Master’s degree in International Relations

Chair of International Organizations and Human Rights

The Fight Against Terrorism and Disinformation: An Analysis on the Contribution of International Organizations

Prof. Donato Greco

Thesis Supervisor

Prof. Alessandro Orsini

Thesis Co-Supervisor

Matilda BARI

Matr. 654372

Candidate

Academic Year 2024/2025

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

AI	Artificial Intelligence
AI/ML	Artificial Intelligence/Machine Learning
APIs	Application Programming Interfaces
ASEAN	Association of Southeast Asian Nations
CFREU	Charter of Fundamental Rights of the European Union
COE-DAT	Center of Excellence Defense Against Terrorism
CSEP	Civil Society Empowerment Programme
CTC	Counter-Terrorism Committee
CTED	Counter-Terrorism Committee Executive Directorate
DSA	Digital Service Act
e.g.	*exempli gratia* (or “example given”)
ECHR	European Convention on Human Rights
ECtHR	European Court of Human Rights
EDAP	European Democracy Action Plan
EEAS	European External Action Service
etc.	*et cetera*
EU	European Union
GIFCT	Global Internet Forum to Counter Terrorism
HTS	Hay’at Tahrir al-Sham
ICC	International Criminal Court
ICCPR	International Covenant on Civil and Political Rights
ICJ	International Court of Justice
ICT	Information and Communication Technology
IOM	International Organization for Migration
IS	Islamic State
ISIL	Islamic State of Iraq and the Levant
KPIs	Key Performance Indicators
MDM	Misinformation, Disinformation, Malinformation

MIL	Media and Information Literacy
ML	Machine Learning
NATO	North Atlantic Treaty Organization
OSCE	Organization for Security and Co-operation in Europe
RAN	Radicalization Awareness Network
RES	Resolution
RFOM	Representative on Freedom of the Media
SG	Secretary General
STL	Special Tribunal for Lebanon
StratCom COE	Strategic Communications Center of Excellence
TCAP	Terrorist Content Analytics Platform
TEU	Treaty on European Union
TFEU	Treaty on the Functioning of the European Union
TFTP	Terrorist Finance Tracking Programme
UN	United Nations
UNCTT	United Nations Counter-Terrorism Center
UNDP	United Nations Development Programme
UNGA	United Nations General Assembly
UNOCT	United Nations Office of Counter-Terrorism
UNODC	United Nations Office on Drugs and Crime
UNSC	United Nations Security Council
VCLT	Vienna Convention on the Law of Treaties

INTRODUCTION

1. Abstract

This thesis analyzes the contribution of international organizations – the United Nations, the European Union, NATO, and the OSCE – in addressing online misinformation, disinformation, and malinformation (MDM) in the context of counterterrorism. By evaluating their legal foundations, strategic frameworks, and institutional practices, the study explores how each organization integrates the fight against MDM into its broader security and human rights agendas. The research finds that while significant normative and political efforts have been made – particularly through the EU's regulatory tools and the UN's global strategies – the overall impact remains limited by fragmented mandates, non-binding instruments, and varying member state commitments. NATO and the OSCE have made strides in conceptual framing and capacity-building, but their actions are constrained by consensus-driven structures and limited jurisdiction. The thesis concludes that international organizations contribute meaningfully to shaping the global discourse on MDM and terrorism, yet their operational capacities need reinforcement through stronger coordination, technological adaptation, and rights-based accountability.

2. Scope and General Objectives

This thesis focuses on analyzing misinformation, disinformation, and malinformation within the broader framework of counterterrorism efforts in the international sphere. More precisely, it investigates how international organizations tackle the increasing threat posed by false, misleading, or harmful information that may threaten security and undermine effective counterterrorism strategies. The research will concentrate on the Euro-Atlantic region only. This is reflected in the choice of the key international organizations that will be analyzed, namely the United Nations (UN), the European Union (EU), the Organization for Security and Co-operation in Europe (OSCE), and the North Atlantic Treaty Organization (NATO).

The objective of this research is to assess the legal, institutional, and policy responses taken by organizations to address and manage misinformation, disinformation, and malinformation in counterterrorism, and understand what limitations and effectiveness each intervention presents. The interest in this topic stems from its present-day relevance given the ways that the information landscape is rapidly changing; however, this thesis also seeks to understand the impacts that these ecosystems can have on security, and whether multilateral responses can sufficiently protect both security and fundamental rights. Therefore, this thesis seeks a comprehensive and critical evaluation of the legal, institutional, and policy frameworks established by the UN, EU, NATO, and OSCE to combat MDM in the counterterrorism sphere. In doing so, it will address the following research questions: To what extent do International Organizations and contribute to addressing the challenge of online disinformation, misinformation, and malinformation in counterterrorism efforts? Are the adopted measures effective and sufficient? What are the primary limitations and critiques of each international organization? And how might international cooperation in this area be enhanced?

There are a total of four chapters. The first one, of a more conceptual nature, is an introduction chapter explaining further the disinformation phenomenon and its link with terrorism. To give concrete examples of that link, two case studies will be analyzed: on one hand, how jihadists use disinformation in their propaganda; on the other, how Russia uses disinformation and uses the word "terrorist" on their target to justify their action and bolster its own image.

The second chapter examines the UN, focusing on its normative foundations and institutional efforts to combat disinformation, misinformation, and malinformation in counterterrorism. It starts by discussing the legal framework, particularly Articles 19 and 20 of the International Covenant on Civil and Political Rights, and assesses relevant UN initiatives such as the Strategy and Plan of Action on Hate Speech, and the developing Global Principles for Information Integrity, along with the expected Code of Conduct. Then the chapter explores the roles of the UN Security Council and other bodies, including the UN Global Counter-Terrorism Strategy and UNESCO.

The third chapter focuses on the EU's contributions, which follows the same logic of the previous chapter by starting with a legal and institutional reflection on whether the EU has the necessary competences by analyzing its foundational treaties and relevant documents. Following this, the chapter examines policy initiatives such as the Commission's Communication on Preventing Radicalization, contributions from the European External Action Service, and the Strategic Compass addressing hybrid threats. Furthermore the EU strategies and agendas will be analyzed, notably the 2005 Counter-Terrorism Strategy and the 2020 Security Union Strategy and Counterterrorism Agenda. Nonetheless, the chapter also engages with key regulatory mechanisms including the EU Internet Forum, the Digital Services Act, the Code of Practice on Disinformation, and the EU Code of Conduct on Disinformation.

The fourth and final chapter addresses Euro-Atlantic security organizations, specifically the OSCE and the NATO. It analyzes the OSCE's role through its Copenhagen Document and the work of the Representative on Freedom of the Media, while assessing NATO's contributions through its Centres of Excellence, its Counter Hybrid Threats Strategy, and its collaborative efforts with the EU, including an initial investigation into NATO's mandate concerning counterterrorism and disinformation. The thesis ends with some final Considerations, providing a synthesis of key and reflects on the effectiveness, synergies, and shortcomings of these organizations' responses, aiming to provide answers to the original research question.

3. Methodology

This thesis employs a methodology that critically analyzes the roles and actions of international organizations in tackling disinformation, misinformation, and malinformation in the context of counterterrorism. The analysis involves reviewing official reports, legal frameworks, and institutional strategies, evaluating both the extent of their involvement in counterterrorism and information management, as well as their adherence to international human rights law.

Special emphasis is placed on the security-freedom of speech dichotomy, examining how these organizations balance the fight against informational threats with the degree of protection of fundamental rights. Additionally, this thesis mostly focuses on official documents, supporting a qualitative approach throughout the paper.

4. Preliminary Concepts and Definitions

Initially, it is essential to clarify key concepts and significant terms that will be used throughout the thesis to prevent any confusion or misunderstandings, as well as to outline the characteristics associated with each definition. There will be two main definitions: disinformation and similar phenomena, and the concept of terrorism. Before delving into the specific meanings assigned to these concepts in this work, it is vital to first examine their origins and evolution. Therefore, each definition will begin with a brief historical overview to facilitate a clearer understanding of how their meanings in this thesis differ – or are similar – from previous interpretations.

4.1 Difference between Disinformation, Misinformation, and Malinformation

Disinformation and misinformation are most commonly discussed in the literature, as can be seen, for example, in the European Commission's Communication on the European Democracy Action Plan (hereafter EDAP), which gave the following definition of the two terms:

> «*Disinformation is false or misleading content that is spread with an intention to deceive or secure economic or political gain, and which may cause public harm. Misinformation is false or misleading content shared without harmful intent though the effects can be still harmful.* »[1]

Meanwhile, the UN defines disinformation as having the following characteristics: «information that is inaccurate, intended to deceive and shared in order to do serious harm ».[2] This stems from the definition given by the Special Rapporteur on the promotion and protection of the right to

[1] European Commission, *Tackling online disinformation*, Digital Strategy. 2024.

[2] UN Secretary General, *Countering disinformation for the promotion and protection of human rights and fundamental freedoms (A/ 77/287)*, United Nations Digital Library, 2022, p. 2.

freedom of opinion and expression (Irene Khan) and the United Nations Educational, Scientific and Cultural Organization. The latter defined it as «false or misleading content that can cause specific harm – irrespective of motivations, awareness or behaviours »,[3] and Khan gives a similar definition stating that it is «false information that is disseminated intentionally to cause serious social harm ».[4] Thus, both the EU and UN have a similar conception of disinformation, albeit the EU also considers the possibility for economic and/or political gain out of spreading disinformation.

As opposed to disinformation, which the first known use of this word was found in the *Medicine Lodge* newspapers[5], and misinformation, which according to Oxford English Dictionary was coined by Abraham Fleming in his work «Holinshed's Chronicles » (1587)[6], the term malinformation is quite recent: it was created in Council of Europe report named « Information Disorder: Toward an interdisciplinary framework for research and policy making » (2017) by Claire Wardle and Hossein Derakhshan which defined malinformation as « when genuine information is shared to cause harm, often by moving information designed to stay private into the public sphere »[7]. In this same report, it is interesting to notice that for each term, the authors decided to illustrate them with the example of the 2017 French Presidential election: misinformation, as seen in the spread of false rumors (e.g., the death of a second policeman) during the breaking news of the Champs-Élysées attack[8]; disinformation, as seen by fabricated stories such as Macron's alleged offshore account and coordinated 'Twitter raids' spreading false claims (e.g., Saudi Arabia funding Macron)[9]; and malinformation, demonstrated by the leak of Macron's private emails just before the run-off vote[10].

[3] BONTCHEVA et al., *Balancing Act: Countering Digital Disinformation While Respecting Freedom of Expression*, United Nations Educational, Scientific and Cultural Organization, Paris, 2020, p. 25.
[4] KHAN, *Disinformation and freedom of opinion and expression (A/HRC/47/25)*, United Nations Digital Library, 2021, p. 4.
[5] Medicine Lodge Cresset. *Issue of 13th February 1887*, Medicine Lodge, 1887, p. 3.
[6] Oxford University Press, *Misinformation*, Oxford English Dictionary, 2024.
[7] WARDLE, DERAKHSHAN, *Information Disorder: Toward an interdisciplinary framework for research and policy making*, Council of Europe, 2017, p. 5.
[8] *Ivi.*, p. 21.
[9] *Ibidem*.
[10] *Ibidem*.

Thus, we can group the different definitions above mentioned into a simple one for each word in the following table:

Table 1. Definitions and Examples

	Definition	**Example**
Misinformation	A **false information** but with **no malintent**.	False Connection; Misleading Content

Disinformation	A **false information** with the **intent** to deceive and potentially cause harm.	False Context;
		Imposter Content;
		Manipulated Content; Fabricated Content

Malinformation	An **information based on reality** but manipulated or removed from its context with the **intent** to deceive and cause harm.	Leaks; Harassment; Hate speech

Source and definitions: Bari M. with examples from Wardle and Derakhshan, Information Disorder: Toward an interdisciplinary framework for research and policy making, Council of Europe report DGI(2017)09, 2017.

The words in bold in Table 1 allow us to notice the two variables being falseness of information (variable A) and malintent (variable B). As such, it could be noted that misinformation and malinformation are at polar opposites, while disinformation could be defined as a hybrid of the two having half of each characteristic. These two variables can be grouped in a matrix, as seen in Table 2, which allows a better comprehension of the relation between variable A and variable B.

Table 2. Matrix between the variables of falseness and malintent

	False information	True information
No malintent	**Misinformation**	**Reality**
Malintent	**Disinformation**	**Malinformation**

Source: Bari M.

With all that being said, it is imperative to create an umbrella term that would allow us to talk about the three concepts without making a synecdoche (simply using one word for the three of them), or, repeating each concept together which would be lengthy. The importance of separating the individual definitions and when talking about them together comes from the fact that without an umbrella term, there would be further confusion and unclarity when the differences between the terms are already subtle and often wrongly used as synonyms. As such, when discussing the phenomenon of misinformation, disinformation, and malinformation together this paper will refer to them as MDM. The words' placement choice is not casual; this order is to purposely keep the relation of the two pole opposites while disinformation acts as their hybrid. Furthermore, the decision to not use the term "fake news" when referring to MDM is because this thesis wants to distance itself from the journalistic and dramatic vocabulary, and instead opt for a more scientific and objective approach of the phenomenon.

4.2 Defining Terrorism

The term "terrorism" has been used extensively in both political and legal debates, yet it remains one of the most disputed concepts in the international sphere. As such, despite so many efforts, a universally accepted definition is still lacking. But why is it so complicated to define a phenomenon so often discussed? This difficulty is due to political factors and the complexities involved in differentiating terrorism from other violent forms – such as rebellion, insurgency, and state actions – as well as the possible connotations behind it.

The term itself originated from the “Reign of Terror” (1793–1794) during the French Revolution, where it referred to state-sponsored oppression by the Jacobins[11]. Initially associated with state control, its meaning however evolved in the 19th century to include non-state actors – e.g. Russian anarchists and revolutionaries were denominated to be “terrorists” during the times of Imperial Russia[12]. Another, most clamorous, example would be the origin of World War I with the assassination of Archduke Franz Ferdinand by the ethnic separatists. Likewise, in the 1940s, a similar application was used when Jewish militants employed violence to expedite Israeli independence, notably through the killing of UN mediator Count Bernadotte[13]. However, during and post-World War II, the definition was changed with a mix of state actions and military actions, such as the Allied “terror bombing” campaigns[14], blurring the lines between state and non-state violence. In the 20th century, terrorism became linked with anti-colonial efforts and the ideological struggles of the Cold War. Then, the late 20th century marked the emergence of religious terrorism, represented by groups like Al-Qaeda[15]. And now, unlike earlier movements focused on territorial or political aims, contemporary terrorism seems to be marked by indiscriminate tactics and transnational goals.

[11] SAUL, *Defining Terrorism in International Law*, Oxford University Press, Oxford, 2006, p. 1.

[12] KOUFA, *Terrorism and Human Rights: Progress Report by Special Rapporteur Kalliopi K. Koufa*, United Nations Digital Library, 2001, p. 11.

[13] ICJ, *Reparation for Injuries case*, ICJ Reports 174, 1949.

[14] WALZER, *Just and Unjust Wars: A Moral Argument with Historical Illustrations*, New York, 2006, p. 323 – 325.

[15] SAUL, *Defining Terrorism in International Law (2006)*, cit., p. 2.

As intuited from the history of the term, defining terrorism always seemed to be quite a Sisyphean task: always so close to agreeing on a final definition but never achieving it as it is filled with semantic instability and ideological contention. Most remarkable is the term's fluidity which enables its political weaponization: states label adversaries as "terrorists" to delegitimize them, while avoiding the term to refer to allies. For instance, Israel described Palestinian refugees in Sabra and Shatilla as terrorists[16], while the UK hesitated to label France's bombing of the Rainbow Warrior as terrorism to avoid diplomatic friction[17]. Thus, the term's overbroad application further complicates legal clarity.

Jurisprudential disputes center on state vs. non-state violence. Ben Saul explains well this dichotomy: developing states emphasize "state terrorism" (e.g., colonial repression), while developed states focus on non-state actors[18]. So, anti-colonial movements challenge traditional definitions by framing violence as legitimate self-determination. Moral judgment is inherent in the term and thus, labeling an act as "terrorism" implies some sort of condemnation, complicating further the neutral legal categorization. Besides that, criminal law's demand for precision clashes with terrorism's subjectivity, risking violations of the *nullum crimen sine lege principle*. And so, without a clear definition, states exploit ambiguity to justify arbitrary prosecutions or human rights abuses.

Early legal efforts to define terrorism were hindered by ideological divides. The 1937 League of Nations Convention for the Prevention and Punishment of Terrorism marked the very first attempt at a definition, criminalizing acts intended to provoke terror in specific populations or the general public[19]. However, the Convention failed due to limited ratification and because of the arrival of the Second World War[20]. Post-1970, sectoral treaties addressed specific acts without defining terrorism broadly: The 1970 Convention for the Suppression of Unlawful Seizure of Aircraft targeted aircraft hijacking; the 1979 Hostages Convention defined hostage-taking; the 1997 Terrorist Bombings Convention criminalized explosive attacks in public spaces. All these treaties,

[16] SAUL, *Defining Terrorism in International Law (2006)*, cit., p. 3.
[17] *Ibidem*.
[18] *Ivi.*, p.2.
[19] League of Nations, *Convention for the Prevention and Punishment of Terrorism*, 1937, Article 1, p. 6.
[20] SAUL, *Defining Terrorism in International Law*, New York School of Law, 2021, pp. 3–4.

and more, avoided defining the term "terrorism" – or even mentioning it at all for some conventions – reflecting states' reluctance to engage in broader conceptual debates.

After 9/11, UN Security Council Resolution 1373 (2001) required nations to combat terrorism[21] but lacked a clear definition, leading to varying domestic interpretations. Also the International Criminal Court (ICC) Statute from 1998 did not include terrorism, this time due to a lack of consensus[22]. In the regional context, the 1998 Arab Convention allowed armed resistance against foreign occupation[23], while the 2002 EU Framework Decision defined terrorism seemingly excluding state actions as it's states that terrorism «seriously jeopardises human rights, threatens democracy, and aims notably to destabilise legitimately constituted governments and to undermine pluralistic civil society »[24].

Similarly, international and judicial bodies have approached the issue of terrorism with caution. The International Court of Justice (ICJ)'s *Reparation for Injuries* Case (1949) implicitly acknowledged terrorism by affirming that states must protect UN agents, referencing to the failed prevention of the assassination of Count Bernadotte[25]. Similarly, in the *Lockerbie* Case (1992), the ICJ avoided defining terrorism, concentrating instead on jurisdictional issues, nevertheless, the United Nations Security Council implicitly recognized as terrorism the act of violence on civil aviation[26]. The Kadi v. EU (2008) case contested the UN sanctioning process, arguing that terrorist associate listing lacked clear definition and due process in terrorist designations thus exposing the dangers of politicized listings[27].

It seems that quite a variety of conduct has been considered "terrorist acts" by the literature and judicial cases: while the *Lockerbie* case represents the conduct of the sabotage of a civilian aircraft,

[21] UNSC, *Resolution 1373*, 2001, p. 2.

[22] SAUL, *Defining Terrorism in International Law (2021)*., cit.

[23] League of Arab States, *The Arab Convention For The Suppression of Terrorism*, 1998, Preamble, p. 1.

[24] Council of Europe, *Guidelines on Human Rights and the Fight Against Terrorism*, 2002, Preamble(a). p. 3.

[25] ICJ, *Reparation for Injuries Suffered in the Service of the United* Nations, Advisory Opinion of 11 April 1949, ICJ Reports 174, 1949, p. 177 – 184.

[26] SAUL, *Defining Terrorism in International Law (2006)*, cit., p. 226.

[27] *Ivi.*, p. 232.

in the *Reparation for Injuries* Case, it's the targeted political assassination. However, there are other examples of different conduct.

In the *Tehran Hostages* and *Nicaragua* cases, the ICJ refrained from employing the term "terrorism," opting instead to frame the incidents in terms of violations of diplomatic immunity and non-intervention, respectively.[28] In the first case, it's the taking of hostages; in the latter, it's the indirect use of force by supporting the Contras, who committed acts of violence (including attacks on civilian populations and destruction of infrastructure). In contrast, the 2004 *Wall Advisory Opinion* marked a notable shift: although the ICJ majority remained legally silent on defining terrorism,[29] separate and dissenting opinions – particularly that of Judge Kooijmans – offered a nascent legal concept.[30] Kooijmans identified deliberate and indiscriminate attacks on civilians as the core element of terrorism and acknowledged such acts as international crimes condemned irrespective of motive.[31] As such, in the *Wall Advisory Opinion,* it's the indiscriminate attacks on civilians.

However, it seems that an important turning point in judicial and international was marked by the UN Special Tribunal for Lebanon and the Security Council Resolution 1566. In 2011, the Appeals Chamber of the UN Special Tribunal for Lebanon (STL) sought to address the longstanding ambiguity surrounding the definition of terrorism under international law. The Tribunal outlined a three-part framework highlighting (1) the commission (or intent) of violent acts[32], (2) the aim to instill fear or pressure authorities[33], and (3) the need for a transnational aspect[34]. This effort tried to unify a fragmented legal framework regarding terrorism. Although its ruling was not universally applicable, it still marked a crucial move towards harmonizing different national approaches and providing an example for future rulings related to terrorism.

[28] SAUL, *Terrorism in Customary International Law*, Oxford University Press, Oxford, 2008, pp. 251 – 252.
[29] *Ivi.*, pp. 252 – 253.
[30] *Ibidem*.
[31] *Ibidem*.
[32] UN Special Tribunal for Lebanon (Appeals Chamber), *Interlocutory Decision on the Applicable Law: Terrorism, Conspiracy, Homicide, Perpetration, Cumulative Charging*, 2011, p. 73.
[33] *Ibidem*.
[34] *Ibidem*.

The STL's ruling dealt with the issue of defining terrorism in a way that balances state sovereignty with international legal consistency. In developing its definition, the Tribunal referenced existing international agreements on terrorism, such as the 1999 Terrorist Financing Convention and the UN Draft Comprehensive Convention, while diverging from the narrower definition in UN Security Council Resolution 1566 (2004). But most importantly, the Tribunal chose not to include political motivation as a key component, as it found no solidified state practice or *opinio juris* advocating for this requirement.[35] This choice not only clarified the definition by steering clear of subjective interpretations of political motives but also expanded the definition of terrorism to include actions that have typically been viewed as outside its scope.

A more critical examination of the STL's reasoning reveals both its strengths but also its limitations: the Tribunal relied heavily on a comparative survey of national terrorism statutes, asserting that these domestic laws demonstrated a "widespread consensus" on the constitutive elements of terrorism[36]. However, a closer analysis reveals significant differences among these legal frameworks. National legislation defines terrorism variably in relation to civil war, public disorder, constitutional subversion, or threats to territorial integrity, often expanding the definition beyond international concerns. Ben Saul points out that the STL's blending of domestic and transnational terrorism undermines its argument[37], stating that it neglects the jurisdictional distinctions that shape state responses to counterterrorism. Furthermore, he criticizes that by selectively citing a narrow range of national laws (37), the Tribunal's conclusion overlooks the wider and more varied legal landscape that continues to resist a unified definition[38]. Ben Saul also comments that the STL's dependence on Security Council Resolution 1566 to substantiate its definition is equally problematic, as the resolution does not endorse a general customary norm of terrorism but rather links the crime to existing sectoral treaty offenses, thus reinforcing a more limited legal framework[39]. Besides, he notes that other references the Tribunal uses, such as General Assembly resolutions and regional treaties, similarly fail to provide a universally accepted

[35] *Ivi.*, par. 106, p. 68.
[36] *Ivi.*, par. 92, p. 57.
[37] SAUL, *Defining Terrorism in International Law (2021).*, cit.
[38] *Ibidem*.
[39] *Ibidem*.

definition[40]. In fact, he points out that recent decisions from national courts, especially in the UK with the case *Al-Sirri v Secretary of State for the Home Department*, reaffirm the absence of a universally recognized legal standard[41], further highlighting the lack of state consensus needed to classify terrorism as a distinct crime under customary international law. The Tribunal's attempt to establish a definition of terrorism demonstrates a practical approach to fill existing legal voids but raises issues concerning the principle of legality (*nullum crimen sine lege*). By formulating a crime based on inconsistent state practices, the Tribunal risks undermining legal certainty and broadening the definition to encompass acts lacking ideological intent. Furthermore, its omission of purely domestic terrorism creates an arbitrary division that fails to accurately represent the nature or seriousness of these acts. Ultimately, this jurisprudential approach highlights the ongoing challenge of balancing practical legal needs with clear conceptual definitions.

