessary power to the – artificially created – state '*Leviathan*' so that it can guarantee peace and security for the people. The alternative – the sovereignty-less state of nature '*Behemoth*' – must be overcome. Therefore, according to Hobbes, the state has a very clear legitimising task: to end the "war of all against all" and to ensure peace and order. This is achieved through its indivisible sovereignty and the associated monopoly on the use of force. Although Hobbes no longer justifies sovereignty on religious grounds, his concept still reflects the idea of (divine?) omnipotence as the basis or even prerequisite for security, peace, and order.

With Bodin and Hobbes, we have two theories of sovereignty that provided a (religiously inspired) response to (entirely religiously motivated) civil war situations in France and England, respectively. If we now consider John Locke's

concept of sovereignty, we should take a look back at Machiavelli, whose pagan Roman-inspired idea of sovereignty (based on the golden years of the Roman Republic) shows a certain degree of tolerance, as we will see in a moment.

In his major work *“Two Treatises of Government”* from 1689, John Locke set out his concept of sovereignty: he draws on the medieval idea that rulers have *executive* rather than *legislative* power: the people have natural rights (life, liberty, property) that the state must not infringe under any circumstances. The law and thus the power emanates from the people, the state has executive power. In order to argue this division of sovereignty, Locke developed the concept of the ‘separation of powers’: *legislative* power with the people, *executive* power with the state, the ‘limited’ sovereign. Thus political liberalism was born, the concept of power that diametrically contradicts the absolute concept of sovereignty since esp. Bodin and Hobbes. Machiavelli learned from ancient Rome that constitutionally institutionalised tension management can be beneficial and conducive to a political community. Total civil war on blood-soaked streets is just as reprehensible as a unified dictatorial state of oppression. John Locke took a mediating, reasonable position here, drawing demonstrably on the Catholic philosophers of the Spanish “School of Salamanca”. It is an ancient legacy of Spanish political philosophy (since 1188) that every political king needs a spiritual authority above him that judges whether his earthly actions are consistent with divine natural law – and thus decides on the legitimacy of monarchical actions.

Jean-Jacques Rousseau radicalised the notion of sovereignty even further, referring back to Bodin and Hobbes: the entire sovereign power, the ‘*volonté générale*’ lies now with “the people” – according to Rousseau in his main political work: *“Du contrat social”* from 1762. Sovereignty is again undivided here, no checks and balances⁵, it is in a certain sense ‘absolute’, because it is based on the one ‘*volonté générale*’, which never errs and constitutes the

5 A concept that can be traced from Plato’s *“Nomoi”* through John Locke and Montesquieu to current constitutional theory. Rousseau contrasted this with what he called the irresistible dictatorship of the people’s reason – a concept that played a central role not only in the form of Robespierre in France, but also in the Bolshevik Revolution, among others, especially when the people were regarded as being represented in a supernatural way by the “Jacobins” or “Bolsheviks.” Here, Rousseau was even more cautious and spoke conceptually in favour of direct popular sovereignty – but his successors turned this into the “dictatorship of the party.”

core of the 'social contract'. One could therefore say that Bodin's traditional understanding – the absolute sovereignty of the ruler – has been transformed in Rousseau into the equally absolute sovereignty of the people, which ultimately leads to the bloody concept of the 'religion civile' (*religion politique*), which demands unconditional obedience on pain of execution (Jacobinism). With Rousseau, all the 'medieval' barriers to the power of the people were thus torn down – thereby paving the way, intellectually speaking, for the totalitarianisms of his time – something Nietzsche had already warned against: 'If God is dead, then everything is permitted'. The masses can thus begin their triumphant march and transform history into a raging torrent of blood. Anticipating this development, the philosophers of German Idealism set out to counter it, at least in intellectual and argumentative terms – but as Nietzsche would later remark: 'A fine waste'.⁶

1.2.2 On the Ethical Concept of Sovereignty (1750–1850)

In the course of his 'Copernican Revolution', Immanuel Kant also placed the concept of sovereignty on a new – post-Rousseauist – footing: it is no longer a question of negotiating heteronomy (state vs. subjects), but of legitimising sovereignty through autonomy: it is a question – both for the individual and the state – of self-legislation, of self-determination, of freedom. We can therefore recognise in Kant's '*Perpetual Peace*' in 1795 and in the '*Metaphysics of Morals*' in 1797 that a rational republic is based on the autonomous cooperation of rational beings, with the state serving as an institution for the phenomenal facilitation of people's noumenal freedom. This is where it finds its legitimising foundation and also its limits. Immanuel Kant's central contribution to modern sovereignty theory consists of the following points: on the one hand, Kant deepened the social contract from superficial political utility into moral obligation. It is a dictate of practical reason that the coexistence of free, self-determining rational beings must be based on a contract that does not negate the foundations of its validity. Within the scope of this contract, the contracting parties must remain free and equal. This can only be ensured if this contract does not provide for the rule of persons, but of laws; Kant thus returns – as Montesquieu did before him – to Plato's '*Nomoi*'. These laws are enacted by representatives of the people for the people – here we see the great departure from Rousseau, who asserted direct popular sovereignty, which

⁶ In (better) German original: "Ein schönes Umsonst."

then – without moral restraint – tipped over into the terror of Jacobin parties, something that can still be observed today in communist states.

Johann Gottlieb Fichte is a particularly exciting author in this context because his 'ethical idealism' had a massive influence on his metaphysical and political thinking (1796/97: *Foundations of Natural Law*). Fichte understands law – and thus the state – not only as an ex post guarantor of noumenal freedom (as Kant had understood it), but as a transcendental condition of the possibility of *Being* in general: the I cannot exist without the sphere of mutual recognition of free beings, a sphere that must be realised as a state constituted by Natural Law. However, such a constitutional state (which Fichte still conceives as a *social contract*) is not possible without its transcendental condition, namely the morality of the free beings that constitute it. This morality, in turn, has its origin in an original freedom that has – in turn – its source in God, i.e. in the Absolute. Fichte thus combines the moral substance that Kant understood as 'freedom in itself' with Hobbes' existential need for security. The Legal State requires morality in order to make existence possible at all. Compared to his predecessors Bodin, Hobbes and Rousseau, Fichte undergoes a radical change of perspective (which Kant had already prepared but not implemented): for him, sovereignty is not the power to do whatever one wants, but the guarantee of the legal space in which freedom is possible in the first place. Just as law is metaphysically justified as a necessity of being, sovereignty is just as radically justified as a necessity of real law (space). For him, this is the core of 'political sovereignty' as he understood it in 1796/97 in his natural law doctrine. In *'The Closed Commercial State'* of 1800, he goes one step further and conceives of 'economic sovereignty' as a condition for the possibility of successful 'political sovereignty'. Ultimately, both economic and political sovereignty serve the sole purpose of securing human freedom. Fichte thus formulated essential elements of modern sovereignty theory as early as 1800.

Friedrich Wilhelm Joseph Schelling, in his main political treatise *'Philosophical Investigations on the Essence of Human Freedom'* from 1809 (in disagreement with his teacher Fichte) did not declare the legal relationship, but freedom directly as the ontological condition of the possibility of *Being*: 'Man is only free when he knows himself in his actions as the reason for his actions.' Moreover, later in his *'Philosophy of Revelation'* (1831/1841), Schelling explains: 'It is not the state that is the first thing, but freedom – the free being from which the state springs.' For Schelling, sovereignty is thus derived from freedom, or to

put it another way, since freedom is metaphysically founded for Schelling: the sovereignty of the state is ultimately a metaphysically founded power, not politically, not liberally, not morally, not institutionally, not legally, but directly metaphysically founded. The radical nature of this metaphysical justification of statehood is reminiscent of Plato's *'Politeia'*, when the aim was to institutionalise the 'Idea of the Good'. Schelling's radical concept of freedom was to have a decisive influence on the doctrine of sovereignty and also on political theory in general in the late 19th century (political romanticism): whereas Fichte still based state order on social contract theory and thus on rational grounds, Schelling went beyond this and based state sovereignty on human freedom as a functional derivative. This led Schelling to reduce the sovereignty of the state to a mere functional servant of human freedom, thereby completely instrumentalising the exercise of state power. In contrast to Fichte, Schelling also placed mystical freedom before state order in terms of logical justification. Fichte had still based law, and thus the state, on metaphysics and had also captured freedom ontologically as a manifestation of being. For Schelling, the state – and thus the law – is completely subject to freedom and therefore subordinate to it. Here, echoes of Plato's *'Politeia'* and Augustine's *'De Civitate Dei'* can be discerned.

Georg Wilhelm Friedrich Hegel – as in almost his entire oeuvre – took up and refined central ideas of Fichte and Schelling. This is particularly visible in the area of state sovereignty: while Fichte still speaks of the liberal constitutional state as requiring morality as a transcendental enabling condition, Hegel goes one step further (1820: *'Elements of the Philosophy of Right'*) and argues that the true state must be a 'moral state', a state that rests in and institutionalises established morality. This approach points to Schelling's influence. Unlike Fichte and Hegel, Schelling did not develop a fully formulated theory of the state, but rather formulated a radical critique of the state, criticising statehood in meta-legal and thus also meta-rational terms, in an anti-Enlightenment, romantic manner, as it were. Hegel thus goes so far back in tradition that he re-arrives at Bodin, who for his part regarded sovereignty as anchored in the Catholic and not Enlightenment-tradition: 'The state is the reality of the moral idea – the moral spirit as the manifest, self-revealing, substantial will.' It is not difficult to recognise the medieval concept of the self-revealing 'divine' in the 'moral spirit'. Nevertheless, Hegel is the Aristotelian of the German Idealists, whereas Schelling is the Platonist: Hegel was unable to achieve his radicalism

because he attributed a substantial morality to the established tradition, which was inconceivable for Schelling, who referred to a Patristic order of freedom.

1.2.3 On the Constitutional Concept of Sovereignty (1850–1950)

It is well known that the impressive systems of German Idealism collapsed around 1850, leading to a step back towards the 'political reality' of sovereignty, which consisted in the development of Constitutionalism. We would like to briefly mention and introduce the most relevant authors:

Hans Kelsen: Sovereignty in the *'Pure Theory of Law'*: 'Sovereign is whoever has the highest legally established norm-setting competence'; according to art. 1 B-VG (Austrian Constitution), the people are therefore the sovereign. For Kelsen, the state is identical with the legal system. There can therefore only be such a thing as 'sovereignty' within the positivist system of norms, within the constitutional order. Kelsen therefore attempts to tame the realpolitik power problem through positive-normative regulation. This is why Kelsen states in his *'General Theory of State'* in 1925, "The idea of a super-positive, extra-legal sovereignty contradicts the *'Pure Theory of Law'*." Kelsen thus has an understanding of sovereignty that is neither political, nor moral, nor even metaphysical. His sovereign is to be found in the people, who can exercise their 'supreme norm-setting competence' without any restrictions, as long as it takes place on the ground and within the framework of the constitution – cf. art. 44 (3) B-VG.

Carl Schmitt: Sovereignty in *'Political Theology'*: 'The sovereign is the one who decides on the state of exception', according to Schmitt in 1922. For Schmitt – as for Kelsen – the unrestricted decision is at the centre of the theory of sovereignty, but not the unrestricted expression of arbitrariness of the people on the basis of a value-free constitution, but the equally unrestricted expression of arbitrariness of the real politically powerful, on whose decision the law (or also the legal system) is subsequently based. Like Kelsen, Schmitt is also a positivist – although not a norm positivist, but a power positivist. Therefore, the sovereign cannot be bound by the law, which owes itself to his decision in the first place. In the light of his notorious theory of the political as a friend-foe relationship, it is to be understood that for Schmitt the *decision* made in fact therefore represents the essence of the political, and the law is only a function of the political; once again one must correct, as with

the most positive Kelsen: both Schmitt and Kelsen always speak of 'law', but mean the mere 'legislation' – and this legislation, both positivists agree, can have any content.

Rudolf Smend: Sovereignty in the *'Theory of Integration'*: Sovereignty is the expression of a constitution that proves to be a living order, according to Smend in *'Constitution and Constitutional Law'* (1928); it is not about abstract norm-setting that may be morally deduced or induced by power politics, but about the fact that law is the expression of a living moral-traditional order. Here one can sense the Aristotelian recourse to Hegel, for whom concrete morality as the essence of the 'State' and thus of sovereignty was also decisive. Characteristic of Smend, the 'creator' of the German *Grundgesetz* in 1949, was the idea that the constitution and its law represent more a process of constant integration than a fixed state of moral legality at a specific point in time. Smend's thinking thus contains a very modern, if not cutting-edge momentum: sovereignty is understood as a momentum of a communicative social integration process.

If we look at the political development of the concept of sovereignty, it always revolves around one core concept – and that is the free decision over either others or oneself within the framework of a political community. This also gives rise to the subordinate distinction between 'internal' and 'external' sovereignty, if we look at it from an international rather than an internal political perspective. Nonetheless, these differentiations are only subordinate and do not differ significantly from the concept. At its core, it is always about free self-determination, even if the circle of those who make up the 'self' can differ substantially, depending on the organisation of the respective political order. In methodological terms, sovereignty always has to do with communication: as Aristotle already knew, man must communicate with his peers in order to make the *Logos* manifest. His teacher Plato, for whom communication ('dialectic') is the only path to truth, takes an even more radical view. Therefore, anyone who wants to take the first step towards sovereignty must be certain of himself as a rational being and must therefore communicate, be it with his peers or with his tradition. The development of the Will of the respective rational being is then based on this rational insight. In this respect, Kant has already stated in the *'Groundwork of the Metaphysics of Morals'* (1785) that there can be nothing nobler and better on earth than a 'Good Will'. Kant made a clear distinction between *wish* and *will*: the mere wish is worthless, the will

is the wish combined with the maximum possible effort to realise it. Whether an action is ultimately 'good' can hardly be answered for, but one can answer for having realised the good will to the best of one's ability. This is the core of autonomy, of responsible freedom – and thus, according to Kant, the core of human dignity.

Finally, we must refer back to another classic, Carl von Clausewitz, who said in *'On War'* (1832) that the essence of war is to *break the political will* of the enemy. In our context, this means that every war is directed against the political sovereignty of the enemy; it is about breaking their political will, their political sovereignty, i.e. destroying it. You have achieved a 'victory' when you have destroyed the enemy's sovereignty - this can be done by destroying its military, its economy, its infrastructure, or also by massively influencing public opinion and thus the political will of the respective community, which ultimately has democratic feedback. It must be clear: however chameleon-like war may appear, its essence is always the same: it is about breaking the political will, according to Clausewitz, who for his part is philosophically located between Kant and Fichte. Political sovereignty is thus inseparably linked to the political will and thus to the endeavour, according to Fichte, to defend this freedom and thus sovereignty with one's full strength, because it constitutes the essence of being a worthy human being.

At this point, we make a leap in time to the present: since our communication, but also our knowledge and our ideas and thus also our morally innermost being, namely our will, are now digitally mediated, we are standing on a precipice: if Arthur Schopenhauer correctly understood that our world presents itself as *'Will and Imagination'* (1819), then it is not just about political sovereignty and its mediation, it is also not about our sovereign decision for responsible freedom, but it is even our 'world' in general, which we can only want and thus know digitally to a limited extent. We need to understand this phenomenon from the ground up, a burden of the concept that our ancestors were unable to relieve us of. This modern momentum of sovereignty is problematised under the buzzword *'Digital Sovereignty'*, which refers to a different form of 'sovereignty': it is not the opposite of, or even a replacement for, the classical concept of sovereignty, but rather a further ongoing development of it, so that we should 'retrospectively' refer to classical sovereignty as *'Analogue Sovereignty'*.

1.3 On Notion and Nature of ‘Digital Sovereignty’

As already mentioned in the introduction, we must distinguish here between the *Oikos* and *Logos* dimensions of digital sovereignty in the information space. This is not to suggest that, in an ideal-typical sense, this distinction did not also exist in the traditional realm of analogue sovereignty – it was simply not so clearly differentiated, as the “infrastructural relevance” had not yet come to light so strikingly. Nowadays, however, modern information technology compels us to overcome the traditional complacency in this regard and to make a precise distinction. The following discussion will therefore first focus on the *Oikos* dimension, which is, after all, the subject of most intense debate and is often already misunderstood as the overarching phenomenon of ‘digital sovereignty’. We must therefore subsequently also address the *Logos* dimension of digital sovereignty.

Oikos-Dimension

The *Oikos*-Dimension of “Digital sovereignty” refers to the ability of individuals, organisations, or states to use, shape, and control the digital technologies, data, and infrastructures on which they depend – independently, competently, and self-determinedly. It concerns hardware as well as software and can be understood on **three different levels**:

1. Individual Digital Sovereignty

The ability to use digital technologies **consciously, securely, and critically**. This includes data protection, media literacy, control over one’s own data (“data ownership”), and understanding how algorithms shape experience. This means your personal right to control your digital footprint. It’s often called **Self-Sovereign Identity**.

- **Data Privacy:** Knowing exactly who has your data and being able to delete it.
- **Portability:** The ability to move your photos, contacts, and files from one platform to another without being “locked in.”
- **Autonomy:** Not being manipulated by algorithms you do not understand or did not consent to.

2. Organisational Digital Sovereignty

The capacity of companies, but similarly of schools or public administrations to **avoid complete dependency** on external digital providers. That involves control over IT infrastructure, software, and data storage, the use of open standards and open-source solutions and the ability to operate or migrate digital services independently. For companies in particular, this is about risk management and independence from “Big Tech.”

- **Cloud Independence:** Ensuring that if a service provider (like AWS or Microsoft) changes their terms or prices, the business is not paralysed.
- **Open Source:** Using software that the company can inspect, modify, and run locally.
- **Intellectual Property:** Keeping proprietary secrets safe from the platforms they are hosted on.

3. State Digital Sovereignty

This is the most debated level and concerns a state's or supranational entity's (like the EU) ability to **develop, regulate, and protect critical digital infrastructures and technologies independently**. That includes the control over data spaces, AI models, and network infrastructure, legal frameworks (e.g., GDPR, Data Governance Act) and **strategic autonomy in global technology policy**. It is the idea that a country should be able to govern the digital world within its borders, just as it does the physical world.

- **Data Residency:** Laws requiring that a citizen's data stays on servers located within the country (e.g., GDPR in Europe).
- **Technological Independence:** Reducing reliance on foreign technology (like chips from the US or 5G tech from China) to avoid political leverage.
- **Digital Infrastructure:** Owning the cables, satellites, and data centers that keep the country running.

For further argumentation, we will stress this community- or state-aspect. In the past, the internet was treated like a global “Wild West” with no borders. However, as digital life became *real* life, classical power has evolved and shifted heavily toward a few massive corporations and a couple of powerful nations. The naive dream of “communication free of domination” was over – the 1968 revolution was devoured by its own grandchildren.

Reflecting on the digital world of today, we have to admit that it contains, among other things, two dimensions of “power” which shape the open space

for “Digital Sovereignty”: the power of – illiberal – states and the power of – illiberal – companies. In this context, “illiberal” means entities that follow or respect neither the Kantian⁷ nor the Fichtean⁸ notion of Freedom. As we mentioned before: the common “notion” of Sovereignty is freedom, liberty, autonomy – or in short: self-determination. The EU was originally founded in order to safeguard exactly this “self-determination” of the ‘European way of life’.⁹ In the 20th century, the European integration was only about political and economic self-determination. In the 21st century, a new dimension was added: the need for digital self-determination. Therefore, it is not a surprise that the EU in particular is working on the topic of “Digital Sovereignty”¹⁰: In the European Union, “Digital Sovereignty” is no longer just a theoretical concept – it is a central pillar of law and industrial policy.

As of **2026**, the focus has shifted from only setting rules to also building capacity. Here is how it is defined and applied in this specific context of the concept of “Open Strategic Autonomy”. The EU defines digital sovereignty as the ability to act independently in the digital world while staying open to global trade. It is driven by the fear of being caught between the US (software/cloud dominance by private big companies) and China (hardware/infrastructure dominance by state controlled big companies). Therefore, the EU has identified three spheres of action:

- **Concerning Rules:** Besides the wellknown GDPR, the landscape now includes the AI Act (regulating artificial intelligence based on risk) and the Data Act, which forces companies to allow users to port their data away from proprietary “walled gardens”.
- **Concerning Infrastructure:** The EU is currently moving toward “Hard Sovereignty.” Initiatives like the EURO-3C Project (launched early 2026) aim to create a federated European cloud and 5G/6G network to reduce reliance on non-EU providers¹¹.

7 Based on his concept of Categorical imperative (1785)

8 Based on his concept of Mutual recognition (1796)

9 This was the founding idea of Richard von Coudenhove-Kalergi, which he expressed, in his prophetic text: Pan-Europa from 1923.

10 <https://digital-strategy.ec.europa.eu/en>

11 <https://digital-strategy.ec.europa.eu/en/news/commission-announces-eu75-million-euro-3c-project-build-federated-telco-edge-cloud-infrastructure>

- **Concerning cloud sovereignty:** This is a key term in 2026. It refers to cloud services where the data, the operators, and the legal jurisdiction are all strictly within the EU (e.g., the AWS European Sovereign Cloud).
- The **EU Cloud and AI Development Act** (proposed in Q1 2026) is the latest attempt to harmonise these goals.

Logos-Dimension

As mentioned above, however, it is now also important to address the *Logos* dimension of ‘digital sovereignty’, which concerns the spirit, the will, the ideas, the narratives, the fears, the motivations, the images, that are triggered or formed within the information space.

Digital Sovereignty must – in its *Logos*-Perspective – therefore be understood just as comprehensively as it corresponds to its nature: on the one hand, it is about *sovereignty of action*, but also about *sovereignty of knowledge* and thus ultimately *sovereignty of will*:

Sovereignty of Action: This is about the ability to handle digital instruments autonomously and therefore to be able to act. It is about independence in the field of hardware and, above all, software. This aspect represents, as it were, the transition from the *Oikos* to the *Logos* dimension of sovereignty, the focus is no longer (merely) on the existence of hardware and software, but on their relevance to action; in doing so, we move beyond the pure *Oikos* realm and come to view it as a kind of condition of possibility for action. The US is dominant in this area¹² and EU Europe is only just beginning to develop¹³ ‘instrumental sovereignty’, for example by overcoming the monopoly-like dominance of Microsoft, Meta and Google. However, this question also affects

12 See for instance: “Europe’s digital ecosystem remains heavily dependent on non-EU software and cloud providers. This study maps these dependencies, as well as the geopolitical and economic risks they raise. It finds that US firms dominate all major software layers, exposing Europe to strategic vulnerabilities.” https://www.europarl.europa.eu/thinktank/en/document/ECTI_STU%282025%29778576

13 A key driver behind this EU-led European engagement is the war in Ukraine and the resulting geostrategic upheavals – including the potential disintegration of NATO. Suddenly, it becomes clear that supply chains and markets must be assessed not only in terms of profit but also from a geostrategic perspective. The same applies to, for instance, the energy supply, any area of relevance to security policy in the event of a crisis or conflict that can be deployed as instruments of sovereignty relevance.

all digital products, such as control elements for all areas of life that we have 'digitalised' – the pressing question must be answered everywhere: Who controls these systems and therefore our ability to act? Consider the work: Barnett, *'The New Pentagon Map'*, already from 2005: a distinction is made between the 'connected world' and the 'unconnected world' – the 'connected world' no longer requires conventional kinetic means of combat – a digital shutdown from a distance is sufficient to break the sovereignty of a state. The prerequisite for the possibility of action is knowledge... without knowledge, there can be no action, only behaviour. We will therefore now turn our attention to the phenomenon of knowledge.

Sovereignty of Knowledge: This is about the ability to inform oneself autonomously and adequately about events in the present and the past. What has happened, what is actually the case. How can the apparent facts be appropriately contextualised so that they can also be 'understood'? For example, how can we correctly assess Russian policy over the last 150 years if we do not take into account the 'Crimean War' of 1850? It is about the narrative, about the story, about the creation of meaning, so that we can know what is happening at all, so that we are enlightened. Both Russia and the US are global dominant actors in this area of narrative control. It is therefore a matter of regaining 'cognitive sovereignty' in the information space, which is currently massively digitalised. The importance of knowledge in this context of sovereignty cannot be overstated: if knowledge is lacking or distorted, if narratives contradict one another, or if citizens have no guiding principles whatsoever to enable them to assess the flood of incoming information independently in any way, if all this is lacking or only exists in a rudimentary form, one – for example, as an enemy – can be virtually certain that the community under attack will be unable to mount any relevant defensive action. So, without knowledge, there can be no defence. If, therefore, one reduces or diminishes the knowledge base, or undermines the population's hermeneutic media literacy, then such a community is an easy victim; it does not even realise that it should be defending itself right now. However, what is the prerequisite for knowing something, for learning something, or for trying to understand it? It may well be that a state brings the hardware and software under its control (which is what the EU is currently doing); it may well be that the education system begins to present content relevant to the state, politics and security in a way that is both didactically sound and critical of the media (insofar as this is not already happening, but that is another matter), but something crucial is still

missing and a potential enemy knows this too (at least if they have studied Schopenhauer): it is all about the will. This is the be-all and end-all of the whole matter, so let us now turn our attention to this will:

Sovereignty of Will: This is about our moral integrity, about what we should actually *want* based on our instrumental possibilities and our cognitive level of knowledge. This will is decisive for the question of victory or defeat, is at the core of any meaningful security policy, because there can be no knowledge without will, and no action without knowledge – and those who do not act do not defend themselves and have already lost themselves, their freedom and, with it, their sovereignty. The central Clausewitzian question in *‘Security and War’* is always whether it is possible to break the political will. However, what can we still authentically will if we have already forfeited our sovereignty, our independence from foreign power, that is, our autonomy, partly in the realm of instruments and almost entirely in the realm of information?

The will, according to Kant, is a desire that is realised with all available means. What happens when we have hardly any means at our disposal and our desire can hardly be grasped because it has been distorted by manipulated information? Here our sovereignty is being undermined and eroded on three levels, in truth, a war is being waged in the sense of Clausewitz, which is operating ‘hybrid’ on several levels simultaneously: at the level of instruments (which we no longer have under our control, even if we should still have wishes), at the level of information (which makes it almost impossible for us to formulate adequate wishes), – and thus at the level of integrity, of the will that has been broken before it could even be conceived – just think of the problem of the will to defend oneself in EU-Europe – one is defeated before the war has begun, which is very reminiscent of Sun Tzu, who said: *“1. The best victory is to win without fighting” and “10. The highest in the art of war is to subdue the enemy without giving him the fight”* which brings us to China, the *“Smiling Third”* in global power politics. This finding has been highly topical over the last 5–10 years; indeed, as early as the war in Ukraine (which, in truth, has been in a heated phase since 2014, albeit not yet predominantly kinetic as of 2022), the reliability of the transatlantic security architecture was called into question by the US’s initially hesitant support for Ukraine. Recent events in 2025 and 2026 under the Trump administration mean that we must already strategically anticipate the end of this transatlantic security architecture. In other words, the US and the EU are drifting apart, or rather

the EU is attempting to emancipate itself from the US hegemony – and this creates a situation of at least three poles: US – EU – China.

To better understand the developments and shifts in US policy, it is necessary to engage with the thinking of the philosopher who, alongside Leo Strauss and Carl Schmitt, is currently the most influential figure within the so-called US “establishment”: René Girard, the French cultural philosopher who taught and conducted research at Stanford University from 1981 until his death in 2015. Why is René Girard so crucial, both in terms of his personal background and his theoretical arguments, when considering sovereignty in the information sphere? His mimetic theory can also be understood as a theory of information and interaction, which establishes a clear *Logos* link to our overarching theme – the question of the will and its formation.

1.4 On René Girard and his Mimetic Concept of Sovereignty

René Girard's theory of mimetic desire posits that human desires are not autonomous or innate but are fundamentally *imitative*: people desire objects, status, or experiences primarily because they observe others – whom Girard calls “models” – desiring them. This imitation, or mimesis, creates a triangular relationship between the subject (the desirer), the model (the one being imitated), and the object of desire. For instance, a person might want a particular car not for its intrinsic qualities but because an admired figure seems to value it, leading to unconscious borrowing of desires from social surroundings. This process often escalates into rivalry, as the subject and model compete for the same thing, fostering envy, conflict, and a cycle of escalating tension within groups. Girard argued that this mimetic nature explains much of human behaviour, from personal relationships to cultural formations, and it underpins broader societal dynamics such as violence and the need for resolution mechanisms.

Extending beyond individual psychology, Girard's theory incorporates the “scapegoat mechanism” as a way societies manage the crises born from unchecked mimetic rivalry. When rivalries proliferate and threaten to dissolve social cohesion – creating a “mimetic crisis” of undifferentiated violence –

the group redirects its aggression onto a single victim, often an outsider or marginalised figure, whom they collectively blame and expel or sacrifice. This act restores peace by unifying the community against a common enemy and the victim is later mythologised as both monstrous (causing the chaos) and sacred (bringing harmony through their removal). Girard saw this as the origin of religion, rituals, and myths, where prohibitions (like taboos) prevent mimetic escalation, and sacrifices reenact the foundational violence to maintain order.

On Sovereignty in the Shadow of Mimetic Desire

Girard's mimetic framework reframes political power and authority as deeply intertwined with the control of mimetic violence and the scapegoat process. Sovereignty, in this view, emerges from the need to monopolise and ritualise violence to prevent chaotic rivalries from destroying the social fabric – echoing but critiquing thinkers like Hobbes, who saw the sovereign as ending the “war of all against all”. The sovereign figure often embodies the ambivalence of the scapegoat: revered as a god-like unifier who channels collective desires and rivalries, yet always risks becoming the next victim if they fail to maintain order. In archaic societies, kings were sometimes sacrificial figures, substitutes for the original scapegoat, blending power with vulnerability.

In modern contexts, sovereignty manifests through legal and political systems that secularise this mechanism – laws prohibit unchecked mimesis (e.g., by regulating competition), while the state acts as a “cathartic” ritualiser, identifying and punishing scapegoats (criminals, enemies) to provide societal release without remystifying violence. This protects the sovereign from direct sacrifice but permits the scapegoating of outsiders, as seen in phenomena such as nationalism or pogroms. Girard's theory challenges modern notions of autonomous sovereignty by revealing it as contingent on mimetic dynamics, where true escape from rivalry requires transcending imitation through non-violent models (e.g., in Christian ethics, which Girard saw as exposing scapegoating's injustice).

Overall, mimetic desire underscores sovereignty's fragility: it is not an absolute, rational construct but a precarious balance against the ever-present threat of mimetic collapse, influencing discussions in political philosophy on how power sustains itself amid human imitation and conflict.

On Digital Sovereignty in the Shadow of Mimetic Desire

When viewed through the lens of René Girard's mimetic theory – which posits that desires are not innate but imitated from models, often escalating into rivalry and social crises – digital sovereignty appears precarious and shadowed by amplified mimetic dynamics in online spaces. Algorithms and platforms accelerate mimesis by curating content that encourages imitation of desires, from consumer trends to ideological stances, eroding individual and collective control over one's desires and digital identity. Social media, for example, fosters a “crisis of sameness” where users are homogenised through standardised profiles and constant exposure to others' curated lives, blurring distinctions and intensifying rivalry over status, attention, or validation. This mimetic contagion, driven by features such as infinite scrolling and algorithmic recommendations, commodifies relationships and desires, turning users into unwitting imitators in a flattened hierarchy that risks broader societal fragmentation without shared models or truths to anchor them. The result is a subtle loss of personal sovereignty, as desires become externally mediated by tech platforms.

On a collective level, mimetic theory illuminates how digital spaces precipitate crises akin to Girard's scapegoat mechanism, where escalating rivalries – such as online swarms, cancel culture, or ideological echo chambers – culminate in the ritualistic targeting of individuals or groups to restore temporary unity. In the internet age, algorithms perpetuate these patterns by amplifying viral imitation, from memes to outrage cycles, potentially leading to apocalyptic escalations if unchecked. This ties directly to sovereignty challenges, as states and societies grapple with mimetic influences that transcend borders, eroding national control over narratives and behaviours. Big tech companies exacerbate this by engaging in “mimetic diplomacy”, imitating state diplomatic practices – such as establishing “permanent representations” to international bodies or proposing accords like *Microsoft's Digital Geneva Convention* – to influence regulations and assert authority in digital governance. This emulation blurs lines between corporate and state power, complicating efforts to achieve digital sovereignty, as tech giants replicate sovereign-like behaviours to shape rules on data, cybersecurity, and multilateralism.

1.5 Polemological Epilogue

In conclusion, one must not overlook the radical paradigm shift taking place within our topic – the digital sovereignty in the information space in the shadow of René Girard: the key players in this mimetic formation are private or non-state actors. Our previous – modern – concepts of sovereignty, by contrast, were state-centred (since Machiavelli, who, after all, also ‘nationalised’ war and took it out of the hands of private mercenaries). Are we now facing a fateful epochal break (Friedrich Merz) that runs far deeper than a mere turning of the times (Olaf Scholz)? Will the mimetic manipulation of the will – and thus of knowledge – ultimately shatter sovereignty – achieving a ‘victory’, as it were, by non-kinetic means, in a war that cannot even be grasped in analogue terms? It is therefore necessary, in order to protect analogue sovereignty – which is referred to, among other places, in the Federal Constitutional Act and in the Defence Act – to take action in the realm of digital sovereignty.

Whilst in the analogue age it was reasonable and prudent to defend sovereignty when it was under attack, in times of *polemos* in the information space, a wait-and-see approach is a form of capitulation. We must take a proactive approach here (strengthening resilience), for we are operating in the realm of the transcendental, if the conditions enabling analogue sovereignty in the information space are eliminated, it will be too late to begin defending against the ensuing attack on ‘les misérables’, the broken wills, there will then be nothing left to defend. Therefore, we must start much earlier here, in the defence of the sovereignty of the will – a very profound rationale for a “Geistige Landesverteidigung” (National Defence of the Mind) conceived in this way.

This makes it clear that “Umfassende Landesverteidigung” remains highly relevant, but that, due to digitalisation and the vast scope of the information landscape, ‘Geistige Landesverteidigung’ has undergone an enormous resurgence in importance, far exceeding its mere classification as a sub-section of the school subject ‘Civics’. It is time to recognise the signs of the times.



2

Matthias Postl:

The Information Environment

A Comparative Analysis
and Contextualisation of
Existing Concepts



The Information Environment – A Comparative Analysis and Contextualisation of Existing Concepts

Matthias Postl is a research associate at the Cyber Documentation and Research Centre of the Central Documentation Department at the Austrian National Defence Academy. His research focuses on the analysis of systemic conflicts and their implications for the foreign and security policy of Austria and the European Union. Another area of focus is disinformation and its impact on democratic systems, as well as the use of open-source information gathering methods (OSINFO).

2.1 Introduction

In the digital age, the ability to structure perception, meaning and collective interpretation is becoming a central dimension of political power. Information conflicts, strategic narratives and algorithmically mediated visibility are increasingly shaping the conditions under which social reality is interpreted and political decisions are prepared. Against this backdrop, the concept of the ‘information space’ is gaining analytical and political significance, without, however, being uniformly defined.

The information space is understood differently in various academic and political contexts. Whilst technical perspectives link it closely to cyberspace and conceive of it as the infrastructural environment of digital systems, approaches from the social and communication sciences focus on the production, mediation and interpretation of meaning. The information space thus encompasses not only the technical prerequisites of communication, but in particular its content, effects and social embeddedness. It describes a framework in which

information takes effect as a process, as an object and as knowledge, and in which technical, social and cognitive dimensions overlap.

Despite this growing relevance, there is no consistent, cross-actor definition of the information space. Rather, political strategies and security policy concepts show that key international actors conceptualise and shape the information space in different ways. In political and strategic documents of the European Union, it is understood as a normatively structured space of democratic public sphere, which is to be stabilised through regulation, transparency and institutional safeguards. In contrast, NATO and the United States describe the so-called ‘information environment’ as an operational framework of influence in which perception, interpretation and behaviour can be specifically influenced. In the People’s Republic of China, the information space appears as a component of state security architecture and as a cognitive space of order, whilst in the Russian Federation a security policy perspective dominates, which conceives of the information space as an arena of geopolitical competition and cognitive influence.

Although the term is gaining central importance, there has so far been no systematic comparative presentation of these different concepts, nor an integrative theoretical classification. The information space is frequently used as an analytical given without sufficiently differentiating its various logics of meaning. This is where the present chapter begins, examining how key security policy actors – in particular the US, Russia, China, NATO and the European Union (EU) – define and shape “their” information space.

The aim is to systematically record and present the different conceptions of the information space held by the selected actors. This is intended to contribute to the conceptual clarification and **interdisciplinary grounding of the information space**.

2.1.1 Distinction from cyberspace

The information space must be analytically distinguished from cyberspace, even though the two are closely intertwined. Whilst cyberspace is primarily understood as a technical space of networked information technology systems, comprising infrastructure, networks and digital architectures, the information space refers to the broader framework of content, meanings and societal effects of communication. Cyberspace thus forms the infrastructural

basis upon which processes of information production, dissemination and interpretation unfold, and can be understood as a functional subcomponent of the information space (Hennecken, n.d., p. 5; Buchanan, 2020, pp. 23–25; Giles, 2021, pp. 12–14; European Commission, 2023, p. 8).

This distinction points to two complementary analytical perspectives: Technology oriented approaches focus on issues of IT security, system integrity and network architecture, whilst perspectives from the fields of communication and social sciences examine the dynamics of meaning making, information circulation and reception. However, these two levels cannot be viewed independently of one another. Technological infrastructures shape the conditions of communicative processes by influencing visibility, reach and forms of interaction, whilst social requirements and usage practices, in turn, influence the development and regulation of technological systems (Sørensen & Kocksch, 2019, pp. 45–47; Dolata, 2017, pp. 12–15; Barth et al., 2023, pp. 102–105).

The information space cannot therefore be reduced to cyberspace, but encompasses it as an infrastructural dimension within a broader, multi-layered framework of effects in which technological, social and cognitive processes are inextricably linked.

2.2 Method

At the start of the systematic review, following the approach of Brocke et al. (2009), abstracts in academic databases were screened to gain an initial overview of the research field and to derive relevant search terms. Relevant articles were then obtained in full text. The search terms were divided into primary and secondary keywords, with the latter comprising specific and thematically related terms.

As a rule, the theoretical foundation of academic work is based on primary academic literature (Döring & Bortz, 2016). Other types of sources, such as newspaper articles, internet sources or grey literature such as government reports, are included only sporadically due to the highly dynamic nature of the topic. Their use is based exclusively on a justified selection and careful examination of content.

To ensure full traceability, completeness and transparency, the standardised Prisma 2020 checklist for systematic literature reviews, developed by Page et al. (2021), was used to conduct the systematic literature search.

As part of the selection process, potentially relevant sources were first systematically recorded and duplicate entries removed. This was followed by an initial content screening based on title and abstract to exclude articles with no apparent relevance to the research question. In a further step, the remaining works were examined in full text and evaluated against defined criteria, with unsuitable studies being excluded and the respective reasons for exclusion stated. Finally, only those sources that met the defined requirements were included in the final analysis. The screening, examination and evaluation of the sources were carried out by the team at the Cyber Documentation and Research Centre of the National Defence Academy.

2.3 What is information?

Information is one of the key terms of the 21st century. We encounter it in everyday contexts, from data transmission on the internet and navigation systems to theories about knowledge, meaning and reality. Nevertheless, there is no single definition of what ‘information’ actually is.

This conceptual openness is no accident. Information is a phenomenon that can take various forms: as a physical medium such as a book or a data set, as the meaning people derive from messages, or as the effect communication produces in a social context. It thus combines material aspects (such as writing or code) with immaterial meanings (such as knowledge or intention). Consequently, information always exists in the tension between what is conveyed and what is understood.

2.3.1 Historical origins and interdisciplinary perspectives

The word ‘information’ derives from the Latin *informatio* and originally meant ‘formation’, ‘shaping’ or ‘instruction’. For Aristotle, *informatio* referred to the form given to a substance (*hyle*) – that is, the process through which something takes shape. Information was thus a formative force, not transmitted content.

In modern times, the term became increasingly restricted to the transmission of signs and messages – particularly in the wake of developments in printing technology, telegraphy and computer technology. Claude Shannon’s mathematical theory of information was pivotal in this regard. From this perspective, information is a quantifiable quantity: it reduces uncertainty when a signal is received – for example, how many bits are needed to transmit a message. The meaning of the message plays no role in this (Shannon & Weaver, 1949).

The philosopher and information ethicist Rafael Capurro, who has dedicated decades to the study of the cultural and philosophical content of the concept of information, criticises this purely technology focused perspective. He speaks of a ‘narrowing’ of the concept, which, whilst making sense in mechanical engineering, does not do justice to the human experience of meaning and interpretation (Capurro, 2005, p. 353). Instead, he proposes viewing the concept of ‘information’ in the context of its cultural and historical background (Capurro, 2024). For him, information is not simply ‘there’, but arises through a process of interpretation. He emphasises that information must not only be transmitted but also understood, and that this understanding depends on cultural, linguistic and social contexts (Capurro, 2005, p. 355). Capurro therefore describes information as a ‘process of meaning’ between humans, the world and language.

The philosopher Luciano Floridi, one of the most influential thinkers in the field of the philosophy of information, on the other hand, addresses the question of how information shapes our understanding of the world. In doing so, he makes a fundamental distinction between mere data and information. Whilst data can be isolated signs or values, it only becomes information when these data have a recognisable structure, can be meaningfully interpreted and actually say something true about the world. For Floridi, semantic information therefore consists of “well-structured, meaningful and true data” (Floridi, 2011, p. 260).

He also proposes rethinking the concept of knowledge. Knowledge is not merely a correct opinion with justification, but consists in the ability to understand, classify and explain information (Floridi, 2011, pp. 270–272). In his concept of *Informational Structural Realism* (ISR), he goes even further: the world consists, at its core, of structured informational relations – of patterns, states and relations (Floridi, 2011: 339–347).

In information science, too, the term ‘information’ is ambiguous and is understood there primarily in functional and practice-oriented terms. A particularly influential contribution in this regard comes from the information scientist **Michael Buckland**, who distinguishes between three different meanings of information:

- Information as a process: refers to the transmission of content, e.g., when someone sends a message and another person receives it.
- Information as knowledge: the understanding or insight that arises from information.
- Information as an object: refers to material or digital carriers of information such as books, files, data or artefacts (Buckland, 2017, pp. 22–24).

This three-way division shows that information **cannot be reduced to a single form**. It moves between **thing and meaning**, between **technical structure** and **social significance**. This ambiguity becomes particularly apparent in the digital age, where information can be **stored on a massive scale** and **transmitted in real time**, for example, when a text file (object) is accessed via the internet (process) and understood in the reader’s mind (knowledge).

At the same time, Buckland points out that information is **not informative in and of itself**: a document is only information if it **has meaning for someone**, that is, if it can be read and understood in the right context (Buckland, 2017, p. 26). In this way, his approach aligns with Capurro’s understanding of information as an interpretative process: information arises **not only through technology**, but also through the interplay of signs, context and understanding.

The social and cultural sciences take a similar view. Here, information is understood as a culturally interpreted, socially contextualised process of negotiation. Meaning arises only through language, media and social practices, that is, through what is shared and understood (Capurro, 2024).

Whilst many approaches view information as a sign or content, systems theory understands it as a driver of structural change. For Wolfgang Hofkirchner and Wolfgang Fleissner, information arises when a system – biological, technical or social – reacts to stimuli in such a way that its structure evolves (Hofkirchner et al., 1995, p. 284). Information is not a fixed content, but a process of self-organisation.

2.3.2 Conclusion and classification of the concept of information within the information space

Information is not a clearly defined entity, but a concept with many meanings. What all these perspectives have in common is that information arises where something takes effect: as a signal in technology, as structured meaning in philosophy, as a functional model in information science, as a new order in systems theory, and through interpretation and contextualisation in the social and cultural sciences.

In this sense, the information space can be understood as the domain in which information not only exists, but also becomes effective as a technical process, as the production of meaning, as a cultural practice, or as a systemic dynamic. It encompasses the material, symbolic and social dimensions of information processing and interpretation. The information space is thus not an abstract container for data, but a living space of action in which knowledge is created, shared, negotiated and transformed.

2.4 The EU information space – a classification

2.4.1 How the information space became a field of action

‘Information space’ is not a term defined in EU law. Nevertheless, within a few years, a clearly delineated policy area has emerged that addresses precisely this space. The EU information space is thus not a terminologically fixed concept, but a functionally developed one, and has arisen from political responses to digital transformation, disinformation and geopolitical influence.

It is striking that the EU itself uses different terms, such as ‘*information space*’, ‘*information environment*’ or ‘*data space*’, depending on whether security policy, analytical or economic-regulatory aspects are in the foreground. The terminological difference thus points to differing institutional perspectives at an early stage (Ertl, 2025).

2.4.2 Protecting the democratic public sphere as a starting point

A turning point came in 2018. With its ‘Action Plan against Disinformation’, the European Commission responded to coordinated manipulation campaigns,

particularly in the context of elections (European Commission, 2018). The initial focus was on rapid information dissemination, strategic communication and cooperation with platforms. The information space was viewed here as a vulnerable democratic communication space, threatened by targeted deception.

Two years later, the perspective shifted significantly. The “European Action Plan for Democracy” (European Commission, 2020) no longer viewed disinformation merely as an isolated problem of false content, but as a structural threat to democratic integrity. Transparency in political advertising, the protection of free elections and media pluralism were formulated as cornerstones of democratic order. The information space was thus no longer seen merely as a communication space worthy of protection, but as a space for discourse in which the political shaping of the European public sphere takes place.

Hannah Arendt already describes the concept of the ‘public sphere’ as a ‘space of appearance’ in which political actors become visible and collective action becomes possible. For her, the public sphere is not an abstract ideal, but a space of perception in which political reality can only emerge through visibility. Without this space, democracy cannot exist (Arendt, 1958, pp. 198–199).

Habermas expands this perspective to include the ‘conditions of rational deliberation’ (Habermas, 1992, pp. 349–398). Democratic legitimacy requires verifiable information, argumentative exchange and institutionalised criticism. The public sphere is thus not merely visibility, but also a process of reason-guided understanding.

The public sphere is not an abstract principle, but is mediated institutionally. Democratic visibility does not arise of its own accord, but through organised communication structures. Schrijvers et al. (2026, pp. 16–17) therefore describe the media as democratic infrastructure. They not only facilitate the transmission of information, but also structure collective perception and judgement-forming. If these structures are destabilised, the capacity for deliberation in democratic societies comes under pressure.

The ‘European Media Freedom Act’ (Regulation (EU) 2024/1083) translates this insight into EU law. Editorial independence, transparency of ownership structures and the protection of journalistic sources are harmonised across the EU. Here, media freedom appears not merely as an individual right of defence,

but as a systemic condition for the stability of democratic order and, at the same time, as an integral part of the single market.

If we apply these considerations at the European level, the EU information space appears to be a structural prerequisite for democratic decision-making. Measures such as transparency obligations for media and platforms, regulatory requirements for political advertising, or support for fact-checking initiatives do not serve merely to avert risks. Rather, they stabilise the epistemic foundations of public communication.

The 'information space' is thus, first and foremost, a normative arena for democratic legitimacy. It forms the institutionally framed environment in which political visibility, criticism and collective judgement become possible, and, as an institutionally secured structure underpinned by the key elements of a pluralistic media landscape, it enables the democratic public sphere.

It is precisely this structural dimension, however, that raises a fundamental question: under digital conditions, who decides what becomes visible in this space?

In parallel with the "European Action Plan for Democracy", the "European Data Strategy" developed the concept of common European data spaces (European Commission, 2020a). It became clear at this point, if not before: the information space is not only a space for political discourse, but also a space for data and infrastructure. Information does not circulate in a vacuum, but via platforms, cloud infrastructures and algorithmic systems.

2.4.3 Platform architecture as a normative structure

The Digital Services Act (Regulation (EU) 2022/2065) marked a further qualitative step. For the first time, the systemic risks posed by large social media platforms were addressed legally. The information space was now understood as an algorithmically structured architecture in which visibility, attention and economic incentives are technically mediated. The 'Data Governance Act' (Regulation (EU) 2022/868) and the 'Digital Decade 2030' policy programme (Decision (EU) 2022/2481) expanded this perspective to include issues of digital sovereignty and infrastructural autonomy.

With the weakening of traditional media due to the technological transformation of the gatekeeper function, driven by the rapid development of social

media platforms, the structure of the information space is once again undergoing a fundamental shift. The public sphere is no longer mediated exclusively through institutions, but is encoded through infrastructure. Digital platforms now form the supporting walls of this space: through recommendation logic, ranking algorithms and moderation rules, they determine what becomes visible and what remains in the background.

The gatekeeper theory offers a classic point of reference here. Lewin (1947, pp. 12–14) and White (1950, pp. 386–388) describe editorial teams as selection bodies that decide which information gains public relevance. Visibility has always been the result of journalistic, editorial assessment and normative criteria.

With platformisation, this power of selection is fundamentally shifting. Algorithmic systems are taking on gatekeeper functions, not through deliberative weighing up in accordance with Habermas' theory of communicative action (1981, pp. 385–386), but through 'data-driven' optimisation of attention. 'Visibility', as described by Arendt (1958, pp. 198–199), is not abolished, but technically organised.

This alters the structure of the information space. What was once institutionally mediated is now architecturally encoded. The order of the space shifts from editorial responsibility towards algorithmic structuring. The public sphere is no longer merely institutionally framed, but infrastructurally pre-structured.

Zuboff (2019, pp. 8–12) describes this constellation as surveillance capitalism, in which data extraction, behavioural prediction and the attention economy have a structure-forming effect. Political communication is thereby embedded in economic incentive systems. Under these conditions, polarisation is no coincidence, but a potential side effect of structural optimisation logics.

Lessig (1999, pp. 6–8) illustrates with his thesis 'Code is Law' that digital architecture has a normative effect. The rules of visibility are implemented not only legally but also technically. Architecture does not replace the law; it complements and, in part, reshapes it. Whoever shapes the architecture shapes the conditions of the political public sphere.

However, where architecture becomes politically effective, a new form of strategic vulnerability arises at the same time. The 'Digital Services Act'

(Regulation (EU) 2022/2065) intervenes precisely at this structural level. It regulates not primarily content, but systemic risks. Transparency obligations, risk assessments and audit mechanisms stabilise the architecture on which democratic public life depends. Platform regulation thus appears as an intervention in the very structure of the information space itself.

In this context, Deibert (2020, pp. 18–22) argues that digital infrastructures are not naturally occurring spaces, but rather power constellations that have evolved over time. Their political anchoring in democratic norms is therefore not a technical detail, but a central question of how modern societies are structured. European platform regulation can be understood as an attempt to place digital architecture once again at the service of public accountability.

The ‘information space’ is thus not a neutral area, but a constructed environment. Its stability depends not solely on content, but on the structural order of its walls, passageways and lines of sight.

Where the architecture of a space is structured, it can also be used strategically. The technical malleability of digital communication environments opens up opportunities for targeted intervention. The information space thus becomes not only a place of democratic public sphere, but also a field of strategic influence.

2.4.4 Hybrid logics of influence in the information space: Foreign Information Manipulation and Interference (FIMI)

Foreign Information Manipulation and Interference (FIMI) describes targeted influence on public discourse and political processes by foreign actors. This involves not only the dissemination of disinformation, but also the targeted distortion and manipulation of existing information in order to deepen social divisions and undermine trust in institutions. In contrast to traditional disinformation, which is often carried out by domestic actors, FIMI is characterised by targeted, strategic influence from outside. The aim of these manipulations is to jeopardise political stability, exacerbate social tensions and secure geopolitical advantages for the manipulating actor. FIMI can be carried out by various actors, including state bodies, intelligence services or state-affiliated organisations. Non-state groups such as business interest groups or extremist networks can also act as actors. Furthermore, many states make use of so-called proxy

organisations, which act as seemingly independent actors, to conceal their manipulation campaigns and avoid international repercussions (EEAS, 2024; German Federal Government, 2023; Huntress, 2024; ENISA, 2024).

With the development of the FIMI concept by the European External Action Service, the information space was finally explicitly defined as a geopolitically contested domain (EEAS, 2023; 2024; 2025). In this context, the terms ‘information space’ and ‘information environment’ dominate, describing the strategic vulnerability of the democratic public sphere. Manipulation is no longer understood merely as a “content-related” challenge, but as “coordinated strategic behaviour by external actors”. The EU’s Strategic Compass enshrined this perspective within the security policy framework of hybrid threats (Council of the EU, 2022).

Howard and Bradshaw use the concept of *computational propaganda* to analyse the targeted use of digital tools for political influence (Bradshaw & Howard, 2019, pp. 2–4; Howard, 2018, pp. 3–5). Automated accounts, data-driven targeting and algorithmic amplification make it possible to efficiently feed narratives into existing communication structures. Whyte & Etudo (2025, pp. 1–3) further demonstrate that cyber operations and narrative influence campaigns represent distinct but combinable logics. Technical infiltration and discursive manipulation are not separate spheres, but can reinforce one another.

However, this security policy framework is not itself neutral. Cavelti (2013, pp. 110–112) demonstrates that cybersecurity does not merely describe technical threat prevention, but is a political process of problem definition. Security discourses structure perceptions, prioritise risks and legitimise institutional consolidation. The European conception of FIMI is also an expression of such a political classification.

The strategic objective of such operations lies less in immediate persuasion than in long-term destabilisation. Robbins demonstrates that disinformation campaigns aim to undermine trust in institutions, deepen social divisions and generate uncertainty (Robbins, 2020, p. 32). At the same time, Berk (2025, pp. 2–3) relativises linear assumptions about impact and points to the complexity of pluralistic information environments. Impact rarely unfolds directly, but rather through a gradual shift in interpretative frameworks and structures of trust.

The information space thus proves to be a multi-layered structure of discourse, technology and strategy. Its architecture facilitates democratic public discourse, yet at the same time creates structural vulnerabilities. It is precisely this ambivalence that explains why the European Union is not merely responding through communication, but seeks to stabilise the space itself institutionally through resilience.

2.4.5 Resilience as an organisational principle

If the information space is both a space of possibility for democratic public sphere and a field of strategic influence, its stability becomes the focus of political shaping. The European Union responds to this not with ad hoc interventions, but with a structural logic of resilience.

Beck's (1986) concept of the risk society provides a theoretical framework for this. Modern politics is increasingly organised around the anticipation and management of systemic risks. Hazards are not merely dealt with after they have occurred, but are calculated preventively, processed institutionally and contained through regulation. Security becomes an ongoing task of organised precaution (Beck, pp. 25–30; Ertl, 2022, pp. 255–262).

The 'European Democracy Shield' (2025) institutionalises this logic within the context of the information space (European Commission, 2025). Early warning systems, coordination mechanisms and structured cooperation between Member States are intended to increase the resilience of digital communication structures. Here, resilience does not mean isolation, as propagated by other global actors, but rather the stabilisation of open structures under conditions of permanent vulnerability.

This approach differs fundamentally from sovereignty-centred notions of the state of exception. Whilst Schmitt (1932, pp. 26–27) conceives of political unity through the identification of enemies and the state of exception, the European Union relies on legal binding, transparency and institutional consolidation. Protection is not achieved by suspending norms, but by their systematic reinforcement.

Resilience thus becomes the organising principle of the information space. It is intended to 'remain open' without being structurally destabilised. Stability arises not through disconnection or closure, but through regulated resilience.

Yet stability alone is not enough. In a globally fragmented order, resilience becomes a question of infrastructural autonomy.

2.4.6 Digital sovereignty in a fragmented order

Resilience stabilises the space from within. Digital sovereignty determines who controls its supporting structures. In a globally networked information order, infrastructure is not a technical footnote, but a geopolitical resource.

Bratton describes the digital order as a layered architecture, ranging from cloud infrastructures through networks to the user level. Whoever controls these layers shapes the conditions of political agency. The information space is thus not merely a space for communication, but simultaneously an infrastructural power structure (Bratton, 2016, pp. 5–11).

The ‘European Data Strategy’ (European Commission, 2020a), the ‘Data Governance Act’ (Regulation (EU) 2022/868) and the ‘Digital Decade 2030’ policy programme (Decision (EU) 2022/2481) aim to strengthen this autonomy. Data spaces refer to regulatory frameworks within which data is to be made secure, interoperable and usable across borders. The term ‘*digital environment*’ or ‘*digital ecosystem*’, in turn, refers to the broader framework of transformation in which infrastructure, competencies and *governance* interact.

In this context, digital sovereignty does not mean isolation, but rather the strategic capacity to act within interconnected systems. Vardanyan & Kocharyan (2023) interpret it as an expression of European autonomy within a fragmented global data order. Morse (2022) describes this fragmentation as competition between normative orders of different regulatory models from various actors such as the US, China or Russia. Bradford analyses the global impact of European regulation and demonstrates that, under certain conditions, European regulation can exert global influence – the so-called ‘Brussels Effect’ (Bradford, 2020, pp. 54–56).

The ‘information space’ thus becomes a space of normative projection. It is not merely the subject of regulation, but a strategic instrument. Its architecture conveys not only democratic deliberation, but also European conceptions of order into a globally contested information landscape.

This brings together the previously developed levels – public sphere, infrastructure, architecture, security and sovereignty – into an integrated spatial structure.

2.4.7 Terminological consolidation: concepts of space in the European context

The European Union does not use uniform terminology for the space analysed here. Rather, it differentiates according to the political context.

The term '*information space*' appears primarily in the context of security policy and democracy. It refers to the communication space in which strategic influence operations take place, public opinion is shaped and democratic processes can potentially be destabilised (EEAS, 2023; Council of the European Union, 2022). In this usage, the focus is on the space as a politically contested domain.

In contrast, the term '*information environment*' is used primarily in an analytical and operational context. The term is used to describe threat scenarios, hybrid patterns of influence or military contexts of action, and is therefore more process-oriented. This terminology is found particularly in security and defence policy contexts and is compatible with the NATO framework, in which the information environment is defined as a military-operational domain (Council of the European Union, 2022).

The European Data Strategy has also established the term '*data space*'. This refers to clearly defined, regulatory-structured data infrastructures designed to ensure economic value creation, interoperability and European autonomy (European Commission, 2020a; Regulation (EU) 2022/868). Here, the focus is not on discourse but on infrastructure.

Finally, transformation programmes such as the 'Digital Decade 2030' use the terminology '*digital environment*' or '*digital ecosystem*' (Decision (EU) 2022/2481). These terms mark the broader technological and societal context of digital development.

The term '*information sphere*', on the other hand, is not institutionally enshrined in official EU documents. It has no consistent political or legal function in the European context.

From an analytical perspective, therefore, the term ‘EU Information Space’ proves to be an integrative umbrella term. It combines dimensions of security policy, democracy and infrastructure without being reduced to a single functional logic. It describes not an isolated policy area, but a structured regulatory architecture.

2.4.8 EU information space

The development of the information space in the EU can thus be described as a transition: from countering individual cases of disinformation, through safeguarding democratic integrity, to the comprehensive governance architecture of a transnational information space. It constitutes the central sphere of action and interpretation for European democracy in the digital age. It is not a metaphorical description of digital communication, but politically shaped regulatory architecture.

If we imagine a room in a house in which we live, then the information space encompasses the democratic public sphere, which forms the legally and institutionally framed structure of the space through the presence and visibility of political actors (Arendt, 1958, pp. 198–199) and reason-guided communication (Habermas, 1992, pp. 385–386), in which democratic decision-making takes place. This gives the space the purpose for which it is designed, such as the living room, where family life takes place. The technical equipment of the space, such as an internet connection, and the organised communication structures of the interior design, such as televisions or computers, form part of the existing democratic infrastructure of the space, which is shaped by a pluralistic democratic media order (Schrijvers et al., 2026) and information and data networks.

Both are currently being redesigned, particularly in terms of interior architecture, by external digital platforms; today, they partly unintentionally form the load-bearing walls of this space and reshape the infrastructure. The smartphone can be cited as an example here, as it has now partly replaced the television and computer (Lessig, 1999, pp. 6–8; Zuboff, 2019; Deibert, 2020, pp. 18–22). As these load-bearing walls were not constructed by experts but designed according to instructions from a leading social media platform, they exhibit construction defects and must be repaired through platform regulations, which could once again alter the architecture of the information space.

The open architectural design of the space (with open windows and doors), which promotes public democracy within the space through its permeability, simultaneously creates a structural vulnerability to hybrid strategic actors (Whyte & Etudo, 2025; EEAS 2023; 2024; 2025; Dunn Cavelty, 2013, pp. 110–112; Hanley, 2024; Berk, 2025, pp. 2–3). These vulnerabilities should be countered by construction methods that promote resilience, such as the installation of alarm systems and security doors (Beck, 1986, pp. 25–30). In order to achieve the necessary degree of independence and self-determination – that is, sovereignty – through resilience, care must be taken to ensure that the infrastructure of the space is constructed as autonomously as possible, i.e. as burglarproof as possible. Furthermore, only those who actually live in the house should be given the key to the door and the password for the alarm system (Bratton, 2016, pp. 5–11; Bradford, 2020, pp. 29–33; Morse, 2022; Vardanyan & Kocharyan, 2023). Should the EU model prove functional, it could exert a normative influence and become an example of successful architecture for other ‘houses’ of lived democracy (Bradford, 2020, pp. 29–33).

2.5 The NATO Information Space – Operational Logic, Perception Management and Transatlantic Interdependence

2.5.1 From communication space to operational space

Whilst the information space in the European Union has developed as a regulated space of democratic public life, within NATO it has appeared from the outset as an integral part of military operations. This difference points to two distinct starting points: the EU shapes the architecture of the information space and anchors it normatively in law, infrastructure and public life, whilst NATO utilises it as a strategic environment of operation.

NATO doctrines, in particular AJP-01 (NATO, 2022), AJP-10 (NATO, 2023a), AJP-10.1 (NATO, 2023b), AJP-10.3 (NATO, 2024) and AJP-3.10.1 (NATO, 2015) make it clear that information is not understood as a passive element, but as a means that can be deployed in a targeted manner to influence perception, interpretation and behaviour.

In Clausewitz's sense, the information environment is thus part of the pursuit of political objectives in conflict (Clausewitz 1832/2008: Book I, Chapter 1). It is not merely a space accompanying military operations, but an independent sphere of strategic action in which impact is achieved through the targeted shaping of information.

2.5.2 The NATO 'Information Environment': Structure and Definition

Unlike the EU, NATO systematically uses the term '*information environment*'. This is described in its doctrines as the totality of actors, information and technical systems that influence perception and decision-making processes (NATO, 2022; NATO, 2023b).

This definition fundamentally shifts the focus: it is not the space itself that is central, but the processes that unfold within it. The information space thus appears less as a clearly defined structure and more as a dynamic web of influences in which information circulates, is interpreted and takes effect.

NATO (2025) elaborates on this understanding across three closely interlinked levels:

- **Physical dimension** → the material basis of the space, i.e. infrastructure, networks and technical systems
- **Informational dimension** → the content, data and narratives that move within this space
- **Cognitive dimension** → the perception, interpretation and decision-making of individuals and groups

This three-part division makes it clear that the information space should not be understood as a static space, but as a multi-layered structure in which physical structures, communicative content and cognitive processes are inextricably linked.

The analytical focus here is on the cognitive level. Information only unfolds its strategic significance where it alters perceptions, shifts interpretations and influences decisions.

The information space thus appears not primarily as a normatively ordered space, but as a sphere of influence in which strategic effects unfold. Whilst

in the European context communicative processes constitute the space itself, here they are understood as operational elements that are deployed in a targeted manner within this framework.

From a NATO perspective, the information environment does not therefore describe a clearly defined, structured space, but rather a dynamic network of interactions in which information only develops its strategic relevance through ongoing interaction and influence. The term *'environment'* should not be understood as part of a higher-level information space, but rather as an independent, process-oriented perspective on the same reality.

2.5.3 The information space as a space of power and discourse

The processes described above can be understood theoretically as discursive dynamics. Information does not exert its effect in isolation, but always within contexts of meaning that determine how events are perceived and categorised.

This perspective can be captured using Foucault. Discourses determine what is considered relevant, credible and true in a given context (Foucault, 1971, pp. 11–17; 1972, pp. 49–55). They thus structure not only communication, but the conditions under which reality is interpreted.

In the context of NATO, this logic is put into practice. Strategic communication is understood as the coordinated use of all means of communication to specifically influence perceptions, attitudes and behaviour (NATO, 2023a). Discourses thus appear not as an open process of opinion forming, but as a malleable factor in strategic action.

The information space is consequently not a neutral communication space, but an environment in which meaning arises through process, is shaped discursively and can be specifically influenced. Whilst in the European context discourse is understood as the foundation of the political public sphere, here it becomes an instrument of strategic influence.

This dual perspective also points to a further distinction that becomes particularly clear when compared with the US. Whilst meaning is understood in the European context primarily as the result of deliberative processes and appears in the NATO context as a factor that can be strategically influenced,

in US approaches it is frequently conceived explicitly as an operationalisable instrument of control. There, meaning is not only malleable, but can be specifically produced and scaled.

The information space is therefore seen as a process-driven environment in which meaning is not only interpreted but actively produced and deployed, albeit within different political and strategic logics.

2.5.4 Perception management: framing and agenda setting as operational mechanisms

These discursive processes take concrete form in the information environment through specific mechanisms of perception management. Information exerts its influence not solely through its content, but through the way in which it is selected, weighted and interpreted.

These processes can be clarified using the concepts of agenda setting and framing. Agenda setting describes the ability to bring issues to the fore and determine their priority within public communication (McCombs & Shaw, 1972, pp. 176–177). Framing, in turn, structures how these issues are interpreted by highlighting certain interpretative frameworks and obscuring others (Goffman, 1974, pp. 21–25; Entman, 1993, pp. 52–53).

In NATO doctrine, these mechanisms are not merely described but are utilised operationally. Information operations aim to shape information environments in such a way that certain topics come to the fore and certain interpretations become more likely (NATO, 2023b). Military Public Affairs complement this logic by focusing attention and stabilising trust through credible communication (NATO, 2024).

However, these mechanisms do not remain at the level of abstract models, but can be understood as concrete interventions in the information environment. By deliberately highlighting certain topics and obscuring others, attention is structured and the framework is set within which events can be perceived at all. Agenda setting thus determines what is discussed, whilst framing influences how it is discussed.

Empirical analyses show that these processes are systematically employed in real-world conflict contexts. Wilbur (2023) demonstrates that topics are delib-

erately prioritised to direct attention towards strategically relevant content. Berk further demonstrates that interpretative frames do not operate uniformly in complex information environments, but are received and processed differently depending on the context, audience and platform (Berk, 2025, pp. 2–3).

The information environment thus becomes a process in which meaning does not arise by chance, but is steered along specific interpretative paths. Agenda setting and framing work together in this process: they structure perception in such a way that certain interpretations become compatible and can coalesce into overarching narratives.

2.5.5 Narratives as a strategic structuring of meaning

The mechanisms of perception management described above coalesce within the information environment to form overarching interpretative frameworks. When topics are deliberately brought to the fore and interpretations are steered, stable contexts of meaning emerge that link individual pieces of information and give them direction.

In NATO doctrine, these structures are referred to as narratives. Narratives describe coherent interpretative patterns that place events, actors and objectives within a comprehensible context, thereby enabling orientation within complex information environments (NATO, 2023a).

Crucially, narratives do not simply emerge retrospectively, but are actively shaped. They bundle individual themes and frames into overarching structures of meaning that determine how events are interpreted and what conclusions are drawn from them.

Empirical analyses show that these processes are significantly amplified by digital technologies. Howard demonstrates that narratives in digital communication environments can not only be disseminated, but also systematically scaled up and automatically amplified. Platforms, networks and algorithmic systems act as amplifiers, making certain interpretations visible whilst pushing others into the background (Howard, 2018, pp. 2–4).

The information environment thus appears as a setting in which meaning is not only created and influenced, but also stabilises along strategically established patterns of interpretation. Narratives function as a connecting structure: they

place individual pieces of information within a broader context and make them relevant to perception and action (Ertl, 2022, pp- 179–183).

The information environment thus takes the form of a strategic competition of interpretations, in which different narratives collide, overlap and vie for dominance.

2.5.6 Operationalisation: The information environment as a strategic arena

In the NATO context, the processes of controlling meaning and perception described above do not remain at the analytical level, but are deliberately translated into operational practice. The information environment thus becomes a concrete arena for military and political influence.

This influence is not exerted in isolated instances, but through the coordinated interplay of various instruments. Strategic communication serves as an overarching framework within which different forms of communication are coordinated to achieve coherent effects (NATO, 2023a).

Within this framework, information operations take on the task of purposefully shaping the information environment. They intervene in the processes described above by prioritising topics, influencing interpretative frameworks and reinforcing narratives (NATO, 2023b). Military Public Affairs complement this approach by building trust through credible and transparent communication, thereby increasing the receptiveness to their own messages (NATO, 2024).

Finally, Psychological Operations operate at the cognitive level. They aim to directly influence attitudes and behaviour by addressing perceptions, emotions and decision-making processes (NATO, 2015).

Crucially, these instruments do not operate in isolation but function within the same framework of influence. They intervene at different levels of the information environment, ranging from the visibility of issues through their interpretation to the anchoring of meaning in the form of narratives.

The information environment thus appears not merely as an object of analysis, but as a strategic arena for action in which processes of perception, interpretation and decision-making can be specifically shaped.

2.5.7 Comparison with the EU: information space and information environment

A comparison of the European Union and NATO shows that both actors address the same subject area, but conceptualise it in fundamentally different ways.

The European Union describes the information space as a normatively structured space shaped by legal, institutional and infrastructural measures. At its core lies the stabilisation of democratic public life. Communication and discourse are not means to an end here, but constitutive elements of the space itself: public life arises through visibility, deliberation and institutionalised criticism.

In contrast, NATO understands the information environment not as a space to be shaped, but as a network of effects in which information, actors and processes of perception interact with one another. Discursive processes do not form the basis of the space here, but rather its operational level. They are deployed specifically to influence perception, meaning and behaviour.

The difference can thus be summarised precisely: whilst in the European context discourse constitutes the space, in the NATO context it operates within an environment understood in terms of processes. The EU conceives of the information space as an architecture, NATO as an ongoing process.

However, this difference does not imply a contradiction, but rather points to complementary perspectives. The EU creates the structural and normative conditions within which information circulates, whilst NATO analyses and specifically influences the dynamics of this framework.

The information space and the information environment thus appear as a shared sphere of action that is addressed from different perspectives: as a regulatory structure, as an environment of impact, and, in a broader view, as an operational control field.

2.5.8 Reference definition: The NATO information environment

NATO's understanding of the 'information space', which is viewed as an information environment, thus differs fundamentally from a space-centred perspective. The focus is not on the order of a structured communication space, but on the analysis and shaping of a dynamic framework of effects.

The information environment refers to the totality of actors, information, technical structures and cognitive processes within which perception arises, meaning is formed and decisions are influenced (NATO, 2022; NATO, 2023b). It is not a demarcated space, but a process-oriented environment in which information only unfolds its strategic relevance through interaction.

Within this framework, various mechanisms interplay: issues are brought to the fore or sidelined (agenda setting), interpretative frameworks shape how they are understood (framing), and overarching narratives bring these processes together to form stable structures of meaning. These dynamics are amplified by digital technologies and can be specifically scaled and controlled (Howard, 2018, pp. 2–4).

The information environment is thus not a neutral communication space, but a strategic context for action in which perception, meaning and behaviour can be actively influenced. Strategic communication, information operations, public affairs and PSYOPS intervene in a coordinated manner at different levels of this framework to achieve coherent effects (NATO, 2023a; 2023b; 2024; 2015).