Besides, there is an immense debate in the literature on the configurability of terrorism as an autonomous international crime. Antonio Cassese, another relevant scholar for international law, has contributed significantly to this debate. While acknowledging the definitional challenges in international law, he contends that it is possible that serious acts of terrorism (such as the 09/11 events and similar cases) could fall under the existing category of crimes against humanity, provided that they «meet the requirements of that category of crimes »[42] and that «no special account should be taken of one of the specific features of terrorism, namely the intent to spread terror among civilians ».[43] On a similar path, James D. Fry argues that terrorism shouldn't be considered an autonomous crime but that it should be prosecuted under already existing categories of international crimes (e.g., crimes against humanity or even genocide)[44] for a more pragmatic and policy-oriented approach.

[40] *Ibidem*.
[41] *Ibidem*.
[42] CASSESE, *Terrorism is also disrupting some crucial legal categories of international law*, European journal of international law, 2001, p. 995.
[43] *Ibidem*.
[44] FRY, *Terrorism as a Crime Against Humanity and Genocide: The Backdoor to Universal Jurisdiction*, UCLA Journal of International Law and Foreign Affairs, 2002, pp. 198 – 198.

However, in later research, Cassese opts for a more nuanced view: he delineates terrorism as a distinct international crime in peacetime,[45] characterized by its transnational nature, objective crime even under national legal system, political or ideological motives, and intent to spread terror or coerce states.[46] In times of conflict, terrorism is characterized by actions against civilians and individuals not taking part in the "armed hostilities", and the aim to spread terror.[47] While terrorism may, under certain circumstances, qualify as a crime against humanity (especially when it forms part of «widespread or systematic attacks against civilians »)[48] its legal status in armed conflict remains contested. Moreover, Cassese concludes with the possible evolution of terrorism in armed conflict from a subcategory of war crimes into an autonomous international crime, similarly to how genocide started as a subcategory of crimes against humanity into its own category of crime.[49] So, although terrorism *per se* is not codified as an autonomous crime in the Rome Statute of the International Criminal Court or in any other international treaty, there is a growing scholarly and jurisprudential basis, as seen with Cassese, for its recognition as an international crime under customary international law.[50] Besides, while tribunals like the International Criminal Tribunal for the former Yugoslavia in the *Galic case* did not explicitly declare terrorism *per se* a customary crime, the elements used in their legal reasoning align closely with Cassese's proposed definition.[51]

With all that being said, in this thesis, terrorism will be defined in having the following characteristics:

(i.) is a threat to international and/or national peace and security;

(ii.) is a criminal act with the intent to commit serious harm (objective element);

[45] CASSESE, *The Multifaceted Criminal Notion of Terrorism in International Law*, Journal of International Criminal Justice, 2006, p. 957.
[46] *Ibidem*.
[47] *Ibidem*.
[48] *Ibidem*.
[49] *Ivi*., p. 958.
[50] DE LONDRAS, *Terrorism as an international crime*, edited by SCHABAS et al., *Routledge Handbook of International Criminal Law*, Oxfordshire, 2010, pp. 175 – 176.
[51] Ibidem.

(iii.) has the intent to spread terror among a population, specific people or group of persons (*dolus specialis*), and/or intimidate to compel a subject of international law to act or to abstain from acting (subjective element).

However, since this research is subscribed within the world of digital MDM as well, we must also consider also the possible and important criteria of the transnational scale since the Internet is something that can be accessed globally, hence it is hard to restrict any content only nationally. Besides, while terrorism is typically associated with non-state actors, the question of state terrorism remains contentious, as international law primarily governs state conduct under different legal frameworks. Nonetheless, in this thesis, state terrorism will be considered since we are focusing on disinformation as a phenomenon and, based on recent world events, states can take part in aiding or inciting terrorism through disinformation.

In conclusion, this will be the definition of terrorism used in this paper, following the characteristics mentioned above, which come from different sources: the decision not to strictly follow a source is because digital disinformation doesn't fit perfectly into any of the definitions found in the given sources. Besides, it underscores the struggle to define such a concept since it can be so differently intended on different occasions. Regardless, the definition proposed here mostly aligns with the Security Council Resolution 1566 (2004). However, this thesis does not believe that this is the ultimate and precise definition of terrorism since, for example, it excludes harm against everything other than humans, such as, for example, infrastructure, property, and environment, since this paper only discusses the realm of information. It is simply the definition that will be intended in this paper, and clarifying the definition is necessary for a good comprehension of the thesis overall.

CHAPTER I

The Phenomenon of Terrorism and the Role of Disinformation

1. Introduction

The digital era has not merely multiplied the conducts through which MDM circulates; as we will see in this chapter, the rise of social media has fundamentally restructured the *modus operandi* of MDM, rendering deception more agile, pervasive, and resistant to traditional countermeasures. Classical propaganda relied on a massive logistical framework to build falsehoods through a number of centralized media. In contrast, modern MDM is exploiting the networked design of the Internet. The connective architecture of social media, algorithmic curation, and consumer surveillance can be blended, and each provides ways to target specific audiences with manipulative content, which can be done at scale and silently. Disinformation and malinformation are characterized by a clear intent to mislead through deception. They are developed and manipulated to shape the audience's perception. As analyzed later on in this chapter, one of the basic strategies behind is misdirection;[52] while clear misdirection can be achieved through fake identities by means

[52] KIVIMÄKI, *Open-Source Information for Intelligence Purposes: the Challenge of Disinformation*, edited by ARCOS, CHIRU, IVAN, *Routledge Handbook of Disinformation and National Security*, Abingdon, 2023, p. 285.

of fake accounts,[53] automated bots,[54] and cover organizations,[55] the practice also engages in clouding the idea of fact by overloading the information space through competing and conflicting ideas.

Additionally, this chapter will delve on how digital MDM is a fundamental aspect of hybrid threats,[56] enhancing cyber intrusions and espionage. The strategic pattern of hacking, selective leaking, and narrative amplification, as illustrated by the 2017 Macron email leaks, shows how stolen information can be weaponized to alter public discourse. Even more importantly, these strategies are not limited to state actors; non-state groups, including terrorist organizations like Islamic State of Iraq and the Levant (ISIL),[57] have adopted similar approaches, integrating false narratives and artificial amplification to exert influence.[58] Amplification is key. While pre-digital MDM depended on controlled distribution and elite support, virality now replaces credibility. Social media algorithms favor emotionally charged content, especially fear and outrage, thereby speeding up the dissemination of manipulative materials. By taking advantage of the principle of "social proof",[59] which in the context of social media is demarked by the number of likes and shares,[60] MDM propagators artificially boost engagement through bots and coordinated networks to create a misleading sense of consensus.[61] Besides, even individuals who always fact-check information they read are vulnerable to being influenced in emotionally charged moments,[62] thus MDM have a pervasive power on everyone.

[53] NIMMO et al., *Secondary Infektion*, Graphika Report, 2020, p. 4.

[54] KIVIMÄKI*, Open-Source Information for Intelligence Purposes: the Challenge of Disinformation*, cit., p. 287.

[55] U.S. Department of Treasury, *Treasury Escalates Sanctions Against the Russian Government's Attempts to Influence U.S. Elections*, U.S. Department of Treasury, 2021.

[56] KIVIMÄKI*, Open-Source Information for Intelligence Purposes: the Challenge of Disinformation*, cit., p. 285.

[57] *Ivi.*, p. 287.

[58] MILTON, *Truth and lies in the Caliphate: The use of deception in Islamic State propaganda*, edited by MALTBY et al, *Media, War & Conflict*, Sage Journal, London, 2020, p. 231.

[59] CIALDINI, *Influence: The Psychology of Persuasion (Revised Edition)*, Collins business, New York, 2006, pp. 115 – 116.

[60] KIVIMÄKI*, Open-Source Information for Intelligence Purposes: the Challenge of Disinformation*, cit., p. 287.

[61] *Ibidem*.

[62] SALING et al., *No One Is Immune to Misinformation: An Investigation of Misinformation Sharing by Subscribers to a Fact-Checking Newsletter*, PLOS ONE, 2021, pp. 1–12.

This chapter serves as an introductory framework designed to provide the conceptual foundations necessary for understanding the nexus between MDM and its link with terrorism. This chapter builds upon a comprehensive synthesis of key theoretical insights. The handbook proved particularly apt as a vital foundational text, providing a comprehensive yet focused examination of disinformation, which shapes the analytical perspective used throughout this thesis.

This chapter consists of three main sections. The first section, "Strategic Threat of Disinformation", outlines the progression of disinformation strategies, moving from the so-called "fourth wave" to new trends that indicate the emergence of a potential fifth wave, marked by the use of deepfakes and artificial intelligence/machine learning manipulation. Additionally, it explores the triadic theory of offense, defense, and deterrence in the realm of informational warfare. The second section, "Terrorism and Disinformation", shifts to the security implications of MDM, evaluating its effects on both national and international stability. There will be a specific focus placed on the role of propaganda in terrorist communications and how MDM can be utilized to promote terrorist objectives. This will be supported by two case studies – jihadist organizations and the Russian Federation – illustrating how different actors exploit MDM within the context of terrorism. The chapter concludes with a summary of the key findings, setting the stage for the following analytical sections of the thesis.

Introducing a preliminary consideration of the phenomenon of disinformation in order to provide a conceptual and operational framework is essential for understanding the dynamics through which disinformation interacts with terrorism; thus, the first section aims to outline the evolving characteristics of disinformation as a strategic threat, analyzing its new forms, the technological tools employed, and its impact in information warfare contexts. It is only from this framework that it is possible to address, in the following sections, the core of the chapter: the intersection of terrorism and disinformation, with a focus on security effects, propaganda, and the instrumental use of MDM by state and non-state actors.

2. Strategic Threat of Disinformation

2.1 The fourth wave of disinformation

The proliferation of open-source information has long been central to intelligence analysis, yet digitalization has radically transformed its nature. The explosion of user-generated content and commercially available data has reshaped the information landscape–rendering it both more accessible and more vulnerable to manipulation. Open-source intelligence (OSINT) has been going through an evolution, with different generations.[63] Open-source information is defined as «publicly available information which can lawfully be acquired either through a collection activity or purchase »,[64] as such it could be defined as "second-hand information".[65] It can be found in literature or broadcast media, and now even on the digital sphere.[66] The issue is no longer the mere existence of open source information, which has historically been used in the intelligence domain,[67] but how it could be structured, propagated, and instrumentalized to shape perceptions and policy outcomes. Hence why it needs to be verified or neutralized.

Thomas Rid, a political scientist and scholar known to be one of the experts on cybersecurity, intelligence, and hacking, delineates four waves of disinformation: each reflecting shifts in geopolitical strategy and technological capacity. The first wave started in the 1920s during the Great Depression, with the rise of the radio.[68] The second wave, which happened after World War II is characterized by disinformation becoming "professionalized" by national intelligences.[69] The

[63] WILLIAMS, BLUM, *Defining Second Generation Open Source Intelligence for the Defense Enterprise*, RAND Corporation, California, 2018, p. 40.
[64] KIVIMÄKI*, Open-Source Information for Intelligence Purposes: the Challenge of Disinformation*, cit., p. 283.
[65] *Ibidem*.
[66] *Ivi.*, p. 284.
[67] KIVIMÄKI, *Open-Source Information for Intelligence Purposes: the Challenge of Disinformation*, cit., p. 283.
[68] RID, *Active Measures: The Secret History of Disinformation and Political Warfare*, Profile Books, New York, 2020, p. 7.
[69] *Ibidem*.

third, in the 1970s, institutionalized further these operations, with the Soviets vastly using it.[70] The fourth wave, from the 2010s, is defined by digital platforms and rapid information diffusion.[71]

Thus, with the rise of social media and its user-generated content, open-source information and disinformation have become more complex phenomena than ever before. And with this new era, participatory propaganda enables individuals to become both consumers and inadvertent amplifiers of it, making it even more pervasive in our everyday lives.[72] Understanding these historical and theoretical trajectories is critical in counterterrorism. Disinformation, whether aimed at electoral manipulation,[73] social destabilization,[74] or for any terrorist purposes, remains a potent strategic tool. As MDM evolves with technological innovation, countermeasures must equally advance.

2.2 A new fifth wave of disinformation with Deepfake Disinformation?

The advent of deepfake, also known officially as synthetic media,[75] disinformation marks a significant evolution in information warfare, fusing artificial intelligence (AI) and machine learning (ML) to produce deceptive content of unparalleled realism.[76] When talking about this phenomenon, the terms are often fused into Artificial Intelligence/Machine Learning (AI/ML). To put it simply, the AI/ML is the computer process, while deepfakes are the output of that process. While the mechanism and science behind the logic of AI/ML is fascinating, this thesis will not

[70] *Ibidem*.
[71] *Ibidem*.
[72] ASMOLOV, *The Effects of Participatory Propaganda: From Socialization to Internalization of Conflicts*, Journal of Design and Science, 2019, pp. 6 – 15.
[73] See the theory of five stages of election meddling. AALTOLA, *Democracy's Eleventh Hour: Safeguarding Democratic Elections Against Cyber-Enabled Autocratic Meddling*, Fiia Briefing Paper(226), 2017, pp. 3 – 4.
[74] GIANNOPOULOS, SMITH, THEOCHARIDOU, *The Landscape of Hybrid Threats: A conceptual model,* Publications Office of the European Union, Luxembourg, 2021, p. 32.
[75] VENEMA, *Deepfake Disinformation: How Digital Deception and Synthetic Media Threaten National Security*, edited by ARCOS, CHIRU, IVAN, *Routledge Handbook of Disinformation and National Security*, Abingdon, 2023, p. 175.
[76] *Ibidem*.

explain the technicalities of it, instead remaining at a conceptual level and focusing on how it represents a threat to security. Unlike conventional MDM image modification, which manipulates text or static images with softwares like Photoshop, deepfakes generate hyper-realistic audio-visual imitations that can easily fog the boundaries between fact and fabrication, thereby challenging the capacity to uphold informational integrity as a whole. As scholar Agnes E. Venema – expert in deepfakes, disinformation, and security – explains it simply, the threatening factor about this type of visual MDM is:

> «*While disinformation spread through text is dangerous, people are inclined to believe what they see even more, as will be discussed later in this chapter. The saying 'a picture paints a thousand words' was already being questioned in light of Photoshop, but it takes a new turn now that video can be artificially altered as well and lifelike synthetic faces can be computer generated.* »[77]

Deepfakes harness Generative Adversarial Networks (GANs), wherein two AI models iteratively refine their synthetic media to learn data better and faster[78] with unsupervised learning (without any human input).[79] This escalating sophistication enables the fabrication of credible speeches, falsified confessions, or manipulated war footage—tools capable of inciting geopolitical tension or undermining judicial proceedings. As AI/ML tools become increasingly accessible, deepfake production is no longer confined to state actors; non-state groups can now also exploit the technology for any reason, ranging from political subversion to financial fraud.[80]

Beyond immediate tactical applications, deepfakes pose a fundamental security risk. The proliferation of synthetic media facilitates the so-called "Liar's Dividend", the ability to dismiss

[77] *Ivi.*, p. 178.

[78] *Ivi.*, p. 176.

[79] RUSSELL, NORVIG, *Artificial Intelligence, A Modern Approach (Third Edition*), Pearson, London, 2014, p.694.

[80] See auditory deepfake.

authentic evidence as fake, manipulating reality and thus undermining trust in audio-visual material.[81] Agnes E. Venema emphasizes how deepfake disinformation not only distorts digital discourse but exacerbates real-world instability, as witnessed in events like the 2019 attempted coup in Gabon.[82] She also explains the difficulty in responding to these threats: traditional fact-checking mechanisms are ill-suited to detect synthetic media of such realism.[83] Although efforts are underway to develop AI-driven detection tools, watermarking systems, and authenticity protocols, these measures struggle to match the accelerating evolution of deepfake generation.[84] Besides, as she mentions, dismantling MDM might mean that it's already too late.[85] As Venema argues, countering this phenomenon demands an integrated approach with a mix of technological solutions,[86] the need for updated legal frameworks, and public awareness strategies.[87]

Ultimately, the challenge of deepfake disinformation transcends mere technology; it is a battle over the very concept of truth, with far-reaching implications for democratic resilience and international security. Hence why the title of this section, as the rise of this new technology could very well represent the fifth wave of disinformation.

2.3 Offense, defense, and deterrence in informational warfare

In the evolving landscape of digital conflict, information warfare is no longer a peripheral concern but a central component of national security strategy. Governments, intelligence agencies, and international organizations have come to recognize that the ability to manipulate, defend, and deter within the informational space is as critical as traditional military capabilities. This recognition has permitted H. Akin Ünver and Arhan S. Ertan to theorize a strategic framework encompassing three

[81] CHESNEY, CITRON, *Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security*, University of California Berkeley School of Law, 2019, p. 1785.

[82] VENEMA, *Deepfake Disinformation: How Digital Deception and Synthetic Media Threaten National Security*, cit., p. 184.

[83] *Ivi.*, p. 185.

[84] *Ibidem*.

[85] *Ibidem*.

[86] *Ibidem*.

[87] *Ivi.*, p. 186.

key dimensions: offense, defense, and deterrence.[88] These elements define the way states engage in, counter, and seek to prevent disinformation campaigns in an increasingly networked global order.

Ünver and Ertan state that disinformation has often been framed as a tool of authoritarian regimes, yet they remind that democracies have also engaged in organized information manipulation,[89] using digital disinformation to shape foreign perceptions and achieve strategic objectives.[90] The rationale behind offensive information warfare is clear: controlling narratives means controlling political and security outcomes. Thus, there are two strategies: the Attacker (being disinformation) and the Defender (the target of disinformation).[91]

Nevertheless, they seem to agree that disinformation (as the Attacker), while it is pervasive in critical situations like elections or emotionally charged moments, has a good pay-off only for short-term victory.[92] Thus, its advantage is quite limited in time, especially if the disinformation of the Attacker gets debunked, causing the perpetrator to be affected by a "reputational penalty".[93] Their game theory model, which was conceptualized to illustrate the balance between offence and defense, shows that there are two possible dynamics: disinformation attacks must have higher short-term benefit than reputational costs; and since the Attacker's benefit is only short-term, then its advantage is not absolute.[94]

However, the Defender can also use a strategy against the attack, being Deterrence,[95] thus preventing disinformation campaigns from being launched in the first place. However, for this

[88] ÜNVER, ERTAN, *The Strategic Logic of Digital Disinformation: Offence, Defence And Deterrence in Information Warfare*, edited by ARCOS, CHIRU, IVAN, *Routledge Handbook of Disinformation and National Security*, Abingdon, 2023, p. 192.

[89] *Ibidem*.

[90] See example France and Russia with the disinformation campaigns in Mali. The Guardian, *Rival Disinformation Campaigns Targeted African Users, Facebook Says*, The Guardian, 2020.

[91] ÜNVER, ERTAN, *The Strategic Logic of Digital Disinformation: Offence, Defence And Deterrence in Information Warfare*, cit., p. 194.

[92] *Ivi.*, p. 195.

[93] *Ivi.*, pp. 194 – 195.

[94] *Ivi.*, pp. 197 – 202.

[95] *Ivi.*, p. 202.

strategy to work, two conditions must be met. The first one is that the Defender lets its "debunking capacity" be known to the Attacker successfully;[96] the second one is that the Defender can demonstrate that it can quickly gain support from the international audience.[97] While the model is recognized by its creators as not perfect,[98] it is a solid conceptual basis for viewing the strategy of disinformation attacks.

Besides, one could think that legal and regulatory measures have also emerged as potential deterrence tools. Sanctions targeting individuals and entities involved in disinformation campaigns, as well as platform-based restrictions on malicious actors, would actually impose more tangible costs on the Attacker. However, it would still be a complex scenario as digital information warfare can be of a global and decentralized nature, allowing the Attacker to evade such countermeasures.

Finally, this offense-defense-deterrence framework by Ünver and Ertan perfectly demonstrates the challenges international actors face in navigating the complexities of digital MDM. To effectively counter these threats, international organizations must refine their approaches to digital disinformation, which will be seen in the later chapters.

3. Terrorism and Disinformation

3.1 Impact on national and international security

The weaponization of MDM has long served as a tool of geopolitical competition,[99] yet in the digital age, its impact on national and international security has reached unprecedented heights.[100]

[96] *Ibidem*.
[97] *Ibidem*.
[98] *Ibidem*.
[99] IVAN, *Protective Factors Against Disinformation*, edited by ARCOS, CHIRU, IVAN, *Routledge Handbook of Disinformation and National Security*, Abingdon, 2023, p. 295.
[100] CHIRU, BULUC, *An Ethical Understanding of Military Strategic Communication, Public Relations, And Persuasion*, edited by ARCOS, CHIRU, IVAN, *Routledge Handbook of Disinformation and National Security*, Abingdon, 2023, p. 264.

At the national level, disinformation erodes trust in institutions, disrupts electoral processes, and deepens social divisions. From Cold War-era Soviet *dezinformatsiya* (disinformation)[101] to contemporary manipulation campaigns with the mechanisms remain varied as analyzed so far, with the disinformation attacks, the AI/ML deepfakes, and more. Since the digital era, disinformation thrives in anonymity, masking its origins through covert dissemination.[102] MDM has threatened multiple times the national security: the 2016 U.S. presidential election witnessed foreign interference via social media manipulation,[103] and similarly, France's 2017 elections faced cyberattacks and disinformation targeting political candidates.[104] As the American ex-director of national intelligence Dan Coats emphasizes, these operations use information manipulation and propaganda on social media to weaken democratic resilience by worsening division at the sociological but also political level.[105] However, D. Coats is not the only one to explicitly consider MDM as a threat to national security. In fact several states and organs,[106] such as the US State Department,[107] and scholars like Rubén Arcos and Cristina M. Arribas,[108] who are respectively experts on international relations and communication sciences, consider this to be the case.

Beyond borders, disinformation also destabilizes international security. With digital MDM, there is a progressive blur between national and international in the field of security implications.[109]

[101] KRZAK, *Operational Disinformation of Soviet Counterintelligence during the Cold War*, International Journal of Intelligence and CounterIntelligence, 2022, pp. 2 – 3.

[102] LUKITO, *Digital Disinformation, Electoral Interference and Systemic Distrust*, edited by ARCOS, CHIRU, IVAN, *Routledge Handbook of Disinformation and National Security*, Abingdon, 2023, p.123.

[103] TUDORACHE, *A Perception Management Take on Propaganda as Political Warfare*, edited by ARCOS, CHIRU, IVAN, *Routledge Handbook of Disinformation and National Security*, Abingdon, 2023, p.135.

[104] *Ibidem*.

[105] COATS, *Transcript: Dan Coats Warns The Lights Are 'Blinking Red' On Russian Cyberattacks*, NPR, 2018.

[106] ÜNVER, ERTAN, *The Strategic Logic of Digital Disinformation: Offence, Defence And Deterrence in Information Warfare*, cit., p. 194.

[107] ARQUILLA et al, *Russian Strategic Intentions: A Strategic Multilayer Assessment (SMA) White Paper*, SMA White Papers, NSI Inc., 2019, pp. 70 – 71.

[108] ARCOS, ARRIBAS, *Anticipatory Approaches to Disinformation, Warning and Supporting Technologies*, edited by ARCOS, CHIRU, IVAN, *Routledge Handbook of Disinformation and National Security*, Abingdon, 2023, p. 401.

[109] VENEMA, *Deepfake Disinformation: How Digital Deception and Synthetic Media Threaten National Security*, cit., p. 181.

On the specific matter of deepfakes, Venema emphasizes that they represent a significant and emerging threat to international security due to their ability to destabilize inter-state relationships, undermine diplomatic trust, and spark armed conflict through deliberate deception. As she states:

> «*Where relations between states are precarious and honest brokers are attempting to mend the peace, a well-timed and believable deepfake of one of the members of the negotiation team or a deepfake social media post can undermine the trust in the entire process. Particularly in high-stakes negotiations with low levels of trust to begin with, it may be hard if not impossible to convince counterparts in the negotiation that a deepfake is, in fact, fake.* »[110]

In addition to diplomatic disruption, Venema points out that deepfakes can be exploited in false-flag operations,[111] where an entity fabricates evidence, such as a fake video of an attack or a recording planning aggression, to justify military action under the guise of self-defense.[112] These strategies are not just theoretical; the example Venema uses in her research is when Nazi Germany's use of manipulated information, thus justifying their invasion of Poland behind that pretext.[113] Therefore, deepfakes introduce a disruptive element into the framework of international security, heightening the risks of conflict escalation and compromising the integrity of inter-state relations. Hence, scholars like Venema have a negative view of the phenomenon, stating that «the future looks grim »[114] in light of disinformation's alarming capacity to disrupt diplomacy and destabilize peace processes.

[110] *Ivi.*, p. 182.
[111] *Ibidem*.
[112] *Ibidem*.
[113] *Ibidem*.
[114] *Ivi.*, p. 186.

3.2 Propaganda and Terrorism

As analyzed so far, disinformation has long been a strategic instrument wielded by both state and non-state actors; especially in the digital era, its capacity to manipulate narratives, erode institutional credibility, and undermine public trust has intensified, posing significant challenges to national and international security. Terrorist organizations, in particular, have recognized information as a weapon, leveraging propaganda and disinformation to recruit and justify violence,[115] but also to exploit societal vulnerabilities of targeted groups to ensure they would be more compliant to accept their MDM.[116]

Before the digital era, this sort of propaganda was only possible through local and niche outlets.[117] Today, propaganda has evolved into more sophisticated campaigns exploiting social media and online echo chambers.[118] Thus, disinformation is not auxiliary to terrorism; it is central to its operational strategy for propaganda.[119]

To demonstrate that propaganda serves to systematically construct, disseminate, and consolidate extremist narratives, it is important to quickly look at an example of terrorist group analysis. The study of Taliban videos by W. Mehran, senior lecturer expert in extremists' media strategies and terrorism, is quite illustrative as it highlights how the convergence of text, imagery, and audio-visual elements increases persuasive impact, translating such propaganda especially effective in

[115] PALMERTZ, PAMMENT, *Asymmetrical Conflict in the Information Domain—The Case Of Russia*, edited by ARCOS, CHIRU, IVAN, *Routledge Handbook of Disinformation and National Security*, Abingdon, 2023, p. 95.

[116] ARCOS, ARRIBAS, *Anticipatory Approaches to Disinformation, Warning and Supporting Technologies*, cit., p. 401.

[117] KIVIMÄKI, *Open-Source Information for Intelligence Purposes: the Challenge of Disinformation*, cit., p. 287.

[118] RICHARDS, *The Use of Discourse Analysis in Propaganda Detection and Understanding*, edited by ARCOS, CHIRU, IVAN, *Routledge Handbook of Disinformation and National Security*, Abingdon, 2023, p. 391.

[119] *Ivi.*, 387

shaping both individual and collective attitudes.[120] For example, their beheading videos, deliver powerful “visual facts” and thus strongly influencing public opinion.[121] Because of this, it shuts down political debate, thereby facilitating broader acceptance of counter-violence measures.[122] This use of shock-inducing imagery exemplifies how terrorist propaganda is designed to disrupt rational discourse and compel emotional responses.

Finally, the relationship between terrorism and propaganda is both intrinsic and strategic. Propaganda is both a communicative supplement and a crucial operational instrument, allowing terrorist groups to manipulate narratives, justify violence, and recruit followers. It functions as a tool for persuasion and as a means of ideological combat, solidifying extremist ideologies across various contexts and media channels.

3.3 In what ways does MDM facilitate terrorist strategies?

As such, the use of MDM has long been an integral component of modern terrorism, functioning not merely as a tool for deception but as a strategic force multiplier that amplifies extremist narratives. Whether disseminated by state actors seeking to exploit terrorism for geopolitical gains or by non-state groups looking to radicalize and recruit, MDM is a mechanism that extends beyond the battlefield and into the cognitive domain. As will be seen in this section, the symbiotic relationship between terrorism and MDM reveals a fundamental reality: the success of violent extremist movements is not solely measured by attacks carried out but also by the narratives they manage to spread at an international level. Understanding how disinformation supports terrorist agendas is thus not just an analytical exercise; it is a necessary step toward countering it, since understanding the phenomenon is the key to being able to fight against terrorism more efficiently.

[120] MEHRAN et al., *Deep Analysis of Taliban Videos: Differential Use of Multimodal, Visual and Sonic Forms across Strategic Themes*, Studies in Conflict and Terrorism, 2021, pp. 1–21.
[121] *Ibidem*.
[122] *Ibidem*.

The selection of jihadist groups and Russia as case studies reflects two distinct but complementary manifestations of MDM in the context of terrorism. These two cases illustrate both the bottom-up and top-down applications of MDM in terrorism. Together, they offer a comprehensive lens through which to analyze the evolving nexus between terrorism and information warfare and give concrete examples of what, up until now, was rather conceptual.

3.3.1 The case of Jihadists

The convergence of disinformation and terrorism constitutes a substantial threat to both national and international security. While propaganda has historically underpinned radicalization, contemporary jihadist groups have refined this tactic, harnessing digital disinformation as both a recruitment mechanism and a tool to undermine state legitimacy. The digital age has transformed the operational landscape As Ingelevič-Citak and Przyszlak observe, the internet facilitates not only propaganda dissemination but also mobilization, training, and resource acquisition.[123]. Furthermore, they engineer parallel realities where disinformation serves simultaneously as an ideological weapon and a tactical asset. For such reasons, Lakomy, a scholar expert in online terrorism and information warfare, defined the Islamic State (IS) as "propaganda machine".[124] Specific analyses of jihadist organizations, such as Al-Qaeda and the Islamic State (IS), provide further evidence of propaganda's central role. An example would be Conroy and Al-Dayel's study on ISIS's "No Respite" video illustrates how propaganda embeds a calculated combination of factual elements and distorted data to manipulate perceptions of history and society.[125] Such communication strategies are deliberately propagandistic, blending selective statistics with emotional appeals to shape the cognitive frameworks of target audiences. Similarly, Lakomy's examination of IS's digital magazines demonstrates how these media leverage simplified historical

[123] INGELEVIČ-CITAK and PRZYSZLAK, *Jihadist, Far-Right And Far-Left Terrorism in Cyberspace – Same Threat and Same Countermeasures?*, International Comparative Jurisprudence, 2020, p. 159.
[124] LAKOMY, *Recruitment and Incitement to Violence in the Islamic State's Online Propaganda: Comparative Analysis of Dabiq and Rumiyah*, Studies in Conflict and Terrorism, 2019, p. 577.
[125] CONROY and AL-DAYEL, *Identity Construction through Discourse: A Case Study of ISIS's No Respite Video*, Studies in Conflict and Terrorism, 2020, p. 13.

narratives and justification of violent jihad by choosing selective religious references to justify, thus framing it as a religious duty.[126]

Nevertheless, not only are terrorist narratives specifically pervasive, but they have a wide range of action too: Chiluwa's study reveals how local jihadist groups interweave their narratives with those of the global jihadist movement, reinforcing universal ideological themes while preserving contextual resonance.[127] This intertextual strategy fortifies both local and global objectives, demonstrating a calculated layering of propaganda to sustain and expand extremist influence.

Jamil Ammar, a scholar expert on violent disinformation and balancing free speech and religious freedom, delineates three principal disinformation techniques employed by three jihadist groups being the Organization for the Liberation of the Levant, or Hay'at Tahrir al-Sham (HTS), in Syria, the IS in Iraq and the LVU ("*Lagen med särskilda bestämmelser om vård av unga"* or also called "The Swedish Care of Young Persons Act")[128] campaign in Sweden.

In the first case, the HTS, is quite interesting as they consider themselves to be moderate with a "new media strategy",[129] to which, according to two jihadi scholars Mohamad Al-Maqdisi and Abdulmonem Mustafa Halimah, is based on: (1) the foundational source of Jihad and Islamic principles;[130] (2) a more tame public communication without the display of atrocities, without referencing other jihadi groups like, for example, al-Qaida, and finally, avoiding unrealistic or unreasonable plans to institute a Caliphate.[131] While all these points are respected, Ammar states that the latter is not always followed, although HTS has maintained a strong hold of territorial integrity in the north of Syria.[132]

[126] LAKOMY, *Recruitment and Incitement to Violence in the Islamic State's Online Propaganda: Comparative Analysis of Dabiq and Rumiyah*, cit., p. 577.
[127] CHILUWA, *The Discourse of Terror Threats: Assessing Online Written Threats by Nigerian Terrorist Groups*, Studies in Conflict and Terrorism, 2016, pp. 318 – 338.
[128] Advokatfirman Segerström, *Legal advice – The Swedish Care of Young Persons Act (LVU*), Advokatsegerstrom.
[129] AMMAR, *Disinformation: The Jihadists' New Religion*, edited by ARCOS, CHIRU, IVAN, *Routledge Handbook of Disinformation and National Security*, Abingdon, 2023, p. 112.
[130] *Ivi.*, p. 113.
[131] *Ibidem*.
[132] *Ibidem*.

The second case, the IS, focuses mostly on spreading disinformation campaigns with a combination of two disinformation techniques, which were coined by Martin Innes: "Fogging" and "Flooding". The first technique consists of «the spreading of multiple interpretations of an event related, in our case, to IS or to its supporters or allies »[133] to create confusion. The second is the mass posting of content on multiple platforms to ensure that this confusing context remains.[134] These disinformation tactics are amplified by the extensive use of social media like Telegram, where IS utilizes a structured network of five distinct channel types for different functions:[135] starting from class A, which are the tamer channels with journalistic content covering real political and military events without overt IS branding or graphic content, to class E which are the type of temporary channels containing the most graphic content.[136] Ammar notes that Class A channels contain two further disinformation tactics that are, firstly, the masking of IS operations as accidents, and secondly, the framing of state action against IS as unjust persecution.[137]

The third case analyzed by Ammar is the LVU campaign in Sweden, which was happening at the time of the publication of his research. Now that a couple of years have passed since this publication, it is possible to take a distant approach to the overall phenomenon. The LVU disinformation campaign constitutes a targeted disinformation effort aimed at eroding trust in Swedish social services by falsely alleging systematic misuse of the LVU to forcibly remove children, particularly from immigrant and Muslim families.[138] Emerging in 2021 and intensifying through 2022 and 2023,[139] the campaign proliferated across social media platforms, with amplification by radical Islamist groups, thereby exacerbating distrust towards Swedish authorities;[140] an issue flagged as a substantial security risk by both the Swedish Security Service and the Swedish Defense University.[141] The Swedish National Agency for Education reported a tangible spillover effect on the education sector, noting heightened anxiety among students, staff, and guardians, alongside a broader erosion of confidence in public institutions. However, Ammar

[133] *Ivi.*, p. 114.
[134] *Ibidem*.
[135] *Ibidem*.
[136] *Ivi.*, p. 114 – 116.
[137] *Ivi.*, p. 114.
[138] *Ivi.*, p. 116.
[139] GIANDOMENICO J. et al., *Disinformation landscape in Sweden (V2)*, EUDisinfoLab, 2025, pp. 3– 4.
[140] *Ibidem*.
[141] *Ibidem*.

notes that this campaign was so pervasive that not only did it affect radical Islamist groups, but also scholars believed in it, thus endorsing these efforts.[142] This worries Ammar as he reminds of the case of Samuel Paty, where the spread of disinformation led to the beheading of a French teacher.[143] To conclude on the LVU campaign, the disinformation rapidly transcended national borders, extending into Finland and Denmark.[144] A report from EUDisinfo informs that, in response to these developments, the Swedish government has allocated a four-year funding package to the National Board of Health and Welfare, explicitly designed to counter and mitigate disinformation targeting social services.[145]Ammar concludes these three analyses, stating that:

> «*The Jihadist new media strategy intentionally combines both hate speech and disinformation. This new media strategy raises novel challenges to national security. Disinformation that does not meet the threshold of illegal speech—such as IS disinformation strategies used to construct and transmit disinformation under the guise of journalistic credentials or the orchestrated campaign again the Swedish social services, as mere examples — often stays online for an extended period of time*. »[146]

In his conclusion, he then further explains that both peaceful Salafi groups and terrorist organizations like IS exploit a shared sense of injustice to spread disinformation with similar persuasive efficacy despite divergent methods of violence, blurring distinctions between them.[147] Ammar supports the idea that by invoking real grievances, even non-violent actors reinforce false narratives, complicating further countermeasures since ideological alignment drives both the spread and resistance to debunking.[148] Besides, as seen from the LVU case, not all individuals who

[142] AMMAR, *Disinformation: The Jihadists' New Religion*, cit., p. 116.
[143] *Ibidem*.
[144] EUDisinfoLab, *Disinformation landscape in Sweden (V2)*, cit., pp. 3– 4.
[145] *Ibidem*.
[146] AMMAR, *Disinformation: The Jihadists' New Religion*, cit., p. 117.
[147] *Ibidem*.
[148] *Ibidem*.

fall for these disinformation campaigns are extremists; even common folk may end up believing this false information. Furthermore, these new media strategies by jihadists make it even harder to detect propaganda as they no longer use the display of graphic images or violent terminology extensively. The complexity of this challenge has been further exacerbated by geopolitical events such as Russia's invasion of Ukraine, especially as *Meta* platforms had revised their term of service rules, allowing for calls to violence against Russia and Putin, thus allowing hate speech.[149] This, in turn, reinforces extremist narratives of victimization and perceived double standards. Ammar ultimately warns that unless these dynamics are critically addressed, efforts to mitigate extremist propaganda and its violent consequences will remain tenuous.[150]

3.3.2 The Case of Russia

The instrumentalization of *dezinformatsiya* by the Russian Federation is not a recent occurrence.[151] Rather, it constitutes a continuation of a longstanding doctrine rooted in Soviet-era *aktivnye meropriyatiya* (active measures), where deception, propaganda, and psychological operations functioned as core mechanisms to destabilize adversaries, distort public discourse, and advance both political and military objectives.[152] As will be seen in this section, these practices have not merely persisted but evolved post-Cold War era.

At the conceptual core of Russia's disinformation architecture lies the methodology encapsulated by what Randolph H. Pherson et al. in their paper "Historical Disinformation Practices Learning from the Russians" call the "Five D's", being: Dismiss, Distract, Deflect, Distort, and Distrust.[153] This systematic approach has been operationalized to obscure Russia's involvement in

[149] *Ivi.*, p.118.

[150] *Ibidem*.

[151] HOSAKA, *Cold War Active Measure*s, edited by ARCOS, CHIRU, IVAN, *Routledge Handbook of Disinformation and National Security*, Abingdon, 2023, p. 45.

[152] *Ibidem*.

[153] PHERSON, LABRINY, DIORIO, *Historical Disinformation Practices: Learning from the Russians*, edited by ARCOS, CHIRU, IVAN, *Routledge Handbook of Disinformation and National Security*, Abingdon, 2023, p. 61.

international conflicts,[154] legitimize contentious domestic and foreign policies,[155] and erode Western political cohesion.[156] Empirical illustrations of this tactic are plentiful. A notable example in Randolph H. Pherson et al.'s paper is the downing of Malaysia Airlines Flight MH17, over Ukraine. Despite overwhelming evidence[157] tying the missile system directly to Russia,[158] more specifically the Russian 53rd Anti-Aircraft Missile Brigade,[159] Russia embarked on a sustained disinformation campaign to obscure its involvement using all five D's.[160] As Randolph H. Pherson et al. analyze:

> «*Russia tried to dismiss the allegations and deflect blame by refocusing attention and claiming it had evidence of Ukraine's involvement. The press conferences held two years after the accusations and four years after the crash served to distract the international media and the public from other more recent or ongoing crimes. Russian officials sought to deflect the international media and the public from other more recent or ongoing crimes. Russian officials attempted to distort the facts by stating that the missile launcher was never returned to Russia despite video evidence of the missile launcher's transport that had been released two years earlier. Russian propagandists tried to instill doubt and distrust in the investigation by dismissing the video footage as fake and propagating other Digital Disinformation.* »[161]

[154] See involvement with Assad Regime and Russia's use of disinformation to Distort, Dismiss, Deflect, Distrust. *Ivi.*, p. 65.

[155] See Transnistria operations and Russia's use of disinformation to Distrust, Distort and Distract. *Ivi.*, p. 62.

[156] See involvement with Assad Regime. *Ivi.*, p. 65.

[157] The paper reports evidence of satellite imagery, photographs, forensic evidence, intercepted communications, and recovery of missile fragments. *Ivi.*, p. 67.

[158] *Ibidem*.

[159] *Ivi.*, p. 68.

[160] *Ivi.*, pp. 68 – 69.

[161] *Ivi.*, p. 69.

Russian disinformation efforts have not been confined to conflict zones, as they have also systematically targeted Western democratic institutions and processes. The Kremlin's interference in the 2016 United States presidential election[162] and its influence campaigns during the UK's Brexit referendum[163] illustrate their deliberate strategy of intensifying societal cleavages through cyber-enabled information operations. Another famous example, discussed earlier in this chapter, is the disruption of Emmanuel Macron's 2017 presidential campaign in France[164]. Although these operations achieved variable tactical success, their strategic significance lies in their demonstration of Russia's adaptive and iterative disinformation methodology. Besides, also public health crises have served as strategic vectors; disinformation with Operation INFEKTION (falsely attributed the origin of HIV/AIDS to United States bioweapons research)[165] or even recently, during the COVID-19 pandemic, Russian media outlets have been active in spreading hostile narratives and MDM.[166] These narratives are congruent with Russia's broader strategic objective of diminishing international cooperation while simultaneously advancing its own state-sponsored alternatives. Moreover, a notable manifestation of Russia's relationship with terrorism is the Kremlin's appropriation of the term "terrorism" as a tool of delegitimization and obscuring underlying objectives – e.g., branding all Syrian opposition groups as terrorists.[167]

Not only do Russian media outlets spread disinformation, but also public figures; President Vladimir Putin's article, which appeared on "National Interest" talking about the legacy of the Second World War, was repurposed to bolster state propaganda and framing contemporary military aggression as a continuation of Russia's historical role as a liberator.[168] Domestically, the manipulation of Russian public opinion through disinformation remains foundational to regime stability. An example of this would be the full-scale invasion of Ukraine in 2022. State-controlled

[162] *Ivi.*, p. 73.
[163] *Ivi.*, p. 76.
[164] *Ivi.*, p. 77.
[165] KIVIMÄKI, *Open-Source Information for Intelligence Purposes: the Challenge of Disinformation*, cit., p. 286.
[166] European Commission, *Tackling coronavirus disinformation*, Strategy and Policy, 2020.
[167] PHERSON, LABRINY, DIORIO, *Historical Disinformation Practices: Learning from the Russians*, cit., p.65.
[168] BJOLA, MANOR, *The Use and Abuse of History by Russian Embassies on Twitter*, edited by ARCOS, CHIRU, IVAN, *Routledge Handbook of Disinformation and National Security*, Abingdon, 2023, pp. 154 – 157.

media systematically suppress dissent and construct narratives of external threat to consolidate domestic support. The Ukrainian conflict exemplifies this perfectly, as «Russian President Vladimir Putin's aggressive disinformation convinced most Russians that Ukraine had initiated hostilities against Russians and that the conflict was not a war but a 'special military operation'.»[169]

As such, Russia has progressively refined a potent capacity to manipulate public opinion through online disinformation. Although Russia's efforts remain a considerable threat, R.H. Pherson et al. argue that emerging defensive strategies, as seen in the West's reaction to disinformation during the Ukraine invasion, show increased resilience and proactive measures.[170] However, as discussed in this section, Russian meddling in elections and the spread of anti-Western narratives indirectly contribute to the undermining of democratic governance, which in turn, diminishes trust in liberal democratic institutions; thus aligning, even if unintentionally, with the strategic aims of terrorist groups.

4. Conclusion

In conclusion, this chapter has demonstrated that the phenomenon of disinformation, now accelerated by digital technologies, artificial intelligence, and machine learning, is a complex and escalated threat to national and international security. The evolution from early, centralized propaganda to the current era of decentralized, algorithm-driven manipulation marks a paradigm shift in how falsehoods are crafted, disseminated, and received. The advent of deepfakes and synthetic media signals a potential fifth wave of disinformation, where the boundaries between reality and fabrication become increasingly blurred, undermining public trust and complicating detection and response efforts. In the realm of information warfare, the strategic interplay of offense, defense, and deterrence in information warfare, as theorized by Ünver and Ertan, highlights the complex dynamics between Attackers and Defenders, stressing the need for adaptive and integrated countermeasures. Furthermore, the chapter has illustrated how both terrorist

[169] PHERSON, LABRINY, DIORIO, *Historical Disinformation Practices: Learning from the Russians*, cit., p.60.
[170] *Ivi.*, p.78.

organizations and state actors have misused disinformation techniques to achieve their strategic objectives, disrupt society, and destroy the reasoning and legitimacy of institutions. Finally, understanding the conceptual and practical linkages between MDM and terrorism is essential for developing effective strategies to safeguard democratic resilience and global security in the face of rapidly evolving informational threats. However, now that the conceptual basis has been laid down, it would be interesting to analyze what is the factual contribution of international organizations in the fight against MDM, which will be analyzed in the following chapters.

CHAPTER II
The Role of the United Nations in Countering MDM and Terrorism

Summary: 1. Introduction. – 2. The legal basis and main initiatives. – 2.1 Art. 19 and Art. 20 International Covenant on Civil and Political Rights. – 2.2 UN's Strategy and Plan on Action on Hate Speech. – 2.3 The Verified Initiative. – 2.4 Resolution GA 75/309. – 2.5 United Nations Global Principles for Information Integrity and the upcoming Code of Conduct – 3. The role of the Security Council in the fight against MDM and Terrorism. – 4. Other United Nations Offices and Projects. – 4.1 CTED-ICT4Peace Joint Projects – 4.2 UN Global Counter-Terrorism Strategy. – 5. UN Education, Scientific and Cultural Organization – 6. Conclusion.