The NATO approach therefore conceives of the ‘information space’ not as a stable order, but as an ongoing process. Meaning here does not arise primarily through deliberative negotiation, but is specifically shaped, reinforced and deployed within the framework of strategic communication.

In interaction with the European Union, it becomes clear that these are not competing but complementary perspectives. Whilst the EU shapes the information space as a normatively structured architecture of democratic public sphere, NATO addresses its dynamic processes as a strategic framework of effects.

As the reference concept for this study, the NATO information environment can thus be understood as that process-oriented perspective on the information space in which the interactions between information, perception and behaviour are analysed and deliberately shaped in order to achieve political and military objectives.

2.6 The US information environment: From information control to structured meaning-making

The development of the US concept of the *information environment* (IE) since the mid-1990s reflects a fundamental shift in the understanding of political power in the information age. Unlike the EU information space, the information environment is viewed as a cross-domain military operational space. Whilst early doctrines treated information primarily as a technical resource, it is now understood as a socio-cognitive framework of effects in which the production of meaning, perception and behaviour are systematically interlinked. This shift is not merely terminological in nature, but rather an expression of a conceptual shift: from the control of information systems towards the shaping of structured processes of meaning.

The development of the US information environment can be explained as a transition from an infrastructural control model to a model of structured meaning production. The doctrine implicitly follows social constructivist and cognitive psychological insights, which place the strategic relevance of 'meaning' at the centre.

2.6.1 Information as strategic infrastructure

The *Joint Doctrine for Command and Control Warfare* (1996) describes the information environment for the first time primarily as a network of military and civilian information systems, the protection or disruption of which generates operational effects (Joint Chiefs of Staff, 1996). 'Information' is understood here as a strategic resource: control over communication channels therefore implies control over decision-making and command processes. *JP 3-13* (1998) provides the first formal definition of the '*information environment*' as an aggregate of individuals, organisations and systems that collect, process or disseminate information (Joint Chiefs of Staff, 1998). This definition is standardised in the Department of Defence Dictionary of Military and Associated Terms (2001) (Joint Chiefs of Staff, 2001). In this phase, an infrastructural paradigm dominates. Information appears as an object circulating between actors. Strategic impact is generated primarily through the control, disruption or manipulation of information flows. Meaning, in the sense of socially constructed interpreta-

tion, plays a subordinate role. Power is understood in functional-instrumental terms: whoever controls information systems gains operational advantages.

2.6.2 The cognitive opening

A conceptual shift took place with the 2006 revision. Since then, the doctrine has incorporated the three-dimensional model: physical, informal and cognitive (Joint Chiefs of Staff, 2006). At this point, it is clear if not before, that information does not automatically have an effect, but is interpreted by people. The 2012 revision expands on this cognitive dimension (Joint Chiefs of Staff, 2012). This shift can be classified as constructivist. Berger and Luckmann argue that social reality is not objectively given, but is constructed and stabilised through social processes (Berger & Luckmann, 1966, pp. 33–36). Information acquires political relevance not based on its objective content, but based on the meanings attributed to it. With this insight, the doctrine begins to recognise implicitly that strategic impact does not depend solely on technical control, but on the ability to influence processes of interpretation. Information is increasingly understood as a medium for the structured production of meaning.

2.6.3 The Information Environment as a Structured Framework of Meaning

The decisive paradigm shift occurs with *Joint Publication 3-04: Information in Joint Operations* (2022). The information environment is now defined as an aggregate of social, cultural, linguistic, psychological, technical and physical factors that influence how individuals and automated systems construct ‘meaning’ and act on that basis (Joint Chiefs of Staff, 2022).

This definition fundamentally shifts the focus: the emphasis is no longer on the transmission of information, but on the conditions under which meaning is ascribed to it. ‘Meaning’ here does not refer to the objective content of information, but to the socially constructed interpretation of information. ‘Meaning’ arises from the interplay of narrative orders, cultural contexts, individual experiences and social dispositions. It is relational, context-dependent and action-guiding. Strategic narratives structure political reality by embedding events within coherent interpretative frameworks (Miskimmon, O’Loughlin & Roselle, 2013, pp. 2–5). Wilbur demonstrates that such narratives gain salience through media diffusion processes and shape political agenda structures (Wilbur, 2021, pp. 147–148). At the individual level, narratives are mediated

through frames that structure perception (Gamson, 1975, p. 604). Cognitive processes often proceed heuristically and intuitively, as the dual-process model suggests (Kahneman, 2011, pp. 20–21). Information is processed according to learned evaluation patterns. These evaluation patterns are socially shaped.

Bourdieu's 'concept of habitus' describes incorporated dispositions that structure perception and action (Bourdieu, 1977, pp. 72–95). Generative schemata organise practical logic and limit the compatibility of certain interpretations. Symbolic power consists in making interpretations appear self-evident (Bourdieu, 1977, pp. 164–171). 'Meaning' is therefore neither freely available nor arbitrarily manipulable, but embedded in social structures. The strategic relevance of 'meaning' is ultimately demonstrated in its influence on the actions of the target audience. JP 3-04 emphasises this explicitly (Joint Chiefs of Staff, 2022).

Meaning becomes politically effective when it influences behaviour. Reinforcement mechanisms such as strategic communication serve to stabilise such orientations (Berk, 2025, pp. 2–4). Howard therefore emphasises that strategic communication must be considered in an integrated manner, as perception, reputation and action are closely interlinked (Howard, 2018, pp. 2–4). At the same time, the impact of information operations is structurally limited. Whilst international anarchy, institutional checks and pre-existing dispositions reduce the strategic impact of disinformation, Lanoszka also emphasises the security policy relevance of cognitive vulnerability and refers to structural conditions (Lanoszka, 2019, pp. 24–28). The information environment thus appears as a setting in which power is structured, but not as an unlimited space for manipulation.

2.6.4 The current definition of the 'information environment' in 2022

The consistent use of the term '*information environment*' since 1998 (Joint Chiefs of Staff, 1998; 2022; Office of the Chairman of the Joint Chiefs of Staff, 2025) makes it clear that this is not a territorially defined 'space', but rather a relational framework of conditions.

The US information environment refers to the overall framework of individuals, organisations and technical systems that generate, collect, process and disseminate information or act on the basis of it, as well as the social, cultural,

psychological and technological conditions under which these information processes take place and meaning is generated (Office of the Chairman of the Joint Chiefs of Staff, 2025; Joint Chiefs of Staff, 2022).

The historical development of the US information environment shows a clear trajectory: from the control of technical information systems, through the recognition of cognitive processes, to a theoretically implicit conception of structured meaning production.

The 2022 definition sums up this development. The information environment is neither a territory nor a normative public sphere, but a socio-cognitive framework of influence. Within it, political influence is mediated through the structuring of perception and interpretation.

The US doctrine thus reflects a central insight: in the information age, power operates less through physical coercive means than through the organisation and determination of orders of meaning.

2.7 The Chinese information space as a cognitive space of order

2.7.1 From sovereignty to the question of order

The Chinese conception of the information space differs fundamentally from Western models. Whilst the United States defines the *information environment* primarily as a military operational space and the European Union regards the information space chiefly as a democratic sphere of influence that must be protected, a more comprehensive, regulatory perspective is developing in China. Here, the information space is understood not merely as a technological infrastructure or sphere of communication, but as an integral part of national sovereignty and social stability.

The underlying ideological structure of this concept lies in the state-institutionalised narrative of the so-called ‘Century of National Humiliation’. The period from the Opium Wars of 1839 to the founding of the People’s Republic in 1949 is interpreted in official white papers as a phase of external subjugation, territorial fragmentation and political powerlessness (State Council Information

Office, 2011; State Council Information Office, 2019). This historical experience is not treated as a closed chapter of the past, but rather as a normative frame of reference for the current security architecture.

Security thus appears as a preventive category: it is intended to prevent structural inferiority or external dominance from becoming possible again. In this context, the guiding principle of 'national rejuvenation' takes on security-related significance for the enduring sovereignty of China vis-à-vis external actors. Political stability, technological autonomy and ideological coherence are understood as prerequisites for historical rehabilitation. Unlike in the US, where global freedom of action is paramount, the Chinese discourse emphasises the avoidance of renewed vulnerability (State Council Information Office, 2019).

2.7.2 Institutional integration and conceptual transformation

The concept of the "holistic national security approach" institutionalises this historical semantics. The policy document published in 2022 defines national security as a comprehensive organising principle that systematically links political stability, ideological coherence, cultural identity, digital infrastructure and technological self-reliance (Central Propaganda Department of the Communist Party of China, 2022). Here, security is not conceived in isolated sub-areas, but as an integrative structure, that encompasses state leadership, social stability and technological sovereignty in equal measure.

This broadening of the concept of security paves the way for the conceptual transformation from a technical information space to a cognitive space. During the phase of 'informationised warfare' in the 1990s and 2000s, the focus was on control over information flows, networks and C4ISR (Command, Control, Communications, Computers, Intelligence, Surveillance, and Reconnaissance) systems. The information space was primarily understood as a technological operational space. This phase was strongly influenced by Western military doctrines and aimed at efficiency and decision-making superiority (State Council Information Office, 2015).

With the 2015 military strategy, the focus shifted to intelligent or AI-supported warfare. Data is not only collected but also processed algorithmically. It is no longer the volume of information alone that is decisive, but algorithmic power of selection. Whoever determines relevance structures decision-making processes (State Council Information Office, 2015).

From around 2018, the concept of cognitive space finally comes to the fore. This marks a qualitative shift: from control over information channels to the management of frames of reference. Information control concerns content and distribution channels; the management of meaning concerns patterns of interpretation, emotional dispositions and constructions of identity. The scope of state power expands from the technical to the symbolic realm (Wu, 2022, p. 5).

2.7.3 Cognitive security and warfare

With the establishment of the term ‘cognitive security’, perception itself becomes an object of protection in security policy. In Chinese literature, ‘cognitive security’ refers to the stabilisation of social interpretative frameworks, ideological coherence and political loyalty (Liang, 2023, pp. 2–3). It forms the interface between political security, ideological security and social stability within the “overall state security concept” (Central Propaganda Department of the Communist Party of China, 2022).

The operational dimension is referred to as cognitive warfare. It is not primarily aimed at short-term deception, but at the long-term structuring of patterns of perception, narratives and attributions of identity (Yan, 2023, pp. 6–7). In this context, Lin (2023, pp. 7–8) speaks of a national system of cognitive security, which encompasses institutional mechanisms for stabilising social orders of meaning.

Historically, three phases of development can be distinguished. Firstly, the phase of technical information dominance (approx. 1990s–2014), in which information infrastructure and networking were central (State Council Information Office, 2015). Secondly, the phase of algorithmic control dominance (2015–2018), in which artificial intelligence and data-driven decision-making processes became strategically central (Zeng, 2023, pp. 9–10). Thirdly, the phase of cognitive order dominance (from around 2018), in which structures of perception and meaning themselves are defined as security-relevant (Yu, 2023, pp. 4–5).

A structural difference emerges when compared to the US, Russia and the EU: whilst the US treats the information environment primarily as a domain for military operations, the EU regards it as a space for democratic resilience, and Russia emphasises informational influence operations, China systematically integrates the cognitive space into its state security architecture (Yan, 2023, pp. 10–11).

The Chinese information space is therefore to be understood as a historically standardised, institutionally integrated and technologically structured cognitive space of order, in which perception, meaning and collective disposition are treated as strategic resources. Its logic of protection and shaping serves to stabilise political legitimacy, social cohesion and national sovereignty (Lin, 2023, pp. 8–9).

The decisive shift is thus not merely technological, but one of power politics: from control over information to the structuring of meaning production. The Chinese information space is less a battlefield than a field of order, a space in which historical experience, ideological stability and technological control are brought together (Wu, 2022, p. 5; Liang, 2023, pp. 3–4).

2.8 The information space in the strategic thinking of the Russian Federation

The Russian concept of the information space is deeply rooted in security and geopolitical traditions of thought. In the strategic discourse of the Russian Federation, information is not primarily understood as a means of communication, but as a resource of power. The information space refers to the security policy perspective on the entirety of information-related structures and processes. It constitutes an independent dimension of state capacity for action, which possesses strategic relevance alongside military, economic and territorial factors (Panarin, 2006, pp. 34–38; Manoilo, 2015, pp. 52–58).

Terminologically, a distinction must be made between the ‘information sphere’ and the ‘information space’. Whilst the information sphere, in academic works – particularly in Manoilo’s – describes the functional totality of infrastructure, actors, communication processes and regulatory mechanisms (Manoilo, 2015, pp. 48–55), the concept of information space predominates in state documents and strategic analyses. The concept of space does not denote a different ontological category, but rather a security-policy interpretation of this structure. It implies the ability to be demarcated, strategic relevance and a need for protection. The information space is thus the strategic framing of the information sphere from the perspective of state sovereignty (Doctrine of Information Security, 2016).

Panarin, a long-standing professor at the Diplomatic Academy of the Russian Ministry of Foreign Affairs and a leading figure in the Russian geopolitical school, explicitly regards information as a geopolitical resource. In his conception, information power is an independent category of power capable of stabilising or destabilising political systems (Panarin, 2006, pp. 27–33). Modern conflicts are therefore not decided solely by physical force, but by the ability to shape patterns of interpretation and influence structures of perception. The information space appears as an arena of international competition for interpretative sovereignty.

Manoilo expands this perspective by describing the information sphere as an integrated system that links technical infrastructure, media institutions and normative paradigms (Manoilo, 2015, pp. 52–58). When viewed from a security policy perspective, this structure appears as an information space, the stability of which requires state control and protective measures. Information thus functions not only as a technical flow of data, but also as a vehicle for social and political order.

This theoretical foundation gives rise to an expanded understanding of security. Security is not defined exclusively in territorial terms, but cognitively and normatively. The protection of the information space consequently also encompasses the protection of societal structures of interpretation and perception (Panarin, 2006, pp. 39–45).

This logic is reflected in the Russian understanding of information warfare. Komov defines information warfare as a coordinated set of measures aimed at achieving information superiority (Komov, 1996, pp. 76–78). In the further development of this approach, a distinction is made between information warfare as a form of escalation and information confrontation as a permanent structural state (Manoilo, 2015, pp. 67–72). Information confrontation describes a continuous competition between states within the information space for narrative dominance and opportunities for cognitive influence.

The traditional distinction between war and peace is thereby relativised. Information measures are regarded as a standard instrument of state policy. The information space appears as a permanent arena of conflict in which political, media and psychological means are employed. Military escalation is interpreted as an intensification of an already existing confrontation.

The cognitive dimension plays a central role. Klintsova (2023) describes information warfare as the targeted manipulation of perceptions conveyed through the mass media. Narratives, enemy stereotypes and symbolic attributions structure collective perception and can generate social polarisation. Security is thus understood as the protection of mental structures of order within the information space.

These considerations tie in with Messner's concept of modern conflicts, in which societies themselves become the theatre of operations (Messner, 1960, pp. 112–118). The influencing of collective structures of consciousness is defined as a security policy objective.

The 2016 Information Security Doctrine codifies these theoretical assumptions normatively. Information security is defined as the protection of individuals, society and the state from internal and external information threats (Information Security Doctrine, 2016). In this context, the information space is treated as a security-relevant area of state responsibility. The protection of intellectual and moral values is explicitly named as a security objective. In conjunction with the National Security Strategy 2021, external information intervention is interpreted as a potential threat to social stability (National Security Strategy, 2021).

This security-theoretical logic gives rise to the concept of information sovereignty. Information sovereignty refers to the ability to regulate digital infrastructure autonomously and to maintain the national information space independently in the event of a crisis. The concept of the 'sovereign internet' is an expression of this state-centred vision of order.

Overall, a coherent theoretical framework emerges. The information sphere describes the functional structure of information-related processes; the information space denotes their security policy framework. From this perspective, the information space is understood as a strategic arena of power and conflict in which geopolitical competition, normative identity and cognitive control are intertwined. Security encompasses not only territorial integrity but also the protection of societal structures of perception and interpretation. The information space thus appears as a distinct dimension of modern security strategy.

Against this background, the further question arises as to how far the Russian information space actually extends. The arguments presented so far suggest that the information space cannot be conceived exclusively in territorial terms, but within Russian strategic logic, extends beyond the borders of the Russian Federation.

The Russian Federation's 2023 Foreign Policy Concept provides a key framework for this. It explicitly describes Russia as a 'civilisational community' that extends beyond its national territory and encompasses the 'Russian World' ("russkij mir"). This encompasses not only citizens, but also Russian-speaking communities, cultural ties and historical identity spaces abroad. At the same time, the promotion of the Russian language, support for compatriots abroad and the strengthening of Russia's position in the 'global information space' are defined as explicit foreign policy objectives.

This results in an analytically relevant distinction between two levels of the information space. Firstly, there is a territorially limited information space that refers to the Russian Federation in the narrower sense. This is subject to legal regulation, technical control and security-related shielding, as expressed in particular in the doctrine of information security and the concept of the 'sovereign internet'. In this 'narrow' information space, a defensive logic prevails, aimed at stability, control and protection against external influence.

Secondly, an expanded information space can be identified that extends beyond the national territory. This encompasses, in particular, the states of the so-called 'near abroad', Russian-speaking publics, and transnational communication spaces in which Russian media, narratives and cultural reference systems operate. The foreign policy concept explicitly emphasises the need to strengthen Russia's position in the global information space and to promote an "objective perception of Russia abroad". Furthermore, supporting compatriots abroad and maintaining a shared cultural identity are defined as strategic tasks.

This distinction points to a dual logic in Russian information policy. Whilst the internal information space is primarily conceived as a protective space, the external information space appears as a sphere of influence. In this expanded space, information is actively employed as a foreign policy instrument to shape narratives, strengthen cultural ties and expand opportunities for polit-

ical influence. The information space is thus not strictly bound to territorial sovereignty, but follows a civilisational and linguistic-cultural logic.

This simultaneously corrects the assumption of a purely defensive orientation of Russian information policy. Rather, it becomes apparent that Russia pursues a significantly more active and, in some cases, offensive strategy within the expanded information space. The foreign policy framework makes it clear that exerting influence in the information space is regarded as a legitimate instrument of international politics, particularly in the context of geopolitical competition and the confrontation with the West.

Overall, the Russian information space can thus be understood as a multi-layered concept: as a territorially limited, regulatable space on the one hand, and as a transnational sphere of influence on the other. This dual structure is central to understanding Russian information policy and at the same time constitutes a significant difference from Western, particularly European, conceptions of the information space.

2.9 Classification and comparison of concepts in the information space

2.9.1 Psychological Operations (PSYOPS)

In a military context, psychological operations (PSYOPS) refer to planned and targeted measures that make deliberate use of communication to influence the perceptions, attitudes and behaviour of selected target groups, thereby contributing to the achievement of political and military objectives (NATO, 2015, p. 1).

In the current NATO understanding, PSYOPS form part of a behaviour-centric approach aimed at bringing about targeted changes in the cognitive dimension of target groups, in particular changes in perceptions, attitudes and decision-making behaviour (NATO, 2024, p. 3). They thus operate primarily within the cognitive dimension of the information environment, which encompasses beliefs, emotions, values and decision-making processes (U.S. Department of the Air Force, 2023, p. 2).

The information environment represents a complex fabric of social, cultural, psychological, technical and physical factors that shape how information is generated, disseminated and processed (U.S. Department of the Air Force, 2023, p. 6). PSYOPS intervene specifically in these structures by influencing patterns of interpretation and steering perception processes.

Within NATO, PSYOPS are embedded within a broader approach to strategic communications (StratCom). StratCom serves the coordinated use of communication and action to specifically influence perceptions and behaviour in the information domain (NATO, 2017, p. 1). The operational implementation of these objectives takes place within the framework of information operations, which combine various instruments to achieve coherent effects on target groups (NATO, 2019a, p. 2).

Within this system, PSYOPS represent a specific instrument based on the systematic analysis and targeted engagement of defined target groups (NATO, 2015, pp. 2–3). The aim is to achieve effects such as demoralisation, behavioural adjustment or support for one's own actions through coordinated communication measures (NATO, 2024, p. 5). Their effectiveness depends largely on the credibility of the messages and the perception of the senders.

The impact of PSYOPS unfolds primarily through strategic narratives. Narratives structure perception by placing events, actors and actions within meaningful contexts, thereby prescribing patterns of interpretation (Roselle et al., 2014, p. 71). They define how the international order is interpreted, which actors appear legitimate and which actions are deemed necessary. PSYOPS make targeted use of these mechanisms to influence perception and behaviour not just in isolated instances, but also in the long term.

Furthermore, the mechanism of action can be explained by the framing approach. Communication highlights certain aspects of perceived reality and obscures others, thereby creating interpretative frames that systematically structure perception and evaluation (Jecker, 2017, p. 356). PSYOPS specifically target such frames in order to steer the classification of information in line with strategic objectives.

With digitalisation, the possibilities for psychological influence have expanded considerably. Modern influence operations are characterised by high volume,

the use of multiple communication channels, and the rapid, continuous dissemination of content (Paul & Matthews, 2016, pp. 2–3). Repetition and the use of various communication channels increase the visibility and perceived credibility of messages. Even contradictory or factually incorrect information can be effective, provided it appeals to existing beliefs or triggers emotional reactions (Paul & Matthews, 2016, pp. 5–6).

This mechanism can be explained by findings from the fields of communication and social sciences. Strategic communication is fundamentally based on persuasion, that is, the deliberate influencing of attitudes and behaviour through communication. Its impact depends not only on the content of the information, but also on context, repetition, the credibility of the source, and the cognitive dispositions of the recipients.

However, the conceptual classification of psychological influence operations varies considerably among international actors. Whilst NATO continues to use the term PSYOPS and defines it clearly in its doctrine, other actors draw on alternative concepts that describe partly similar mechanisms of action but are framed within different normative frameworks.

In the United States, psychological operations are increasingly conducted under the term ‘Military Information Support Operations’ (MISO). These aim to influence specifically the emotions, motives and behaviour of selected target groups in order to support military objectives (U.S. Department of Defence, 2010, pp. 1–3). MISO are an integral part of information operations, in which various instruments of influence are deployed in a coordinated manner.

Russia, by contrast, rarely uses the term PSYOPS and locates comparable measures within the concept of information warfare (“informatsionnaya voyna”), which is characterised by high-volume, multi-channel and continuous influence (Paul & Matthews, 2016, pp. 2–3). This form of influence is also characterised by repetition and a loose connection to objective reality, thereby shaping perception in a targeted manner (Paul & Matthews, 2016, pp. 5–6).

With the concept of the “Three Warfares”, China pursues a holistic approach that systematically combines psychological, media and legal mechanisms of influence. Psychological warfare aims to influence perception, behaviour and social sentiment and is employed not only in times of conflict but continuously

in peacetime (Lee, 2014, pp. 203–204). The aim is to influence perceptions and decision-making processes in the long term and to gain strategic advantages even before an open military conflict arises.

The European Union takes a markedly different approach to psychological influence. The term ‘PSYOPS’ is largely avoided and replaced by concepts such as strategic communication, countering disinformation and societal resilience. The focus here is not on directly influencing behaviour, but on protecting democratic processes and ensuring an informed public (European Commission, 2018, p. 1). Strategic communication serves as a tool for identifying and countering disinformation, as well as for strengthening societal resilience (European Commission, 2018, pp. 1–2, 10–11).

Overall, it is evident that psychological influence is a central element of modern conflict management in the information domain, though its specific form varies greatly. Whilst military actors such as NATO and the US focus more on directly influencing behaviour, Russia and China pursue more comprehensive, long-term strategies to control perception and information flows. The European Union, by contrast, relies primarily on defensive measures and normative frameworks. Psychological influence operations are thus not only operational instruments, but also an expression of differing strategic cultures, security policy objectives and normative conceptions of order.

2.9.2 Propaganda and disinformation

Propaganda and disinformation refer to forms of strategic communication aimed at deliberately influencing the perceptions, attitudes and actions of political publics, and differ in terms of their reach, operational logic and relationship to truth. Whilst propaganda can be understood as an overarching form of strategic influence communication, disinformation refers to a specific technique within this spectrum that is based on the deliberate dissemination of false or misleading information. For a theoretically precise classification, the post-truth politics approach is also relevant, as it focuses less on the existence of lies and more on a fundamental erosion of epistemic standards and a growing indifference towards truth.

A classic work of propaganda research is provided by Harold D. Lasswell, who defines propaganda as “the control of collective attitudes through the manipulation of meaningful symbols” (Lasswell, 1927, p. 627).

Propaganda thus appears as a targeted attempt to influence societal patterns of interpretation and perception and is closely linked to political power. This functional perspective is systematised by Garth S. Jowett and Victoria O'Donnell, who describe propaganda as a deliberate and systematic attempt to shape perceptions, manipulate cognitions and control behaviour (Jowett & O'Donnell, 2019, p. 6) The focus is thus on intentionality, systematisation and purposefulness as constitutive features of strategic communication.

For understanding the impact of such processes, Walter Lippmann's analysis of public opinion is of central importance. Lippmann argues that individuals do not grasp political reality directly, but act based on 'pictures in our minds'. This 'pseudo-environment' he describes arises through selective perception, stereotypical patterns of interpretation and media mediation, and significantly structures political action. Particularly relevant is his distinction between 'news' and 'truth': whilst news merely reports events, the function of truth is to reveal connections and provide guidance for political action. This makes it clear that public opinion is structurally susceptible to manipulation (Lippmann, 1922).

The propaganda model developed by Edward S. Herman and Noam Chomsky (2023) offers a structural extension of this understanding. Here, the media are not seen as neutral intermediaries, but as institutions that contribute to the reproduction of existing power relations through their choice of topics, weighting, framing and the limitation of the scope of debate. Propaganda thus arises not exclusively from deliberate communication strategies, but also from systemic filtering mechanisms within modern media systems (Herman & Chomsky, 2023).

In contrast, disinformation refers to a narrower form of strategic communication characterised by deliberate deception. In security policy research, it is understood as an integral part of modern information operations. Keir Giles describes disinformation as part of a form of conflict that operates deliberately below the threshold of open military confrontation ('grey zone'). The aim is to influence political systems, weaken social cohesion and steer decision-making processes in the interests of the actor involved. Disinformation is particularly effective in digital communication spaces, as it specifically targets broad sections of the population and can thus indirectly influence political processes. A key feature of modern disinformation is that it does not primarily aim to persuade, but rather to generate uncertainty, doubt and polarisation. Simply destabilising

trust in institutions can be enough to impair the capacity for political action. Furthermore, disinformation is often disseminated via seemingly independent actors who, consciously or unconsciously, propagate narratives and thereby enhance their credibility (Giles, 2022).

Sergei Prozorov's post-truth approach offers further theoretical insight. This demonstrates that disinformation becomes particularly effective when the distinction between truth and falsehood itself loses its significance. In such a context, statements are primarily perceived as opinions of equal value, causing objective criteria of evaluation to lose relevance. Argumentative debate is replaced by an expressive logic in which statements are no longer assessed according to their truth content, but according to their political function (Prozorov, 2021).

In summary, it can be stated that propaganda, disinformation and post-truth represent distinct yet closely interlinked analytical levels. Propaganda functions as an overarching concept of strategic influence communication, encompassing both deliberate manipulation and structural communication processes. Disinformation constitutes a specific operational technique geared towards deception. The post-truth approach complements this perspective with an epistemic dimension by describing the conditions under which disinformation can take effect. Taken together, these approaches enable a nuanced analysis of modern information spaces, in which political power is increasingly exercised through the control of perception and interpretation.

2.10 Comparative Matrix of the Information Environment

Comparative overview of the differing understandings of the information environment within the EU, NATO, the United States, Russia, and China.

Category	EU	NATO	USA	Russia	China
Fundamental Understanding of the Information Environment	Democratic, normatively regulated information, public communication and infrastructure environment	Strategic-operational information environment for perception and effects management	Socio-cognitive structure of meaning and perception	Sovereignty-centred security-political information environment	State-integrated cognitive security, order, governance and regulatory environment for societal stability
Type of Environment/ Terminology	Regulated information, public communication and data environment	Strategic information environment as an operational domain	Cognitive information environment as a sphere of influence	State-protected internal information environment and external sphere of influence	Cognitively integrated security and governance environment
Guiding Logic	Resilience and protection of democratic public discourse	Strategic effects and perception management	Production of meaning and behavioural influence	Information security and state-centred logic of stability and control	Societal cohesion and cognitive regulatory stability
Understanding of Security	Protection of democratic integrity and resilience	Protection of strategic capacity for action and operational effectiveness against hybrid, information-related and cognitive influence activities	Protection of strategic capacity for action and influence	Protection of state stability against external influence	Protection of cognitive security, ideological coherence and societal stability
Role of Information	Resource of democratic public discourse and societal resilience	Instrument of strategic effects and operational influence	Medium of cognitive meaning-making and behavioural steering	Instrument of security-political control and geopolitical influence	Instrument of state-directed perception management, societal cohesion and political stabilisation
Perceived Threats	Democratic destabilisation and FIMI	Hybrid influence operations, strategic destabilisation, and cognitive and information-related attacks	Strategic manipulation of perception and behaviour	External influence and information warfare	Ideological destabilisation and societal fragmentation
Strategic Use of Narratives	Protection of democratic public discourse and resilience	Narrative influence to generate strategic effects	Narrative production of meaning and behavioural influence	Geopolitical narratives and information control	Ideological coherence and societal stabilisation

Category	EU	NATO	USA	Russia	China
Cognitive Dimension	Protection of public deliberation and societal resilience	Operational perception and behavioural management	Cognitive attribution of meaning and behavioural influence	Psychological influence and societal destabilisation	Cognitive security and societal perception order
Political Regulatory Function	Stabilisation of pluralistic-democratic public discourse	Safeguarding strategic stability within the alliance framework	Establishment of strategic dominance in meaning and action	Safeguarding state sovereignty and geopolitical order	Stabilisation of party-state and societal order
Operationalisation of the Information Environment	Platform regulation, FIMI defence and resilience measures	Strategic communication, information operations, psychological operations and cognitive warfare	Strategic communication, influence operations and narrative meaning management	Information warfare, narrative control and influence operations	Cognitive security architecture, algorithmic societal governance and ideological stabilisation
Relationship between Technology and Control	Regulated technology governance for the protection of democratic public discourse	Technologically enabled effects and networking capabilities	Data- and platform-supported influence and analytical capabilities	Security-centred control and surveillance technologies	Integrated algorithmic societal and security governance
Dominant Logic of Conflict	Protection of democratic public discourse against disinformation, manipulation and information-related destabilisation	Coordinated hybrid, information-related and cognitive conflict activities for the generation of strategic effects	Competition over meaning, perception and influence	Permanent information and influence conflict	Long-term safeguarding of societal and ideological stability
Safeguarding and Expansion of Information-Related Sovereignty	Regulatory safeguarding of digital agency and data sovereignty	Securing strategic information and effects superiority through networking and operational integration	Expansion of global influence and information capabilities through platform, data and technological power	Protection of state information sovereignty through control, information security and defence against influence operations	Safeguarding and expansion of state information and control sovereignty through digital integration, platform governance and cognitive security architecture

Table 1: Understandings of the information environment

3

Alexandra Duca:

Cognitive Warfare

A Legal Assessment





Cognitive Warfare – A Legal Assessment

Alexandra Duca has worked in the International Law Department of the Austrian Federal Ministry of Defence since 2021. She holds a degree in Law as well as in Business Law and has specialised in Public International Law. Her work focuses on new technologies and domains as well as modern means and methods of warfare, including AI, autonomy, disinformation, cyber operations, and outer space. She regularly serves as part of the Austrian delegation at the United Nations in Geneva, New York and Vienna representing Austria's interests in different international fora. She also frequently gives guest lectures at the University of Vienna and Graz as well as at the Theresian Military Academy in Wiener Neustadt.

International law is by its nature technology agnostic, which means that its rules and principles apply regardless of the tools and means employed by states. Rather, the law aims to regulate behaviours and activities, thereby providing the necessary flexibility to govern also the use of new technologies, including non-kinetic ones. The core challenge therefore does not lie in determining whether the existing law applies at all, but in how it applies to these new technologies and conducts.¹⁴

The following chapter intends to unpack this question and address how activities that could be conducted as part of cognitive warfare need to be measured against the different legal standards in international law. In the course of this, it will become evident that a case-by-case assessment is often required in order to determine the lawfulness of a certain behaviour, since the specific circumstances are pivotal thereto. It is for this reason and to enable a more

¹⁴ For instance, states have repeatedly reaffirmed the applicability of international law to cyber-activities (Cyber Law Toolkit, 2025).

in-depth analysis of the legal implications that an example-based approach was adopted throughout this chapter.

The following remarks aim at underlining the importance and flexibility of existing international law. When applied correctly and in good faith, it can significantly contribute to countering and minimising threats related to the cognitive domain.

3.1 Relevant Obligations under International Law

In order to examine how international law applies to acts referred to under the concept of cognitive warfare, one first needs to identify which rules and principles are relevant. Naturally, the applicability of areas of international law will vary depending on the strategic situation prevailing. In principle, two situations can be distinguished: peacetimes and war times. While the primarily applicable body of law in times of war is international humanitarian law (IHL), in peacetimes states need to adhere to other legal frameworks, rules and principles, including international human rights law (IHRL), the principle of state sovereignty, the prohibition of intervention and the prohibition on the threat or use of force.

Nevertheless, even after the clarification of the applicable legal framework, the question remains, what precisely it is that one intends to assess and measure against these existing rules. Cognitive warfare is generally understood to refer to a variety of activities that can be undertaken to exploit vulnerabilities in human cognition. These may include disinformation campaigns, psychological operations, or cyber- and AI-enabled influence activities that target other states' populations, governmental institutions, or critical infrastructures.

For the purpose of the legal analysis, it is necessary to break down this bundle of activities. International law consists of general principles under which specific circumstances must be subsumed and assessed on a case-by-case basis. This includes examining a specific conduct, its specific background and intention, as well as the causally related consequences thereto. Yet, at the same time, it appears necessary to not always view situations purely

individually, but to take a holistic approach. Sometimes individual acts may not themselves reach a certain threshold under international law, but when viewed together they could.

The following chapter will further elaborate on fundamental rules and principles of international law, such as state sovereignty, the prohibition of intervention, the due diligence obligation and the prohibition on the threat or use of force as well as their application to acts of cognitive warfare.

3.1.1 State sovereignty

The principle of sovereignty is a cornerstone of the international legal order. It is enshrined in Article 2 (1) of the UN Charter, affirming the “sovereign equality” of all states. According to this principle, which is also a binding rule of customary international law¹⁵, a state has the exclusive authority to govern its territory and to exercise sovereignty over infrastructure thereon and its own nationals.

Generally, an activity violates the sovereignty of another state when it violates a state’s territorial integrity, as well as when it constitutes an interference with or usurpation of an inherently governmental function of a state. Cyber operations, disinformation campaigns and influence operations conducted remotely across borders challenge the traditional understanding of sovereignty. However, states have recognised that digital intrusions and cyber activities can constitute a violation of state sovereignty not only if they result in physical damage or injury, but may also do so if they lead to loss of functionality of a state’s critical infrastructure or if they prevent access to essential governmental services, be it only temporarily. The mere traversing of data packages through the ICT infrastructure of another state, without negative effects on the functionality thereof, does not itself violate the principle of sovereignty (Republic of Austria, 2024), whereas e.g. tampering with the electronic ballot system of another state would constitute a clear interference with inherently governmental functions and hence a violation of sovereignty (Cyber Law Toolkit, 2025).

15 Customary international law refers to international obligations arising from established international practices and results from a general and consistent practice of states that they follow from a sense of legal obligation (cf. Article 38 para. 1 lit. b Statute of the International Court of Justice).

3.1.2 Prohibition of Intervention

The prohibition of intervention is a rule of customary international law that is closely linked to the principle of sovereignty. It prohibits coercive interference by one state in the internal or external affairs (*domaine réservé*) of another state. The International Court of Justice (ICJ) has further clarified what the internal or external affairs of a state are in the Nicaragua case. Accordingly, a state is free to decide on its “choice of a political, economic, social and cultural system and the formulation of foreign policy” (International Court of Justice, 1986). The second element, coercion, does not necessarily need to take place in a forcible manner; rather, a state needs to seek to compel another state to change its behaviour with respect to its internal or external affairs, e.g. by forcing it to act in an involuntary manner or to involuntarily refrain from acting in a particular way (Republic of Austria, 2024).

Cognitive warfare techniques often aim at manipulating public opinion, electoral processes or governmental decision-making. Depending on the specific circumstances, these techniques are likely to transgress the *domaine réservé* of a state. The assessment of whether such activities are coercive in nature is more complicated and needs to be made on a case-by-case basis.

Although the element of coercion in this context is still debated among international lawyers, by now, a consensus in scholarship and state practice has emerged, that

- tampering with electoral infrastructure (e.g. cyber-attacks on vote counting machines that result in a miscount);
- the direct manipulation of the results of an election and
- the interference with a state’s ability to hold elections

would constitute a coercive interference with the *domaine réservé* of another state, irrespective of the means used by the interfering state (Kriener, 2023).

States can also indirectly disrupt a state’s ability to hold elections by engaging in activities directed at voters, e.g. by engineering voter suppression. One example would be that a state gives improper instructions on social media about voting in another state, such as the wrong location, or if a state circulates false information online regarding how and when to request, complete, and mail-in, absentee ballots (Schmitt, 2021). Merely influencing the target state by persuasion or propaganda or causing a nuisance without any particular goal, however, is insufficient to qualify as coercion (Schmitt, 2016, p. 747).

Nevertheless, the greatest success in affecting elections has been achieved by influencing the way voters think, rather than tampering with actual vote tallies. If a state engages in activities to influence the electorate's attitude towards a particular candidate or issue on the ballot, the legal assessment will be more complex. It seems reasonable to argue that (cyber) operations that deprive the electorate, or a substantial number of individual voters, of information bearing on the election, can be characterised as coercive. This would be the case, e.g. if a state conducted denial of service attacks against a campaign's social media presence. If a foreign state pushes information regarding candidates or issues to the electorate, a distinction needs to be made. While the publication of truthful information will most likely not violate international law, disinformation campaigns could manifest as coercive, depending on the specific circumstances (Schmitt, 2021).

In this regard, one could argue that false information has a coercive effect in that it manipulates the rationality of decision-makers. All rational decisions are based on facts. Presenting an inaccurate factual basis therefore restricts the freedom of choice by compelling those affected to shift their normative assessment of the options available. If this inaccurate factual basis is disseminated by state actors or actors attributable to a state, and there is a direct link between this intervention and the involuntary change in behaviour of another state, this constitutes a violation of the non-intervention rule by the former (Baade, 2020, pp. 166–67). Consider the following example: a foreign state that creates a deep fake in an election campaign in which a candidate purportedly admits to egregious criminal behaviour. The video is released just before Election Day with no time to counteract its effect. Since it will most likely have a decisive influence on the voting behaviour of citizens, such a conduct constitutes a prohibited intervention within the meaning of this rule (Schmitt, 2021, p. 750).

Another example for a violation of the prohibition of intervention would be a large-scale campaign launched by state A against the government of state B that spreads disinformation about the alleged corrupt business practices of state B's government, intended to sow distrust within the population of state B. The campaign eventually causes the government of state B to resign, resulting in a governmental crisis (Republic of Austria, 2024, para. 6).

3.1.3 Due Diligence

The due diligence principle in international law imposes an obligation on states to ensure that their territory is not knowingly used (both by other states and non-state actors, including private corporations) for activities contrary to the rights of other states (International Court of Justice, 1949, ICJ Rep 4, 22). A state that is or should be aware of such activities has the obligation to take all reasonable measures that are appropriate to prevent the violation and end the activity. The due diligence obligation is therefore an obligation of conduct, not one of result. However, it does not require that a state monitors activities on its territory and due diligence may not be taken as an excuse to infringe on human rights (Schmitt, 2016).

The due diligence principle constitutes an “objective” international standard of vigilance and care which is reasonably expected from an average “good government” that is mindful of its international obligations and acting in good faith, in the light of the potential risks in a given situation (Republic of Austria, 2024, para. 10). Accordingly, the required reasonable measures to be taken by a state under the due diligence obligation will vary depending on the situation and the respective state’s capacities and therefore need to be determined on a case-by-case basis. In the cyber context, putting in place a minimum of governmental ICT infrastructure and capacity to enable the exercise of diligence with regard to cyber activities on one’s own territory, might be considered a first reasonable measure to be taken by states under the due diligence obligation.

In the context of cognitive warfare, this principle is especially pertinent, as activities under this umbrella often aim at exploiting the non-kinetic nature of cyberspace and the existing challenges in the attribution of internationally wrongful acts to states. Such challenges mainly arise due to the increasing number of non-state actors engaging in such activities and the use of proxies. Even though the problem of attribution amounts to a technical and factual challenge rather than a legal one, the due diligence obligation does not require attribution for invoking state responsibility and therefore circumvents this problem. It offers a normative basis to hold states accountable and encourages international cooperation to mitigate harmful effects.

How a violation of the due diligence principle could occur, is demonstrated in the following example: State A is aware of a criminal hacker group operating

on its territory and targeting other states. The group is conducting a cyber-attack against state B, encrypting all the files of state B's Ministry of Foreign Affairs and requesting a ransom payment. State B is able to trace the activity to state A, but is unable to clearly attribute the attack to state A. It notifies state A of the cyber activity by the group emanating from state A's territory. State A rejects the allegation and refuses to investigate the conduct on its territory. Thereby, state A violates its due diligence obligation (Republic of Austria, 2024, para. 11).

3.1.4 Prohibition of the Threat or Use of Force

According to Article 2 (4), UN Charter states “shall refrain in their international relations from the threat or use of force against the territorial integrity or political independence of any state, or in any other manner inconsistent with the Purposes of the United Nations”. This obligation is a peremptory norm of customary international law. The only exceptions to it are the use of force in cases of self-defence (in accordance with Article 51 UN Charter) or when authorised by a UN Security Council mandate. In 1996, the ICJ confirmed that the prohibition set out in Article 2 (4) UN Charter applies to any use of force, regardless of the weapons employed (International Court of Justice, 1996, ICJ Rep 226, 39).

Although there is no authoritative definition of, or criterion for, “threat” or “use of force”, states have clarified that economic coercion and mere political pressure do not amount to a threat or use of force within the meaning of Article 2 (4) UN Charter (UN Special Committee on Friendly Relations, 1970). While this provision was indeed originally designed with the idea to prohibit the use of kinetic force by states, non-kinetic acts do not automatically fall outside its scope. Quite on the contrary, by now the majority of states has adopted a so-called “effects-based approach” to the interpretation of this rule. The terminology of “scale and effects” was first introduced by the ICJ in 1986 stating that those are the criteria to be considered when determining whether particular actions amount to an armed attack (International Court of Justice, 1986). Thus, an activity conducted by non-kinetic means is understood to constitute a threat or use of force if its scale and effects are or would be comparable to those of a kinetic threat or use of force. (Schmitt, 2016, Rule 69).

Therefore, a non-kinetic activity, e.g. a cyber-activity that leads to injury, death or significant physical damage constitutes an unlawful use of force. One

concrete example for that would be that state B conducts a cyber-activity against an electric power station situated in state A, on which a large part of state A's population and governmental infrastructure relies. As a result, the power station is severely damaged, causing a blackout for several hours until power supply can be restored. Such a cyber-activity would constitute an unlawful use of force (Republic of Austria, 2024, para. 7).

An activity, however, does not necessarily need to lead to physical damage in order to be qualified as a use of force. Depending on the specific circumstances and the severity of the consequences, an operation that merely causes deactivation of or prevents access to critical infrastructure of a state could also constitute an unlawful use of force. Similarly, disinformation campaigns may constitute an unlawful use of force if they lead to consequences comparable to those of kinetic attacks. This could be the case, for instance, if state A launches a public campaign in state B, spreading fear and terror and calling for acts of violence. As a direct result, the population of state B begins to attack each other and several people are killed.¹⁶

On the contrary, non-destructive cyber psychological operations intended solely to undermine confidence in a government or a state's prohibition of e-commerce with another state designed to cause negative economic consequences do not qualify as use of force (Schmitt, 2016, Rule 69 para.3).

The determination whether or not a situation constitutes a threat or use of force may potentially be more straightforward in terms of kinetic attacks, being very likely to manifest in physical damage or injury. Non-kinetic acts on the other hand require a more in-depth analysis with regard to their legal classification. However, this should not imply a weakness or obsolescence of this rule when dealing with the contemporary reality of threats. Quite on the contrary, the international community, by interpreting this obligation in a manner that takes into account new technological developments, has once again confirmed their commitment to the existing legal framework and emphasised its flexible, coherent and forward-looking nature.

16 A public campaign attributable to a state that calls for acts of violence against a national, ethnic, racial or religious group with the intention to destroy this group, could also amount to public incitement to commit genocide in violation of Art. III (c) of the Genocide Convention (cf. Republic of Austria, 2024, 9).

3.1.5 International Humanitarian Law

International Humanitarian Law (IHL) applies to acts of cognitive warfare whenever such acts are conducted in the context of an armed conflict. Information operations have been used to influence or mislead military adversaries or the civilian population during war since ancient times, however the scale, speed and reach of those operations has drastically changed in the last years, increasing the risk of harm to civilians and protected persons under IHL. It is therefore all the more important to examine the existing IHL framework in that context.

Generally, IHL allows for ruses of war, including acts that are intended to mislead an adversary or to induce him to act recklessly, but only if no rule of international law applicable in armed conflict is infringed thereby. Additionally, civilian and military leadership of a party to an armed conflict must refrain from ordering or encouraging IHL violations by their own forces. This prohibition extends to using digital platforms or messengers to encourage violence against civilians, civilian objects, or wounded, sick or detained enemy soldiers (Rodenhäuser & D'Cunha, 2023).

Additionally, acts or threats of violence with the primary purpose to spread terror among the civilian population are explicitly prohibited under IHL, irrespective of the means used. Therefore, e.g. the hacking into communication networks to propagate false air raid alarms, if such activities are designed to keep the inhabitants in a state of terror or to displace them, would be prohibited under IHL.

Even though spreading fear and terror among the enemy armed forces is not explicitly prohibited under IHL, e.g. threatening to kill, rape, torture or ill-treat captured or wounded soldiers would constitute a violation of IHL. Using information operations to diffuse these messages is consequently also prohibited (Rodenhäuser & D'Cunha, 2023).

3.2 Possible Measures in Response to an Internationally Wrongful Act

3.2.1 State Responsibility and Countermeasures

Generally, a state can be held responsible under international law for acts or omissions that are attributable to it and that constitute a breach of an international legal obligation to another state. The customary international law rules of state responsibility are largely codified in the International Law Commission's Articles on Responsibility of States for Internationally Wrongful Acts (ARSIWA) (International Law Commission, 2001). Accordingly, an activity is attributable to a state if it is conducted by a state organ, a person of governmental authority or by a non-state actor acting on the instructions of, or under the direction or effective control (International Court of Justice, 2007, ICJ Rep 43, para. 400) of a state. It must be pointed out that state responsibility can be invoked for conducts including kinetic as well as non-kinetic means and the abovementioned rules of state responsibility apply regardless of the means employed.

In the context of cognitive warfare, states may be held responsible for a variety of acts. State responsibility could be invoked, for instance, when a state violates the principle of non-intervention as described in Chapter 3.1.2. Similarly, a state may be held accountable if it allows its territory or infrastructure to be used by hackers to carry out unlawful cyber operations against another state, thereby violating their due diligence obligation.

However, attribution remains one of the most significant hurdles in invoking state responsibility. Cognitive warfare often involves cyber operations and disinformation spread via social media, encrypted channels, or anonymous networks, making it difficult for a state to prove that a specific act is attributable to one or more states. Without this technical attribution however, the legal attribution to a state, which naturally depends on and builds upon the technical one, becomes impossible. This issue is of a factual nature rather than a legal one and can therefore not be addressed through legal measures, even though the due diligence obligation may provide a certain (legal) remedy. Ways to encounter attribution difficulties could include international cooperation, capacity-building and strengthening resilience.

If an internationally wrongful act is attributable to a state or if a state has acted in violation of its due diligence obligation, the victim state may have the right to take countermeasures in order to compel the responsible state to cease its unlawful conduct and to make full reparation. When doing so, the victim state needs to make sure that countermeasures are proportionate to the harm suffered and do not involve the use of force referred to in Article 2 (4) UN Charter (Article 50 and 51 ARSIWA). Furthermore, a state that takes countermeasures must notify the offending state and suspend them once the unlawful activity has ended.

In the context of cognitive warfare, potential countermeasures could include, depending on the suffered violation, public (political) attribution and diplomatic measures, imposing sanctions, defensive information operations (e.g. factual rebuttal campaigns, strategic communications to neutralise false narratives) or cyber countermeasures (e.g. blocking or retaliating against malicious cyber operations as long as they do not amount to the use of force). Preventive measures, such as exploiting vulnerabilities in other states' ICT networks or placing "logic bombs" therein, cannot be justified under the law of state responsibility (Republic of Austria, 2024, para. 9).

3.2.2 The Threshold of Armed Attack and the Right of Self-Defence

A state may exercise its inherent right of individual or collective self-defence if an armed attack within the meaning of Article 51 UN Charter occurs. There is no general definition of an "armed attack" in international law, but the ICJ has clarified that not every use of force constitutes an armed attack in the sense of Article 51 UN Charter (International Court of Justice, 1986). Generally, it is recognised that force must be used on a relatively large scale, be of sufficient gravity and have a substantial effect in order to meet the criteria of an armed attack. Both, state and non-state actors can commit armed attacks.

The central issue in the context of cognitive warfare is whether non-physical operations, such as cyberattacks, disinformation and psychological manipulation could under any circumstances be qualified as an armed attack. As with the use of force, states generally approach this question by adopting an effects-based approach. Thus, a cyber-operation constitutes an armed attack if, as in the kinetic context, it causes significant death or injury to persons, or substantial material damage or destruction (Republic of Austria, 2024, para. 7).

This is, however, also context-dependent and subject to a case-by-case analysis. An armed attack can also consist of a series of attacks (International Court of Justice, 2003, ICJ Rep 4, para. 64).

At the same time, conversely this means that those acts that do not cause significant death or injury or substantial material damage or destruction do not amount to an armed attack in the sense of Article 51 UN Charter. Therefore, as Article 51 UN Charter is currently understood by the international community, psychological and disinformation operations do not constitute an armed attack, if they do not causally lead to the abovementioned effects or are not accompanied by kinetic operations causing the required effects. This is notwithstanding the fact that such operations may constitute other violations of international law.

If an armed attack according to Article 51 UN Charter occurs, the state suffering the attack has the right to self-defence. The use of force in self-defence must be necessary and proportionate in order to end the attack.

3.3 Human Rights Implications

States are bound by their human rights obligations as laid down in international instruments, such as the International Covenant on Civil and Political Rights (ICCPR) and the International Covenant on Economic Social and Cultural Rights (ICESCR). Moreover, regional instruments may be applicable, such as, for European states the European Convention on Human Rights (ECHR) and its protocols, the Charter of Fundamental Rights of the European Union (EUCFR), the European Social Charter and the Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data.

The same rights that people have offline must also be protected online (Human Rights Council, 2012). While states are the primary obligated parties, business enterprises, regardless of their size, industry, operational context and structure, are also required to respect human rights. States are therefore also under an obligation to ensure respect for human rights with regard to (cyber) activities by third parties, including large corporations.

Within their jurisdiction, states must hold third parties, including business enterprises, accountable if their operations, products or services lead to human rights abuse through harmful activities (Republic of Austria, 2024, para. 13).

Cognitive warfare can affect human rights, as it may target the foundation of democratic participation, individual autonomy and psychological well-being. The manipulation of information, the spread of disinformation, and the strategic use of psychological operations, be it through the use of cyber means or not, may violate the freedom of expression, including the freedom to seek, receive and impart information (which includes the right to internet access), the right to privacy, as well as the right to freedom of association. Any interference with human rights, be it online or offline, requires a legal basis, must pursue a legitimate aim, and must be necessary and proportionate to the aim pursued (UN ICCPR Human Rights Committee, 2011).

States must take a proactive role in protecting their citizens from human rights violations potentially caused by cognitive warfare. This may include not only holding perpetrators accountable, but also implementing safeguards against the misuse of digital platforms and promoting media literacy to mitigate the effects of disinformation.

3.3.1 Right to Freedom of Expression

According to Article 10 ECHR and Article 11 EUCFR, everyone has the right to freedom of expression. This right shall include freedom to hold opinions and to receive and impart information and ideas without interference by public authority and regardless of frontiers. It may only be restricted when provided by law and deemed necessary (Article 19 ICCPR).

In the context of cognitive warfare, the spread of disinformation and the manipulation of public narratives often undermine individuals' ability to form opinions based on accurate and truthful information. The spread of false or misleading information may influence political decisions, elections and public opinion, depriving individuals of the ability to access factual and trustworthy information, which could impair their freedom of expression and their right to make informed decisions. Similarly, state-sponsored or private sector manipulation of media outlets or on social media platforms can severely limit access to diverse viewpoints, curbing the free exchange of ideas and information.

3.3.2 Right to Privacy and Data Protection

Everyone has the right to respect for his private and family life, his home and his correspondence. No one shall be subjected to arbitrary or unlawful interference with his privacy, family, home or correspondence, nor to unlawful attacks on his honour and reputation. Everyone has the right to the protection of the law against such interference or attacks (Article 17 ICCPR).

Moreover, Article 8 EUCFR states that everyone has the right to the protection of his or her personal data. Such data must be processed fairly for specified purposes and based on the consent of the person concerned or some other legitimate basis laid down by law. Everyone has the right of access to data, which has been collected concerning him or her, and the right to have it rectified.

Cognitive warfare operations often involve the collection, analysis, and exploitation of personal data to construct personalised disinformation campaigns aimed at influencing individuals' thoughts and behaviours. Such operations can rely on social media platforms to gather personal data for psychographic profiling. This personal data may then be used to target individuals with tailored propaganda, political ads, or fake news, often without their consent or knowledge. Such conducts, depending on the specific circumstances, may raise serious concerns as to their compatibility with the right to privacy as well as the right to personal data protection.

3.3.3 Right to Integrity

According to Article 3 EUCFR, everyone has the right to respect for his or her physical and mental integrity. The manipulation of individuals' perceptions, beliefs, and emotions through psychological operations, targeted misinformation, and propaganda can potentially lead to significant mental harm. Cognitive warfare often aims to undermine public trust in institutions, destabilise societal norms, and even provoke fear and division. Individuals targeted by cognitive warfare operations, particularly disinformation campaigns, may experience cognitive dissonance, stress, and anxiety as they try to reconcile conflicting information. While a violation of the right to integrity will need to be assessed on a case-by-case basis, such a violation may appear particularly likely, when cognitive warfare operations deliberately induce massive psychological distress or aim at manipulating and playing individuals off against each other on a large scale.

3.3.4 Right to Participate in Political and Public Affairs

Article 25 ICCPR guarantees the right of citizens to vote, to hold public office, and to take part in public debates on matters that affect their communities and nations. Cognitive warfare operations could affect this right when interfering with democratic processes of free and fair elections.

Disinformation campaigns targeting voters, candidates, or political parties can influence the outcome of elections by distorting the truth, spreading fear, and confusing voters about their choices. Moreover, operations targeting vulnerable or marginalised groups can suppress voter turnout or manipulate the voting behaviour of certain demographics. By exploiting societal divisions and using disinformation to create distrust in the electoral system, cognitive warfare can erode the integrity of democratic participation. States therefore need to ensure that their citizens can freely participate in the political process, without the choices available being distorted and them being deprived of the opportunity to make informed decisions.

3.4 Conclusion

From a legal point of view, two main takeaways can be identified when it comes to cognitive warfare. Firstly, the existing legal framework can be considered robust, comprehensive, carefully developed as well as technology agnostic. As such, it is fully applicable to new and emerging technologies, including in the context of cognitive warfare. Existing law provides both, general rules and principles that states need to adhere to in their international relations as well as the possibility to hold them responsible in case of violations. The fact that international law by its very nature always requires a careful analysis of the specific circumstances and a case-by-case assessment when applied does not make its limits blurry, nor should this be taken as a reason to consider it outdated or obsolete; rather, this should be an incentive for states to continue discussing how international law applies in this field and identifying further common ground.

Secondly, current key issues in the cognitive domain, like attribution, the increasing number of non-state actors and the deliberate non-compliance with existing law by certain actors, despite naturally having legal implications, do

not have their roots in the legal sphere. Therefore, these issues cannot be addressed through legal means, including through the adoption of new legal instruments that some call for. Instead, those challenges would require other measures to be taken, including technical, political and educational ones. Apart from that, in view of the current geopolitical situation, the likelihood for the conclusion of a new universal treaty in the near future, particularly in an area that is as sensitive as the cognitive domain, is very low.

In conclusion, currently the best way for states to strengthen legal certainty and maintain the protection afforded by the law in view of these new threats and challenges is to commit to and apply the existing legal framework in good faith. This includes, among other things, that states continue to engage in in-depth discussions with other states and stakeholders on how the existing law applies to activities in the realm of cognitive warfare and to build upon existing convergences. In addition, states should consider taking further appropriate (non-legal) measures on a national, regional or international level, such as the exchange of good practices, capacity-building, political commitments and educational programmes.

4

Philip Schäfer:

Protection and Civil Liberties in the Digital Information Environment



Protection and Civil Liberties in the Digital Information Environment

Dr Philip Jan Schäfer is the Co-Director of the German Wargaming Center. He is a senior research fellow for the Chair of Political Theory at the Helmut-Schmidt University / University of the German Federal Armed Forces Hamburg. He previously held the role of lecturer for special tasks in the Department of Politics and Society at Bielefeld University. Prior to this, he served as a scientific advisor at the Bundeswehr Centre for Information Work. He completed his diploma in political science at the Free University of Berlin and went on to earn his doctorate at the University of Cologne where his dissertation on the effects of environmental change on security policy in the Arab MENA states was awarded the grade magna cum laude. During his doctoral studies, Schäfer began working in communications and headed the state campaign for sustainable mobility for the Baden-Württemberg Ministry of Transport. In his subsequent role as Assistant Professor in the Global Studies Department at Korea University in Sejong, Schäfer taught and conducted research on mobility issues and systems theory.

4.1 Introduction

Digital connectivity has made the information space a field of relevance to security policy, in which state protection mechanisms, societal resilience and individual civil liberties are inextricably linked. Against this backdrop, the Federal Ministry of Defence commissioned a political science study on the governance of the digital information space, with the aim of analysing and contextualising the key areas of tension in modern security policy in the digital age. The project brief defines two key questions for this purpose, which form the analytical framework of the study and provide the structural basis for the research report.

Both guiding questions are constitutive of the overall project because they address the two fundamental perspectives of democratic security architecture in the digital age. These are the normative perspective (the balance between protection and freedom) and the structural perspective (the risks of securitisation and the unintended effects of military instruments). They determine which theoretical concepts, political strategy documents and empirical developments must be examined, thereby structuring the entire research logic. The question of balance opens up the normative perspective on democratic requirements, whilst the question of securitisation highlights the structural consequences of security policy measures. Together, they enable a comprehensive, politically robust assessment of how democratic states can protect the digital information space without jeopardising the foundations of a free, open and pluralistic communication order.

4.2 First key question

The first key question – *How can the balance between protection interests and civil liberties be maintained in the digital information space?* – addresses one of the fundamental challenges facing modern democracies. It touches on the normative core of state action, for digital security can only be legitimate if it stabilises, rather than restricts, fundamental democratic rights. The question is central because the digital information space is simultaneously communications infrastructure, a forum for social debate and a field of action for external threats. State protective measures against disinformation, cyber-attacks or hybrid influence therefore directly intervene in those areas where freedom of expression, transparency and political participation are exercised. The key question ensures that security measures follow democratic logic and do not inadvertently become mechanisms for controlling or regulating public communication. It thus functions as a normative anchor that structures all further analytical steps and reveals the tension between protection and freedom, both theoretically and empirically.

4.3 Policy recommendations regarding the first guiding question

Conceiving security as an architecture of fundamental rights

Security measures in the digital sphere must not be conceived as instruments of control, but rather as a means of safeguarding fundamental rights and freedoms.

Make Privacy by Design and Security by Design mandatory

Privacy by Design and Security by Design must be made mandatory in procurement, IT strategies and standard contracts; only solutions with transparent, auditable security standards should be permitted.

Translating the European legal framework into operational practice

The GDPR, NIS2 and the Cyber Resilience Act require specialised monitoring and advisory bodies within ministries, as well as standardised risk, impact and fundamental rights assessments for new digital technologies.

Strengthening democratic resilience through education

Ministries should establish internal training programmes on digital security and the protection of fundamental rights, whilst also providing low-threshold information services for citizens.

Ensuring the ethical integrity of digital transformation

Interdisciplinary ethics advisory boards and mandatory ethics reviews in projects and decision-making processes ensure that digital measures remain socially and constitutionally accountable.

Making strategic use of Austria's role in the EU

Austria can contribute its tradition of freedom-based resilience to European governance approaches, thereby helping to ensure that European security mechanisms shape security as a prerequisite for democratic freedom, rather than as a means of control.

4.4 Second key question

The second key question – *What undesirable or unintended effects may arise from the deployment of military forces, how can these be counteracted, and how should the risk of securitisation be addressed?* – broadens the perspective to include the second central dimension: the institutional and discursive risks of security policy interventions in the digital sphere.

Militarised protection logics can unintentionally lead to permanent intervention routines, whereby temporary measures give rise to long-term control practices. This guiding question therefore focuses not only on specific security policy measures, but also on the way in which digital problems are defined. If phenomena in the digital space are primarily described as threats, there is a risk that the scope for political action will be permanently expanded or shifted by security policy logics. This question is therefore central to securing democratic stability in the long term and embedding security policy practices within a democratic framework.

4.5 Policy recommendations on the second key question

Institutional control rather than structural concentration of power

New security structures (e.g. Cyber Commands, Strategic Communications Units) must include sunset clauses, regular evaluations and safeguards to protect fundamental rights.

Actively prevent mission creep

Link security measures in the information domain to parliamentary mandates, clear definitions of objectives, exit criteria and independent evaluations, including for non-kinetic measures.

Strengthen communicative de-securitisation

Government communication should present threats in a sober and risk-based manner. Emphasise democratic resilience through education, media literacy and pluralism; keep military actors in the background in terms of communication.

Position Austria as a freedom-oriented actor in the EU

Austria should consciously distinguish itself within the EU as a champion of a security culture grounded in the rule of law and freedom, through fundamental rights impact assessments, transparency mechanisms and alliances with like-minded states.

Balancing protection and freedom in the digital information space

At the heart of this report lies the key question of how state protection mechanisms can be designed to effectively ward off digital threats whilst maintaining the balance between protection interests and the freedoms of open public communication.

4.6 Key findings on the balance between protection and freedom in the digital information space

In the digital information space, security cannot be achieved without freedom, and civil liberties cannot be upheld without security measures. The analysis shows that state protection against digital threats remains democratically legitimate only if it does not restrict civil liberties but actively stabilises them.

The balance between protection interests and fundamental rights is therefore not a political compromise or a negotiation over the extent of losses of freedom, but a structural relationship of interdependence. Security requires oversight under the rule of law, transparent responsibilities and technical standards that embed fundamental rights into digital systems before interventions become necessary. At the same time, freedom requires state protection mechanisms that safeguard the digital information space against manipulation, strategic influence and forms of attack, without regulating public communication.

The present findings illustrate that democratic resilience does not arise primarily through control, but through trustworthy and transparent security architectures.

Security measures must therefore be designed – technically, institutionally and communicatively – in such a way that they do not restrict freedom, but rather enable it. Only on this basis can the digital information space remain both protected and free.

4.7 The central role of the question of balance in democratic security architecture

The question of striking a balance between security interests and civil liberties in the digital information space is of fundamental importance, as digital security threats and state countermeasures are intensifying simultaneously.

State protection against disinformation, cyberattacks and strategic influence is increasingly portrayed as a security policy necessity. However, this also increases the risk that such measures themselves will infringe upon freedom of communication. The key question therefore concerns not only how digital security should be established, but also how democratic scope for action can be preserved.

In the digital information space, security measures affect not only technical infrastructures but also public communication spaces. They encroach upon spheres that enable political opinion-forming, social debate and democratic participation. If state security rationales are not clearly delimited, there is a danger that the safeguarding of democratic communication will give rise to the regulation of public debate. The key question therefore determines whether security architectures protect democracy or inadvertently restrict it. It determines whether state responses to threats become means of democratic stabilisation or mechanisms of control and surveillance.

Austria has a tradition of security policy that is more focused on civil protection, the rule of law and political neutrality than on military intervention. Historically, the country has tended to view security policy more as the protection

of the population and less as an extension of political power. At the same time, as a digitally connected state, Austria is exposed to significant threats from disinformation and hybrid influence. This situation creates a tension. Austria must develop security measures without thereby coming into conflict with its own culture of the rule of law.

Whereas other states rely on expansive security frameworks, Austria must develop solutions that work not through control, but through democratic stabilisation.

The question of balance is therefore not merely a technical or legal challenge, but a matter of political identity. Austria can only consistently implement security measures if they remain visibly compatible with fundamental democratic principles. This key question thus plays a decisive role in determining what kind of security state Austria wishes to be and what limits it sets on its own executive.

Within the European Union, Austria finds itself caught between two trends. On the one hand, the EU is increasingly developing regulatory measures to secure the digital information space, for example through legal requirements for platforms, cooperative security mechanisms and strategic resilience programmes. On the other hand, Austria maintains a strong focus on data protection, fundamental rights and state restraint in dealing with public communication.

This position can become a strategic advantage if Austria plays an active role in shaping European security architectures. The country could use its normative focus to legitimise security policy measures more strongly through fundamental rights and interpret European digital policy not merely as a defence mechanism, but as a means of democratic stabilisation. However, if the question of balance is not consciously addressed, there is a risk that Austria will merely adopt European security mechanisms rather than helping to shape them in line with its democratic tradition.

This makes the central question one of Europe's capacity to act. It determines whether Austria will become an active shaper of a freedom-oriented security culture or merely a recipient of external security policy standards.

4.8 Research

At the heart of the project lies the fundamental question of how the balance between security interests and civil liberties can be maintained in the digital information space. This issue is highly relevant against the backdrop of growing political, societal and technological challenges. On the one hand, there are measures to protect against disinformation, cyber threats or foreign influence; on the other hand, it is essential to safeguard core democratic fundamental rights such as freedom of expression, freedom of information and data protection.

4.9 Definition of securitisation

In the theory of the Copenhagen School, particularly as developed by Barry Buzan, Ole Wæver and Jaap de Wilde, securitisation refers to a discursive process through which a topic is transferred into the realm of security policy thinking by means of language. Crucially, a phenomenon is not recognised as a threat on the basis of objective characteristics, but is only constructed as an existential danger through a performative ‘speech act’. In this sense, ‘something only becomes a security problem by being designated as such’. This process simultaneously marks a shift from regular politics towards a state of emergency. If a threat is portrayed as existential, this legitimises extraordinary measures that are no longer subject to the usual rules of open political debate and democratic decision-making procedures. Security is thus not a given, but is constructed politically and linguistically.

In this theoretical framework, a political or social sphere, such as the digital information space, is considered to have been ‘securitised’ when several characteristics converge.

Firstly, the issue must be explicitly framed as relevant to security policy. Secondly, a discourse must emerge that portrays the survival of collective identities – such as the state, society or the democratic public sphere – as being under threat. Thirdly, extraordinary measures must appear justified, such as interventions by military or intelligence actors. This is typically accompanied by

dramatic language that emphasises urgency, crisis or the lack of alternatives, thereby crowding out regular political negotiation. The Copenhagen School's summary formulates this succinctly: *securitisation is the successful speech act through which an intersubjective understanding emerges within a political community to treat something as an existential threat*.

4.10 Methodological overview

To answer the key questions, a multi-stage research design was developed that combines theoretical, empirical and document-analytical elements. The starting point was a *systematic literature review* covering political science debates on protection interests, civil liberties and digital security in the European context. Through the structured evaluation of relevant research contributions, key concepts, lines of conflict and analytical perspectives were identified, which serve as the theoretical basis for further analysis.

Based on this literature, the research topic was *precisely operationalised*. The original guiding questions were translated into verifiable analytical categories that guide the subsequent empirical evaluation. This included the definition of relevant dimensions such as state powers of intervention, institutional responsibilities, technical security logics, guarantees of civil liberties, and discursive processes of securitisation. This operationalisation enables a systematic and comparable analysis of various political and strategic documents.

Building on this, an *analysis* was conducted of *state and multilateral strategy documents*, which serve as key sources of political self-description. The corpus of documents comprised security policy strategies, cyber and cyberspace concepts, as well as relevant European and national policy documents dating from 2010 onwards. Through qualitative document analysis, passages were identified in which states define risks, justify powers of intervention, normatively situate civil liberties, or define military roles. This evaluation reveals which security policy models prevail in the digital space and how they structure the relationship between protection and freedom.

The combination of these three methodological steps – systematic literature review, analytical operationalisation and document-based strategy analysis –

yields a robust overall evaluation. It enables both a theoretically sound and empirically verifiable assessment of how democratic states shape the digital information space, what risks arise in the process, and which security architectures can have a stabilising effect in the long term.

4.11 Systematic literature review

To facilitate a well-founded academic examination of this topic, a systematic review was carried out as a first step.

The aim was to provide a structured overview of existing political science discourses in which protection interests and civil liberties in the digital sphere are placed in relation to one another. To this end, the original guiding question was analytically refined and narrowed down to a testable research question. The revised research question is: *How is the balance between protection interests and civil liberties in the digital information space addressed in political science literature within the context of the European Union, with a particular focus on Austria?*

This formulation allows for the systematic examination of both overarching European perspectives and national particularities. Of particular interest here is which protection interests (e.g. with regard to disinformation, cyberattacks or the influencing of democratic processes) are emphasised, and in what way civil liberties in the digital space are conceptually situated and evaluated. In addition, we examined whether and how political science literature takes into account security policy frameworks, for example in the context of NATO or the EU Security Strategy.

In preparation for the systematic review, the subject area was precisely defined. The focus is on the tension between protection and freedom in the digital information space, understood as a specific arena of political negotiation within democratic systems. The geographical focus is on the European Union, with Austria being examined in depth as a case study. At the same time, overarching security policy considerations, such as the role of NATO, are taken into account where they are addressed in political science discourse.

In terms of discipline, the project is clearly situated within the field of political science. Detailed legal analyses, such as those concerning the interpretation of specific legal norms or the doctrine of individual statutory provisions, are deliberately excluded. Similarly, purely legal commentaries are not systematically included, as the focus lies on political science arguments, conceptualisations and empirical analyses. Both German- and English-language academic sources are taken into account in order to systematically capture different academic perspectives and discursive contexts.

Following the definition of the research question's scope, a systematic search strategy was developed to identify political science literature addressing the tension between protection interests and civil liberties in the digital information space. The aim was to create a structured foundation for subsequent analysis, covering both thematically relevant concepts and the geographical and disciplinary framework. To this end, key thematic and spatial terms were first identified. At the thematic level, a broad spectrum of terms was considered, encompassing both protection dimensions (e.g. disinformation, cyber threats, protection of democracy) and civil liberties (e.g. freedom of expression, fundamental rights, and digital sovereignty). In addition, terms relating to the regulation of digital public spheres (e.g. platform regulation, content moderation, EU Digital Services Act) were integrated. These terms were supplemented by a selection of geographical and political reference points. The focus is on the European Union, with particular attention paid to Austria. The security policy context, in particular the role of NATO and cyber policy cooperation at EU level, has also been incorporated into the search strategy.

The search formulas combine these terms using logical operators (AND, OR) to achieve results that are as relevant as possible, yet can be narrowed down. The formulations have been deliberately designed to cover both German and English-language specialist literature.