1. Introduction

Since its creation, the United Nations (UN) has attempted to play an important role in the international context and in addressing global issues such as terrorism and hate speech. Recognizing that these two issues are intrinsically interconnected, since hate speech often incites acts of radicalization and violence, the UN has sought to develop a complete, multilayered approach within the organization as it pertains to its founding principles contained within the UN Charter, namely - the promotion of international peace and security, the advancement of human rights, and the safeguarding of fundamental freedoms. This not only enables the UN to engage with such a complex issue but also obliges it to adopt a proactive stance against hate speech and the dissemination of MDM, particularly when these phenomena contribute to violent extremism or the recruitment of violent extremists in the affected countries.

The UN's dedication to these issues can be found under the great umbrella of UN's issues/campaigns. In fact, the organization has multiple core areas of focus which all can be found as their general main areas of action, two of which should be brought to attention being

“Countering Terrorism” and “Hate Speech”. The inclusion of these issues as one of their primary themes under the UN’s umbrella is not only a reflection of the understanding of the growing concern from the international community, but a call for a coordinated effort to mitigate these phenomena. Terrorism, as the UN has consistently been clear about, is not only a threat to international and national security but is also a fundamental menace to international peace[171] and stability,[172] challenging the rule of law and human rights. Likewise, the growth of online platforms and the use of MDM has allowed an increase in hate speech which in turn intensified tensions within and in between societies causing these platforms to potentially foster a fertile environment in which extremist ideologies can flourish.

Considering these circumstances, the UN has expanded its scope and engagement in both countering terrorism and countering hate speech to address directly the role of MDM[173] and online radicalization[174] in current conditions of modern security threats. By tackling these two fields, the UN aims to strengthen overall global resilience to terrorism and its broader aims of security, peace and human rights, so that it remains achievable within an increasingly digitized environment with asymmetric threats.

To analyze the UN's contribution on the fight against MDM and terrorism, this chapter adopts a structured approach that examines both normative foundations and institutional initiatives. It begins by outlining the legal basis and principal frameworks that underpin the UN’s action, with particular emphasis on the balance between freedom of expression and limitations aimed at countering hate speech and incitement. This first section discusses Article 19 and Article 20 of the International Covenant on Civil and Political Rights, followed by key initiatives such as the UN Strategy and Plan of Action on Hate Speech, the Verified Initiative, General Assembly Resolution

[171] UNSC, *Security Council resolution 2178 on threats to international peace and security caused by foreign terrorist fighters*, 2014, p. 1.

[172] MOHAMMED, *Strengthening Regional Cooperation and Institution Building to Address the Evolving Threat of Terrorism in Africa*, United Nations Statements, Abuja, 2024, p. 1.

[173] KUMAR, *UN unveils 'Global Principles' to combat online misinformation, hate speech*, Business Standard, 2024, p. 1.

[174] UNGA, *As Terrorists Exploit Online Platforms, Sixth Committee Delegates Urge Holistic, Global Collaboration with Governments, Civil Society, Internet Providers*, United Nations Press Releases, 2019, pp. 1 – 3.

75/309, and the recently adopted United Nations Global Principles for Information Integrity alongside the forthcoming Code of Conduct. The chapter then turns to the role of the Security Council in addressing the nexus between terrorism and MDM, particularly through its counterterrorism mandates and resolutions. Finally, it examines the contribution of other UN bodies and projects, including the work of UNESCO, the CTED-ICT4Peace joint initiatives, and the UN Global Counter-Terrorism Strategy, thereby providing a comprehensive overview of the multilevel efforts undertaken by the United Nations to mitigate the threats posed by manipulated information in the context of global security.

2. The UN legal basis and main initiatives

2.1 Art. 19 and Art. 20 International Covenant on Civil and Political Rights

While the UN seems to have broadened its scope by wanting to include the phenomenon of MDM and counterterrorism, it is important to analyze if the UN have the legal basis for contrasting MDM. To answer this, it is crucial to introduce the International Covenant on Civil and Political Rights (ICCPR). The ICCPR has been adopted by the United Nations General Assembly (UNGA) with Resolution 2200A (XXI), allowing it to enter into force by March 1976. This covenant not only represents a big step in favor of international law and human rights, but it provides a critical legal framework for specifically the balancing between freedom of expression and the need to counter harmful content. Namely, Article 19 and Article 20 of the ICCPR are the ones concerning this important provision, which also articulates the parameters within which states may regulate hate speech or MDM while still upholding fundamental freedoms.

Article 19

1. *Everyone shall have the right to hold opinions without interference*
2. *Everyone shall have the right to freedom of expression; this right shall include freedom to seek, receive and impart information and ideas of all*

> *kinds, regardless of frontiers, either orally, in writing or in print, in the form of art, or through any other media of his choice.*
>
> 3. *The exercise of the rights provided for in paragraph 2 of this article carries with it special duties and responsibilities. It may therefore be subject to certain restrictions, but these shall only be such as are provided by law and are necessary:*
>
> a. *For respect of the rights or reputation of others;*
>
> b. *For the protection of national security or of public order (ordre public), or of public health or morals.*

Article 20

> 1. *Any propaganda for war shall be prohibited by law.*
> 2. *Any advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence shall be prohibited by law.*

As such, Article 19 of the ICCPR focuses on guaranteeing the right to the freedom of expression as it is a fundamental right everyone is entitled to have, however, in paragraph 3 it does allow restriction if necessary. In fact, it states that it may be required to restrictions in other to «respect of the rights or reputation of others » and for «the protection of national security or of public order or of public health or morals », but that only if it meets the two criteria of necessity and is already provided by law. The carefulness of allowing restriction to only certain cases[175] seems to try and limit any abuse of power of states, whether intentional or as an accident, and thus upholding this human right. Article 20 further refines the scope of freedom of expression by establishing a clear obligation for states. It requires the prohibition of «any advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence ». However, while these two articles establish what could be considered a legal foundation for the countering for hate speech and MDM, their implementation varies across member states. The ICCPR establishes key

[175] UNGA, *International Covenant on Civil and Political Rights*, Art. 19(3), 1966, p. 11.

principles that must guide any limitation of rights, such as the necessity of the measure, its basis in law, and its relevance to the protection of national security.

However, while states that ratify the Convention are bound to respect these principles, the ICCPR does not prescribe specific methods for their implementation, leaving states discretion in determining how to give effect to them. With that being said, the broad language employed in these articles also leaves room for interpretation which could result in divergent enforcement practices and thus in global inconsistency. For example, some States may invoke these articles to justify overly restrictive measures of diverging opinions, thus suppressing dissent, by considering them necessary for the protection of national security and *ordre public*. This raises concerns about the potential misuse of regulatory frameworks under the guise of combating terrorism[176], although the same could be said for hate speech and MDM. This proves that in this field, as many others, needs not only for careful monitoring by neutral third-party international organizations like the UN, but also that the texts are clear in their meaning and will, as to ensure that the fight against counterterrorism and MDM are not excuses to restrain civil liberties by malicious interpretation of the provisions.

As the only international organization with truly global reach,[177] one of the UN's key strengths lies in its capacity to address issues within its mandate by convening a diverse range of actors. These include states, technology companies, NGOs, other international organizations such as the EU, and many others, fostering cooperation and bringing these issues to the attention of the international community. By promoting cooperation across international, regional and national platforms, the UN creates the basis of comprehensive strategies that allows multi-layered responses that integrate many fields from legal to educational approaches. Nevertheless, the extent to which these efforts can take hold does depend on the willingness of each member state to implement and actively defend international norms. While it is important to consider that States should always act

[176] NÍ AOLÁIN, *Visit to Kazakhstan: Report of the Special Rapporteur on the promotion and protection of human rights and fundamental freedoms while countering terrorism*, UN Digital Library, 2020, p. 9.

[177] «As the world's only truly universal global organization, the United Nations has become the foremost forum to address issues that transcend national boundaries and cannot be resolved by any one country acting alone. » UN, *Global Issues*, United Nations, p.1.

according to the general principle of “good faith”,[178] and that the opposite cases should be considered as exceptions in international law, the reluctance to actively defend or implement a provision could be due to specific circumstances, e.g. *force majeure* or simply other priorities. Thus, a persistent challenge, in this regard, is the absence of a robust enforcement mechanism in the UN system. The ICCPR indeed places legal obligations on the ratifying states following Article 51 of said Convention[179] and the principle of *pacta sunt servanda* – which the latter is not only customary international law but also is found in Article 26 of the Vienna Convention on the Law of Treaties[180] (VCLT) – however, it doesn’t provide a direct enforcement structure beyond the periodic reviews done by the Human Rights Committee (CCPR). The fact that it relied upon voluntary compliance along with a peer review mechanism diminishes the UN’s ability to promote consistency across jurisdictions, and that is all without the interpretation argument explained earlier which further complicates the efforts to establish a cohesive and global response to hate speech and MDM.

Beyond issues with the ICCPR, another obstacle to effective implementation would come from the absence of a universally accepted definition of hate speech. As seen from the introduction, even the UN states that there is no agreement on a definition for the MDM, and hate speech is no different. The subjective nature of what constitutes «incitement to discrimination, hostility, or violence »[181] makes the enforcement of this provision more of a challenge and thus increases further the risk of inconsistent application across national jurisdiction. Some legal systems may offer more narrow definitions where the emphasis is on direct incitement to certain forms of harm, while other legal systems might apply broader definitions of hate speech that may include expressions that, while controversial or offensive, do not necessarily constitute incitement. This wide range of different interpretations raises concern regarding the potential for overreach, especially but not limited to authoritarian or semi-authoritarian political regimes, where hate

[178] REINHOLD, *Good Faith in International Law*, UCL Journal of Law and Jurisprudence, 2013, pp. 40–63.
[179] UNGA, *International Covenant on Civil and Political Rights*, cit., Art. 51(3), p. 24.
[180] International Law Commission, *Vienna Convention on the Law of Treaties*, Art. 26, 1969, p. 11.
[181] UNGA, *International Covenant on Civil and Political Rights*, cit., Art. 20, p. 11.

speech restriction can be used as tools for political repression rather than genuine efforts to combat a harmful rhetoric.

Before start the analysis of concrete UN projects and efforts in regard to terrorism and MDM, it is important to resume the situation: as the UN continues to work on how it tackles these issues and refine its strategies for addressing hate speech and MDM more effectively, it needs to navigate the definitions and enforcement challenges while continuing to commit to uphold fundamental freedoms. With better and generally agreed definitions, the UN can reinforce its role as a guardian of both security and human rights in this new digital age.

2.2 UN's Strategy and Plan on Action on Hate Speech

One of the first system-wide initiatives which was specifically aimed at addressing the issues of hate speech comprehensively is the United Nations Strategy and Plan on Action on Hate Speech, which was launched in 2019. It defined the phenomenon of hate speech as «any kind of communication in speech, writing or behavior, that attacks or uses pejorative or discriminatory language with reference to a person or a group on the basis on who they are, in other words, based on their religion, ethnicity, nationality, race, colour, descent, gender or other identity factor ».[182] Besides giving a working definition, what is most notable about this initiative is that it focuses on the recognition that hate speech, and particularly in the digital age, is a powerful fuel to discrimination, social fragmentation[183] and even, in more extreme cases, violence and terrorism.[184] However, the link between hate speech with MDM and terrorism might be confusing. To explain it, it is crucial to consider a UN factsheet released in the context of the #NoToHate campaign which started in 2023, which explains the relation between hate speech and MDM:

[182] United Nations, *United Nations Strategy and Plan of Action on Hate Speech*, 2020, p. 8.

[183] GUTERRES, *Foreword by the Secretary-General United Nations Strategy and Plan of Action on Hate Speech*, 2020, p. 3.

[184] *Ibidem*.

> «Mis- and disinformation and hate speech are related but distinct phenomena, with certain areas of overlap and difference in how they can be identified, mitigated and addressed. All three pollute the information ecosystem and threaten human progress. »[185]

The UN, therefore, focuses on the area where the dimensions of MDM and hate speech overlap and intersect. The text goes on to explain that hate speech has been considered to be a precursor to many of the recent mass atrocities[186] that the world has been faced with, which is not to be surprised of since public discourse as a tool for own goal has been used extensively during the course of history. In all of this, social media is just the new platform where these activities take place. Finally, the document explains that «as history continues to show, hate speech coupled with disinformation can lead to stigmatization, discrimination and large-scale violence »[187]. It emphasizes that, beyond being a persistent challenge, preventing disinformation could contribute to addressing these serious issues of widespread violence and stigmatization.

This Strategy and Plan on Action is built on international human rights frameworks[188], including the ICCPR and Rabat Plan of Action,[189] to maintain the overall UN's path of balancing freedom of expression with the necessity to counter harmful discourse. The approach chosen by the UN is deeply rooted in the need to promote a preventive and united response,[190] as such, the strategy does not simply seek to react to instances of hate speech but aims to build a better and comprehensive framework that addresses its root causes, societal manifestations and potential long-term consequences. The emphasis is given on commitments such as monitoring and analyzing patterns

[185] UN, *NoToHate Factsheets for the United Nations Strategy and Plan of Action on Hate Speech*, 2023, p. 1.

[186] UN, *Secretary-General Launches United Nations Strategy and Plan of Action against Hate Speech, Designating Special Adviser on Genocide Prevention as Focal Point*, UN Press Releases, 2019.

[187] UN, *#NoToHate Factsheets for the United Nations Strategy and Plan of Action on Hate Speech*, cit., p.2.

[188] GUTERRES, *Foreword by the Secretary-General United Nations Strategy and Plan of Action on Hate Speech*, cit., p. 3.

[189] UN, *United Nation Strategy and Plan on Action on Hate Speech*, cit., p. 8.

[190] *Ivi.*, p. 3.

of hate speech,[191] engaging with media and digital platforms[192] and by supporting member states in aligning their national policies with the international human rights standards.[193] This first initiative seems to close the gap between principle and practice with this coordinated effort which spans the different dimensions of legal, educational and technological areas as to ensure that the fight against hate speech does not undermine fundamental freedoms. Hence, by framing hate speech as both a human rights issue but also as a broader challenge to international peace and security, the UN reinforces the interconnection between speech regulation, social stability and democratic resilience.

At the core of this Strategy and Plan on Action are a total of thirteen commitments, some of which are already mentioned, which all serve as guiding principles for action. The commitments establish the need for a holistic and collaborative approach, one that brings together legal frameworks, social interventions, and policy developments. More precisely, the commitments are: (1) strengthening the monitoring and analysis of hate speech, (2) addressing the root causes of hate speech through education and advocacy, (3) supporting the victims, (4) establishing strategic relationships with relevant actors, (5) working with new and traditional media to counter online hate narratives, (6) using technology, (7) harvesting education as a tool against hate speech, (8) working with civil societies to address these root causes, (9) Aiding with advocacy, (10) Creating guidance for better external communications, (11) having partnerships, (12) Educating the UN staff, (13) supporting the Member States. The strategy also commits to develop laws in accordance with international human rights standards,[194] and to ensure that efforts to respond to hate speech avoid the risk of being misused for political repression. Finally, this plan seeks to embed these commitments within a broad multi-sectoral strategy, premised on the belief that this will achieve a sustainable and complementary approach for addressing hate speech[195] and its impact as an ever-evolving phenomenon, online and offline.

[191] *Ivi.*, p. 25.
[192] *Ivi.*, p. 33.
[193] *Ivi.*, p. 50.
[194] *Ivi.*, p. 4.
[195] Ivi, p. 8.

Despite its ambitious and theoretically comprehensive framework, the strategy faces significant challenges in practical implementation. While international law clearly prohibits incitement to violence and discrimination, the broader scope of the UN's strategy requires a degree of subjectivity. This, in turn, creates the risk that certain regimes may misuse hate speech regulations as tools of political repression against dissenting opponents. Moreover, although the UN acknowledges the growing role of digital platforms in amplifying hate speech, engaging with technology companies presents significant complexities. The commercial logic that drives these companies often runs counter to regulatory efforts. As a result, balancing profit-oriented priorities with the obligations of a joint initiative concluded with another international organization, for example, can prove challenging for them. The reliance on self-regulation within the private sector, as well as the lack of binding enforcement mechanisms at the international level, proves the existence of a fundamental tension between combating harmful discourse and indolently letting it happen as the private sector might have different priorities. This example tells of how the UN has to play the role of the balancer and defender of rights with actors who are not only states but also act in the private and commercial sectors, guided by different principles and logic than a state would.

Finally, it is important to highlight the UN Strategy and Plan of Action on Hate Speech as a significant initiative in the global effort to counter harmful discourse that fuels discrimination and violence. It represents the first official recognition of this phenomenon at the UN level. Given the UN's global reach, the initiative has drawn considerable attention from the international community. However, it is not the only initiative in this field; rather, it is one of the first in a series of measures that have emerged recently and are likely to continue in the future.

A critical appraisal of the United Nations Strategy and Plan of Action on Hate Speech, as elaborated by Audrey Fino, reveals both normative advancements and substantive limitations inherent in the document's legal and operational framework. On the one hand, Fino recognises the

Strategy and its accompanying Guidance as «welcome additions »[196] that have contributed meaningfully to enhancing institutional capacities within UN field missions and among Member States. By embedding the Rabat Plan of Action and ARTICLE 19's "hate speech pyramid" into its operational logic, the Strategy proposes a graduated typology of hate speech that could, in theory, facilitate proportionate and rights-compliant responses.[197] Moreover, its deliberate expansion of identity factor – encompassing not merely nationality, race, and religion, but also gender, sexual orientation, language, and political opinion – aligns with an evolutive reading of international human rights law and reflects a commendable attempt to respond to contemporary societal complexities.[198]

Nevertheless, Fino's critique systematically exposes critical legal and conceptual shortcomings that undermine the Strategy's coherence and legal certainty. Most critical among these is the Strategy's adoption of an expansive and imprecise definition of hate speech, which fails to clearly distinguish between forms of expression legitimately subject to prohibition under Article 20(2) of the ICCPR and those that fall within the protective ambit of freedom of expression. Instead it using vague language like «incitement to discrimination, hostility and violence », without aligning clearly with the previously mentioned ICCPR article, nor Article 4 of the International Convention on the Elimination of all forms of Racial Discrimination (ICERD).[199] This imprecision, Fino contends, risks destabilising the delicate balance enshrined in international law between countering harmful speech and safeguarding fundamental freedoms.[200] Even more ambiguity comes from the interaction between the Rabat Test and the Article 19(3) necessity and proportionality test, which remains insufficiently articulated, leaving practitioners uncertain about their cumulative or sequential application.[201]A second substantive weakness identified is the Guidance's extension of the scope of Article 20(2) ICCPR to include identity factors beyond those explicitly enumerated in the treaty text, absent clear justification grounded in treaty law. Fino warns that such an expansive interpretation, applied to a provision that constitutes a restriction on free speech,

[196] FINO, *A critique of the UN Strategy and Guidance on 'Hate Speech': Some Legal Considerations*, edited by FORTIN et al., *Netherlands Quarterly of Human Rights*, 2023, p. 191.
[197] *Ivi.*, pp 191 – 192.
[198] *Ivi.*, p. 194.
[199] *Ibidem*.
[200] *Ivi.*, p. 196.
[201] *Ivi.*, p. 198.

contravenes the well-established principle of narrow construction of limitations clauses. Equally problematic is the omission of Article 20(1) ICCPR from ARTICLE 19's report in which the Strategy has been based on despite its obvious relevance to identity-based incitement and its historical significance in the drafting of the Covenant.[202] Fino comments that perhaps it is the reason why there is no mention of propaganda for war.[203] Finally, Fino criticizes the underestimation of the harm posed by certain categories of expression – such as Holocaust denial and deliberate disinformation – which the Strategy relegates to the lower levels of the pyramid despite evidence of their capacity to inflict significant societal damage.[204]

2.3 The Verified Initiative

A year later, in 2020 during the emergency of the global pandemic caused by COVID-19, the UN launched the "Verified Initiative" as an answer to the imminent threat of the spread on MDM[205] about the virus SARS-CoV-2 responsible for the pandemic, but also common MDM of the period such as MDM about vaccines, and about various governments (e.g. Chinese Government). As such, this initiative was launched as an effort to counter this by promoting reliable, fact-based information. This UN answer emerged as a direct response to what has been defined as "infodemic",[206] wherein conspiracy theories, manipulated content, and misleading claims proliferated across digital platforms, exacerbating public health crises while also eroding the trust in institutions and governments. Although the initiative initially focused on countering MDM during the global health emergency of COVID-19, its broader implications extended into areas crucial to global security. In particular, it has contributed to efforts to combat both extremism and terrorism.

[202] *Ivi.*, p. 207.
[203] *Ibidem*.
[204] *Ivi.*, pp. 206 – 207.
[205] United Nations Department of Global Communications, *'Verified' initiative aims to flood digital space with facts amid COVID-19 crisis*, United Nations, 2020, p.1.
[206] Verified, *Our Mission*, Verified, p. 1.

As such, the aim of the Verified Initiative is to leverage a network-based model to flood the social media and digital sphere with accurate and verified information through a combination of trusted sources, civil society actors, and digital platforms.[207] The logic here is to partner with the private sector of social media companies and non-governmental organizations to attempt to disrupt the virality of harmful content and replace it instead with data-driven, contextually relevant messaging that supports public resilience against misinformation. Notably, this strategy aligns with the broader UN objective aimed at strengthening both media literacy, and digital awareness but also counter-narrative production as a tool to mitigate the influence of extremist propaganda. Nevertheless, the effectiveness of the Verified Initiative remains contingent on several factors, including the similar problem as the UN's Strategy and Plan on Action on Hate Speech and so the willingness of social media companies to cooperate meaningfully with regulatory efforts, but also a factor linked specifically with social media: the "attractiveness" of false information. In fact, the most contagious and viral information seems to be the one that triggers strong emotions such as anger or sadness. Adding to that, a study[208] in 2021 discovered that negatively manipulated news are the most viral type of news, especially the ones that stem from factual elements which are employed to evoke these strong emotions: it wouldn't be a surprise if news containing malinformation are the most popular and viral information shared on social media. Thus, while the initiative has successfully reached over a billion people globally[209] (1,4 billion as reported by Purpose, a collaborator to the project), it still remains a challenge to counter false narratives online as, especially malinformation, seems to be more contagious than true information. Moreover, two other main challenges persist. The first issue concerns platform accountability and algorithmic transparency. MDM campaigns, particularly those associated with extremist groups, can exploit the very mechanisms that drive online engagement, making it challenging to limit their spread without risking infringements on fundamental freedoms. The second issue is that, while these types of initiatives are backed by UN agencies and other collaborators, their dependency on voluntary partnerships limits their enforcement capabilities, much like other UN-led MDM countermeasures. Nevertheless, the initiative represents an important step forward for the UN, as it provides a

[207] United Nations Department of Global Communications, *'Verified' initiative aims to flood digital space with facts amid COVID-19 crisis*, cit., pp. 1 – 4.

[208] CORBU et al., *Fake news going viral: The mediating effect of negative emotions*, Media Literacy and Academic Research, 2021, pp. 58–87.

[209] Verified, *Our Mission*, cit. p. 1.

concrete response in the fight against MDM. However, it does not explicitly address the issue of terrorism, focusing instead on the link between extremism and MDM. As a result, significant work remains to be done in countering this particular phenomenon.

2.4 Resolution GA 75/309

The General Assembly, a year later after the launch of the Verified Initiative, decided to remind of the danger of MDM in its Resolution 75/309 in the context of «Promoting interreligious and intercultural dialogue and tolerance in countering hate speech ».[210] Acknowledging the evolving nature of digital communication, shaped by the ongoing transformation of the digital sphere and the increasing prevalence of false narratives, this resolution calls for a better international cooperation among the Members States, capacity-building initiatives, and state-led efforts to counter online threats effectively, and in particular, it «invites Member States to support, in accordance with relevant international obligations, transparent and accessible systems to identify, track, collect data and analyze trends on hate speech, both in person and in digital contexts, at all national levels, as appropriate, to support effective responses ».[211]

The resolution recognizes the increasing implications of MDM on global security, pointing to its propensity to inspire divisions in society. It also mentions the United Nations Strategy and Plan of Action on Hate Speech and considers positively its efforts in creating partnerships which also aids in the promotion of tolerance that this resolution strives for.[212] But besides that, it explicitly condemns «any advocacy of hatred that constitutes incitement to discrimination, hostility or violence, whether it involves the use of print, audiovisual or electronic media, social media or any other means »:[213] as a result, the role of digital platforms, among others, in both dissemination and mitigation of the spread of false narratives also receives attention, with States and technology

[210] UNGA, *Resolution 75/309: Promoting interreligious and intercultural dialogue and tolerance in countering hate speech*, 2021, p. 2.
[211] *Ivi.*, p. 5.
[212] *Ivi.*, p. 4.
[213] *Ibidem*.

companies are encouraged to engage in coordinated efforts to identify and combat harmful content. This resolution, while reaffirming that freedom of expression remains the basis of democratic systems of governance,[214] emphasizes that the deliberate spread of false information – outlined in the previously identified Art.20 ICCPR – presents a pressing opposition that needs to be addressed and framed within legal and framework approaches, consistent with international human rights norms.[215] In fact, it stresses the need for «the dissemination of factual, timely, targeted, clear, accessible, multilingual and science-based information, and emphasizes the need for all Member States to stand together to address the challenge of disinformation and misinformation ».[216]

However, despite representing a positive step towards countering this phenomenon, the resolution continues to suffer from the same recurring shortcomings identified in the previously analyzed UN procedures. Furthermore, while the resolution strongly condemns the deliberate spread of falsehoods designed to incite violence or foster extremism, it does not establish clear criteria for distinguishing between harmful MDM and legitimate political discourse. As with other UN initiatives, this resolution is not exempt from the risk of misuse arising from the absence of a clear definition of MDM. This is particularly concerning in jurisdictions where broadly interpreted MDM laws have been used to suppress dissenting voices and restrict freedom of expression. A final critique to be added to this resolution relates to the lack of definition of MDM, since it mentions disinformation and misinformation but neither it explains the difference, nor mentions at all the phenomenon of malinformation which instead is confused with the other two. Despite this, Resolution 75/309 seems to remain a significant step in the global conversation on MDM and counterterrorism. Even if this resolution stems from the COVID-19 emergency, these principles found in Art.19 and Art.20 ICCPR are universally recognized for all cases, especially in something as important as terrorism and hate speech. The resolution contributes to ongoing discussions about the need for multilateral governance mechanisms that balance security imperatives with human rights protections which is done with the formal recognition of the role of digital misinformation in exacerbating security threats. Its emphasis on international cooperation,

[214] *Ivi.*, p. 2.
[215] *Ibidem*.
[216] *Ibidem*.

capacity-building, and the engagement of civil society actors reflects a broader commitment to addressing MDM while ensuring that regulatory responses do not inadvertently undermine democratic freedoms.

As of now, there appears to be a lack of scholarly articles that directly analyze United Nations General Assembly Resolution 75/309. However, the resolution has been the subject of discussion among various stakeholders, including member states and international organizations, highlighting differing perspectives on its implications for freedom of expression and the definition of hate speech. The most striking example is the UK, which expressed concerns that certain elements of the resolution could suggest limitations on freedom of expression beyond what is established in international human rights law.[217] They opposed any future attempts to agree on new definitions of hate speech at the UN level, including at the proposed conference in 2025.[218]

Finally, we could consider this resolution to be one of the many bases on which further resolutions and action plans will be built on as it is a very important document produced by the General Assembly on the issues of MDM and terrorism.

2.5 United Nations Global Principles for Information Integrity and the upcoming Code of Conduct

The latest UN initiative on these issues was finalized in June 2024, when the United Nations Global Principles for Information Integrity were adopted as a foundational framework to guide the development of a forthcoming code of conduct aimed at countering significant threats that continue to plague the international community, notably issues of MDM. These Global Principles acknowledge the role of manipulated information in not only deepening social divisions but also

[217] Foreign, Commonwealth & Development Office, *Promoting interreligious and intercultural dialogue and tolerance in countering hate speech: UK statement at the UN General Assembly*, Government of UK official website, 2023.
[218] *Ibidem*.

in inciting violence[219] and eroding trust in democratic institutions.[220] Their overarching aim is to strengthen the capacity of international organizations to mitigate digital threats, while simultaneously safeguarding human rights and freedom of expression.[221]

This initiative is an important reminder that digital platforms play a dual role as both amplifiers of harmful content and vehicles for counter-narratives that promote democratic resilience. As such, the document envisions a structured framework for engagement between governments, tech companies, news media, and more,[222] encouraging cooperation in tackling information pollution while preventing the misuse of regulatory tools to suppress legitimate dissent. Drawing from existing UN frameworks, such as the Strategy and Plan of Action on Hate Speech,[223] this normative framework document seeks to establish common standards for platform accountability, transparency, content moderation, and more recommendations.

The UN Global Principles for Information Integrity is divided into two main parts: the first one being the 5 principles for information integrity (societal trust and resilience; healthy incentives; public empowerment; independent, free, and pluralistic media; transparency and research); the second named "Calls to Action" which gives a general framework for different stakeholders such as news media, Artificial Intelligence (AI) actors, technology company, UN, States and political actors, fact-checking organization and networks etc.

In a policy brief written by the Secretary General (SG) António Guterres in preparation for these General Principles for Information Integrity, the SG sets out some important potential principles for the purpose of aiding Member States in the creation of these principles. However, in his policy brief, the SG had defined the difference between disinformation, misinformation, and hate speech,[224] which seems to be absent in the UN Global Principles for Information Integrity. In fact,

[219] UN, *United Nations Global Principles For Information Integrity: Recommendations for Multi-stakeholder Action*, United Nations, 2024, p. 9.
[220] *Ivi.*, p. 8.
[221] *Ivi.*, p. 3.
[222] *Ivi.*, p. 18.
[223] *Ivi.*, p. 5.
[224] GUTERRES, *Information Integrity on Digital Platforms*, Our Common Agenda, 2023, p. 5.

misinformation is even mentioned only once in the introduction paragraph of the UN Global Principles For Information Integrity, never to be written again. Not only it is quite a missed opportunity for clarity of comprehension and informational integrity – which this document is all about – but the absence of difference between the three important topics seems to be something they forgot to add since it was literally the first passage in the policy brief. Not to mention that the latter even states that «while there are no universally accepted definitions of these terms, United Nations entities have developed working definitions »,[225] hence there would be no reason to exclude them from the final product.

Even if many topics and ideas transferred from the SG policy brief to the final product such as the call for technology companies, policymakers, and many more actors to collectively develop safeguards against the malicious use of digital media, it is clear that this UN Global Principles for Information Integrity is more a set of guiding principles and recommendations for actors than anything else. It does set out 5 important principles that the document follows and respects, but it lacks a jurisdictional and binding aspect that would make this framework more impactful to UN members. Recommendations are very useful for the actors to understand the issue and follow a certain behavior, however, it is still the biggest issue with such UN projects: it all depends on voluntary commitments of these actors. The policy brief intelligently mentions the European Digital Service Act (DSA) which regulates the obligations of these digital services and is directly applicable across all the European Union (EU) member states.[226] In fact, the difference between the UN Global Principles for Information Integrity and the DSA is that the latter is an EU regulation and as such «they come into force and are legally binding without any action on the part of member states »,[227] following Article 288 of the Treaty on the Functioning of the European Union (TFEU), which is something that the UN cannot impose on its member states, which hence could be considered less effective than the EU.

[225] *Ibidem*.
[226] *Ivi.*, p. 16.
[227] Thomson Reuters, *Direct applicability (EU)*, Thomson Reuters Practical Law, p. 1.

As mentioned earlier, building on the foundation laid by the Global Principles, the forthcoming UN Code of Conduct for Information Integrity on Digital Platforms is expected to translate these broad normative aspirations into concrete, actionable guidelines for a wide range of stakeholders, including governments, technology companies, civil society, and media actors.[228] Its objective is to establish a coherent framework of voluntary commitments aimed at safeguarding the integrity of the information ecosystem while upholding international human rights standards, notably freedom of expression[229]. By operationalizing the principles of transparency, accountability, and public empowerment, the Code of Conduct aspires to foster greater resilience against the spread of MDM[230], thereby contributing to a more trustworthy and inclusive digital environment at the global level.

3. The role of the Security Council in the fight against MDM and Terrorism

The United Nations Security Council (UNSC) is one of the primary bodies responsible for maintaining international peace and security and is described by Chapter V of the UN Charter. The UNSC has played a vital role in addressing terrorism and its associated threats, including the spread of terrorism narratives. Acknowledging the evolving nature of terrorism in the digital age, the UNSC has adopted several resolutions and frameworks to tackle these issues, particularly through Resolution 1373 (2001) and Resolution 2354 (2017), as well as the Comprehensive International Framework to Counter Terrorist Narratives (S/2017/375).

Overall, the UN, thus including the SC, views terrorism as one of the most serious threats to international peace and security,[231] emphasizing that it cannot be justified under any circumstances[232] and that threats to international peace and security need to be fought by all

[228] GLOBSEC, *Policy Brief:* United *Nations Code of Conduct for Information Integrity on Digital Platforms*, GLOBSEC Centre for Democracy & Resilience, 2023, pp. 2–3.
[229] *Ibidem*.
[230] *Ivi.*, p. 4 – 6.
[231] UNSC, *Resolution 2354 on implementation of the Comprehensive International Framework to Counter Terrorist Narratives*, 2017, p.1.
[232] *Ibidem*.

means.[233] Resolution 1373 (2001), in particular, was primarily directed at traditional counterterrorism measures – such as countering terrorist financing,[234] denying safe havens to them,[235] and promoting international collaboration,[236] but also opened the door for future subsequent efforts to address more nuanced aspects of terrorism, namely ideological and technological aspect while also underscoring that states could «find ways of intensifying and accelerating the exchange of operational information, especially regarding actions or movements of terrorist persons or networks; forged or falsified travel documents; traffic in arms, explosives or sensitive materials; use of communications technologies by terrorist groups (…) »[237] as per clause 3(a) of said resolution.

However, it is by 2017 that the UNSC has recognized the critical role of narratives in terrorism. Resolution 2354 introduces a comprehensive framework to counter terrorist narratives (document number S/2017/375), not only to physically counter terrorism but this time giving an important focus to the communications. In particular, it highlighted how terrorist groups exploit online platforms and communication technologies to spread distorted narratives, recruit supporters, mobilize resources, and garner sympathy.[238] The UNSC stressed that countering these narratives requires a multi-faceted approach involving states, international organizations, religious leaders, civil society, educational institutions, and more.[239]

With that being said, the UNSC in this Resolution 2354, identifies terrorist narratives as a strategic tool used by groups like ISIL (Da'esh) and Al-Qaida to misrepresent religion and justify violence[240] but also «that terrorist craft distorted narratives that are based on the misinterpretation and misrepresentation of religion to justify violence, which are utilized to recruit supporters and Foreign Terrorist Fighters (FTFs), mobilize resources, and garner support from sympathizers, in particular by exploiting information and communications technologies, including through the

[233] UNSC, *Resolution 1373*, 2001, pp. 1–3.
[234] *Ivi.*,1(a), p. 2.
[235] *Ibidem.*, 2(c). p. 2.
[236] *Ivi.*, 3(b), 3(c), 3(e), p. 3.
[237] *Ibidem.*, 3(a), p. 3.
[238] UNSC, *Resolution 2354 on threats to international peace and security caused by terrorist acts*, cit., p.2.
[239] *Ibidem*.
[240] *Ibidem*.

Internet and social media ».[241] Both terrorist goals are thus achieved by narratives that are disseminated through social media and other digital platforms to radicalize individuals and incite violence, which is why the Council has expressed concern over how such MDM fosters extremism and undermines global security. Hence why the UNSC has contributed to the counter-terrorism manner by adopting the guidelines of Resolution 2354 which accepted the Comprehensive International Framework to Counter Terrorist Narratives (S/2017/375).[242] The latter, while it provides guidelines for states and organizations to counter terrorist propaganda, it also ensures that all these actors not only fight this phenomenon in a coordinated manner but guarantees that they follow principles of international law[243] as per Article 1 of the UN Charter but also of other Covenant and Treaties mentioned in clause 4 (ICCPR e.g.). So, not only do they need to rebut terrorist messages but also amplify credible alternatives that address vulnerabilities exploited by terrorist groups. While these resolutions seem to focus mostly on security, they don't undermine the respect of human rights aspects as seen by the following passage:

> «Recalling the right to freedom of expression, reflected in Article 19 of the Universal Declaration of Human Rights adopted by the General Assembly in 1948 ("the Universal Declaration"), and recalling also the right to freedom of expression in Article 19 of the International Covenant on Civil and Political Rights adopted by the General Assembly in 1966 ("ICCPR") and that any restrictions thereon shall only be such as are provided by law and are necessary on the grounds set out in paragraph 3 of Article 19 of the ICCPR.»[244]

Finally, the UNSC recognizes the threat of terrorism narratives and the MDM's use as strategic terrorist tools, however, it seems that there is a lack of an update on addressing countering terrorist

[241] *Ibidem*.
[242] *Ibidem*.
[243] UNSC, *Comprehensive International Framework to Counter Terrorist Narratives (S/2017/375)*, 2017, p. 5.
[244] UNSC, *Resolution 2354*, cit., p. 2.

narratives. In fact, RES 2354 remains the most comprehensive framework adopted by the UNSC in this regard with the endorsement of the Guidelines. Thus, while they did work on the issue, an update of these Guidelines would be necessary in order to keep it up to date to the newer technologies available in order to counter terrorist narratives more effectively.

4. Other United Nations Office and their Projects

The UNSC isn't the only organ that can have a role in fighting terrorism, in fact, the UN has an entire architecture only devoted to this problem: the UN counterterrorism architecture comprises distinct yet interconnected entities addressing disinformation-terrorism linkages through specialized mandates. Established under the UN Security Council Resolution 1371 (2001), the Counter-Terrorism Committee (CTC) has a function as the primary oversight body monitoring state compliance with counterterrorism obligations. Its subsidiary office, the Counter-Terrorism Committee Executive Directorate (CTED), operationalizes this mandate through technical assessments and capacity-building coordination, acting as a broker between states requiring assistance and those offering expertise.

However, the CTC and CTED must not be confused with a similar yet distinct office: the UN Office of Counter-Terrorism (UNOCT). It was created in 2017 as the system-wide coordinator under the General Assembly Resolution 71/291, which tasked it with implementing the UN Global Counter-Terrorism Strategy's four pillars. The UNOCT has some differences compared to the other two offices since, unlike CTED's monitoring role, UNOCT focuses on programmatic execution through its UN Counter-Terrorism Center (UNCTT). The latter delivers hands-on technical assistance funded by voluntary contributions. So, the UNOCT operates under the Secretariat with broader political advocacy and resource mobilization responsibilities compared to the CTED office which remains a subsidiary of the Security Council. Why it's important to not confuse these different offices is because not only they would be easy to get them mixed up due to their similar names, but also because this institutional differentiation creates a certain division of labor. The CTED/CTC provides both normative frameworks and compliance assessments, while

UNOCT/UNCTT implements practical counter-disinformation measures through partnerships like e.g. The UNOCT's Global Programme on Preventing Violent Extremism. Finally, the UN Global Counter-Terrorism Coordination Compact further harmonizes their efforts through inter-agency working groups.

4.1 CTED-ICT4Peace Joint Projects

The landmark collaboration between CTED and ICT4Peace Foundation marked a paradigm shift in addressing terrorist exploitation of digital technologies. Launched in 2015, this initiative created structured dialogue mechanisms between governments, UN agencies, and technology companies such as social media companies in order to «prevent the use of ICTs (Information and communication technology) for terrorist purposes, while respecting human rights and freedom of speech ».[245]

A second joint project was launched in April 2016, this time focusing on a better «understanding of current industry responses to terrorist use of their products and services, particularly with regard to content and-operational related issues and identify practices and experiences »,[246] thus on the engagement of the technology sector, more precisely in the response to terrorist use of ICTs. A first comprehensive report can be found in «Private Sector Engagement in Responding to the Use of the Internet and ICT for Terrorist Purposes: Strengthening Dialogue and Building Trust » which was published few months after the second joint project.

Lastly, another project that is worth bringing to attention is the «Tech Against Terrorism » initiative from 2017, which is still joint between UN CTED and ICT4Peace, and notably its work has been recognized by two UNSC resolutions.[247] This initiative operates at the intersection of digital governance and counterterrorism, providing a structured response to the exploitation of online

[245] ICT4Peace, *More support needed for smaller technology platforms to counter terrorist content*, ICT4Peace Foundation, 2018, p. 1.

[246] ICT4Peace, *UN and ICT4Peace engage with private sector on responding to terrorist use of ICT*, ICT4Peace Foundation, 2016, p. 1.

[247] ICT4Peace, *UN Security Council recognises ICT4Peace's work with the United Nations*, ICT4Peace Foundation, 2018, p. 1.

platforms by extremist actors. Additionally, this initiative partners with technology companies, to strengthen their ability to detect, disrupt, and remove terrorist content while also preserving fundamental freedoms.[248] Recognizing that terrorist organizations are increasingly reliant on algorithmic amplification, encrypted communication, and decentralized networks,[249] Tech Against Terrorism advances the development of transparent content moderation policies, strong threat assessments, and improved information sharing among digital service providers.

The initiative encourages the timely identification and removal of verified extremist content which is done thanks to the Terrorist Content Analytics Platform (TCAP). Besides that, it also provides recommendations for policymakers with position papers analyzing and describing issues with the proper recommendations, e.g. on the spread of terroristic and extremist content online.[250] The Tech Against Terrorism initiative decided to take an approach without reinforcing overly securitized digital governance frameworks that may undermine freedom of expression by promoting multi-stakeholder collaboration. However, its success remains contingent on various factors, such as the cooperation of tech companies, the adaptability of moderation technologies, and the ongoing challenge of distinguishing between legitimate political discourse and content linked to extremist ideologies, particularly in politically contested environments.

4.2 UN Global Counter-Terrorism Strategy

The United Nations Global Counter-Terrorism Strategy could be considered the UN most prominent and important project on the topic of counterterrorism[251] as it is constantly mentioned. It was adopted by the General Assembly with Resolution 60/288 during the month of September of 2006.[252] Overall, it is a comprehensive framework for addressing terrorism at the international

[248] HATTOTUWA et al., *High-Level Panel on Digital Cooperation: Reflections and Recommendations from the ICT4Peace Foundation*, ICT4Peace Publishing, 2018, p. 14.
[249] Tech against Terrorism, *Position Paper: Content Personalisation and the online dissemination of terrorist and violent extremist content*, Tech against Terrorism, 2021, p. 1.
[250] *Ibidem*.
[251] BORISOVICH-ORISHEV et al., *The UN role in combating international terrorism: achievements and challenges*, Revista de investigaciones Universidad del Quindio, 2022, p. 269.
[252] UNGA, *The United Nations Global Counter-Terrorism Strategy (Resolution 60/288)*, 2006, p. 1.

level, but in particular, it has been designated to integrate for multiple tasks: preventive measures,[253] capacity-building efforts,[254] and human rights protections.[255] The strategy reflects a multilateral commitment to counterterrorism that transcends purely security-based approaches. This Strategy acknowledges that that terrorism thrives in conditions of political instability, economic deprivation, and social marginalization, and by doing that it situates counterterrorism within a broader agenda that seeks to address root causes, enhance state capacities, and uphold the rule of law. As such, it advances a complete approach that brings together preventive, operational, institutional, and normative dimensions of counterterrorism.

The Strategy is structured around four foundational pillars, which are the following:

> Pillar I: «Measures to address conditions conducive to the spread of terrorism »,[256] which includes also tackling root causes such as poverty, political exclusion, and lack of education.
>
> Pillar II: «Measures to prevent and combat terrorism »,[257] focusing on law enforcement, border security, and countering terrorist financing.
>
> Pillar III: «Measures to build States' capacity to prevent and combat terrorism and to strengthen the role of the United Nations system in that regard ».[258]
>
> Pillar IV: «Measures to ensure respect for human rights for all and the rule of law as the fundamental basis of the fight against
>
> terrorism ».[259]

[253] *Ivi.*, Annex II, pp. 5 – 7.
[254] *Ivi.*, Annex III, pp. 7 – 8.
[255] *Ivi.*, Annex IV, p. 9.
[256] UNGA, *Report of the Secretary-General: Activities of the United Nations system in implementing the United Nations Global Counter-Terrorism Strategy*, 2023, p. 3.
[257] *Ivi.*, p. 4.
[258] *Ivi.*, p. 6.
[259] *Ivi.*, p. 7.

Alongside these pillars, the Strategy doesn't forget, and in fact, emphasizes, state institutions to enhance their resilience against terrorist threats[260] and ensure human rights-based counterterrorism policies[261] that do not erode civil liberties in the pursuit of security.

According to the SG report «Activities of the United Nations system in implementing the United Nations Global Counter-Terrorism Strategy » (A/77/718), these UN offices have worked together in the context of the UN Global Counter-Terrorism Strategy, contributing to various activities.

Notably, the United Nations Office on Drugs and Crime (UNODC), in South Asia and South-East Asia, promoted the good practices in preventing violent extremism and assisted in the creation of regional networks of practitioners.[262] Furthermore, it has supported countering child recruitment by terrorists by supporting not only their rehabilitation, but also their reintegration in Indonesia, Iraq, and Nigeria.[263] Besides the work of UNODC, twenty nations have received aid from the United Nations Development Programme (UNDP) to create national action plans aimed at combating violent extremism. Within the Association of Southeast Asian Nations (ASEAN), UNOCT's activities are like the latter as the office aided in monitoring national and regional plans of action, as well as evaluating them.[264] However, the UNDP goes a step further, as the office promoted in forty country alternative narratives in order to counter extremist and terroristic content, all of which in the name of tolerance and openness, both on the digital sphere and offline.[265] But since violent extremism doesn't only come from narratives, UNDP (and the International Organization for Migration (IOM) with Bosnia and Herzegovina and Morocco) has thought to aid individuals by including mental health and psychosocial support in fifteen countries.[266]

[260] UNGA, *The United Nations Global Counter-Terrorism Strategy: eighth review (Resolution 77/298)*, 2023, p. 2.
[261] *Ibidem*.
[262] *Ivi.*, p. 4.
[263] *Ibidem*.
[264] *Ibidem*.
[265] *Ibidem*.
[266] *Ibidem*.

To assess the effectiveness of the Strategy, the GA conducts a biennial review.[267] This is necessary due to the shifting nature of terrorist threats which adapts as the technologies evolve. As mentioned earlier, UNOCT is the UN structure that plays a central role in coordinating its implementation, working in close collaboration with UNCCT and CTED. These entities oversee the operational dimensions of counterterrorism, monitoring state compliance with Security Council resolutions and guiding member states in strengthening their legal, institutional, and operational frameworks. With such a complete and structured oversight of offices, the UN strives to ensure that global counterterrorism efforts remain dynamic, adaptable, and responsive to evolving security landscapes, particularly in light of technological advancements that have expanded the scope of terrorist threats beyond traditional battlefields.