The systematic review was conducted using the relevant academic database, *Scopus*. This database offers comprehensive coverage of political science literature with interdisciplinary links, thereby providing a sound basis for further analysis. Clear inclusion and exclusion criteria were defined to ensure the relevance of the content and the academic quality of the literature included. Only academic publications in German or English were considered, in order to systematically capture both national and international perspectives. The

time frame was primarily set to include publications from 2010 onwards in order to reflect current political, technological and societal developments in the digital sphere as well as in the security policy debate. The focus was on works with a clear disciplinary basis in political science or related disciplines such as international relations, security policy or media policy. Both empirical studies and theoretical-conceptual contributions were included, provided they addressed issues relating to digital public spheres, platform regulation, disinformation, cybersecurity or fundamental democratic norms such as freedom of expression, data protection and digital sovereignty. Particular emphasis was also placed on works with an institutional link to the European Union and security policy cooperation structures such as NATO. Conversely, purely legal or technical-IT publications without a discernible political science analytical focus were excluded. Comments, editorials or journalistic articles lacking a scholarly foundation were also not considered. Literature with an exclusively national focus outside the EU context, particularly without a transatlantic dimension, was likewise excluded. Furthermore, articles without an abstract or with insufficient transparency regarding their content were not considered, nor were duplicates or incomplete publications.

To select and categorise the identified literature, the abstracts of all relevant articles were first reviewed. On this basis, a preliminary thematic classification was carried out. Subsequently, the selected articles were read in full to identify central lines of argumentation, theoretical reference points and dominant strands of discussion. The aim was not only to record individual case studies or findings, but also to map the overarching academic discourse, for instance with regard to fundamental controversies, recurring issues or conceptual breaks within the field of research. This approach enabled an initial structured engagement with the breadth and depth of the literature, as well as providing an analytical framework for further analysis.

4.12 Thematic search strands and key findings

The systematic literature review was based on three thematic search strands that reflect various dimensions of the tension between protection interests and civil liberties in the digital information space. Each search strand was operationalised using specific search formulas that combined key terms from

the fields of civil liberties, disinformation, and the protection of democracy, cybersecurity and European governance. In total, several hundred relevant publications were identified and evaluated on the basis of specialist relevance criteria.

The first search strand¹⁷ focused on the connection between civil liberties and the regulation of digital platforms. The search formula combined terms such as ‘freedom of expression’, ‘civil liberties’ and ‘fundamental rights’ with concepts of platform governance such as ‘content moderation’, ‘digital public sphere’ or ‘digital public sphere’. This was supplemented by geographical and institutional references to the European Union, as well as political science perspectives on governance and regulatory issues. This search strand resulted in a selection of eleven academic articles, two book chapters and one research note. The analysed contributions make it clear that the balance between protection interests and civil liberties cannot be achieved either through comprehensive deregulation or through heavy-handed state control. Instead, they advocate a nuanced regulatory framework grounded in the rule of law, which strengthens transparency, accountability and democratic oversight of platforms without restricting users’ autonomy. Key principles in this regard include non-discriminatory moderation practices, transparent interventions and institutional mechanisms to control platform power.

The second search strand¹⁸ focused on the topic of disinformation and the protection of democracy. The search terms included ‘disinformation’, ‘fake news’ and ‘misinformation’, as well as concepts relating to the digital public sphere and online-based political communication. The search focused on the European context, in particular the European Union and Austria. In this area, 32 academic articles, 10 book chapters, 3 monographs and 1 conference paper were identified. The literature consistently shows that disinformation is by no means an exclusively technical problem, but rather a complex social, political and legal phenomenon. European Union measures – including reg-

17 (“freedom of expression” OR “civil liberties” OR “fundamental rights”) AND (“digital platforms” OR “content moderation” OR “digital public sphere”) AND (“EU” OR “European Union” OR “Österreich” OR “Austria”) AND (“political science” OR “governance” OR “policy analysis”)

18 (“disinformation” OR “Fake News” OR “misinformation”) AND (“digital public sphere” OR “digital media” OR “online discourse” OR “information space” OR “digital platforms”) AND (“European Union” OR “EU” OR “Europe” OR “Austria”)

ulatory requirements, voluntary commitments by platforms and educational programmes – are described as a necessary but challenging mix. Their success depends crucially on whether they adequately protect fundamental rights, effectively address digital power asymmetries and empower the public to navigate the digital public sphere safely and with sound judgement.

The third search strand¹⁹ examined the relationship between security policy measures and fundamental rights in the digital sphere. The search formula linked terms such as ‘cybersecurity’, ‘cyber defence’ and ‘cyber resilience’ with concepts such as ‘fundamental rights’, ‘civil liberties’, ‘data protection’ and ‘privacy’, and incorporated security policy frameworks such as the European Union, NATO and international security cooperation. With 137 specialist articles, 91 conference papers, 75 book chapters and 5 academic books, this strand yielded the most extensive body of publications. The literature engages in intensive discussion on how security measures in the digital space must be designed to be effective whilst simultaneously safeguarding civil liberties. There are frequent calls for a risk-based, proportionate and constitutionally legitimate approach that integrates technological safeguards such as ‘Privacy by Design’ and ‘Security by Design’ into the development process of digital systems. The European legal framework – such as the GDPR, NIS2 or the AI Act – is highlighted as an ambitious attempt to combine security, innovation and the protection of fundamental rights. At the same time, the literature points to the danger that authoritarian states are increasingly using digital technologies for the control and restriction of civil liberties and are thereby driving forward the international fragmentation of digital standards.

Across all three lines of inquiry, it becomes clear that the balance between protection interests and civil liberties in the digital sphere can only be sustained if security measures are understood not as instruments of control, but as a protection framework grounded in fundamental rights. This requires, in equal measure, forward-looking regulation, embedded technical security mechanisms, public education and continuous democratic oversight. The

19 (“cybersecurity” OR “cyber threats” OR “cyberattacks” OR “cyber defence” OR “cyber resilience” OR “Cybersicherheit” OR “Cyberbedrohung”) AND (“fundamental rights” OR “civil liberties” OR “human rights” OR “Grundrechte” OR “Freiheitsrechte” OR “data protection” OR “Datenschutz” OR “privacy”) AND (“European Union” OR “EU” OR “NATO” OR “security cooperation” OR “European security” OR “international security” OR “Common Security and Defence Policy”)

literature makes it clear that security and freedom in the digital information space are not set up as opposites, but as structurally linked prerequisites for democratic stability. Taken together, the analysed papers demonstrate that the balance between protection and freedom in the digital space must be understood not as a static goal, but as a continuous process of negotiation. Platform regulation, cybersecurity and the protection of democracy are closely interlinked fields of action that can only be successfully shaped through integrative, rule-of-law-based and values-oriented strategies. It is crucial to ensure both technical resilience and normative orientation, not least in order to maintain Europe's position as a model of liberal-democratic digital policy in the international context.

4.13 Development of the analytical framework and compilation of the strategic document corpus

Following the completion of the systematic literature review, the validation and fine-tuning of the analytical framework marked the methodological start of the second project phase.

To this end, pilot analyses were first carried out on selected strategy documents, including the *EU Global Strategy* (2016), the *German National Security Strategy* (2023) and the *Austrian Security Strategy* (2024). The document corpus was compiled in parallel with the refinement of the methodology. This comprised national and multilateral security policy strategies dating from 2010 onwards, including documents from EU Member States, NATO member states and selected neutral countries. The corpus was systematically catalogued according to origin, year, institutional context and security policy focus. The aim was to create a representative basis for comparison that reflected both spatial and temporal variation. Where necessary, all documents were machine-translated and converted into an analysable text format so that they could be evaluated using computer-assisted methods in the subsequent phase.

This methodological and technical preparation provided a robust foundation for systematic analysis, establishing a reliable link between theoretical concepts and empirical text data.

Strategy documents, as the subject of this study, not only provide guidance in fields of action relating to security policy, but also implicitly determine the role that freedom and protection can play within state security architectures. Their significance thus lies less in concrete measures than in the politically legitimised pre-structuring of possible and excluded solutions. They determine which risks are recognised as relevant to security policy, which roles of actors are strengthened or limited, and which resources are to be made available to counter digital threats. This has a crucial implication for democratic theory, as strategic documents guide action even before operational decisions, regulatory acts or technical measures are formulated.

In this sense, security policy strategies function both as best practices and as a normative selection. They not only highlight what is deemed important, but also which alternative approaches to security – such as more control-oriented models or radically freedom-oriented deregulation – are deliberately excluded.

The balance between protection interests and civil liberties is therefore not negotiated only during subsequent implementation, but is politically predetermined at the level of strategic framework-setting. These documents reveal whether states view security in the digital sphere as the protection of democratic freedom of communication or as the control of public communication flows, and whether civil liberties are regarded as a stabilising asset or merely as a legal constraint on security policy interventions.

The way in which threats are described, responsibilities allocated and objectives formulated determines whether security policy leads to an expansion of democratic capacity for action or to the structural securitisation of public communication. The analysis of strategic framework documents thus enables an empirically grounded assessment of how democratic states attempt to strike a balance between protection and freedom, and to what extent this balance is shifted institutionally, technically or in terms of power politics towards control-oriented approaches. The insights gained therefore lie less in explicit statements about civil liberties and more in the identification of a political trajectory that is effectively shaped through resource allocation, structures of responsibility and normative frameworks.

4.14 Analysis of security policy strategy documents in the European context

An analysis of European security strategies clearly shows that the digital information space is no longer understood today as a purely civilian sphere of communication. States increasingly describe it as a space of political competition, strategic influence and hybrid threats. Consequently, security policy is shifting towards the protection of communication and information, with direct implications for civil liberties.

Although all the states examined define this balance as a core democratic task, they choose different institutional, technical and power-political approaches to ensure both protection and freedom. It is precisely here that it becomes apparent whether security architectures promote democracy or veer towards control-oriented logics.

Democratic states ground security normatively in freedom. Even at the normative level, the question of balance is not framed as a trade-off between two competing goods, but mutual interdependence. In its *White Paper on Security Policy and the Future of the Bundeswehr* (2016), Germany states that security policy serves the “defence of human rights, freedom, democracy, the rule of law and international law” and explicitly includes the digital sphere. Austria, for its part, defines in the *Austrian Security Strategy Paper* (2013) and again, even more precisely, in the *Security Strategy of the Republic of Austria* (2024), that security can only be guaranteed on the basis of a democratic state governed by the rule of law that protects “human rights, fundamental freedoms and, in particular, freedom of expression and freedom of the press”. The *EU Strategic Compass* (2022) also emphasises that the Union defends “citizens, values and interests”.

In all cases, freedom is understood not as a limitation on the scope for action in security policy, but as its democratic prerequisite. Security is legitimised not in spite of fundamental rights, but through them.

At the same time, the digital space is being comprehensively securitised. Virtually all strategies now describe the information space as a *conflict zone* used for disinformation, cyberattacks and covert influence. Germany already

recognised the internet as a ‘conflict space’ in its *White Paper* (2016), in which “ideologies of hatred and violence” are spreading. *Austria’s Security Strategy* (2024) refers to the cyber and information domain as an “increasingly contentious arena” in which malicious cyber activities, cybercrime and hybrid operations “jeopardise security and public communication”. The *EU Strategic Compass* (2022) goes a step further. Information manipulation is defined as a security policy attack that must be countered with a dedicated set of tools, ranging from hybrid task forces and cyber diplomacy to sanctionable mechanisms against foreign influence. Security policy thus transcends the traditional protection of military infrastructure and extends into societal communication spaces.

The balance between state protection against digital threats and the safeguarding of democratic freedoms in the information sphere is not achieved through case-by-case assessments, but through institutional architectures.

The strategies operationalise protection and freedom primarily through three structural elements.

- Constitutional safeguards (fundamental rights, proportionality, international standards)
- Technical and organisational resilience (privacy/security by design, critical infrastructure, cyber defence)
- Protection of democracy and strategic communication (combating disinformation, empowering the public)

Austria links these elements particularly consistently by defining security policy as a project for the rule of law and society that integrates both state and non-state actors (*Security Strategy* 2013, 2024). In its *National Security Strategy* (2023), Germany speaks of ‘integrated security’, which combines military, infrastructural and digital protective measures with societal resilience. The EU, for its part, is developing complex toolbox approaches that encompass not only protection but also punitive responses against foreign actors (*EU Strategic Compass* 2022). This makes it clear that democracy is not only intended to be defended but also to be made capable of acting, including the ability to intervene in the information sphere.

However, the question is where this capacity for action tips over into control mechanisms. This becomes particularly apparent in documents that define threats very strongly in terms of external enemies. *Lithuania's Threat Assessment* (2025) describes the information space almost exclusively as a “theatre of operations for hostile states”, particularly Russia and Belarus, whose propaganda, sabotage and hybrid tactics are to be countered using security state measures. Civil liberties hardly feature as active assets to be protected in such threat narratives, but remain in the background. The situation is similar in the UK's *National Security Strategy* 2025, which calls for a “hardening and sharpening” of national security. The text explicitly states that in the future “difficult compromises, potentially including between values and interests” will be necessary. This marks the first time it has been openly stated that civil liberties could, if necessary, be restricted in favour of security policy priorities – a formulation absent from continental European documents.

This reveals three types of democracy in terms of security policy balance.

- Freedom-based resilience strategies (Austria, traditionally Germany).
Freedom as an active security resource. Strong rhetoric on the rule of law.
- Governance and toolbox models (EU, partly Germany). Democratic legitimacy remains central, but institutions create instruments of power for enforcement.
- Securitisation/hardening strategies (Lithuania, increasingly the UK).
Information as an adversarial operational field. Greater readiness to intervene than explicit safeguarding of freedom.

The balance between security and freedom thus arises not as the result of individual decisions, but as a structure that can be predetermined by institutions, technical standards and security architectures. Strategy documents are therefore less political declarations of intent than normative policy directions. They determine whether democratic public spheres are designed as communication spaces to be protected, or as controllable fields of national defence against enemies. In this early framing, it becomes clear whether security policy strengthens freedom or absorbs it.

4.15 A typology of democratic security architectures in the digital information space

A typology is indispensable for deriving recommendations for political action, as it makes the underlying assumptions of security policy decisions transparent.

Rather than formulating abstract demands, it reveals what kind of security architecture a state pursues, which normative models shape this architecture, and by what institutional and technical means protection and freedom are organised. The typology thus reveals whether recommendations are based on freedom-oriented resilience, on state-enhanced capacity for intervention, on control-oriented defence, or on cooperative security.

This makes political recommendations not only comprehensible but also verifiable, as they clearly refer to the strategic logics on which their implementation is based. Anyone wishing to shape security policy must therefore first understand which model of security and freedom a state is actually pursuing. A typology provides both the analytical basis and the normative orientation for this.

An examination of European security strategies shows that democratic states today define the digital information space as a field of action relevant to security policy, yet decide in highly divergent ways on the extent to which state protection mechanisms may interfere with freedom of communication.

These differences cannot simply be understood as a spectrum between greater freedom and greater control. They each reflect specific conceptions of how security is created and what role democratic freedom of communication plays in this process.

Five ideal-typical models of state security architecture can be distinguished, which determine whether the state actively strengthens freedom, merely protects it, or possibly overrides it with security logic.

In the first type, civil liberties are not the object of state protection, but the very foundation of security itself. States of this type render security policy meaningless if it does not strengthen the capacity for action of a free public

sphere. This is particularly evident in Austria. The 2013 security strategy defines national security as the protection of high standards of fundamental rights and links this to civil society-led prevention. The 2024 security strategy builds on this by explicitly grounding security in freedom of expression, freedom of the press and freedom of information. Here, civil liberties are not a risk to be mitigated, but a democratic instrument of protection. This logic is also evident in the 2016 German White Paper on Security Policy. The defence of human rights, freedom, democracy and the rule of law is defined as the core of security policy, and the challenges of the digital sphere are to be addressed through social openness. Security thus arises not through control, but through stability under the rule of law and pluralism.

The second type maintains that freedom and security are of equal importance, but no longer regards freedom as a sufficiently stabilising factor. Protection becomes a task that must be actively enforced. These states therefore build up organisational, legal and technical capabilities to protect the digital space without fundamentally restricting its public nature. The European Union exemplifies this logic. The 2022 Strategic Compass defends citizens, values and interests, combining this normative stance with concrete instruments of political intervention. The Hybrid Toolbox, the Cyber Diplomacy Toolbox and measures against foreign information manipulation are expressions of a security architecture that not only permits defensive capabilities but also institutionalises them. With its 2023 National Security Strategy, Germany is moving in this direction by formulating integrated security and strengthening state coordination and the capacity for intervention in the digital space. Freedom remains central here, but is safeguarded by the state's ability to respond to threats in a targeted manner.

The third type understands security not primarily as the protection of the conditions for freedom, but as defence against hostile operations. The digital information space appears primarily as a theatre for external attacks. States such as Lithuania represent this model particularly clearly. The 2025 Threat Assessment describes information manipulation as an instrument of hostile states, and the state's response consists primarily of analysis, surveillance and countermeasures. Civil liberties remain as a fundamental concept, but they lose their operational significance. A similar shift can be observed in the UK's 2025 National Security Strategy. Whilst the document acknowledges normative principles, it explicitly states that difficult compromises between

values and interests may be necessary in the future. Freedom is thus no longer understood as a prerequisite for security, but as an element that can be set aside in an emergency. Protection becomes a control-oriented defence logic in which freedom remains symbolic but loses strategic influence.

In the fourth model, the focus is not on national intervention capabilities but on international cooperation. States that follow this model often have limited military resources or political traditions that view security interventions in society with scepticism. Protection is organised here through multilateral norms, international security systems and cooperation with the private sector. Ireland illustrates this clearly in its National Cyber Security Strategy for 2019 to 2024, which promotes resilience primarily through international cooperation and the development of societal capabilities. Similarly, the Swiss Cyber Strategy of 2018 describes security primarily as a matter of technical infrastructure, economic stability and international coordination.

The state hardly expands its powers of intervention, but rather strengthens participation, cooperation and the legal framework. Security thus arises less through control than through shared responsibility.

The fifth type shifts the focus from internal security to the geopolitical dimension of democratic order. Here, freedom is not merely defended but understood as a strategic resource embedded in global competition. States and organisations following this model view digital security policy as a means of projecting normative power. The European Union assumes this role by using data protection standards such as the GDPR, platform regulation such as the Digital Services Act, and diplomatic mechanisms such as the Cyber Diplomacy Toolbox to establish liberal norms in the digital sphere. The 2022 Strategic Compass explicitly describes this task as a systemic competition between authoritarian and democratic models.

Security is thereby not only guaranteed domestically, but also championed internationally as a European value proposition. Freedom is asserted globally, not merely protected.

4.16 An overview of ideal-type security architectures

Type I – Security derived from democratic structures

States following this model assume that security can only be achieved if democratic foundations remain stable and pluralistic. Civil liberties are understood not as a constraint, but as an active resource for protection. The central logic is to strengthen pluralism, fundamental rights and societal resilience in order to make the digital information space resilient.

Type II – Joint enforcement of security and freedom

Here, freedom and security are understood as objectives of equal importance, which must, however, be protected in practice. Freedom remains central in normative terms, but is safeguarded by the state's capacity for intervention and coordination. The core element is the institutional development of structures that enable the state to respond to digital threats without structurally restricting democratic public life.

Type III – The information space as a theatre of operations for hostile actors

In this model, the digital information space is primarily understood as an adversarial theatre of operations. While civil liberties remain a formal objective, they lose their operational significance. The core mechanism is characterised by surveillance, countermeasures and the state's security defence capabilities. Protection here arises primarily through control and readiness to respond to external threats.

Type IV – Security through international cooperation and societal resilience

States of this type rely less on national intervention and more on international cooperation, multilateral standards and societal resilience. Civil liberties are protected primarily through cooperation based on the rule of law, shared standards and the involvement of the private sector. The focus is on building technical, organisational and societal capabilities.

Type V – Security as a policy of global norms and values

In this model, freedom is understood as a geopolitical resource, the protection and export of which form part of a global competitive dynamic. Security policy aims to assert democratic standards internationally and establish them as a regulatory model. The central mechanism lies in the export of democratic norms, global regulation and the aspiration to shape the digital space normatively.

An analysis of these five types shows that the balance between security and freedom is not determined solely by specific decisions regarding interventions. It is predetermined by the fundamental architecture of state security strategy.

States that view security as an expression of democratic resilience actively strengthen freedom. States with toolbox-oriented strategies safeguard freedom through institutionalised defensive capabilities. Securitisation models risk restricting it. Cooperative models relieve the burden on state intervention agencies, whilst normative leadership models shift the balance into a geopolitical arena of tension. A democracy thus decides not only how it protects freedom, but also what relationship between freedom and security it enables in the first place.

4.17 Policy recommendations

Based on a systematic review of the literature and an analysis of European security strategies, it can be concluded that a democratic actor can only maintain the balance between security interests and civil liberties in the digital information space if security measures are understood as a framework for fundamental rights rather than as an instrument of control, and are underpinned by concrete, institutionally embedded practices.

For government departments, this means, firstly, mandatorily integrating technical security and the protection of fundamental rights into the design and procurement of digital systems by embedding principles such as Privacy by Design and Security by Design in IT procurement guidelines, tendering procedures and standard contracts, and only approving solutions that meet

transparent, auditable security standards. In parallel, service providers and IT staff should be specifically trained and certified in these approaches.

Secondly, the literature, as well as a review of strategy documents, shows that the existing European legal framework comprising the GDPR, NIS2 and the Cyber Resilience Act only protects fundamental rights if it is implemented operationally, for example through specialised monitoring and advisory bodies at ministerial level that both exercise oversight and support public authorities in the lawful and proportionate application of the law, as well as through the introduction of standardised risk, impact and fundamental rights assessments for new digital technologies.

Thirdly, democratic resilience requires systematic education and awareness-raising programmes. Ministries should establish internal training programmes on digital security and the protection of fundamental rights, whilst also providing publicly accessible, barrier-free information portals that offer citizens concrete guidance on the safe use of digital services and the protection of personal data.

Fourthly and finally, both academic literature and policy documents suggest that the management of the digital transformation must be structured in an interdisciplinary and ethically sound manner. This can be facilitated by internal ministerial ethics committees that assess digital projects in terms of their social and fundamental rights implications, as well as standardised ethics checks that are integrated into decision-making processes in a manner analogous to budgetary or sustainability audits.

For Austria within the EU, this presents an opportunity to make strategic use of its tradition of freedom-based resilience. By consistently implementing such fundamental rights and ethical principles within its own administration, whilst at the same time incorporating them into European governance and toolbox approaches, Austria can help ensure that European security mechanisms in the digital information space do not veer towards control-oriented approaches, but rather to shape security as a prerequisite for democratic freedom.

4.18 Risks of military intervention and strategies against securitisation

At the heart of this report lies the key question of what undesirable or unintended effects may arise from the deployment of military forces in the digital space, how these can be counteracted, and how the associated risk of securitisation can be responsibly mitigated.

4.19 Key findings on addressing securitisation in the digital information space

In the digital information space, security arises not only through state measures, but also through the very way in which risks are defined.

The analysis shows that digital phenomena are increasingly described as threats, the defence against which is said to justify extraordinary measures. This securitisation alters the scope for political action, as security logics legitimise interventions designed to deal with exceptional circumstances. It therefore not only reinforces protective measures but also expands the powers of state actors. The risk lies in these logics becoming permanent, with one-off interventions giving rise to enduring control practices. The risk of securitisation cannot therefore be addressed only once measures have been implemented, but must be tackled right from the stage of problem definition. The findings illustrate that democratic security architectures remain stable only if security measures are legitimised not through language and crisis rhetoric, but through comprehensible procedures, transparent responsibilities and limited powers of intervention.

Securitisation must therefore not be prevented, but controlled. Only in this way can security policy remain an instrument of democratic stabilisation and avoid developing into a permanent logic of control over public communication.

4.20 The strategic significance of the securitisation issue for democratic security architecture

The question of how to deal with securitisation determines whether security logics will permanently govern the digital information space. Security measures are intended to ward off threats, but must not themselves become interventions in free communication that regulate the political public sphere. The question is therefore not whether state responses are necessary, but how they are limited and controlled. Digital threats create political pressure to act, yet democratic security architectures must ensure that this pressure does not lead to a permanent expansion of executive power. Securitisation thus becomes not a technical problem, but a democratic challenge. The answer to this question determines whether security measures remain exceptional decisions or become established as standard instruments. It determines whether security architecture protects freedom or structurally alters it. The risk of securitisation therefore concerns not individual measures, but the fundamental form of state power in the digital sphere.

Austria has a strong security culture rooted in the rule of law, which traditionally binds state intervention to standards of proportionality. Its historical focus on civil protection and neutrality limits expansive security rationales and fosters expectations that state measures remain legitimate, subject to scrutiny and transparently limited. However, this fundamental stance is challenged when digital threats are portrayed as existential and security policy measures are increasingly regarded as indispensable. Austria must therefore not only develop security measures, but also at the same time prevent these measures from evolving into permanent powers that are no longer explicitly legitimised. The question of securitisation thus determines whether Austria maintains its security culture based on the rule of law or adopts security policy normalisations that do not fit with its own tradition. For Austria, it becomes a question of political identity. It determines whether security remains an expression of democratic restraint or becomes an instrument of state control.

Within the European Union, mechanisms for regulating and controlling digital communication processes are increasingly emerging. This development is driven by hybrid threats, but also by the demand for European capacity to act. In

this context, Austria faces the decision of whether to merely adopt European security architectures or to actively act as a freedom-oriented shaper.

As a state with a strong data protection profile and a tradition of executive power limited by the rule of law, Austria could assume a role within the EU that does not weaken security policy measures but rather safeguards them democratically. However, if the issue of securitisation is not consciously addressed, Austria could become a mere user of European security logics without having a say in their normative direction. The issue therefore determines whether Austria contributes to democratically balancing the European security architecture or whether it adopts external control logics that alter the character of its own security system.

In the European context, the issue of securitisation thus means not only protection against threats, but also protection against security policy dominance as a permanent structure.

4.21 Study

The sub-project focuses on the question of what unintended or undesirable effects may arise from the deployment of military forces, and how these can be addressed either preventively or reactively. Particular attention is paid to the concept of securitisation, that is, the political tendency to define certain areas as security issues and thereby treat them under exceptional conditions, with special powers of intervention and outside normal democratic procedures. This issue is highly relevant today, as military operations are increasingly taking place in hybrid contexts comprising both civilian and military components, and their legitimacy is increasingly being debated in democratic societies.

4.22 Methodological overview

To answer the second research question, a research design was developed that combines political science theory-building, a systematic literature review and a discourse analysis based on strategy documents. The starting point

was a systematic literature review that collates political science research on the unintended effects of the use of military force and on the concept of securitisation. For this purpose, the research question was narrowed down analytically to the question of how the unintended consequences of military operations and security logics are addressed in the specialist literature – initially focusing on international missions by the EU, NATO and the UN, and later expanded to include national and alliance defence, statewide security concepts and ‘Total Defence’ approaches.

The literature review followed a structured search strategy in the Scopus database. Four thematic concept fields were used: military intervention and defence policy (e.g. military intervention, security and defence policy), unintended effects (unintended consequences, mission creep, blowback), securitisation and security policy discourses (securitisation, emergency politics, militarisation), and mechanisms of democratic control (civilian control, democratic oversight, parliamentary accountability). This was supplemented by geographical and institutional references (EU, NATO, UN) as well as a disciplinary focus on political science, international relations, security studies, and peace and conflict research. The analysis included academic publications in German and English from 2010 onwards with clear relevance to political science; purely legal, technical and computer-science-related, journalistic or historical-descriptive works with no connection to current security policy debates were excluded.

On this basis, the identified contributions were analysed in terms of content in order to identify key problem areas relating to unintended effects (e.g. mission creep, blurring of roles, and the de-democratisation of discourse) as well as theoretical approaches to securitisation. The Copenhagen School (Buzan, Wæver, and de Wilde) served as the guiding conceptual framework for understanding security logics as a discursive practice. Security problems arise not solely from objective threats, but from their linguistic construction as an existential danger and the associated legitimisation of extraordinary measures.

Subsequently, national and multilateral security strategies were systematically analysed to determine how the digital information space and the deployment of military forces are framed in terms of security policy. Based on securitisation theory, an analytical grid comprising eight indicators was developed (including problematisation of the information space, existential threat semantics, nature of countermeasures, institutional anchoring, normative legitimisation

patterns, and training references). These indicators were used for the coded evaluation of European and Euro-Atlantic strategy documents in order to make the extent and form of securitisation comparable and to identify different security architectures – ranging from freedom-oriented resilience models to defence-centred hardening strategies – in ideal-typical terms.

The combination of systematic literature review, theory-guided operationalisation of securitisation and indicator-based analysis of strategy documents yields robust overall findings regarding the unintended effects that the use of military force in the digital domain may have, how security logics become institutionally entrenched – and which political, legal and societal mechanisms are suitable for democratically limiting the risk of securitisation.

To structure the academic debate, the central research question for the systematic literature review was refined analytically. The aim is to examine how unintended effects of the use of military force and the risk of securitisation are addressed in the political science literature, particularly in the context of international missions by the EU, NATO or the UN. This focus allows for the systematic identification of both conceptual and empirical contributions. The analysis centres on the secondary effects of military operations, such as destabilisation, social fragmentation, security-policy-legitimised democratic deficits, or the strengthening of repressive actors. At the same time, it examines how security-policy narratives are strategically employed in political decision-making processes, for instance to expand the state's scope for action or to legitimise extraordinary measures. Finally, the project explores which political, institutional or societal mechanisms are suitable for counteracting such dynamics. These include, amongst others, parliamentary limits on mandates, instruments of democratic control, security policy transparency, civil society participation, and the deliberate integration of alternative (civilian) security concepts.

As the project progressed, the original research question was expanded to systematically examine defence policies in the narrower sense. The resulting research question is therefore: *How are the unintended effects of the use of military force and the risk of securitisation addressed in the political science literature within various security policy contexts, particularly with regard to international missions, national and alliance defence, and national security strategies?*

This broadening of the scope of investigation enables a nuanced analysis of various fields of action in security policy. On the one hand, the focus will be on military missions abroad under multilateral mandates, such as those of the European Union, NATO or the United Nations. On the other hand, the focus will be on national and alliance defence, particularly in light of changing threat scenarios and new defence strategies within the context of European security policy. Furthermore, national defence approaches will be analysed, such as those discussed under the banner of 'Total Defence' or in concepts of civil-military resilience. Finally, the question is also examined of the extent to which security policy discourses and measures, in times of peace as well as in times of crisis, contribute to the securitisation of previously civilian or societal domains, for example in migration policy, public health situations or digital infrastructure. This systematic broadening of perspectives allows for a comparative examination of both the immediate and structural effects of the deployment of military forces and sharpens our focus on long-term transformation processes within democratic security architectures.

In preparation for the systematic literature review, a structured search strategy was developed that enables the identification of both theoretical and conceptual works on the concept and practice of the securitisation process as well as empirical studies on the unintended effects of the military use of force. The aim is to systematically search the political science literature for works that address security policy measures, their side effects and potential control mechanisms across various contexts.

4.23 Systematic literature review

The search strategy centres on four thematic concept fields. Firstly, key terms from the field of security policy and military interventions were included, such as 'military intervention', 'peacekeeping', 'military deployment', 'EU missions', 'NATO operations', 'external action' and 'security and defence policy'. Secondly, terms were included that refer to undesirable side effects and unintended consequences of military measures, such as 'unintended consequences', 'side effects', 'blowback', 'mission creep', 'escalation', 'civil-military relations' or 'destabilisation'.

A third cluster focuses on the concept of securitisation and its political effects. Here, search terms such as 'securitisation', 'security discourse', 'exceptionalism', 'emergency politics', 'militarisation' or 'politicisation of security' were included in the formulation. Finally, terms were also incorporated that relate to control and counter-mechanisms through which democratic societies can safeguard themselves against excessive security policy interventions, including 'civil oversight', 'democratic control', 'parliamentary accountability', 'mandate limitations', 'normative frameworks', 'legitimacy' and 'exit strategies'.

In addition, geographical and institutional reference points were included to focus the search results on security-related contexts with a European or international dimension. These include, amongst others, the terms 'European Union', 'EU', 'NATO', 'UN', 'United Nations' and 'Common Security and Defence Policy' or 'CSDP'. In terms of academic disciplines, the search was limited to works in the fields of political science, international relations, security studies, and peace and conflict research.

The individual search terms were systematically combined using logical operators (AND, OR). The search queries were formulated in both German and English to capture both national and international discourses. The actual search was conducted in the multidisciplinary academic database Scopus, which ensures broad coverage of relevant political science literature.

To further narrow down the relevant corpus of academic works, specific inclusion and exclusion criteria were established to ensure the quality, academic rigour and thematic relevance of the analysed literature. Only academic publications in German or English were considered, in order to systematically include both national and international discourses. The temporal focus was on publications from 2010 onwards in order to adequately capture current developments in security policy, theoretical approaches and discursive dynamics. Works with clear political science relevance were included, including those from related disciplines such as international relations, peace and conflict studies, or security studies. Both empirical studies, such as case analyses or comparative studies, and theoretical-conceptual and normative works were considered, provided they dealt with at least one of the following subject areas: military interventions, national and alliance defence, national defence strategies (such as 'Total Defence' or resilience approaches), unintended effects

of the use of military force, the concept of securitisation, or institutional and societal control mechanisms vis-à-vis the exercise of power in security policy.

Conversely, publications with no discernible connection to political science were excluded, in particular purely legal or technical-IT works lacking a discourse-analytical or political-theoretical context. Comments, editorials or journalistic articles lacking an academic foundation were also excluded. Literature with an exclusively national focus outside the European or transatlantic context was only considered if an explicit comparison or point of reference to the EU or NATO was established. Furthermore, contributions were excluded if they were published without an abstract or if their structure – for instance, regarding research questions, methodology or theoretical framework – lacked sufficient transparency to allow for a systematic analysis. Duplicates and works with a purely historical focus were also excluded, provided they did not demonstrate a connection to current security policy debates. These criteria ensure a robust academic foundation for the systematic evaluation and guarantee that the analysis focuses on works that make a substantial contribution to research into the use of military force, security policy discourses and democratic governance in the European and international context.

4.24 Thematic search strands and key findings

The systematic literature review addressing the second research question is based on three thematic search strands, which map out different dimensions of unintended effects of the use of military force and the securitisation of security policy discourses. In total, several hundred relevant academic publications were identified and evaluated on this basis.

The first search strand²⁰ focused on the question of which unintended effects arise in the context of national and alliance defence. To this end, terms relating to the side-effects of military policy – such as ‘unintended consequences’, ‘mission creep’, ‘blowback’, ‘side-effects’, ‘collateral effects’ or ‘public criticism’ – were combined with terms from defence and deterrence policy such as ‘territorial defence’, ‘national defence’, ‘collective defence’, ‘military readiness’, ‘alliance defence’ and ‘defence policy’, as well as with political science terms such as ‘security policy’, ‘defence policy’, ‘international relations’ and ‘security studies’. This search yielded 145 academic articles, 50 book chapters, 16 conference papers, 21 books, 7 reviews and 3 editorials. The analysed literature shows that the deployment of military forces within the framework of modern security and defence policy can be associated with a range of undesirable or unintended effects. Processes of alienation between politics and the public are frequently described, particularly when security policy decisions are not communicated transparently or justified in a comprehensible manner. In democratic systems, this undermines trust in state institutions and weakens the social legitimacy of security policy measures.

Furthermore, numerous contributions point to fear- and outrage-driven dynamics that are amplified by dramatised threat rhetoric, for instance in the context of cyber risks or hybrid threats. A security-centred policy that primarily imbues threats with emotional weight hinders rational deliberation and opens the door to measures that are politically difficult to control. Moral ambiguity is discussed particularly critically in the context of offensive cyber strategies, such as so-called “hackbacks”, where the line between defence and retaliation becomes blurred. Added to this is the danger of ‘mission creep’, i.e. the creeping expansion of mandates: the original mandate of a military operation

20 (“unintended consequences” OR “negative effects” OR “mission creep” OR “blowback” OR “controversy” OR “criticism” OR “public backlash” OR “risks” OR “unforeseen impact” OR “secondary effects” OR “unintended consequences” OR “side effects” OR “collateral effects” OR “negative repercussions” OR “consequential risks” OR “public criticism”) AND (“territorial defence” OR “national defence” OR “collective defence” OR “homeland defence” OR “strategic posture” OR “deterrence” OR “military readiness” OR “armed forces” OR “NATO Article 5” OR “national defence” OR “alliance defence” OR “national defence” OR “military readiness” OR “defence policy” OR “deterrence” OR “NATO”) AND (“security policy” OR “defence policy” OR “political science” OR “international relations” OR “strategic studies” OR “security studies” OR “Sicherheitspolitik” OR “Verteidigungspolitik” OR “Politikwissenschaft” OR “internationale Beziehungen” OR “Sicherheitsstudien” OR “strategische Studien”)

is watered down or overstretched over time, which creates strategic ambiguity and increases the risk of unintended escalations. As countermeasures, the literature recommends, among other things, transparent political communication, fact-based risk communication in dealing with disinformation and information operations, the expansion of institutional checks and balances, and clearly defined operational objectives with verifiable exit strategies. In theory, reference is often made to securitisation theory to illustrate how the construction of threats can lead to exceptional political powers and why security policy decisions should be negotiated in open, pluralistic discourses rather than under a state of emergency. Constitutional control mechanisms, a pluralistic media landscape, data protection principles and continuous reflection on the limits of legitimate security practices are identified as key prerequisites for a democratic balance between security interests and civil liberties.

The second search strand²¹ focused on securitisation in the context of national security strategies. The search terms combined concepts such as ‘securitisation’, ‘security discourse’, ‘emergency politics’, ‘militarisation’ or ‘state of exception’ with concepts of national defence and resilience such as ‘total defence’, ‘civil-military cooperation’, ‘resilience’ or ‘societal preparedness’, as well as with European and transatlantic reference points such as ‘European Union’, ‘EU’, ‘NATO’ and ‘Europe’. This strand comprised 12 academic articles, 5 book chapters, 2 books, a conference paper and a review. The literature makes it clear that the deployment of military forces in non-traditional fields of operation – such as healthcare, migration policy or internal security – is associated with specific risks.

The third search strand²² was explicitly dedicated to the protection of democracy and oversight in military defence policy. The search strategy combined terms such as ‘civilian control’, ‘democratic oversight’, ‘parliamentary

21 (“securitisation” OR “Versicherheitslichung” OR “security discourse” OR “emergency politics” OR “militarisation” OR “state of exception”) AND (“total defence” OR “gesamtverteidigung” OR “civil-military cooperation” OR “resilience” OR “societal preparedness”) AND (“European Union” OR “EU” OR “Austria” OR “NATO” OR “Europe”)

22 (“civilian control” OR “democratic oversight” OR “parliamentary accountability” OR “mandate limitations” OR “transparency”) AND (“military defence” OR “national defence” OR “alliance defence” OR “total defence” OR “armed forces”) AND (“democracy” OR “rule of law” OR “fundamental rights” OR “democratic legitimacy”)

accountability', 'mandate limitations' or 'transparency' with terms relating to military defence (such as 'military defence', 'national defence', 'alliance defence', 'total defence', 'armed forces') and democratic theory concepts such as 'democracy', 'rule of law', 'fundamental rights' and 'democratic legitimacy'. This strand comprised 10 academic articles, 5 book chapters, 2 books and one review. The literature shows that unintended effects of military operations in civilian, humanitarian or societal contexts are not only operational, but also above all discursive and structural. Three problem areas stand out: firstly, the securitisation of societal challenges and the accompanying de-democratisation of political discourse, as highlighted by Aradau and Huysmans, amongst others. Political complexity is being replaced by operational measures, and technocratic solutions are supplanting deliberative processes. Secondly, the blurring of the boundary between civilian and military roles, for instance in development cooperation, humanitarian aid or health protection, which weakens civilian institutions, politicises humanitarian principles and undermines trust in neutral actors. Thirdly, the shift in political power relations in favour of executive decision-makers who use military instruments to steer political processes. This can undermine parliamentary control, deplete public spaces and legitimise authoritarian practices under the guise of 'security'. The literature cites as counterstrategies the clear functional and institutional separation of civilian and military action, the expansion of critical security research, the integration of postcolonial and feminist perspectives, the promotion of alternative narratives, and deliberately designed de-securitisation of social problems.

The academic contributions compiled in the Scopus document make it clear that the deployment of military forces in non-traditional fields of operation, such as healthcare, migration or internal security, is associated with a range of undesirable and unintended effects. These can be divided into several problem areas, which are critically examined in academic discourse.

A key area of concern relates to the securitisation of societal challenges. When complex social or political issues such as migration, pandemics or extremism are viewed primarily through the lens of security policy, there is a risk that they will be treated as threats rather than as political or humanitarian challenges. This security-oriented approach can reinforce problematic narratives, lead to sweeping measures and ultimately supplant democratic negotiation processes. Several contributions therefore advocate 'de-securitisation' approaches that specifically aim to depoliticise or de-dramatise of such discourses.

Another problem area lies in the normalisation of military presence in the civilian sphere. Particularly in the context of crises such as the COVID-19 pandemic, it became clear how quickly military actors encroach upon areas of responsibility traditionally handled by civilian institutions. This can not only lead to a shift in institutional responsibilities, but also carries the risk of long-term structural changes in which the military as a “problem solver” is embedded in everyday life, with potential implications for the state’s monopoly on the use of force, the rule of law and public trust.

Added to this is the problem of role confusion and overreach. The involvement of the military in police or administrative tasks often leads to friction, inefficient parallel structures and conflicts of competence. Particularly in multilateral operations or disaster relief, the contributions emphasise the need for clear mandates, limited deployment periods and coordinated ‘multi-agency’ approaches to ensure the effectiveness and legitimacy of state action.

A fourth risk arises from the legitimacy issues and political repercussions of military action in civilian contexts. The contributions show that authoritarian tendencies, mistrust of state action and the restriction of civil society spaces can be fostered by the security policy framing of sensitive issues. To counteract these dynamics, reference is made to the importance of public participation, parliamentary oversight and transparent communication.

Overall, it becomes clear that the undesirable effects mentioned do not necessarily occur, but that conscious engagement with the structural, normative and institutional conditions of military action in civilian contexts is required. The contributions therefore emphasise measures such as strengthening civil society control mechanisms, a clear separation of military and civilian tasks, ethically reflective training, and inter-institutional cooperation. Only through such countermeasures can it be ensured that military deployment does not become a standard instrument of political action and that the balance between security and freedom, efficiency and democracy is maintained.

4.25 An examination of the militarisation of strategic security documents

Following the systematic review of the literature, the study examines the extent to which the digital information space is securitised in national and multilateral security strategies. The focus here is on a discursive-sociological perspective, as shaped in particular by the Copenhagen School (Wæver, Buzan, and de Wilde). Rather than an essentialist understanding of security, this approach assumes that security is not objectively given, but is constructed linguistically and politically – by actors who frame certain issues as existential threats and thereby legitimise extraordinary measures.

This theoretical framework allows strategic documents to be read not merely as planning instruments or administrative routines, but as acts of ‘securitisation’. Strategy papers construct threat scenarios, identify key assets to be protected, assign specific roles to actors, and define the framework for legitimate state action. At the same time, they have a reciprocal effect on the organisations that adopt them: they shape self-image and mission profiles, influence training, communication and personnel policy, and thus contribute to the development of institutional security cultures. Strategies thus have a dual effect – discursive externally and structural internally.

It is precisely this action-guiding and order-establishing function that makes strategic documents a suitable subject of investigation for analysing securitisation trends in the digital information space. The link to the preceding systematic literature review stems from the mutual complementarity of theoretical and practical perspectives. Whilst the review systematically captures the academic debate on security interests, civil liberties and securitisation, the subsequent document analysis focuses on how these tensions are actually articulated in the official strategies of state and multilateral actors. In particular, the study examines indications of whether and in what ways the information space is being politicised in terms of security policy, for instance through the description of hybrid threats, disinformation, influence operations or cyber interventions as security-relevant challenges.

The choice of the Copenhagen School as a theoretical framework stems from its broad establishment in security studies, its empirical applicability and its

conceptual rigour. Its categories make it possible to systematically capture the discursive dynamics of strategic texts and render them comparable. At the same time, its sociological foundation opens up the possibility of critically reflecting on how institutional security cultures emerge and what consequences this may have for democratic order, public communication and civil liberties in the digital sphere.

To systematically capture these dynamics, an analytical framework has been developed that operationalises key dimensions of securitisation, ranging from the framing of certain phenomena as problems, through their portrayal as threats, to institutional responses, patterns of normative legitimisation, and links to training and organisation. This framework forms the basis for a comparative analysis of documents across different states, organisations and periods, and makes it possible to trace empirically the degree and form of the securitisation of the digital information space.

4.26 Operationalisation of securitisation in the digital information space

In order to systematically examine securitisation in the digital information space, these theoretical concepts must be precisely operationalised. The digital space differs from traditional security sectors – such as the military, territory or energy supply – due to its intangible, connected and asymmetrical structure. Threats do not manifest physically, but rather in the form of disinformation, digital influence operations, strategic destabilisation or attacks on decision-making processes and public debates.

This combination of the civil sphere of communication and the field of security policy intervention makes the digital information space a particularly sensitive and productive field of study for securitisation processes. The operationalisation is based on the discursive logic of the Copenhagen School, which focuses on the linguistic construction of threat and the legitimisation of extraordinary measures and applies this logic to the specific characteristics of security policy strategy documents. The analysis therefore examines not only whether the information space is addressed as a security policy issue at all, but also which terms and images are used to identify risks; whether

threats are portrayed as existential; whether military or intelligence measures appear necessary; whether urgency, crisis or the lack of alternatives are emphasised; and to what extent the information space is embedded within the national security architecture. Equally central is the question of whether institutional structures are being created or expanded – such as specialised units for strategic communication, cyber defence or hybrid threats – and which normative frameworks are invoked to legitimise these measures, for example the protection of democratic values, truth or public debate. Finally, the study examines whether the security-policy framing of the information space is also reflected in education, training and professionalisation, for instance through the integration of strategic communication into military curricula or official training programmes.

These dimensions enable a systematic empirical analysis of the securitisation of the information space by capturing both the discursive and the institutional levels. They reveal whether security logics are merely established rhetorically or are already organisationally embedded, thereby demonstrating whether securitisation is an episodic political reflex or a structurally defining component of modern security architecture.

4.27 Indicators of the securitisation of the information space

The problematisation of the information space

This is evident in terms such as ‘information environment’, ‘disinformation’, ‘hybrid threats’, ‘cyber-enabled influence’ or ‘hostile narratives’, which identify the information space as a security-relevant domain.

Portrayal as an existential threat

Evidence can be found in phrases such as ‘existential threat’, ‘destabilising’, ‘undermines democracy’ or ‘erodes trust’, which portray fundamental concepts of social order as being at risk.

Security logic and measures

Visible in the demand for or portrayal of measures by the military, intelligence services or state control apparatuses, as well as in terms such as 'strategic communication' or 'information resilience'.

Rhetoric of a state of emergency and the lack of alternatives

Key phrases such as 'must be countered', 'urgent', 'critical', 'decisive action required' or references to a 'whole-of-government approach' point to a logic of urgency and crisis.

Embedding in national security architecture

The information domain is portrayed as part of national defence, hybrid warfare or national security planning.

Institutionalisation

References to the establishment of new structures such as information warfare units, strategic communication units, cyber defence commands or inter-ministerial coordination bodies.

Normative orientation and patterns of legitimisation

Emphasis on values such as democracy, truth or the protection of public debate, often linked to phrases such as 'protecting democratic values' or 'truth'.

Integration into training and professionalisation

References to training, curricula or courses on strategic communication or information operations within the armed forces or government agencies.

4.28 The securitisation of the digital information space in European strategy documents

An analysis of European security strategies shows that the digital information space is increasingly no longer understood as a civilian sphere of communication, but rather as a security-relevant arena of conflict. Digital communication is portrayed as a potential arena for attack, linked to the rhetoric of existential threats and associated with security policy or military countermeasures. This

shift in the interpretative framework can be described as securitisation in the sense of the Copenhagen School, whereby language transforms political spheres of action into a state of exception, thereby legitimising extraordinary measures.

This shift is particularly evident in Germany. The 2016 White Paper explicitly describes digital communication as a security policy challenge for the first time. It states that it offers opportunities “to influence public opinion, for example through hidden attempts to sway discussions”. This wording shifts the focus from safeguarding open public discourse to countering covert influence. The 2023 National Security Strategy reinforces this logic by classifying disinformation within a series of existential crisis scenarios: “The COVID-19 pandemic, cyberattacks, disinformation campaigns – all these threats show how vulnerable we are.” By equating it with pandemics and cyberattacks, information policy is legitimised as a sphere of action within security policy.

An alternative model can be found in Austria. The 2024 Security Strategy explicitly defines security in terms of democratic freedoms: “human rights and fundamental freedoms, in particular the freedom of expression, are the main yardsticks.” Here, the digital space is addressed as a risk, but not as a reason for extraordinary security measures, rather as a basis for strengthening the rule of law. Securitisation remains limited; protection is seen as safeguarding freedom, not as control.

Strategies in Southern and Eastern Europe take a different approach. Spain’s 2017 National Security Strategy describes hybrid threats as a key danger to the state and society: “hybrid threats, a combination of conventional and unconventional threats aimed at destabilising our way of life”. The information domain is thus classified as part of “unconventional threats” and is treated not as a communications issue but as a defence issue.

The European Union goes particularly far in its Strategic Compass (2022), which not only identifies threats but also establishes concrete instruments of intervention to sanction disinformation. The text states: *“We will firmly respond to foreign information manipulation and interference, as we did through our decisive and coordinated action against Russia’s disinformation campaign.”* The terms “firmly” and “decisive action” point to a logic of action with no alternative, which legitimises security policy measures.

In Eastern Europe, finally, a distinctly clear case of securitisation emerges. Poland's 2020 National Security Strategy describes information operations as quasi-military means: "multi-faceted and comprehensive actions using non-military means (including: cyberattacks, disinformation) to destabilise the structures of Western states and societies". The information space is conceived here not as a social medium, but as a weapon. Lithuania's Threat Assessment 2025 further sharpens this perspective by describing Russia as an intentional adversary of free communication: "Russia seeks to become a digital totalitarian surveillance state", coupled with the aim of "full control over the internal information space". This existential framing ("totalitarian") creates a security policy obligation to defend against such threats and merges information protection with the defence of geopolitical identity.

Analysis of further European and transatlantic documents also confirms this trend. Austria's 2013 Security Strategy already identifies "cybercrime, cyberattacks, the misuse of the internet for extremist purposes" as "serious new challenges". The Dutch 2023 Security Strategy describes hybrid activities that encompass "the spread of propaganda and disinformation to influence public opinion". In 2024, Sweden states that "digital platforms offer new opportunities to disseminate information and plan violent acts", which is "amplified" by influence operations. Estonia warns that "international propaganda and the spreading of disinformation damage the security and defensive capability of Estonia". Finally, Portugal formulates a classic security policy framework by stating that state actors pursue strategic objectives through "interference and disinformation operations in a digital environment" and "endanger essential services". At the multilateral level, NATO permanently establishes the information domain as a defence sector and announces: "We will expedite our digital transformation, adapt the NATO Command Structure for the information age and enhance our cyber defences".

Across all documents, securitisation can be observed where

- the information space is designated as a strategic field of operations,
- threats are described as existential to the state, democracy or infrastructure,
- security policy or military countermeasures are legitimised,
- clear, normatively charged dichotomies (democracy vs. hostile manipulation) are constructed.

At the same time, the scope and normative limits of this securitisation vary. Whilst Austria (2013, 2024) and Portugal (2019–2023) do address the information space as a problem, they firmly embed it within fundamental rights and the principles of the rule of law; by contrast, Sweden (2024), Estonia (2023) and NATO (2022), for example, shift the communications sphere deeper into a defence logic. The digital information space is thus increasingly being imbued with security policy significance in Europe and within the Euro-Atlantic alliance and, in part, removed from the realm of democratic negotiation – precisely the process that the Copenhagen School describes as securitisation.

A comparative analysis of strategic security documents shows that European states are increasingly defining the digital information space as a conflict area relevant to security policy and thereby transferring it into the logic of securitisation. Drawing on the Copenhagen School, communication, digital infrastructure and disinformation are no longer understood primarily as political, social or media challenges, but as potential instruments of hostile influence. This discourse shifts information policy from the realm of regular political negotiation into a sphere legitimised by security policy.

4.29 Indicator-based criteria of securitisation

Coding according to eight indicator-based criteria of securitisation (including problematisation, existential threat semantics, security measures, institutional anchoring, and normative legitimisation) highlights clear differences in the extent of this security-policy framing.

The information space as a defence sector

The most pronounced forms of securitisation are found in Lithuania, Estonia and NATO. In these documents, the digital information space is explicitly framed as a domain of attack by hostile states, requiring exceptional, militarily organised countermeasures. The Lithuanian Threat Assessment (2025) states that Russia is striving to establish a “digital totalitarian surveillance state” and therefore seeks “full control over the internal information space”. The information space is thus imbued with existential significance and defined as the centre of geopolitical conflicts. NATO 2022 continues this logic at the alliance level by identifying “cyber, space, and hybrid and other asymmetric threats”

as central collective security tasks. The high frequency of these terms in the text reflects a lasting institutionalisation of security policy, including military responsibilities and command structures. Here, the digital space appears not as a public sphere, but as an operational battlefield.

Intervention in the face of existential risks

Poland, Sweden, Portugal and the European Union also closely link digital communication and disinformation to existential threats to democratic stability. The 2020 Polish strategy refers to “multi-faceted and comprehensive actions using non-military means (including: cyberattacks, disinformation) to destabilise the structures of Western states and societies”. The EU goes further and enshrines concrete instruments of defence: “We will firmly respond... through decisive and coordinated action against Russia’s disinformation campaign.” Here, securitisation is not merely discursive but operational. It leads to institutional mechanisms (Hybrid Toolbox, StratCom units) that define the information space as the object of active security policy. It is noteworthy, however, that this securitisation is always justified normatively, by reference to democratic values and European identity. Security is thus not formulated in opposition to freedom, but declared its prerequisite.

Security logic in balance

Germany (2016), the Netherlands (2023) and Spain (2017) fall within the middle range. They do clearly identify disinformation as a threat, for example, when Germany refers to “Hidden attempts to sway discussions” or Spain describes, “destabilising our way of life”. At the same time, these states refrain from using rhetoric associated with a state of emergency. Security measures appear legitimate, but not the only option. A preventive, coordinating and cooperative tone is typical. In these cases, securitisation remains limited and does not lead to a comprehensive military-institutional overhaul of communication policy.

Freedom as a security principle

Austria provides the counter-model. The 2024 Security Strategy states: “Human rights and fundamental freedoms, in particular the freedom of expression, are the main yardsticks.” This phrasing confirms that the information space is problematised, yet without existential threat narratives or far-reaching security-state measures. Austria projects security not as control over communication, but as the protection of the conditions of freedom within which communication can take place. Unlike Baltic or Polish strategies, Austria does

not define threats as a cause for special measures, but as a mandate for stabilisation under the rule of law.

4.30 How security strategies reinterpret the information space

The coding of European strategy documents shows that the digital information space is increasingly understood not as a civil infrastructure for public communication, but as a site of strategic contestation. This shift does not arise from objective threat situations, but is generated discursively, through the way in which states speak about disinformation, cyber activities and digital influence. From this perspective, the documented formulations illustrate not only which threats are perceived, but also what normative and institutional architecture the state is building to respond to them.

From this, five key patterns can be identified that reveal a gradual securitisation of the information space.

Firstly, it is evident that the degree of securitisation rises in direct proportion to the characterisation of the digital sphere as a tool of hostile states. Where disinformation is described as a political problem, information policy remains part of democratic governance. Where, on the other hand, it is perceived as a foreign-controlled operation – as is the case in Poland, NATO or Lithuania – the discourse shifts towards defensive logic. The more digital communication is perceived as an attack, the more likely it is to be treated as a matter of defence policy. In this sense, it is not only the problem that changes, but also the responsibility. What used to be the task of media policy, education and civil society becomes the task of the security apparatus.

Secondly, narratives of existential threat have a decisive rhetorical effect. Terms such as ‘destabilising our way of life’ or warnings of a ‘digital totalitarian state’ frame the digital space not as a risk, but as a threat to the survival of the political order. Such scenarios create a sense of urgency that transcends the realm of normal politics and legitimises measures that would otherwise be considered disproportionate. The Copenhagen School describes precisely

this moment as the transition from politics to security: the state of emergency is not declared, it is linguistically constructed.

Thirdly, it is crucial that securitisation does not stop at rhetoric. It only becomes permanent when it is institutionally embedded, through cyber commands, hybrid defence structures, strategic communication centres or EU toolboxes. These structures create new responsibilities, new resources and specialised expertise. As a result, there is a long-term shift in who protects social communication, who monitors it, and who decides when to intervene. Institutionalisation is thus the point at which securitisation becomes a security architecture.

Fourthly, it is striking that security documents rarely frame their measures as restrictions on freedom. Instead, freedom becomes a resource that must be defended. Values of freedom legitimise interventions, not in contradiction to them, but by reference to them. Democracy is defined here not as an open space for communication, but as a protected asset that must be shielded from external influence. This pattern of argument shifts freedom from the principle of openness towards a duty of control. Civil liberties are not abolished, but instrumentalised for security policy purposes.

Fifthly, the comparative analysis reveals two fundamentally different security cultures in Europe. States such as Austria (and to some extent Portugal) rely heavily on freedom-oriented resilience models, in which security arises from the rule of law, pluralism and social empowerment. Here, security is tied to freedom, not to control. Other states, particularly the Baltic states, Poland and NATO, are developing defence-oriented hardening models that conceive of the information space as a geopolitical theatre of operations and place military countermeasures at the centre. This distinction points to different political self-conceptions.

The question is not merely how to react, but what one actually believes needs protecting – communication as a freedom or communication as infrastructure.

4.31 Implications for democratic theory

Security and freedom become unbalanced where information policy is replaced by security logic. It is not the intervention itself that endangers freedom, but the structure within which it becomes possible. When security is defined through control mechanisms, threat rhetoric and institutional defence capabilities, the media and communications sphere is gradually removed from the political arena of negotiation. Decisions regarding public speech, discourse and influence are then no longer negotiated democratically, but managed through security policy. It is precisely this process of transformation that the Copenhagen School describes as securitisation.

Conversely, it is evident that where security is organised through the rule of law, transparency, pluralism and fundamental technical rights, freedom does not become a vulnerability but a prerequisite for resilience. In such models, the digital information space remains a political arena and does not become a militarily administered territory. This is precisely why the nature of the security architecture is crucial. It is not just a question of how we protect, but from what and in whose name.

4.32 Policy Recommendations

An analysis of the strategy documents shows that securitisation does not begin with the measures themselves, but already in the definition of the problem. Where disinformation is framed as a tool of hostile states and as an 'existential' threat, responsibility shifts from media, education and social policy to security agencies. This is precisely what is evident in strategies such as those of Poland, Lithuania, NATO or the EU Strategic Compass. The resulting recommendation is that democratic actors should describe the digital information space primarily as a political arena and only secondarily as a security arena. Problem formulations should clearly identify risks, but not rhetorically elevate them to a state of emergency. For Austria, this means maintaining the approach set out in the 2024 Security Strategy, orienting security towards "human rights and fundamental freedoms, in particular the freedom of expression", and ensuring that emerging digital strategies reflect

this principle. Where threats are described, it should always be specified which political, social and legal means are to be prioritised before military or intelligence instruments are deployed.

The literature on the unintended effects of the use of military force shows that military involvement in civilian spheres frequently leads to a blurring of roles, the normalisation of a state of emergency, mission creep and a silent increase in the power of executive actors. Analysis of strategic documents confirms similar trends in the digital information space, particularly where the information space is defined as an 'operations field' for hostile actors. This gives rise to several risks. Military involvement in cyber defence, information operations or the protection of critical infrastructure can blur the line between internal and external security, reinforce the perception that communication is a field of military planning, politically devalue civilian institutions (regulatory authorities, media oversight bodies, civil society), and in the long term lead to the normalisation of military presence in crisis and conflict situations.

This can only be addressed if military forces are consistently deployed in a subsidiary capacity within the digital domain. Austria should strictly limit military tasks in the digital information domain to traditional protective functions, such as providing technical support for critical infrastructure or the defence against clearly defined cyberattacks, whilst explicitly leaving the design of communication spaces in civilian hands. Any involvement of the armed forces should be subject to three conditions: a clear mandate with a narrowly defined purpose, a time limit, and civilian command responsibility.

The wording of strategic documents shows that securitisation becomes particularly entrenched when it is translated into structures: Cyber Commands, StratCom Units, toolboxes, dedicated coordination centres. This creates a 'security architecture' that continues to operate even after the original crisis situation has subsided. Democratic actors should therefore provide for binding sunset clauses for all new security structures in the digital space; introduce regular independent evaluations that assess not only effectiveness but also the impact on fundamental rights and political side-effects; and formulate the mandates of cyber or information security structures in such a way that they are explicitly bound to fundamental rights, freedom of expression and data protection. In practical terms, this means that when participating in EU instruments such as the Hybrid Toolbox or mechanisms to counter foreign

information manipulation and interference, Austria should work to ensure that all measures are subject to rule-of-law oversight mechanisms, transparency requirements and ex-ante fundamental rights impact assessments. Austria can draw on its own security strategy principles to systematically ask in Brussels what democratic safeguards each new security mechanism in the information space contains.

The literature on the unintended effects of military force deployment repeatedly highlights that mission creep, escalation through reactivity and the creeping expansion of mandates are central problems. An analysis of European strategies shows that precisely these dynamics threaten the digital space once the information space is established as a sector of security policy. Security and defence measures in the digital domain should always be tied to a clear parliamentary mandate, including regular reporting on effectiveness, side effects and implications for fundamental rights. Mandates require defined objectives and exit criteria, including for non-kinetic measures such as information campaigns or offensive cyber capabilities. The involvement of civil society expertise and independent research in evaluation processes should be institutionally guaranteed in order to limit self-referentiality in security policy.

For Austria, this means that the National Council should systematically have a say in all security policy measures in the digital information space. One model would be parliamentary reports on cyber and information security measures that explicitly address potential securitisation effects. In this way, Austria would effectively extend its own tradition of the rule of law to this new field.

The analysis shows how heavily key documents rely on existential threat narratives and thus fulfil the Copenhagen logic, according to which “it is by labelling something a security issue that it becomes one”. At the same time, there are counter-examples such as Austria, where security is explicitly linked to civil liberties.

This points to an active de-securitisation strategy. Government communications should present threats in a sober, risk-based and nuanced manner, without resorting to alarmist or sweeping narratives. When addressing disinformation and hybrid threats, it should always be made clear that democratic resilience is strengthened not primarily through control, but through education, media literacy, pluralistic debate and the rule of law. Military actors should remain

in the background in terms of communication, particularly when it comes to public communication, content moderation or platform regulation.

Austria has a particular role to play here. It can bring the emphasis on freedom of expression, press freedom and freedom of information, which is enshrined in its strategies, into European processes and systematically press for EU documents to always emphasise the role of democratic public discourse and civil society actors alongside the description of threats. Austria could, for example, advocate for EU strategies on disinformation to routinely refer to 'democratic resilience' and 'public deliberation' instead of merely 'countering threats'.

The comparative analysis reveals two security cultures in Europe: liberal, rule-of-law-based resilience models and defence-oriented hardening models. Austria clearly stands on the side of the former. The strategic question is therefore not whether Austria adopts European security mechanisms, but how. Austria should position itself as an advocate of a security culture oriented towards freedom and fundamental rights within EU bodies and in the further development of the DSA, NIS2, Cyber Defence Policy and Hybrid Toolbox. Based on its own strategies, it can propose systematically linking EU security instruments to fundamental rights and democracy impact assessments, particularly where military or semi-military capabilities are integrated into information-related matters. Austria can seek allies among states with similar cultures, such as those with freedom-oriented and cooperation-based models, and thus form a visible "liberal-democratic security caucus" within the EU that combines the capacity for action in security policy with structural restraint.

4.33 Conclusion

The balance between protection interests and civil liberties in the digital information space is maintained where threats are described precisely but not in existential terms, where security measures are embedded in transparent procedures under the rule of law, where military forces are deployed strictly on a subsidiary and limited basis, where institutional structures are regularly scrutinised for mission creep and the expansion of control, and where freedom is not understood as a cost factor of security but as its actual resource.

For Austria within the EU in particular, this means that security in the digital information space must be organised as a democratic project. Military forces can help to ward off threats, but must not become permanent actors in the regulation of public communication. The risk of securitisation cannot be completely avoided, but it can be managed through language, institutions and a security culture that does not administer freedom, but enables it.

In doing so, Austria would not only defend its own security culture based on the rule of law, but also establish it as a benchmark in European debates. It could help ensure that the European security architecture in the digital information space does not veer towards the permanent dominance of security-driven approaches, but instead defines security as a prerequisite for democratic freedom.

5

Mikael Weissmann
and James Pamment:

Governance of the Information Environment. Concrete Measures for Cognitive Security and Cognitive Warfare

The Role of a Ministry
of Defence (MoD)



Governance of the Information Environment. Concrete Measures for Cognitive Security and Cognitive Warfare – The Role of a Ministry of Defence (MoD)

Docent Dr Mikael Weissmann is Senior Lecturer at the Department of Systems Science for Defence and Security at the Swedish Defence University, Associate Professor in War Studies, and Deputy Director of the Swedish Defence University Centre for Land Warfare. He is also an Affiliated Researcher at the Swedish Defence University Centre for Wargaming and co-convenes the Hybrid Threats Research Group, an interdisciplinary initiative focused on identifying, analysing, and responding to hybrid threats. His research focuses on security, strategy, intelligence, hybrid threats, and contemporary warfare.

James Pamment (PhD Stockholm University, 2011) is Director of the Lund University Psychological Defence Research Institute. He is editor emeritus of the journal *Place Branding and Public Diplomacy* and former Head of the Department of Strategic Communication. Previous affiliations include the Carnegie Endowment for International Peace, Swedish Defence University, the EU-NATO Hybrid Threats Centre of Excellence, University of Texas at Austin, University of Southern California, and Oxford University. Pamment researches the analysis and countering of hostile foreign interference, such as information influence operations and hybrid threats. As Director of the Psychological Defence Research Institute, he leads a team that provides research-based support to the Swedish Psychological Defence Agency and other governmental organisations. He has advised, trained, or produced research at the intersection of security policy and communication on behalf of UN and EU institutions, NATO, tech companies, and governments.

5.1 Introduction & Scope

This report examines how a Ministry of Defence in a small state can respond to and defend against **hybrid threats and cognitive warfare** so as to ensure sovereignty in the information space and contribute to **cognitive security**. More specifically, this report focuses on practical measures – *how a Ministry of Defence can organise, equip, and coordinate with partners to detect, withstand, and counter hybrid threats and cognitive warfare, thereby protecting sovereignty in the information space*.²³

The report takes as its point of departure the fact that different terms originating in the cognitive dimension are today increasingly being used to understand and address challenges of defence and military policy. This trend is evident both in the context of Austria and in the broader NATO context.²⁴ In the context of Austria, in April 2025, the SI – Directorate General for Defence Policy held a symposium on “Cognitive – Warfare, Security and Resilience” to initiate a whole-of-state and whole-of-society discourse.²⁵ In the NATO context, these discussions, while ongoing for some time, culminated in December 2025 with the publication of the NATO Chief Scientist’s 2025 Report on Cognitive Warfare, which was released to consolidate and advance the Alliance’s conceptual and operational understanding of the cognitive dimension.²⁶

With its focus on concrete measures for cognitive security, cognitive warfare, and hybrid threats, this report addresses three interrelated phenomena, all of which are located, at least in part, within the cognitive domain. In order to clarify this foundation, it is necessary to define the cognitive domain

23 During the research, analysis and writing process and in the generation of figures, a number of generative AI and AI-assisted technologies have been used for support, ChatGPT, Gemini and Grammarly. The authors carefully reviewed, edited, and validated all results and content to ensure their accuracy and integrity and take full responsibility for the final published work.

24 See, inter alia: “NATO Cognitive Warfare Concept” (18 June 2025), 1372084/2025, as well as the Austria-led MCDC Project “Countering Cognitive Warfare”.

25 See “Kognitiv – Kriegsführung, Sicherheit und Resilienz; Symposium 09.04.2025, 09:00-16:30; -Einladung”, S90500/17-GDVPol/2025.

26 NATO Science and Technology Organisation, Cognitive Warfare: NATO Chief Scientist Research Report (Brussels: STO - NATO Science & Technological Organisation, 2025), accessed March 10, 2026, <https://www.sto.nato.int/wp-content/uploads/chief-scientist-report-cognitive-warfare-final.pdf>.

and the cognitive realm more precisely. **Cognition** here includes “mental activities and relates to the issue of how humans in their mental constitution can relate (mental representation/cognitive semantics) to things or content (such as information) in the world (outside of themselves)”, which in turn “are essential building blocks of conceptions of truth, convictions, desires and will” and “constitute the basis of decision-making and acting (from potentiality to actuality).”²⁷ In this sense, cognition forms the precondition for meaningful action and political agency.

With the cognitive realm being “understood as the freedom-enabling area in which feelings, perception, understanding, knowledge, convictions and values, hence cognitive systems of individuals, individual groups, the entire population as well as institutions and the state, exist, and is the area of subjective and objective decision-making based on informed and justified reasoning.”²⁸ The cognitive realm thus constitutes a foundational sphere for democratic governance and sovereign statehood.

When analysing hybrid threats and cognitive warfare and cognitive security, they can be understood as being part of a continuum of competition, where cooperation, rivalry, confrontation, and armed conflict between states and non-state actors involve the exercise of power and the assertion of interests across the entire spectrum of available means and instruments.²⁹ (See figure 1.)

27 S90500/62-GDVPol/2025, Hybrid threats and cognitive realm: Definition of defence policy terms and their relationships (English, unpublished), 1. See, by way of comparison. Thomas Metzinger, ed., *Grundkurs Philosophie Des Geistes - Band 1: Phänomenales Bewusstsein* (Mentis Verlag, 2009), 18f. (cited in Ibid., 1.)

28 S90500/62-GDVPol/2025, Hybrid threats and cognitive realm: Definition of defence policy terms and their relationships (English, unpublished), 1. citing S90516/1-GDVPol/2022, *Hybride und zukünftige Konfliktaustragung in der kognitiven Domäne*, Positionspapier GDVPol.

29 S90500/62-GDVPol/2025, Hybrid threats and cognitive realm: Definition of defence policy terms and their relationships.

Continuum of competition

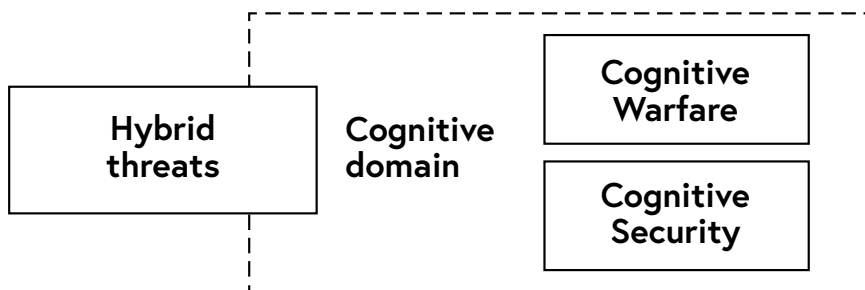


Figure 1: Continuum of competition

Source: S90500/62-GDVPol/2025, Hybrid threats and cognitive realm: Definition of defence policy terms and their relationships (English, unpublished).