In fact, after these biennial reviews, there have been some reviews of the UN Global Counter-Terrorism Strategy. The latest one is the eighth review which was conducted in June 2023 during the General Assembly's 77th session due to the growing complexity of contemporary terrorist threats, particularly in the digital sphere.[268] The growing focus of terrorist entities on new technologies, such as encrypted communications, virtual currencies, and AI-enhanced propaganda, required a reevaluation of counterterrorism approaches. Besides the new technologies, this review renewed the emphasis on the protection of human rights and gender-sensitive considerations in each pillar of the strategy, inspired by the uneven effects of terrorism and counterterrorism on marginalized groups.

The 8th review also advanced processes for the development for improved monitoring and evaluation[269] in order to assist in guaranteeing that counterterrorism measures are applied fairly to all states, especially what developing nations bear compared to others in equitable enforcement. Due to the exploitation of digital platforms by terrorists in ungoverned spaces, the review also

[267] UN Office of Counter-Terrorism, *United Nations Global Counter-Terrorism Strategy*, United Nations Global Counter-Terrorism Website, 2023, p. 1.
[268] *Ibidem*.
[269] UNGA, *The United Nations Global Counter-Terrorism Strategy: eighth review (Resolution 77/298)*, cit., p. 4.

called for action among collaborating states[270] to limit the spread of extremist content, but always without guaranteeing freedom of expression and the right to privacy.

A missing part in this 8th review, compared to the SG report, is the emphasis on the geopolitical context of terrorism which pointed to a surge and intensification of terrorist activity in Africa and other fragile regions. This may be because since the digital sphere is accessible non-depending on the location, it doesn't need to state specifically a geopolitical context.

Finally, increased capacity-building support in these contexts was identified as a critical priority, particularly for border security, law enforcement, and behavioral analysis,[271] etc. The review also noted some advancements in counterterrorism capacity and program work; especially work that utilizes behavioral insights[272] to prevent terrorism and counter-radicalization capacity through technological ways of facilitating information-sharing among national security assistance and agencies, among others.

There is an upcoming review of the strategy in 2026 which will provide an opportunity to further refine the UN's global counterterrorism approach, particularly as artificial intelligence, cyber warfare, and climate-related security risks reshape the threat landscape. At the time of writing, the content of the forthcoming review remains speculative. However, it is reasonable to anticipate an emphasis on enhancing international cooperation and aligning counterterrorism strategies with evolving global realities. Such a focus would reflect the growing need for a more adaptive and preemptive approach to the challenges currently faced. The increasing weaponization of AI-derived MDM, autonomous threats by terrorists, and radicalization pathways due to climate change will likely require new regulations and enhanced cooperation among both state and non-state actors. This review will be an occasion for perhaps a redefinition of the UN's strategy or a confirmation of the path it is currently following. Future research will be able to analyze it and comment on its new points to add to the vast literature surrounding the topic.

[270] *Ivi.*, Pillar I (25), p. 12.
[271] *Ivi.*, p. 7.
[272] *Ivi.*, p. 3.

5. UN Educational, Scientific and Cultural Organization

A quick comment on the UN Educational, Scientific and Cultural Organization (UNESCO) is necessary as it’s worth mentioning for its essential role in addressing the crossroad between terrorism and the proliferation of MDM. UNESCO is known for its central role and expertise in media literacy and digital resilience after all. As such, the organization aims to strengthen communities against the manipulative tactics of extremist actors who use online spaces to radicalize individuals and disrupt social cohesion through targeted initiatives. In fact, at the heart of the strategy employed by UNESCO is the Media and Information Literacy (MIL) project, which promotes critical thinking skills, media literacy, informational resilience, and ethical consumption of information.[273] MIL wants to empower individuals to determine which sources are credible or misleading, acting as protective counter-narratives to extremist recruitment mechanisms reliant on disinformation to create ideological divides. The MIL project also produced global frameworks, pedagogical toolkits, and integrated curricula to support educational locales around the world in embedding media literacy into both formal and informal educational working environments. As such, by strengthening digital resilience at the level of individuals themselves, UNESCO enhances societal capacity to resist the spread of terrorist propaganda while safeguarding fundamental freedoms.

Another significantly important initiative is “Social Media 4 Peace” – an initiative aimed directly at addressing the ways hate speech, extremist discourse, and polarization fueled by disinformation operate online. Implemented in Bosnia and Herzegovina, Kenya, Indonesia, and Colombia,[274] the project recognizes that countering radicalization requires contextualized approaches that engage with the sociopolitical realities of each region. As such, the project facilitates multi-stakeholder engagement with social media platforms, civil society organizations, and legislators to help improve the moderation of content while, of course, preserving freedom of expression. The development of counter-narratives that promote tolerance and social cohesion represents the main

[273] FRAU-MEIGS, *Media and information literacy*, UNESCODOC Digital Library, 2023, p. 4.

[274] UNESCO, *Evaluation of the project “Social Media 4 Peace”*, United Nations Global Marketplace, 2024, p. 1.

effort of this initiative which, unlike reactive censorship models, it prioritizes constructive engagement with online communities, empowering local actors to shape inclusive digital spaces rather than merely suppressing harmful content.

For example, a significant partnership with UNESCO in the context of this initiative, Article 19 led in-depth research focusing on specific geopolitical contexts (Bosnia and Herzegovina, Indonesia, and Kenya) for the creation of a handbook named "Social Media 4 Peace Handbook" which they describe as «a guide to content moderation and freedom of expression ».[275] The handbook not only defines relevant international human rights standards and definitions of MDM and hate speech but also describes and critically analyzes the content moderation in practices, pointing out the strengths and weaknesses of the moderation and framework.

In conclusion, the approach adopted by UNESCO seems to align itself with the UN and its various organs' counterterrorism agenda. However, rather than prioritizing multi-stakeholder partnerships or international cooperation, UNESCO seeks to address the root of the problem by acting at the individual level. In doing so, the organization complements traditional security-oriented measures, ensuring that counterterrorism strategies are not exclusively dependent on surveillance and content suppression. Instead, by fostering informed digital citizenship and promoting media pluralism, UNESCO enhances democratic resilience against the manipulative tactics of extremist groups.

6. Conclusion

The UN's contribution to countering MDM in the context of terrorism is characterized by a comprehensive, multilayered approach grounded in international law, human rights standards, and broad multilateral cooperation. Through foundational legal instruments such as Articles 19 and 20 of the ICCPR, and key initiatives including the UN Strategy and Plan of Action on Hate Speech, the Verified Initiative, Resolution 75/309, and the Global Principles for Information Integrity, the

[275] Article 19, *Social Media 4 Peace: A handbook to support freedom of expression*, Article 19 Website, 2023, p. 1.

UN has progressively acknowledged the intricate links between manipulated information and violent extremism. Additionally, the Security Council's resolutions and frameworks, notably Resolution 2354 and the Comprehensive International Framework to Counter Terrorist Narratives, further highlight the organization's capacity to integrate both security imperatives and the protection of fundamental freedoms. Complementary efforts by specialized agencies and partnerships, such as UNESCO's media literacy initiatives and the CTED–ICT4Peace joint projects, underscore the UN's ability to foster multi-stakeholder collaboration and promote resilience at institutional and societal levels.

The principal strength of the UN's approach lies in its normative authority, convening power, and emphasis on balancing counterterrorism measures with the safeguarding of human rights. However, persistent limitations undermine the full effectiveness of these efforts. The absence of universally accepted definitions of MDM and hate speech creates legal ambiguity, increasing the risk of inconsistent implementation and potential misuse by states to suppress dissent. Moreover, the predominantly non-binding nature of UN initiatives, reliant on voluntary commitments and lacking robust enforcement mechanisms, constrains their operational impact, particularly when compared to binding regional instruments such as the European Union's Digital Services Act. Finally, the challenges of engaging private sector actors, whose commercial incentives often diverge from regulatory objectives, further complicate the implementation of counter-MDM strategies. Aside from these shortcomings, the UN's evolving frameworks and sustained normative leadership position it as a key actor in promoting an international response that strives to reconcile security, digital resilience, and human rights in an increasingly complex information landscape.

CHAPTER III

The European Union and its efforts against MDM and Terrorism

Summary: 1. Introduction. – 2. Does the European Union have the necessary competences? – 2.1 The EU's Legal framework. – 2.1.1 The EU foundational treaties. – 2.1.2 The Council of Europe's European Convention on Human Rights. – 2.1.3 Charter of Fundamental Rights of the European Union. – 2.2 The EU Commission's Communication on Preventing Radicalization Leading to Violent Extremism. – 2.3 European External Action Service and its contribution. – 2.4 The EU Strategic Compass against hybrid threats. – 3. EU Strategies and Agendas. – 3.1 The European Union Counter-Terrorism Strategy (2005). – 3.2 Security Union Strategy and Counter Terrorism Agenda (2020). – 4. EU Internet Forum (2015). – 5. Digital Services Act (2022). – 5.1 Code of Practice on Disinformation (2022). – 5.2 The EU Code of Conduct on Disinformation. – 6. Conclusion.

1. Introduction

It's important to first define what the European Union is: the EU is a political and economic union which is currently composed of 27 member states,[276] post-Brexit. Since it is quite a unique case in the international scene, it's always been quite difficult to clearly define what the EU truly is. Depending on the various schools of thought, the EU could be considered a supranational organization but also an international organization of a regional type. Nevertheless, what is clear, though, is that the EU is a very important subject of international law, hence why it is crucial to discuss the EU's contribution against terrorism and MDM.

For a brief history parenthesis, it was established after World War II to foster economic cooperation with the belief that countries, if they traded together, they would be less likely to engage in conflict.[277] The latter was the reason for all these countries to create a Union of European countries for Europeans wanted to avoid repeating the horrors of the conflict they had to face during both

[276] EU, *EU countries*, European Union Official Website, p. 1.

[277] EU, *The Schuman Declaration*, European Union Official Website, 1950, p. 1

the World Wars. Over time, what started out as simply this economic union, it evolved into a multifaceted organization with growing in size with all the new members as well as growing in responsibilities with new competencies attributed to it. Soon the EU started to manage trade, environmental policies, security, and various other fields. Its governance structure includes key institutions such as the European Commission, the European Parliament, and the European Council, all of which work collaboratively to enact legislation and coordinate policies across the member states.

This chapter offers a comprehensive examination of how the EU addresses the intersection between online MDM and terrorism, by combining an analysis of its legal foundations with an assessment of key strategies and initiatives. The discussion opens with an exploration of the Union's competences, focusing on the legal frameworks that ground its capacity to act in the fields of counterterrorism, radicalisation prevention, and resilience against hybrid threats. This includes an appraisal of relevant policy communications from the EU Commission as well as the role of external action instruments. Building on this groundwork, the chapter then analyses the principal strategies that have shaped the EU's counterterrorism agenda over time, paying particular attention to how these frameworks have evolved to incorporate challenges stemming from the online information environment. Then there's a focus on more specific and concrete initiatives and regulatory measures, notably the EU Internet Forum and the Digital Services Act, complemented by the highly relevant instruments such as the Code of Practice on Disinformation and the EU Code of Conduct.

2. Does the European Union have the necessary competences?

However, it is important to examine the relevance of MDM and counter-terrorism to the EU, as well as the reasons why the EU can be considered a significant actor in these domains. An important reason is due to their impact on societal stability, democratic integrity but also security: MDM campaigns can undermine trust in institutions, polarize societies and amplify extremist ideologies that could fuel terrorism but also nationalistic sentiments which would crumble the trust in the EU institutions. Given the EU's unique transnational nature, which enables cross-border coordination and the mobilization of collective resources, failure to act effectively in this domain would constitute a significant limitation in addressing these challenges. Furthermore, the EU's emphasis given on hybrid threat mitigation – outlined in frameworks like the Strategic Compass, which will be explained later – underscores its commitment to countering foreign information manipulation and interference. Hence, in these domains, the EU's role is indispensable as it provides a unified platform for intelligence sharing, strategic communication, and partnerships with international organizations to combat global security threats effectively.

2.1 The EU's Legal framework

2.1.1 The EU foundational treaties

As previously noted, the EU has expanded its competencies beyond its original economic focus. It is therefore essential to first assess whether the EU possesses the necessary competences to address these matters. The core of its authority lies in Article 83 of the Treaty on the Functioning of the European Union (hereby as TFEU), which used to be the Article 31 of the Treaty on European Union (TEU):

> *«The European Parliament and the Council may, by means of directives adopted in accordance with the ordinary legislative procedure, establish minimum rules concerning the definition of criminal offences and sanctions in the areas of particularly serious crime with a cross-border dimension resulting from the nature or impact of such offences or from a special need to combat them on a common basis. These areas of crime are the following: terrorism, […] ».*[278]

[278] EU, *Consolidated version of the Treaty on the Functioning of the European Union*, Official Journal of the European Union, Art. 83, 2009, pp. 34 – 35.

This article thus enables the EU to establish minimum rules for terrorism-related crimes. This provision serves as the foundation for measures such as the 2017 Counter-Terrorism Directive, which obliges Member States to ensure the swift removal of terrorist content hosted within their jurisdiction.[279] However, there may be cases where jurisdictional ambiguities persist like, for example, when content is hosted outside a member's state territory which can still be accessed through social media, thereby exposing regulatory gaps. For this reason, Article 23 stipulates that, if removal of the content is not feasible, an alternative measure is to block access to it.[280]

As such, the authority of the EU in counterterrorism remains circumscribed by the foundational treaty provision that reaffirms the primacy of the member states in matters of national security. Article 4(2) TEU explicitly designates national security as the exclusive competence of member states, while Article 72 TFEU further reinforces their sovereign prerogatives in upholding law and order, as well as safeguarding national security. Hence, this legal framework delineates the role of the EU as complementary rather than overriding, necessitating a careful balance between supranational coordination and national autonomy. In fact, this is in line with the principle of subsidiarity which the EU is based on according to Article 5(1) TEU.[281]

It is important to note that the principle of subsidiarity does not relate to the division of powers between the EU and the Member States, but rather to the effective exercise of those powers. A distinction must also be made between exclusive and shared competences within the EU framework.

Exclusive competences, as defined in Article 2(1) TFEU, apply when «when the Treaties confer on the Union exclusive competence in a specific area, only the Union may legislate and adopt

[279] European Parliament and Council of the European Union, *Directive (EU) 2017/541 of the European Parliament and of the Council of 15 March 2017 [on combating terrorism and replacing Council Framework Decision 2002/475/JHA and amending Council Decision 2005/671/JHA]*, Official Journal of the European Union, 2017, p. 4.

[280] *Ibidem*.

[281] EU, *Consolidated version of the Treaty on the European Union*, Official Journal of the European Union, Art. 5(1), 2009, p. 6.

legally binding acts, the Member States being able to do so themselves only if so empowered by the Union or for the implementation of Union acts ».[282] Article 3(1) TFEU provides the list of these exclusive competences. Shared competences are defined in Article 2(2) TFEU. In this case, «when the Treaties confer on the Union a competence shared with the Member States in a specific area, the Union and the Member States may legislate and adopt legally binding acts in that area. The Member States shall exercise their competence to the extent that the Union has not exercised its competence. The Member States shall again exercise their competence to the extent that the Union has decided to cease exercising its competence ».[283] Article 4(3) TEU provides the list of the shared competencies.

Hence, the principle of subsidiarity is not valid in matters of exclusive competence of the Union – since only the Union (but not the member states) can intervene – but in matters of shared competence, as well as those concerning supporting, coordinating, or supplementing actions. As such, the principle of subsidiarity is essentially aimed at safeguarding the scope of operations by member states, thus placing a curb on potentially excessive intrusiveness from the EU. In fact, Article 5(3) TEU seems to be written in quite a restrictive manner, with respect to the Union's intervention. It establishes two conditions that must be met for its application to be justified. The first condition is the inadequacy of action at the national level to achieve the objectives pursued.[284] The second condition concerns the 'added value' that EU intervention would provide, taking into account the scale or effects of the action in question.[285]

Furthermore, the solidarity clause under Article 222 TFEU solidifies the framework by adding a collective response and thus requiring the EU to deploy all available measures, «including the military resources made available by the Member States »,[286] to aid member states against terroristic threats or help prevent it, for a better-coordinated intervention. Besides, the EU's

[282] EU, *Consolidated version of the Treaty on the European Union*, Art. 2(1), cit., p. 5.
[283] *Ibidem*.
[284] EU, *Consolidated version of the Treaty on the European Union*, Art. 5(3), cit., p. 6.
[285] *Ibidem*.
[286] EU, *Consolidated version of the Treaty on the Functioning of the European Union*, Art. 222, cit., p. 102.

capacity to coordinate effectively, enabled by the solidarity and cooperation clauses and articles, is an important element in the fight against the terroristic threat.

2.1.2 The Council of Europe’s European Convention on Human Rights

Another important legal basis for the EU is the European Convention on Human Rights (ECHR), or also officially known as European Convention for the Protection of Human Rights and Fundamental Freedoms. While it is highly important to remember that the ECHR wasn’t stipulated by the EU but by the Council of Europe,[287] it remains a vital instrument to the EU. In fact with the Lisbon Treaty, which modified the TEU, the EU is legally bound to adhere and respect the content of the ECHR.[288] According to Article 6 (2) of the TEU, «the Union shall accede to the European Convention for the Protection of Human Rights and Fundamental Freedoms. Such accession shall not affect the Union's competences as defined in the Treaties ».[289] Besides, Article 6 (3) confirms this, stating that « fundamental rights, as guaranteed by the European Convention for the Protection of Human Rights and Fundamental Freedoms and as they result from the constitutional traditions common to the Member States, shall constitute general principles of the Union's law ».[290]

Thus, this thesis will also take into consideration the ECHR and particularly Article 10, which introduces a necessary equilibrium between counterterrorism imperatives and fundamental freedoms.[291] This provision stresses the importance of the freedom of expression in clause 1, stating that «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 ».[292] But while it enshrines freedom of expression, the

[287] Council of Europe, *A Convention to protect your rights and liberties*, Council of Europe Official Website.
[288] European e-Justice, *Protecting fundamental rights within the European Union*, European e-Justice Portal, p. 2.
[289] EU, *Consolidated version of the Treaty on the European Union*, Art. 6(2), cit., p. 7.
[290] EU, *Consolidated version of the Treaty on the European Union*, Art. 6(3), cit., p. 7.
[291] Council of Europe, *European Convention on Human Rights (as amended by Protocols No. 15)*, Art.10, 2021, p. 12.
[292] *Ibidem*.

jurisprudence of the ECHR acknowledges that speech inciting violence or threatening national security may be subject to lawful restrictions in Article 10, clause 2: «the exercise of these freedoms, since it carries with it duties and responsibilities, may be subject to such formalities, conditions, restrictions or penalties as are prescribed by law and are necessary in a democratic society, in the interests of national security, territorial integrity or public safety, for the prevention of disorder or crime, […] ».[293] Similarly to this, Directive 2017/541 grants the EU the power to mandate content removal, albeit subject to proportionality constraints in order to mitigate the risk of overreach or abuse of powers.[294]

Nonetheless, while the exclusion of certain expressions from the protection of Article 10 through the application of Article 17, also called "abuse clause", may appear as a prudent approach to curbing harmful content, this path presents notable challenges that cannot be overlooked. The European Court of Human Rights (ECtHR) has progressively broadened the scope of Article 17, initially intended to counter explicit calls for totalitarian regimes, to also deal with cases of Holocaust denial, promotion of dictatorship, and the dissemination of anti-Semitic and Islamophobic content.[295]

However, the Court's jurisprudence has been marked by inconsistencies, as demonstrated in *Perinçek v. Switzerland*, where denial of the Armenian genocide was treated differently from Holocaust denial, and Article 17 was not applied.[296] Importantly, Shattock points out that when Article 17 is invoked, the Court circumvents its established "three-pronged test" under Article 10 (2): assessing legality, legitimacy of aim, and necessity in a democratic society.[297] He considers that extending Article 17 to disinformation risks exacerbating this ambiguity, particularly given the conceptual vagueness that characterizes disinformation itself. The lack of clear and consistent definitions surrounding terms such as "fake news" further complicates the legal landscape,

[293] *Ibidem*.

[294] European Parliament and Council of the European Union, *Directive (EU) 2017/541 of the European Parliament and of the Council of 15 March 2017 on combating terrorism and replacing Council Framework Decision 2002/475/JHA and amending Council Decision 2005/671/JHA*, cit., p. 4.

[295] See *Roj TV A/S v. Denmark*, 2018; *Pavel Ivanov v. Russia*, 2007.

[296] SHATTOCK, *Should the ECtHR Invoke Article 17 for Disinformation Cases?*, EJIL:Talk! (Blog of the European Journal of International Law), 2021, pp. 4 – 5.

[297] *Ibidem*.

particularly given the ECtHR's own inadvertent use of the phrase in *Brzezinski v. Poland* (2019).[298] Many have criticized the use of the term "fake news" by the Court, as for example Ó Fathaigh, who observes that both Council of Europe[299] and European Commission[300] reports have strongly cautioned against its use.[301]

Finally, Shattock concludes stating that, should national authorities impose sanctions for disseminating disinformation, the ECtHR must adjudicate such cases under the structured proportionality assessment of Article 10 (2), since «failure to do so, and to instead relegate disinformation to Article 17 cases, could create unnecessary uncertainty in this ever-challenging area ».[302]

2.1.3 Charter of Fundamental Rights of the European Union

Although the Charter of Fundamental Rights of the European Union (CFREU) does not explicitly refer to terrorism within its provisions, it nevertheless articulates fundamental principles relevant to the balance between security and individual freedoms. Notably, Article 11 (1) enshrines the right to freedom of expression and information, affirming that: «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 ».[303] Furthermore, Article 52 (1) of the Charter delineates the framework under which the exercise of these rights and freedoms may be limited, stipulating that: «any limitation on the exercise of the rights and

[298] *Ibidem*.

[299] WARDLE, *Information Disorder: Toward an interdisciplinary framework for research and policy making*, cit., p. 5.

[300] European Commission: Directorate-General for Communications Networks, Content and Technology, *A multi-dimensional approach to disinformation: report of the independent High level Group on fake news and online disinformation*, Publications Office of the European Union, 2018, p. 3.

[301] Ó FATHAIGH, *Brzeziński v. Poland: Fine over 'false' information during election campaign violated Article 10*, Strasbourg Observers, 2019, pp. 2 – 3.

[302] SHATTOCK, *Should the ECtHR Invoke Article 17 for Disinformation Cases?*, cit., p. 4.

[303] EU, *Charter of Fundamental Rights of the European Union*, Official Journal of the European Union, 2009, Article 11(1), p. 11.

freedoms recognised by this Charter must be provided for by law and respect the essence of those rights and freedoms. Subject to the principle of proportionality, limitations may be made only if they are necessary and genuinely meet objectives of general interest recognised by the Union or the need to protect the rights and freedoms of others ».[304] Hence, limitations are permissible solely under strict legal conditions and if they respect the criteria of necessity and proportionality.

Besides, Article 52 (3) explicitly establishes a direct relationship between the Charter and the ECHR, stating that «in so far as this Charter contains rights which correspond to rights guaranteed by the Convention for the Protection of Human Rights and Fundamental Freedoms, the meaning and scope of those rights shall be the same as those laid down by the said Convention. This provision shall not prevent Union law providing more extensive protection ».[305] Thus, this clause not only reaffirms the foundational principles established in the ECHR, but the last sentence is crucial as it further clarifies that it ensures a minimum standard of protection harmonized with the ECHR, while allowing the EU to grant higher safeguards where appropriate.

Hence, while the Charter omits direct references to terrorism or MDM, it nonetheless creates the basis for a legal architecture that is pertinent when addressing measures enacted in the context of MDM and terrorism. Unfortunately, there is no specific case law or jurisdictional opinion from the Court of Justice of the European Union (CJEU) on the matter of disinformation. In fact, a study commissioned by the European Parliament mentions that «the Court of Justice of the European Union (CJEU)'s media jurisprudence is underdeveloped and relies on ECHR case law ».[306]

As a last comment for this section, it's important to remember that the EU measures against terrorism must operate within the framework of fundamental freedoms, much like seen previously with the UN and its Charter. In fact, the ECHR binds the member states and the ECtHR sets additional legal parameters for counterterrorism efforts. As such, compliance with these fundamental freedoms balances out the need for security and the protection of civil liberties in the EU.

[304] *Ivi.*, Article 52 (1), p. 21.

[305] *Ibidem*., Article 52 (3).

[306] BAYER et al., *The fight against disinformation and the right to freedom of expression*, Think Tank European Parliament, 2021, pp. 19–20.

2.2 The EU Commission's Communication on Preventing Radicalization Leading to Violent Extremism

Much like Article 222 TFEU, the EU approach seems to be overall not only reactive but also preventive which can also be noticed in the 2016 Communication on Preventing Radicalization Leading to Violent Extremism: in this communication, in fact, the EU has a facilitative role in addressing radicalization even if the domain primarily belongs within the competence of member states.[307] This aligns with the principle of proximity which is recognized in Article 1 TEU: given that local actors (member states) maintain greater proximity to at-risk communities (citizens), they are best positioned to implement early detection and prevention measures.[308] However, the EU still has the important role of strengthening and facilitating cross-border cooperation and disseminating best practices, which is necessary considering the transnational character of online radicalization and terrorism narratives.

In itself, the Communication on Preventing Radicalization Leading to Violent Extremism identifies seven strategic priorities through which the initiatives from the EU complement the national efforts: (1) advancing research and evidence-based policymaking,[309] (2) countering terrorist propaganda and hate speech online,[310] (3) mitigating radicalization within the prison system,[311] (4) promoting inclusive education and European values,[312] (5) strengthening societal resilience,[313] (6) enhancing security measures,[314] and (7) addressing the international dimensions

[307] European Commission, *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions Supporting the Prevention of Radicalisation Leading to Violent Extremism* (COM/2016), 2016, p. 2.
[308] EU, *Consolidated version of the Treaty on the European Union*, Article 1, cit., p. 4.
[309] European Commission, *Communication Supporting the Prevention of Radicalisation Leading to Violent Extremism* (COM/2016), cit., p. 3.
[310] *Ivi.*, p. 5.
[311] *Ivi.*, p. 8
[312] *Ivi.*, p. 9.
[313] *Ivi.*, p. 11.
[314] *Ivi.*, p. 13.

of radicalization.[315] Out of these seven efforts, the MDM would be the second priority, proving that the EU recognizes its importance since it gave the phenomenon its own pillar. An important member of this approach is the Radicalization Awareness Network (RAN) Centre of Excellence which has the role of a knowledge-exchange platform in which educators, officials, and others are reunited to exchange information.[316] Through this multi-agency Centre of Excellence, the RAN provides practical guidance (e.g. standardized handbooks and intervention toolkits) to support the development of effective prevention strategies.

2.3 European External Action Service and its contribution

An important role on the fight against MDM and counterterrorism is done by the European External Action Service (EEAS), which is the EU's diplomatic service and is in charge of implementing the Union's foreign and security policy, which in this case includes also efforts to counter MDM. With their own initiatives, the EEAS monitors and responds to foreign information manipulation which threatens the EU security and democracy.[317] One of its best tools against that threat is the Strategic Communication (StratCom) Task Forces: they are the ones who continuously analyze and counter disinformation campaigns. The East StratCom Task Force, for instance, focuses specifically on Russian disinformation efforts,[318] while other units target information threats originating from regions such as the Western Balkans and the Middle East. All these combined efforts from the different StratCom Forces are part of what is called "Mitigating hybrid threats", which consists of tactics that blend conventional and unconventional means to undermine stability, including the deliberate spread of disinformation by state and non-state actors. Notably, the EU Commission defines hybrid threats as «when, state or non-state, actors seek to exploit the vulnerabilities of the EU to their own advantage by using in a coordinated way a mixture of measures (i.e. diplomatic, military, economic, technological) while remaining below the threshold

[315] *Ivi.*, p. 14.
[316] *Ivi.*, p. 5.
[317] Strategic Communications, *Information Integrity and Countering Foreign Information Manipulation & Interference (FIMI)*, European Union External Action website, 2025.
[318] EUvsDisinfo, *'To Challenge Russia's Ongoing Disinformation Campaigns': Eight Years of EUvsDisinfo*, News and Analysis, EUvsDisinfo, 2023, p. 1.

of formal warfare »[319] while also giving as an example for hybrid threats the «hindering of democratic decision-making processes by massive disinformation campaigns, using social media to control the political narrative or to radicalize, recruit and direct proxy actors ».[320] In this sense, the EEAS also works close to the member states to make sure that they have the necessary intel and can answer these threats coordinately to other member states.

But beyond its internal actions, as said earlier, the EEAS is the EU's diplomatic service and as such it has external initiatives. The EEAS, in fact, collaborates with international organizations such as the UN and the North Atlantic Treaty Organization (NATO) to give support and promote collaboration to counter terrorism-related MDM.[321] The logic is that if there are multiple partnerships between subjects of international law such as international organizations, then the collective resilience against propaganda is strengthened. Thus, the EEAS's engagement with these institutions proves the EU's commitment to align its information security efforts within the more general international counterterrorism framework.

2.4 The EU Strategic Compass against hybrid threats

Finally, the adoption of the EU Strategic Compass in March 2022 marked a significant step and is worth mentioning as part of the general provision of the Union. It focuses particularly on the hybrid threats, which also include those linked to terrorism, and to strengthen the overall security architecture of the region. Foreign information manipulation is directly defined by the Strategic Compass as a pressing security challenge,[322] which in turn needs to be given more attention and to reinforce both the intelligence capacities and situational awareness to counteract these harmful narratives. The document actually talks about disinformation being a vector of hybrid warfare,[323] and as such is highlighted as a priority area for intervention: that's why the Strategic Compass

[319] European Commission's Directorate-General for Defence Industry and Space, *Hybrid Threats*, Directorate-General for Defence Industry and Space, p. 1.
[320] *Ibidem*.
[321] *Ibidem*.
[322] Council of the European Union, *Strategic Compass for Security and Defence*, 2022, p. 22.
[323] *Ivi.*, p. 5.

focuses on improving the coordination among member states,[324] but also EU institutions such as the just-mentioned EEAS,[325] and other partners[326] – e.g. NATO, UN, OSCE and more. This Strategy also recognizes the role of strategic communication in neutralizing MDM from extremists,[327] hence why it advocates for the spread of counter-narratives to dismantle these MDM and hopefully the terroristic propaganda or hate speech. Furthermore, the Council of the European Union acknowledges that extremist groups do exploit social media and digital platforms to disseminate their ideology, attack democracies, and recruit individuals,[328] hence why the EU sought to align its information security policies with its counterterrorism framework which ranges from leveraging cyber defense tools to engaging with technology companies to disrupt online extremist networks.

A report was issued in 2024 on the Strategic Compass demarking the policy advancements, notably with the implementation of the updated policies, such as the EU Policy on Cyber Defence.[329] The latter aims to mitigate digital vulnerabilities exploited by both state actors and terrorist organizations.

In conclusion, the general provisions demonstrate that the EU has a growing recognition that disinformation and terrorism are becoming increasingly interconnected, and they can pose a challenge, which requires a cohesive and proactive response.

[324] *Ivi.*, p. 26.
[325] *Ivi.*, p. 41
[326] *Ivi.*, p. 53
[327] *Ivi.*, p. 40.
[328] Council of the European Union, *EU measures to prevent radicalisation*, Consilium Europa, 2022, p. 1.
[329] EEAS, *Annual Progress Report on the Implementation of the Strategic Compass for Security and Defence*, European Union External Action, 2024, p. 8.

3. EU Strategies and Agendas

The main initial strategy was the 2005 EU Counter-Terrorism Strategy, since its adoption the political and legal framework has undergone significant transformations to reflect the evolving nature of security threats because of the increasing role of digital platforms in terrorist activities.

Especially with the integration of counter-disinformation measures in the EU security policies, it shows that the Union tries to adapt to address contemporary hybrid threats, also including the online sphere, as effectively as possible. As digital propaganda, MDM, and extremist content dissemination became more prominent, the EU progressively refined its approach, culminating in the regulatory and strategic frameworks introduced in 2020 and beyond.

3.1 The European Union Counter-Terrorism Strategy (2005)

In 2005 the EU formulated a response to the Madrid bombings[330] and established a foundational structure to avoid a similar event from happening again and thus the EU Counter-Terrorism Strategy was established. It's structured around four interdependent pillars:[331] (1) prevention; (2) protection; (3) pursuit; (4) response. The first pillar, prevention, sought to tackle the root causes of radicalization by promoting education initiatives, community engagement, and counter-narrative programs, all with the aim of dissuading individuals from engaging in extremist ideologies, or even letting these propaganda and distorted narratives get to them in the first place. The second pillar focused on protection measures which were developed to strengthen border security, enhance surveillance of critical infrastructure, and implement technological solutions to enhance public safety. The third pillar made sure that cross-border law enforcement cooperation could operate as intended, especially through the help of EU institutions such as Europol's intelligence-sharing capabilities and Eurojust's role in judicial coordination. And finally, the last

[330] Council of the European Union, *The European Union Counter-Terrorism Strategy*, Consilium Europa, 2005, p. 6.
[331] *Ivi.*, p. 3.

pillar being response, focused on crisis management and ensuring that post-attack interventions also included victim support mechanisms and emergency preparedness protocols.

In the end, while this framework provided by the EU Counter-Terrorism Strategy helped to construct the base of the counterterrorism strategy of the Union, it is a product of its time since, if we had only this strategy, it would be incomplete nowadays: especially as it does not contain specific provisions for addressing digital disinformation and the dissemination of extremist content online. Still, it is one of the foundational bases for the EU's counterterrorism and counter-MDM strategy.

3.2 Security Union Strategy and Counter Terrorism Agenda (2020)

Since the EU recognized the evolving nature of terrorist tactics, they decided to introduce the Security Union Strategy which was launched in 2020 until 2025. The logic is to enhance the resilience against emerging threats including the convergence of terrorism and disinformation. Thankfully, this strategy is built upon the foundational base of the 2005 framework but this time it incorporated a new dimension that was tailored to the digital age. Again, the identification and mitigation of hybrid threats became the central priority,[332] leading to the intention to create an EU research Hub for a better EU security ecosystem.[333] In fact, in the Seventh Progress Report, it is mentioned the EU Knowledge Hub for the Prevention of Radicalisation[334] which appears to be the result of this planned project. Similarly to this, there is the EU Hybrid Risk survey, tasked with detecting and countering cyber-enhanced MDM campaigns. Thus, it includes an expansion of protection measures and rigorous stress tests for critical infrastructure to ensure that sectors such as energy and transportation remain resilient against cyber-physical attacks. It has been given

[332] European Commission, *EU Security Union Strategy: connecting the dots in a new security Ecosystem (Press release)*, European Commission, 2020, p. 2.

[333] *Ibidem*.

[334] European Commission, *Communication from the Commission to the European Parliament and the Council on the Seventh Progress Report on the implementation of the EU Security Union Strategy*, 2024, p. 8.

significant importance to digital investigations, as they utilize artificial intelligence (AI) tools[335] to oversee encrypted communications among terrorists and explore dark web networks. Furthermore, formal partnerships with international organizations were planned[336] to establish and enhance online content moderation, thereby strengthening the collaboration between law enforcement agencies and major technology platforms in the identification and elimination of extremist content.[337]

In parallel to the Strategy, the EU introduced a refined approach structured around the principles of anticipation, prevention, protection, and response: the Counter-Terrorism Agenda. It is similar to the EU Counter-Terrorism Strategy, however, the Agenda does present some differences – and if not – it only reinforces the system.

A key component of the Agenda is improving threat anticipation through intelligence sharing and risk assessment. Member States are encouraged to contribute high-quality information to the EU Intelligence and Situation Centre (EU INTCEN) to enhance situational awareness.[338] Furthermore, the European Commission is dedicated to sending advisory missions to support national authorities in assessing risks to critical infrastructure, utilizing the expertise of EU Protective Security Advisors.[339] Investments in security research and technological advancements further strengthen the EU's proactive ability to detect emerging threats before they arise. The second step, preventing radicalization, is another essential element of the EU's counter-terrorism strategy. It confirms its acknowledgment of the influence of digital platforms in spreading extremist ideologies and thus calls for the urgent implementation of regulations to remove terrorist content online.[340] In this field, the EU Internet Forum (EUIF) is critical in developing content moderation standards to combat the dissemination of extremist narratives.[341] Additionally, the Agenda also recognizes the

[335] European Commission, *EU Security Union Strategy: connecting the dots in a new security Ecosystem (Press release)*, cit., p. 2.
[336] *Ivi.*, p. 1.
[337] *Ivi.*, p. 9
[338] European Commission, *A Counter-Terrorism Agenda for the EU: Anticipate, Prevent, Protect, Respond (COM 2020/795)*, 2020, p. 3.
[339] *Ibidem*.
[340] *Ibidem*.
[341] *Ibidem*.

importance of the need for social resilience by promoting educational, cultural, and youth initiatives that foster inclusion and help prevent radicalization. It also mentions prison radicalization, with efforts aimed at rehabilitating and reintegrating radicalized individuals after release.[342] In order to make knowledge learning quicker and more effective, it thought to centralize expertise, and thus, it is in this document that the Commission has proposed to create the EU Knowledge Hub dedicated to preventing radicalization.[343] The EU's counter-terrorism strategy also enhances protective measures for public spaces and critical infrastructure.[344] In light of recent attacks targeting crowded and symbolic sites, the Commission is committed to improving security through the principle of 'security-by-design,'[345] which includes both architectural resilience and strategic urban planning to reduce vulnerabilities.[346] A proposed EU Pledge on Urban Security and Resilience aims to support collaboration at the city level and provide access to funding for protective initiatives.[347] Additional measures focus on securing critical infrastructures like transportation hubs and power stations, while improvements in aviation security—potentially including EU-wide regulations for in-flight security officers—are also being considered[348]. Regarding operational support and response mechanisms, the Agenda points towards police collaboration and judicial coordination. The "future" EU Police Cooperation Code seeks to simplify cooperation among law enforcement agencies in Member States, enhancing their collective capability to address terrorist threats.[349] This – of course – means that the exchange of information is important, particularly concerning encrypted communications, where the Commission aims to balance privacy rights with the need for lawful access in counter-terrorism investigations. The Terrorist Finance Tracking Programme (TFTP) remains a vital tool for dismantling terrorist networks, with Europol supporting counter-terrorism financial investigators in disrupting funding streams.[350] Furthermore, intelligence from conflict zones is incorporated into EU databases to identify and monitor returning foreign terrorist fighters.[351] However, victim

[342] *Ibidem*.
[343] *Ibidem*.
[344] *Ibidem*.
[345] *Ibidem*.
[346] *Ibidem*.
[347] *Ibidem*.
[348] *Ibidem*.
[349] *Ibidem*.
[350] *Ibidem*.
[351] *Ibidem*.

protection also constitutes an integral part of the EU’s counter-terrorism approach. The Commission focuses on strengthening victims’ rights, as well as improving access to compensation and psychological support for those affected.[352]

Finally, the Agenda further strengthens international cooperation, especially with the Western Balkans, North Africa, and critical regions in Asia, to combat terrorism financing, illegal arms trafficking, and the movement of foreign fighters.[353] Most interesting for the topic of this thesis are partnerships with organizations like Europol and the European Public Prosecutor’s Office which enhance even more the EU’s capacity to respond effectively to terrorist threats.[354] A small comment on Europol is that its role in counterterrorism has been significantly strengthened by the Agenda. With this enhanced mandate, the agency can collaborate more closely with private companies to detect and prevent terrorist activities on digital platforms. Europol, considering the report to the Agenda, is now better equipped to analyze large and intricate datasets, enhancing its ability to identify potential threats.[355] However, the improved collaboration is not only local but with international partners, who enable law enforcement agencies to effectively combat terrorism beyond the EU's borders.

Thus, continuing the EU Security Union Strategy, the Counter-Terrorism Agenda builds on existing legal frameworks that address issues like terrorist financing, firearm access, and cross-border law enforcement cooperation. By combining preventive strategies with solid enforcement measures, the Agenda aims to create a robust security infrastructure that can adapt to modern terrorist threats while protecting fundamental rights and freedoms within the EU.

The EU's Security Union Strategy (2020) has redefined counterterrorism efforts through a coordinated framework and targeted digital measures, merging regulatory requirements with operational tactics to effectively neutralize threats. As noted by the report from 2024 aiming to explain the progress made by the implementation of the Security Union Strategy, the 2020

[352] *Ibidem*.
[353] *Ibidem*.
[354] *Ibidem*.
[355] *Ibidem*.

Counter-Terrorism Agenda implemented mandatory legislative measures, particularly Directive (EU) 2017/541,[356] fully adopted by all Member States by 2024, which criminalizes terrorist training, travel, and financing. Concurrently, Regulation (EU) 2021/784 led to the removal of 500 items of terrorist content from online platforms between June 2022 and April 2024.[357] Specifically, the regulation binds hosting states to, if they receive an order for removal from the authorities of the member state, take down completely the content – or at least remove access – by blocking within one hour.[358] The aftermath of the 2023 Hamas attacks demonstrated the effectiveness of the regulation, as it enabled the rapid removal of antisemitic and jihadist propaganda before its dissemination. The Strategy acknowledges that conflicts occurring outside the EU can influence the threat of terrorism within its borders, thereby increasing the risk to the Union.[359] Beyond removal orders, the EU has leveraged financial and technical resources to strengthen Member States' cybersecurity capabilities. To support the Digital Europe Programme, EUR 84 million has been allocated to support cybersecurity initiatives, where they also embrace the application of AI in Security Operation Centres with the introduction to post-quantum cryptography.[360] These complement the work of the European Cybersecurity Competence Centre,[361] which ensures that technological advancements in threat detection and mitigation benefit businesses, SMEs, and public administrations across the EU.

The Security Union Strategy, following the overall EU logic on threat prevention, also underscores the importance of preemptive measures to combat radicalization, particularly through the Radicalisation Awareness Network (RAN). The report reminds that the network, which includes over 6,500 practitioners, has been instrumental in shaping policy responses and best practices to counter violent extremism.[362] In fact, the EU has integrated RAN into the EU Knowledge Hub for the Prevention of Radicalisation as of June 2024, striving to enhance intersectoral collaboration,

[356] European Commission, *Seventh Progress Report on the implementation of the EU Security Union Strategy*, cit., pp. 6–7.
[357] *Ivi.*, p. 8.
[358] *Ibidem*.
[359] *Ivi.*, p. 9.
[360] *Ivi.*, p. 4.
[361] *Ibidem*.
[362] *Ivi.*, p. 8.

provide strategic foresight, and improve the effectiveness of counter-radicalization policies.[363] A quick comment, since the report brings it up, on financial measures which remain central to the EU's counterterrorism framework: as of 2020, the Internal Security Fund has allocated EUR 30 million to the PROTECT programme, to help the safeguarding of public spaces and places of worship, including synagogues and mosques (and additionally they decided to fund EUR 5 million more specifically reserved to counter rising antisemitic threats).[364] Concurrently, regulatory reforms have tightened financial oversight to prevent terrorist organizations from exploiting EU funds, with new provisions in the revised Financial Regulation (2023) barring entities convicted of «incitement to hatred » from accessing EU financial support.[365]

Moreover, still according to the report, the EU has intensified its scrutiny of artificial intelligence (AI) and its potential exploitation by terrorist groups. The EU Internet Referral Unit at Europol has been proactive in identifying and referring extremist content across over 300 platforms,[366] bolstered by the enforcement of the Digital Services Act (applicable from February 2024). The EU Internet Forum (EUIF) has further reinforced these efforts by engaging technology companies in discussions on AI-generated extremist propaganda, aiming to align industry standards with regulatory expectations. Finally, the EU's commitment to countering terrorism and disinformation extends to law enforcement coordination and judicial responses. Remarkably, the report states that in 2022 alone, «28 terrorist attacks were completed, failed, or foiled »[367] across the member states, leading to the arrest of 380 individuals, 14 of whom were linked to terrorist financing.[368] Thus, Eurojust has played an essential role in coordinating 203 investigations, including eight Joint Investigation Teams.[369] Additionally, measures such as systematic checks in the Schengen Information System have been reinforced to prevent undetected entries of foreign terrorist fighters into the EU.

[363] *Ibidem*.
[364] *Ivi.*, p. 7.
[365] *Ivi.*, p.8.
[366] *Ivi.*, p. 9.
[367] *Ibidem*.
[368] *Ibidem*.
[369] *Ibidem*.

4. EU Internet Forum (2015)

The EUIF has been mentioned several times already in this thesis but not quite fully explained. It was established in 2015 and is an essential actor in the EU's strategy to counteract the exploitation of digital platforms by terrorist organizations while simultaneously addressing the broader challenges of disinformation. It integrates content removal mechanisms with proactive counter-narrative initiatives, and by doing that, the Forum illustrates the evolving role of international organizations in governing information within counterterrorism frameworks. Its dual approach underscores the necessity of balancing security imperatives with fundamental rights considerations, a challenge inherent to the regulation of digital spaces.

A central component of the Forum's operational framework is the development and deployment of technical initiatives to identify and eliminate terrorist content as well as the proper establishment of the Global Internet Forum to Counter Terrorism (GIFCT) which also aids in the fight, but this time in an international setting. The establishment of a hash database, maintained in collaboration with major technology firms such as Google, Facebook, and Microsoft, enabled the creation of digital fingerprints for over 40,000 hashes by 2017.[370] This automated detection mechanism facilitates cross-platform content removal, preventing the dissemination of extremist material. In fact, the Director of Global Policy Management of Facebook – Monika Bickert – stated that:

> «*The use of AI and other automation to stop the spread of terrorist content is showing promise. Today, 99% of the ISIS and Al Qaeda-related terror content we remove from Facebook is content we detect before anyone in our community has flagged it to us, and in some cases, before it goes live on the site. We do this primarily through the use of automated systems like photo and video matching and text-based machine learning. Once we are aware of a piece of terror content, we remove 83% of subsequently uploaded copies within one*

[370] European Commission, *Fighting Terrorism Online: Internet Forum pushes for automatic detection of terrorist* propaganda *(Press release)*, 2017, p. 2.

> *hour of upload. We recognize that we can always do more and are looking forward to strengthen our collaboration with the European Commission and others as we deepen our joint commitment to combating terrorism.* »[371]

This statement not only proves of the active participation of these technological firms, but their commitment to the cause and to collaboration with the European Commission in the context of the EU Internet Forum. Similarly, EMEA President of Google – Matt Brittin – which represent another of the major industry in the technological field, stated that:

> «*Addressing violent extremism is a critical challenge for us all and we're committed to being part of the solution, together with governments and civil society. We've made significant progress in 2017, deploying machine learning technology, strengthening our enforcement and expanding our partnerships with experts. We've improved the speed and accuracy of our removals - today, 98 percent of the videos we remove for violent extremism on YouTube are flagged to us by machine-learning algorithms, up from 75 percent just a few months ago, and we've a goal to bring the number of people working to tackle problematic content across Google to 10,000 in 2018. The EU Internet Forum has played a vital role in driving collaboration on these issues and helped lay the foundation for the Global Internet Forum to Counter Terrorism, where we partner with other tech companies to curb the spread of terrorist material online.* »[372]

[371] *Ivi*., p. 1.
[372] *Ivi.,* p. 2.