Hybrid threats are here, following the definition of the Hybrid CoE in Helsinki, understood as an umbrella term for

“...harmful activities that are planned and carried out with malign intent. They aim to undermine a target, such as a state or an institution, through a variety of means, often combined. Such means include information manipulation, cyber-attacks, economic influence or coercion, covert political manoeuvring, coercive diplomacy, or threats of military force. Hybrid threats describe a wide array of harmful activities with different goals, ranging from influence operations and interference all the way to hybrid warfare.”³⁰

With hybrid threats, the focus moves from territorial possession or physical destruction towards asserting one’s own interests through defamation, control, or domination of the behaviour of the opponent.³¹

30 Hybrid CoE, “Hybrid Threat as a Concept,” accessed March 11, 2026, <https://www.hybridcoe.fi/hybrid-threats-as-a-phenomenon/>. Also see Mikael Weissmann, “Conceptualising and Countering Hybrid Threats and Hybrid Warfare: The Role of the Military in the Grey Zone,” in Weissmann et al., Hybrid Warfare.

31 S90500/2-DionSihPol/2021, Sicherheits- und verteidigungspolitische Herausforderungen durch hybride und zukünftige Formen der Konfliktaustragung cited in S90500/62-GDVPol/2025, Hybrid threats and cognitive realm: Definition of defence policy terms and their relationships.

Cognitive warfare, in turn, can be understood as synchronised and systematic activities designed to shape perception, judgement, and decision-making among individuals, groups, institutions, and wider populations by deliberately manipulating and influencing both conscious and subconscious cognitive processes. They target the informational aspect and the integrity of cognitive processes. It refers to “a strategy that focuses on altering how a target population thinks – and through that how it acts.” [Bold in original]³² It “focuses on attacking and degrading rationality, which can lead to exploitation of vulnerabilities and systemic weakening.”³³ It should be noted that cognitive warfare does not happen in a vacuum, but is “...conducted in synchronization with other instruments of power, can affect attitudes and behaviour by influencing, protecting, or disrupting individual and group cognition to gain advantage.”³⁴ It can also target subconscious processes through direct methods of intervention in physical/mental processes due to developments in fields like biotechnology and neurotechnology, for example, through substances, neurological implants, or through attacks on intersections such as brain-computer interfaces.

Cognitive security, the defensive counterpart, is a concept that sits at the intersection of national security, information operations, psychology, and digital technologies. It focuses on safeguarding the cognitive processes – beliefs, perceptions, decision-making – of individuals and groups from manipulation, deception, and undue influence, especially in the context of modern information environments.³⁵

-
- 32 Oliver Backes and Andrew Swab, “Cognitive Warfare: The Russian Threat to Election Integrity in the Baltic States” (Belfer Center for Science and International Affairs, Harvard Kennedy School, Cambridge, MA, 2019), https://www.belfercenter.org/sites/default/files/pantheon_files/2019-11/CognitiveWarfare.pdf, 8.
 - 33 Allied Command Transformation, “Cognitive Warfare,” no date, accessed April 25, 2026, <https://www.act.nato.int/activities/cognitive-warfare/>.
 - 34 Allied Command Transformation, “Cognitive Warfare: Strengthening and Defending the Mind,” published April 5, 2023. Accessed September 9, 2025, <https://www.act.nato.int/article/cognitive-warfare-strengthening-and-defending-the-mind/>.
 - 35 Bay S., Bertolin G., Biteniece N., Christie E., Dek A., Fredheim R., Gallacher J.D., Kononova K., Marchenko T., Responding to Cognitive Security Challenges (Riga: NATO Strategic Communications Centre of Excellence, 2019); Hilka Grahn and Toni Taipalus, “Defining Comprehensive Cognitive Security in the Digital Era: Literature Review and Concept Analysis,” *Journal of Information Warfare* 24, no. 2 (2025); Rand Waltzman, “The Weaponization of Information: The Need for Cognitive Security” (Rand, Santa Monica, Calif., 2017), https://www.rand.org/content/dam/rand/pubs/testimonies/CT400/CT473/RAND_CT473.pdf

Cognitive security can thus be understood as the capacity to withstand cognitive warfare and other forms of malign influence by protecting human cognition at both individual and collective levels. Being very similar to psychological defence,³⁶ both concepts capture measures aimed at strengthening the ability to resist psychological and cognitive warfare and thereby withstand manipulation and attempts at malign influence at the individual, societal, as well as institutional levels. In the Austrian context, *the Geistige Landesverteidigung (Intellectual national defence)*³⁷ represents a crucial instrument.

The importance of not only hybrid threats but also cognitive warfare in the military sphere has intensified in recent years, especially in light of experiences from Ukraine and beyond. Cognitive warfare has also gained growing attention within NATO, highlighting both its military dimensions and the critical role of the defence and security sector, including the armed forces. In fact, NATO increasingly frames cognitive warfare as a distinct domain of competition that targets human perception and decision-making rather than physical assets.³⁸ This underscores the need to understand the role of the Ministry of Defence and the armed forces in small non-NATO EU member states such as Austria.

Taken together, these conceptual clarifications provide the analytical foundation for the remainder of this report and frame the subsequent discussion of a Ministry of Defence's role in safeguarding sovereignty in the information space.

Against this backdrop, this report is structured around three key themes: the role of holistic governance in strengthening resilience against influence opera-

36 Jesper Falkheimer and James Pamment, eds., *Psychological Defence and Information Influence: A Textbook on Theory and Practice* (Psychological Defence Agency, 2025); James Pamment and Elsa Isaksson, *Psychological Defence: Concepts and Principles for the 2020s*, MPF report series 6/2024 (Karlstad: Psychological Defence Agency, 2024).

37 cf. Austrian Constitution, https://www.ris.bka.gv.at/Dokumente/ErV/ERV_1930_1/ERV_1930_1.html, (Accessed on 14 November 2025).

38 In NATO, the Allied Command Transformation manages the Cognitive Warfare Exploratory Concept, which is part of the larger Warfare Development Agenda. <https://www.act.nato.int/activities/cognitive-warfare/> (last accessed 11 March 2026). Also see report from the First NATO scientific meeting on Cognitive Warfare (France) in 21 June 2021: Bernard Claverie and François Du Cluzel, ““Cognitive Warfare”: The Advent of the Concept of “Cognitics” in the Field of Warfare,” in *Cognitive Warfare: The Future of Cognitive Dominance*, ed. Bernard Claverie et al. (NATO Collaboration Support Office, 2022).

tions and the contributions a Ministry of Defence can make; the identification of priority target groups for engagement; and the instruments, skills, and tasks needed to meet these. The report pays particular attention to hybrid threats, including cognitive warfare, and underscores the need for a comprehensive defensive toolkit.

This report is organised into three substantive parts. The first part situates the problem within the broader security environment by examining holistic governance, the changing character of hybrid threats, and the place of cognitive warfare within the peace-crisis-war continuum. It then develops the report's two core analytical models – the Hybridity Blizzard Model and the Hourglass Model – in order to clarify both the complexity of the threat environment and the specific role of a Ministry of Defence as an analytical, coordinating, and enabling actor within a whole-of-society defence system.

The second part turns from diagnosis to practice by outlining the capacities, instruments, and tasks available to a Ministry of Defence, and by introducing an ethical-legal decision framework centred on what a MoD could do, should do, and may legitimately carry out. The final part then identifies the principal target audiences for MoD engagement and considers how different stakeholder groups may need to be approached in different ways.

5.2 Delimitations

It is important to clarify that this report does not address the governance of the information environment in its entirety. The information environment is broader than the remit of any single ministry and encompasses political leadership, civilian authorities, regulatory bodies, the private sector, civil society, the media, educational institutions, and the general public. While holistic governance necessarily spans this full spectrum, the present report is deliberately limited to the responsibilities, authorities, and capabilities of the Ministry of Defence and entities under its command.

Accordingly, when referring to “governance of the information space” or to a “whole-of-society approach”, this report does so from the vantage point of the Ministry of Defence as one – albeit central – actor within a wider system.

A MoD cannot and should not seek to assume functions that properly belong to civilian authorities, civil society, or political leadership. Rather, its role is to contribute its specific capabilities – including intelligence, strategic analysis, military planning, and strategic communication – in support of a broader, constitutionally grounded framework of democratic governance.

This limitation does not diminish the importance of holistic governance; on the contrary, it underscores the necessity of clearly delineated mandates, robust inter-agency coordination, and respect for legal and democratic boundaries. The analysis that follows therefore focuses on what a Ministry of Defence can and should do within the information environment, while recognising that effective cognitive security ultimately depends on coordinated action across the whole-of-government and whole-of-society spectrum.

5.3 Holistic governance, hybrid threats and the role of a Ministry of Defence

How can holistic governance in the information space strengthen protection against influence operations, and what contribution can a Ministry of Defence make to this?

The international security environment has become increasingly contested and unpredictable. The international order is characterised by asymmetric conflicts and great power competition, with countries such as Russia, China, Iran and North Korea pursuing strategies to achieve national goals by developing and combining unconventional, less resource-intensive means for competing with the West on the global stage. Asymmetric strategies are also utilised by non-state actors. Centrally, these strategies increasingly include cognitive warfare – deliberate efforts to target perception, attention and trust of individuals and groups in order to shape their behaviour and erode societal cohesion.

Consequently, security challenges arising from hybrid threats that exploit and combine kinetic and non-kinetic tools – ranging from cyberattacks approaching or amounting to sabotage, sabotage of public and private property as well as of critical infrastructure, GPS jamming, and even assassination plots, to coercive diplomacy and economic dependencies – have ascended on security

agendas across the globe including in Europe. These salient concerns continuously present new challenges for liberal democracies such as Austria, their Ministries of Defence, Armed forces and intelligence agencies, for the broader security and defence sector, as well as for their societies as a whole. Cognitive warfare, in turn, operates in the cognitive domain. It is a human-domain layer of hybrid threats, amplifying the effects of other tools by undermining trust in institutions, corrupting information environments, and inducing maladaptive decisions among leaders and publics.

Whether hybrid threats, cognitive warfare, or other forms of influence operations, they do not respect borders and, in fact, seek to exploit borders, divisions, and rules. The role of Ministries of Defence has also been increasingly highlighted, as it has become apparent in the last years that the boundary between peace and war is increasingly blurred, with hybrid attacks occurring not only with growing frequency but also often crossing or skirting the traditional threshold of war. While attribution – at least publicly – may not always be possible, the number and severity of incidents have risen sharply. These include drone interceptions, airspace violations, cyberattacks approaching or amounting to sabotage, sabotage of public and private property, as well as critical infrastructure.

Part of the role of MoDs is closely linked to intelligence, as the intelligence community stands at the forefront of identifying and analysing hybrid threats and cognitive warfare in the current security environment. It is also an important part of building capacity for cognitive security, enabling the protection of human cognition across the population, at both individual and collective levels, to withstand cognitive warfare and other forms of malign influence. It has an important role in coordinating and integrating the information acquired among key societal actors affected by these threats or responsible for countering them. This includes detection and attribution of cognitive operations, timely warning of attempts to manipulate public sentiment or decision-making, and support for resilience and psychological defence through actionable assessments and tailored communication.

The prevalence of hybrid threats and cognitive warfare poses a set of challenges for the intelligence community, which relate directly to the nature of these threats. Hybrid threats are complex because they can target a wide range of societal vulnerabilities across the PMESII spectrum (political, military,

economic, social, information, and infrastructure) using methods that may not be susceptible to early detection. The same is the case with cognitive warfare, though it specifically targets the cognitive dimension seeking to exploit social fault lines and biases.

5.4 The bigger picture – Holistic governance and whole-of-society defence as a prerequisite for success

“Whole-of-society” defence should form the governance backbone for a smaller state. Countering hybrid and cognitive threats demands continuous, interactive loops across – and within – a wide array of **actors in the public, private, and civil spheres**.

This perspective points to the need for strong, practical coordination mechanisms and underscores the Ministry of Defence’s role as an engine and a crucial force in whole-of-society defence, as part of holistic governance, especially as attacks increasingly approach or cross the war threshold and the use of kinetic or near-kinetic means such as drones becomes more frequent. In short, the MoD’s convening role in holistic governance rests on its unique capabilities and command of actors with such capabilities (e.g. armed forces, military intelligence).

In this report, we argue that two core process models may work as enablers:

Firstly, the **Hybridity Blizzard Model** explains why coordination must span both short-term warning and response and long-term resilience-building.³⁹ It is well suited to capture the dynamics across the full conflict spectrum, including peace, crisis, war, and the grey-zone framing, and it clarifies the MoD's role as an engine within the system.

Second, the **Hourglass Model** positions the MoD to:

- aggregate identification and analysis of threats and attacks, as well as own and adversaries' vulnerabilities,
- aggregate analysis and tailor dissemination to relevant actors within the public, private as well as civil spheres, and within the security and defence sector, including the MoD itself, the armed forces and intelligence and security services, and
- ensure effective reception and absorption among key actors across organisations and spheres.⁴⁰

The model highlights bottlenecks and the need for feedback-rich, whole-of-society practice. The hourglass also provides a way to understand how intelligence must be aggregated and tailored to secure absorption among consumers, including non-traditional ones.

The model highlights bottlenecks and the need for feedback-rich, whole-of-society practice. The hourglass also provides a way to understand how intelligence must be aggregated and tailored to secure absorption among consumers, including non-traditional ones.

39 Mikael Weissmann, "Framtida Hotbilders Påverkan För Säkerhetstjänsterna," in *Framtidens Säkerhetstjänst I Totalförsvaret*, Swedish Defence University Report Series (2024), <https://fhs.diva-portal.org/smash/get/diva2:1846156/FULLTEXT01.pdf>; Niklas Nilsson, Mikael Weissmann, and Björn Palmertz, "Hybrid Threats and the Intelligence Community: Priming for a Volatile Age," *International Journal of Intelligence and CounterIntelligence*, 2026, <https://doi.org/10.1080/08850607.2024.2435265>. Also, see Mikael Weissmann, Niklas Nilsson, and Björn Palmertz, "Moving Out of the Blizzard," in Weissmann et al., *Hybrid Warfare*.

40 Nilsson, Weissmann and Palmertz, "Hybrid Threats and the Intelligence Community: Priming for a Volatile Age."

Before turning to these models, we briefly establish the broader context by illustrating the complexity of the security environment, which spans **peace, crisis, war and the grey-zone**.

5.5 The complexity of the security environment: peace, crisis, war and the grey-zone

A useful way to understand current and future hybrid threat and cognitive warfare scenarios today and in the future is to take the idea of the grey zone as an analytical starting point. The idea of the grey zone can best be understood as

“[...] a collective term for antagonistic threats and risks in the borderland between war and peace. These are characterised by variation and an unclear context, resulting in considerable uncertainty. The individual threats have both civilian and military dimensions.” [authors’ own translation]⁴¹

We would argue that the concept remains essential, even as hybrid attacks increasingly approach – or cross – the threshold of war and the use of kinetic or near-kinetic means (e.g., drones) becomes more frequent. At the same time, the traditionally distinct crisis stage before war has become increasingly blurred, and “peace” is increasingly permeated by persistent hybrid activity and cognitive operations.

Accordingly, the grey-zone should be reconceptualised not merely as a space between war and peace, but as a persistent operating condition in which attacks, interference, and influence are continuously present.

The two models below illustrate this shift. First, the traditional view depicts the grey-zone as the area between war and peace, where a broad repertoire of strategic measures and instruments – ranging from diplomatic and economic means, to unconventional and military means – is employed to deter, coerce, or influence (see Figure 2). These means, in turn, can be expected to target

41 Col Dr Anders Palmgren, 2017-01-25 cited in Weissmann, “Framtida hotbilders påverkan för säkerhetstjänsterna,” 41.

vulnerabilities across the PMESII spectrum (political, military, economic, social, information, and infrastructure). Second, the reconceptualised view presents the grey-zone as an enduring condition across the peace-crisis-war continuum, requiring whole-of-society awareness, continuous sensing, and coordinated responses (see Figure 3).

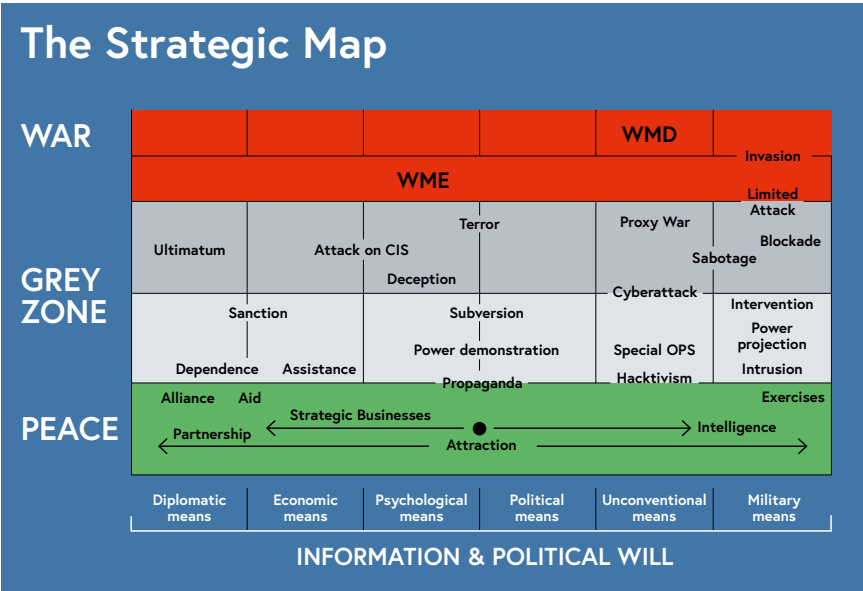


Figure 2: The Strategic Map (Traditional version)

Source: Developed from Col Dr Anders Palmgren (Swedish Armed Forces), based on Sweden's Military Strategic Doctrine (MSD 16) (Swedish Armed Forces 2016): 35. The author has added "War", "Grey Zone" and "Peace" from the Swedish Original.

In the post-2022 reality shaped by Russia's full-scale invasion of Ukraine, the grey-zone has expanded to encompass almost everything short of all-out conventional war. Actions that would have been regarded as acts of war before 2022 are now increasingly blurred into the grey-zone. In this environment, "latent" or "manifest" conflict often provides the most accurate description of relations between states that are not outright allies, while the traditional "crisis" phase has narrowed considerably. The boundaries between manifest conflict, crisis, and acts of war have become indistinct. Even situations that appear latent may still involve sudden spikes of high-intensity attacks that are close to or transcend the acts of war threshold. Moreover, such activities now occur simultaneously across multiple domains and by multiple means, as will be illustrated in the next section with the Hybridity Blizzard Model.

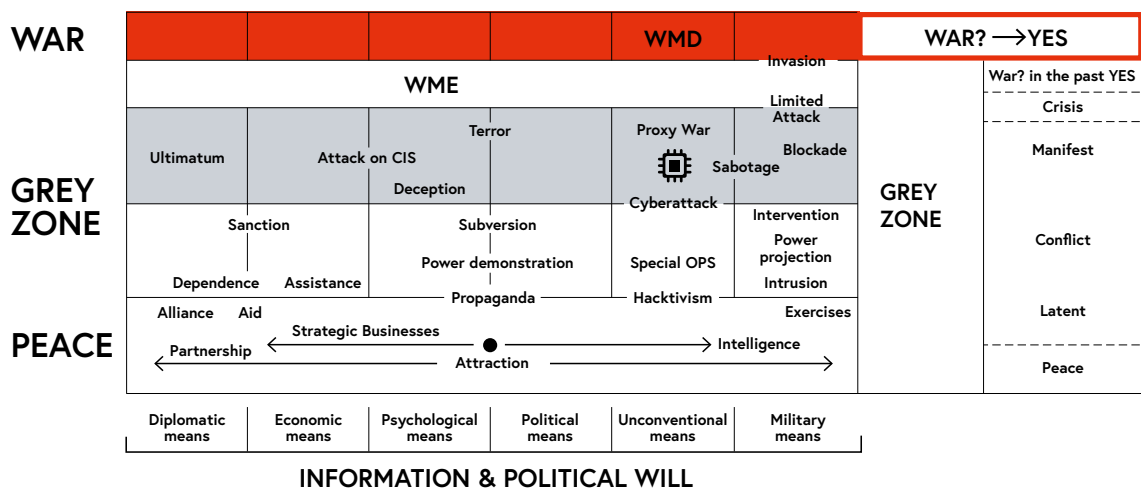


Figure 3: The Strategic Map Re-conceptualised.

Source: Authors' model, developed from Figure 2.

5.6 The complexity of hybrid threats and cognitive warfare

To capture the dynamics of hybrid threats and cognitive warfare, it is essential to consider the aggressor–defender relationship across both the short-term and long-term perspectives. We argue for illustrating this relationship in the form of the “Hybridity Blizzard Model”, which schematically outlines how the time and actor dimensions interact and co-evolve. Both dimensions matter because hybrid threats are neither one-off events nor activities that can be cleanly separated from their context (Figure 4).⁴² The model distinguishes between us as the defenders and our adversaries as the aggressor. It furthermore distinguishes between two temporal dimensions, “short-term” and “long-term”. In the short term, the relationship consists of a battle with ongoing, reciprocal interactions between the aggressor’s hybrid threats and cognitive warfare and the defender’s responses and countermeasures. This is a continuous process with no predetermined beginning or end.

42 Weissmann, “Framtida hotbilders påverkan för säkerhetstjänsterna”; Nilsson, Weissmann and Palmertz, “Hybrid Threats and the Intelligence Community: Priming for a Volatile Age”. Also, see Weissmann, Nilsson and Palmertz, “Moving out of the blizzard.”

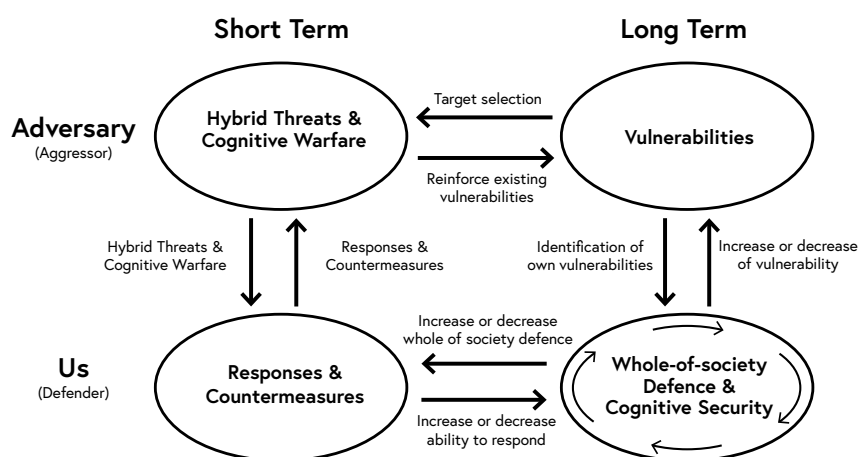


Figure 4: Hybridity Blizzard Model. Developed from Weissmann, Nilsson, and Palmertz (2021); Nilsson, Weissmann, and Palmertz (2026).

In the long-term, the focus shifts to our (the defenders') vulnerabilities and to the resilience and psychological defence built to mitigate them to ensure cognitive security and an efficient whole-of-defence system. The aggressor seeks to identify and exploit vulnerabilities. The defender's systematic identification of its own vulnerabilities enables the building of an efficient whole-of-defence system including ensuring cognitive security. These efforts change the vulnerability landscape – typically reducing exposure, but potentially also creating new vulnerabilities that the aggressor can in turn identify and exploit.

There is also a bidirectional link between the short-term and long-term perspectives. When the aggressor designs hybrid methods, identified vulnerabilities are central to target selection; by striking at weaknesses, existing vulnerabilities may deepen and become focal points for subsequent attacks. On the defender's side, short-term responses and countermeasures are conditioned by the existing whole-of-society defence system and cognitive security. In turn, those responses feed back into the effectiveness and resilience of the overarching social defence system (strengthening or, if misapplied, eroding it). In short, both sides are engaged in a continuous back-and-forth between short-term action and long-term adaptation. This interplay spans the PMESII spectrum and reinforces the need for whole-of-society coordination.

Hybrid threats and cognitive warfare tend to thrive in conditions of ambiguity, overload, and denial. The image of a blizzard is useful here: the target is

buffeted from multiple directions by countless attacks of all severities that cannot be neatly separated, or localised, impeding timely and effective defensive action (see Figure 5). In principle, many more arrows could be added to the model. On the short-term side – where the snowstorm analogy is most vivid – the environment is especially chaotic:

- threats cannot always be identified – they remain ambiguous, unattributed, or even indeterminate to the defender; and
- there is a non-trivial risk of false positives.

This ambiguity is problematic in its own right – akin to “crying wolf,” where repeated false alarms risk eroding trust in warnings and response mechanisms. This risk is particularly acute in a Ministry of Defence context, where the disclosure and communication of intelligence are closely tied to strategic decision-making and to the trust placed in the MoD, the armed forces, and ultimately the government itself. Yet it can also be deliberately engineered as part of a broader strategy. For instance, are you observing a proxy actor, or is the “proxy” operating independently, outside any larger design? Are anomalies in the power grid or the financial infrastructure manifestations of hybrid action, or are they simply glitches? Is a perceived threat to military medical services or food supply a deliberate hybrid operation, or merely a routine hazard with no hostile actor behind it?

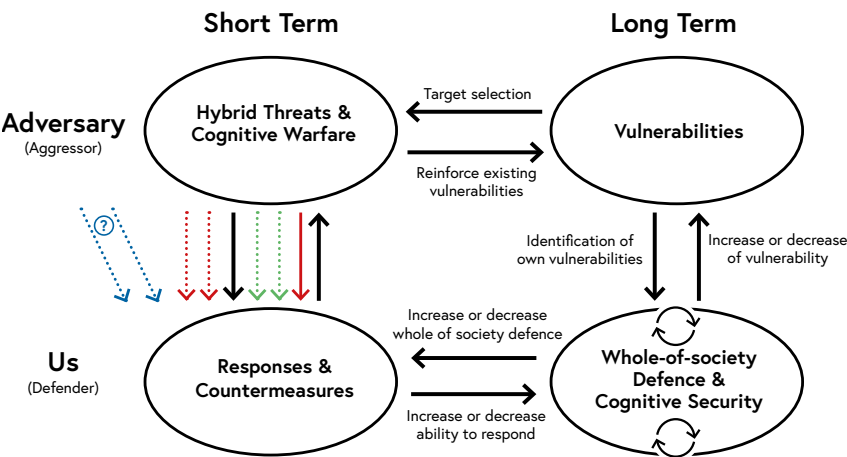


Figure 5: Hybrid threats and cognitive warfare. Developed from Weissmann, Nilsson, and Palmertz (2021); Nilsson, Weissmann, and Palmertz (2026).

The coloured and dotted arrows in the bottom left part of the figure represent attacks that are, from right to left, identified (hybrid threats and cognitive warfare that are detected), **ambiguous** (hybrid threats and cognitive warfare that are not detected), **unattributed** (anonymous or proxy), and **undetermined/unknown** (it is not possible to ascertain if hybrid measures).

5.7 MoD as an engine in the whole-of-society defence system (The Hourglass Model)

Building societal resilience in the face of broadly defined hybrid threats and cognitive warfare requires an equally broad response in the form of a whole-of-society approach to defence and cognitive security. There are several layers to this concept that become especially salient in the discussion of hybrid threats. The main responsibility for countering and preventing the societal impact of antagonistic threats lies with governments and competent government agencies, not least the MoD and entities under its command. Depending on the type of threat concerned, relevant agencies can consist of the armed forces, police authorities, and emergency management agencies, etc. However, given the composite and synergistic nature of these threats, government responses require a purposeful fusion of a wide variety of functions – including regional and local government, cultural and educational institutions, healthcare, energy and transportation, financial institutions, regulators, research, journalism – many of which are not usually considered part of the security apparatus.

The aim is to integrate state responses through a whole-of-government approach supported by horizontal arrangements that avoid stovepipes in information-sharing, decision-making and capability utilisation. From a MoD perspective, the convening role is to enable productive partnerships and preparedness across all sectors and actors, including at the federal government level as well as regional and local levels, and in the public, private as well as the civil sectors. It is also important to not forget the general public. In addition, a MoD needs the same collaboration and preparedness within its own structure and in relation to those under its command.

The **Hourglass** outlined below **illustrates** the three critical processes in the interaction between **the MoD** and key actors as well as within the MoD and its commanded entities. The successful execution of these processes is **crucial to realising the MoD's potential** in forging a whole-of-society approach to countering hybrid threats and cognitive warfare.

The *Hourglass Model* positions a MoD to

- aggregate the identification and analysis of threats and attacks, as well as own and adversaries' vulnerabilities,
- to aggregate analysis and tailor dissemination to
 - key actors across the public, private and civil spheres – including within the MoD the security and defence sector – as well as
 - relevant parts of the MoD and entities under its command (e.g. sections of the armed forces, military intelligence and military security services), and
- ensure effective reception and absorption among key actors across organisations and spheres, including within and its subordinate commands.

The model highlights bottlenecks and the need for feedback-rich, whole-of-society practice (see Figure 6).⁴³

43 The hourglass model presented here is a development of the "Intelligence Analysis Interaction Model" outlined in Nilsson, Weissmann and Palmertz, "Hybrid Threats and the Intelligence Community: Priming for a Volatile Age."

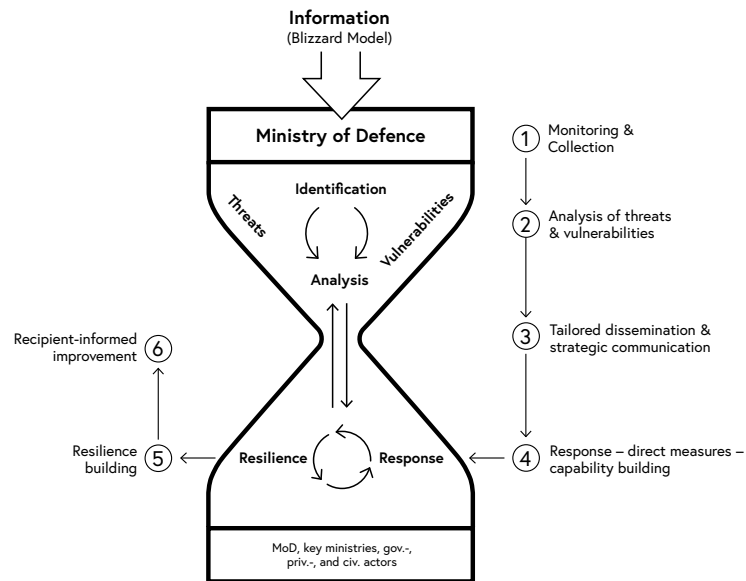


Figure 6: The MoD hourglass model. Author's model.
Drawing from Nilsson, Weissmann, and Palmertz (2026).

The model also provides a way to understand how intelligence **must be aggregated and tailored to secure absorption** among non-traditional consumers, which is an area where the MoD and military intelligence should be expected to play an increasing role as hybrid threats and cognitive warfare intensify, including in the information sphere, and the role of military means and actors expands. Europe is moving closer to open conflict. Furthermore, as the war in Ukraine has shown, there is a general shift in intelligence towards greater disclosure and strategic communication, rather than keeping everything secret. In parallel, OSINT (Open Source Intelligence) has expanded, with more non-traditional actors contributing to intelligence and a wider whole-of-society recipient base; central government and the military cannot do it alone.

In this report, the Hourglass Model is used in two complementary ways to clarify the role of a Ministry of Defence: the MoD Internal Hourglass and the MoD System Hourglass (see Figure 7).

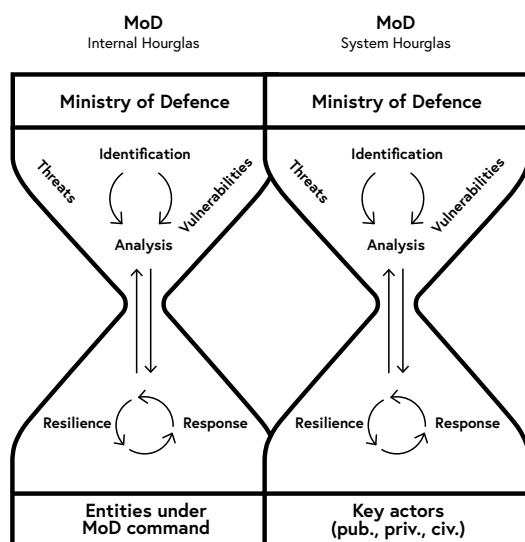


Figure 7: The MoD Internal and System Hourglass. Author's model.
Drawing from Nilsson, Weissmann, and Palmertz (2026).

The former focuses on how the MoD organises its own processes within the ministry and entities under its command. One could in principle expand the internal model to include the Ministry of Interior and/or the federal government in the hourglass. If so, the logic and processes would still be the same. However, for the purpose of this report, focusing on the MoD itself is most relevant. In contrast, the latter targets how the MoD connects its defence intelligence and capabilities to other ministries and government actors, including public agencies at the national, regional and local levels and key actors in the private and civil sectors.

The two models can be thought of as working in parallel with the MoD as the common denominator.

In both versions, the hourglass clarifies the role of the Ministry of Defence through five core tasks/roles. The difference is who the recipient/audience/consumer is.

1. *Identification and analysis* of threats and vulnerabilities,
2. *Aggregation of findings and tailoring of dissemination and strategic communication* to relevant consumers (including non-traditional consumers and possibly the general public),

3. *Actors' response*, ranging from direct counter measures to capability development for the future),
4. *Resilience building*, focusing on learning and building long-term capabilities and structures to enhance both own and system-wide (whole-of-society) defence.
5. Finally, *recipient-informed improvement*, by ensuring feedback channels to the MoD to ensure enhanced learning from previous communications and the needs of the recipients, to continuously calibrate the system for the future.

In both models, information flows in at the top; the MoD monitors and collects information to analyse it in the upper bowl.

1. Analysis of threats and vulnerabilities, where the MoD identifies and analyses antagonistic threats and national vulnerabilities. This process is represented in the upper part of the Hourglass, depicting how the vast flow of information could be narrowed down in order to identify threats and vulnerabilities in the haystack of information. These are then analysed and transformed from information into actionable intelligence and recommendations to be tailored for dissemination and strategically communicated within and outside the MoD structure, as well as to the public.
2. The tailored dissemination and strategic communication denotes a bottleneck in the interaction between the MoD and the key audiences within the ministry, the security and defence sector as well as across public, private and civil actors. This part of the model is not only a bottleneck but also a necessary conduit to optimise policy and counter-measures in response to hybrid threats and cognitive warfare that are informed by the best possible knowledge.

Two components stand out as critical with regard to optimal communication of intelligence: aggregated analysis and tailored dissemination. In the same way that aggregated data is needed in order to produce good analyses, aggregated analysis is required to produce the best possible assessment of a question or an attack, threat or vulnerability in focus. Tailored dissemination is required in order to meet the particular needs and requirements of the specific target audience. Since responses to hybrid threats and cognitive warfare can potentially involve a wide range of actors, many of which are not included in the conventional audience and consumer base of the MoD,

dissemination must be calibrated to sectoral context and maturity, including inexperienced consumers.

3. The response and resilience (reception and absorption?) building among key audiences denotes the implementation process whereby new insights and needs broadly feed into a whole-of-society approach to devise responses to identified and anticipated hybrid threats and cognitive warfare. This is represented by capability development in the lower part of the Hourglass.

Ideally, the three processes of analysis – dissemination – reception should be interactive rather than a one-way street. On the defence side, the MoD requires feedback from actors on their needs and requirements as well as incident reporting in order to enhance its ability to identify threats and vulnerabilities and to refine monitoring and analyses, and development of capabilities and instruments offered. Feedback is also an important component of tailoring dissemination. Yet this reciprocal process rarely functions optimally: time constraints, institutional stovepipes and cultural discrepancies frequently present barriers to systematic feedback to the MoD, while channels into the often guarded MoD environment are limited. This is particularly so in the case of intelligence and security services.

Linking back to the Hybridity Blizzard Model and the challenges facing the MoD in the contemporary operational environment, the “targeting and vulnerability reinforcement” dimension is catered for by the upper part of the Hourglass, where the MoD identifies and analyses threats and vulnerabilities. These analyses then need to be channelled to relevant actors to facilitate responses and resilience, providing the necessary input for successful capability development among key actors inside and outside the MoD, as shown by the lower part of the Hourglass. This dissemination and strategic communication layer between the analysis level and the receivers (audiences) is a critical bottleneck in the decision and response cycle. Here, sophisticated and timely aggregated analysis delivered through tailored dissemination and strategic communication to relevant actors is crucial.

The success of the Hourglass model depends on the optimal execution of the three critical processes: MoD analysis of threats and vulnerabilities, aggregated analysis and tailored dissemination and strategic communication, and the

reception and absorption among key actors to ensure successful response and resilience building. To counter hybrid threats, including information operations and cognitive warfare, building societal resilience requires a whole-of-society approach, where the responsibility is not limited to either the government or the MoD, but also includes numerous organisations across the public and private sphere nationally and internationally. Within government, the MoD plays a pivotal role in identifying vulnerabilities and establishing situational awareness, but the nature of the threats and responses requires a reassessment of what role a MoD can and should have, including what information and intelligence and military capabilities and instruments that can be shared, how and for what purpose.

5.8 Role of the MoD

Three core roles can be identified for the Austrian MoD (see Table 1):

- Firstly, **BMLV acting alone** (including entities under its command). This category of action is self-explanatory.
- Secondly, **BMLV with others**. Here, the role of the MoD is established within a national system consisting of military and civilian stakeholders, each with their own areas of responsibility. Increasingly, thinking about whole-of-society security positions each stakeholder as a centre of excellence for its own domain. In theory, each stakeholder can advise (not command) others within their area of expertise; for example, a transport authority might need advice from the national cyber centre, or municipal authorities from the media regulator. Here, the MoD has a unique role and mandate to strengthen the stakeholder network through its enhanced knowledge of the physical security domain.
- Thirdly, **BMLV working with international partners** (e.g. EU, NATO centers). Likewise, international actors have created stakeholder networks centred on sharing information and experience. EU institutions, complemented by centres of excellence and ad hoc coalitions, facilitate international sharing of best practices, research, training, as well as opportunities for collective analysis, attribution, and deterrence.

Alone	With others	With international partners
BMLV acting alone (mandated roles) ⁴⁴	Government – National/Regional Local – Public/private/civil spheres	State partners, International organisations and partners, EU

Table 2: Roles of the MoD

5.9 MoD's contribution to whole-of-society defence and cognitive security

The analysis above clarifies how the Ministry of Defence (MoD) and entities under its command can anchor holistic governance against hybrid threats including – especially cognitive warfare and influence operations – by operationalising the Hourglass logic both within the MoD (MoD Internal Hourglass) and across society (MoD System Hourglass). In both versions, information flows in at the top, is narrowed through analysis, and then expanded into tailored actions and capability development, with feedback closing the loop.

The MoD's contribution to whole-of-society defence and cognitive security is concentrated in five core roles:

1. **Identify & analyse:** Detection and analysis of threats and of own/adversary vulnerabilities to generate actionable insight.
2. **Aggregate & tailor (dissemination/StratCom):** Fusion of findings across domains and calibrated dissemination to the right consumers – including non-traditional stakeholders and, when appropriate, the general public – recognising this layer as a critical bottleneck.
3. **Enable responses:** Facilitation of responses and countermeasures among public, private, and civil actors, while directing immediate MoD actions where required.

⁴⁴ MoD itself, Austrian Armed Forces (Bundesheer), Foreign military intelligence (Heeresnachrichtenamt (HNA), Military security & counterintelligence (Abwehramt)

4. **Build resilience:** Help institutionalise lessons, strengthen structures, and develop capabilities for durable cognitive security within the MoD and across the system.
5. **Close the loop (recipient-informed improvement):** Ensure systematic feedback from recipients to refine monitoring, analysis, dissemination, capabilities, and instruments over time. This area should not be underestimated, as all too often there is a lack of feedback on information, intelligence, and recommendations. There is also a frequent tendency to neglect after-action review and feedback when more immediate challenges demand attention. Furthermore, functional channels for feedback are often lacking, making it difficult to ensure that insights from recipients are systematically captured and integrated.

Executed together, these roles allow the MoD to act not only as a key contributor but also as an engine of whole-of-society defence – integrating government, private and civil sectors as well as the public into a feedback-rich practice that enables and enhances the ability to respond and overall cognitive security. This is essential as Europe faces intensifying hybrid activity, including in the information and cognitive domains, where timely, aggregated analysis and tailored communication – received, absorbed, and acted upon – determine whether resilience is merely asserted or actually achieved.

5.10 MoD Instruments: Capacity, Instruments and Tasks

What instruments are available to a Ministry of Defence, what capabilities and tasks are required?

When looking more concretely at what the MoD in a small European state such as Austria can contribute in terms of concrete actions against hybrid threats such as cognitive warfare and information operations, there is a broad range of options. We are, at this point, value-neutral, and not focused on any specific state, or on the mandate and capabilities of the Austrian MoD; instead we will list a wide array of available instruments.

In this study, we do not assess whether they are ethical, whether they should be done, or whether they overlap. We present a range of possibilities – some available today, others that could be developed in the future – without drawing conclusions about whether they should be carried out, their legality or compliance with regulations, or whether Austria should or can pursue them.

Hence, below you will find an outline of examples of what the MoD could contribute today or in the future, followed by a list in the form of capabilities (what the MoD needs to make instruments work), instruments (what the MoD can wield to deliver effects), and tasks (what the MoD can credibly do and contribute).

Thereafter, the possible MoD capabilities, instruments, and tasks will be discussed from the perspective of what is available, what should (and could) be developed, and the ‘carry out’ aspect, including ethical judgements, assessments, and decisions that need to be made.

5.11 What a Ministry of Defence Can Contribute to Holistic Governance and Security in the Information Sphere

A number of possible examples/areas of MoD contribution to Holistic Governance and Security in the Information Sphere can be identified:

- National Information-Environment Picture (NIEP)
- All-Source Fusion & Prioritised Threat & Vulnerability Assessments (TVAs)
- Crisis Narrative Cell/National Security or Defence Communications (strategic communication)
- Public Advisory System
- Platform & Media Liaison (limited to issues related to national security)
- Sector Hardening Programs
- Rapid Support Cadres
- Exercises & Drills (“hybridity blizzards”)
- Education, Training & Awareness
- Standards, Tooling & Procurement Enablement
- International Alignment & Lift
- Recipient-Informed Improvement (closed loop)

- Community & Local-Leader Networks
- Operational Research & Experimentation
- Outcome-Centred Metrics

5.12 Ethical-Legal Decision Framework ("Could/Should/Carry Out")

The development and use of any tool – including those outlined above – can be assessed in an Ethical-Legal Decision Framework ("Could/Should/Carry Out"). The framework above outlines MoD:

- Capabilities,
- Instruments and
- Tasks.

These now need to be separated into:

- available,
- should (and) could, and
- the "carry out" aspect,

including the ethical judgements, assessments, and decisions that need to be made. One thing is what we *could do*, and another is what we *should do*. For a liberal Western democracy like Austria, the decisive test is not only whether we can act but also whether actions are ethically and legally defensible, proportionate, and consistent with our values; this is what separates us from authoritarian adversaries such as Russia or Iran. Legitimacy – not just effectiveness – is a strategic advantage. If we adopt the same tools and tactics as our adversaries, we risk winning the contest while forfeiting what we seek to protect. The real question, therefore, is arguably not *how far we can go*, but *how far we should go*.

This question applies whether a measure is known to the public or undertaken in secret. While not every action needs to be public, every action must be ethically and legally defensible; secrecy may alter how measures are carried out; secrecy may alter if a tool is used or measure carried out, but the question *how far we should go* remains as relevant.

More concretely, we here propose a triad of questions to formalise the framework into tools (see Table 3):

1. **COULD**: could the tool be developed? Here there is a need for a **legality and feasibility** check; is it feasible for a small state, and can it be done within existing legal frameworks and mandates? If not, is **that** **by** design **and** purpose, or can it **be** changed?
2. **SHOULD**: should the tool be developed? Here a layer of ethics, values, reputational risk, **and** second-order effects **needs** to be added, and questions **asked and pondered**. This applies both to capabilities that **the** MoD already **has** and **to** new capabilities it **considers** or should consider acquiring.
3. **CARRY OUT**: shift focus **to** the actual **authorisation** and implementation phase of the tool. Here the authorisation chain becomes important, including questions **about** transparency levels (public/attributed/unattributed) **as well as** safeguards **and** logging (including oversight).

More concretely, the ‘carry out’ phase should always include asking the following three questions before developing any tool, wielding an instrument, or fulfilling a task:

- Are we allowed to do this?
- Should we do this?
- If we should and can, do we need special permission or a change e.g., in laws, funding, or staff?

A recurring concern that should be mentioned in relation to building cognitive security is the risk of ‘securitisation of society’. There is a risk that an expanded security framing will gradually normalise exceptional measures, blur the boundary between defence and domestic governance, and shift everyday democratic life towards suspicion and control. If measures intended to protect citizens are perceived as surveillance, propaganda, or political policing, they can erode trust in public institutions, chill legitimate speech, and deepen polarisation – thereby creating the very vulnerabilities that cognitive operations seek to exploit. In the cognitive domain, legitimacy is not a constraint external to security; it is a core security asset that adversaries actively target.

To manage these risks, measures should be designed around clear boundaries and safeguards, including tests of necessity and proportionality and – not least – the separation between capability development and employment

(‘could/should/carry out’). Independent oversight is important, including parliamentary, judicial, and data-protection scrutiny, and transparency and auditability need to be ensured where possible. It is also important to take data minimisation and retention rules into account, alongside purpose limitation, access controls, and documented authorisation chains. A preference for enabling measures (preparedness, coordination, trusted communication, education and inoculation) should normally be maintained over coercive interventions, even where deterrence requires credible offensive capabilities. The latter should be developed, held, and governed in ways that prevent routine domestic application and ensure strict authorisation, logging, and review. In practice, the credibility of cognitive security policy depends not only on effectiveness against adversaries, but on sustained public confidence that the state is protecting democratic space rather than shrinking it.

Triad for formalising the proposed framework into a tool

Could?	legality & feasibility checks
Should?	ethics, values, reputational risk, second-order effects
Carry Out?	authorisation chain, transparency level (public/attributed/unattributed), safeguards & logging (including oversight), other organisational changes

Table 3: Triad for proposed formalising the proposed framework into a tool

Having set out the ethical–legal challenge, the next step is to outline a practical tool/ structured framework for assessing the role of the Ministry of Defence to support decision-making. This model organises

- capabilities,
- instruments, and
- tasks into three stages:

what is available, what the MoD should (and could) employ, and what it should actually carry out.

5.13 Capacity-instruments-task model for MoD

Against authoritarian adversaries who recognise few limits, liberal democracies must manage defence contributions according to different standards. A Ministry of Defence must know not only what it can do, but also what it should do – and what it must not do if it is to preserve the values it seeks to defend. To make this distinction operational, we present a framework that organises MoD

- capabilities,
- instruments, and
- tasks into three stages:

what is available, what should (and could) be developed and employed, and what should – or should not – be carried out (see Table 4).

This model is deliberately neutral with regard to specific countries or ethical judgements; it does not prescribe particular choices. It is not the role of researchers to decide what a Ministry of Defence should or should not do. Those decisions belong to the MoD itself, acting within its mandate, national law, and democratic oversight. However, as outlined above, the questions must be asked: what is desirable, what is appropriate, and what is defensible? The framework is intended as a structured way of asking those questions, and as an enabler of more transparent and legitimate decision-making.

Ultimately, effective defence against hybrid threats, cognitive warfare, and information operations depends not just on the tools a state has at its disposal, but on how it chooses to use them. Whether or not a measure is adopted in the end is the responsibility of the Ministry of Defence. From our perspective, the critical issue is that the questions are posed, debated, and answered within a clear ethical and legal framework – not what any single answer may be.

	Capability	Instrument	Task
Available (current practice)	Strategic communication about public policy. Ability to warn public about threats.	Free or low-cost analysis tools, official accounts, annual report about national security threats	Analysing and informing public opinion about security issues
Should & could?	“Shaping” the debate: Develop a strategic analytical capability, track main debates, & proactive engagement with main voices.	Corporate-level analysis tools, work (including payments) with influencers & intermediaries. More visible profile for some figures within MOD.	Analyse social media discourse and sentiment and intervene with facts (sometimes through intermediaries) where appropriate.

	Capability	Instrument	Task
Carry Out (Ethical considerations)	Continuous scanning of the information environment to understand debate could be seen as eavesdropping & undue influence over public policy.	Are we allowed to do this? Should we do this? If we should and can, but it is against regulations and laws, should we do it anyway?	What new permissions, resources are required? E.g. exceptions from GDPR. Who should we partner with?

Table 4: Weighing up the potential for the MoD to take a more proactive approach to public debates about the hybrid threat

Such a tool could be used to weigh up many issues about the future direction of the MoD in relation to more proactively protecting the country from hybrid threats. For example, topics to weigh up could include:

- Could/should the MOD (perhaps in partnership with other institutions) ...
- Develop a national information campaign (with before/after measurement) about the will to defend/spirit of resistance. Perhaps with a focus on minority groups in Austria who traditionally do not engage with the MoD.
 - National effort to explain the wider societal role of the MoD (“More than War”) in schools and particularly as a career path.
 - Convene key nongovernmental societal actors with closed-door briefings about hybrid threats relevant to their sectors.
 - Develop a series of position papers outlining the MoD’s view on different forms of hybrid threats (as a more specialised complement to annual security risks reporting). Could be an indirect focus (i.e. “trends in Europe that Austrians should be aware of”).
 - Conduct a society-wide information campaign about certain themes under the hybrid umbrella (e.g. a focus on risks to critical infrastructure from sabotage, or on how certain behaviours fuel transnational organised crime).
 - Approach universities about security threats to the research sector from hostile states (e.g., dual use, IP exfiltration, undue influence, espionage).
 - Train NGOs, journalists, and researchers in open source analysis centred on hybrid, explaining the MoD’s role in this work more clearly and offering points of contact.
 - Run simulations and exercises to test response capabilities to hybrid threats. Use simple tools such as MoD-branded card decks (hybrid threat attack and response gaming) to support engagement.

5.14 Target audience

Which target audiences should the MOD address with its measures?

So far, we have developed the idea of a MoD acting as a centre of excellence within national and international systems of stakeholders that encompass the whole-of-society. Note that from this perspective, the idea of “audience” and “stakeholder” become interchangeable; audiences must be perceived as active players, valued partners, and genuine stakeholders in the work of the MoD. The natural question follows, which audiences should the MoD prioritise in this work?

Simultaneously, the MoD does not and should not have a comprehensive interaction with all audiences on all issues. Such partnerships can equally be characterised by what they are not. The MoD does not need to tell journalists how to report, the private sector how to sell things, or politicians how to govern. Within its areas of expertise, however, the MoD is an authority with unique perspectives to share.

At the most abstract level, these audiences can be grouped into types and issues, based on priorities. We suggest that an engagement plan be developed on the basis of an analysis of organisational priorities, stakeholder networks (MoD-internal, national, and external), needs as expressed by stakeholders, and gaps and vulnerabilities identified by the MoD throughout the network. For example:

- Analysis based on internal surveys may suggest that internal MoD audiences (soldiers, commanders, civilians, reservists) differ in how they understand and explain the MoD’s mission. An organisational priority may therefore become a branding campaign centred on promoting a sense of belonging based on a single identity for all internal audiences.
- Informal feedback from civil society groups may suggest that the MoD is perceived as distant and uncommunicative. Based on MoD priorities, individual staff could be tasked with building better relationships with key partners based on mutual advantage (“we should talk!”).

Defining the exact target audiences depends somewhat on the MoD's objectives, national culture, and long-term patterns of interaction. However, as a starting point, the following groups of audiences could be considered:

- Internal MoD audiences (extended members of the MoD community)
- National governmental audiences
(political level, departments, agencies, local government, etc)
- National civil society audiences (public, health, youth, education, media, NGOs, CI operators, platforms)
- National private sector audiences
(industry, tech, security, critical infrastructure)
- International private sector audiences
(e.g. tech suppliers such as social media platforms)
- International specialist audiences
(allies/partners, diaspora, neutral/third-country)
- International adversarial audiences
(government, business, civil society, civilians in hostile countries)

Segmenting these broader groups into smaller targets reveals significant differences in modus. For example, within national audiences are both the political level and local government level, which obviously have to be handled very differently. We recommend developing an overarching strategy across the three questions covered in this chapter. This involves deciding on where the MoD can contribute within a holistic system, ensuring that the contribution is framed as a capability to be weighed up both ethically and pragmatically, and developing tailored approaches to different audiences/stakeholder groups.

6

André K. Rodarte
and Lena Frischlich:

Preservation and Promotion of Democratic Resilience

A Scientific Basis for
Policy Recommendations



Preservation and Promotion of Democratic Resilience – A Scientific Basis for Policy Recommendations

Lena Frischlich is an Associate Professor at and Vice-Director of the Digital Democracy Centre at the University of Southern Denmark. She has a background in Media Psychology and Communication Science. Her research examines how digitalisation creates new opportunity structures for manipulation in digital environments and how to strengthen democratic resilience.

André K. Rodarte is a postdoctoral researcher at the Digital Democracy Centre, University of Southern Denmark. His research explores the role of digital media in representative politics, with a specific focus on new forms of propaganda, populism, and threats to democratic institutions. He holds a Ph.D. from the University of Texas at Austin and an M.Phil. from the University of Cambridge.

Acknowledgements

We would like to thank Carolin Iber for her contributions to the project's analyses as a student assistant. We are also grateful for Miriam Riis Diemer's valuable feedback and for the Digital Democracy Centre's unwavering support.

6.1 Executive Summary and Key Takeaways

The PRORESIL project sought to

1. Identify groups susceptible to misinformation and understand why. To do so, it examined four analytical levels of abstraction: The micro-level of the individual media user, the meso-level 1 of social groups and collectives to which these users belong, the meso-level 2 of (digital) media environments in which both individuals and collectives interact and inform themselves, and the macro-level of institutions and states. These four levels constitute the Complexity Framework, which theoretically oriented this project.

2. It further sought to map interventions aimed at increasing resilience in the academic literature, comparing these interventions with how state actors responded to misinformation in recent crises across the world.
3. Finally, it examined public preferences concerning strategic communications and countermeasures to misinformation. Ultimately, the pursuit of these objectives facilitates fostering democratic resilience against misinformation, particularly during crises.

To reach these objectives, the project combined three studies:

- A scoping review of the English-language academic literature on susceptibility to misinformation (N = 663 abstracts);
- a content and textual analysis of academic articles that suggested interventions against misinformation (N = 130 articles), combined with a narrative case review of state actor interventions against misinformation in recent international crises;
- a survey of the Austrian adult population (N = 1,985 participants) with an embedded conjoint experiment testing the public preferences for different regulatory and communicative measures.

Overall, the project answered the following seven research questions (RQ's):

- **RQ1:** Which levels of the Complexity Framework have gained the most attention and to what extent is the complexity of the phenomenon reflected in the academic literature?
- **RQ2:** Which topics of misinformation susceptibility have been studied across levels?
- **RQ3:** Which levels of the complex framework have received the most attention when scholars discuss interventions to curb the spread of misinformation?
- **RQ4:** How do state responses to misinformation during recent international conflicts compare to the interventions emphasised in the scholarly literature?
- **RQ5:** What are Austrians' preferences when it comes to official communications and responses to misinformation?
- **RQ6:** Do crises, particularly hybrid warfare, shift citizens' preferences?
- **RQ7:** How are citizens' preferences related to their susceptibility towards misinformation?

6.2 Core Findings

Study 1 documented a massive growth of research on misinformation since 2016, with most of the literature focusing on the US, the UK and China. Answering RQ1, we showed that the literature is heavily concentrated on the micro-level of individual communicators and recipients, with very few studies examining the interaction between different levels of the Complexity Framework through which misinformation is perpetuated. This narrow emphasis on individual susceptibility overlooks the complex cross-level dynamics that shape misinformation, such as international conflicts increasing the “supply” of mis- and disinformation beyond directly affected countries, or generative AI facilitating its production and dissemination. With regard to the topics studied (RQ2), the literature provides mounting evidence that limited cognitive elaboration and motivational factors – such as the confirmation of one’s world-view– increase the susceptibility to misinformation. This cognitive lens is also reflected in a strong emphasis on research examining misinformation through Fact checking. In contrast, studies that account for macro-level factors remain the exception, with few studies considering structural factors or conducting cross-country comparative research.

Answering RQ3, Study 2 extended these findings by showing that the academic literature primarily discusses interventions at the micro-level. In contrast, comparatively few studies address macro-level interventions. In line with the findings of Study 1, the literature underscores the value of cognition-oriented measures to increase individuals’ resilience, such as by confronting media users with small doses of misleading content to activate cognitive “defence systems”. Critically, poorly implemented countermeasures (e.g., broad “Don’t trust anything online” warnings) can produce unintended consequences, such as promoting distrust in all kinds of online information, including reliable news and thereby increasing susceptibility over time. Nevertheless, such risks are rarely considered in the literature. Study 2 also highlights a lack of discussion about who should be responsible for implementing these interventions.

We connect these academic lessons to case studies of state responses during crises, particularly military conflicts. Answering RQ4, our analysis showed that state actors often adopted far more restrictive and systemic approaches to both official communication and the handling of misinformation than those

typically examined by scholars. While this mismatch is partially rooted in ethical challenges, for example when studying restrictive state-level approaches, it underscores the need for a closer exchange between academic knowledge and practical policy-making, as well as a stronger focus on providing evidence that accounts for potential risks of countermeasures.

Study 3 built on the first two studies and developed a Compass of the Information Environment to guide the study of official communications and responses to misinformation. It did so by identifying two dimensions along which communication activities and interventions can vary, namely: (A) proactive versus reactive countermeasures to misinformation and (B) permissive versus restrictive information environments, which concerns citizens' inclusion in communication activities. We used this compass to examine the Austrian population's preferences with regard to official communication activities and the handling of misinformation. Study 3 further explored how widespread these preferences are among this population, and whether they change during crises.

Answering RQ5, the analyses showed that Austrians overall preferred permissive information environments characterised by transparent official communication, the inclusion of citizens' perspectives as valuable information sources, and a strong role for independent news media as the main channel of official communication. Study 3 further showed that Austrians preferred proactive information environments in which misinformation is addressed both through official communication and platform regulation, although preferences diverged regarding which specific regulatory measures would be most desirable. At the same time, participants clearly favoured information environments in which veracity judgements are grounded in expert advice and evidence.

In response to RQ6, the results show that crises meaningfully shift citizens' preferences regarding the information environment. During crises, direct communication by state actors becomes more acceptable. In military crises, this extends to direct communication by military spokespersons as an important addendum to a news-centred information environment. The results also indicate a continued preference for veracity judgements grounded in expert advice, while physical warfare additionally generates support for legislation as a complementary basis for such decisions. These findings reveal that citizens expect state actors to play a more active role during military crises, but in cooperation with scientific and media institutions rather than as their replacement.

Finally, we examined how these preferences differed across the population by identifying groups that were more or less susceptible to misinformation and unpacking how these groups envisioned their preferred information environments. The analyses revealed three distinct participant clusters. The “Institutional Mainstream” cluster was characterised by high trust in and satisfaction with democracy, strong psychological resilience, an ideologically moderate stance, and a more traditional news diet centred on mainstream media. The “Networked Apart” cluster combined moderate democratic trust with lower resilience and a greater reliance on social media and AI chatbots for news, paired with comparatively weaker use of traditional media sources. By contrast, the “Right-wing Authoritarian” cluster emerged as the most ideologically right-leaning group, marked by stronger populist, authoritarian, and conspiratorial attitudes, the lowest trust in and satisfaction with democracy, the weakest consumption of mainstream news, and the strongest reliance on alternative outlets. These three groups also differed significantly in their ability to correctly classify misleading headlines as false, with the Institutional Mainstream performing best, the Networked Apart occupying an intermediate position, and the Right-wing Authoritarian cluster performing substantially worse.

These groups also differed with regard to their preferences for official communication activities and the handling of misinformation. The “Institutional Mainstream” cluster most closely aligned with the broader sample’s preference for active, transparent, and expert-led information environments. The “Networked Apart” cluster shared support for active intervention but showed greater tolerance for lower transparency and weaker commitments to institutional openness. The “Right-wing Authoritarian” cluster was most distinctive in its weaker preference for expert-based truth verification and its greater willingness to accept state-aligned communicators alongside independent news media. This reflects their distrust in epistemic elites – such as scientists and fact-checkers – leading to a preference for truth defined either by state authorities or media users themselves. Yet, across all three groups, one commonality remained: a clear preference for some form of active intervention against misinformation, even where views diverged on transparency, epistemic authority, or the specific tools of governance.

6.3 Implications for Strategic Communication

The PRORESIL project bears important implications for understanding susceptibility to misinformation and for planning strategic communication activities, including within the Ministry of Defence. We summarise the ten main implications in the following.

1. First, understanding susceptibility to misinformation requires a systems perspective that moves beyond an exclusive focus on individual cognition. Across the project, we found that both academic research and intervention design remain heavily concentrated on the micro-level, even though misinformation dynamics unfold through the interaction of individuals, social groups, media environments, and institutional contexts. Our findings show that susceptibility is shaped by individual characteristics, media use patterns, political and ideological views, and trust in democratic institutions and other actors.
2. Second, strengthening resilience against misinformation requires building a trustworthy information environment rather than relying on official communication alone. To prevent potential spirals of distrust and deception (Frischlich & Humprecht, 2021), it is necessary to address the broader systemic roots of distrust. Our findings show that misinformation susceptibility is closely connected to conspiracy mentality, Right-wing Authoritarian attitudes, and patterned forms of digital media consumption. By contrast, resilience depends on an information environment in which official communication is transparent, leaves little room for conspiratorial interpretation, and treats citizens' perspectives as a valuable source of input alongside other institutional actors. In this sense, trust is not produced by messaging alone, but by the perceived openness, inclusiveness, and legitimacy of the wider communication environment.
3. This project underscores the central role of a healthy news media ecosystem. Misinformation constitutes only one, comparatively small, part of a broader information environment, the larger share of which is shaped by citizens' access to credible, reliable, and trustworthy information provided by independent media (see also Altay, 2026). Amid declining economic revenues for journalism and the ongoing transformation driven by generative AI, strengthening healthy information environments

requires supporting free and independent media. The European Democracy Shield (European Commission, 2025a) rightly acknowledges the central role of independent journalism alongside the need to counter disinformation.

4. The role of direct state communication should be calibrated to context rather than treated as a constant strategy. For most participants, continuous direct communication by the Chancellor and/or the military was not desirable under normal circumstances. However, in times of crisis, state communications become more acceptable and can play a role in stabilising the information environment. This shift is relevant because members of the Right-wing Authoritarian cluster were generally more favourable toward state official communication, suggesting that transparent and context-sensitive communication by authorities can serve as one important element within a broader whole-of-society communication approach targeted at vulnerable audiences.
5. AI chatbots should not be treated as substitutes for official communication channels. Our findings suggest that relying on AI chatbots to take over official communication tasks is not aligned with the preferences of the Austrian population. Participants in Study 3 strongly disliked information environments in which official communication was delivered primarily through summaries generated by AI chatbots. This aversion was especially pronounced among members of the Networked Apart cluster, despite this group's comparatively high usage of AI chatbots and digital media for news. Ultimately, even digitally connected audiences do not necessarily regard AI-mediated summarisation as a legitimate replacement for transparent communication by trusted institutional and media actors.
6. Citizens expect misinformation to be actively addressed rather than left to citizens alone. Survey participants consistently preferred information environments in which official communication and platforms take responsibility for addressing harmful misinformation over scenarios in which the task of debunking falsehoods is left entirely to citizens. This stands in sharp contrast to the recent shift by major platforms toward Community Notes and the reduced use of professional fact-checkers, as illustrated by Meta's 2025 policy change (META, 2025). Crucially, public support is related to harmful misinformation and does not include all forms of mostly harmless misinformation that can be found online. Such

nuanced judgements are unlikely to be made by automated moderation systems alone, instead requiring contextual knowledge.

7. While citizens do not support a laissez-faire approach to misinformation, there is still no clear public consensus on the most legitimate intervention tools. Participants consistently rejected information environments in which misinformation was left unchallenged. There were no strong preferences as to whether official communication should focus on prebunking, debunking, or other regulatory interventions – suggesting that combining multiple intervention tools is a viable strategy but also that interventions should be adapted with regard to whether they are directed towards more or less harmful forms of misinformation.
8. Crisis communication strategies must differentiate between hybrid and physical warfare scenarios rather than treating all security crises alike. While the Overton-window for official communication and misinformation governance shifted across the imagined crisis scenarios, citizens clearly did not equate hybrid and physical warfare or display the same tolerance for restrictive measures in both contexts. In hybrid crises, their preference for transparent communication and evidence-based misinformation responses grounded in expert advice was especially pronounced. The implication is that hybrid threats require more openness, institutional pluralism, and expert-led truth verification than communication strategies designed for physical warfare, where temporary restrictions may be more publicly tolerated.
9. Strategic communication against misinformation should be audience-segmented rather than based on a one-size-fits-all model. Our findings show that susceptibility emerges through distinct audience profiles shaped by specific combinations of democratic trust, resilience, ideological orientation, and media habits. Trust in democratic institutions served as a protective factor, while conspiratorial, populist, and authoritarian attitudes, as well as low satisfaction with democracy, stood out as risk factors. Furthermore, psychological strain, social detachment, and digital dependency were found to also increase susceptibility, though to a lesser extent than these more ideological variables. Thus, interventions should be tailored to these distinct susceptibility pathways rather than assuming a single target audience.
10. The PRORESIL report underscores the need for greater democratic scrutiny of how societies respond to misinformation. While much existing research focuses on identifying susceptibility and individual-level

interventions, far less attention has been paid to how citizens evaluate the legitimacy of countermeasures and regulatory responses. These perspectives matter: citizens hold nuanced, context-sensitive preferences regarding transparency, expert involvement, crisis communication, and the proportionality of interventions. Democratically legitimate misinformation governance must not only target citizens. It must incorporate their preferences into the design and evaluation of countermeasures.

6.4 Introduction

According to the World Economic Forum (2025), misinformation – that is, content that is intentionally or unintentionally at odds with the best available evidence (Freiling et al., 2023) – is the most significant immediate risk worldwide. On the one hand, these concerns are grounded in research showing that misinformation contributes to misperceptions (Li & Yang, 2024), leading to ill-informed decisions in crucial areas such as health (Van Der Linden et al., 2025) and politics (Ecker et al., 2024). This is particularly problematic for democratic societies, which rely on the ideal of an informed electorate, and during crises where misperceptions (e.g., how to behave safely) can have severe consequences for both individuals and society. On the other hand, concerns about misinformation stem from the risk that countermeasures could endanger civic rights such as freedom of speech (Jungherr, 2024). Furthermore, measures intended to counter the spread of misinformation can have severe side effects if not planned carefully. For example, broad tips on how to “spot misinformation” can fuel scepticism towards all types of information, including democratic institutions (Hoes et al., 2024), thereby potentially increasing people’s susceptibility over time (Frischlich & Humprecht, 2021).

To understand the susceptibility to misinformation and develop adequate measures to foster democratic resilience that avoid such side effects, it is crucial to account for the complex system through which misinformation propagates, reaches its audience, and can lead to misperceptions. This complex system can be conceptualised along four interconnected levels (Frischlich et al., 2025): the micro-level of individual communicators, their messages, and their audience; the meso-level I of social groups and communities, the meso-level II of digital platforms with their specific characteristics, and the macro-level of

institutions, states and supranational organisations. The current report builds upon this framework to study susceptibility to misinformation and examine measures to foster democratic resilience.

6.5 Definition of Core Constructs

Misinformation and disinformation. For the purposes of this report, we understand misinformation as content that is intentionally or unintentionally at odds with the best available evidence (Freiling et al., 2023). That is, misinformation refers to statements that can be proven and rejected based on objective, measurable indicators. This definition does not include opinions, preferences, or convictions. The personal preferences of a political opponent might be perceived as being more or less incompatible with one's worldview or morally "wrong," but there are no objective indicators which can prove that one's preferences are "superior" (Shoemaker et al., 2004).

Misinformation can refer to different aspects of a message (Quandt et al., 2019): The core of a message (i.e., a specific claim or image that is false or faked), the meta-information (e.g., misleading headlines or ascribing a false source to a claim), as well as the larger context in which the information is embedded (e.g., through biased selection of evidence or the biased framing of events in untrustworthy media outlets). Misinformation can further take on various modalities, such as verbal, written, visual, sound or videos. Misinformation can also be generated with the support of artificial intelligence (AI), though of course not all AI-generated content is automatically misinformation.

Misinformation is part of a larger information disorder (Wardle & Derakhshan, 2017) which includes both unintentionally misleading content (i.e., misinformation in a narrower sense) and strategically deceptive content (so-called disinformation). Misinformation can also emerge through sloppy news coverage – for example, through omitting crucial aspects of a story (Quandt et al., 2019). Dis- and misinformation can intersect and are not always easy to separate clearly: For example, a strategic agent can disseminate a false story (e.g., disinformation about bedbugs in Paris) which is then shared by others without strategic aims. We thus use the intention-agnostic umbrella term "misinformation" in the following.

Misperceptions. Misperceptions are beliefs that are at odds with the best available evidence, that is, they can be falsified by evidence (Nyhan, 2020). This distinguishes them from other types of beliefs and assumptions that cannot be tested or proven (such as political opinions). A systematic synthesis of several studies (so-called meta-analysis) has shown that misinformation can lead to misperceptions (Li & Yang, 2024); however, the average effect is small and not everyone responds to misinformation with misperceptions at all. While misperceptions are often studied focusing on individuals, misperceptions can also be shared in social groups and institutions. For example, it took a long time until Galileo Galilei's defence of the heliocentric model was established as the standard model of the universe. The receptivity of individuals, groups, and institutions to developing misperceptions as a consequence of being exposed to misinformation can be described as susceptibility.

Susceptibility. The susceptibility to misinformation can be rooted in various factors. Here we follow the Differential Susceptibility to Media Effects Model (Valkenburg & Peter, 2013), which suggests that the effects of media messages (such as the effect of misinformation on misperceptions) can be enhanced or reduced by individual differences (e.g., in age or attitudes) and by social or contextual factors. These variables shape whether someone turns to a specific message, pays attention to it, and how he or she responds to it. The cognitive, emotional and behavioural responses then influence which longer-lasting effects a message has, for example, whether someone forms misperceptions. Over time, these longer-lasting media effects then affect further consumption patterns. For example, consuming conspiracy theories increases distrust (Pummerer et al., 2022) and distrust motivates the consumption of alternative counter-news (Frischlich et al., 2023) which further fuel distrust (Andersen et al., 2023) – potentially leading to diametric “spirals of distrust and deception” and reduced levels of democratic resilience (Frischlich & Humprecht, 2021).

Democratic resilience. The term resilience is rooted in the Latin term *resiliere*, that is, to ‘jump back’. In material science, it is used to describe a material's ability to keep its shape after a shock. In psychology, resilience often refers to humans' ability to cope with challenging life circumstances or misfortunes (e.g., <https://www.merriam-webster.com/dictionary/resilience>). More recent conceptualisations describe resilience as a psychological process that includes both manifest resilience, that is, the ability to handle challenges, and generative resilience, that is, protective factors in the environment supporting this

adaptation (e.g., Miller-Graff, 2022). Research in media psychology argued that resilience can be understood as a long-term process that includes the development of protective and promotive resilience factors as well as episodic processes of resilience (Reinecke & Rieger, 2021). In political science, democratic resilience adapts the idea of ‘resistance’ to shocks to the macro-level of democracies and conceptualises the “jumping back” as the “ability of a political regime to prevent or react to challenges without losing its democratic character” (Merkel & Lührmann, 2021, p. 872). For democracies, that means preserving the “commitments to democratic norms and values” (Merkel & Lührmann, 2021) and the process of handling threats to the system while upholding this commitment and preserving these values (Frischlich et al., in press).

6.6 The Complex Framework for Studying Misinformation Dynamics

Both susceptibility and resilience unfold not only on the level of the individual media user, but can also be studied with regard to social groups, digital technologies, democratic institutions, states or supranational organisations such as the European Union or NATO. These levels interact with each other and changes on one level can lead to changes on other levels. For example, political conflicts between states (e.g., Russians full-scale invasion of Ukraine, 2022), regime shifts (such as currently happening in the US), or disruptive technologies (such as the global race between large language models and other forms of artificial intelligence) affect actors and misinformation dynamics on all other levels.

To account for this complexity, some of us have recently suggested a multilevel framework for visualising the complex system through which misinformation propagates, is consumed, and evaluated. This structure is best understood as a complex multilevel network of different prototypical actors and their relations to each other (Frischlich et al., 2025). The Complexity Framework thereby distinguishes four central levels of abstraction: The micro-level of individuals, the meso-level I of social relations and social groups, the meso-level II of digital platforms, and the macro-level of societal institutions, states, and supranational organisations.

On the individual micro-level, misinformation is ultimately a communication process. Different actors, who have been extensively studied with respect to their unique psychological characteristics (Sultan et al., 2024), create, disseminate, and consume misinformation. On this level of abstraction, it is the communicative interplay between a specific communicator (or sender), the content that is communicated, and the recipient of this communication that shapes misinformation dissemination and effects. For example, people with higher levels of cognitive reflection or a larger motivation to form accurate perceptions are less likely to perceive misinformation as being accurate or share misinformation (Pennycook & Rand, 2021). Prototypical actors on this level include the creators and spreaders of misinformation (who can be the same, but do not have to), the recipients as well as those passively observing misinformation and those actively countering it.

At the meso-level I, all these individual actors are part of larger collectives that interact with each other and respond to misinformation based on their collective identities (Ziemer & Rothmund, 2024). Collective actors on this level include closely knit primary groups (such as families or peers) but also looser secondary groups (such as workplaces) and larger categories (such as gender, race, or religion). Research inspired by Social Identity Theory (Tajfel & Turner, 1986) and Self-Categorisation Theory (Turner et al., 1987) has shown that the more people identify with their social groups, the more they feel, think and act based on this specific social identity. For example, people are more likely to believe misinformation that paints their “in-group” (e.g., their political party) in a positive light or casts a negative view on the “out-group” (e.g., the political enemy, see Faragó et al., 2020).

At the meso-level II, digital platforms’ socio-technical configurations shape how these communicative dynamics unfold (e.g., Burton et al., 2024). For example, low moderated “alternative” or “fringe” social platforms host more unwarranted conspiracy theories than social media platforms with more intense moderation (Schatto-Eckrodt et al., 2024).

Finally, all prior levels are embedded in the macro-level of societies, where media systems (Humprecht et al., 2020) and political regimes (Hunter, 2024) affect misinformation flows. For example, when governments turn towards illiberalism, they often first attack the media and scientific institutions, thereby

shifting the “supply” of evidence-based information citizens have access to, thereby affecting trust and information search patterns among the population.

Critically, research and action related to misinformation, though abundant, often focuses only on specific aspects of the Complex Framework. For example, our analysis showed that research that uses the military metaphors of “information warfare” often focuses on the macro and meso-II level, for example, studying attacks on communication systems (e.g., Munro & Mackay, 2019) and focusing on the strategic value of deception as the centre of all warfare (Sun Tzū, ca. 544–496 BC, “The Art of War”). While valuable for understanding inter-state conflicts, this approach underestimates the pivotal role of citizens and non-political motivations for the spread of misinformation. It also fails to account for feedback loops such as individual actions influencing social norms and political decision-makers or the risk of fostering distrust in the population through strategic deception (Frischlich et al., 2025).

The rise in international conflicts over the last decade has reframed misinformation as both an instrument of warfare and a threat to national security. Myanmar offered an early illustration of how misinformation can be instrumentalised to fuel ethnic violence, while similar threats to national security were observed across the globe (Surjatmodjo et al., 2024). More recently, Russian state actors have leveraged hyper-partisan media, automated accounts, and coordinated online networks to obscure military objectives, sow confusion, and induce information paralysis (Mandić & Klarić, 2023).

6.7 The PRORESIL Project

The PRORESIL project draws on the Complexity Framework to

- (1) identify groups susceptible to misinformation and understand why.
- It further (2) sought to map interventions aimed at increasing resilience in the academic literature, comparing these interventions with how state actors responded to misinformation in recent crises around the world.
- Finally, (3) it examined public preferences concerning strategic communication and countermeasures to misinformation. Ultimately, the pursuit

of these objectives facilitates democratic resilience against misinformation, particularly during crises.

We sought to achieve these objectives through three separate studies: a systematic review of the scholarly literature summarising the current state of evidence regarding factors that modulate susceptibility to misinformation (scoping review, Study 1); a detailed review of interventions discussed, developed, and evaluated in the scholarly literature (manual content analysis, Study 2a); a qualitative review of communicative interventions across different contexts (case studies, Study 2b); and a survey of the Austrian adult population (N = 1,985 participants) with an embedded experiment testing the population's preferences for different regulatory and communicative countermeasures (conjoint experiment, Study 3).

6.8 Study 1 – Understanding Susceptibility to Misinformation

The spread of misinformation and people's belief in misinformation are complex phenomena that can be analysed across multiple levels of factors influencing susceptibility. To provide a comprehensive overview of the current state of knowledge about these factors, we conducted a systematic review of the scholarly literature. We created a dataset of 663 academic papers from a range of disciplines, from research on hybrid warfare to health science, psychology, and political communication. Our analysis is oriented by two research questions:

- **RQ1:** Which levels of the Complexity Framework have gained the most attention and to what extent is the complexity of the phenomenon reflected in the academic literature?
- **RQ2:** Which topics of misinformation susceptibility have been studied across levels?

We answered these questions by conducting a scoping review of academic studies on misinformation. Scoping reviews are systematic methods to compile and summarise the findings established within a body of literature. They are especially well suited to synthesise evidence from different fields or different

types of studies (Munn et al., 2018). For this study, we followed the PRISMA Extension for Scoping Reviews (Tricco et al., 2018) typically used to synthesise medical evidence. We combined this systematic review with a quantitative and computational content analysis, which enabled us to describe central topics examined across these levels and highlight gaps in the literature. Together, these analyses helped us to reach the first objective of our study: Understanding factors that increase susceptibility across the different levels of the Complexity Framework.

6.8.1 Database and Screening Procedure

We created a theoretically informed search term for this task that included four different blocks of keywords: the first block included keywords related to misinforming content (e.g., misinformation, disinformation, fake news, information warfare etc.), the second block included keywords referring to digital channels (e.g., social media, specific platforms such as YouTube or TikTok, artificial intelligence), the third included keywords conveying information about susceptibility (e.g., susceptibility, risk factor), and the fourth block contained terms related to misinformation effects (e.g., misperceptions). The different blocks were entered such that only papers that included keywords from all four blocks were included for further analyses (details in Table 1A in the Appendix).

Figure 8 outlines this data collection three-part process. In the first step (Identification), search terms were used to identify potentially relevant academic contributions on the two largest international interdisciplinary academic databases, the Web of Science and Scopus, as well as two discipline-specific databases that address the micro, macro, and the meso-levels (PsycINFO, Political Science Complete). We employed the systematic review tool COVidence for abstract screening and the automatic exclusion of duplicate texts.

In the next step (Screening), all identified records were manually screened via COVidence. We filtered out duplicates, retracted papers, and articles that did not meet our a priori inclusion and exclusion criteria (see Table 2A in the Appendix). Among them, we excluded papers not in English or German, whose manuscripts were not available, or when they referred to misinformation simply as a buzzword.

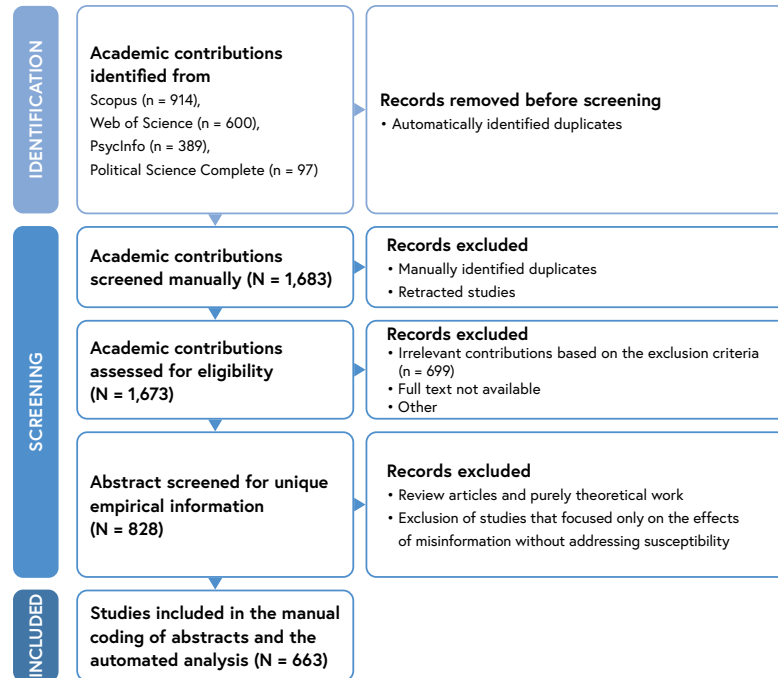


Figure 8. Screening process of the scoping review as a flow diagram. N = number of.

This procedure yielded a corpus of 828 manuscripts, whose abstracts were read to assess their empirical value (that is, whether they provided new evidence on misinformation susceptibility). Abstracts which did not refer to any empirical research, introduced scoping reviews, or did not help to explain the factors behind the spread or belief in misinformation (e.g., papers studying misinformation as a risk factor for vaccine hesitancy) were excluded. This resulted in a total of N = 663 abstracts for the final analyses.

6.8.2 Scholarly Research on the Susceptibility to Misinformation (RQ1)

Methodological approach

To answer our first research question and describe which levels of the Complexity Framework were studied in the literature, we conducted a manual quantitative content analysis (Lacy et al., 2015; Neuendorf, 2001). Based on this analysis, we assigned each scholarly contribution in our database to the levels of the framework, with articles that covered more than one level being included in all relevant categories. The codebook for this assignment was developed based on an initial coding of 30 abstracts by two members of the research team. We established the reliability of the coding based on 50 random abstracts, as

reported in Table 5, before one of the authors coded all remaining abstracts in the database. Using our codebook allowed us to successfully categorise 97% of abstracts in the corpus along the levels of the Complexity Framework.

Code	Description
Country	Country(ies) studied in the paper.
Platform(s) or media examined	Digital Platform(s) examined. E.g., Facebook. If not about digital platforms code “offline” or include media type (e.g., TV, Radio).
Producer characteristics	If the paper concerns how the producer’s characteristics facilitate the spread of / belief in misinformation, write 1. Includes demographic descriptors or network characteristics such as being labelled an opinion leader.
Recipient characteristics	If the paper concerns how the recipient’s characteristics facilitate the spread of / belief in misinformation, write 1. E.g., Demographic descriptors, psychological traits, etc.
Content factors	If the paper concerns how the characteristics of the message facilitate the spread of / belief in misinformation, write 1. E.g., this includes comparisons between video and textual messages.
Group factors	If the paper concerns how social group factors facilitate the spread of / belief in misinformation, write 1. E.g., this includes partisanship and out-versus-in group variables.
Organisational factors	If the paper concerns how organisations facilitate the spread of / belief in misinformation, write 1. E.g., this includes businesses, NGOs, parties, media organisations.
Platform factors	If the paper concerns how platforms’ design or affordances facilitate the spread of / belief in misinformation, write 1. E.g., this includes comparisons between allowing anonymity or the engagement reactions available.
Cross-country or media systems comparisons	If the paper concerns how the characteristics of a country or media system facilitate the spread of / belief in misinformation, write 1. E.g., this includes comparisons between legislations, countries, and media systems more broadly (liberal, authoritarian, etc.).
Suggested interventions	If the study suggests or tests a solution to misinformation, write 1.

Table 5: Codebook for manual content analysis.

Code	Percent Agreement	Brennan-Prediger'-coefficient
Producer Characteristics	0.90	0.81
Recipient Characteristics	0.87	0.73
Content Factors	0.87	0.73
Group Factors	0.87	0.73
Organisational Factors	0.98	0.96
Platform Factors	0.88	0.77
Cross-country or media systems comparisons	0.98	0.96
Suggested interventions	0.81	0.62

Table 6. Inter-coder Reliability Scores.

Note. Percentage agreement expresses how often both coders assigned the same academic contribution to the same category. Brennan-Prediger's coefficient is particularly well suited when some codes are rare. In both cases, values of 0 indicate no agreement and 1 indicates perfect agreement.

6.8.3 Results

The academic literature predominantly studies the US, China, and the UK.

To understand the development of misinformation research, we examined the number of academic contributions per year and their geographical information. Our results show that most misinformation research has focused on the United States, followed by China and the United Kingdom, in line with prior studies (Blair et al., 2024). Large parts of Europe, South America and Africa remain underrepresented, as shown in Figure 9. It is important to note that 313 of the abstracts did not specify the country of study, which limits this analysis and may lead to an underestimation of research conducted in North America.

Countries Most Frequently Mentioned in the Abstracts

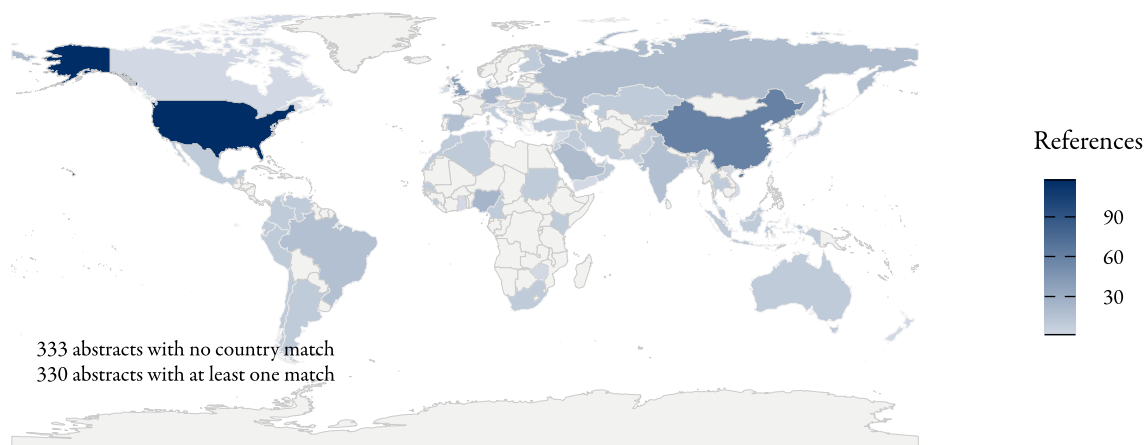


Figure 9: Most frequently studied countries in misinformation research. The figure displays the results of a manual identification of countries mentioned as the objects of study in our corpus. Among the papers analysed, 350 reported geographical information.

Research on the susceptibility to misinformation is on the rise.

Research on misinformation has grown exponentially in recent years. Between 2008 and 2017, scholarly output on the topic remained minimal, with no more than two publications per year. Publication counts rose sharply thereafter, reflecting growing interest in the topic following the US elections of 2016 (Freelon & Wells, 2020), surpassing one hundred papers in 2021, and have remained above that level since. Figure 10 displays the evolution of publications in this area and the level of the Complexity Framework studied – already showing that the vast majority of academic contributions focused on the micro-level.

Research on the susceptibility to misinformation is on the rise

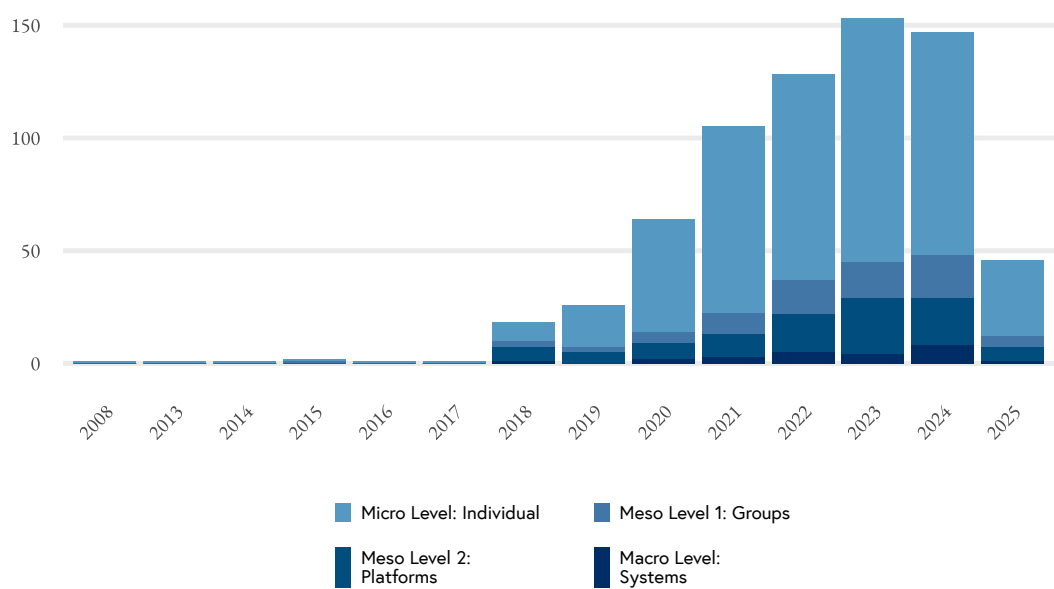


Figure 10: Number of papers per year by Complexity level of analysis. The figure reports the count results of our content analysis based on 663 paper abstracts.

Academic contributions overwhelmingly focus on micro-level factors.

The results of our manual content analysis show that most studies examined misinformation at the individual micro-level, with 77% (N = 510) of coded papers studying the creators, spreaders, or recipients of misinformation. Within this level, 72% of these studies examined recipients’ characteristics, while 41% examined producers. For example, Qerimi and Gërguri (2022) linked economic well-being to the propensity to believe false information, arguing that individuals with lower levels of education and income are more likely to believe it.

The second most common category referred to studies concerned with content factors (37%, N = 244), followed by 15% focused on platform-related variables at the second meso-level (N = 120). This percentage drop, visually displayed in Figure 11, indicates that the literature overwhelmingly investigates misinformation through individual-level determinants, particularly recipients’ susceptibility. Only a small number of studies examine collective or systemic factors, such as group dynamics (12%, N = 78) or system-level features of the

broader media environment. Though overall less prevalent, studies on digital platforms and their characteristics have gained ground in recent years, with more than fifty being published since 2023.

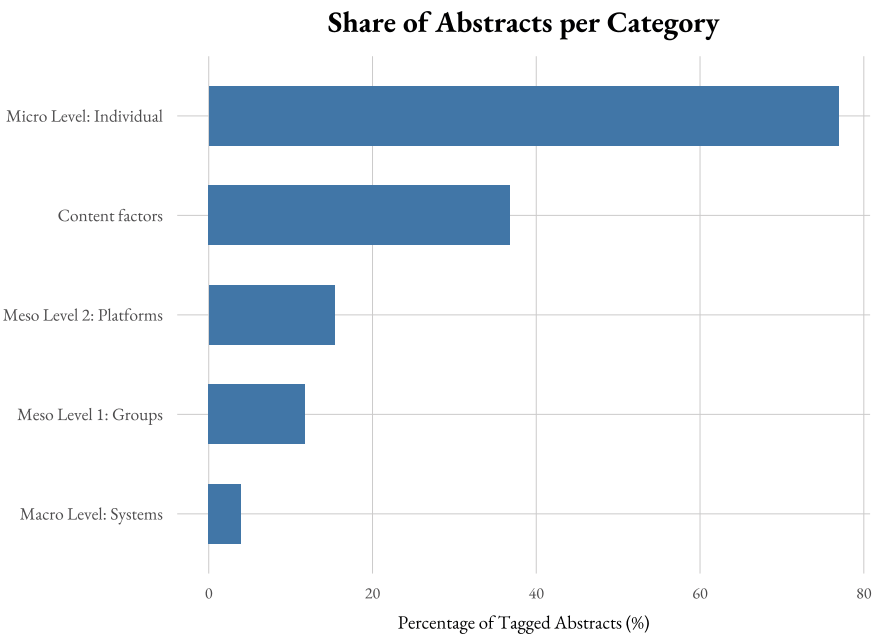


Figure 11: Share of abstracts per level of analysis. The figure reports the percentage results of our content analysis, including 663 paper abstracts.

Few academic contributions account for multiple levels.

Misinformation can be examined through variables that cut across multiple levels of analysis, but the majority of studies focused on a single level (58%, N = 387). 141 papers addressed both individual and content-related variables. Craig and Vijaykumar’s study (2023) exemplified this approach by examining the effectiveness of corrective messages against COVID-19 misinformation across different socio-demographic groups. Their study showed that corrective information presented after exposure to misinformation (so-called debunking) can reduce misperceptions; however, the effect can be short-lived if debunking is immediately followed by additional misinformation. Furthermore, they found that young adults were particularly susceptible to perceiving misinformation as truthful and credible.

Among the studies that analysed more than one level, typically a maximum of two levels were considered. This often occurred in studies examining content moderation policies, which involve both content and platform factors. Figure 12 visualises the frequency with which studies engaged with more than one analytical level as a heatmap.

Category Co-occurrence Heatmap

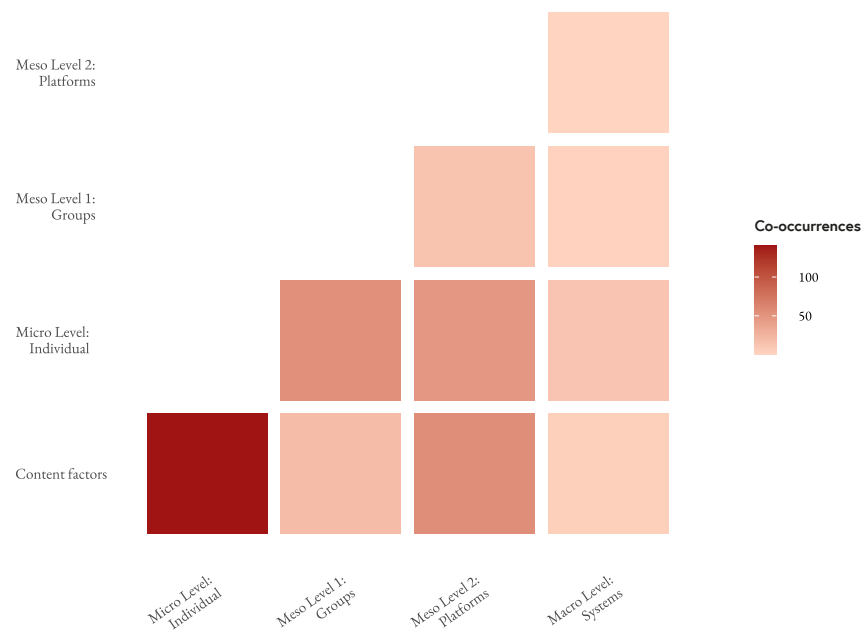


Figure 12. Heatmap expressing how often abstracts address more than one level (co-occurrence). Darker colours indicate more academic contributions mentioning the level on the horizontal and the vertical axis. Co-occurrences were identified in 276 abstracts (41 %).

6.8.4 Most Prevalent Topics in the Academic Literature (RQ2)

Methodological Approach

To understand which topics of misinformation susceptibility have been studied across the different Complexity levels, we employed a (semi-) automated computational approach using structural topic modelling (STM). The underlying logic of this approach is that words that often occur together can be under-

stood as related to the same topic, and that each document (or, in this case, scholarly article) potentially contains a mixture of topics (Günther & Quandt, 2016). Structured topic modelling is well suited to extract these topics within short texts, such as the abstracts analysed here (Roberts et al., 2019).

We improved our structured topic model by leveraging recent advances in machine learning technology. We did so by using a large language model (namely, OPEN AI's GPT) via its advanced programming interface to extract sentences from each abstract that explicitly referred to the paper's empirical analysis. Finally, we ran our structured topic model on these textual segments. The research team assigned interpretive labels to the identified topics based on exemplary abstracts and representative keywords. Methodological details, including how we determined the optimal numbers of topics and the GPT prompts, are provided in the Appendix.

6.8.5 Results

Main topics across levels

In this section, we move from our classification of academic articles to examining the topics most studied within each Complexity level. For each level, we focus on the three topics with the highest prevalence across the abstracts assigned to it, prioritising the most prominent themes in the field over exhaustive description. The exception to this is studies about social groups (meso-level I), which were not numerous enough for topic modelling of their abstracts.

Table 7 shows the three most prominent topics for each level, the keywords that formed the topic (in order of their frequency and exclusivity for the topic, Roberts et al., 2019) and the probability that an academic contribution addressed the respective topic. For a visualisation of topic proportions, see the Appendix.

Category	Label	Keywords	Theta
Micro-level: Individuals	Fake News Sharing	fake, news, emoticon, story, share, post, reason, article	0.202
	Promoting resilience	correct, misinform, mispercept, inocul, message, toward, exposure, skeptic	0.168
	Misinformation Spreaders	disinform, bot, network, twitter, troll, opinion, group, active	0.165
Content variables	Fact-checking effects	fact-check, label, reduce, headline, warn, false, post, tag	0.158
	Correction of misperceptions	mispercept, debunk, message, individual, exposure, inocul, toward, find	0.139
	Disinformation Strategies	disinform, public, play, strategy, generate, suspect, use, frequent	0.127
Meso-level 2: Platforms	Twitter Corrections	twitter, content, moderate, conspiracy, recommend, theory, tweet, spread	0.172
	Platform engagement features	misinform, pandemic, social, negative, threat, can, media, covid	0.163
	Heuristic cues	number, influence, credible, people, disseminate, trust, media, understand	0.161
Macro-level: media systems	Country specifics	misinform, country, impact, also, show, source, use, social	0.261
	Systemic roots	media, influence, politics, differ, source, social, can, use	0.231
	Trustworthy systems	news, credible, differ, source, belief, find, countries	0.191

Table 7: Summary of Structural Topic Modelling Results.
Notes. Theta estimated prevalence of a topic within abstracts assigned to a given level of the Complexity Framework. Values can range from 0 to 1 and indicate the proportion of text in those abstracts associated with the topic. Higher values indicate that the topic occupies a larger share of the abstracts related to a level. For example, Theta = 0.20 indicates that, on average, approximately 20 % of the textual content in abstracts classified at the micro-level is associated with the topic “fake news sharing.”

Cognitive and motivational processes dominate research at the micro-level.

At the micro-level, fake news sharing emerges as the most prevalent topic of analysis. This topic encompasses studies that investigate individual-level factors explaining why people share false information, often focusing on the

cognitive characteristics underlying such behaviour. One central factor influencing the susceptibility of recipients on this level is their cognitive reflection (Ross et al., 2021; Stecula & Pickup, 2021). Research, specifically in Psychology, has shown that individuals with a more analytical (as opposed to a more intuitive) thinking style (Pennycook & Rand, 2020), those who engage in careful deliberative thinking and are motivated to strive for accuracy, believe less in misinformation and are likely to share such content less (Pennycook et al., 2021).

Other studies have used the uses and gratifications framework (Katz et al., 1973), which postulates that media use is driven by expected gratifications and experiences with obtained gratifications, thus a motivational process. For example, Omar et al.'s study (2024) showed that fake news sharing in a Malaysian cross-sectional survey was predicted by factors such as the seeking of social status but also altruistic motivations.

A substantial share of studies also examines efforts to promote resilience to misinformation through pre- and debunking, as reflected in the recurrent use of terms such as correction, inoculation, and scepticism (e.g., Lewandowsky and van der Linden, 2021) which also tap into the idea of susceptibility being rooted in cognitive processes and a more thorough and critical engagement with content thus being able to reduce susceptibility. The third most prevalent topic concerns misinformation spreaders, encompassing research on bot activity, trolling, and the broader manipulation of public opinion formation through strategic actors.

Content is studied mostly through the lens of fact-checking.

Studies concerned with content variables primarily contained research on fact-checking effects. They presented research about the employment of tags, warnings, or labels on social media to combat misinformation. The second most recurring topic within this category is concerned with the correction of misperceptions, and focused on debunking strategies, introducing questions about truth judgement and inoculation (Mourali & Drake, 2022). This kind of research was commonly conducted through experiments, as in the case of Meiske et al.'s paper (2024), revealing that corrective comments can be effective in reducing belief in misinformation. Finally, this fact-checking category also contained a series of studies about disinformation strategies, including research about the rhetorical tools and strategies behind the spread of problematic information.

On the meso-level II, researchers primarily studied Twitter and platform logics.

Moving from the micro to the meso-level, the most prevalent platform-related topic concerns Twitter corrections. Studies in this category primarily examined Twitter's [now X] algorithmic recommendation systems and content moderation policies. The focus on this platform is symptomatic of an "API age", a period during which researchers still had extensive access to this platform's data (Mimizuka et al., 2025). Consequently, the dominance of this topic is likely to decline, giving way to the other two themes identified at this level: platform engagement features and heuristic cues, which capture how platform design and engagement metrics affect users' information processing and credibility judgments (e.g., Kim & Lim, 2025).

The few contributions at the macro-level highlight country-specificities and underscore the need to account for systemic aspects.

At the macro-level, we identified a small subset of studies examining the role of regulatory frameworks, institutions, and media systems in shaping misinformation dynamics. It is important to stress that only 25 studies in the entire dataset were assigned to the macro-level of analysis. This finding indicates that the existing literature seldom incorporated structural variables in their research design, a gap which limits our understanding of how nation-wide factors facilitate the spread of and belief in misinformation. The few already existing studies do, however, provide meaningful starting points for studying susceptibility on the level of states and institutions.

When studying country-specifics, scholars examined whether socioeconomic factors correlate with the prevalence of misinformation across contexts, as in the case of Cha, M. et al. (2021)'s study, which revealed that countries with lower Gross Domestic Product (GDP) tend to report higher perceptions of misinformation prevalence in the media. Some studies were grouped under systemic roots, revealing a scholarly interest in the relationship between political systems and misinformation. This interest is exemplified by Hunter's study (2024) revealing that authoritarian regimes are more likely to disseminate disinformation abroad on social media than democratic regimes.

The third common topic, trustworthy systems, addressed cross-country variations in news and source credibility. These studies expand the scope of misinformation research by adopting broader comparative and transnational

perspectives, for example highlighting the central role of trustworthy news media environments as protective factors (Humprecht et al., 2020).

6.8.6 Summary of Core Findings Study 1

Our first study demonstrates that susceptibility to misinformation is currently studied mostly on the micro-level of individual creators, spreaders, and recipients of misleading information. Relatively few studies have explored the wider group, platform, political, or the media system contexts within which these individuals are embedded. There are also very few studies that account for the complexity of the phenomenon or study links and feedback loops between the different levels. Our analysis underscores the importance of moving beyond individual factors and studying misinformation as a phenomenon shaped by media infrastructures and wider institutional and social conditions. For example, little research examined the role of social groups and intergroup processes in the creation and consumption of misinformation (for a similar argument, see Ziemer & Rothmund, 2022). While these findings echo past critiques of the field's disproportionate attention to individuals (Kont et al., 2025; Schüz & Jones, 2024), our framework allowed us to highlight critical areas of misinformation research – particularly platforms' algorithmic recommendation systems, content moderation policies, and engagement metrics.

At the micro-level, the academic literature contributes to our understanding of the cognitive processes shaping susceptibility to media manipulation. Studies in this area suggest that fostering resilience to misinformation requires strengthening individuals' cognitive skills through media literacy, fact-checking, and related interventions. They also underscore the importance of motivation, as engagement with misinformation is not driven solely by perceptions of its accuracy but also by the cognitive and social needs it may satisfy.

Research examining content-level factors largely approaches misinformation through the lens of fact-checking. Studies in this area investigate the effectiveness of corrective interventions in reducing belief in false information, with experimental research showing that such interventions can indeed show promising results. However, this focus on factual correction tends to treat misinformation as a problem of veracity, overlooking how this content may promote propaganda and media manipulation even when not easily categorised as simply true or false. Moreover, this body of research pays comparatively little attention to inattentional sharing, whereby users disseminate misleading

content not because they believe it to be accurate, but because they engage with information quickly and with limited scrutiny.

Relatively few studies adopt an intergroup or collective perspective on misinformation dynamics. At the meso-level, the literature has largely focused on social media platforms and algorithmic incentives. This work shows that the diffusion of misinformation is facilitated by the interaction between platform design and cognitive biases, such as attention to emotionally charged content. However, research examining broader structural factors remains limited. The few existing contributions nevertheless highlight the importance of media trust, political institutions, and socio-economic conditions in shaping misinformation dynamics.

International conflicts across the globe heightened concerns surrounding misinformation as a societal threat. Foreign actors exploit social media to pursue political and military objectives, rendering misinformation a strategic instrument of statecraft (Livingston & Miller, 2025). Fostering democratic resilience will require accounting for these topics and moving beyond individual susceptibility factors. The literature has yet to broaden its inquiries to examine how wider political and media systems can help to mitigate the spread of misinformation.

This line of inquiry is essential for scholars and policymakers seeking to keep pace with developments in this area, given the sweeping initiatives undertaken by state actors during crises. From internet shutdowns to the promotion of digital sleuthing to document war crimes, states have combated misinformation with limited scrutiny of their actions, the evidence motivating them, or their effectiveness. As we discuss in this report, developing effective responses to manipulation attempts requires accounting for these initiatives and the complex system through which misinformation circulates. The next chapter addresses this limitation by discussing interventions designed to foster democratic resilience (Study 2). Based on the insights we gained in Study 1 and 2, we will then examine how citizens evaluate these countermeasures to the spread of misinformation.

6.8.7 Limitations and Pathways for Future Research

When interpreting the results of Study 1, some important limitations should be kept in mind as they limit the generalisability of our findings to other contexts

and warrant further research. First, the current study is based on a systematic review of academic contributions written in English or German, therefore it does not include research written in other languages. This limitation is compounded by the fact that the majority of these articles studied a few countries in the Global North, as previously reported. Research on misinformation interventions has already underscored the need to test the applicability of interventions in the Global South (Blair et al., 2024) as interventions effective in the Global North might not have the same effects in other contexts. This points towards potential differences in susceptibility factors between different contexts and political or cultural environments.

Furthermore, our content analysis and structural topic modelling were based on the abstracts of these academic contributions. Though this methodological decision allowed us to produce insights based on a broad database, its drawback is that it risks overlooking information which was not reported in the abstracts. Finally, although the Complexity framework provides a useful way to organise research across levels of analysis, the feedback loops and reciprocal dynamics through which misinformation processes unfold require more scholarly attention and were not properly discussed here. Future research should therefore examine how these levels interact over time.

6.9 Study 2 – Tackling Misinformation

In this chapter, we shift our focus from how scholars have examined the spread of misinformation to the interventions proposed to counter it. The analyses in this chapter are guided by the following research questions:

- **RQ3:** Which levels of the complex framework have received the most attention when scholars discuss interventions to curb the spread of misinformation?
- **RQ4:** How do state responses to misinformation during recent international conflicts compare to the interventions emphasised in the scholarly literature?

We answered these questions by means of a quantitative and thematic content analysis of 130 academic contributions, which explicitly stated their intention to discuss solutions to misinformation and through a narrative review of case studies,

in which we examined how governments across the world communicated and responded to misinformation in recent international conflicts. The qualitative content analysis allowed us to synthesise research on countermeasures (Golovchenko et al., 2018) and map suggested interventions to the Complex system through which misinformation propagates – thereby identifying gaps where potential solutions lie unseen. An in-depth thematic coding of the articles further allowed us to extract detailed insights and identify patterns underlying current intervention strategies.

The review of case studies contrasted the academic discourse with actually implemented communication strategies during recent intergroup conflicts and other crises. We focused on crises because they are a “fertile ground” for the flourishing of mis- and disinformation (Allport & Postman, 1946), at a time when citizens urgently seek out trustworthy information (Perse & Lambe, 2016). Combined, both analyses allowed us to address our second research objective and develop a classification model of misinformation countermeasures ranging from reactive to proactive on one axis and from restrictive to permissive approaches on the second axis. This model accounts for both the provision of official communication and the countering of misinformation.

6.9.1 Why study interventions against misinformation?

Discussions of countermeasures have grown in importance as manipulations of the information environment have evolved substantially. Contemporary propagandists build on longstanding persuasive techniques while expanding their repertoire to exploit the affordances of digital media environments. This repertoire includes deploying automated accounts, such as bots (Marwick & Lewis, 2017), leveraging networks of amateur content producers to disseminate misleading narratives (Christopher & Matthews, 2016), and coordinating activity to crowd out or overwhelm rival voices (Till, 2020).⁴⁵

Contests over the interpretation of public affairs are not confined to social media platforms. Misleading claims can be amplified through legacy news organisations, in a phenomenon commonly described as “trading up the chain” (Lukito et al., 2019): In India, for example, television channels falsely reported

45 The use of amateur content producers by propagandist is key both in the dissemination of misinformation as well as in its production, as can be exemplified by misinformation campaigns in Taiwan: <https://www.aljazeera.com/news/2021/5/28/disinformation-goes-viral-as-taiwan-battles-new-covid-surge>.

that the army had entered Pakistan following a wave of online misinformation (Kumar, 2025). Information warfare has also extended beyond content dissemination to target technological infrastructures, including through cyberattacks and internet shutdowns (Dutsyk, 2022).

To curb the threat of misinformation, scholars have debated a range of interventions aimed at fostering resilience. These interventions can be mapped to different levels of the Complex system through which misinformation propagates. For example, scholars have studied the role of fact-checking skills and media literacy (e.g., Alon et al., 2024) and message-based responses on the micro-level of recipients and content. Christopher and Matthews (2016) examined the role of debunkings and showed that retractions or refutations have weaker corrective effects once problematic information has circulated. Others have developed computational tools to detect misleading online content (Wirtschafter et al., 2024), while again other work has emphasised the importance of complementing technical solutions with user-centred approaches. These arguments underscore that the spread of misinformation is a complex phenomenon spanning multiple levels of influence. Critically, countermeasures are not without risks and ethical controversies, as regulation may come at the cost of freedom of expression or surveillance creep (Surjatmodjo et al., 2024).⁴⁶ As Hoes et al. (2024) discussed, fostering scepticism towards misinformation in the population may also unintentionally lead to scepticism towards trustworthy information sources.

In sum, there is a growing need to synthesise research on misinformation countermeasures (Golovchenko et al., 2018). It is crucial to consider how crises shape implementation, given that extreme events have provided a pretext for states to rein in digital media and expand their control over information flows. This chapter addresses both gaps by reviewing scholarly discussions of countermeasures and comparing them with interventions implemented by states during crises. In doing so, we assess the extent to which interventions target the different levels of the Complexity system through which misinformation circulates.

46 Surveillance creep refers to a covered expansion of surveillance as the result of the introduction of convenient technological artefacts.

6.9.2 How scholars discuss interventions against misinformation (RQ3)

Methodological approach

To answer our third research question, we conducted quantitative and qualitative manual content analyses. From the database created for Study 1, we identified a subset of manuscripts explicitly stating an intention to discuss interventions against misinformation. Twenty percent of the papers (N = 130 out of 663) met this criterion. We then performed a content and thematic analysis of the full papers. We first assigned the intervention to the different levels of the Complexity framework, again allowing papers to be coded for more than one level. For instance, manuscripts examining the effectiveness of content moderation policies and fact-checking labels were both coded at the meso-level II of digital platforms and for content variables. Table 8 presents the codebook used in this analysis:

Code	Description
Producer characteristics	If the paper's intervention concerns a media producer's characteristics, write 1.
Recipient characteristics	If the paper's intervention concerns a recipient's characteristics, write 1.
Content factors	If the paper's intervention concerns the characteristics of messages, write 1.
Group factors	If the paper's intervention concerns social groups, write 1.
Organisational factors	If the paper's intervention concerns organisations, write 1.
Platform factors	If the paper's intervention concerns a platform's design or affordances, write 1.
Cross-country or media systems comparisons	If the paper's intervention concerns the characteristics of a country or media system, write 1.
Suggested interventions	If the study suggests or tests a solution to misinformation, write 1.

Table 8: Codebook for manual content analysis.

We complemented this investigation with a detailed thematic analysis to examine

- whether the academic contribution assigned responsibility to a specific actor for carrying out the intervention, and if so, which actor was deemed responsible.

- Whether the academic contribution only discussed or actually tested or evaluated an intervention.
- Whether the academic contribution mentioned (potential or observed) risks associated with the respective intervention.

We excluded articles from the analysis that did not test, propose or clearly discuss an intervention against misinformation. This was particularly the case for articles that mentioned an intervention in the abstract without further discussion, or for papers that only briefly mentioned interventions to be explored in future research.

6.9.3 Results

Micro-level interventions are the most common type of interventions discussed in the academic literature.

Approximately half of the coded manuscripts discuss interventions at the individual-micro-level, as shown in Figure 13. These studies focus on messaging strategies related to misinformation, often accompanied by guidance on how individuals can recognise and respond to misleading content. Several of the discussed interventions can be classified as so-called prebunking (Lewandowsky & van der Linden, 2021) – that is, they are intended to be employed before the consumption of misinformation to strengthen resilience against it. These interventions are often based on the theory of cognitive inoculation, which postulates that confronting individuals with small doses of a potentially persuasive message together with a warning that they might fall for it activates their “cognitive immune system” and helps them to defend themselves against the persuasion (for an overview, see Pfau & Szabo, 2003). For example, Zerback et al. (2021) discussed how media users become more resilient to manipulation attempts if they are instructed in advance about the exact arguments being deployed in these attempts, but they warn that official communications must be continuous rather than episodic.

Scholars also examined media literacy programs. These efforts included, for example, initiatives aimed at enhancing fact-checking skills through public information campaigns and school-based curricula (Lemaire et al., 2025; Rodrigo et al., 2024). Recent contributions have also addressed emerging threats associated with AI-generated misinformation and deep fakes (Sanchez-Acedo et al., 2021) and the need for adapting media literacy skills to the constantly changing digital environments.

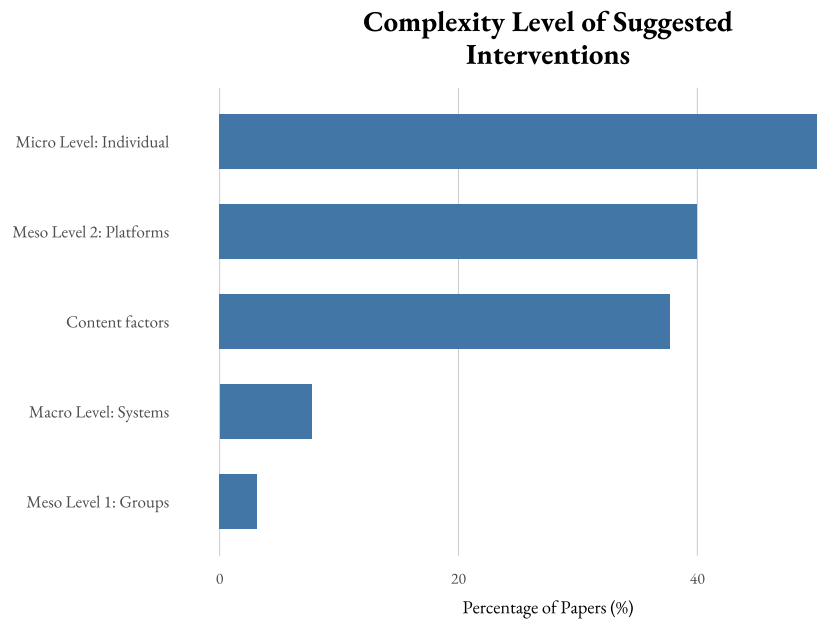


Figure 13: Share of papers proposing solutions per Complexity level. The figure reports the percentage results of our content analysis, including 130 manuscripts.

The second most frequently discussed category of interventions is Meso-Level II, which focuses on the role of digital platforms in the spread of misinformation. These interventions typically involve modifying the presentation of online content, which is why they span both platform- and content-level dimensions: approximately 40% of the manuscripts were coded as addressing both levels. One of the most commonly discussed interventions is the use of warning labels attached to misleading posts. Proponents argue that such labels, particularly when applied to content suspected of being amplified by bots, can reduce perceived credibility and decrease users’ willingness to engage with the content (Lanius et al., 2021). Tokita (2021), in turn, notes that flagging mechanisms can serve as inputs for additional platform-level interventions, such as introducing friction into the sharing process to slow the spread of flagged content. Other proposed interventions include the implementation of credibility badges or reputation scores attached to sources or user accounts (Prike et al., 2024). Finally, a set of studies discusses the role of behavioural norms in social media environments, for example in relation to accuracy prompts, that is, small nudges reminding people that others prefer accurate content (Pennycook et al., 2021).

A Research Gap in Meso-, Macro-, and Multi-Level Interventions.

The few studies tackling meso-level interventions were mostly concerned with platform-focused interventions, with very few academic contributions (3%, $n = 4$) examining group-level interventions at Meso-level I. In these studies, partisanship emerged as a key factor shaping the effectiveness of fact-checking strategies. Lee (2024) finds that corrections originating from cross-partisan sources can be more effective in fostering audience acceptance than those from co-partisan sources, highlighting the importance of perceived independence and credibility. Similarly, Zhang and Schmierbach (2024) argue that the effectiveness of corrections depends on group dynamics, including whether corrective messages are perceived as originating from one's in-group or an out-group. Nevertheless, few studies propose interventions that directly target group-level processes themselves. Instead, partisanship is typically treated as a moderating factor that constrains the effectiveness of existing interventions, rather than as a locus of intervention in its own right.

Macro-level interventions represent another comparatively neglected category, with only 7% of academic contributions ($n = 9$) addressing this level. When discussed, these kinds of interventions are typically framed as organisational rather than structural solutions. Some studies examine how corporations can develop protective measures to mitigate reputational damage resulting from misinformation (Boman, 2021). Others focus on organisational routines aimed at building trust (Mahdi et al., 2022) or on reactive strategies, such as issuing responses to false content and engaging directly with actors responsible for disseminating it (Hancock, 2020; Kim & Lim, 2025). Beyond the corporate contexts, a smaller number of studies address broader institutional and national-level strategies. Among them, Rasheed and Naseer (2021) discuss coordinated cybersecurity efforts as part of comprehensive defence strategies against disinformation. Aro (2016) highlights the potential role of legal interventions, including the use of legal frameworks to prosecute coordinated disinformation campaigns and the strengthening of regulatory requirements for social media platforms to rein in fake accounts.

Overall, our findings confirm prior research by Kozyreva et al. (2024), who noted that individual level approaches have been most commonly examined, while system-level approaches are underexplored despite their potential to make the social or digital ecosystems less conducive to the spread of misinformation. Measures such as implementing labels and guidelines on digital platforms or

promoting diverse media systems account for multiple levels of the Complexity framework. Yet, studies spanning more than one level of the Complexity framework remain exceptions in the literature, as Figure 14 demonstrates:

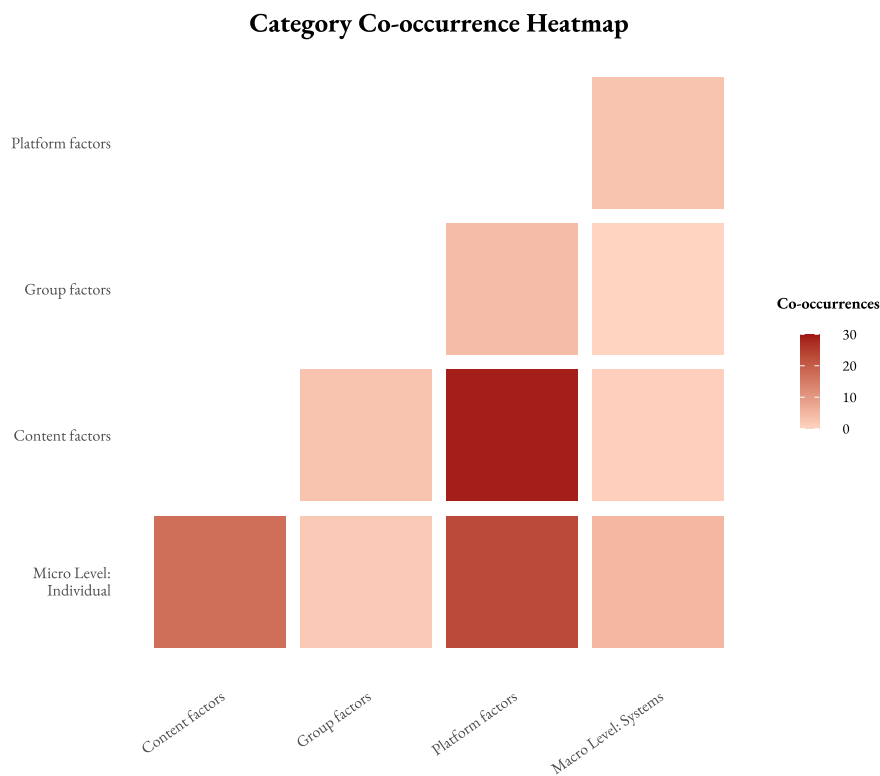


Figure 14: Heatmap expressing how often articles discussing interventions addressed more than one level (co-occurrence). Darker colours indicate more academic contributions mentioning the level on the horizontal and the vertical axis. Co-occurrences were identified in 59 articles (45%).

Proposals for interventions often ascribe responsibility (and power) to platforms – overlooking political and civic influence.

When it comes to the question of who is considered responsible for implementing countermeasures, the thematic analysis showed that scholars overwhelmingly assign responsibility to social media platforms, while comparatively few scholars specify directly the roles that policymakers, civil society organisations, or other actors should play – even though these are likely the actors and institutions most relevant for implementing, for example, media literacy programs at scale (see Table 9).

Category	Count	%
Specifies who should carry out the intervention	86	66.2
Empirically Tested	54	41.5
Mentions Risk	26	20

Table 9. Summary of the thematic analysis.

Notes. Based on 130 analysed academic contributions.

Approximately 66% of the coded manuscripts identified a responsible actor, although these attributions were often articulated in broad terms. Social media companies emerged as the actors most frequently mentioned, reflecting their ability to modify platform architectures, implement algorithmic adjustments, and introduce technical countermeasures. Policymakers were the second most commonly referenced category, yet scholars rarely specified the relevant level of governance or policy mechanisms.

As shown in Table 10, references to other actors were substantially less frequent. These included private companies and marketing professionals, public health authorities, and journalists or news media organisations. Overall, the literature assigns primary responsibility to platform companies, while offering limited accounts of the roles that public institutions, civil society actors, or media organisations might play in addressing misinformation. Critically, this also limits the scholarly evidence regarding the (in-) effectiveness of communicative measures implemented by state actors. While scholarly debates have increasingly examined interventions to curb misinformation, far less attention has been paid to how governments respond to these threats in practice during crises.

Responsible Actor	Count
Social Media Companies	49
Policymakers	37
Private companies and marketing professionals	18
Public Health Authorities	15
Journalists or news media	13

Table 10: Top 5 actors deemed responsible for carrying out misinformation interventions.

Notes. This summary is based on the 130 manuscripts coded in this study. N = absolute number of mentions.

Proposals for interventions rarely discuss potential side effects and risks of interventions.

Although scholars have raised important concerns regarding countermeasures against misinformation, only 20 % of the coded manuscripts addressed potential risks associated with the interventions they discussed. The risks predominantly discussed included backfire effects occurring with debunking strategies targeting false content. This was either due to debunkings repeating the false claim with the correction resulting in its amplification (Lee, J. et al., 2023) or because corrections were based on low-credibility sources (Chen et al., 2022; Lee, Y.-I. et al., 2024). Scholars found that resistance against such corrections was particularly noticeable among sceptics or partisan users (Gounaridis & Newell, 2024; Sharevski et al., 2022).

The second most commonly discussed risk is related to media literacy and inoculation practices, with scholars noting that they may lead to generalised distrust of news overall, including accurate content (Ferrucci & Hopp, 2023; Gimpel et al., 2021). Finally, source labels or credibility badges were discussed as potentially resulting in over-trust or premature dismissal of content (Prike et al., 2024).

6.9.4 How do state responses to misinformation during recent international conflicts compare to the interventions emphasised in the scholarly literature (RQ4)?

Methodological approach

The second part of this chapter moves beyond the academic literature to documenting state responses to misinformation during crises. In light of the recent increase in international conflicts worldwide, we compiled case studies on how governments have sought to regulate and at times control the information environment in these contexts. Examining state responses provides an opportunity to assess the extent to which real-world interventions align with those emphasised in the academic literature, as well as to identify forms of intervention that remain underexplored in scholarly discussions. In response to RQ4, our analysis focuses on interventions implemented within the last three years, allowing us to capture recent developments in how governments confront information threats under conditions of geopolitical tension. The selection of case studies was guided by the goal of maximising variability (including Pakistan, India, Ukraine, Russia), thereby providing insights that can be transferred to several contexts.

We identified these cases through a review of reports published by international news media, government officials, and civil society organisations. The analysis is not intended to provide comprehensive documentation of all state responses undertaken during this timeframe. Its purpose is to map the spectrum of interventions that have emerged in recent crises, highlighting the diversity of approaches and identifying measures that have gained prominence in recent years.

6.9.5 Results

Media literacy as a necessary skill for every citizen.

One area of convergence between scholarly discussions and state practice concerns is the investment in media literacy initiatives. Governments have increasingly adopted media literacy as part of broader national security strategies, particularly in countries exposed to sustained information warfare. For example, Ukraine, Latvia, Georgia, and Moldova underwent “securitisation of the media literacy as part of a national security and defence” strategy (Denisa-Liepniece, 2021, p. 4). These efforts included the establishment of dedicated state agencies, curricular reforms, and partnerships with civil society organisations. Ukrainian authorities produced daily informational content directing citizens to verified sources and promoted initiatives such as the “Filter” project, designed to strengthen news verification skills among domestic and international audiences (Dutsyk, 2022). Similarly, Poland’s Ministry of Defence and the Academic Centre for Strategic Communication launched the #Fejkoodporni (#Fakeproof) campaign, a cross-media literacy intervention aimed at raising awareness of misinformation as a wartime threat (Glapiak, 2023). Nordic countries, for instance Sweden, distributed preparedness leaflets advising citizens how to navigate information environments during crises (Henley, 2018), while Estonia’s “AI Leap” initiative focused on strengthening public capacity to identify AI-generated content. These initiatives echo scholarly emphasis on educational interventions targeting individual resilience through literacy.

Broadening the security frameworks, Finland implemented a national media and education policy in 2019 acknowledging that the “primary objective of media education is to promote people’s willingness and ability to act actively and responsibly in media culture, and thus reinforce the desired future trends,” (Salomaa & Palsa, 2019, p. 5). Media literacy has also been part of Finland’s national curriculum since 2014, which defines multiliteracy – that is, the ability to obtain, process, and verify information – as a central transdisciplinary competency in all school subjects.

The European Union considers media literacy a “crucial skill for citizens of all ages, enhancing their critical thinking and ability to better navigate in the digital media environment and identify disinformation” (European Commission, n.d.). The Audiovisual Media Services Directive (AVMSD, to be revised 2026) requires member states and video sharing platforms to actively promote media literacy (European Commission, n.d.). The European Democracy Action Plan (EDAP), the Media and Audiovisual Action Plan (MAAP), as well as the Digital Education Plan, all recognise the need to foster digital and media literacy. Furthermore, the Commission’s Media literacy expert group meets regularly to identify, document, and extend good practices, coordinate EU policies and support initiatives and build synergies between Member States.

Media literacy initiatives as part of security frameworks have a long history, traceable back to anti-Nazi propaganda bulletins distributed in the US during World War II (Sproule, 2001). Over time, however, they have evolved through the incorporation of participatory practices (Mimizuka et al., 2026). In the Philippines, media literacy workshops have trained citizens to detect misinformation and improve online safety (DW, 2024). In Georgia, the Media Development Foundation has developed workshops that bring citizens together to produce articles using fact-checking skills (Zach, 2024). The German initiative “Journalism makes school” (n.d.) and the European NGO LieDetectors bring professional journalists to schools to foster news and misinformation literacy (n.d.). Media literacy initiatives are not always state-sponsored but often driven by civic society, non-governmental organisations, or journalistic media.

Fact-checking and debunking initiatives.

State authorities also play a consequential role when it comes to the debunking of misinformation. One example is a website created by the U.S. Department of State to catalogue Russian disinformation, presented as exposing “Russian lies about the current crisis and its causes – and the truth.” This case illustrates how governments use the internet not only to counter hostile narratives but also to position themselves as information authorities. Yet, it is important to note that such a pursuit of epistemic authority has become increasingly diffused and contested. For example, the ultra-conservative and highly contested American Heritage Foundation accused the EU Disinformation center (established to debunk misinformation and information operations) of “centralizing control” and institutionalizing “top-down ‘media-literacy’” (McCarthy, 2026).

Governments have expanded institutional mechanisms to monitor online content. In India, authorities empowered the Press Information Bureau to evaluate the veracity of online information and mandate its removal, while executive orders required social media platforms to block accounts under threat of fines or imprisonment. Policies that place the state in a gatekeeping role have sparked intense debates about censorship and government overreach ('Assault on free speech', 2023). These interventions position the state as a central gatekeeper of the information environment, highlighting a divergence from scholarly literature, which has emphasised educational and platform-based interventions rather than coercive forms of information control.

Including citizens' perspectives.

Countermeasures are not only implemented (and contested) by state actors but also emerge through the inclusion of citizens and grassroots initiatives. For example, Ukrainians formed a coalition with hundreds of volunteers, the so-called PR Arming, connecting war witnesses and communication professionals to international media. Initially operating through Telegram and later Slack, this coalition sought to publicise its perspective on the full-scale invasion and shape international news agendas. Other grassroots efforts have focused on independent fact gathering, as in the case of Mediazona and BBC Russian compiling war casualties records to estimate the number of Russian soldiers killed in Ukraine (Basmal, 2025; Klingová, 2024). Indeed, domestic and foreign organisations increasingly produce war accounts by leveraging civilian online content production.⁴⁷ Rather than merely strengthening individual resilience, these cases signal that grassroots actors participate in the production and dissemination of war accounts, highlighting their role as active participants in shaping the information environment during crises.

Entertaining and social media optimised official communication.

To reach their target audience, especially younger citizens, state actors increasingly also adapt to social media logics (Klinger et al., 2024). In Taiwan, state ministers boosted their media visibility by hiring comedians and tapping into platform vernaculars, suggesting that a "humour over rumour" tactic may

47 Amnesty International produced numerous reports based on open-source evidence, as exemplified by their investigation of "Raqqa 'death trap'" in: <https://www.amnesty.org/en/latest/press-release/2019/04/syria-unprecedented-investigation-reveals-us-led-coalition-killed-more-than-1600-civilians-in-raqqa-death-trap/>. Access on 01/20/2026.

be effective during crises (Quito, 2020). Recent conflicts suggest that even military spokespersons can effectively adapt to social media communication logics and achieve high visibility online. During the 2025 India–Pakistan conflict, Air Vice Marshal Aurangzeb Ahmed of the Pakistan Air Force appeared in several viral clips, helping circulate official messaging to domestic youth audiences (The Nation, 2025).

Restrictions of information.

Not all official communications are conducted through relationships of cooperation and the inclusion of citizens or media literacy initiatives. Governments also establish guidelines for what they conceive as illegal communication and increase control over information flows. For example, the United Arab Emirates regulates influencer communication by making the spreading of rumours or unverified information a crime punishable by law. After the 2026 attacks of Iran on the United Arab Emirates, several European Social Media influencers based in the country posted videos with the same message “You live in Dubai, aren’t you scared?” followed by the answer “No, because I know who protects us” combined with images of the Prime minister of the United Arab Emirates and his son, the Minister of Defence. This episode heightened concerns about the authenticity of online content and what was allowed to be posted on social media (Grenier, 2026).

Crises may often lead to tensions in the relationship between governments and news media. In the US, the military has restricted journalists’ access to the Pentagon and required approval prior to the public release of information, with the risk of losing press passes for noncompliance. In the 2026 Middle Eastern war, journalists in Israel and Iran faced censorship over their reporting, which included restrictions against recording city skylines during missile attacks and their impact sites (AFP, 2026). Such restrictions highlight the persistence of tensions around press freedom, especially when reporting exposes governmental contradictions or weakens public support (Kirat & Slamene, 2024).⁴⁸

48 International law (ICCPR) permits restrictions on freedom of expression to protect national security, public order, and related interests, but these limits remain contested when applied to journalism. More about the Pentagon decision can be found in: <https://www.theguardian.com/us-news/2025/sep/20/pentagon-journalists-restrictions-pledge>. Access on 01/20/2026.

Recent crises prompted states to adopt coercive and regulatory interventions aimed at asserting greater control over media environments, interventions which remain underemphasised in the scholarly literature. In 2023 alone, crises were associated with 283 internet shutdowns across 39 countries, underscoring the extent to which governments rely on infrastructural control to manage information flows (Access Now, 2025). Such shutdowns have been used to restrict international scrutiny of authoritarian crackdowns, as illustrated by Iran's response to protest movements (Amnesty International, 2026). Beyond infrastructural disruption, states have also pursued regulatory and discursive strategies targeting specific actors. For instance, Russia has designated U.S. technology companies as "foreign agents," while other governments have sought to delegitimise civil society organisations engaged in countering online harms, as in the case of efforts to label Germany-based NGO HateAid as extremist (hateaid, 2025).

Measures to reduce information flows are also implemented by democratic actors. In Ukraine, new legislation prohibited the import of printed materials from Russia and banned Russian media outlets from operating domestically. Under martial law, the criminal code was amended to impose criminal liability for disseminating unauthorised information about military operations (Dutsyk, 2024). Similar legislation can be found in Estonia, where new regulatory frameworks allow for the suspension of foreign state-aligned broadcasters accused of disseminating propaganda. As another example of a rather intense intervention, the European Union banned Russian state-sponsored alternative counter-news Russia Today and Sputnik in direct response to Russia's full-scale invasion of Ukraine, 2022 (Council of the European Union, 2022).

These findings indicate convergence as well as critical divergence between scholarly discussions and state responses to misinformation. Governments across the globe have adopted media literacy programs, reflecting scholarly emphasis on strengthening individual resilience. However, state responses extend beyond educational interventions to include strategic communication, institutional adaptation, legal sanctions, and direct control over information infrastructures. These interventions remain comparatively underexamined in academic literature. This gap suggests that scholarly discussions may underestimate the extent to which regulatory and coercive tools are implemented during crises to manage information environments.

6.9.6 Summary of Core Findings Study 2

This chapter examined interventions against misinformation through a review of scholarly discussions and an analysis of countermeasures implemented against the background of international crises. Figure 15 presents a heuristic map of these interventions across the levels of the Complexity Framework, distinguishing between proactive measures – designed to operate prior to exposure to misinformation, including both direct prebunking efforts and broader systemic strategies aimed at fostering resilience – and reactive measures, which are implemented after exposure has occurred.

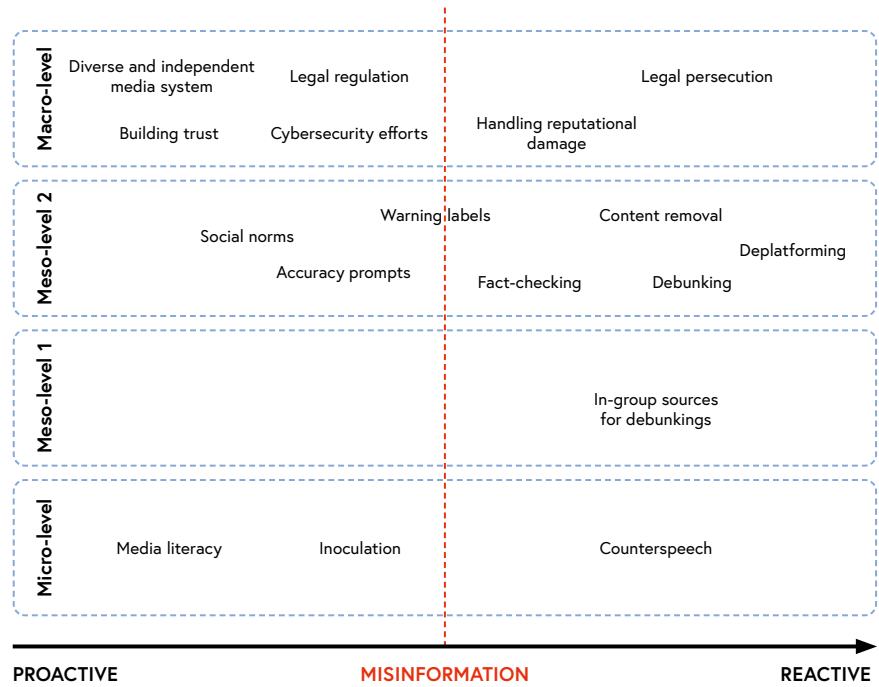


Figure 15. Heuristic visualisation of countermeasures across the Complexity Framework.

Our results indicate that the scholarly literature is heavily concentrated in the lower-left quadrant of Figure 15, emphasising proactive, micro-level interventions aimed at strengthening individual resilience. Media literacy and educational initiatives dominate this space, emerging as the most common type of intervention in our content and thematic analysis. To a lesser degree, the literature also engages with meso-level approaches focusing on platform governance and content moderation. Consistent with the findings from the previous chapter, macro-level interventions received comparatively limited attention in the examined

literature and, when addressed, were often articulated as broad institutional recommendations rather than empirically grounded analyses of state practice.

Responsibility for implementing countermeasures is overwhelmingly attributed to social media companies, with policymakers, civil society organisations, and other institutional actors receiving marginal attention and remaining insufficiently specified in terms of their roles. Moreover, the literature devotes limited consideration to the potential risks, unintended consequences, and normative trade-offs associated with these interventions, including their implications for democratic accountability and freedom of expression.

By comparison, our analysis of state responses during recent international conflicts reveals a broader distribution across the map in Figure 8, extending beyond the lower-left quadrant into both reactive and higher-level interventions. Governments have implemented extensive media literacy initiatives, often in collaboration with civil society organisations, suggesting a degree of convergence between scholarly recommendations and state practice in this domain. Nevertheless, these initiatives also illustrate how media literacy interventions may move upward and rightward within the framework, as they become embedded in national security strategies rather than remaining purely educational. This repositioning highlights a closer alignment between civil society actors and state strategic objectives than is typically acknowledged in the literature. Moreover, such efforts do not necessarily reflect best practices from inoculation research and may overlook potential side effects of highly dramatised campaigns, including the risk of eroding trust in online information more broadly, including established news coverage (Jungheer & Rauchfleisch, 2024).

State responses also extend into the upper-right quadrant of Figure 15, encompassing reactive and macro-level interventions that remain largely absent from scholarly discussions. Governments have asserted greater control over information environments by restricting journalistic access, banning media outlets, imposing internet shutdowns, and introducing legislation enabling criminal sanctions for unauthorised information dissemination. Such measures often entail the expansion or reconfiguration of institutional authority, raising important concerns regarding accountability and the concentration of power. Despite their prevalence, these interventions remain comparatively underexamined and lack systematic research on their negative consequences, including

the possibility that they erode trust in state institutions and, in turn, increase susceptibility to misinformation (Frischlich & Humprecht, 2021).

Limitations.

The current study is based on the full articles in our database that explicitly mentioned interventions in their abstracts. As such, it is limited to scholarship that explicitly addresses the susceptibility to digital misinformation flows.

The analysis focuses on English and German publications, which may lead to underrepresentation of research produced in other linguistic and regional contexts.

Third, the identification of case studies relied on international news coverage and publicly available reports, which further constrains the scope of our analysis. Because such sources are disproportionately available in English, this chapter's discussions are skewed toward highly visible or geopolitically salient cases.

6.10 Study 3 – Democratic Resilience Against Misinformation

The previous chapter documented a proliferation of measures to curb misinformation, arguing that these interventions have received little scrutiny in the past. This chapter addresses this by providing new evidence on citizens' preferences regarding how misinformation should be addressed, particularly under crisis conditions.⁴⁹

A central challenge for measures intended to promote democratic resilience is their public acceptability. Interventions that are not perceived as legitimate or trustworthy may fail to correct misperceptions, even when they provide

49 We align our study with the literature on conflict communications by focusing on communication efforts designed to foster democratic resilience; these efforts, in turn, are constitutive of wider official communications aimed at informing the public, "with the ultimate goal of enhancing public self-efficacy and fostering enhanced cooperation" (Glapiak, 2023). These strategies target security threats, described by NATO as "no longer bound by geography and can have impact on a sub-state or worldwide basis. They are not even bound by terrestrial limits and may manifest themselves in space or cyberspace" (2011, p. 115)

the public with accurate information (Stubenvoll & Matthes, 2023). From a policy-making perspective, the range of measures considered acceptable at a given moment can be described as the Overton window. This concept was originally proposed by Joe Overton (Russell, 2006) to describe the range of policies that can be implemented without jeopardising public support and, consequently, viability. This window is not fixed: it can expand or contract, making it critical for policymakers to understand which policies are likely to be accepted or contested.

In this study, we examine the range of official communication efforts and measures against misinformation (which we refer to as their “information environment”) that citizens deem (un-) preferable, i.e., that jointly lie within, respectively outside of their Overton window. Research has shown that citizens have clear preferences when it comes to how digital platforms should be designed (Stock et al., 2024), regulated (Riedl et al., 2022), and how misinformation on these platforms should be moderated (Wiesner, 2025).

This Study builds on this work and examines citizens’ preferences regarding their information environments, and whether these preferences shift during crisis conditions. It does so by addressing the following research questions:

- **RQ5:** What are Austrians’ preferences when it comes to official communications and responses to misinformation?
- **RQ6:** Do crises, particularly hybrid warfare, shift citizens’ preferences?
- **RQ7:** How are citizens’ preferences related to their susceptibility towards misinformation?

Methodological Approach

To answer our research questions, we conducted a conjoint choice experiment (Hainmueller et al., 2014) embedded in an online survey, which was sent out to the access panel by Marketagent.

Sample

A total of $N = 4,485$ individuals clicked on the survey link and $n = 2,013$ completed the questionnaire. Most of those who did not finish the questionnaire failed to give informed consent on the first page. Another 559 participants failed an attention check or finished the questionnaire in less than a minute (indicating automated participation), being excluded from the sample. 23 asked

for their data to be deleted before the analysis (“right to be forgotten”) and were thus also not included in the analysis.

The final sample consisted of $N = 1,985$ participants, of whom 1,001 were women, 985 were men, and one identified as diverse. Participants were on average 47.63 years old ($SD = 16.81$) and were broadly representative of the Austrian population in terms of regional distribution. Regarding education, the largest group had completed professional training ($n = 564$, 28%), followed by those with a high-school degree (“Matura”, $n = 494$, 28%) and those with a university degree ($n = 498$). A minority ($n = 262$, 13%) reported lower educational degrees (e.g., completion of middle school). In terms of employment status, most participants ($n = 1,079$, 54%) were currently employed, while 5.29% ($n = 105$) were searching for a job. The majority of respondents were born in Austria and had parents who were also born in the country, whereas 7.20% ($n = 145$) reported a migration background.

Design

The survey instrument is provided via open access via the Open Science Framework.⁵⁰ The study employs a conjoint experimental design, which accounts for how individuals form attitudes based on multiple attributes simultaneously (Hainmueller et al., 2014). This approach reflects the reality that attitudes and choices are typically formed based on combinations of features rather than isolated characteristics. For example, voters might prefer one political candidate over another because of his or her policy agenda, gender, charm, among other attributes which can hardly be understood in isolation.

This chapter examines citizens' preferences regarding official communication activities and measures to counter misinformation, which combined form the “information environment” in which citizens are embedded. Study 2 characterised this information environment along two dimensions. First, we described official communications as more or less permissive, depending on the extent to which they enable actors to participate in them. Second, we described communications as proactive or reactive, depending on whether they are implemented before or after exposure to misinformation. These two

50 Preregistration can be found here: <https://aspredicted.org/aq7j3w.pdf>, the survey instrument can be accessed here https://osf.io/bqkru/overview?view_only=788404e1031b40f7a0945f2c7de5006c.

dimensions constitute a “compass” which allows us to navigate and describe information environments.

The Compass of the Information Environment

To characterise information environments, our compass accounts for official communication strategies and countermeasures to misinformation as facets of the information environment. These initiatives are situated along the two dimensions previously explained: proactive – reactive, permissive – restrictive (see Figure 16). These dimensions capture both the timing of interventions and the extent to which actors are incorporated into communication processes. The compass identifies six attributes along these dimensions.

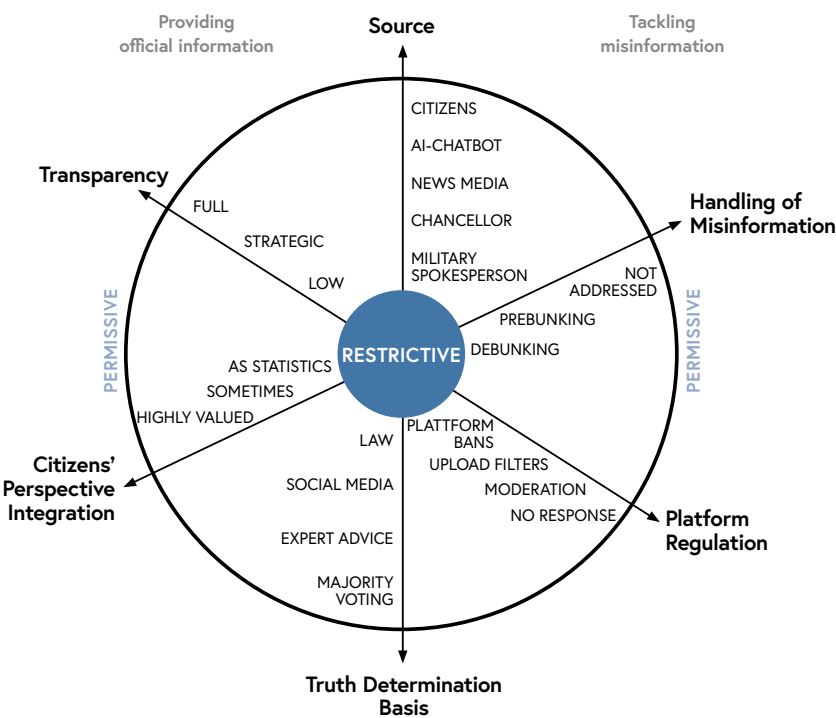


Figure 16: Compass of the Information Environment.

The first attribute concerns the **(1.) source of official communication**. Military spokespersons were positioned at the restrictive end of the spectrum, followed by the chancellor, news media organisations, AI-chatbots and, at the inclusive-end of the spectrum, citizens. The second attribute is **(2.) level of transparency**. Official communication is characterised as restrictive when it selectively discloses information, sharing essential government decisions

while withholding details to preserve social cohesion or morale. Transparent communication involves the comprehensive disclosure of information, even when it may risk public concern. Official communications also vary as to the extent to which they integrate **(3.) citizens' perspectives**. Communication is more restrictive when their perspectives are sidelined in favour of state officials', journalists', and experts' statements, appearing only through statistics. At the inclusive end, citizens' perspectives are treated as equally relevant and are directly incorporated alongside those of institutional actors.

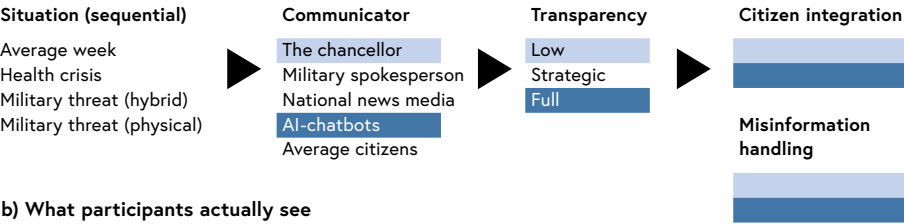
The next attribute concerns the **(4.) handling of misinformation**. Approaches that prioritise debunking content are positioned toward the restrictive end of the spectrum, as they typically result in binary judgments (e.g., "true" vs. "false"). Less restrictive approaches emphasise fostering media literacy and equipping citizens with skills to verify information. At the permissive end, individuals are made responsible for distinguishing between factual and false content on their own. Approaches to **(5.) social media regulations** were considered restrictive if they involved outright bans of platforms or the screening of user content. Less restrictive measures include content moderation that responds to posts only after they have been published. At the permissive end, a laissez-faire approach entails the absence of regulatory intervention (Frischlich et al., 2019).

The last attribute concerns **(6.) who determines what constitutes falsehoods**. The most restrictive arrangements involve legal regulation, where determinations are made through formal legislation. This is followed by decisions made by social media platforms and less restrictive approaches which rely on evidence-based assessments or majority judgments, which distribute authority more broadly across actors.

Choosing one's preferred information environment.

This compass oriented the conjoint experiment. Participants were presented with two vignettes, each describing a different information environment. Each vignette described six characteristics of these environments, and these characteristics were randomly drawn from the compass, as visually displayed in Figure 17.

a) Design of conjoint experiment



b) What participants actually see

E.G., Physical Military threat: Which of the following scenarios of communication do you prefer to learn about current events?		Official communication			Addressing misinformation		
		Communi-cator	Transpa-rency	Citizen integration	Mis-information	Regulation	Truth base
	Vignette A	Chancellor	Low	As statistic	Prebunking	Blocked	Law
	Vignette B	AI-chatbot	High	As valuable perspective	Debunking	Labelled	Majority votes

Social media regulation

Basis for determining truth

Figure 17. Visualisation of the construction process of the vignettes in the conjoint experiment. The dice indicate that the attribute levels (e.g., the specific communicator, such as military spokesperson) were drawn at random. Each vignette contained six attributes: (1) Source of Communication; (2) Transparency Level; (3) Citizen perspective integration; (4) Misinformation handling; (5) Platform regulation; (6) Truth base.

For example, a participant might be presented with a first vignette (“Vignette A”) describing an environment in which official communication is primarily delivered by the chancellor, limited to essential information, and incorporates citizens’ perspectives only in statistical forms. In our conceptualisation, this would be considered a “restrictive” information environment. The same participant would simultaneously be shown a second vignette (“Vignette B”), in which official communication is summarised through AI chatbots, information is disclosed more fully, and citizens’ perspectives are treated as equally valuable alongside those of experts, journalists, and political actors. This vignette would represent an inclusive information environment. Both vignettes also varied in how misinformation is addressed, platforms are regulated, and how truthful content is discerned from falsehoods. The participant would then be asked which of the two they preferred.

Examining the Role of Crises

To understand how the Overton window might shift during a crisis (RQ6), we examined the role of different crises scenarios. Prior to each set of vignettes, participants were instructed to imagine that the country was experiencing one of four situations: an average week, a health crisis, a hybrid military threat, or a physical military attack.⁵¹ These situations were presented sequentially, such that participants evaluated information environments under each condition. They all started with an “average week” situation, with the subsequent situations presented in randomised order. Respondents expressed their preferences across all four situations.

6.10.1 Results: Austrians’ general preferences (RQ5)

To understand Austrians’ general preferences towards their information environment (RQ5), we first focused on their “average week” situation preferences. These were calculated as the probability that a vignette would be chosen if it featured a specific attribute level (e.g., high transparency). This probability receives the name of Marginal Means (MMs). It indicates participants’ favourability towards that attribute level, while averaging over all other characteristics in the information environment.⁵² Because respondents chose between two vignettes, the MMs ranges from 0 to 1. A MM greater than 0.5 indicates that the attribute level increases the probability of a profile being chosen, thus signalling preference among participants, while a MMs below 0.5 indicates a reduced likelihood relative to other options.

Austrians favour transparent, news media-led information environments that incorporate citizens’ perspectives.

With regard to the source of official communications, Austrians show a clear preference for being informed by the news media. Vignettes in which official information is provided primarily via national news media coverage were the most likely to be selected (being chosen 64% of the time on average,

51 The public health crisis was introduced as: “Austria is in the midst of a national medical crisis that is affecting the lives of many citizens (comparable to the COVID-19 pandemic).” The hybrid conflict was introduced as: “Imagine Austria is involved in a non-violent military conflict with another country (‘hybrid conflict’). There have been various cyber-attacks and espionage cases in recent months.”

52 Operationally, marginal means are computed across every possible combination an attribute appears with across vignettes, that is why it is “averaged over” those other characteristics.

i.e., $MM = 0.64$). The preference for news media was statistically significant. Information environments in which official communication came alternatively via AI-chatbots were chosen only 36 % of the time, indicating a strong aversion to this option. Table 11 reports these results in greater detail.

Participants leaned moderately towards information environments where official communications do not withhold information from them, but their preferences regarding high ($MM = 0.53$) or strategic ($MM = 0.51$) transparency did not differ statistically significantly from each other. The clearest finding concerns their least favoured option: Austrians demonstrate unfavourability towards information environments that restrict information to most essential government decisions, which was only chosen 47 % of the time. This aversion aligns with participants' preferences for permissive information environments that highly integrate citizens' perspectives, treating them as valuable as state officials', journalists', and experts' perspectives ($MM = 0.55$).

Austrians reject “laissez-faire” approaches to misinformation.

Our findings reveal that Austrians prefer misinformation to be actively addressed. Participants favoured information environments where citizens are taught prebunking strategies ($MM = 0.54$) and that fact-checking initiatives are part of official communications ($MM = 0.53$). Both options were significantly preferred over no official engagement, which was chosen only 44 % of the time. While the results point to clear preferences concerning addressing misinformation, they were less conclusive about social media regulations. We found no significant preference for content moderation, upload filters, or platform bans relative to the baseline. However, participants demonstrated a clear aversion to platforms not reacting to misinformation, an option chosen only 44 % of the time. This suggests that while Austrians are divided on the method of regulation, they do reject a “hands-off” approach.

Finally, participants clearly preferred scenarios in which “what counts as misinformation is determined by expert advice (e.g., science, fact-checking)” (58 %) over all other options. In contrast, they disliked scenarios in which “what counts as misinformation is determined by majority vote (e.g., through community notes, online rating systems)” (46 %). Table 11 details these results.

Dimension	Level	MM	SE	95% CI	n Profiles	n Respondents
Source of official communication	News media	0.636	0.017	[0.602, 0.670]	780	705
	Chancellor	0.547	0.018	[0.513, 0.582]	795	716
	Citizens	0.501	0.018	[0.465, 0.536]	763	690
	Military	0.457	0.017	[0.424, 0.491]	846	748
	AI-chatbot	0.363	0.017	[0.329, 0.396]	786	702
Transparency	High transparency	0.525	0.014	[0.498, 0.552]	1316	1109
	Strategic transparency	0.506	0.014	[0.479, 0.533]	1306	1085
	Low transparency	0.470	0.014	[0.443, 0.496]	1348	1113
Integration of citizens	Participatory	0.547	0.014	[0.521, 0.574]	1358	1144
	Only as statistic	0.485	0.014	[0.457, 0.512]	1302	1094
	Sometimes	0.466	0.014	[0.439, 0.493]	1310	1107
Handling of misinformation	Prebunking	0.542	0.014	[0.515, 0.569]	1304	1094
	Debunking	0.526	0.014	[0.499, 0.553]	1301	1102
	No engagement	0.435	0.013	[0.409, 0.461]	1365	1151
Platform regulation	Moderation	0.528	0.016	[0.496, 0.560]	951	839
	Upload filter	0.519	0.016	[0.488, 0.550]	973	847
	Platform bans	0.511	0.015	[0.481, 0.541]	1071	927
	No regulation	0.442	0.016	[0.411, 0.473]	975	849

Dimension	Level	MM	SE	95% CI	n Profiles	n Respondents
Basis for truth determination	Expert advice	0.575	0.016	[0.544, 0.606]	1000	888
	Legal basis	0.502	0.016	[0.471, 0.533]	1012	881
	Platforms	0.461	0.016	[0.430, 0.492]	981	864
	Majority	0.461	0.016	[0.429, 0.492]	977	858

Table 11: Marginal means for attribute levels in “average week” scenario.

Notes. This table reports the probability a vignette would be chosen when the respective level was included, with this probability being measured through marginal means (MM). SE = Standard error, measure the statistic uncertainty of the MM estimates. CI = 95% confidence interval, meaning that, if the experiment were repeated many times, the true population mean would fall within this range 95% of the time. If the 95% CIs of two levels do not overlap, the difference in preference is statistically significant at the $p < 0.05$ level. N profiles = the number of vignettes in which an attribute level was present, n recipients = the number of participants that saw at least one vignette featuring the attribute level.

In sum, Austrians preferred an information environment characterised by permissiveness, favouring official communication provided via national news media, high levels of transparency, and the active inclusion of citizens’ perspectives. They disliked leaving misinformation unchallenged, though they were divided as to how it should be addressed. Preferences for proactive and reactive measures were diffused, with both being favoured over inaction. Our results showed support for initiatives that call out misinformation, empower citizens to detect it, and leverage expert advice for labelling it. These findings indicate that information environments where official communications are transparent, inclusive, and maintain a mandate for intervention lie within the Overton window (see Figure 18).

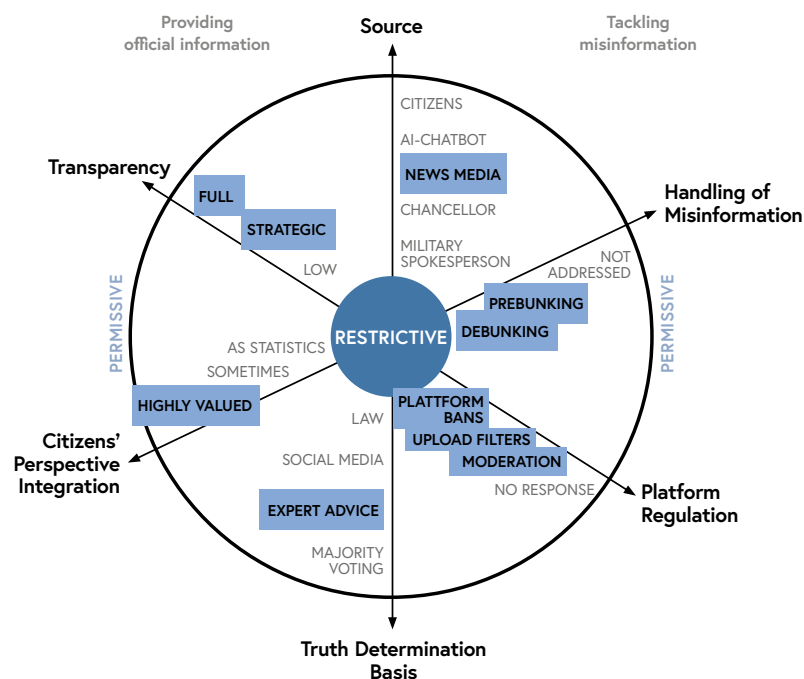


Figure 18: Overton window for official communications and misinformation regulation in an average week. Attribute levels which are preferred are highlighted in blue, while those which are disliked (i.e., lie outside of the Overton window) are presented in grey.

6.10.2 Results: Austrians' Preferences in Times of Crises (RQ6)

Next, we investigate whether these preferences change under crisis conditions (RQ6). To this end, we compare the results obtained under an “average week” situation to preferences under health and military crises. This section thus explores the causal effects of crises in how Austrians form their preferences concerning information environments.

These causal effects are estimated through so-called Average Marginal Component Effects (AMCEs, Hainmueller et al., 2014). Unlike Marginal Means, which show absolute probabilities, AMCEs represent the change in the probability that a vignette will be selected when shifting from a baseline category to a different attribute level (e.g., from high to low transparency). We use the most preferred attribute levels – identified in the previous section – as our baseline categories. AMCEs are expressed as percentage point differences relative to these baselines. For example, an AMCE of -0.27 (or -27 pp) for an information environment where official communication is provided by an AI-chatbot in-

stead of the news media means that, all else being equal, participants were 27 percentage points less likely to select that vignette. These estimates are visually displayed in Figure 19, with more details in the Appendix.

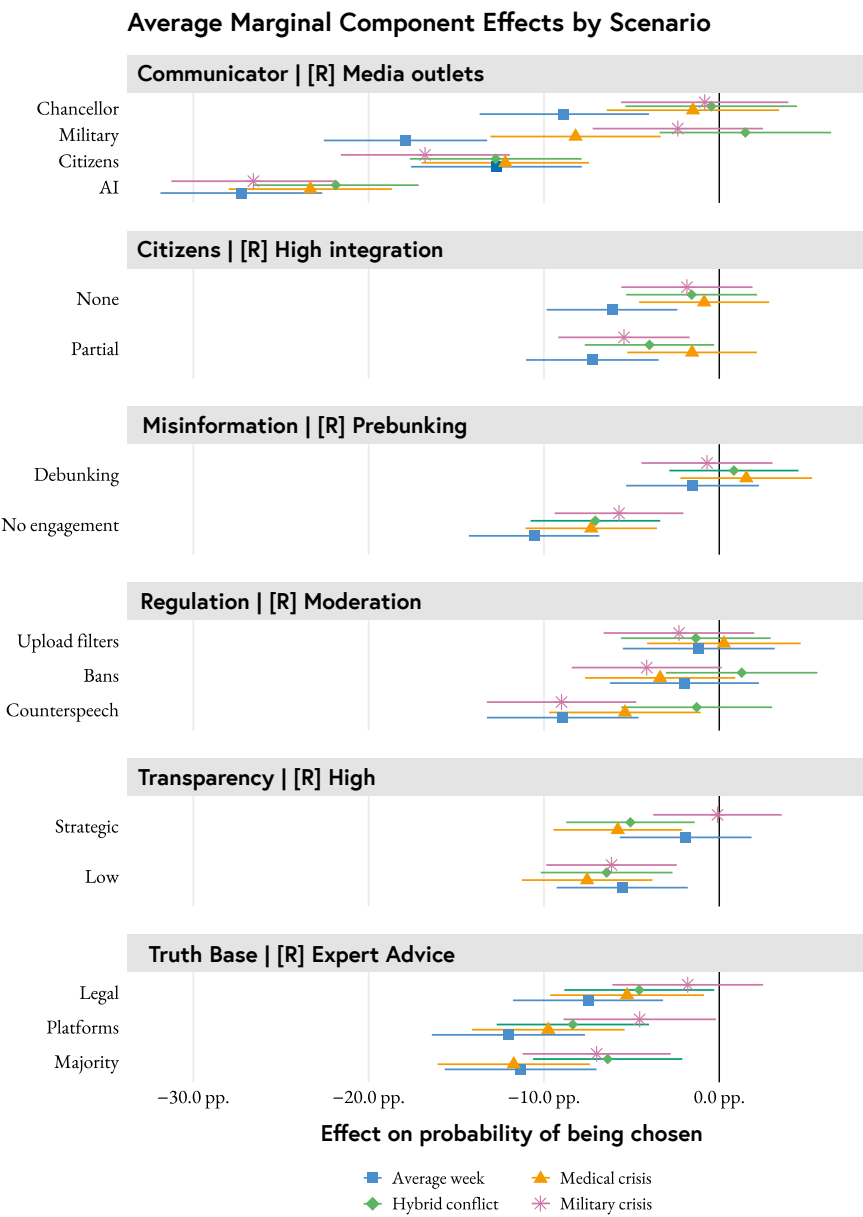


Figure 19. Average Marginal Component Effects by Situation. Average Marginal Component Effects are expressed as percentage-point changes in the probability that a profile is chosen relative to a baseline. The reference category is written in the grey boxes and indicated by the horizontal line. Reference categories were picked based on the marginal mean estimates presented in Table 11. Horizontal lines surrounding AMCEs estimates indicate the 95% confidence intervals.

During crises, state communication matters – but only during military crises is military communication widely accepted.

In average weeks, information environments in which state officials communicate directly with citizens were preferred significantly less than those where communication is provided via news media. Shifting the main source of official communication from the news media to the Chancellor reduced the probability of a scenario being chosen by 9 percentage points. A shift to a military spokesperson reduced that probability by 18 pp, while AI-chatbots were the least popular, leading to a 27 pp decrease. However, we found important shifts in these preferences during crises. During hybrid or physical warfare, scenarios featuring a military spokesperson were no longer statistically less likely to be chosen than those featuring the news media. This acceptance does not extend to medical crises, where military communicators remain significantly less favourable than news media (-8 pp).

Across crisis situations, participants were statistically as likely to choose information environments where official communication came directly from the Chancellor as those where it came from the news. These shifts suggest that citizens expect state authorities to assume a more active role in communicating official information during crises, even as news media remain a preferred source. For military spokespersons, this increased public acceptance is restricted to military contexts.

Transparency matters – especially in times of hybrid warfare.

In an average week, information environments that reduce transparency in favour of social cohesion were significantly less attractive than those that were fully transparent (-6 pp). This pattern holds during physical warfare: while low transparency is penalised (-6 pp), there is no statistically significant preference for full transparency over an approach centred on motivating citizens. These results shifted during crises, when low (-8 pp) and strategic transparency became significantly less favourable than full transparency (-6 pp). This underscores a heightened public demand for unfiltered official communication during crises. It also suggests that while citizens may grant the state “strategic” discretion during traditional physical warfare, they remain sceptical of information-withholding in the face of hybrid threats or public health emergencies.

Citizens value individual perspectives – but accept expert-focused communication during military crises.

In an average week, participants showed a strong preference for “participatory” environments where citizens’ perspectives are valued equally to those of officials and experts. Information environments that included citizens’ perspectives only through statistics (-6 pp) or “sometimes” while prioritising experts (-7 pp) were selected significantly less often. This suggests that during stable times, integrative communication is a vital tool for building trust and a sense of inclusion.

In all crisis scenarios, however, the preference for high citizen integration was no longer statistically significant. Participants became more inclined to choose official communications that focus on the statements of state officials, journalists, and experts. Information environments where citizens’ perspectives were sometimes included, while experts were preferred, were disliked in both military situations (hybrid warfare, -4 pps.; physical warfare - 5 pps.), suggesting that either an expert oriented or an integrative communication strategy is more favourable during conflicts. During medical crises, the integration strategy did not significantly shift preferences, suggesting a larger tolerance for different communication strategies during such times.

Citizens want official communication to address misinformation.

Across both average weeks and crisis scenarios, participants significantly favoured measures against misinformation over inaction. Scenarios in which misinformation was not officially addressed – focusing only on current events – were chosen significantly less often than those featuring pre-bunking strategies (with AMCEs ranging from -6 pp to -11 pp). Across all examined situations, there was no statistically significant difference in preference between “pre-bunking” (media literacy initiatives) and “debunking” (including fact-checking). This reveals that citizens support both empowering the population to detect misinformation alongside maintaining a clear mandate for official fact-checking and debunking activities.

Citizens support platform regulation over leaving it to users themselves to call out misinformation – but are split under conditions of hybrid warfare.

During average weeks, medical crises, and physical warfare, citizens significantly disliked “no-response” scenarios where platforms did not react to misinformation, leaving it to other users to refute (Counterspeech). Compared to a standard

moderation baseline, this “no- response” approach reduces the probability of a vignette being chosen by 5 to 9 percentage points (pp). Participants showed no clear preference for standard moderation over more proactive or restrictive measures, such as upload filters or banning platforms that fail to remove harmful content. Notably, these reactive scenarios explicitly included references to “harmful misinformation”, therefore the support for such measures is likely contingent on demonstrating actual harm. Taken together, these results confirm that while there is a firm mandate for platform intervention, the specific method of regulation is less critical to the public than the requirement that action be taken.

Importantly, this consensus disappears during hybrid warfare. In this situation, there was no statistically significant preference for moderation over a “no-response” approach. As shown in Figure 19, the 95% confidence intervals for all regulation types cross the zero line, indicating that the clear mandate for intervention observed in other crises is absent here. These findings pose a challenge to regulators, as hybrid wars are associated with more hesitant preferences regarding platform regulation.

What is considered misinformation should be determined based on expert advice and scientific evidence.

Across all examined situations, information environments where misinformation is determined by expert advice (e.g., science, fact-checking) are preferred over other approaches. During an average week, participants significantly disliked scenarios where misinformation was defined by legal regulations (-7 pp), majority opinion (e.g., community notes; -11 pp), or social media platforms (-12 pp). This pattern persists during medical and hybrid crises. Physical warfare presented the only exception: in this situation, citizens did not significantly prefer expert determination over legal regulations, suggesting a higher level of tolerance for state-led definitions in times of physical conflicts.

In sum, Austrians’ general preferences shifted during crises –compared to the general preference for news-based, inclusive and transparent information environments in which misinformation was determined based on expert advice and addressed in official communication. Participants showed a higher tolerance for elite-focused communication during times of threat, sometimes at the expense of citizen participation. This includes acceptance of direct communication from the Chancellor and, specifically during international conflicts, the military. Hybrid warfare led to a heightened demand for full disclosure in official communications,

a rejection of a “laissez-faire” approach to misinformation regulation but unclear and divided preferences regarding platform regulation.

6.10.3 Citizens’ susceptibility to misinformation and their preferences for different information environments (RQ7)

The conjoint experiment was embedded in a larger online survey that included measures for variables that have been found to be related to individuals’ susceptibility to misinformation. In line with the Complexity framework, we accounted for variables at the individual micro-level, meso-level I of social relations and collective identities, the meso-level II of digital platforms, and the macro-level. To do so, we focused on participants’ perceptions related to these different levels: For example, we examined participants’ satisfaction towards macro-level democratic institutions, not the state of Austrian democracy itself. The full questionnaire is provided via the open science framework (https://osf.io/bqkru/overview?view_only=788404e1031b40f7a0945f2c7de5006c).⁵³

Measurements

We compiled a survey instrument in German and English, the two languages in which the survey was available. We used preliminary exploratory factor analysis and internal consistency analyses to check the reliability of the variables scales and adapted them, if needed. All scales, reliability, and mean values are provided in Table 12.

On the micro-level of individual media users, we included two blocks of variables:

- We included socio-demographic variables (age, gender, profession, educational level and migration background) to describe participants. To account for subjective perceptions, we also included a measure for perceived status in society (Adler et al., 2000).
- We also examined psychological resilience. As indicators for participants’ general resilience and psychological susceptibility to crises, we measured trait resilience (Sinclair & Wallston, 2004), life satisfaction (Cheung & Lucas, 2014), and perceived control over one’s life (Lachman & Weaver, 1998).

53 Missing values for this analysis (3.7% in total) were imputed using a random forest algorithm (MissForrest) which takes the other answers in the survey to predict the most representative replacement for the missing value. Both the Normalized Root Mean Square Error (NRMSE) and the percentage of falsely classified cases (PFC) were below 0.01, speaking for an excellent imputation result.

Additionally, we accounted for participants' need for cognition. Research has shown that a more intuitive or emotional instead of a more analytical cognitive style increases the susceptibility to misinformation (Pennycook & Rand, 2021), so, we measured participants' need for cognition (Beißert et al., 2014).

With regard to the meso-level 1, we focused on political and ideological attitudes. Based on prior research on the "supply" of misinformation (e.g., Humprecht, 2019; Hunter, 2024; Törnberg & Chueri, 2025), we measured participants' political ideology (from left to right, Wasmer et al., 1996), their level of Right-wing Authoritarianism (Duckitt and Sibley, 2010), their social-dominance orientation (Aichholzer, 2019) and their level of conspiracy mentality (Bruder et al., 2013) and populist attitudes (Akkerman et al., 2014). Moreover, we accounted for the level of loneliness as an indicator of how satisfied participants were with their social relationships.

With regard to the meso-level 2 of digital information environments, we measured participants' media diets. We asked participants

- for their use of different news channels (e.g., the TV, social media) in the last week (Gadringer et al., 2025).
- We also asked participants how often they would use "social media and video platforms" and whether they would use specific platforms (e.g., Facebook, YouTube).

With regard to the macro-level, we measured participants' satisfaction with democracy and their trust towards a range of institutional actors (Frischlich & Humprecht, 2021). Drawing from the Eurobarometer measure (OECD, 2017), these institutional actors included the government, political parties, the media or the armed forces. Based on a factor analysis, we grouped these institutions as democratic institutions (the government, the parliament, and political parties), non-elected actors (the police, armed forces, and ordinary citizens), NGOs, science, and established news media. We also measured participants' overall satisfaction with democracy (Wike et al., 2022).

Measures tapping into the susceptibility to misinformation.

To explore the validity of our selected susceptibility factors, we asked participants how often in the last months they had seen someone on social media share misinformation (Boulianne & Humprecht, 2023) as opposed to information "they disagree with" (modelled after Ribeiro et al., 2017). We also evaluated

participants' ability to detect misleading and factual headlines using the MIST test (Maertens et al., 2023). These measures were presented at the end of the survey, to avoid alerting participants to misinformation in advance of the conjoint experiment. Based on preliminary reliability analyses, we examined participants' ability to detect misinformation based on their results in identifying four misleading headlines as false. Table 12 shows an overview of all items including their reliability scores (where available), their mean value or proportion in the sample and their standard deviation.

		Variable	α	M	SD
micro-level	Socio-demographics	age	–	47,63	16,18
		subjective SES	–	5,87	1,60
	Cognitive Style	NFC (revised index)	0.7	2,69	0,96
	Psychological resilience	psychological resilience	0.86	3,37	0,82
		life satisfaction	–	3,59	0,89
		perceived control	–	3,67	0,90
social meso-level	Political & ideological attitudes	political ideology	–	6,10	2,24
		populism	0.84	3,61	0,80
		conspiracy mentality	0.86	3,34	0,87
		SDO (revised index)	0.67	2,79	1,01
		RWA submission	–	2,73	1,12
		RWA conservative	–	2,09	1,22
		RWA aggression	–	2,83	1,12
	Loneliness	UCLA	–	2,22	0,84
macro-level	satisfaction with democracy		–	3,10	1,00
	trust	trust in non-traditional actors	0.84	2,13	0,93
		trust in democratic institutions	0.88	2,61	0,95
		trust in unelected actors	0.71	3,31	0,81
		trust in NGOs	–	2,82	0,99
		trust in science	–	3,67	1,11
		trust in established news	–	2,86	1,09

		Variable	α	M	SD
	Misinformation	perceived exposure misinformation	–	2,85	1,13
		perceived exposure counterattitudinal content	–	3,30	1,08
		MIST (fake detection)	0.64	3,03	1,17

Table 12: Descriptives and reliabilities for different scales in the survey

Notes. M = Mean, SD = standard deviation, α = Cronbachs α , values > .80 are good, > .70 acceptable, > .60 questionable, values >.50 indicate unreliable scales. “--” reliability could not be computed as this was a single item measure. Multi-item scales were revised to improve internal consistency and reliability and aggregated to a single mean value.

For NFC, the reverse coded items did not load onto the same factor and did not form a reliable factor on their own. We thus formed a new index with two items directly expressing need for cognition (“I like it when my life is full of tricky problems that I have to solve” and “I would prefer more complicated problems to simple ones”).

RWA showed an unsatisfactory internal reliability (Cronbachs’ alpha = .53) and a lack of a coherent factor structure in an exploratory factor analysis. We thus used single items to measure the three core-dimensions of right-wing authoritarianism. Items were selected based on their loading on distinct factors (authoritarian submission: “What our country needs most is discipline, with everyone following our leaders in unity”; conventionalism: “God’s laws about abortion, pornography, and marriage must be strictly followed before it is too late”, authoritarian aggression: “Our society does NOT need tougher government and stricter laws” (reverse coded).

For SDO, one item tapping into egalitarian ideas (“We should do what we can to equalise conditions for different groups”) and loaded only weakly on the same factor. This item was removed, improving internal consistency.

For trust, exploratory factor analyses revealed three subdimensions which we labelled trust in non-traditional actors (trust in “big-social media companies”, “influencers”, “alternative news” and “AI”), trust in democratic institutions (trust in “the government”, “parties”, and “politicians”), and trust in unelected actors (trust in “the police”, “the military”, and “average citizens”), reliability analyses furthermore indicated that a superordinate factor would also match the data structure. Trust in science, the media, and NGOs showed cross-loadings and was thus analysed separately.

For misinformation and truth discernment, an exploratory factor analysis identified two factors. However, the reliability of the factor measuring participants’ ability to identify correct news was unsatisfying (Cronbachs’ α < .45). Thus, we focused only on the items measuring participants ability to correctly classify false statements as false. Items were aggregated to sum index, counting the correct answers across items.

Susceptibility to misinformation is Related to Micro-, Meso-, and Macro-levels from the Complexity Framework.

This section assesses whether participants’ ability to detect misinformation is related to the micro-, meso-, and macro-level variables under study, thereby testing the validity of our measurements. To this end, we calculate correlations using Spearman’s ρ . The resulting values can range from - 1 to +1, with positive results indicating that increases in one variable (e.g., age) are associated with increased values in the other variable (e.g., misinformation detection). Negative values indicate the inverse association between the variables, while a null result indicates no association at all. Table 13 shows these correlation results. It confirmed that variables related to the micro, meso- and macro-levels are associated with participants’ perceived exposure to misinformation as well as with their ability to correctly identify misinforming headlines as “false”. Confirming the validity of the chosen susceptibility factors on each level, perceived exposure and misinformation detection skills showed distinct association patterns and were also distinct from perceptions of being exposed to counter- attitudinal content in digital environments.

Block	Measurements	Perceived exposure to misinformation	Perceived exposure counterattitudinal content	Mis-information detection
Micro-level				
Socio-demographics	age	-0.24***	-0.18***	0.19***
	male	0.03	0.01	0.11***
	migration background	0.06**	0.03	-0.07**
	higher education	0.02	0.10***	0.18***
	lower education	-0.01	-0.02	-0.04
	subjective SES	0.01	0.04	0.09***
Cognitive style	NFC (revised index)	0.05*	0.02	-0.07**
Psychological resilience	Psychological resilience	0.13***	0.10***	-0.19***
	life satisfaction	-0.07**	-0.05*	0.14***
	perceived control	-0.09***	-0.06**	0.11***

Block	Measurements	Perceived exposure to misinformation	Perceived exposure counterattitudinal content	Mis-information detection
Meso-level I				
Political and ideological attitudes	political ideological	-0.02	-0.01	-0.23***
	populism	0.01	0.03	-0.23***
	conspiracy mentality	0.09***	0.04	-0.42***
	SDO (revised index)	0.08***	0.05*	-0.19***
	RWA submission	0.08***	0.00	-0.25***
	RWA conservative	0.06**	-0.05*	-0.31***
	RWA aggression	0.03	0.02	-0.07**
Loneliness	UCLA	0.13***	0.10***	-0.19***
Meso-level II				
Traditional media channels	TV	-0.11***	-0.04	0.14***
	radio	-0.05*	0.01	0.14***
	newspaper online	0.19***	0.25***	0.02
	newspaper offline	0.04	0.11***	0.18***
	broadcasters online	-0.06**	-0.05*	0.17***
New media channels	social media	0.07**	0.13***	0.14***
	AI chatbots	0.13***	0.18***	0.00
Alternative news	alternative news	0.06*	0.07**	-0.16***

Block	Measurements	Perceived exposure to misinformation	Perceived exposure counterattitudinal content	Mis-information detection
Social media platforms	Facebook	0.16***	0.19***	0.00
	YouTube	0.16***	0.22***	0.02
	TikTok	0.16***	0.13***	-0.16***
	Instagram	0.19***	0.22***	-0.04
	X/Twitter	0.06*	0.07**	0.00
	WhatsApp	0.18***	0.23***	0.03
	Telegram	0.06**	0.08***	-0.05*
	alternative social	0.04	0.06*	-0.09***
Macro-level				
Democracy satisfaction	Democracy satisfaction	-0.07**	-0.05*	0.29***
Trust	trust in non-traditional actors	0.05*	-0.01	-0.14***
	trust in democratic actors	-0.02	-0.04	0.24***
	trust in unelected actors	-0.10***	-0.06**	0.11***
	trust in NGOs	0.05*	0.00	0.11***
	trust in science	0.01	0.04	0.31***
	trust in established news	-0.02	-0.03	0.30***

Table 13: Correlations between the susceptibility variables and participants susceptibility to misinformation.

Notes. N = 1,985. Reported are the zero-order Spearman ρ correlation coefficients.

*** $p < .001$, ** $p < .01$, * $p < .05$. Correlations with $\rho > 0.1$ = small, $\rho > 0.3$ = medium, and $\rho > 0.5$ = large associations.

6.10.4 Identifying Susceptible Target Groups

Methodological Approach

We identified susceptible groups through cluster analyses, which are statistical procedures for grouping cases based on a set of predefined variables (Field & Graham, 2002). In this case, the set of variables relates to different levels of the Complexity framework. Given their distinctiveness, we implemented a two-step clustering procedure: First, we ran separate cluster analyses for each level of the Complexity framework. In the second step, we identified how often participants ended up in the same cluster across these analyses using an ensemble classification algorithm (Ghosh & Acharya, 2011). This algorithm groups participants together based on their similarity across cluster analyses, leading to more robust results. We report full details of this clustering analysis in Appendix B.

The analysis proceeded in three stages. First, participants were assigned to clusters based on the hierarchical clustering solution. Second, we estimated a series of multinomial logistic regression models to identify the variables that best characterised membership in the clusters across analytical levels. Predictors that showed limited explanatory power were not retained in the integrated visualisation of cluster membership. Third, we examined whether the clusters differed in their ability to detect misinformation using analyses of variance (ANOVAs).

6.10.5 Three Clusters: Institutional Mainstream, Networked Apart, and Right-Wing Authoritarian.

The clustering analysis identified three distinct participant profiles. Cluster 1, labelled the “Institutional Mainstream” ($n = 898$, 45.2%), is characterised by the highest levels of democratic trust, satisfaction with democracy, psychological resilience, and subjective socioeconomic status. Members of this cluster are ideologically moderate, displaying a slightly right-leaning orientation alongside comparatively lower levels of populist and conspiratorial attitudes. They also show the strongest reliance on legacy news sources, particularly television and offline newspapers, while reporting comparatively lower use of social media for news. Socio-demographically, this cluster is older than the other groups, highly educated, and least likely to report a migration background. Detailed information on cluster composition is provided in Appendix B.

Cluster 2 (n = 467, 23.5%) brought together participants who report moderate democratic trust and satisfaction with democracy and whose ideological profiles closely resemble those of Cluster 1. What most clearly distinguishes this group is its greater reliance on social media and AI chatbots for news, paired with comparatively weaker use of traditional media sources. At the same time, this cluster scores substantially lower on resilience and perceived control, while reporting the highest levels of loneliness. Because they are digitally connected but report feeling socially detached, we label this cluster “Networked Apart.” We also observed that this cluster includes the highest proportion of women, is the youngest group, and reports the lowest subjective socioeconomic status.

Cluster 3, labelled the “Right-wing Authoritarian” (n = 620, 31.2%), stands out as the most ideologically right-leaning group. Its members are distinguished by the highest levels of populist, authoritarian, and conspiracy mentality, alongside the lowest trust in and satisfaction with democracy. In terms of media habits, this cluster reports the weakest consumption of mainstream news sources and the highest reliance on alternative news outlets. Socio-demographically, its members are less likely to have higher education and more likely to report a migration background than other clusters’ members (though the overall share of participants with a migration background was low both in the survey and this cluster).

Multinomial regression analyses confirmed the distinctiveness of these clusters, as Figure 20 shows. Taking the “Institutional Mainstream” as the baseline category, the multinomial models reveal two pathways away from the mainstream. The first, represented by Cluster 2 (“Networked Apart”), is driven by psychological vulnerability traits, namely: lower resilience and perceived control coupled with greater loneliness. The second, represented by Cluster 3 (“Right-wing Authoritarian”), is shaped by right-wing ideological orientation, heightened populism, authoritarianism, conspiracy mentality, and markedly lower democratic trust and satisfaction, pointing to an anti-establishment disposition.

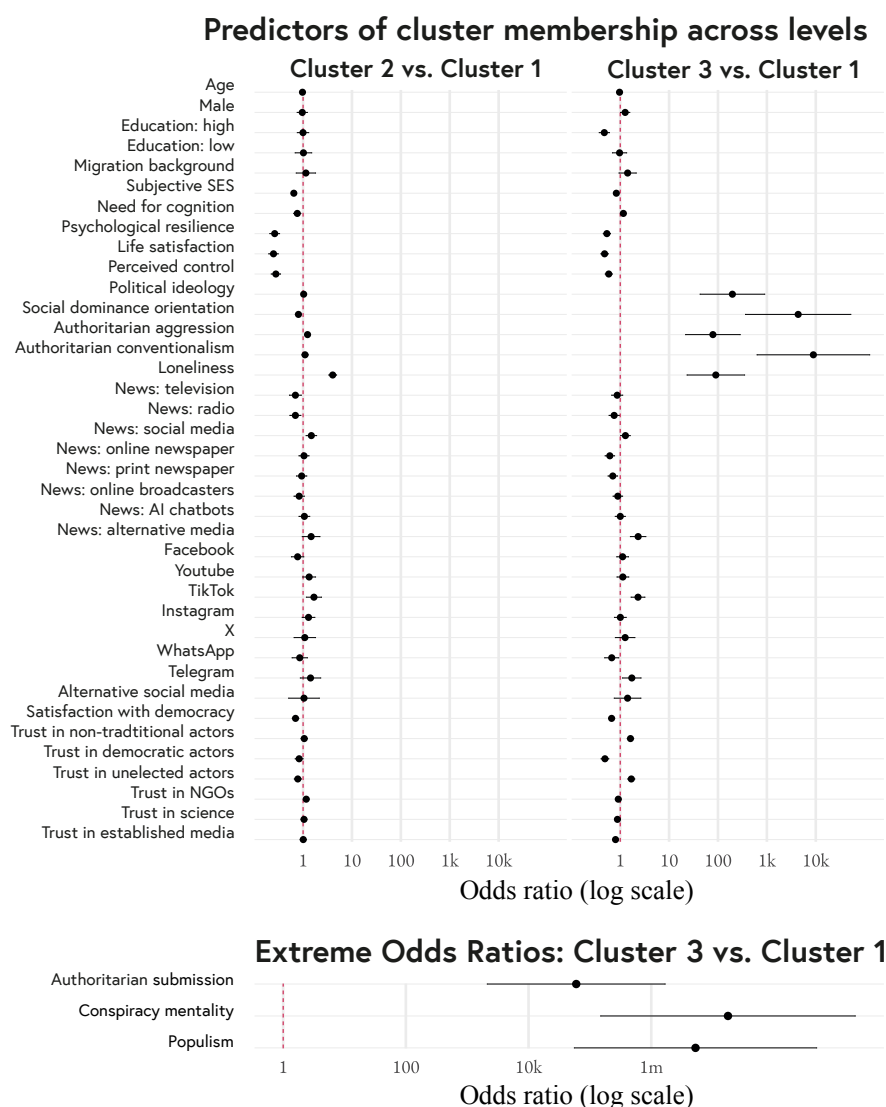


Figure 20: Predictors of cluster membership. Points represent odds ratios (ORs) from multinomial logistic regression models, and horizontal bars indicate 95% confidence intervals. The upper panels display the contrasts between Cluster 2 and Cluster 1 and between Cluster 3 and Cluster 1 on a logarithmic scale. The dashed vertical line marks the null value of OR = 1, indicating no change in the likelihood of cluster membership relative to Cluster 1. Odds ratios greater than 1 indicate a higher likelihood of assignment to Cluster 2 or Cluster 3 rather than Cluster 1, whereas values below 1 indicate a lower likelihood. Because conspiracy mentality, populism, and authoritarian submission yielded exceptionally large odds ratios in the Cluster 3 versus Cluster 1 comparison, they are shown separately in the lower panel to preserve readability. The overall model showed strong explanatory power (Nagelkerke's pseudo $R^2 = .71$), underscoring the relevance of the displayed variables for characterising the consensus clusters. SES = socioeconomic status.

Next, we tested if there are significant differences across these clusters' ability to detect misinformation, comparing these results with their perceived exposure to misinformation and counter-attitudinal content (see Figure 21). Our tests (Welch's ANOVAs) revealed significant differences across the three clusters for all outcomes. The clearest pattern emerged for misinformation discernment: the "Institutional Mainstream" cluster performed best, correctly classifying most of the headlines as false, followed by "Networked Apart," while the "right-wing authoritarian" cluster scored substantially lower than both. For perceived exposure to misinformation, both "Networked Apart" and "Right-wing Authoritarian" respondents reported higher levels than "Institutional Mainstream," with no significant difference between the two former clusters. A similar but weaker pattern appeared for perceived exposure to counter-attitudinal content, where only the "Right-wing Authoritarian" cluster differed significantly from "Institutional Mainstream".

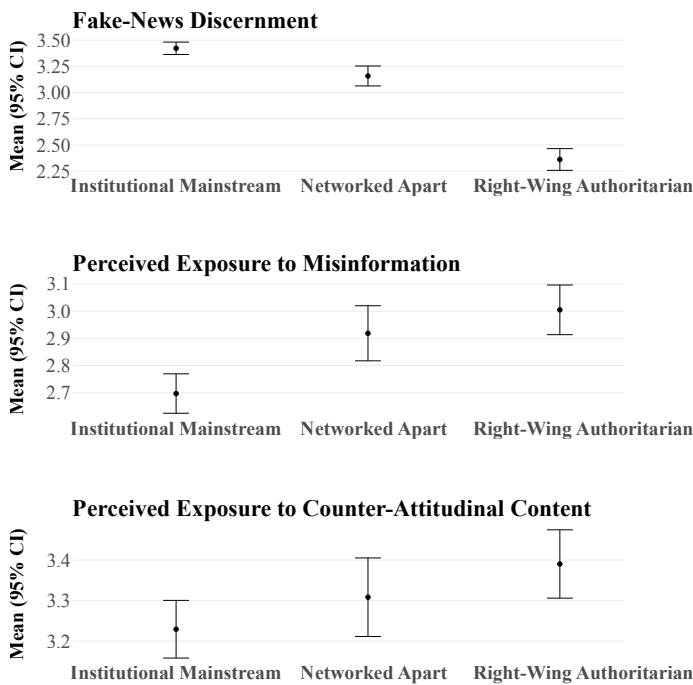


Figure 21: Mean differences in misinformation-related outcomes across clusters. Points represent cluster means and error bars indicate 95% confidence intervals. Significant differences emerged across the three clusters for fake-news discernment, perceived exposure to misinformation, and perceived exposure to counter-attitudinal content (Welch's ANOVA, $p < .05$ for all outcomes).

Taken together, these findings suggest that the “Right-wing Authoritarian” cluster was the most susceptible to misinformation. Not only did this group perform substantially worse in distinguishing false from accurate headlines, but its members also reported the highest perceived exposure to misinformation and counter-attitudinal content. By contrast, the “Institutional Mainstream” cluster combined the strongest misinformation discernment with the lowest perceived exposure to misinformation, suggesting a comparatively more resilient target group. The “Networked Apart” cluster occupied an intermediate position, scoring below the mainstream group in discernment while resembling the more vulnerable cluster in perceived exposure to misinformation. These patterns indicate that ideological orientation alone is unlikely to fully explain misinformation susceptibility. Rather, broader differences in how individuals relate to institutions, political networks, and oppositional information environments may also shape vulnerability.

6.10.6 How are citizens’ preferences related to susceptibility towards misinformation? (RQ7)

The preferences of members of the “Institutional Mainstream” cluster closely align with the sample’s broader preference for active, transparent information environments. This group shows a strong preference for receiving official communication through news media rather than AI (-29.5 pp), citizens (-20.0 pp), or the military (-10.7 pp), while expressing no significant difference between the Chancellor and news media as communicators. They also favour high transparency, prefer expert-based truth verification over legal, majoritarian, or platform-based approaches, and reject both government inaction toward misinformation and a counterspeech-only regulatory model. Indeed, this is the cluster with the strongest aversion to governmental non-engagement with misinformation. Although they prefer stronger citizen inclusion, only the contrast between high and partial integration reaches statistical significance. Overall, this cluster appears especially supportive of institutional and expert-led approaches to official communication.

The “Networked Apart” cluster’s preferences appear less permissive and institutionally demanding in comparison to cluster 1. They maintain a strong preference for having news media and the Chancellor as an official source rather than AI. They are similarly averse to allowing misinformation to go unchallenged, penalising government non-engagement (-4.4 pp). Yet, they impose fewer penalties on restrictive information environments. The cluster’s

most distinctive feature is its comparatively higher tolerance for opacity in official communication: neither low transparency nor strategic transparency significantly differs from high transparency, suggesting that openness is less central to their evaluation of information environments. Their preference for high citizen integration is statistically significant, but the magnitude of these effects remains modest. Overall, this cluster appears less committed to transparency and expert-centred governance, while still supporting some active intervention against misinformation.

The “Right-wing Authoritarian” cluster’s preferences are particularly distinct from those of the “Institutional Mainstream”, although notable overlaps remain. Members of this cluster show no statistically significant differences in their preferences for news media, the Chancellor, and even the military as official communicators, indicating a greater willingness to accept state-aligned sources alongside the news media. Their preferences regarding the foundations of truth are also weaker than those of the Institutional Mainstream: they do not significantly distinguish expert evidence from legal or majoritarian standards, and only penalise platform-based truth definitions (-4.7 pp). Nevertheless, this cluster overlaps with the Institutional Mainstream in important ways. Members favour high over low transparency, dislike scenarios in which misinformation is not addressed at all, and prefer high citizen integration over the partial integration of exemplary citizens – preferences that distinguish both clusters from their “Networked Apart” counterparts. In sum, the “Right-wing Authoritarian” cluster is distinctive for its weaker commitment to expert truth standards and its greater tolerance for information environments in which news media share their role as communicators with state sources.

Across the three clusters, we identified a lack of clear preferences regarding the tools of misinformation governance. Participants appear divided over how policymakers and social media companies should handle misinformation – whether through prebunking, debunking, or different forms of platform governance. Yet, they clearly provide a mandate for some form of intervention, as indicated by their consistent aversion to information environments in which misinformation is either left unchallenged or addressed only through counterspeech. Figure 22 visualises these differences, with the full AMCE estimates reported in Appendix B.

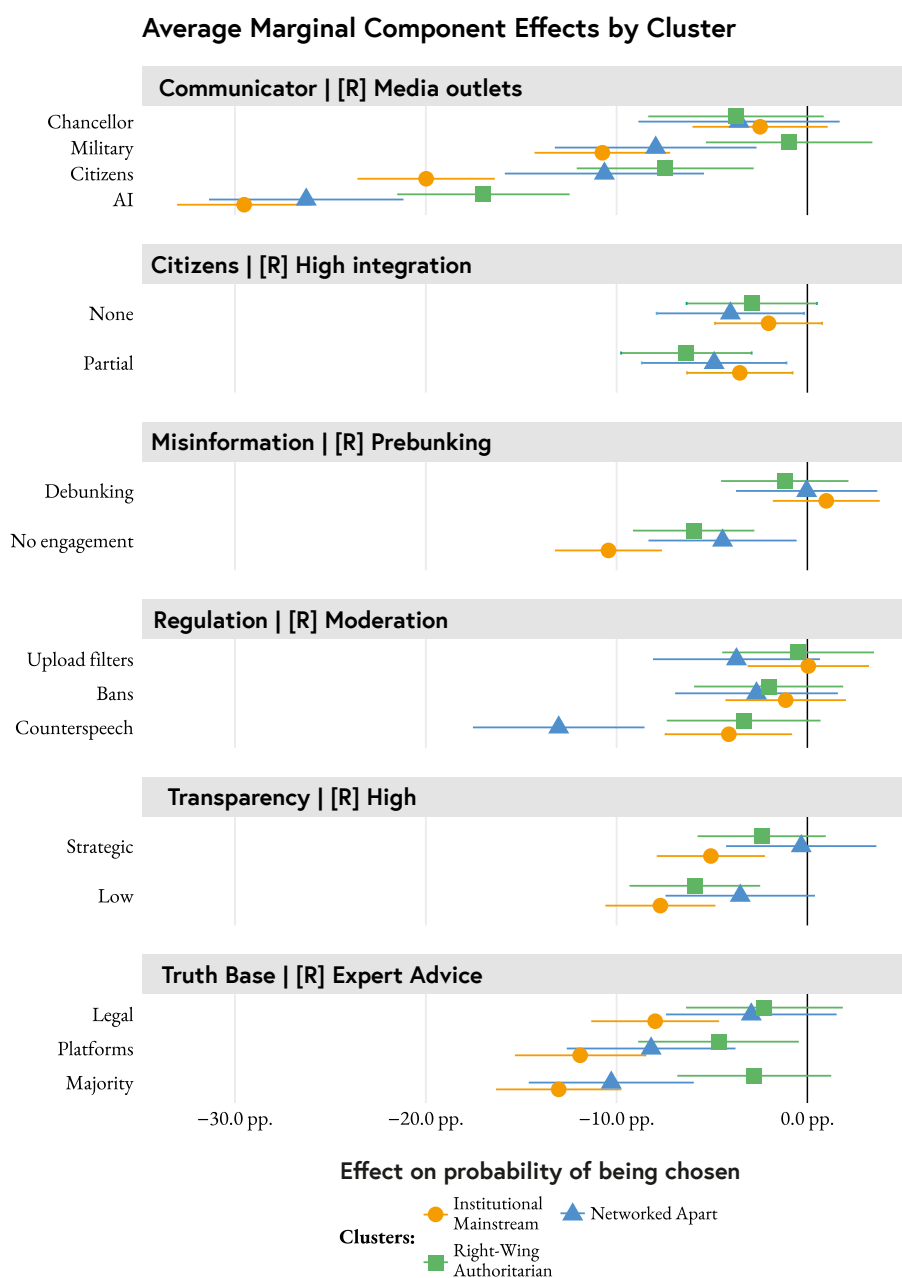


Figure 22: Average Marginal Component Effects by Cluster. Average Marginal Component Effects are expressed as percentage-point changes in the probability that a profile is chosen relative to a baseline. The reference category is written in the gray boxes and indicated by the horizontal line. Reference categories were picked based on the marginal mean estimates presented in Table 11. Horizontal lines surrounding AMCEs estimates indicate the 95% confidence intervals.

6.10.7 Summary of Core Findings Study 3

This chapter examined public preferences regarding how misinformation should be addressed, particularly under crisis conditions. We conceptualised these preferences as components of broader information environments, focusing on both official communication and responses to misinformation. Across the sample, Austrians showed a clear preference for active and transparent information environments in which news media play a central role, citizens' perspectives are meaningfully incorporated, expert advice guides truth verification, and misinformation is proactively addressed through prebunking and fact-checking. At the same time, while participants consistently rejected laissez-faire approaches in which misinformation is left unchallenged, they were less unified regarding the specific platform governance tools that should be used.

These preferences became more contingent under crisis conditions. In military and hybrid threat scenarios, participants became more willing to accept political and military spokespersons alongside news media, indicating a stronger expectation that state authorities assume a visible communicative role during crises. Importantly, however, this shift did not displace the media: political authorities were preferred as complements rather than substitutes to journalistic communication. Transparency preferences remained comparatively stable across situations, with participants consistently rejecting low-transparency environments, except for a limited tolerance of strategic discretion during physical warfare. These findings suggest that crises expand the acceptable role of the state in communication, but do not fundamentally weaken expectations of transparency and media involvement. Notably, we find no public support for the integration of official communication in AI chatbots serving as a primary source of official communications across all situations studied.

The chapter also demonstrated that these preferences are not evenly distributed across the population. The cluster analysis identified three distinct clusters in the population – “Institutional Mainstream,” “Networked Apart,” and “Right-wing Authoritarian” – that differed in their support for expert-led governance, transparency, and the role of institutional actors in defining truth. Yet, across these groups, one commonality remained: a clear preference for some form of intervention against misinformation, even when views diverged on the exact tools or epistemic authorities involved.

Finally, we showed that these orientations are meaningfully related to misinformation susceptibility. The “Institutional Mainstream” cluster performed best in distinguishing false from accurate headlines and reported the lowest perceived exposure to misinformation, suggesting a comparatively more resilient information profile. By contrast, the “Right-wing Authoritarian” cluster performed substantially worse on discernment tasks and reported the highest perceived exposure to misinformation and counter-attitudinal content. These findings demonstrate that susceptibility is shaped not simply by ideology, but by how individuals relate to institutions, political networks, and oppositional information environments.

6.11 General Discussion of Limitations and Directions for Future Research

Our studies had several limitations that should be considered when interpreting the results or generalising them to other contexts. First, regarding the collected data and, consequently, the completeness and generalisability of our analyses, Studies 1 and 2 focused heavily on the English-language academic literature. While English remains the dominant language in academia, future research incorporating other languages is needed to provide a more nuanced understanding of specific contexts.

Moreover, while our scoping review provided insights into the current state of academically scrutinised evidence, we only included published work. Research has shown that research findings that are statistically significant are published more easily which can lead to an overestimation of relationships between, for example, specific susceptibility factors and the response to misinformation (so-called publication bias, e.g., Sutton et al., 2000). Future research implementing meta-analytical techniques that also include non-published findings is thus needed for more robust insights. Another limitation affecting the generalisability of our findings is that Study 1 focused only on the abstracts of the academic texts due to the temporal scope of this project. Future research examining full papers in detail could provide additional insights.

In Study 2, our examination of state actor responses to recent crises adopted a broader lens across contexts, which required a narrative case-study approach. More systematic comparisons across contexts and countries are needed to

understand which country-level factors and crisis characteristics are associated with different communication strategies and misinformation countermeasures implemented by state actors. Study 3 focused on the Austrian population aged 18 years and older. Future research replicating our findings in other contexts is needed to better understand citizens' preferences across different socio-political and cultural environments and to assess how far our results generalise beyond Austria, especially to contexts in the Global South or under authoritarian regimes.

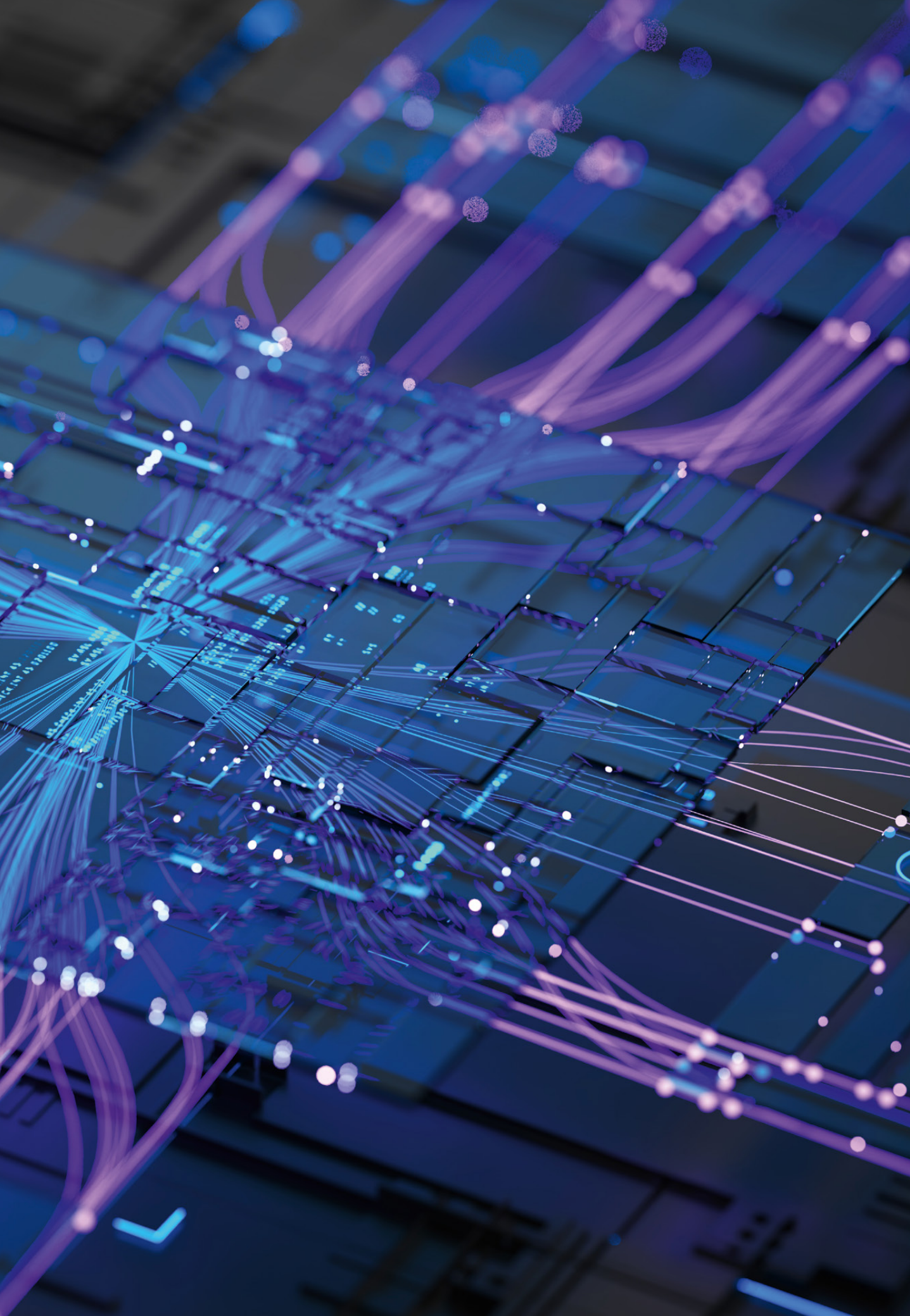
With regard to our methodological approach, and thus the reliability of our findings, the coders in Study 1 agreed in 81% of the papers on whether the papers suggested interventions. However, once corrected for chance, agreement declined to 62 % (as measured by Brennan-Predigers κ). Following Landis and Koch (1977), this still falls within the range of substantial agreement (0.61–0.80, p. 165), but future research should nonetheless aim to further improve the reliability of this category. Our detailed full paper analyses in Study 2, however, confirmed that all studies identified as discussing interventions did actually discuss interventions in some form.

In Study 3, several well-established measures showed unsatisfactory levels of internal consistency or failed to confirm the factor structures expected from the literature. While we adapted the scales accordingly and relied on measures that reliably captured the central dimensions of the constructs under study, future research is needed to examine factor structures and measurement invariance in the Austrian context in greater detail and, if necessary, adapt established scales to the current socio-cultural environment. Critically, we also observed questionable internal consistency for the MIST test (Cronbach's $\alpha = .64$), which falls below the desirable threshold ($\alpha > .70$) for measuring participants' truth- and misinformation-discernment skills. While we primarily used this measure to validate the selection of susceptibility indicators included in Study 3, future research should replicate these exploratory findings using alternative measures adapted to and validated in the Austrian context.

Nonetheless, our studies provide meaningful starting points for future research, for example on the complexity of misinformation dynamics, state actors measures to counter misinformation across countries, citizens' susceptibility to misinformation taking variables related to different levels into account, or citizens' perspectives on desirable information environments and the political and communicative measures intended to shape them.

7

Conclusion



Conclusion

Is the concept of sovereignty viable in the digital realm, and what aspects does it encompass? (Chapter 1)

Key ideas

Sovereignty is a fundamental political concept with a long and rich intellectual tradition. Over time, a range of thinkers, including Niccolò Machiavelli, Jean-Jacques Rousseau, and Immanuel Kant, have developed a variety of reflections on the concept and its different dimensions. Considered in historical perspective, these developments revolve around a central idea: the capacity for autonomous decision-making over oneself or others within the framework of a political community. In this context, the concept of digital sovereignty is highly relevant. It constitutes neither the opposite of nor a replacement for the classical understanding of sovereignty; rather, it represents a further and ongoing evolution of the concept itself.

Data and information as well as digital infrastructure constitute essential factors in the context of hybrid conflicts and hybrid warfare. In this sense, Digital Sovereignty can be conceptualised as entailing a so-called “*Oikos*” and a “*Logos*” dimension. The *Oikos* dimension thereby refers to the ability of individuals, organisations or states to use, shape and control material digital technologies, data and infrastructure on which they depend. The *Logos* dimension encompasses the immaterial will, the ideas, the narratives, the fears and the motivations that are formed within the information environment. In the *Oikos* dimension, the meaning of Digital Sovereignty is the capacity to independently develop, regulate and protect critical digital infrastructure and technologies. In the *Logos* dimension, Digital Sovereignty means sovereignty of will and knowledge; that is, the ability to access content independently and without conscious manipulation. Thus, the conditions for individual and collective capacity to act are located in the *Logos* dimension.

Implications

From a historical-developmental perspective, the concept of sovereignty has been subject to continual transformation. In this regard, it reflects both the context-specific self-understanding of a given era and the capacity of political actors to assert influence and shape prevailing norms and convictions.

Consequently, developments in the digital age do not constitute a rupture with this tradition; rather, they require active engagement with their specific characteristics and their political classification. Sovereignty, in particular, provides a suitable conceptual framework for a reason-based relationship between the state, its liberty-enabling institutions, and its citizens. Within a democratic polity oriented towards the common good, all elements constituting the information environment therefore bear considerable responsibility for the defence of freedom. This includes the issue of the infrastructural dimension of the information environment, which should translate to measures for greater autonomy in critical areas.

How do we define “information”, and how do global actors such as the EU, the US, NATO, China, and Russia differ in their conceptions of the information space? (Chapter 2)

Key ideas

Information is an ambiguous term with material, symbolic and social aspects. From a political perspective, information is a means of communication as well as a powerful resource. The analysis of political strategies and security policy concepts that were examined in this study illustrates clear differences in the understanding of the term “information environment” among the big actors of the EU, the US, NATO, Russia and China. NATO locates the information environment within military operations and views it as a strategic operating environment. The US define the information environment primarily as a domain of military operations; the EU considers it mainly a malleable democratic sphere to be protected; China understands the information environment not merely as technological infrastructure or a space for communication but as an integral element of state sovereignty and societal stability; Russia differentiates between the inner information space, which is primarily seen as a protected space, and, following a civilisational, linguistic-cultural logic, an outer information space as a sphere of influence.

Implications

Adequate defence policy measures thus require consideration and understanding of the different conceptualisations of the information environment.

How should cognitive warfare be assessed from a legal perspective? (Chapter 3)

Key ideas

From a historical viewpoint, information operations are not a novelty, but the continuation of an ancient phenomenon. However, modern technologies are altering the scale, speed and consequences of manipulative communication techniques to a high degree.

Cognitive warfare affects numerous legal norms of international law and human rights law. Generally, due to its technology-agnostic nature, existing international law is flexible enough to regulate non-kinetic acts. Nevertheless, the attribution remains a considerable practical obstacle as anonymous networks and encrypted channels impede obtaining evidence of state responsibility.

Cognitive warfare can infringe upon the principle of non-intervention when it aims to coercively influence public opinion, elections or governmental decisions.

Regarding human rights, the following risks arise:

- The right to freedom of speech is at risk when information manipulation and the strategic use of psychological operations take place.
- Cognitive warfare is often based on the collection, analysis and use of personal data without the knowledge or consent of those affected, which contradicts the right to (data) privacy.
- The manipulation of perception and emotions often intends to provoke societal divisions, fear and mistrust. As the potential for mental harm is high, it is possible that the right to bodily and mental integrity is violated.
- Influencing democratic election processes and decision-making infringes upon the right to political participation.

Implications

Currently the best way for states to strengthen legal certainty and maintain the protection afforded by the law in view of these new threats and challenges is to commit to and apply the existing legal framework in good faith. This includes, among other things, that states continue to engage in in-depth discussions with other states and stakeholders on how the existing law applies to activities in the realm of cognitive warfare and to build upon existing

convergences. In addition, states should consider taking further appropriate (non-legal) measures on a national, regional or international level, such as the exchange of good practices, capacity-building, political commitments and educational programmes.

How can democratic states maintain a balance between security interests and civil liberties? What undesirable or unintended effects may arise from the deployment of military forces, how can these be counteracted, and how can this risk be managed? (Chapter 4)

Key ideas

Digital interconnectedness has transformed the information space into a central domain of security policy, wherein state protective measures, societal resilience, and individual liberties are inextricably linked. An analysis of European security strategies reveals an evolving paradigm: the information space is increasingly no longer viewed merely as a sphere of civilian communication, but rather as a realm of political competition, strategic influence operations, and hybrid threats.

Concurrently, questions regarding the legitimacy of military interventions within hybrid, civil-military contexts in democratic societies have moved into sharp focus. Central to this discourse is the concept of “securitisation”: the political tendency to frame specific domains as existential security issues, thereby shifting their governance from the realm of standard political negotiation to a state of exception characterised by expanded powers of intervention.

Such undesirable effects are not inevitable. To maintain the critical balance between security and freedom, the structural, normative, and institutional frameworks governing military action in civilian contexts must be clearly delineated. Essential countermeasures include:

- The reinforcement of civil society oversight mechanisms;
- A strict separation between military and civilian spheres of responsibility;
- Ethically reflective training standards; and
- Inter-institutional cooperation.

Democratic resilience, therefore, arises not primarily from controlling security architectures, but from education, media literacy, pluralistic debate, and the rule of law.

Implications

The clear separation of tasks and responsibilities is increasingly difficult in practice, given the complexity and intense overlap of these fields. Consequently, beyond the requisite legal regulations, it is highly necessary to establish appropriate coordination formats designed to negotiate borderline cases through discourse. In turn, this depends on a degree of shared strategic culture among the relevant actors.

How can holistic governance in the information space strengthen protection against influence operations, and what contribution can a Ministry of Defence make in this regard? What instruments are available to a Ministry of Defence, and which capabilities and tasks are required? Which target groups should the Ministry of Defence address? (Chapter 5)

Key ideas

Hybrid threats within the information environment can be described using the Hybridity Blizzard model. Hybrid threats resemble a chaotic “blizzard” in which attacks originate simultaneously from multiple directions and cannot be clearly attributed. In the short term, the focus lies on tactical responses aimed at countering immediate attacks; in the long term, however, it concerns the systematic development of resilience and the minimisation of vulnerabilities through a coordinated whole-of-society approach.

The information environment extends beyond the jurisdictional and operational remit of any single ministry and encompasses political leadership, civilian authorities, regulatory bodies, the private sector, civil society, the media, educational institutions, and the broader public. Against this backdrop, a ministry of defence neither can nor should seek to assume functions that belong to civilian authorities, civil society, or the political leadership. Rather, its role is to contribute its specific capabilities – including intelligence, strategic analysis, military planning, and strategic communication – in support of a broader constitutionally grounded framework of democratic governance. The

ministry of defence operates on three principal levels: as an individual actor, in cooperation with national partners, and in cooperation with international partners. Ultimately, effective cognitive security depends on coordinated whole-of-government as well as whole-of-society efforts. In this context, clearly delineated responsibilities and robust inter-agency coordination constitute essential factors.

As hybrid threats are multifaceted and synergetic in nature, state countermeasures require an intentional combination of a number of areas – including local and regional authorities, cultural and educational institutions, healthcare, energy and public transport, financial institutions, regulatory authorities, research and journalism – many of which are not usually considered part of the security apparatus. Here a Ministry of Defence might receive the task of consolidating threat assessments and disseminating them among relevant actors in a targeted manner.

Implications

The tension inherent in a whole-of-government approach to addressing hybrid threats is characterised by a high degree of fluidity in competencies and responsibilities and is frequently accompanied by questions concerning redundancies and checks and balances. Particularly within the information environment, a careful allocation of responsibilities that is mindful of overlaps may be required, as influence over the information environment has a direct impact on people's lived reality and is of central importance for the strengthening of resilience.

Which target groups are vulnerable to influence, and why? What constitutes effective strategic communication? What are protective factors against manipulation in the information space? (Chapter 6)

Key ideas

The literature shows that susceptibility to misinformation is rooted in individual personality characteristics, media consumption patterns, political and ideological views as well as trust in democratic institutions and other actors. Psychological strain, social isolation and digital dependency are further additional relevant factors. To prevent distrust and deceptions, it is necessary to address the systemic roots of distrust in the population. Resilience requires transparent state communication and giving room to the perspectives of citizens. Trust emerges not simply through the content of a state's strategic

communication but also through the perceived openness, inclusion and legitimacy of the overall information environment.

The empirical findings demonstrate a clear general public preference for interventions against misinformation, although opinions vary in their specifics. Austrians prefer transparency in state communication, the inclusion of citizens' perspectives as valuable sources of information, a strong role for independent media and veracity judgments based on expert advice and evidence. They overwhelmingly reject communication via AI chatbots.

Strategic communication should be target-group-specific, context-sensitive and nuanced. Citizens have sophisticated, context-dependent preferences regarding transparency, the involvement of experts, crisis communication and proportionality of interventions. Democratically legitimate governance must not only address citizens but also incorporate their preferences in the conception, implementation and evaluation of countermeasures.

Implications

From a defence policy perspective, strategic communication is a vital instrument in addressing these challenges. However, it must be embedded in whole-of-state ambitions and strategies for tackling hybrid threats and for building resilience.

Summary

In conclusion, Digital Sovereignty constitutes a whole-of-state and whole-of-society endeavour that vastly exceeds the scope of a single ministry. As a uniquely positioned actor, the MoD plays a central role by applying its core competencies – intelligence, strategic analysis and military planning – in a broader context. Achieving cognitive security requires more than isolated measures; it demands a synthesis of civilian education, coordinated national and international collaboration, and nuanced, target-group-specific strategic communication. The relevance of Sovereignty remains undisputed during all (political-historical) developmental steps and applicable to the information environment. Hence, it is imperative to constantly renegotiate and adapt the concept of Sovereignty and its associated issues in the context of autonomy, international law and socio-economic considerations. This can only take place in a democratic dialogue carried by mutual respect.

Appendix: André K. Rodarte and Lena Frischlich: Preservation and Promotion of Democratic Resilience – A Scientific Basis for Policy Recommendations

Appendix A – Table 1A: Search Terms.

	Strings
Content	“Misinform*” OR “disinform*” OR “distorted inform*” OR “falsehood*” OR “lie*” OR “conspiracy*” OR “deep fake*” OR “deepfake*” OR “infodemic” OR “viral misinformation” OR “viral disinformation” OR “information warfare” OR “big lie” OR “psyops” OR “psychological operation*” OR “propaganda” OR “foreign inferences” OR “hybrid war*” OR “information war*” OR “psychological war*” OR “polluted information” OR “information disorder” OR “fake news” OR “alternative news” OR “hyperpartisan news” OR “untrustworthy news” OR “junk news”
Channel	AND “artificial intellig*” OR “social bot*” OR “LLM*” OR “ChatGPT” OR “DeepSeek” OR “Gemini” OR “Groat” OR “Perplexity” OR “Llama AI” OR “Social media” OR “digital social network*” OR “social web” OR “social network* site*” OR “social platform” OR “online social media” OR “online social network*” OR “technological affordance*” OR “across platform*” OR “cross-platform*” OR “dark social” OR “alternative social” OR “fringe communit*” OR “microblog*” OR “twitter” OR “Gab” OR “Parler” OR “Truth Social” OR “Gettr” OR “social network site” OR “Facebook” OR “VKontakte” OR “Tumblr” OR “Instagram” OR “Pinterest” OR “9gag” OR “YouTube” OR “TikTok” OR “Rumble” OR “Bitchute” OR “Twitch” OR “Discord” OR “Steam” OR “4Chan” OR “8kun” OR “Reddit” OR “instant messag*” OR “Snapchat” OR “WhatsApp” OR “Signal messenger” OR “Threema” OR “Telegram” OR “Linkedin” OR “Wechat” OR “Weibo” OR “Tencent QQ” OR “Douyin”
Susceptibility	AND “susceptib*” OR “risk factor” OR “vulnerab*” OR “effect” OR “caus*” OR “cross-country” OR “national diff*” OR “comparativ” OR “crisis”
Multi-level effects	AND “credib*” OR “belief*” OR “believ*” OR “misperc*” OR “persua*” OR “accurac*” OR “percep*” OR “truth percep*” OR “truth discern*” OR “shar*” OR “lik*” OR “commen*” OR “spread*” OR “behav*” OR “radicali*” OR “normalization” OR “mainstreaming” OR “authoritar*” OR “illiberal”

Table 2A: Screening Criteria.

Inclusion criteria	Exclusion criteria
English or German	Other languages
Contains all four search blocks in the title or abstract	Does not contain all four search blocks in the title or abstract
Articles containing at least one empirical study (if unclear based on the abstract, check in full text)	Addresses either block only illustrative (e.g., online experiments using the example of misinformation)
Peer-reviewed articles	Abstract not accessible (e.g., Conference proceedings, where only the book, but not single articles are accessible)
Preprints	Purely theoretical articles
Online accessible book chapters and dissertations	Editorials and opinion pieces
	Pre-registrations
	Conference presentations without full text
	Pure focus on detection tools (e.g., classifiers)
	Other scoping reviews

Figure 1A: Topic distributions.

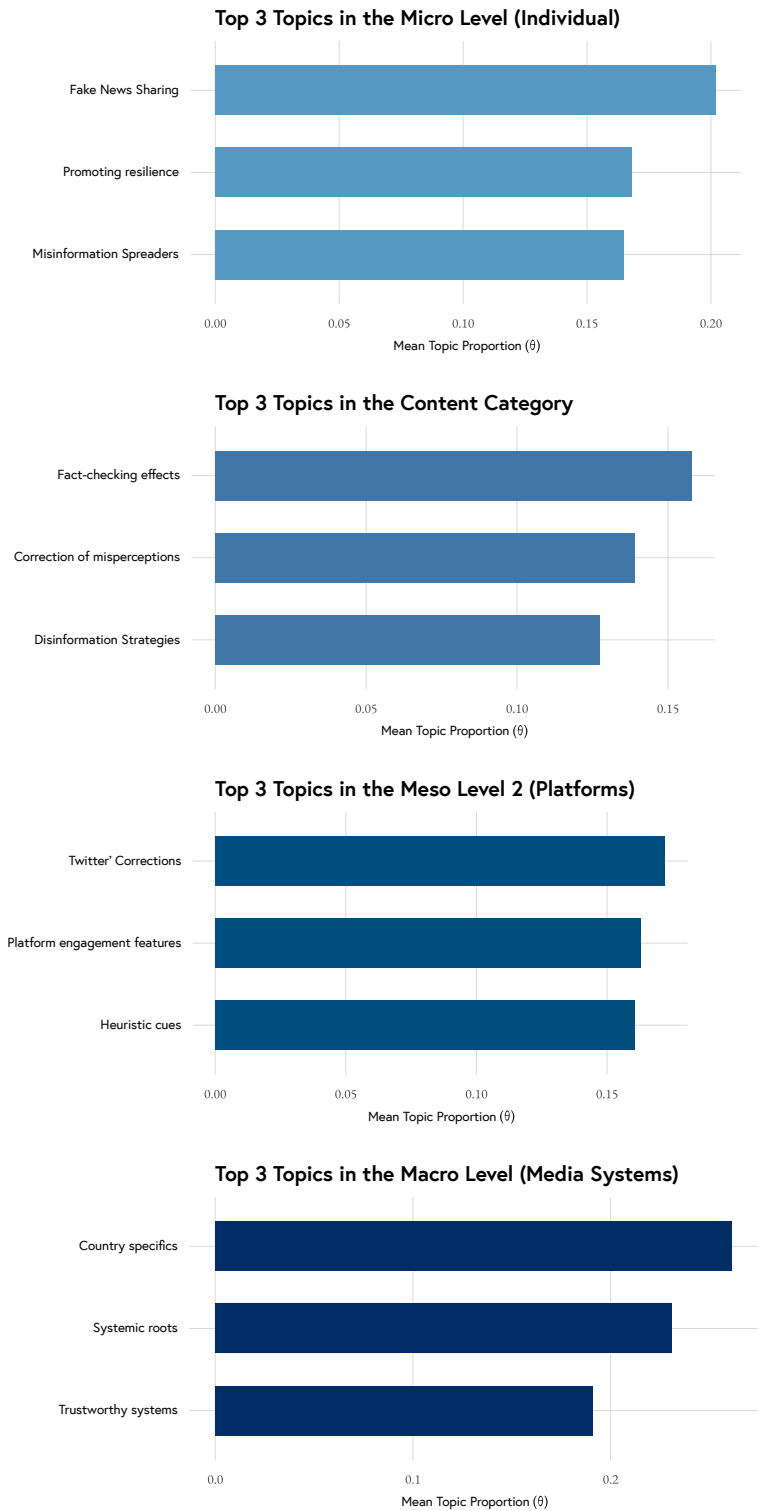


Figure 2A: K-Evaluations for STM.

The number of topics was determined by inspecting the diagnostic plots of held-out likelihood, residuals, semantic coherence, and the lower bound. For the micro-level dataset, both held-out likelihood and semantic coherence peaked at $K = 6$, suggesting that six topics offered the best balance between model fit and interpretability. In the content dataset, performance improved until $K^{\circ} = 8$, where coherence remained high and additional topics added little explanatory value. The meso-level 2 dataset also showed stable results around $K = 6-7$; however, to prioritise clearer thematic consistency and higher semantic coherence, $K^{\circ} = 6$ was chosen. Finally, for the systems dataset, coherence and likelihood measures both indicated that a more compact model with $K^{\circ} = 5$ provided the most adequate number of topics.

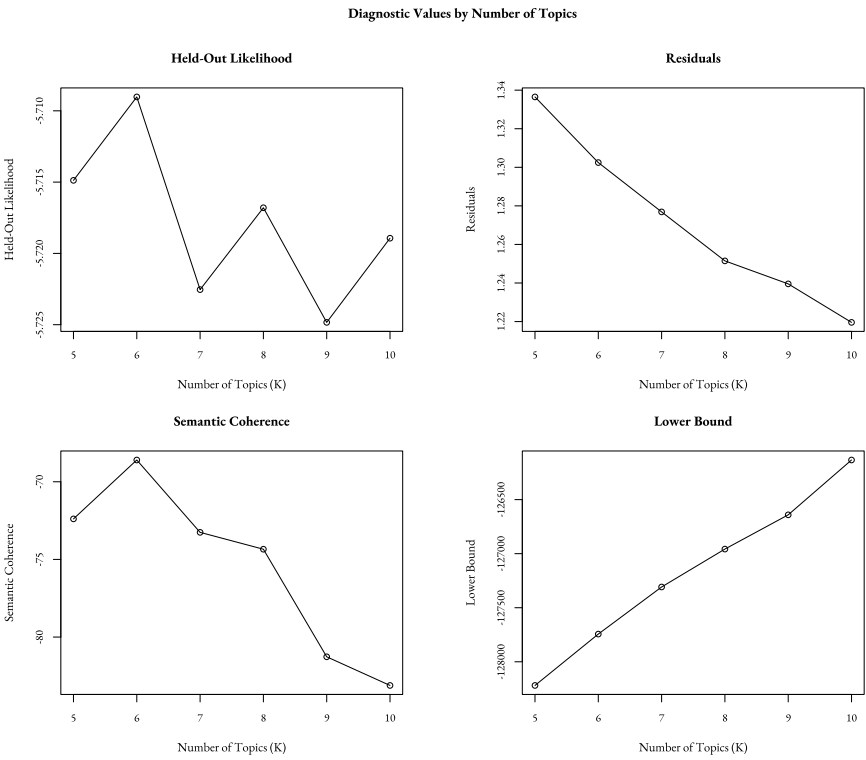


Figure 2Aa: Micro-level.

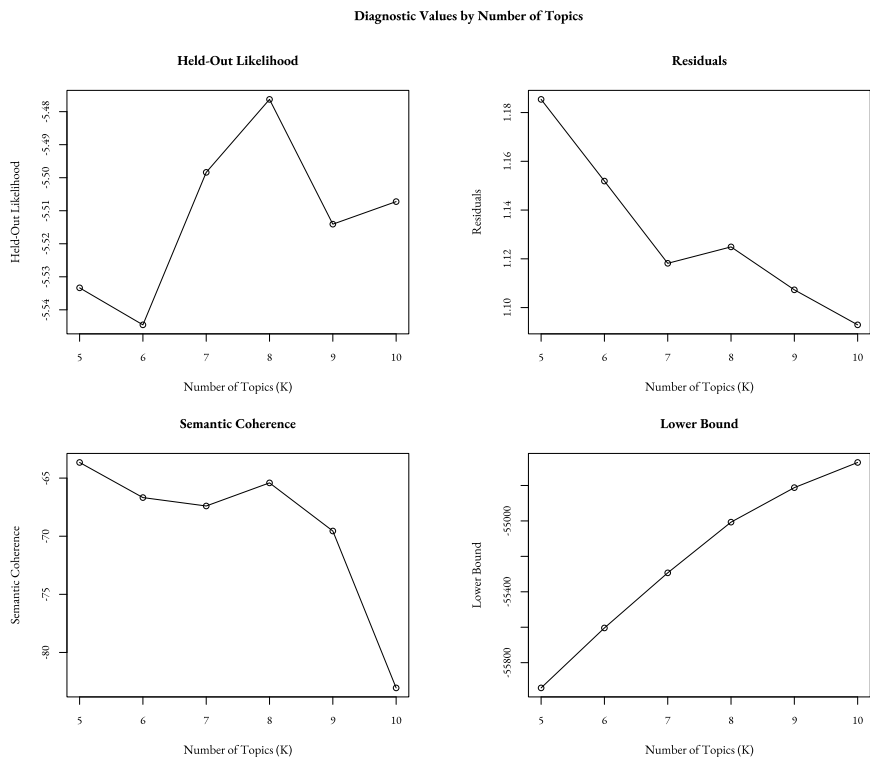


Figure 2Ab: Content level.

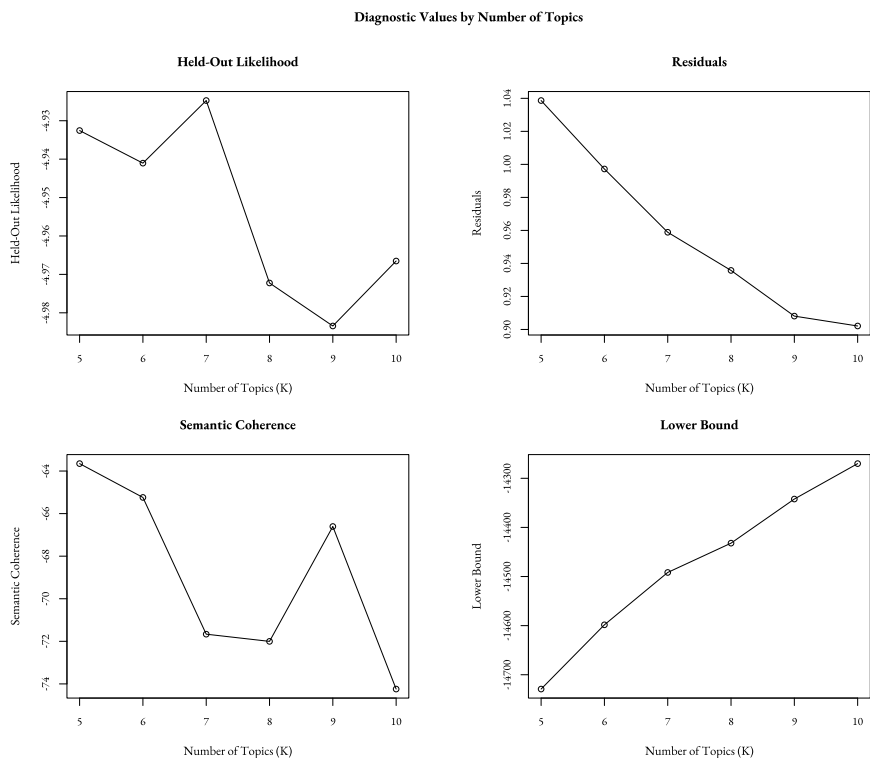


Figure 2Ac: Meso-level II (platforms).

Diagnostic Values by Number of Topics

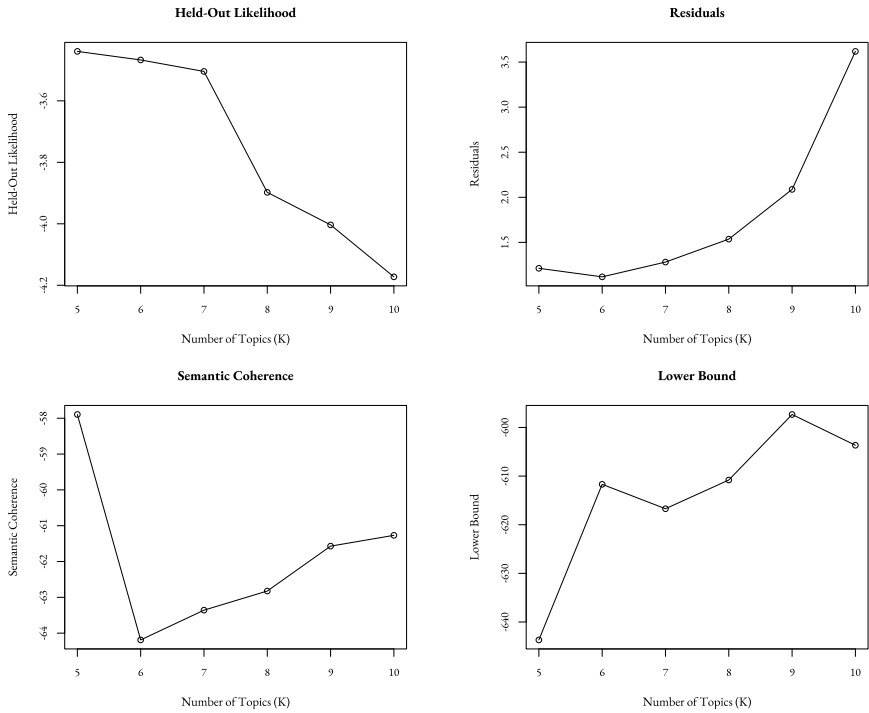


Figure 2Ad: Macro-level.

Figure 3A: GPT-Prompt

```
sanitize_prompt <- function(base_instruction) {  
  paste(  
    base_instruction,  
    Extract ONLY the exact sentences copied verbatim from the abstract  
    referring to such information,  
    Do NOT add headings, introductions, markdown, or explanations.,  
    Return only the copied sentences as plain text separated by spaces.  
  )  
}  
  
prompts <- list(  
  communicator = sanitize_prompt(  
    This abstract presents information about how communicators may  
    facilitate the spread of or belief in misinformation (a communicator  
    is either the source or recipients of these messages).  
  ),  
  content = sanitize_prompt(  
    This abstract presents information about how the content of certain  
    messages may facilitate the spread of or belief in misinformation.  
    Variations in fact-checking labels can be included here.  
  ),  
  group = sanitize_prompt(  
    This abstract presents information about how group factors may  
    facilitate the spread of or belief in misinformation.  
  ),  
  organization = sanitize_prompt(  
    This abstract presents information about how organizations may  
    facilitate the spread of or belief in misinformation.  
  ),  
  platform = sanitize_prompt(  
    This abstract presents information about how social media platforms  
    may facilitate the spread of or belief in misinformation (such  
    information may concern content moderation or anonymity, for  
    example).  
  ),  
  systems = sanitize_prompt(  
    This abstract presents information about the characteristics of a  
    country or media system facilitate the spread of or belief in  
    misinformation.  
  ),  
  interventions = sanitize_prompt(  
    This abstract presents information about interventions that can  
    diminish the spread of or belief in misinformation.  
  )  
)
```

Table 3A

	Comparison (Reference in parentheses)	Scenario	AMCE (SE)	95% CI	p
Source of the official communication	AI (vs. Media outlets)	Average week	-27 pp (0.02)	[-32, -23]	< .001
	Chancellor (vs. Media outlets)	Average week	-9 pp (0.02)	[-14, -4]	< .001
	Citizens (vs. Media outlets)	Average week	-13 pp (0.02)	[-18, -8]	< .001
	Military (vs. Media outlets)	Average week	-18 pp (0.02)	[-23, -13]	< .001
	AI (vs. Media outlets)	Medical crisis	-23 pp (0.02)	[-28, -19]	< .001
	Chancellor (vs. Media outlets)	Medical crisis	-2 pp (0.03)	[-6, 3]	.55
	Citizens (vs. Media outlets)	Medical crisis	-12 pp (0.02)	[-17, -7]	< .001
	Military (vs. Media outlets)	Medical crisis	-8 pp (0.02)	[-13, -3]	< .001
	AI (vs. Media outlets)	Hybrid crisis	-22 pp (0.02)	[-27, -17]	< .001
	Chancellor (vs. Media outlets)	Hybrid crisis	0 pp (0.02)	[-5, 4]	.85
	Citizens (vs. Media outlets)	Hybrid crisis	-13 pp (0.02)	[-18, -8]	< .001
	Military (vs. Media outlets)	Hybrid crisis	1 pp (0.02)	[-3, 6]	.55
	AI (vs. Media outlets)	Military crisis	-27 pp (0.02)	[-31, -22]	< .001
	Chancellor (vs. Media outlets)	Military crisis	-1 pp (0.02)	[-6, 4]	.73
	Citizens (vs. Media outlets)	Military crisis	-17 pp (0.02)	[-22, -12]	< .001
	Military (vs. Media outlets)	Military crisis	-2 pp (0.02)	[-7, 2]	.34

	Comparison (Reference in parentheses)	Scenario	AMCE (SE)	95% CI	p
Integration of citizens	No integration (vs. High)	Average week	-6 pp (0.02)	[-10, -2]	.001
	Partial integration (vs. High)	Average week	-7 pp (0.02)	[-11, -3]	< .001
	No integration (vs. High integration)	Medical crisis	-1 pp (0.02)	[-5, 3]	.65
	Partial integration (vs. High)	Medical crisis	-2 pp (0.02)	[-5, 2]	.41
	No integration (vs. High)	Hybrid crisis	-2 pp (0.02)	[-5, 2]	.41
	Partial integration (vs. High)	Hybrid crisis	-4 pp (0.02)	[-8, 0]	.035
	No integration (vs. High)	Military crisis	-2 pp (0.02)	[-6, 2]	.33
	Partial integration (vs. High)	Military crisis	-5 pp (0.02)	[-9, -2]	.004
Handling of misinformation	Debunking (vs. Prebunking)	Average week	-2 pp (0.02)	[-5, 2]	.43
	No engagement (vs. Prebunking)	Average week	-11 pp (0.02)	[-14, -7]	< .001
	Debunking (vs. Prebunking)	Medical crisis	2 pp (0.02)	[-2, 5]	.42
	No engagement (vs. Prebunking)	Medical crisis	-7 pp (0.02)	[-11, -4]	< .001
	Debunking (vs. Prebunking)	Hybrid crisis	1 pp (0.02)	[-3, 5]	.66
	No engagement (vs. Prebunking)	Hybrid crisis	-7 pp (0.02)	[-11, -3]	< .001
	Debunking (vs. Prebunking)	Military crisis	-1 pp (0.02)	[-4, 3]	.71
	No engagement (vs. Prebunking)	Military crisis	-6 pp (0.02)	[-9, -2]	.002

	Comparison (Reference in parentheses)	Scenario	AMCE (SE)	95% CI	p
Regulation	Bans (vs. Moderation)	Average week	-2 pp (0.02)	[-6, 2]	.36
	Counterspeech (vs. Moderation)	Average week	-9 pp (0.02)	[-13, -5]	< .001
	Upload filter (vs. Moderation)	Average week	-1 pp (0.02)	[-5, 3]	.60
	Bans (vs. Moderation)	Medical crisis	-3 pp (0.02)	[-8, 1]	.12
	Counterspeech (vs. Moderation)	Medical crisis	-5 pp (0.02)	[-10, -1]	.015
	Upload filter (vs. Moderation)	Medical crisis	0 pp (0.02)	[-4, 5]	.91
	Bans (vs. Moderation)	Hybrid crisis	1 pp (0.02)	[-3, 6]	.56
	Counterspeech (vs. Moderation)	Hybrid crisis	-1 pp (0.02)	[-6, 3]	.56
	Upload filter (vs. Moderation)	Hybrid crisis	-1 pp (0.02)	[-6, 3]	.54
	Bans (vs. Moderation)	Military crisis	-4 pp (0.02)	[-8, 0]	.06
	Counterspeech (vs. Moderation)	Military crisis	-9 pp (0.02)	[-13, -5]	< .001
	Upload filter (vs. Moderation)	Military crisis	-2 pp (0.02)	[-7, 2]	.29
Transparency	Low transparency (vs. High transparency)	Average week	-6 pp (0.02)	[-9, -2]	.004
	Strategic transparency (vs. High)	Average week	-2 pp (0.02)	[-6, 2]	.32
	Low transparency (vs. High)	Medical crisis	-8 pp (0.02)	[-11, -4]	< .001
	Strategic transparency (vs. High)	Medical crisis	-6 pp (0.02)	[-9, -2]	.002
	Low transparency (vs. High)	Hybrid crisis	-6 pp (0.02)	[-10, -3]	< .001
	Strategic transparency (vs. High)	Hybrid crisis	-5 pp (0.02)	[-9, -1]	.007
	Low transparency (vs. High transparency)	Military crisis	-6 pp (0.02)	[-10, -2]	.001
	Strategic transparency (vs. High)	Military crisis	0 pp (0.02)	[-4, 4]	.96

	Comparison (Reference in parentheses)	Scenario	AMCE (SE)	95% CI	p
Truth Basis	Legal basis (vs. Evidence)	Average week	-7 pp (0.02)	[-12, -3]	< .001
	Majority opinion (vs. Evidence)	Average week	-11 pp (0.02)	[-16, -7]	< .001
	Platforms (vs. Evidence)	Average week	-12 pp (0.02)	[-16, -8]	< .001
	Legal basis (vs. Evidence)	Medical crisis	-5 pp (0.02)	[-10, -1]	.019
	Majority opinion (vs. Evidence)	Medical crisis	-12 pp (0.02)	[-16, -7]	< .001
	Platforms (vs. Evidence)	Medical crisis	-10 pp (0.02)	[-14, -5]	< .001
	Legal basis (vs. Evidence)	Hybrid crisis	-5 pp (0.02)	[-9, 0]	.036
	Majority opinion (vs. Evidence)	Hybrid crisis	-6 pp (0.02)	[-11, -2]	.003
	Platforms (vs. Evidence)	Hybrid crisis	-8 pp (0.02)	[-13, -4]	< .001
	Legal basis (vs. Evidence)	Military crisis	-2 pp (0.02)	[-6, 2]	.41
	Majority opinion (vs. Evidence)	Military crisis	-7 pp (0.02)	[-11, -3]	.001
	Platforms (vs. Evidence)	Military crisis	-5 pp (0.02)	[-9, 0]	0.41

Table 3A: Average Marginal Component Effects (AMCEs) across scenarios.

Notes. Entries report Average Marginal Component Effects (AMCEs) expressed as percentage-point changes in the probability that a profile is chosen. Estimates are calculated relative to the reference category listed in parentheses for each attribute, which were picked based on the marginal mean estimates presented in Table 11. Confidence intervals are reported at the 95% level.

Appendix B – Cluster analyses.

The cluster analysis required standardising all variables to be measured on a scale. This prevents variables with a very large range (e.g., age) from becoming dominant compared to variables with a very low range (e.g., a 1–5 scale). We also transformed all categorical variables to dummy-coded variables ranging from 0 to 1, (e.g., male gender = 1, means that person identified as male, 0 =

that person did not identify as male). All analyses used hierarchical clustering. This approach is well-suited for explorative analyses such as in the case of this study. Following Kaufman & Rousseeuw (2009), we identified the best cluster algorithm and optimal cluster solution by comparing two different algorithms: divisive clustering (“Diana”), which starts from one large cluster and then splits into smaller clusters; and agglomerative clustering (“Agnes”), which starts with small clusters and aggregates them to larger ones. Hierarchical clustering is based on distance matrices that express how dissimilar participants are to each other. We used euclidean distances for scale-type variables and gower distance matrixes for categorical variables and variables with mixed types. For the final consensus cluster, we build a dissimilarity matrix based on level cluster solutions (Ghosh & Acharya, 2011). This section reports the diagnosis plots for all clustering solutions.

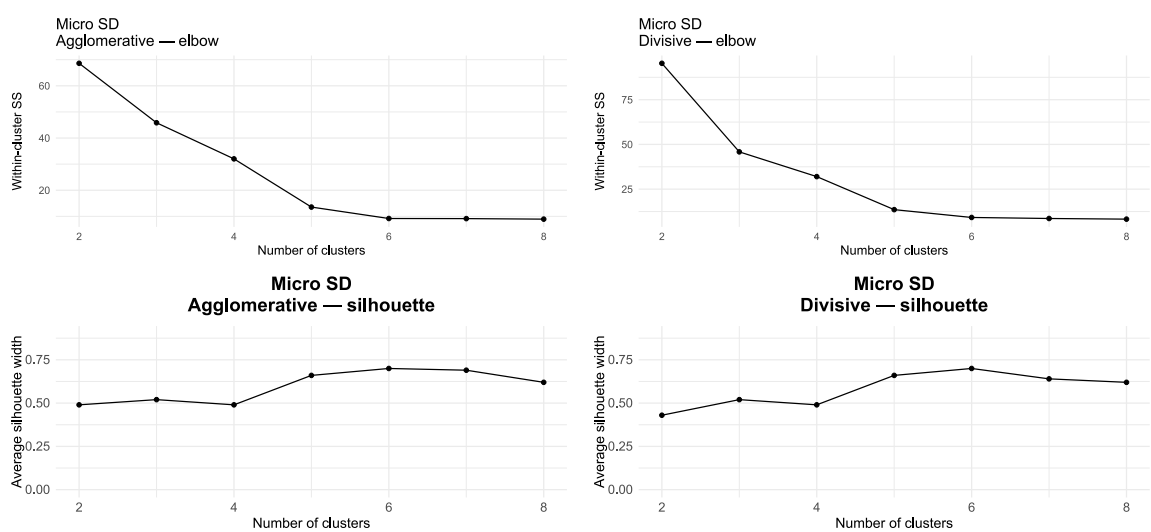


Figure 1B: Cluster diagnostic plots for the micro socio-demographic solution. The figure reports elbow (top panels) and average silhouette width (bottom panels) diagnostics for the agglomerative (left panels; Agnes) and divisive (right panels; Diana) hierarchical clustering solutions based on the micro socio-demographic distance matrix.

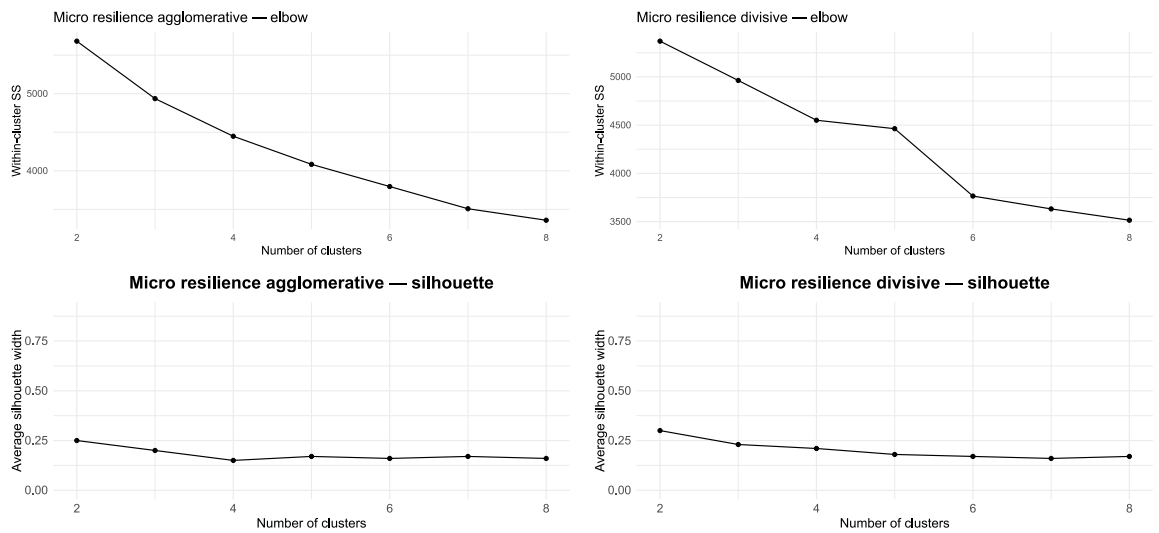


Figure 2B: Diagnostic statistics for selecting the number of micro-level resilience clusters. Elbow and silhouette diagnostics are shown for the agglomerative (Agnes) and divisive (Diana) hierarchical clustering solutions based on the micro-level resilience distance matrix.

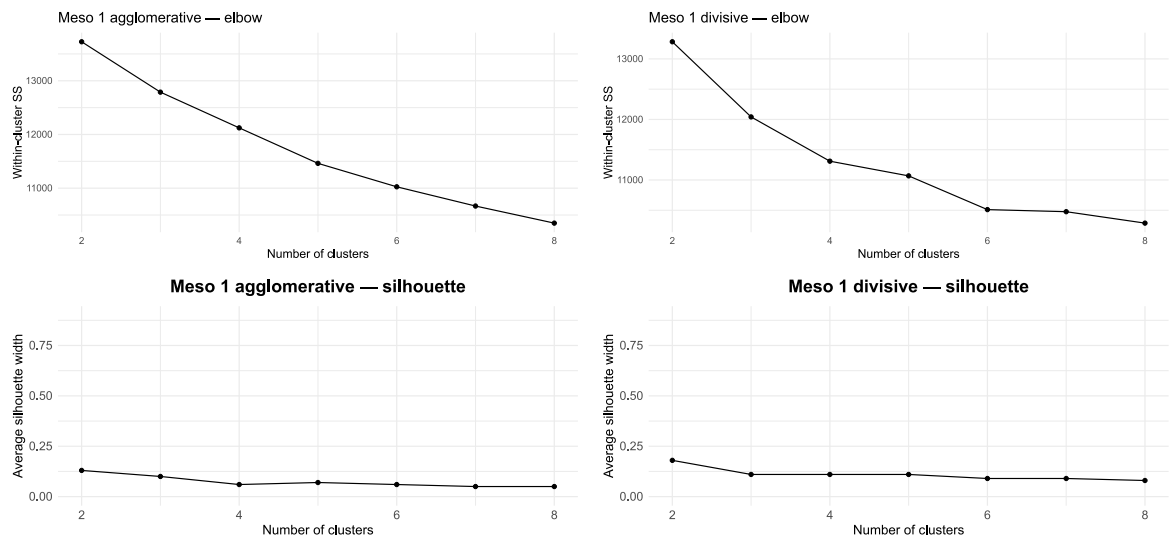


Figure 3B: Elbow and silhouette diagnostics are shown for the agglomerative (Agnes) and divisive (Diana) hierarchical clustering solutions based on the meso-level ideological distance matrix. The elbow plots indicate a gradual reduction in within-cluster variance, while silhouette values remain generally low across solutions.

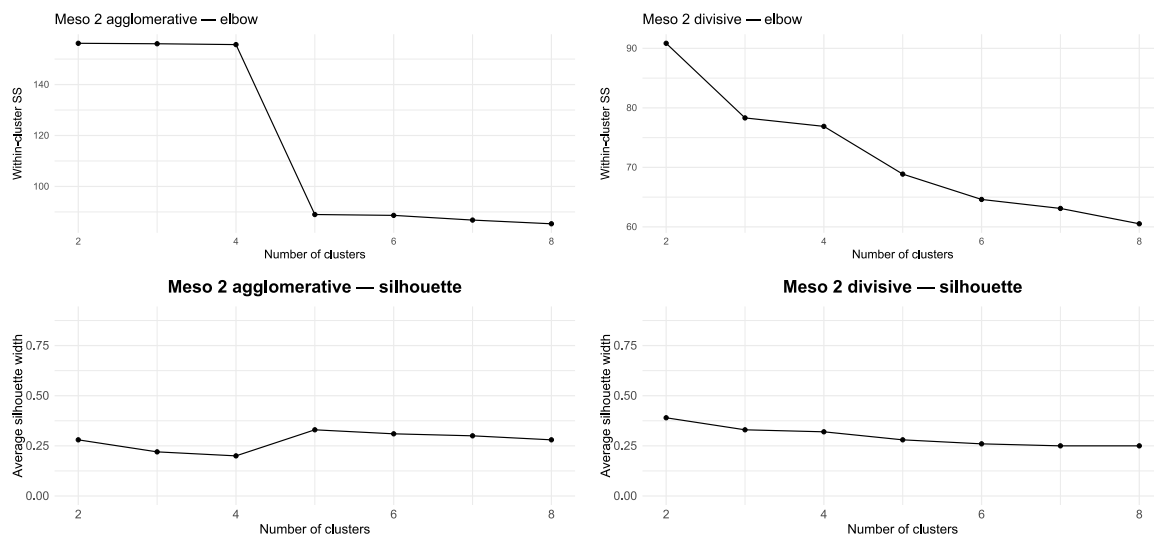


Figure 4B: Diagnostic statistics for selecting the number of meso-level media-use clusters. Elbow and silhouette diagnostics are shown for the agglomerative (Agnes) and divisive (Diana) hierarchical clustering solutions based on the meso-level media-use distance matrix.

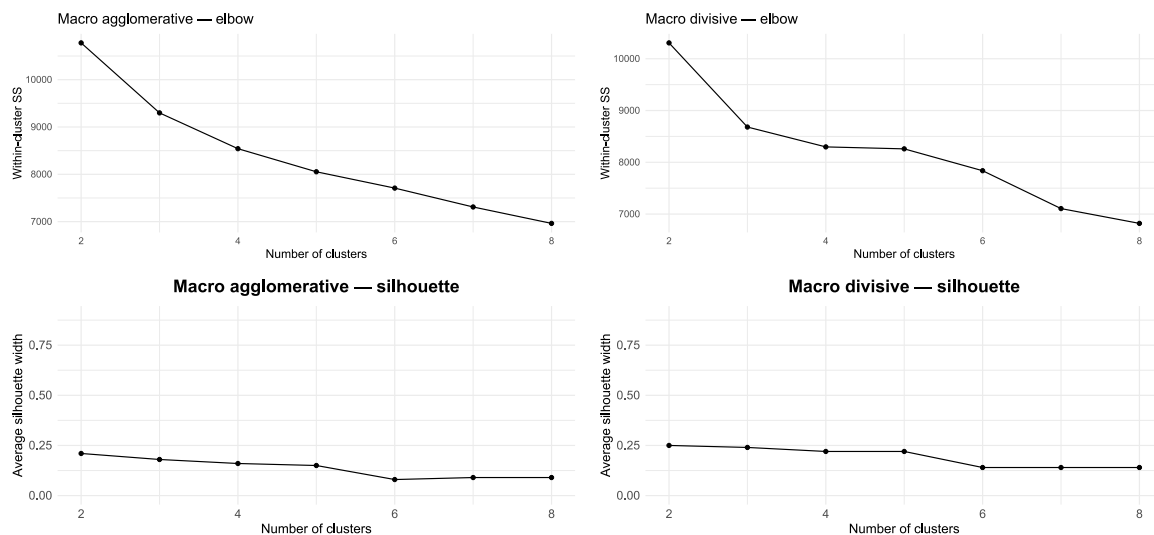


Figure 5B: Diagnostic statistics for selecting the number of macro-level trust clusters. Elbow and silhouette diagnostics are shown for the agglomerative (Agnes) and divisive (Diana) hierarchical clustering solutions based on the macro-level trust distance matrix.

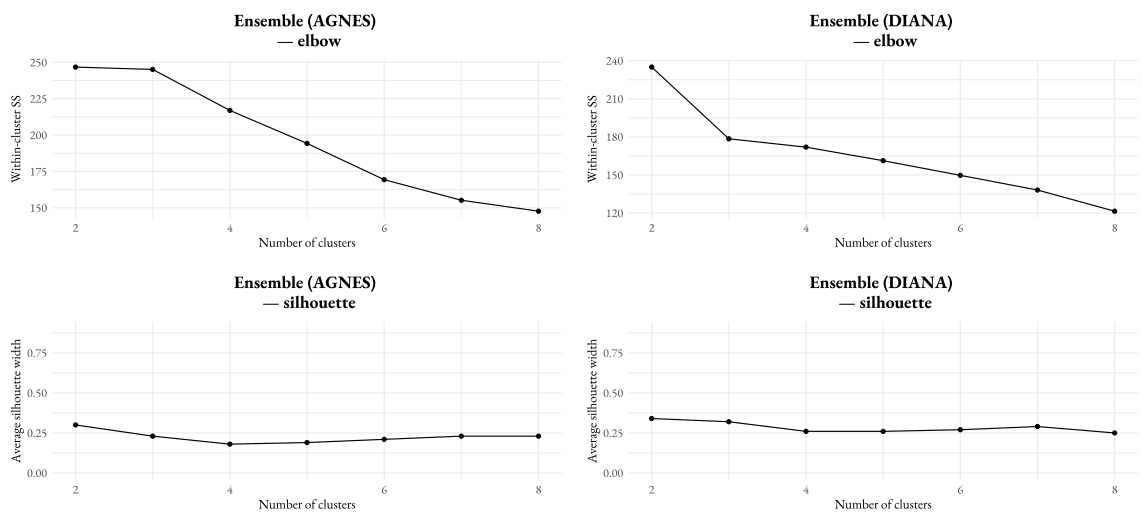


Figure 6B: Diagnostic statistics for selecting the number of ensemble clusters. Elbow and silhouette diagnostics are shown for the agglomerative (Agnes) and divisive (Diana) hierarchical clustering solutions based on the ensemble dissimilarity matrix. The diagnostics support a three-cluster solution, where silhouette values remain comparatively high while additional clusters provide only limited gains in separation and fit.

Table 1B

Variable	Institutional Mainstream	Networked Apart	Right-Wing Authoritarian
Socio-demographic profile			
Age	51.10	44.50	44.90
Male	0.51	0.44	0.52
Higher education	0.56	0.53	0.40
Lower education	0.12	0.12	0.15
Migration background	0.06	0.08	0.09
Subjective SES	6.27	5.22	5.77
Political trust and attitudes			
Democratic trust	2.89	2.61	2.20
Non-traditional trust	2.12	2.11	2.16
Political ideology	5.46	5.34	7.61
Populism	3.33	3.39	4.17

Variable	Institutional Mainstream	Networked Apart	Right-Wing Authoritarian
Conspiracy mentality	2.98	3.07	4.05
RWA submission	2.40	2.29	3.53
Satisfaction with democracy	3.41	3.03	2.70
Media habits			
TV news	0.79	0.69	0.71
Radio news	0.74	0.63	0.63
Social media news	0.59	0.68	0.64
AI chatbot news	0.31	0.35	0.33
Online newspapers	0.50	0.49	0.37
Offline newspapers	0.52	0.45	0.39
Broadcasters online	0.47	0.42	0.41
Alternative news	0.06	0.09	0.13
Psychological profile			
Resilience	3.75	2.72	3.32
Loneliness	1.89	2.65	2.38

Table 1B: Consensus clusters and individual level cluster assignment

Note. Values represent cluster centroids (means) for each variable within the three-cluster ensemble solution. Binary indicators (e.g., gender, education, migration background, and media use) are reported as proportions, where values closer to 1 indicate a larger share of participants within that cluster possessing the characteristic. Higher values on political ideology indicate stronger right-wing self-placement.

Table 2B: AMCEs across clusters.

Attribute	Cluster	Comparison (Reference in parentheses)	AMCE (SE)	95% CI	p
Communicator	Institutional Mainstream	AI (vs. News media)	-29.5 pp (0.02)	[-33.0, -26.0]	< .001
		Chancellor (vs. News media)	-2.5 pp (0.02)	[-6.0, 1.1]	= .17
		Citizens (vs. News media)	-20.0 pp (0.02)	[-23.6, -16.4]	< .001
		Military (vs. News media)	-10.7 pp (0.02)	[-14.3, -7.2]	< .001
	Networked Apart	AI (vs. News media)	-26.3 pp (0.03)	[-31.4, -21.2]	< .001
		Chancellor (vs. News media)	-3.6 pp (0.03)	[-8.9, 1.7]	= .18
		Citizens (vs. News media)	-10.6 pp (0.03)	[-15.9, -5.4]	< .001
		Military (vs. News media)	-8.0 pp (0.03)	[-13.2, -2.7]	= .003
	Right-Wing Authoritarian	AI (vs. News media)	-17.0 pp (0.02)	[-21.5, -12.5]	< .001
		Chancellor (vs. News media)	-3.7 pp (0.02)	[-8.3, 0.9]	= .11
		Citizens (vs. News media)	-7.5 pp (0.02)	[-12.1, -2.8]	= .002
		Military (vs. News media)	-1.0 pp (0.02)	[-5.3, 3.4]	= .67
Citizen integration	Institutional Mainstream	No integration (vs. High integration)	-2.0 pp (0.01)	[-4.8, 0.8]	= .15
		Partial integration (vs. High integration)	-3.5 pp (0.01)	[-6.3, -0.8]	= .012
	Networked Apart	No integration (vs. High integration)	-4.0 pp (0.02)	[-7.9, -0.2]	= .041
		Partial integration (vs. High integration)	-4.9 pp (0.02)	[-8.7, -1.1]	= .012
	Right-Wing Authoritarian	No integration (vs. High integration)	-2.9 pp (0.02)	[-6.3, 0.5]	= .09
		Partial integration (vs. High integration)	-6.3 pp (0.02)	[-9.8, -2.9]	< .001

Attribute	Cluster	Comparison (Reference in parentheses)	AMCE (SE)	95% CI	p
Misinformation handling	Institutional Mainstream	Debunking (vs. Prebunking)	1.0 pp (0.01)	[-1.8, 3.8]	= .49
		No engagement (vs. Prebunking)	-10.4 pp (0.01)	[-13.2, -7.6]	< .001
	Networked Apart	Debunking (vs. Prebunking)	0.0 pp (0.02)	[-3.7, 3.7]	= .98
		No engagement (vs. Prebunking)	-4.4 pp (0.02)	[-8.3, -0.6]	= .025
	Right-Wing Authoritarian	Debunking (vs. Prebunking)	-1.2 pp (0.02)	[-4.5, 2.2]	= .49
		No engagement (vs. Prebunking)	-6.0 pp (0.02)	[-9.1, -2.8]	< .001
Regulation	Institutional Mainstream	Bans (vs. Moderation)	-1.1 pp (0.02)	[-4.3, 2.0]	= .48
		Counterspeech (vs. Moderation)	-4.1 pp (0.02)	[-7.4, -0.8]	= .015
		Upload filter (vs. Moderation)	0.0 pp (0.02)	[-3.1, 3.2]	= .98
	Networked Apart	Bans (vs. Moderation)	-2.7 pp (0.02)	[-6.9, 1.6]	= .22
		Counterspeech (vs. Moderation)	-13.0 pp (0.02)	[-17.5, -8.5]	< .001
		Upload filter (vs. Moderation)	-3.7 pp (0.02)	[-8.1, 0.7]	= .10
	Right-Wing Authoritarian	Bans (vs. Moderation)	-2.0 pp (0.02)	[-5.9, 1.9]	= .31
		Counterspeech (vs. Moderation)	-3.3 pp (0.02)	[-7.4, 0.7]	= .10
		Upload filter (vs. Moderation)	-0.5 pp (0.02)	[-4.5, 3.5]	= .81

Attribute	Cluster	Comparison (Reference in parentheses)	AMCE (SE)	95% CI	p
Transparency	Institutional Mainstream	Low transparency (vs. High transparency)	-7.7 pp (0.01)	[-10.6, -4.8]	< .001
		Strategic transparency (vs. High transparency)	-5.1 pp (0.01)	[-7.9, -2.2]	< .001
	Networked Apart	Low transparency (vs. High transparency)	-3.5 pp (0.02)	[-7.4, 0.4]	= .08
		Strategic transparency (vs. High transparency)	-0.3 pp (0.02)	[-4.3, 3.6]	= .87
	Right-Wing Authoritarian	Low transparency (vs. High transparency)	-5.9 pp (0.02)	[-9.3, -2.5]	< .001
		Strategic transparency (vs. High transparency)	-2.4 pp (0.02)	[-5.8, 1.0]	= .16
Truth base	Institutional Mainstream	Legal basis (vs. Evidence)	-8.0 pp (0.02)	[-11.3, -4.6]	< .001
		Majority opinion (vs. Evidence)	-13.0 pp (0.02)	[-16.3, -9.8]	< .001
		Platforms (vs. Evidence)	-11.9 pp (0.02)	[-15.3, -8.5]	< .001
	Networked Apart	Legal basis (vs. Evidence)	-2.9 pp (0.02)	[-7.4, 1.5]	= .20
		Majority opinion (vs. Evidence)	-10.3 pp (0.02)	[-14.6, -6.0]	< .001
		Platforms (vs. Evidence)	-8.2 pp (0.02)	[-12.6, -3.8]	< .001
	Right-Wing Authoritarian	Legal basis (vs. Evidence)	-2.3 pp (0.02)	[-6.4, 1.9]	= .28
		Majority opinion (vs. Evidence)	-2.8 pp (0.02)	[-6.8, 1.2]	= .18
		Platforms (vs. Evidence)	-4.7 pp (0.02)	[-8.9, -0.4]	= .030

Bibliography

1 Christian Stadler: On Notion and Nature of 'Sovereignty' between Mimesis and Might – A Polemological Reflection on Digital Sovereignty in the Information Space.

Colloquium on Violence and Religion (1990) International Association of Scholars of Mimetic Theory [Online]. Available at: <https://violenceandreligion.com/mimetic-theory-2/> [Accessed: 14 May 2026].

Girard, R. (1965) *Deceit, Desire, and the Novel: Self and Other in Literary Structure*. Baltimore: Johns Hopkins University Press.

Girard, R. (1977) *Violence and the Sacred*. Baltimore: Johns Hopkins University Press.

Girard, R. (1987) *Things Hidden Since the Foundation of the World*. Stanford: Stanford University Press.

Haven, C. (2009) 'History is a test. Mankind is failing it: René Girard scrutinises the human condition from creation to apocalypse', *Stanford Magazine* [Online]. Available at: <https://stanfordmag.org/contents/history-is-a-test-mankind-is-failing-it> [Accessed: 14 May 2026].

Palaver, W. (2013) *René Girard's Mimetic Theory*. East Lansing: Michigan State University Press.

Palaver, W. (2024) 'On the Dangers of a Return to Constantinianism' [Online]. Available at: <https://acireland.ie/on-the-dangers-of-a-return-to-constantinianism/> [Accessed: 14 May 2026].

Palaver, W. (2025) 'Desacralize the Katechon, Do Not Create Empires!' [Online]. Available at: <https://acireland.ie/desacralize-the-katechon-do-not-create-empires/> [Accessed: 14 May 2026].

Rabe, A. (n.d.) 'Mimetic Theory and Anthropological Theology: An online course' [Online]. Available at: <https://www.mimesis.academy/mimetic-theory-and-anthropological-theology/> [Accessed: 14 May 2026].

Reese, E. (2018) 'Signs in the Heavens: René Girard on the Apocalypse', *Dominicana Journal* [Online]. Available at: <https://www.dominicanajournal.org/signs-in-the-heavens-rene-girard-on-the-apocalypse/> [Accessed: 14 May 2026].

Thiel, P. (1995) 'The Diversity Myth', *The New Criterion* [Online]. Available at: <https://newcriterion.com/article/the-diversity-myth/> [Accessed: 14 May 2026].

Thiel, P. (2007) 'The Straussian Moment' [Online]. Available at: <https://gwern.net/doc/politics/2007-thiel.pdf> [Accessed: 14 May 2026].

Thiel, P. (2009) 'The Education of a Libertarian', *Cato Unbound* [Online]. Available at: <https://www.cato-unbound.org/2009/04/13/peter-thiel/education-libertarian/> [Accessed: 14 May 2026].

Thiel, P. (2014) *Zero to One: Notes on Startups, or How to Build the Future*. New York: Crown Business.

Thiel, P. (n.d.) *Imitatio – René Girard Research Site*. Thiel Foundation [Online]. Available at: <https://www.thielfoundation.org/> [Accessed: 14 May 2026].

Wandinger, N. (2009) 'Sacrifice in Harry Potter from a Girardian Perspective' [Online]. Available at: <https://www.uibk.ac.at/de/theol/leseraum/texte/819/> [Accessed: 14 May 2026].

2 Matthias Postl: The Information Environment – A Comparative Analysis and Contextualisation of Existing Concepts

‘Problems of Russia’s National Interests in the Information Sphere’ (2022)
Gosudarstvo i Pravo, 4, pp. 93–101.

‘The Development of Cyberterrorism in the Context of the Globalisation of Cyberspace’ (2021) Bulletin of the University of the Ministry of Internal Affairs of Russia, 3, pp. 112–118.

Arendt, H. (1958) *The Human Condition*. Chicago: University of Chicago Press.

Barth, V., et al. (2023) *Digitale Öffentlichkeit und Plattformstrukturen*. Wiesbaden: Springer VS.

Beck, U. (1986) *Risikogesellschaft: Auf dem Weg in eine andere Moderne*. Frankfurt am Main: Suhrkamp.

Berger, P. L. and Luckmann, T. (1966) *The Social Construction of Reality: A Treatise in the Sociology of Knowledge*. New York: Anchor Books.

Berk, N. (2025) ‘The Impact of Media Framing in Complex Information Environments’, *Political Communication*, 42(5), pp. 757–773.

Bourdieu, P. (1977) *Outline of a Theory of Practice*. Cambridge: Cambridge University Press.

Bradford, A. (2020) *The Brussels Effect: How the European Union Rules the World*. Oxford: Oxford University Press.

Bradshaw, S. and Howard, P. (2019) *The Global Disinformation Order*. Oxford: Oxford Internet Institute.

Bratton, B. (2016) *The Stack: On Software and Sovereignty*. Cambridge, MA: MIT Press.

Brocke, J. vom, Simons, A., Niehaves, B., Riemer, K., Plattfaut, R., and Cleven, A. (2009) Reconstructing the Giant: On the Importance of Rigour in Documenting the Literature Search Process. ECIS 2009 Proceedings. 17th European Conference on Information Systems (ECIS). Available at: <https://aisel.aisnet.org/ecis2009/161> [Accessed: 14 May 2026].

Buchanan, B. (2020) *The Cybersecurity Dilemma: Hacking, Trust and Fear Between Nations*. Oxford: Oxford University Press.

Buckland, M. (2017) *Information and Society*. Cambridge, MA: MIT Press.

Capurro, R. (2005) 'Past, present, and future of the concept of information.' *Annual Review of Information Science and Technology*, 40(1), pp. 353–379.

Capurro, R. (2024) 'Information und Ethik.' [Online]. Available at: <https://www.capurro.de/infovorl-kap3.htm> [Accessed: 14 May 2026].

Central Propaganda Department of the Communist Party of China (2022) *Study Outline on the Overall National Security Concept (总体国家安全观学习纲要)*. Beijing: People's Publishing House.

Chinese Communist Party (CCP) (2003) *Regulations on Political Work of the Chinese People's Liberation Army (中国人民解放军政治工作条例)*. Beijing: Central Military Commission.

Clausewitz, C. von (2008) *Vom Kriege*. Berlin: Dümmler.

Council of the European Union (2022) *A Strategic Compass for Security and Defence*. Brussels.

Cyber Huntress (2024) 'Understanding Foreign Information Manipulation and Interference (FIMI): A Growing Global Threat' [Online]. Available at: <https://medium.com/@thecyberhuntress/understanding-foreign-information-manipulation-and-interference-fimi-a-growing-global-threat-e5dd1f726809> [Accessed: 14 May 2026].

Decision (EU) 2022/2481 of the European Parliament and of the Council of 14 December 2022 on the Digital Decade 2030 policy programme.

Deibert, R. (2020) *Reset: Reclaiming the Internet for Civil Society*. Toronto: House of Anansi Press.

Dolata, U. (2017) *Technik und soziale Transformation: Die Entwicklung digitaler Plattformen*. Berlin: Suhrkamp.

Döring, N., and Bortz, J. (2016) *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften*. Berlin, Heidelberg: Springer. Available at: <https://doi.org/10.1007/978-3-642-41089-5> [Accessed: 14 May 2026].

Dunn Cavelty, M. (2013) 'Cybersecurity and the Political', *European Journal of International Security*, 2(1), pp. 1–18.

EEAS (2024) 'Tackling Disinformation, Foreign Information Manipulation and Interference.' [Online]. Available at: https://www.eeas.europa.eu/eeas/tackling-disinformation-foreign-information-manipulation-interference_en [Accessed: 14 May 2026].

ENISA (2024) 'Foreign Information Manipulation and Interference (FIMI) and Cybersecurity – Threat Landscape.' [Online]. Available at: <https://www.enisa.europa.eu/publications/foreign-information-manipulation-interference-fimi-and-cybersecurity-threat-landscape> [Accessed: 14 May 2026].

Entman, R. (1993) 'Framing: Toward Clarification of a Fractured Paradigm', *Journal of Communication*, 43(4), pp. 51–58.

Ertl, P. (2022) 'Abschied von der Aufklärung? Reflexionen zur Zukunft Europas im Anschluss an Kant, Hegel, Foucault & Co', In: Balthasar/Ertl/Fleck/Lech (Hrsg.): *Internationale Perspektiven im 21. Jahrhundert. Recht – Politik – Gesellschaft*, Sramek, Wien 2021, pp. 169–200.

Ertl, P. (2021) 'Ethik der Resilienz. Staatliche Resilienz als (globale) moralische Verantwortlichkeit', in: Peischel, Wolfgang (Hrsg.): *Resilienz im 21. Jahrhundert – Ein Institutsübergreifendes LVA-Projekt*, Miles-Publishers, pp. 253–271.

Ertl, P. (2025) 'Web Life in the Light of Progressus Theory. A Dialog with Laszlo Ropolyi's Philosophy of the Internet', in: Szívós, Mihály & Szántó, Tibor (eds.): Knowledge, Controversy, Virtuality. Hungarian Philosophical Society, L'Harmattan Könyvesbolt, Budapest, pp. 153–164.

European Commission (2018) Action Plan against Disinformation. Brussels.

European Commission (2020) European Action Plan for Democracy. COM (2020) 790 final. Brussels.

European Commission (2020a) A European Data Strategy. COM(2020) 66 final. Brussels.

European Commission (2023) Hybrid Threats: A Comprehensive Approach. Brussels: European Commission.

European Commission (2025) European Democracy Shield. Brussels.

European External Action Service (EEAS) (2023) 1st Report on Foreign Information Manipulation and Interference Threats. Brussels.

European External Action Service (EEAS) (2024) 2nd Report on Foreign Information Manipulation and Interference Threats. Brussels.

European External Action Service (EEAS) (2025) 3rd Report on Foreign Information Manipulation and Interference Threats. Brussels.

Federal Government (2023) 'Was ist FIMI?' [Online]. Available at: <https://www.bundesregierung.de/breg-de/aktuelles/fimi-fake-news-international-2227610> [Accessed: 14 May 2026].

Federal Ministry of Defence (BMVg) (2018) Konzeption der Bundeswehr. Berlin.

Floridi, L. (2011) The Philosophy of Information. Oxford: Oxford University Press.

Foucault, M. (1971) The Order of Discourse. Paris: Gallimard.

- Foucault, M. (1972) *The Archaeology of Knowledge*. London: Routledge.
- Gamson, W.A. (1975) 'Review of Frame Analysis: An Essay on the Organisation of Experience', *Contemporary Sociology*, 4(6), pp. 603–607.
- Giles, K. (2021) *Handbook of Russian Information Warfare*. London: Routledge.
- Giles, K. (2022) *Russia's War on Everybody: And What It Means for You*. London: Bloomsbury Academic.
- Goffman, E. (1974) *Frame Analysis*. New York: Harper and Row.
- Habermas, J. (1981) *Theorie des kommunikativen Handelns*. 2nd ed. Frankfurt am Main: Suhrkamp.
- Habermas, J. (1992) *Faktizität und Geltung*. Frankfurt am Main: Suhrkamp.
- Hennecken, L. (n.d.) *Cyberraum und Informationsraum: Eine konzeptionelle Einordnung*. Unpublished manuscript.
- Herman, E.S. and Chomsky, N. (2023) *Die Konsensfabrik: Die politische Ökonomie der Massenmedien*. 3rd ed. Frankfurt am Main: Westend.
- Hoffmann, O. and Sandhu, S. (2024) *Handbuch Strategische Kommunikation*. Wiesbaden: Springer VS.
- Hofkirchner, W., Fuchs-Kittowski, F., and Fleissner, P. (1995) 'Einheit der Informationswissenschaft.' in: F. Fuchs-Kittowski and P. Fleissner (Eds.), *Informationswissenschaft: Grundlagen, Probleme, Perspektiven*. Konstanz: UVK, pp. 263–292.
- Howard, O. (2025) 'Communication Strategy: Rethinking Strategic Communication', *Military Review*, September–October.
- Howard, P. (2018) 'Computational Propaganda', Oxford Internet Institute Working Paper.

Information Security Doctrine of the Russian Federation (2016) Doktrina informatsionnoj bezopasnosti Rossijskoj Federatsii. Moscow: President of the Russian Federation.

Jecker, C. (2017) Entmans Framing-Ansatz: Theoretische Grundlegung und empirische Umsetzung. Cologne: Herbert von Halem Verlag.

Joint Chiefs of Staff (1996) Joint Publication 3-13.1: Joint Doctrine for Command and Control Warfare. Washington, D.C.

Joint Chiefs of Staff (1998) Joint Publication 3-13: Joint Doctrine for Information Operations. Washington, D.C.

Joint Chiefs of Staff (2001) JP 1-02: Department of Defence Dictionary of Military and Associated Terms. Washington, D.C.

Joint Chiefs of Staff (2006) Joint Publication 3-13: Information Operations. Washington, D.C.

Joint Chiefs of Staff (2012) Joint Publication 3-13: Information Operations. Washington, D.C.

Joint Chiefs of Staff (2022) Joint Publication 3-04: Information in Joint Operations. Washington, D.C.

Jowett, G.S. and O'Donnell, V. (2019) Propaganda and Persuasion. 7th ed. Thousand Oaks: Sage.

Kahneman, D. (2011) Thinking, Fast and Slow. New York: Farrar, Straus and Giroux.

Kim, K. (2025) 'A Study on the Concept of Cyber Cognitive Warfare', European Conference on Cyber Warfare and Security, pp. 241–248.

Klintsova, M.N. (2023) 'Tactics of Conducting Information Warfare during the Special Military Operation', Problems of National Strategy, 29, pp. 56–69.

Komov, S.A. (1996) 'Information War', Military Thought, 7, pp. 75–80.

- Lanoszka, A. (2019) 'Disinformation in international politics', *European Journal of International Security*, 4(2), pp. 227–248.
- Lasswell, H.D. (1927) 'The Theory of Political Propaganda', *American Political Science Review*, 21(3), pp. 627–631.
- Lee, S. (2014) China's 'Three Warfares': Origins, Applications, and Organisations. *Journal of Strategic Studies*, 37(2), pp. 198–221.
- Lessig, L. (1999) *Code and Other Laws of Cyberspace*. New York: Basic Books.
- Lewin, K. (1947) 'Frontiers in Group Dynamics', *Human Relations*, 1(1), pp. 5–41.
- Liang, X. (2023) 'Cognitive Security: The New Foundational Domain of National Security (认知安全 : 国家安全的新域基石)', *Journal of National Security Studies*, CNKI database.
- Lin, K. (2023) 'On the Construction of a National Cognitive Security System', *Journal of Strategic Studies*, CNKI database.
- Lippmann, W. (1922) *Public Opinion*. New York: Harcourt, Brace and Company.
- Manoilo, A.V. (2015) *State Information Policy and the Management of Information Conflicts*. Moscow: MGIMO University.
- McCombs, M. and Shaw, D. (1972) 'The Agenda-Setting Function of Mass Media', *Public Opinion Quarterly*, 36(2), pp. 176–187.
- Messner, E.E. (1960) *Mjatezhevojna*. Buenos Aires: Svit Publishing House.
- Miskimmon, A., O'Loughlin, B. and Roselle, L. (2013) *Strategic Narratives: Communication Power and the New World Order*. London: Routledge.
- Morse, J. C. (2022) 'Fragmentation in Global Governance', *International Organisation*, 76(2), pp. 345–371.
- NAT (2015) *AJP-3.10.1 Psychological Operations*. Brussels.

National Security Strategy of the Russian Federation (2021) Strategija nacional'noj bezopasnosti Rossijskoj Federatsii. Moscow: President of the Russian Federation.

NATO (2015) AJP-3.10.1 Allied Joint Doctrine for Psychological Operations (PSYOPS). Brussels: NATO Standardisation Office.

NATO (2017) MC 0628 NATO Military Policy on Strategic Communications. Brussels: NATO.

NATO (2019a) MC 0402/3 NATO Military Policy on Psychological Operations. Brussels: NATO.

NATO (2022) AJP-01 Allied Joint Doctrine. Brussels.

NATO (2023a) AJP-10 Strategic Communications. Brussels.

NATO (2023b) AJP-10.1 Information Operations. Brussels.

NATO (2024) AJP-10.3 Military Public Affairs. Brussels.

NATO (2024) PO(2024)0440 Military Policy on Psychological Operations. Brussels: NATO.

NATO (2025) Information Environment Briefing. NATO.

Office of the Chairman of the Joint Chiefs of Staff (2025) DoD Dictionary of Military and Associated Terms. Washington, D.C.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., McGuinness L., Stewart L., Thomas J., Tricco A., Welch V., Whiting P. and Moher, D. (2021) 'The PRISMA 2020 statement: An updated guideline for reporting systematic reviews.' *Systematic Reviews*, 10(89). Available at: <https://doi.org/10.1186/s13643-021-01626-4> [Accessed: 14 May 2026].

Panarin, I.N. (2006) *Information War and Geopolitics*. Moscow: Pokolenie.

Paul, C. and Matthews, M. (2016) 'The Russian "Firehose of Falsehood" Propaganda Model.' [Online]. Available at: <https://www.rand.org/pubs/perspectives/PE198.html> [Accessed: 14 May 2026].

Prozorov, S. (2021) 'Post-Truth and the End of Truth', in: Giles, K. (ed.): *Russia's War on Everybody*. London: Bloomsbury Academic.

Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market for Digital Services (Digital Services Act).

Regulation (EU) 2022/868 of the European Parliament and of the Council of 30 May 2022 on European data governance (Data Governance Act).

Regulation (EU) 2024/1083 of the European Parliament and of the Council on safeguarding media freedom and pluralism in the internal market (European Media Freedom Act).

Robbins, J. W. (2020) 'Countering Russian Disinformation', in: Conley, H. A. et al. (eds.) *The Diversity of Russia's Military Power: Five Perspectives*. Washington, DC: Center for Strategic and International Studies (CSIS), pp. 32–39.

Roselle, L., Miskimmon, A. and O'Loughlin, B. (2014) 'Strategic Narrative', *Media, War and Conflict*, 7(1), pp. 70–84.

Russian Federation (2023) *Concept of the Foreign Policy of the Russian Federation*. Decree of the President of the Russian Federation No. 229 of 31 March 2023. Moscow: President of the Russian Federation.

Schmitt, C. (1932) *Der Begriff des Politischen*. Berlin: Duncker und Humblot.

Schrijvers, E. et al. (2026) *Media Matters: Democratic Infrastructure in the Digital Age*. Amsterdam: Amsterdam University Press.

Shannon, C. E. and Weaver, W. (1949) *The Mathematical Theory of Communication*. Urbana: University of Illinois Press.

Smirnova, A. (2025) *Russische Konzepte der Informationskriegsführung*. Berlin: CeMAS – Centre for Monitoring, Analysis and Strategy.

Sørensen, K. H. and Kocksch, L. (2019) *Digital Transformations and Society*. Cheltenham: Edward Elgar.

State Council Information Office of the People's Republic of China (2011) *China's Peaceful Development*. Beijing: SCIO.

State Council Information Office of the People's Republic of China (2015) *China's Military Strategy*. Beijing: SCIO.

State Council Information Office of the People's Republic of China (2019) *China's National Defence in the New Era*. Beijing: SCIO.

Streletsky, Ya.I. (2024) 'Russia's Special Military Operation in Ukraine: Information Warfare of the "Alliance of Democracies"', *Military History Journal*, 2, pp. 45–54.

U.S. Department of Defence (2010) *Joint Publication 3-53: Psychological Operations*. Washington, D.C.

U.S. Department of the Air Force (2023) *Air Force Doctrine Publication 3-13: Information in Air Force Operations*. Washington, D.C.

Vardanyan, L. and Kocharyan, H. (2023) 'Digital Sovereignty in the European Union', *Journal of International Relations and Development*.

White, D. M. (1950) 'The Gatekeeper: A Case Study in the Selection of News', *Journalism Quarterly*, 27(4), pp. 383–390.

Whyte, C. and Etudo, U. (2023) *Finding the Thieves Amongst the Liars: Thinking Clearly About Cyber-Enabled Influence Operations*. Oxford: Oxford University Press.

Wilbur, D. (2023) 'Agenda Setting in Geopolitical Conflict', *Journal of Information Warfare*.

Wilbur, D.S. (2021) 'Propaganda or Not: Strategic Narrative and Agenda Setting', *Journal of Information Warfare*, 20(3), pp. 146–156.

Wu, A. (2022) 'Cognitive Warfare from a Conceptual-Historical Perspective', Journal of Military Theory, CNKI database.

Yan, Z. (2023) 'International Competition in the Cognitive Domain: Origins, Dimensions and Security Logic', International Security Review, CNKI database.

Yu, Y. (2023) 'Mechanisms and Strategic Choices of Cognitive Domain Operations', PLA Military Studies, CNKI database.

Zeng, Q. (2023) 'Weaponisation of Cognition and AI-Enabled Cognitive Warfare', Journal of Informationized Warfare Studies, CNKI database.

Zuboff, S. (2019) *The Age of Surveillance Capitalism*. New York: PublicAffairs.

3 Alexandra Duca: Cognitive Warfare – A Legal Assessment

Baade, B. (2020) 'Informationskriegführung und Fake News', in Kielmansegg, S.G. v., Krieger, H. and Sohm, S. (eds.), *Die Wiederkehr der Landes- und Bündnisverteidigung*. Baden-Baden. Nomos.

Cyber Law Toolkit (2021) 'Category: National position'. Available at: https://cyberlaw.ccdcoe.org/wiki/Category:National_position [Accessed: 13 August 2025].

Cyber Law Toolkit (2025) 'Scenario 01: Election interference'. Available at: https://cyberlaw.ccdcoe.org/wiki/Scenario_01:_Election_interference [Accessed: 30 December 2025].

Cyber Law Toolkit (2025) 'Self-defence'. Available at: <https://cyberlaw.ccdcoe.org/wiki/Self-defence> [Accessed 9 January 2025].

Human Rights Council (2012) *The promotion, protection and enjoyment of human rights on the Internet*. Resolution A/HRC/RES/20/8.

International Court of Justice (1949) *The Corfu Channel Case (United Kingdom v. Albania) (Merits)*. ICJ Rep 4.

International Court of Justice (1986) Case Concerning Military and Paramilitary Activities in and Against Nicaragua (Nicaragua v. United States) (Merits), ICJ Rep 14.

International Court of Justice (1996) Legality of the Threat or Use of Nuclear Weapons (Advisory Opinion). ICJ Rep 226.

International Court of Justice (2003) Oil Platforms (Iran v. United States) (Merits). ICJ Rep 4.

International Court of Justice (2007) Case Concerning Application of the Convention on the Prevention and Punishment of the Crime of Genocide (Bosnia and Herzegovina v. Serbia and Montenegro) (Merits), ICJ Rep 43.

International Law Commission (2001) Articles on Responsibility of States for Internationally Wrongful Acts. UN Doc. A/RES/56/83.

Kielmansegg S. G. v., Krieger H., and Sohm S. (eds.) (2020) Die Wiederkehr der Landes- und Bündnisverteidigung. Baden-Baden: Nomos.

Kriener, F. (2023) 'Intervention, Prohibition of', Max Planck Encyclopedia of Public International Law. Available at: <https://opil.ouplaw.com/display/10.1093/law:epil/9780199231690/law-9780199231690-e1434?prd=EPIL>. [Accessed: 11 May 2026].

Republic of Austria (2024) Position Paper of the Republic of Austria: Cyber Activities and International Law. Available at: [https://docs-library.unoda.org/Open-Ended_Working_Group_on_Information_and_Communication_Technologies_-_ \(2021\)/Austrian_Position_Paper_-_Cyber_Activities_and_International_Law_\(Final_23.04.2024\).pdf](https://docs-library.unoda.org/Open-Ended_Working_Group_on_Information_and_Communication_Technologies_-_ (2021)/Austrian_Position_Paper_-_Cyber_Activities_and_International_Law_(Final_23.04.2024).pdf) [Accessed: 11 May 2026].

Rodenhäuser, T. and D'Cunha, S. (2023) 'Foghorns of war: IHL and information operations during armed conflict', ICRC Humanitarian Law & Policy Blog, 12 October. Available at: <https://blogs.icrc.org/law-and-policy/2023/10/12/foghorns-of-war-ihl-and-information-operations-during-armed-conflict/>. [Accessed: 11 May 2026].

Schmitt, M. N. (2021) 'Foreign Cyber Interference in Elections', *International Law Studies*, 97, pp. 739-764.

Schmitt, M. N. (ed.) (2016) *Tallinn Manual 2.0 on the International Law Applicable to Cyber Operations*. Cambridge: Cambridge University Press.

United Nations (1970) UN 1970 Special Committee on Principles of International Law concerning Friendly Relations and co-operation among States. Summary records of the 110th to 114th meeting, held at the Palais des Nations, Geneva, from 31 March to 1 May 1970. UN Doc. A/AC.125/SR.110-114.

United Nations International Covenant on Civil and Political Rights Human Rights Committee (2011) General comment No.34. CCPR/C/GC/35.

4 Philip Schäfer: Protection and Civil Liberties in the Digital Information Environment

Buzan, B., Wæver, O. and de Wilde, J. (1998) *Security: A New Framework for Analysis*. Boulder, CO: Lynne Rienner Publishers.

Austria

Bundeskanzleramt (2013) *Austrian Security Strategy. Security in a new decade – Shaping security*. Wien: Bundeskanzleramt [Online]. Available at: https://www.bmlv.gv.at/pdf_pool/publikationen/sicherheitsstrategie_engl.pdf [Accessed: 14 May 2026].

Bundesministerium für Inneres (2024) *Österreichische Sicherheitsstrategie 2024*. Wien: BMI [Online]. Available at: <https://www.bmi.gv.at/502/start.html> [Accessed: 14 May 2026].

Belgium

Premier (n.d.) *Nationale Veiligheidsstrategie*. Brussels: Premier [Online]. Available at: https://premier.be/sites/default/files/articles/NVS_Online_NL.pdf

The Centre for Cybersecurity Belgium (2021). Belgium Cybersecurity Strategy 2.0 2021-2025. Brussels: The Centre for Cybersecurity Belgium [Online]. Available at: <https://dig.watch/resource/belgium-cybersecurity-strategy-2-0-2021-2025> [Accessed: 19 May 2026].

Bulgaria

Ministerstvo na otbranata (n.d.) National Security Strategy of the Republic of Bulgaria. Sofia: Ministerstvo na otbranata [Online]. Available at: https://www.me.government.bg/files/useruploads/files/national_strategy1.pdf [Accessed: 14 May 2026].

Croatia

Vlada Republike Hrvatske (2017) National Security Strategy of the Republic of Croatia. Zagreb: Vlada RH [Online]. Available at: <https://www.soa.hr/files/file/National-Security-Strategy-2017.pdf> [Accessed: 14 May 2026].

Czech Republic

Ministerstvo zahraničních věcí (2015) Security Strategy of the Czech Republic 2015. Prag: MZV [Online]. Available at: https://mzv.gov.cz/public/2a/57/16/1375879_1259981_Security_Strategy_CZ_2015.pdf [Accessed: 14 May 2026].

Ministerstvo zahraničních věcí (2023) Security Strategy of the Czech Republic 2023. Prag: MZV [Online]. Available at: https://mzv.gov.cz/file/5161068/Security_Strategy_of_the_Czech_Republic_2023.pdf [Accessed: 14 May 2026].

Národní úřad pro kybernetickou a informační bezpečnost (n.d.) Cybersecurity Strategy of the Czech Republic. Prag: NÚKIB [Online]. Available at: <https://nukib.gov.cz/en/cyber-security/strategy-action-plan/> [Accessed: 14 May 2026].

Denmark

Digitaliseringsstyrelsen (2021) The Danish National Strategy for Cyber and Information Security. Copenhagen: The Danish Government [Online]. Available at: https://digst.dk/media/bxxcnby2/digst_ncis_2022-2024_uk.pdf [Accessed: 14 May 2026].

The security policy analysis group (2022) Danish Security and Defence towards 2035. Copenhagen: The security policy analysis group [Online]. Available at: https://www.fmn.dk/globalassets/fmn/dokumenter/strategi/rsa/-regeringens_security-policy-report_uk_web-.pdf [Accessed: 14 May 2026].

Estonia

Kaitseministeerium (Estland) (2023) Estonia's National Security Concept 2023. Tallinn: Kaitseministeerium [Online]. Available at: <https://kaitseministeerium.ee/en/policies-and-development-defence-capability/policies/security-policy> [Accessed: 14 May 2026].

Estonian Ministry of Economic Affairs and Communications (2024) Cyber-security Strategy 2024-2030. Tallinn: Estonian Ministry of Economic Affairs and Communications [Online]. Available at: https://www.enisa.europa.eu/sites/default/files/ncss-map/strategies/reports/EE_NCSS_2024_en.pdf [Accessed: 14 May 2026].

Siseministeerium (Estland) (2025) Siseturvalisuse arengukava 2025 / Internal Security Strategy 2020-2030. Tallinn: Siseministeerium [Online]. Available at: https://www.siseministeerium.ee/sites/default/files/documents/2025-01/Siseturvalisuse_arengukava_2020_2030_48lk_final_ENG.pdf [Accessed: 14 May 2026].

EU

European Commission (2010) The EU Internal Security Strategy in Action: Five steps towards a more secure Europe. KOM(2010) 673. Brussels: European Commission [Online]. Available at: <https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2010:0673:FIN:EN:PDF> [Accessed: 14 May 2026].

European External Action Service (EEAS) (2016) A Global Strategy for the European Union's Foreign and Security Policy. Brussels: EEAS [Online]. Available at: https://www.eeas.europa.eu/eeas/global-strategy-european-unions-foreign-and-security-policy_en [Accessed: 14 May 2026].

European External Action Service (EEAS) (2022) Ein Strategischer Kompass für Sicherheit und Verteidigung. Brussels: EAD [Online]. Available at: https://www.eeas.europa.eu/eeas/global-strategy-european-unions-foreign-and-security-policy_en [Accessed: 14 May 2026].

Finland

Turvallisuuskomitea (Finnland) (2017) Yhteiskunnan turvallisuusstrategia / Security Strategy for Society. Helsinki: Turvallisuuskomitea [Online]. Available at: https://turvallisuuskomitea.fi/wp-content/uploads/2018/04/YTS_2017_english.pdf [Accessed: 14 May 2026].

Finland's Prime Minister's Office (2024) Finland's Cyber Security Strategy 2024-2035. Helsinki: Finland's Prime Minister's Office [Online]. Available at: <https://julkaisut.valtioneuvosto.fi/server/api/core/bitstreams/e5575ec3-1ce2-4f0a-aaff-50154789cb6d/content> [Accessed: 14 May 2026].

France

Ministère de la Défense (Frankreich) (2013) French White Paper on Defence and National Security. Paris: Ministère de la Défense [Online]. Available at: <https://ccdcoe.org/uploads/2018/10/White-paper-on-defense-2013-1.pdf> [Accessed: 14 May 2026].

Ministère de la Défense (Frankreich) (2017) Revue stratégique de défense et de sécurité nationale 2017. Paris: SGDSN [Online]. Available at: <https://www.vie-publique.fr/files/rapport/pdf/174000744.pdf> [Accessed: 14 May 2026].

Germany

Bundesregierung (2016) White paper 2016 on German Security Policy and the Future of the Bundeswehr. Berlin: Bundesregierung [Online]. Available at: <https://ccdcoe.org/uploads/2018/10/2016-White-Paper-1.pdf> [Accessed: 14 May 2026].

Bundesregierung (2023) Robust. Resilient. Sustainable. Integrated Security for Germany. National Security Strategy. Berlin: Bundesregierung [Online]. Available at: <https://www.nationalesicherheitsstrategie.de/National-Security-Strategy-EN.pdf> [Accessed: 14 May 2026].

Greece

Ministry of Digital Governance National Cybersecurity Authority (2020) National Cyber Security Strategy of Greece 2020-2025. Athens: Ministry of Digital Governance National Cybersecurity Authority [Online]. Available at: https://www.enisa.europa.eu/sites/default/files/ncss-map/strategies/reports/EL_NCSS_2020_en.pdf [Accessed: 14 May 2026].

Hungary

Hungarian Government (2012) National Security Strategy of Hungary. Budapest: Hungarian Government [Online]. Available at: <https://2010-2014.kormany.hu/download/4/32/b0000/National%20Security%20Strategy.pdf> [Accessed: 14 May 2026].

Hungarian Government (2018) Strategy for the Security of Network and Information Systems in Hungary. Budapest: Hungarian Government [Online]. Available at: <https://nki.gov.hu/wp-content/uploads/2020/11/Strategy-for-the-security-of-network-and-information-systems-in-Hungary.pdf> [Accessed: 14 May 2026].

Ireland

Department of Communications, Energy & Natural Resources (2015) National Cyber Security Strategy 2015–2017. Dublin: NCSC [Online]. Available at: https://www.ncsc.gov.ie/pdfs/National_Cyber_Security_Strategy%202015-2017nb.pdf [Accessed: 14 May 2026].

Government of Ireland (2019) National Cyber Security Strategy 2019–2024. Dublin: NCSC [Online]. Available at: https://ccdcoe.org/uploads/2018/10/Ireland_National_Cyber_Security_Strategy-2019-2024-2.pdf [Accessed: 14 May 2026].

Department of Defence (2024) Defence Policy Review 2024. Dublin: Department of Defence [online]. Available at: <https://www.gov.ie/en/department-of-defence/publications/defence-policy-review-2024/> [Accessed 19 May 2026].

Italy

Ministero della Difesa (2015) White Paper for international security and defence. Rome: Ministero della Difesa [Online]. Available at: https://www.surrey.ac.uk/sites/default/files/2024-09/2015_Italy.pdf [Accessed: 14 May 2026].

Agenzia per la Cybersicurezza Nazionale (2022) National cybersecurity strategy 2022–2026. Rome: ACN [Online]. Available at: https://www.acn.gov.it/portale/documents/20119/531899/ACN_EN_Strategy.pdf [Accessed: 14 May 2026].

Latvia

Aizsardzības ministrija (2020) National Security Concept Riga: Aizsardzības ministrija [Online]. Available at: https://www.mod.gov.lv/sites/mod/files/document/NDK_ENG_final.pdf [Accessed: 14 May 2026].

Aizsardzības ministrija (2023) The State Defence Concept 2023-2027. Riga: Aizsardzības ministrija [Online]. Available at: <https://www.mod.gov.lv/sites/mod/files/document/The%20State%20Defence%20Concept%202023-2027.pdf> [Accessed: 14 May 2026].

Ministru kabinets (2023) The National Concept on Strategic Communication and Security of the Information Space. Riga: Ministru kabinets [Online]. Available at: <https://www.mk.gov.lv/en/media/15446/download> [Accessed: 14 May 2026].

Lithuania

Seimas (2017) National Security Strategy of the Republic of Lithuania. Vilnius: Seimas [Online]. Available at: <https://www.newstrategycenter.ro/wp-content/uploads/2019/07/2017-nacsaugstrategijaen.pdf> [Accessed: 14 May 2026].

Ministry of National Defence (2018) National Cyber Security Strategy of Lithuania 2018. Vilnius: KAM [Online]. Available at: https://ccdcoe.org/uploads/2018/10/Lithuania_Cyber-Security-Strategy-2018_English.pdf [Accessed 14 May 2026].

Valstybės saugumo departamentas (2023) Assessment of Threats to National Security 2023 (published 2023). Vilnius: VSD [Online]. Available at: <https://kam.lt/wp-content/uploads/2023/03/Assessment-of-Threats-to-National-Security-2022-published-2023.pdf> [Accessed: 14 May 2026].

Lietuvos Respublikos Vyriausybė (2025) National Threat Assessment 2025. Vilnius [Online]. Available at: https://kam.lt/wp-content/uploads/2025/03/2025-GR-ENG-02-21-El-be-uzraso_.pdf [Accessed: 14 May 2026].

Malta

Ministeru għall-Ekonomija u l-Industrija (2023) National Cybersecurity Strategy of Malta 2023–2026. Valletta: Ministeru [Online]. Available at: <https://economy.gov.mt/wp-content/uploads/2023/05/National-Cyber-security-2023-2026.pdf> [Accessed: 14 May 2026].

NATO

NATO (2010) Strategic Concept 2010. Brussels: NATO [Online]. Available at: <https://www.nato.int/en/about-us/official-texts-and-resources/strategic-concepts/strategic-concept-2010> [Accessed: 19 May 2026].

NATO (2022) Strategisches Konzept der NATO 2022. Brussels: NATO [Online]. Available at: <https://nato.diplo.de/nato-de/01-natostatements/2539668-2539668> [Accessed: 14 May 2026].

Netherlands

National Coordinator for Counterterrorism and Security Ministry of Justice and Security (2019) National Security Strategy 2019. Den Haag: National Coordinator for Counterterrorism and Security Ministry of Justice and Security [Online]. Available at: <https://english.nctv.nl/documents/publications/2019/09/19/national-security-strategy> [Accessed: 14 May 2026].

Government of the Netherlands (2023) Security Strategy for the Kingdom of the Netherlands 2023-2029. Den Haag: Government of the Netherlands [Online]. Available at: <https://www.government.nl/themes/justice-security-and-defence/security-strategy-for-the-kingdom-of-the-netherlands> [Accessed: 14 May 2026].

Norway

Office of the Prime Minister of Norway (2025) National Security Strategy. Oslo: Office of the Prime Minister of Norway [Online]. Available at: <https://www.regjeringen.no/en/documents/national-security-strategy/id3099304/> [Accessed: 14 May 2026].

Poland

Biuro Bezpieczeństwa Narodowego (2014) National Security Strategy of the Republic of Poland. Warsaw: BBN [Online]. Available at: https://www.bbn.gov.pl/ftp/dok/NSS_RP.pdf [Accessed: 14 May 2026].

Biuro Bezpieczeństwa Narodowego (2020) National Security Strategy of the Republic of Poland 2020. Warsaw: BBN [Online]. Available at: https://www.bbn.gov.pl/ftp/dokumenty/National_Security_Strategy_of_the_Republic_of_Poland_2020.pdf [Accessed: 14 May 2026].

Portugal

Governo de Portugal (Portugal) (n.d.) Conceito Estratégico de Defesa Nacional. Lisbon: MDN [Online]. Available at: https://www.defesa.gov.pt/pt/comunicacao/documentos/Lists/PDEFINTER_DocumentoLookupList/Conceito-Estrategico-de-Defesa-Nacional.pdf [Accessed: 14 May 2026].

Conselho Superior de Segurança do Ciberespaço (2019) National Cyber-space Security Strategy 2019–2023. Lisbon: CSSC [Online]. Available at: https://www.enisa.europa.eu/sites/default/files/ncss-map/strategies/reports/PT_NCSS_2019_en.pdf [Accessed: 14 May 2026].

Romania

Administrația Prezidențială (2019) Strategia Națională de Apărare a Țării pentru perioada 2020–2024. Bucharest: Administrația Prezidențială [Online]. Available at: <https://www.newstrategycenter.ro/wp-content/uploads/2019/07/National-Cybersecurity-Strategy-2019.pdf> [Accessed: 14 May 2026].

Administrația Prezidențială (Rumänien) (2020) National Defence Strategy of Romania 2020–2024. Bucharest: Administrația Prezidențială [Online]. Available at: https://www.presidency.ro/files/userfiles/National_Defence_Strategy_2020_2024.pdf [Accessed: 14 May 2026].

Sweden

Swedish Defence Commission's Secretariat (2023) The Swedish Defence Commission's report on security policy, 2023. Stockholm: Swedish Defence Commission Secretariat [Online]. Available at: <https://www.regeringen.se/contentassets/de808e940116476d8252160c58b78bb7/sammandrag-pa-engelska-av-allvarstid-ds-202319.pdf> [Accessed 19 May 2026].

Regeringskansliet (2024) National Security Strategy of Sweden. Stockholm: Swedish Government [Online]. Available at: <https://www.government.se/globalassets/government/national-security-strategy.pdf> [Accessed: 14 May 2026].

Slovakia

Ministerstvo zahraničných vecí (Slovakia) (2021) Security Strategy of the Slovak Republic. Bratislava: MZV SR [Online]. Available at: <https://www.mzv.sk/documents/30297/4638226/security-strategy-of-the-slovak-republic.pdf> [Accessed: 14 May 2026].

Slovakian National Security Authority (2021) National Cyber Security Strategy of the Slovak Republic 2021-2025. Bratislava: NBÚ [Online]. Available at: <https://www.nbu.gov.sk/national-cyber-security-strategy> [Accessed: 14 May 2026].

Slovenia

Republika Slovenija (2019) Resolution on the National Security Strategy of the Republic of Slovenia. Ljubljana: Vlada RS [Online]. Available at: <https://www.gov.si/assets/ministrstva/MO/Dokumenti/ReSNV2.pdf> [Accessed: 14 May 2026].

Republika Slovenija (2020) White Paper on Defence of the Republic of Slovenia 2020. Ljubljana: Ministrstvo za obrambo [Online]. Available at: <https://www.gov.si/assets/ministrstva/MO/Dokumenti/WP2020.pdf> [Accessed: 14 May 2026].

Republika Slovenija (n.d.) Cyber Security Strategy. Ljubljana: Vlada RS [Online]. Available at: https://www.gov.si/assets/ministrstva/MDP/DID/Cyber_Security_Strategy_Slovenia.pdf [Accessed: 14 May 2026].

Spain

Gobierno de España (2013) The national security strategy. Madrid: Presidencia del Gobierno [Online]. Available at: https://www.lamoncloa.gob.es/documents/estrategiaseguridad_baja_julio.pdf [Accessed: 14 May 2026].

Gobierno de España (Spain) (2017) Estrategia de Seguridad Nacional 2017. Madrid: Presidencia del Gobierno [Online]. Available at: https://www.lamoncloa.gob.es/serviciosdeprensa/notasprensa/presidenciadelgobierno/documents/2017-1824_estrategia_de_seguridad_nacional_esn_doble_pag.pdf [Accessed: 14 May 2026].

Gobierno de España (2019) National Cybersecurity Strategy. Madrid: Presidencia del Gobierno [Online]. Available at: https://www.enisa.europa.eu/sites/default/files/ncss-map/strategies/reports/ES_NCSS_2019_en.pdf

Switzerland

Melani (2019) Nationale Cybersicherheitsstrategien im Vergleich – Herausforderungen für die Schweiz. Bern: MELANI [Online]. Available at: https://www.ncsc.admin.ch/dam/ncsc/de/dokumente/strategie/MELANI_Studie_final_AW_18_maer_2019.pdf [Accessed: 14 May 2026].

Center for Security Studies, ETH Zürich (2024) Bulletin zur Schweizer Sicherheitspolitik 2024. Zürich: ETH Zürich [Online]. Available at: https://ethz.ch/content/dam/ethz/special-interest/gess/cis/center-for-securities-studies/pdfs/Bulletin_2024_05_AW.pdf [Accessed: 14 May 2026].

United Kingdom

Cabinet Office (2010) The national security strategy – a strong Britain in an age of uncertainty. London: Cabinet Office [online]. Available at: <https://www.gov.uk/government/publications/the-national-security-strategy-a-strong-britain-in-an-age-of-uncertainty> [Accessed 19 May 2026].

HM Government (2010) Securing Britain in an Age of Uncertainty: The Strategic Defence and Security Review. London: HM Government [online]. Available at: <https://assets.publishing.service.gov.uk/media/5a78da21ed915d0422065d95/strategic-defence-security-review.pdf> [Accessed 19 May 2026].

HM Government (2015) The National Security Strategy and Strategic Defence and Security Review 2015. London: HM Government [online]. Available at: <https://www.gov.uk/government/publications/national-security-strategy-and-strategic-defence-and-security-review-2015> [Accessed 19 May 2026].

Cabinet Office (2021) The Integrated Review 2021. London: HM Government [online]. Available at: <https://www.gov.uk/government/collections/the-integrated-review-2021> [Accessed 19 May 2026].

HM Government (2025) National Security Strategy 2025. London: HM Government [Online]. Available at: https://assets.publishing.service.gov.uk/media/685ab0da72588f418862075c/E03360428_National_Security_Strategy_Accessible.pdf [Accessed: 14 May 2026].

5 Mikael Weissmann and James Pamment: Governance of the Information Environment – Concrete Measures for Cognitive Security and Cognitive Warfare: The Role of a Ministry of Defence (MoD)

Backes, O. and Swab, A. (2019) ‘Cognitive Warfare: The Russian Threat to Election Integrity in the Baltic States. Belfer Center for Science and International Affairs, Harvard Kennedy School. Cambridge, MA. Available at: https://www.belfercenter.org/sites/default/files/pantheon_files/2019-11/CognitiveWarfare.pdf [Accessed: 14 May 2026].

Bay, S., Bertolin, G., Biteniece, N., Christie, E., Dek, A., Fredheim, R., Gallacher, J. D., Kononova, K. and Marchenko, T. (2019) Responding to Cognitive Security Challenges. Riga: NATO Strategic Communications Centre of Excellence.

Claverie, B. and Du Cluzel, F. (2022) ‘“Cognitive Warfare”: The Advent of the Concept of “Cognitics” in the Field of Warfare’, in Claverie, B. et al. (eds.) Cognitive Warfare: The Future of Cognitive Dominance. NATO Collaboration Support Office, pp. 2:1–2:8.

Falkheimer, J. and Pamment, J. (eds.) (2025) *Psychological Defence and Information Influence: A Textbook on Theory and Practice*. Stockholm: Psychological Defence Agency.

Grahn, H. and Taipalus, T. (2025) 'Defining Comprehensive Cognitive Security in the Digital Era: Literature Review and Concept Analysis', *Journal of Information Warfare*, 24(2), pp. 39–59.

Hybrid CoE (n.d.) 'Hybrid Threat as a Concept' [Online]. Available at: <https://www.hybridcoe.fi/hybrid-threats-as-a-phenomenon/> [Accessed 11 March 2026].

Metzinger, T. (ed.) (2009) *Grundkurs Philosophie des Geistes – Band 1: Phänomenales Bewusstsein*. Paderborn: Mentis.

NATO ACT (2023) 'Cognitive Warfare: Strengthening and Defending the Mind' [Online]. Available at: <https://www.act.nato.int/article/cognitive-warfare-strengthening-and-defending-the-mind/> [Accessed: 14 May 2026].

NATO ACT (n.d.) 'Cognitive Warfare' [Online]. Available at: <https://www.act.nato.int/activities/cognitive-warfare/> [Accessed: 14 May 2026].

NATO Science and Technology Organisation (2025) *Cognitive Warfare: NATO Chief Scientist Research Report*. Brussels: STO – NATO Science and Technology Organisation. Available at: <https://www.sto.nato.int/wp-content/uploads/chief-scientist-report-cognitive-warfare-final.pdf> [Accessed 10 March 2026].

Nilsson, N., Weissmann, M. and Palmertz, B. (2026) 'Hybrid Threats and the Intelligence Community: Priming for a Volatile Age', *International Journal of Intelligence and CounterIntelligence*, 39(1), pp. 109–131. Available at: <https://doi.org/10.1080/08850607.2024.2435265> [Accessed: 14 May 2026].

Pamment, J. and Isaksson, E. (2024) *Psychological Defence: Concepts and Principles for the 2020s*. MPF Report Series 6/2024. Karlstad: Psychological Defence Agency.

S90500/62-GDVPol/2025 (2025) Hybride Bedrohungen und die Kognitive Dimension: Verteidigungspolitische Begriffsbestimmung. Unpublished manuscript.

Waltzman, R. (2017) 'The Weaponization of Information: The Need for Cognitive Security'. RAND Corporation. Santa Monica, CA. Available at: https://www.rand.org/content/dam/rand/pubs/testimonies/CT400/CT473/RAND_CT473.pdf [Accessed: 14 May 2026].

Weissmann, M. (2021) 'Conceptualising and Countering Hybrid Threats and Hybrid Warfare: The Role of the Military in the Grey Zone', in Weissmann, M. et al. (eds.) Hybrid Warfare. London: I.B. Tauris, pp. 61–82.

Weissmann, M. (2024) 'Framtida Hotbilders Påverkan För Säkerhetstjänsterna', in Framtidens Säkerhetstjänst i Totalförsvaret. Swedish Defence University Report Series, pp. 38–67. Available at: <https://fhs.diva-portal.org/smash/get/diva2:1846156/FULLTEXT01.pdf> [Accessed: 14 May 2026].

Weissmann, M., Nilsson, N. and Palmertz, B. (2021) 'Moving Out of the Blizzard', in Weissmann, M. et al. (eds.) Hybrid Warfare. London: I.B. Tauris.

Weissmann, M., Nilsson, N., Palmertz, B. and Thunholm, P. (eds.) (2021) Hybrid Warfare. London: I.B. Tauris.

6 André K. Rodarte und Lena Frischlich: Preservation and Promotion of Democratic Resilience – A Scientific Basis for Policy Recommendations

AFP (2026) 'Tightened Restrictions Stifle Press across Middle East, Fogging Coverage of War.', The Times of Israel [Online]. Available at: <https://www.timesofisrael.com/tightened-restrictions-stifle-press-across-middle-east-fogging-coverage-of-war/> [Accessed 27 March 2026].

Akkerman, A., Mudde, C. and Zaslove, A. (2014) 'How Populist Are the People? Measuring Populist Attitudes in Voters.' *Comparative Political Studies*, 47 (9), pp. 1324–1353. Available at: <https://doi.org/10/f6dp5p>.

Al Jazeera (2023) 'Assault on Free Speech': India Editors Slam 'Fake News' Proposal', Al Jazeera [Online]. Available at: <https://www.aljazeera.com/news/2023/1/19/vile-censorship-india-proposed-fake-news-law-for-social-media-egi> [Accessed 27 March 2026].

Allport, G. W. and Postman, L. (1946) 'An Analysis of Rumor.' *Public Opinion Quarterly*, 10 (4), pp. 501–517. Available at: <https://doi.org/10/c5bm3h> [Accessed: 14 May 2026].

Alon, A., Rahimi, I. and Tahar, H. (2024) 'Fighting Fake News on Social Media: A Comparative Evaluation of Digital Literacy Interventions.' *Current Psychology*, 43 (19), pp. 17343–17361. Available at: <https://doi.org/10.1007/s12144-024-05668-4>.

Amnesty International (2026) 'Iran: Authorities Unleash Heavily Militarized Clampdown to Hide Protest Massacres' Amnesty International [Online]. Available at: <https://www.amnesty.org/en/latest/news/2026/01/iran-authorities-unleash-heavily-militarized-clampdown-to-hide-protest-massacres/> [Accessed 6 March 2026].

Andersen, K., Shehata, A. and Andersson, D. (2023) 'Alternative News Orientation and Trust in Mainstream Media: A Longitudinal Audience Perspective' *Digital Journalism*, 11 (5), pp. 833–852. Available at: <https://doi.org/10.1080/21670811.2021.1986412> [Accessed: 14 May 2026].

Aro, J. (2016) 'The Cyberspace War: Propaganda and Trolling as Warfare Tools.' *European View*, 15 (1) June, pp. 121–132. Available at: <https://doi.org/10.1007/s12290-016-0395-5> [Accessed: 14 May 2026].

Basmat, D. (2025) 'Over 119,000 Russian Soldiers Killed in Ukraine Identified by Media Investigation' *The Kyiv Independent* [Online]. Available at: <https://kyivindependent.com/over-119-000-russian-soldiers-killed-in-ukraine-identified-by-media-investigation/> [Accessed 27 March 2026].

Beißert, H., Köhler, M., Rempel, M. and Beierlein, C. (2014) 'Eine Deutschsprachige Kurzsкала Zur Messung Des Konstrukts Need for Cognition [A German Language Short-Scale for Measuring the Construct Need for Cognition].' *GESIS- Working Papers*, 32.

Blair, R. A., Gottlieb, J., Nyhan, B., Paler, L., Argote, P. and Stainfield, C. J. (2024) 'Interventions to Counter Misinformation: Lessons from the Global North and Applications to the Global South.' *Current Opinion in Psychology*, 55 February, p. 101732. Available at: <https://doi.org/10.1016/j.copsyc.2023.101732> [Accessed: 14 May 2026].

Boman, C. D. (2021) 'Examining Characteristics of Prebunking Strategies to Overcome PR Disinformation Attacks.' *Public Relations Review* [Online], 47 (5). Available at: <https://doi.org/10.1016/j.pubrev.2021.102105>.

Boulianne, S. and Humprecht, E. (2023) 'Perceived Exposure to Misinformation and Trust in Institutions in Four Countries before and during a Pandemic.' *International Journal of Communication*, 17, pp. 1–24.

Bruder, M., Haffke, P., Neave, N., Nouripanah, N. and Imhoff, R. (2013) 'Measuring Individual Differences in Generic Beliefs in Conspiracy Theories across Cultures: Conspiracy Mentality Questionnaire.' *Frontiers in Psychology*, 4 (April), pp. 1–15. Available from: <https://doi.org/10/gbftpj> [Accessed: 14 May 2026].

Burton, J. W., Herzog, S. M. and Lorenz-Spreen, P. (2024) 'Simple Changes to Content Curation Algorithms Affect the Beliefs People Form in a Collaborative Filtering Experiment' [Online]. Available at: <https://doi.org/10.31234/osf.io/5yfbt> [Accessed: 14 May 2026].

Cha, M., Cha, C., Singh, K., Lima, G., Ahn, Y.-Y., Kulshrestha, J. and Varol, O. (2021) 'Prevalence of Misinformation and Factchecks on the COVID-19 Pandemic in 35 Countries: Observational Infodemiology Study.' *JMIR Human Factors*, 8 (1) February, p. e23279. Available at: <https://doi.org/10.2196/23279> [Accessed: 14 May 2026].

Chen, X. K., Na, J.-C., Tan, L. K.-W., Chong, M. and Choy, M. (2022) 'Exploring How Online Responses Change in Response to Debunking Messages about COVID-19 on WhatsApp.' *Online Information Review*, 46 (6), pp 1184–1204. Available at: <https://doi.org/10.1108/OIR-08-2021-0422> [Accessed: 14 May 2026].

Christopher, P. and Matthews, M. (2016) 'The Russian "Firehose of Falsehood"' *Propaganda Model*. 2 (7), pp. 1–10.

Council of the European Union (2022) EU Imposes Sanctions on State-Owned Outlets RT/Russia Today and Sputnik's Broadcasting in the EU [Online]. Council of the European Union. Available at: <https://www.consilium.europa.eu/en/press/press-releases/2022/03/02/eu-imposes-sanctions-on-state-owned-outlets-rtrussia-today-and-sputnik-s-broadcasting-in-the-eu/> [Accessed 30 March 2026].

Craig, M. and Vijaykumar, S. (2023) 'One Dose Is Not Enough: The Beneficial Effect of Corrective COVID-19 Information Is Diminished If Followed by Misinformation.' *Social Media + Society*, 9 (2), p. 20563051231161298. Available at: <https://doi.org/10.1177/20563051231161298> [Accessed: 14 May 2026].

Denisa-Liepniece, S. (2021) Report "Media Literacy Sector Mapping in Georgia, Latvia, Moldova and Ukraine". *Baltijas Mediju izcilibas centrs* [Online blog]. Available at: <https://bcme.eu/en/research/report-media-literacy-sector-mapping-in-georgia-latvia-moldova-and-ukraine/> [Accessed 10 April 2026].

Duckitt, J. and Sibley, C. G. (2010) 'Right-Wing Authoritarianism and Social Dominance Orientation Differentially Moderate Intergroup Effects on Prejudice.' *European Journal of Social Psychology*, 24 (November 2009), pp. 583–601. Available at: <https://doi.org/10.1002/per.772> [Accessed: 14 May 2026].

Dutsyk, D. (2022) Transformation of Media Literacy Sphere in Full-Scale War in Ukraine [Online]. Ukrainian Institute of Media and Communication, pp. 1–24. Available at: <https://www.jta.com.ua/news-and-reports/uimk-pidhotuvav-analitychnyy-zvit-pro-transformatsiiu-sfery-mediahramotnosti/> [Accessed 19 January 2026].

Ecker, U., Roozenbeek, J., Linden, S. van der, Tay, L. Q., Cook, J., Oreskes, N. and Lewandowsky, S. (2024) 'Misinformation Poses a Bigger Threat to Democracy than You Might Think.' *Nature*, 630 (8015) June, pp. 29–32. Available at: <https://doi.org/10.1038/d41586-024-01587-3> [Accessed: 14 May 2026].

European Commission (2025a) *Defending Democratic Values in the Digital Age | Shaping Europe's Digital Future* [Online]. Available at: <https://digital-strategy.ec.europa.eu/en/policies/democracy-digital> [Accessed 10 April 2026].

European Commission, (n.d.) *Media Literacy* [Online]. Available at: <https://digital-strategy.ec.europa.eu/en/policies/media-literacy> [Accessed 27 March 2026a].

European Commission, (n.d.) *Revision of the Audiovisual Media Services Directive – AVMSD | Shaping Europe's Digital Future* [Online]. Available at: <https://digital-strategy.ec.europa.eu/en/policies/revision-avmsd> [Accessed 27 March 2026b].

Faragó, L., Kende, A. and Krekó, P. (2020) 'We Only Believe in News That We Doctored Ourselves: The Connection between Partisanship and Political Fake News.' *Social Psychology*, 51 (2) March, pp. 77–90. Available at: <https://doi.org/10/gf876t> [Accessed: 14 May 2026].

Ferrucci, P. and Hopp, T. (2023) 'Let's Intervene: How Platforms Can Combine Media Literacy and Self-Efficacy to Fight Fake News.' *Communication and the Public*, 8 (4), pp. 367–389. Available at: <https://doi.org/10.1177/20570473231203081> [Accessed: 14 May 2026].

Field, A. and Graham, H. (2002) *How to Design and Report Experiments*. London: Sage. Freelon, D. and Wells, C. (2020) *Disinformation as Political Communication*. *Political Communication*, 37 (2) March, pp. 145–156. Available at: <https://doi.org/10.1080/10584609.2020.1723755> [Accessed: 14 May 2026].

Freiling, I., Krause, N. M., Scheufele, D. A. and Brossard, D. (2023) 'Believing and Sharing Misinformation, Fact-Checks, and Accurate Information on Social Media: The Role of Anxiety during COVID-19.' *New Media & Society*, 25, pp. 141–162. Available at <https://doi.org/10.1177/14614448211011451> [Accessed: 14 May 2026].

Frischlich, L., Boberg, S. and Quandt, T. (2019) 'Comment Sections as Targets of Dark Participation? Journalists' Evaluation and Moderation of Deviant User Comments.' *Journalism Studies*, 20 (14), pp. 2014–2033. Available at: <https://doi.org/https://doi.org/10.1080/1461670X.2018.1556320> [Accessed: 14 May 2026].

Frischlich, L. and Humprecht, E. (2021) *Trust, Democratic Resilience, and the Infodemic. Facing up to the infodemic: Promoting a fact- based public discourse in times of crisis.* Tel Aviv, Israel: Israel Public Policy Institute.

Frischlich, L., Kuhfeldt, L., Schatto-Eckrodt, T. and Clever, L. (2023) 'Alternative Counter- News Use and Fake News Recall during the COVID-19 Crisis.' *Digital Journalism*, 11 (1), pp. 80–102. Available at: <https://doi.org/https://doi.org/10.1080/21670811.2022.2106259> [Accessed: 14 May 2026].

Frischlich, L., Olsson, H., Roy, A., Schulze, H., Rhodes, S. and Mansheim, A. (2025) 'The Complexity of Misinformation Extends beyond Virus and Warfare Analogies.' *npj Complexity*, 2 (1) October, p. 29. Available at: <https://doi.org/10.1038/s44260-025-00053-z> [Accessed: 14 May 2026].

Frischlich, L., Schatto-Eckrodt, T. and Clever, L. (in press) 'Democratic Resilience in Times of Dark Participation'. in: Anastasidis, M. and Thimm, C. ed., *Quo vadis digital democracy?* Springer VS.

Ghosh, J. and Acharya, A. (2011) 'Cluster Ensembles.' *Wiley Int. Rev. Data Min. and Knowl. Disc.*, 1 (4) July, pp. 305–315. Available at: <https://doi.org/10.1002/widm.32> [Accessed: 14 May 2026].

Gimpel, H., Heger, S., Olenberger, C. and Utz, L. (2021) 'The Effectiveness of Social Norms in Fighting Fake News on Social Media.' *Journal of Management Information Systems*, 38 (1), pp. 196–221. Available at: <https://doi.org/10.1080/07421222.2021.1870389> [Accessed: 14 May 2026].

Glapiak, A. (2023) 'Implications of the War in Ukraine for the Strategic Communication System of the Polish Ministry of National Defence.' *Security and Defence Quarterly*, 43 (3) September, pp. 22–45. Available at: <https://doi.org/10.35467/sdq/173070> [Accessed: 14 May 2026].

Golovchenko, Y., Hartmann, M. and Adler-Nissen, R. (2018) 'State, Media and Civil Society in the Information Warfare over Ukraine: Citizen Curators of Digital Disinformation.' *International Affairs*, 94 (5) September, pp. 975–994. Available at: <https://doi.org/10.1093/ia/iiy148> [Accessed: 14 May 2026].

Gounaridis, D. and Newell, J. P. (2024) 'The Social Anatomy of Climate Change Denial in the United States.' *Scientific Reports [Online]*, 14 (1). Available at: <https://doi.org/10.1038/s41598-023-50591-6> [Accessed: 14 May 2026].

Grenier, E. (2026) 'Iran Strikes Highlight Dubai Influencers' Free Speech Limits.' *Deutsche Welle [Online]*. Available at: <https://www.dw.com/en/iran-strikes-highlight-dubai-influencers-free-speech-limits/a-76202896> [Accessed 6 March 2026].

Günther, E. and Quandt, T. (2016) 'Word Counts and Topic Models: Automated Text Analysis Methods for Digital Journalism Research.' *Digital Journalism*, 4 (1) January, pp. 75–88. Available at: <https://doi.org/10/gf3g9j> [Accessed: 14 May 2026].

Hainmueller, J., Hopkins, D. J. and Yamamoto, T. (2014) 'Causal Inference in Conjoint Analysis: Understanding Multidimensional Choices via Stated Preference Experiments.' *Political Analysis*, 22 (1), pp. 1–30. Available at: <https://doi.org/10.1093/pan/mpt024> [Accessed: 14 May 2026].

Hancock, T. (2020) 'Three Essays on Malicious Consumer Deviance: The Creation, Dissemination, and Elimination of Misleading Information.' *Dissertation Abstracts International Section A: Humanities and Social Sciences*, 81 (11-A), p. No-Specified.

hateaid (2025) Press Release: 'US Entry Ban on HateAid CEOs. HateAid, 24 December' [Online blog]. Available at: <https://hateaid.org/en/press-release-us-entry-ban-on-hateaid-ceos/> [Accessed 1 April 2026].

Henley, J. (2018) 'Sweden Distributes 'be Prepared for War' Leaflet to All 4.8m Homes.' *The Guardian* [Online]. Available at: <https://www.theguardian.com/world/2018/may/21/sweden-distributes-be-prepared-for-war-cyber-terror-attack-leaflet-to-every-home> [Accessed 27 March 2026].

Hertwig, R. and Lorenz-Spreen, P. (2024) 'People in 26 Countries Prefer Individual Control to Commercial or Governmental Control over Online Choice Architectures' [Online]. Available at: <https://doi.org/10.31219/osf.io/haqu9> [Accessed 15 May 2026].

Hoes, E., Aitken, B., Zhang, J., Gackowski, T. and Wojcieszak, M. (2024) 'Prominent Misinformation Interventions Reduce Misperceptions but Increase Scepticism.' *Nature Human Behaviour*, June, pp. 1–9. Available at: <https://doi.org/10.1038/s41562-024-01884-x> [Accessed: 14 May 2026].

Humprecht, E. (2019) 'Where 'Fake News' Flourishes: A Comparison across Four Western Democracies.' *Information Communication and Society*, 22 (13), pp. 1973–1988. Available at: <https://doi.org/10.1080/1369118X.2018.1474241> [Accessed: 14 May 2026].

Humprecht, E., Esser, F. and Van Aelst, P. (2020) 'Resilience to Online Disinformation: A Framework for Cross-National Comparative Research.' *The International Journal of Press/Politics*, 25 (3) January, pp. 1–24. Available at: <https://doi.org/10/ggjk22> [Accessed: 14 May 2026].

Hunter, L. (2024) 'Regime Characteristics and Online Government Disinformation.' *Journal of Information Technology & Politics*. Available at: <https://doi.org/10.1080/19331681.2024.2445730> [Accessed: 14 May 2026].

Journalismus Macht Schule (n.d.) [Online]. Available at: <https://journalismus-macht-schule.org/> [Accessed: 14 May 2026].

Jungherr, A. (2024) 'Foundational Questions for the Regulation of Digital Disinformation.' *Journal of Media Law*, June, pp. 1–10. Available at: <https://doi.org/10.1080/17577632.2024.2362484> [Accessed: 14 May 2026].

Jungherr, A. and Rauchfleisch, A. (2024) 'Negative Downstream Effects of Alarmist Disinformation Discourse: Evidence from the United States.' *Political Behaviour* [Online], January. Available at: <https://doi.org/10.1007/s11109-024-09911-3> [Accessed: 14 May 2026].

Katz, E., Blumler, J. G. and Gurevitch, M. (1973) Uses and Gratifications Research. *The Public Opinion Quarterly*, 37 (4), pp. 509–523.

Kim, Y. and Lim, H. (2025) 'Alleviating the Bandwagon Effect of Crisis Misinformation on Social Media: Understanding Social Media Users' Bandwagon Perceptions and the Credibility of Crisis Misinformation to Protect Organisational Reputation.' *Communication Studies* [Online]. Available at: <https://doi.org/10.1080/10510974.2025.2485368> [Accessed: 14 May 2026].

Kirat, M. and Slamene, R. (2024) Framing Wars and Conflicts: Manipulation, Misrepresentation, and Distortion. *Migration Letters*, 21 (3) January, pp. 985–1000.

Klinger, U., Kreiss, D. and Mutsvairo, B. (2024) *Platforms, Power, Politics: An Introduction to Political Communication in the Digital Age*. Cambridge, MA, USA: Polity Press.

Klingová, K. (2024) 'War on Words: How to Communicate Ukraine?' [Online]. GLOBSEC. Available at: <https://www.globsec.org/what-we-do/publications/war-words-how-communicate-ukraine> [Accessed 27 March 2026].

Kont, J., Elving, W., Broersma, M. and Bozdauğ, Ç. (2025) 'What Makes Audiences Resilient to Disinformation? Integrating Micro, Meso, and Macro Factors Based on a Systematic Literature Review.' *Communications*, 50 (2), pp. 534–555. Available at: <https://doi.org/10.1515/commun-2023-0078> [Accessed: 14 May 2026].

Kozyreva, A., Lorenz-Spreen, P., Herzog, S. M., Ecker, U. K. H., Lewandowsky, S., Hertwig, R., Ali, A., Bak-Coleman, J., Barzilai, S., Basol, M., Berinsky, A. J., Betsch, C., Cook, J., Fazio, L. K., Geers, M., Guess, A. M., Huang, H., Larreguy, H., Maertens, R., Panizza, F., Pennycook, G., Rand, D. G., Rathje, S., Reifler, J., Schmid, P., Smith, M., Swire-Thompson, B., Szewach, P., Linden, S. van der and Wineburg, S. (2024) 'Toolbox of Individual-Level Interventions against Online Misinformation.' *Nature Human Behaviour*, 8 (6), pp. 1044–1052. Available at: <https://doi.org/10.1038/s41562-024-01881-0> [Accessed: 14 May 2026].

Kumar, R. (2025) "“Truth Is the Casualty”: How Indian Fact-Checkers Debunked False Claims during the India-Pakistan Crisis' Reuters Institute for the Study of Journalism [Online]. Available at: <http://reutersinstitute.politics.ox.ac.uk/news/truth-casualty-how-indian-fact-checkers-debunked-false-claims-during-india-pakistan-crisis> [Accessed 10 April 2026].

Lacy, S., Watson, B. R., Riffe, D. and Lovejoy, J. (2015) 'Issues and Best Practices in Content Analysis.' *Journalism & Mass Communication Quarterly*, 92 (4) December, pp. 791–811. Available at: <https://doi.org/10.1177/1077699015607338> [Accessed: 14 May 2026].

Landis, J. R. and Koch, G. G. (1977) 'The Measurement of Observer Agreement for Categorical Data.' *Biometrics*, 33 (1), pp. 159–174. Available at: <https://doi.org/10.2307/2529310> [Accessed: 14 May 2026].

Lanius, C., Weber, R. and MacKenzie, W. I. (2021) 'Use of Bot and Content Flags to Limit the Spread of Misinformation among Social Networks: A Behaviour and Attitude Survey.' *Social Network Analysis and Mining* [Online], 11 (1). Available at: <https://doi.org/10.1007/s13278-021-00739-x> [Accessed: 14 May 2026].

Lee, J., Kim, J. W. and Yun Lee, H. (2023) 'Unlocking Conspiracy Belief Systems: How Fact-Checking Label on Twitter Counters Conspiratorial MMR Vaccine Misinformation.' *Health Communication*, 38 (9), pp. 1780–1792. Available at: <https://doi.org/10.1080/10410236.2022.2031452> [Accessed: 14 May 2026].

Lee, S. (2024) 'Beyond the Information given: Factors Intensifying or Attenuating Individuals' Susceptibility to Misinformation.' Dissertation Abstracts International Section A: Humanities and Social Sciences, 86 (5-A), p. No-Specified.

Lee, Y.-I., Mu, D., Hsu, Y.-C., Wojdyski, B. W. and Binford, M. (2024) 'Misinformation or Hard to Tell? An Eye-Tracking Study to Investigate the Effects of Food Crisis Misinformation on Social Media Engagement' *Public Relations Review* [Online], 50 (4). Available at: <https://doi.org/10.1016/j.pubrev.2024.102483> [Accessed: 14 May 2026].

Lemaire, M., Ye, S., Le Stanc, L., Borst, G. and Cassotti, M. (2025) 'The Development of Media Truth Discernment and Fake News Detection Is Related to the Development of Reasoning during Adolescence.' *Scientific Reports* [Online], 15 (1). Available at: <https://doi.org/10.1038/s41598-025-90427-z> [Accessed: 14 May 2026].

Lewandowsky, S. and Linden, S. van der (2021) 'Countering Misinformation and Fake News through Inoculation and Prebunking' *European Review of Social Psychology*, 32 (2) July, pp. 348–384. Available at: <https://doi.org/10.1080/10463283.2021.1876983> [Accessed: 14 May 2026].

Li, J. and Yang, X. (2024) 'Does Exposure Necessarily Lead to Misbelief? A Meta-Analysis of Susceptibility to Health Misinformation.' *Public Understanding of Science*, August, p. 09636625241266150. Available at: <https://doi.org/10.1177/09636625241266150> [Accessed: 14 May 2026].

Livingston, S. and Miller, M. eds. (2025) 'Connective Action and the Rise of the Far-Right: Platforms, Politics, and the Crisis of Democracy' [Online]. Oxford University Press. Available at: <https://doi.org/10.1093/oso/9780197794937.001.0001> [Accessed: 14 May 2026].

Lukito, J., Suk, J., Zhang, Y., Doroshenko, L., Kim, S. J., Su, M.-H., Xia, Y., Free-lon, D. and Wells, C. (2019) 'The Wolves in Sheep's Clothing: How Russia's Internet Research Agency Tweets Appeared in U.S. News as Vox Populi.' *The International Journal of Press/Politics*, December, p. 194016121989521. Available at: <https://doi.org/10/ggh6fw> [Accessed: 14 May 2026].

Maertens, R., Götz, F. M., Golino, H. F., Roozenbeek, J., Schneider, C. R., Kyrychenko, Y., Kerr, J. R., Stieger, S., McClanahan, W. P., Drabot, K., He, J. and Van Der Linden, S. (2023) 'The Misinformation Susceptibility Test (MIST): A Psychometrically Validated Measure of News Veracity Discernment.' *Behaviour Research Methods*, 56 (3) June, pp. 1863–1899. Available at: <https://doi.org/10.3758/s13428-023-02124-2> [Accessed: 14 May 2026].

Mahdi, A., Farah, M. F. and Ramadan, Z. (2022) 'What to Believe, Whom to Blame, and When to Share: Exploring the Fake News Experience in the Marketing Context.' *Journal of Consumer Marketing*, 39 (3), pp. 306–316. Available at: <https://doi.org/10.1108/JCM-05-2020-3863> [Accessed: 14 May 2026].

Mandić, J. and Klarić, D. (2023) 'Case Study of the Russian Disinformation Campaign during the War in Ukraine – Propaganda Narratives, Goals, and Impacts.' *National Security and the Future*, 24 (2) July, pp. 97–140. Available at: <https://doi.org/10.37458/nstf.24.2.5> [Accessed: 14 May 2026].

Marwick, A. and Lewis, R. (2017) *Media Manipulation and Disinformation Online* [Online]. Data & Society Research Institute. Available at: <https://datasociety.net/library/media-manipulation-and-disinfo-online/>.

McCarthy, P. (2026) 'European Democracy Shield: A Threat to Free Speech and Democratic Accountability' [Online]. The Heritage Foundation. Available at: <https://www.heritage.org/global-politics/report/european-democracy-shield-threat-free-speech-and-democratic-accountability> [Accessed 6 March 2026].

Meiske, B., Álvarez-Benjumea, A., Andrighetto, G. and Polizzi, E. (2024) 'Nudging Punishment against Sharing of Fake News.' *European Economic Review* [Online], 168. Available at: <https://doi.org/10.1016/j.eurocorev.2024.104795> [Accessed: 14 May 2026].

Merkel, W. and Lührmann, A. (2021) 'Resilience of Democracies: Responses to Illiberal and Authoritarian Challenges.' *Democratization*, 28 (5) July, pp. 869–884. Available at: <https://doi.org/10.1080/13510347.2021.1928081> [Accessed: 14 May 2026].

META (2025) 'Testing Begins for Community Notes on Facebook, Instagram and Threads.' Meta Newsroom, [Online blog]. Available at: <https://about.fb.com/news/2025/03/testing-begins-community-notes-facebook-instagram-threads/> [Accessed 10 April 2026].

Miller-Graff, L. E. (2022) 'The Multidimensional Taxonomy of Individual Resilience.' *Trauma, Violence, & Abuse*, 23 (2), pp. 660–675. Available at: <https://doi.org/DOI:%2010.1177/1524838020967329> [Accessed: 14 May 2026].

Mimizuka, K., Brown, M. A., Yang, K.-C. and Lukito, J. (2025) 'Post-Post-API Age: Studying Digital Platforms in Scant Data Access Times' [Online]. arXiv. Available at: <https://doi.org/10.48550/arXiv.2505.09877> [Accessed: 14 May 2026].

Mimizuka, K., Rodarte, A. K. and Arif, A. (2026) 'On the Fly: How Japanese Social Media "Watchers" Improvise to Counter Problematic Information.' *New Media & Society*, 28 (2) February, pp. 720–738. Available at: <https://doi.org/10.1177/14614448241302311> [Accessed: 14 May 2026].

Mourali, M. and Drake, C. (2022) 'The Challenge of Debunking Health Misinformation in Dynamic Social Media Conversations: Online Randomised Study of Public Masking During COVID-19.' *Journal of Medical Internet Research* [Online], 24 (3). Available at: <https://doi.org/10.2196/34831> [Accessed: 14 May 2026].

Munn, Z., Peters, M. D. J., Stern, C., Tufanaru, C., McArthur, A. and Aromataris, E. (2018) 'Systematic Review or Scoping Review? Guidance for Authors When Choosing between a Systematic or Scoping Review Approach.' *BMC Medical Research Methodology*, 18 (1) December, p. 143. Available at: <https://doi.org/10.1186/s12874-018-0611-x> [Accessed: 14 May 2026].

Munro, I. and Mackay, B. (2019) Information Warfare [Online]. In: *The Blackwell Encyclopaedia of Sociology*. John Wiley & Sons, Ltd, pp. 1–3. Available from: <https://doi.org/10.1002/9781405165518.wbeos1487> [Accessed: 14 May 2026].

Neuendorf, K. A. (2001) *The Content Analysis Guidebook*. [Online]. Available at: https://books.google.de/books?hl=de&lr=&id=nMA5DQAAQBAJ&oi=fnd&pg=PP1&dq=n euendorf+2002+content+analysis+guidebook&ots=pGOor9jxat&sig=w5UcX3Zs- 5lOrJdYUXbw8S0_Coo#v=onepage&q=How Content Analysis Is Done&f=false [Accessed: 14 May 2026].

Nyhan, B. (2020) 'Facts and Myths about Misperceptions.' *Journal of Economic Perspectives*, 34 (3) August, pp. 220–236. Available at: <https://doi.org/10/gg6z5c> [Accessed: 14 May 2026].

OECD (2017) *OECD Guidelines on Measuring Trust* [Online]. OECD. Available at: <https://doi.org/10.1787/9789264278219-en> [Accessed: 14 May 2026].

Omar, B., Apuke, O. and Nor, Z. (2024) 'The Intrinsic and Extrinsic Factors Predicting Fake News Sharing among Social Media Users: The Moderating Role of Fake News Awareness.' *Current Psychology*, 43 (2), pp. 1235–1247. Available at: <https://doi.org/10.1007/s12144-02304343-4> [Accessed: 14 May 2026].

Pennycook, G., Epstein, Z., Mosleh, M., Arechar, A. A., Eckles, D. and Rand, D. G. (2021) 'Shifting Attention to Accuracy Can Reduce Misinformation Online.' *Nature*, 592 (7855), pp. 590–595. Available at: <https://doi.org/10.1038/s41586-021-03344-2> [Accessed: 14 May 2026].

Pennycook, G. and Rand, D. G. (2020) 'Who Falls for Fake News? The Roles of Bullshit Receptivity, Overclaiming, Familiarity, and Analytic Thinking.' *Journal of Personality*, 88 (2), pp. 185–200. Available at: <https://doi.org/10.1111/jopy.12476> [Accessed: 14 May 2026].

Pennycook, G. and Rand, D. G. (2021) 'The Psychology of Fake News.' *Trends in Cognitive Sciences*, 25 (5) May, pp. 388–402. Available at: <https://doi.org/10.1016/j.tics.2021.02.007> [Accessed: 14 May 2026].

Perse, E. M. and Lambe, J. (2016) *Media Effects and Society*. New York: Routledge.

Pfau, M. and Szabo, E. A. (2003) 'Inoculation and Resistance to Persuasion.' in: Seiter, J. S. and Gass, R. H. ed., *Perspectives on Persuasion, Influence and Compliance Gaining*. Boston, USA: Allyn & Bacon, pp. 265–286.

Prike, T., Butler, L. H. and Ecker, U. K. H. (2024) 'Source-Credibility Information and Social Norms Improve Truth Discernment and Reduce Engagement with Misinformation Online.' *Scientific Reports* [Online], 14 (1). Available at: <https://doi.org/10.1038/s41598-024-57560-7> [Accessed 15 May 2026].

Pummerer, L., Böhm, R., Lilleholt, L., Winter, K., Zettler, I. and Sassenberg, K. (2022) 'Conspiracy Theories and Their Societal Effects during the COVID-19 Pandemic.' *Social Psychological and Personality Science*, 13 (1), pp. 49–59. Available from: <https://doi.org/https://doi.org/10.1177/19485506211000217> [Accessed 15 May 2026].

Qerimi, G. and Gërguri, D. (2022) 'Infodemic and the Crisis of Distinguishing Disinformation from Accurate Information: Case Study on the Use of Facebook in Kosovo during COVID-19.' *Information and Media*, 94, pp. 87–109. Available from: <https://doi.org/10.15388/IM.2021.94.56> [Accessed 15 May 2026].

Quandt, T., Frischlich, L., Boberg, S. and Schatto-Eckrodt, T. (2019) 'Fake News'[Online]. in: Vos, T. P. and Hanusch, F. ed., *The International Encyclopedia of Journalism Studies*. Wiley, pp. 1–6. Available at: <https://doi.org/10.1002/9781118841570.iejs0128> [Accessed 15 May 2026].

Quito, A. (2020) 'Taiwan Is Using Humor as a Tool against Coronavirus Hoaxes' *Quartz* [Online]. Available at: <https://qz.com/1863931/taiwan-is-using-humor-to-quash-coronavirus-fake-news> [Accessed 20 February 2026].

Rasheed, M. and Naseer, M. (2021) 'Digital Disinformation & Domestic Disturbance: Hostile Cyber-Enabled Information Operations to Exploit Domestic Issues on Twitter.' *IPRI journal*, 21 (2). Available at: <https://doi.org/10.31945/iprij.210204> [Accessed 15 May 2026].

Reinecke, L. and Rieger, D. (2021) 'Media Entertainment as a Self-Regulatory Resource: The Recovery and Resilience in Entertaining Media Use (R2em) Model' [Online]. in: Vorderer, P. and Klimmt, Christoph ed., *The Oxford Handbook of Entertainment Theory*. Oxford UK: Oxford University Press, pp. 755–779. Available from: <https://doi.org/10.31234/osf.io/fyg49> [Accessed 15 May 2026].

Ribeiro, M. H., Calais, P. H., Almeida, V. A. F. and Wagner, M. J. (2017) 'Everything I Disagree with Is #fakenews': Correlating Political Polarization and Spread of Misinformation.' in: *CEUR Workshop Proceedings*, 2017. vol. 1828. pp. 94–98.

Riedl, M. J., Whipple, K. N. and Wallace, R. (2022) 'Antecedents of Support for Social Media Content Moderation and Platform Regulation: The Role of Presumed Effects on Self and Others.' *Information, Communication & Society*, 25 (11), pp. 1632–1649. Available at: <https://doi.org/10.1080/1369118X.2021.1874040> [Accessed 15 May 2026].

Roberts, M. E., Stewart, B. M. and Tingley, D. (2019) 'Stm: R Package for Structural Topic Models.' *Journal of Statistical Software*, 91 (2), p. 41.

Rodrigo, P., Arakpogun, E. O., Vu, M. C., Olan, F. and Djafarova, E. (2024) 'Can You Be Mindful? The Effectiveness of Mindfulness-Driven Interventions in Enhancing the Digital Resilience to Fake News on COVID-19.' *Information Systems Frontiers*, 26 (2), pp. 501–521. Available at: <https://doi.org/10.1007/s10796-022-10258-5> [Accessed 15 May 2026].

Ross, R. M., Rand, D. G. and Pennycook, G. (2021) 'Beyond "Fake News": Analytic Thinking and the Detection of False and Hyperpartisan News Headlines.' *Judgment and Decision Making*, 16 (2), pp. 484–504.

Russell, N. J. (2006) 'An Introduction to the Overton Window of Political Possibilities.' Mackinac Center [Online blog]. Available at: <https://www.mackinac.org/7504> [Accessed 9 January 2026].

Salomaa, S. and Palsa, L. (2019) *Media Literacy in Finland*. 2019:39. Ministry of Education and Culture.

Sanchez-Acedo, A., Carbonell-Alcocer, A., Gertrudix, M. and Rubio-Tamayo, J. (2021) 'The Challenges of Media and Information Literacy in the Artificial Intelligence Ecology: Deepfakes and Misinformation.' *Communication & Society-Spain*, 37 (4), pp. 223–239. Available at: <https://doi.org/10.15581/003.37.4.223-239> [Accessed 15 May 2026].

Schatto-Eckrodt, T., Clever, L. and Frischlich, L. (2024) 'The Seed of Doubt: Examining the Role of Alternative Social and News Media for the Birth of a Conspiracy Theory.' *Social Science Computer Review*, April, p. 08944393241246281. Available from: <https://doi.org/10.1177/08944393241246281> [Accessed 15 May 2026].

Schüz, B. and Jones, C. (2024) 'Mis- and Disinformation in Social Media: Mitigating Risks in Digital Health Communication.' *Bundesgesundheitsblatt – Gesundheitsforschung - Gesundheitsschutz*, 67 (3), pp. 300–307. Available at: <https://doi.org/10.1007/s00103-02403836-2> [Accessed 15 May 2026].

Sharevski, F., Huff, A., Jachim, P. and Pieroni, E. (2022) '(Mis)Perceptions and Engagement on Twitter: COVID-19 Vaccine Rumors on Efficacy and Mass Immunization Effort.' *International Journal of Information Management Data Insights* [Online], 2 (1). Available at: <https://doi.org/10.1016/j.jjime.2022.100059> [Accessed 15 May 2026].

Shoemaker, P., Tankard, J. and Lasorsa, D. (2004) 'Introduction: The Nature of Science' [Online]. in: *How to Build Social Science Theories*. SAGE Publications, Inc., pp. 1–14. Available at: <https://doi.org/10.4135/9781412990110> [Accessed 15 May 2026].

Sproule, J. M. (2001) *Authorship and Origins of the Seven Propaganda Devices: A Research Note*. 4 (1), pp. 135–143.

Stecula, D. A. and Pickup, M. (2021) 'Social Media, Cognitive Reflection, and Conspiracy Beliefs.' *Frontiers in Political Science* [Online], 3. Available from: <https://doi.org/10.3389/fpos.2021.647957> [Accessed 15 May 2026].

Stock, F., Pelica, S., Ashcroft-Jones, S., Stablum, F., Acosta, C., Aimable, A., Amaro, M., Anderson, P., Brandao, N., Coen, D., Coughlin, C., Crnogorac, M., Dong, A. L., Dönicke, J., Ehlers, R., Elkot, S., Farahat, E., Filippov, N., Gil, S., Jemioło, H., Karczewska, M. W., Kirchhof, N. F., Kudinoor, S., La Rue, O., Meindl, D., Milošič, A., Olayinka, J., Ozdemir, S., Park, Y., Pi, N., Šafranko, N., Saxena, A., Silane, Y., Sloksnath, P. R., Souza-Neto, V., Szaszi, B., Rai, A. A., Vernon, C., Xie, M. W., Xu, R., Yamada, Y., Yetkin, F., Zilahi, E., Ruggeri, K., Sultan, M., Tump, A. N., Ehmann, N., Lorenz-Spreen, P., Hertwig, R., Gollwitzer, A. and Kurvers, R. H. J. M. (2024) 'Susceptibility to Online Misinformation: A Systematic Meta- Analysis of Demographic and Psychological Factors.' *Proceedings of the National Academy of Sciences*, 121 (47) November, p. e2409329121. Available from: <https://doi.org/10.1073/pnas.2409329121> [Accessed 15 May 2026].

Surjatmodjo, D., Unde, A., Cangara, H. and Sonni, A. (2024) 'Information Pandemic: A Critical Review of Disinformation Spread on Social Media and Its Implications for State Resilience.' *Social Sciences-Basel*, 13 (8). Available at: <https://doi.org/10.3390/socsci13080418>.

Sutton, A. J., Duval, S. J., Tweedie, R. L., Abrams, K. R. and Jones, D. R. (2000) 'Empirical Assessment of Effect of Publication Bias on Meta-Analyses.' *BMJ*, 320 (7249) June, pp. 1574–1577. Available at: <https://doi.org/10.1136/bmj.320.7249.1574> [Accessed 15 May 2026].

Tajfel, H. and Turner, J. C. (1986) 'The Social Identity Theory of Intergroup Behaviour.' in: Worchel, S. and Austin, W. G. ed., *Psychology of Intergroup Relations*. Chicago, IL: Nelson-Hall, pp. 7–24.

The Nation (2025) 'AVM Aurangzeb Ahmed Becomes Top Google Search in Pakistan' [Online]. The Nation. Available at: <https://www.nation.com.pk/15-May-2025/avm-aurangzeb-ahmed-becomes-top-google-search-in-pakistan> [Accessed 27 March 2026].

Till, C. (2020) 'Propaganda through 'Reflexive Control' and the Mediated Construction of

Reality'. *New Media & Society* [Online]. Available from: <https://journals.sagepub.com/doi/10.1177/1461444820902446> [Accessed 10 April 2026].

Tokita, C. K. (2021) 'Networks, Collective Behaviour, and Information in Social Systems-from Ant Colonies to Social Media.' *Dissertation Abstracts International: Section B: The Sciences and Engineering*, 82 (12-B), p. No-Specified.

Törnberg, P. and Chueri, J. (2025) 'When Do Parties Lie? Misinformation and Radical-Right Populism across 26 Countries.' *The International Journal of Press/Politics*, January, p. 19401612241311886. Available at: <https://doi.org/10.1177/19401612241311886> [Accessed 15 May 2026].

Tricco, A. C., Lillie, E., Zarin, W., K. O'Brien, K., Colquhoun, H., Levac, D., Moher, D., D.J. Peters, M., Horsley, T., Weeks, L., Hempel, S., A. Akl, E., Chang, C., McGowan, J., Stewart, L., Hartling, L., Aldcroft, A., G. Wilson, M., Garritty, C., Lewin, S., M. Godfrey, C., T. Macdonald, M., V. Langlois, E., Soares-Weiser, K., Moriarty, J., Clifford, T., Tunçalp, Ö. and E. Straus, S. (2018) 'PRISMA Extension for Scoping Reviews (PRISMA-SCR): Checklist and Explanation.' *Annals of Internal Medicine* [Online], September. Available at: <https://doi.org/10.7326/M18-0850> [Accessed 15 May 2026].

Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D. and Wetherell, M. S. (1987) *Rediscovering the Social Group: A Self-Categorisation Theory*. Cambridge, MA, USA: Basil Blackwell.

Valkenburg, P. M. and Peter, J. (2013) 'The Differential Susceptibility to Media Effects Model.'

Journal of Communication, 63 (2), pp. 221–243. Available at: <https://doi.org/10/gdqc28> [Accessed 15 May 2026].

Van Der Linden, S., Albarracín, D., Fazio, L., Freelon, D., Roozenbeek, J., Swire-Thompson, B. and Van Bavel, J. (2025) 'Using Psychological Science to Understand and Fight Health Misinformation: An APA Consensus Statement' *American Psychologist* [Online]. Available at: <https://doi.org/10.1037/e506432023-001> [Accessed 15 May 2026].

Wardle, C. and Derakhshan, H. (2017) 'Information Disorder: Toward an Interdisciplinary Framework for Research and Policy Making Council of Europe Report DGI (2017)09' [Online]. Strasbourg, France. Available at: <https://rm.coe.int/information-disorder-toward-an-interdisciplinary-framework-for-research/168076277c> [Accessed 15 May 2026].

Wasmer, M., Koch, A., Harkness, J. and Gabler, S. (1996), Konzeption Und Durchführung Der "Allgemeinen Bevölkerungsumfrage Der Sozialwissenschaften" (ALLBUS) 1996.' ZUMA-Arbeitsbericht [Online], 96 (08). Available at: http://www.gesis.org/fileadmin/upload/forschung/publikationen/gesis_reihen/zuma_arbeitsberichte/1996/96_08.pdf [Accessed 15 May 2026].

Wiesner, A. (2025) 'Tackling (Misleading) Incivility Online: A User-Centric Evaluation of Different Comment Moderation Strategies.' *Information, Communication & Society*, 0 (0) September, pp. 1–19. Available at: <https://doi.org/10.1080/1369118X.2025.2557548> [Accessed 15 May 2026].

Wirtschafter, V., Pereira, F. B., Bueno, N., Pavão, N., Santos, J. P. O. dos and Nunes, F. (2024) 'Detecting Misinformation: Identifying False News Spread by Political Leaders in the Global South.' *Journal of Quantitative Description: Digital Media* [Online]. Available at: <https://doi.org/10.51685/jqd.2024.007> [Accessed 15 May 2026].

World Economic Forum (2025) 'The Global Risks Report 2025 20th Edition INSIGHT REPORT.' World economic forum.

Zach, E. (2024) 'Learning Fact-Checking amid Georgia's 'foreign Agent' Law' [Online]. Deutsche Welle. Available at: <https://akademie.dw.com/en/fact-checking-amid-georgias-foreign-agent-law/a-70285382> [Accessed 10 April 2026].

Zerback, T., Töpfl, F. and Knöpfle, M. (2021) 'The Disconcerting Potential of Online Disinformation: Persuasive Effects of Astroturfing Comments and Three Strategies for Inoculation against Them.' *New Media and Society*, 23 (5), pp. 1080–1098. Available at: <https://doi.org/10.1177/1461444820908530> [Accessed 15 May 2026].

Zhang, B. and Schmierbach, M. (2024) 'Biased Processing of Political Fact-Checks on Social Media: Testing the Effects of User Comments and Partisan Worldview on Misbeliefs and Political Candidate Evaluation.' *Internet Research* [Online]. Available from: <https://doi.org/10.1108/INTR-06-2024-0890> [Accessed 15 May 2026].

Ziemer, C.-T. and Rothmund, T. (2022) 'Psychological Underpinnings of Disinformation Countermeasures: A Systematic Scoping Review' [Online]. preprint. PsyArXiv. Available at: <https://doi.org/10.31234/osf.io/scq5v> [Accessed 15 May 2026].

Ziemer, C.-T. and Rothmund, T. (2024) 'Psychological Underpinnings of Misinformation Countermeasures: A Systematic Scoping Review.' *Journal of Media Psychology: Theories, Methods, and Applications*, 36 (6), pp. 397–409. Available from: <https://doi.org/10.1027/1864-1105/a000407> [Accessed 15 May 2026].