However, concerns persist regarding the opacity of this system, particularly concerning error rates and the potential implications for fundamental freedoms. In response to these threats, the EU Crisis Protocol (EUCP) – revised in 2023 – introduced real-time coordination cooperation,[373] allowing for swift intervention by EU member states and industry stakeholders during crises, all to avoid phenomena such as the 2019 Christchurch attack. This protocol helps to significantly reduce content takedown times, demonstrating the efficacy of this multilateral crisis management strategy. Finally, beyond technical solutions, the Forum has prioritized civil society engagement as a means of countering disinformation and extremist narratives.

The EU Crisis Protocol Civil Society Empowerment Programme (CSEP) has also played a crucial role in equipping organizations with the necessary projects to challenge terrorist propaganda. Considering the budget which was initially thought to be EUR 10 million for the support to Civil Society Organizations,[374] in the end, the amount provided was EUR 13’840’268 as the maximum grant amount provided for the two projects in 2017 and 2018.[375] From these EUCP funded projects, the highest scoring based on the GAMMA+ Model directed against extremist propaganda were “Resonant Voices Initiative in the EU” (RVIEU)[376] and Online Positive Narratives for the Prevention of Radicalisation within Flemish Education (ONarVla), which scored 42.5 and 42 respectively,[377] mostly above/on average.

5. Digital Services Act (2022)

The EU has taken a leading role in the global fight against online disinformation, implementing a robust regulatory and policy framework that reinforces counterterrorism by addressing the digital sphere used also by extremist groups. The Digital Services Act (DSA), introduced in 2022 along

[373] LUDDEN et al., *Evaluation of impact and effectiveness of counter- and alternative campaigns stemming from the CSEP programme aiming at preventing radicalisation leading to violent extremism and terrorism (Final Report)*, Publications Office of the European Union, 2022, p. 152.
[374] *Ivi.*, p. 179.
[375] *Ibidem*.
[376] *Ivi.*, p. 166.
[377] *Ibidem*.

with the updated Code of Practice on Disinformation and the new EU Commission Code of Conduct on Disinformation (2025): all these Codes reflect a comprehensive strategy that merges enforceable legal standards with voluntary industry collaboration. This initiative falls within the wider regulatory framework known as the *Digital Services Package* and marks the EU's one of most significant actions in order to tackle the complex challenges of digitalization, particularly regarding illegal content, MDM, and the protection of fundamental rights online. The DSA's main goal is to create a safer digital space for consumers and businesses in the EU, which is planned to be achieved through a set of binding regulations focusing on five key objectives: (1) protecting consumers and their fundamental rights;[378] (2) establishing clear and proportionate responsibilities for online platforms and social media companies;[379] (3) addressing illegal content such as hate speech, disinformation;[380] (4) improving transparency in how online services operate;[381] and (5) fostering innovation, growth, and competitiveness in the EU's digital market.[382] Thus, the EU clearly strives for the protection of human rights with its consumer protection policy, rooted in Article 169 TFEU, which is a shared competence and responsibility between the member states and the EU: the EU establishes common rules through ordinary legislative procedure while the member states retain the possibility to adopt or maintain more stringent measures, as long as they don't conflict with EU law and that the EU Commission has been properly informed.[383] Furthermore, to ensure that consumer protection is never forgotten in all EU policies, Article 12 TFEU establishes that it must be considered.[384] Thus reinforcing its horizontal dimension across different regulatory fields.

The scope of application of the DSA is both broad and differentiated. The regulation applies to a wide range of intermediary service providers involving: internet access providers, hosting services such as cloud computing and web hosting, name registrars, online marketplaces, app stores, collaborative economy platforms, social networks, content-sharing platforms, and online travel

[378] Council of the European Union and European Parliament, *Regulation (EU) 2022/2065 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act)*, Official Journal of the European Union, Article 1(1), 2022, p. 41.
[379] *Ivi.*, Chapter II, pp. 44 – 48.
[380] *Ivi.*, Chapter I, Article 3, pp. 42 – 44.
[381] *Ivi.*, Chapter III, pp. 48 – 78.
[382] *Ivi.*, Chapter I. Article 1(1), p. 41.
[383] EU, *Consumer Protection*, EUR-Lex, 2023, p. 1.
[384] *Ibidem*.

and accommodation platforms. However, the DSA introduces a particularly rigorous set of obligations for Very Large Online Platforms (VLOPs) and Very Large Online Search Engines (VLOSEs), defined as those services reaching more than 10% of the EU population[385] – i.e., over 45 million users.[386]

Thus, a significant innovation of the DSA is its multifaceted approach to protecting the rights and interests of EU citizens by creating a strong mechanism to tackle illegal online content, including unlawful goods, services, and harmful information – such as disinformation and hate speech. Users are equipped with improved tools to manage their online experience, primarily through better awareness of the advertisements they see and an easier process to report illegal content or products. Additionally, the DSA promotes collaboration between online platforms and trusted flaggers,[387] who are recognized for their ability to identify harmful material. The regulation also requires online marketplaces to ensure trader traceability[388] in order to enhance accountability within the digital landscape.

Furthermore, the DSA aims to empower users and civil society: this is achieved through various procedural guarantees, including the right for users to contest content moderation decisions and seek compensation via a dedicated dispute mechanism or judicial routes.[389] Moreover, the regulation grants unprecedented access to key data produced by VLOPs for relevant authorities and independent researchers,[390] thus allowing for better monitoring and evaluation of systemic risks in the digital environment. Algorithmic processes also take care of the transparency requirements, especially those related to content or product recommendations.

In addition, the DSA employs a risk-based regulatory approach for platforms. VLOPs and VLOSEs are required to establish risk management systems[391] aimed at preventing the misuse of their services, especially regarding the spread of disinformation and other harmful content. Specifically,

[385] EU, *Digital Services Act*, EUR-Lex, 2022, p. 1.
[386] *Ibidem*.
[387] *Ibidem*.
[388] Council of the European Union and European Parliament, *Regulation (EU) 2022/2065*, cit., Preamble (72), p. 20.
[389] *Ivi.*, Chapter III, Section 3, Art. 20 – 21, p. 53 – 56.
[390] *Ivi.*, Chapter III, Section 5, Art. 40, pp. 70 – 72.
[391] *Ivi.*, Chapter III, Section 5, Art. 34, pp. 64 – 65.

these systems must be subject to follow certain rules – such as independent audits or data sharing[392] – to verify their effectiveness and compliance. However, the regulation could demand platforms to have crisis response plans,[393] meaning that they are needed to respond promptly and effectively to exceptional situations that could worsen the spread of disinformation or amplify terrorist propaganda online. Besides that, specific measures to protect children and limits on the use of sensitive personal data for targeted advertising[394] even further demonstrate the DSA's commitment to a user-focused regulatory framework.

Finally, as the last comment on the DSA, it aims to strengthen the supervisory and enforcement mechanisms in the EU's digital space. Finally, as the last comment on the DSA, it aims to strengthen the supervisory and enforcement mechanisms in the EU's digital space. The regulation has to intend to establish independent Digital Services Coordinators in each Member State, along with the formation of the European Board for Digital Services:[395] they are tasked with overseeing the proper enforcement of the regulation and ensuring effective coordination at the EU level. Importantly, the European Commission, European Board for Digital Services, and Digital Services Coordinators gain additional supervisory authority over VLOPs and VLOSEs,[396] acknowledging their cross-border operations and substantial influence on the integrity of the digital ecosystem.

In circumstances of crisis, the DSA foresees the activation of a Crisis Response Mechanism,[397] through which the European Commission may require VLOPs and VLOSEs to assess the extent to which their services contribute to a serious threat to public security or public health within the EU.[398] Where appropriate, these platforms have to adopt effective and proportionate risk-mitigation measures, as delineated in the regulation,[399] and report to the Commission on their assessments and responses.[400] This innovative mechanism, presented by the DSA, proves the EU's

[392] *Ivi.*, Chapter III, Section 5, Art. 37, pp. 67 – 69.
[393] *Ivi.*, Chapter III, Section 5, Art. 36, pp. 66 – 67.
[394] *Ivi.*, Chapter III, Section 3, Art. 28, p. 60.
[395] *Ivi.*, Chapter IV, Section 3, Art. 61, p. 87.
[396] *Ibidem*.
[397] *Ivi.*, Chapter III, Section 5, Art. 36(1), pp. 66 – 67.
[398] *Ibidem*.
[399] *Ibidem*.
[400] *Ibidem*.

determination to address emergent risks – such as the dissemination of terrorist content or disinformation campaigns – with flexibility and adaptability.

5.1 Code of Practice on Disinformation (2022)

As part of the 2018 Action Plan against Disinformation, the European Commission conducted, in 2020, its first comprehensive assessment of the Code of Practice on Disinformation,[401] initially introduced in 2018 as the first voluntary framework at the global level to tackle online disinformation. The evaluation recognized the Code as an innovative instrument that succeeded in fostering dialogue between platforms, advertisers, and public authorities, while also enhancing transparency regarding platforms' anti-disinformation strategies.[402] Nevertheless, the self-regulatory nature of the Code emerged as its principal weakness, limiting its capacity to deliver consistent and verifiable results. This position was supported by Vice-President Věra Jourová, who recognized the Code's positive impact, but still, he stressed the need to move beyond voluntary measures and instead towards more stringent mechanisms of accountability.[403] Similarly, Commissioner Thierry Breton emphasized the strategic relevance of the Code in defending European values in the digital sphere, while also underlining the shared responsibility of online platforms and the advertising sector in ensuring the integrity of the information ecosystem.[404] Despite initial progress – such as the first baseline reports submitted by signatories including Google, Facebook, X (formerly 'Twitter'), Microsoft, Mozilla, and later TikTok[405] – the Commission identified several persistent shortcomings: among these were the absence of adequate key performance indicators (KPIs);[406] limited access to data for independent scrutiny;[407] vague commitments;[408] insufficient collaboration with the research community.[409] Besides, the COVID-

[401] European Commission, *Disinformation: EU assesses the Code of Practice and publishes platform reports on coronavirus related* disinformation *(Press release*), 2020, p. 1.
[402] *Ibidem*.
[403] *Ibidem*.
[404] *Ibidem*.
[405] European Commission, *Code of Practice on Disinformation*, 2018, p. 1.
[406] European Commission, *Disinformation: EU assesses the Code of Practice*, cit., p. 2.
[407] *Ibidem*.
[408] *Ibidem*.
[409] *Ibidem*.

19 pandemic further tested the Code's capacity, leading to the publication of specific reports on the platforms' efforts to counter specifically coronavirus-related MDM. As such, these measures ranged from promoting verified content and supporting fact-checkers to curbing manipulative behaviors and limiting advertising profits from false narratives. However, the Commission concluded that more systemic and enforceable solutions were necessary.[410] Consequently, it announced the forthcoming European Democracy Action Plan and the DSA package,[411] intended to strengthen the regulatory framework and ensure a more resilient and transparent digital information environment within the Union.

The European Commission released, in 2021, a Guidance on Strengthening the Code of Practice on Disinformation,[412] which outlines a series of strategic recommendations aimed at transforming the existing Code into a more operational and effective instrument against the growing threat of disinformation.[413] Within the broader context of counterterrorism, this initiative assumes particular relevance as online disinformation campaigns increasingly intersect with radicalization processes and the dissemination of extremist narratives. The EU Commission recognizes the existing Code as a pioneering framework at the global level, fostering structured dialogue among relevant stakeholders and enhancing the transparency and accountability of online platforms' policies concerning disinformation.[414] In this perspective, the Guidance identifies several critical areas where reinforcement is required in order to guarantee a coherent and comprehensive application of the Code's principles across all actors and Member States. These areas include the expansion of participation,[415] encouraging a broader range of stakeholders to adopt commitments adapted to their specific services and operational contexts. For example, it further calls for more effective demonetization mechanisms targeting MDM content,[416] thereby disrupting the financial incentives that often sustain the production and dissemination of harmful narratives, including those linked

[410] *Ibidem*.
[411] *Ibidem*.
[412] European Commission, *European Commission Guidance on Strengthening the Code of Practice on Disinformation*, in *Communication to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions*, 2021, pp. 1 – 24.
[413] *Ivi.*, p.1.
[414] *Ibidem*.
[415] *Ivi.*, p. 2.
[416] *Ivi.*, p. 2 & p. 6.

to terrorist propaganda. The integrity of online services is to be strengthened through enhanced safeguards against manipulation techniques[417] – such as fake accounts and bots for example – which are frequently used in the spread of terrorist disinformation campaigns. Moreover, the Commission stresses the importance of empowering users by providing them with the necessary tools and information to identify and report disinformation.[418] Especially, there is a particular emphasis on the expansion of fact-checking coverage and the improvement of data accessibility for researchers: both represent a fundamental step in developing evidence-based policies to counter online disinformation in the context of terrorism.[419] Additionally, the establishment of a robust monitoring framework is envisaged to ensure greater oversight and evaluation of the implementation of the Code's commitments.

A key innovation introduced by the Guidance is the proposal to create a Transparency Centre.[420] This would serve as a centralized platform where signatories disclose the specific policies adopted to enforce the Code, detail their enforcement efforts, and publish all relevant data and metrics linked to the Key Performance Indicators (KPIs). In parallel, the Commission recommends the creation of a permanent task-force, chaired by the Commission itself, to oversee the continuous adaptation and operationalization of the Code.[421] Finally, the Guidance calls upon the signatories of the Code of Practice to engage in a collective effort to revise and strengthen the document in line with the proposed recommendations. A first draft of the reinforced Code was expected to be presented in the autumn of the same year,[422] marking a crucial step forward in the EU's strategy to counter disinformation, including in the sensitive area of terrorism prevention and response. Although there were some concerns regarding the feasibility and potential repercussions of broadening the Code's scope to encompass MDM.[423] The absence of a clear, uniform definition of disinformation across Member States already complicates enforcement efforts.[424] The practical challenge of distinguishing between intentional and unintentional dissemination of false content

[417] *Ivi.*, p. 13.
[418] *Ibidem*.
[419] *Ivi.*, p. 2.
[420] *Ivi.*, pp. 23-24.
[421] *Ivi.*, p. 24.
[422] European Commission, *Guidance on Strengthening the Code of Practice on Disinformation*, *Policy and Legislation*, European Commission Website, 2021, p. 1.
[423] NENADIĆ, *EC's Guidance to Strengthening the Code of Practice on Disinformation: A Mis-take with Mis-information?*, Centre for Media Pluralism and Media Freedom, 2021, pp. 1 – 2.
[424] *Ibidem*.

renders consistent implementation improbable.[425] Moreover, uncertainties persist as to who should determine, and by which criteria,[426] whether a particular piece of content entails a «significant public harm dimension ».[427]

In 2022, the European Commission introduced the Strengthened Code of Practice on Disinformation ("the Code"), following various documents like the Assessment and Guidance. This new Code represents a significant upgrade from its 2018 version since, while the original Code established a foundation for voluntary cooperation among online platforms and stakeholders to combat MDM, the 2022 one aims to actually address the limitations of its predecessor with a more structured, ambitious, and comprehensive approach. The primary goal is to create a safer, more transparent, and accountable online space by enhancing private actors' responsibilities and providing more targeted operational tools to handle disinformation's complexities.[428] The renewed Code distinguishes itself by expanding its scope and refining its commitments. Unlike the 2018 version, which mainly focused on large online platforms, the 2022 Code invites a broader range of participants, including not just social media giants, but also smaller platforms, advertisers, fact-checkers, civil society organizations, and research communities.[429] This strategic expansion acknowledges that disinformation dynamics are not limited to a few dominant platforms; they are dispersed across various actors in the digital ecosystem.[430] Thus, the aim is not only to strengthen self-regulation but also to ensure that all relevant sectors play their role in detecting, preventing, and mitigating disinformation practices.

With 44 commitments and 128 measures,[431] the Code establishes a framework for mitigating the spread of harmful content while preserving fundamental rights. In the counterterrorism domain, all key provisions stand out: (1) the demonetization of disinformation networks disrupts extremist

[425] *Ibidem*.

[426] *Ibidem*.

[427] European Commission, *European Commission Guidance on Strengthening the Code of Practice on Disinformation*, cit., p. 5.

[428] European Commission, *The Strengthened Code of Practice on Disinformation*, 2022, pp. 1–5.

[429] *Ivi.*, p. 26.

[430] *Ivi.*, p. 40.

[431] European Commission, *The 2022 Code of Practice on Disinformation*, European Commission Website, 2025, pp 1 – 4.

financing by restricting their access to advertising revenues, thereby undermining their ability to sustain propaganda operations;[432] (2) political advertising transparency measures[433] expose covert influence campaigns by requiring clear sponsorship disclosures, a critical tool in identifying state-backed or ideologically driven attempts to manipulate public discourse; (3) integrity protocols further enhance security by obligating platforms to eliminate fake accounts and deepfakes,[434] which could be used in radicalization efforts; (4) empowering users by enhancing media literacy but also aid them to identify MDM as harmful, false or misleading information;[435] (5) empowering the research community as they strive to a better cooperation with researchers while also having access to Signatories' data (non-personal data and anonymized) for research on Disinformation;[436] (6) Strengthening the fact-checking community[437] to make sure that true information stay afloat compared to the sea of information found on social media, for example. While these commitments contribute to a more secure online environment, limitations remain, as evidenced by major platforms such as X withdrawing from voluntary obligations, signaling the need for stronger enforcement mechanisms linked to the DSA.[438]

One of the best features of the Strengthened Code lies in the greater precision of its commitments, which are now actually framed around well-defined operational principles. These commitments include, among others, the demonetization of disinformation actors by restricting their access to advertising revenues,[439] the guarantee of transparency in political and issue-based advertising,[440] the promotion of meaningful cooperation with fact-checkers, and the facilitation of researchers' access to relevant platform data.[441] These commitments indicate a clear departure from the rather general and flexible formulations of the 2018 Code, thus opting instead for measurable objectives and verifiable practices. As such, the introduction of Qualitative Reporting Elements (QRE),

[432] European Commission, *The Strengthened Code of Practice on Disinformation*, cit., pp. 5–7.
[433] *Ivi.*, pp. 9-14.
[434] *Ivi.*, pp. 17.
[435] *Ivi.*, pp. 18-25.
[436] *Ivi.*, pp. 26-30.
[437] *Ivi.*, pp. 31-34.
[438] European Commission, *Commission sends preliminary findings to X for breach of the Digital Services Act*, 2024, pp. 1 – 2.
[439] European Commission, *The Strengthened Code of Practice on Disinformation*, cit., pp. 5–8.
[440] *Ivi.*, p.9.
[441] *Ivi.*, p. 34.

Service Level Indicators (SLI),[442] and a strengthened reporting mechanism represents a decisive improvement in terms of accountability. Signatories are now required to regularly report on their progress and provide data that allow for external scrutiny, thus reducing the opacity that had often undermined the credibility of self-regulatory initiatives. Furthermore, the Code also defines with greater precision an important element that many other documents, both EU and UN, missed: a precisely explained terminology. In the Preamble, in fact, they specify that «the Signatories of the present Code of Practice (the "Code") recognize their role in contributing to the fight against Disinformation, which for the rest of the Code is considered to include misinformation, disinformation, information influence operations and foreign interference in the information space (Disinformation) ».[443] Thus, they clarify that even if they are only mentioning ‘disinformation’, they refer to the entire spectrum of information manipulation which includes many other phenomena, such as misinformation. This is highly important especially to this thesis as it proves that the EU is taking a proactive approach not only for disinformation but also for the rest of the MDM, which often are left in the dark.

The new Code's principles reflect a comprehensive grasp of how disinformation spreads online. An important focus is on the transparency of political advertising (chapter III of the Code), seen as crucial for disseminating false or misleading narratives, especially during elections. Besides, the Code mandates that signatories label political ads distinctly, maintain public ad repositories, and allow access to these repositories through Application Programming Interfaces (APIs),[444] enabling external parties to effectively monitor political communication practices. Additionally, the Code introduces measures to combat manipulative tactics that artificially boost disinformation, including bots, fake accounts, deepfakes, and other coordinated inauthentic activities.[445] Another significant aspect of the 2022 Code is the already-mentioned empowerment of users, fact-checkers, and researchers: platforms are urged to create tools that help users easily identify, flag, and report MDM, while also providing greater visibility and support for fact-checkers fighting false narratives. Besides that, improving data access for researchers represents an improvement toward reconciling private data ownership with the public’s interest in examining disinformation issues.

[442] *Ivi.*, p.2.
[443] *Ivi.*, p.1.
[444] *Ivi.*, p.12-14.
[445] *Ivi.*, p. 16.

Platforms are expected to provide anonymized and non-personal data to facilitate independent research while respecting user privacy and data protection, which has been a long-standing request from academia. Collectively, the measures outlined in the Strengthened Code of Practice on Disinformation illustrate the European Commission's intent to move from mere symbolism to actionable steps. More than just a statement of intent, the 2022 Code serves as a strategic tool aimed at achieving measurable and enforceable changes in online information governance. However, its success depends on the commitment of all signatories to actually uphold their obligations in a diligent manner, plus the establishment of effective monitoring systems to ensure compliance. Thus, while the 2022 strengthened Code represents a significant step forward in the EU's approach to regulating MDM, its effectiveness will still mostly rely on the balance between voluntary cooperation and the potential for regulatory intervention, if self-regulation falls short.

5.2 The EU Code of Conduct on Disinformation

In 2025, more specifically on February 13th, the EU advanced its strategy against online MDM by formally anchoring the Code of Practice on Disinformation within the legal framework of the DSA.[446] This development was endorsed by both the European Commission and the European Board for Digital Services (EBDS),[447] effectively elevating the status of the Code from a self-regulatory instrument to an actually recognized Code of Conduct operating under Article 45 of the DSA.[448] Said Article works as the Code will now «contribute to the proper application of »[449] the Digital Service Act, which was already the case with the Code of Conduct on Countering Illegal Hate Speech Online. However, this Code of Conduct will take effects starting from the 1st July 2025, which is the date when the finalized conversion into the DSA framework will happen.[450]

[446] European Commission, *The Code of Conduct on Disinformation*, European Commission Website, 2025, p. 1.
[447] *Ibidem*.
[448] KING, O'KEEFFEE, ENGLISH, *EU's Code of Conduct on Disinformation Integrated into DSA,* Tech Law Blog, 2025, p. 1.
[449] *Ibidem*.
[450] European Commission, *The Code of Conduct on Disinformation*, cit., p. 1.

Although the Code's integration into the DSA does not transform it into a mandatory obligation, its relevance within the Union's regulatory landscape has been considerably strengthened. The Commission – together with the EBDS – outlined a set of recommendations designed to facilitate the effective implementation of the Code's commitments: these recommendations, while not legally binding, reflect the EU's intention to operationalize the Code as an essential tool in addressing systemic risks associated with disinformation. Among these proposals is the refinement of the Rapid Response System – operational since 2024– capable of providing swift action in moments of electoral sensitivity or crisis situations.[451] Moreover, the Commission encouraged regular cooperation within the Code's permanent Task Force and placed particular emphasis on improving the quality and completeness of data reporting from signatories, enabling the development of credible indicators to measure progress in combating disinformation.[452]

Interestingly, the DSA decided to adopt quite a unique regulatory approach as it combines voluntary engagement with subtle forms of regulatory pressure. So, although participation in the Code remains optional, refusing to do so without justified reasons may be taken into consideration by the Commission when evaluating whether a platform has complied with its risk mitigation duties under Article 35 (1)(h) of the DSA.[453] Furthermore, for Very Large Online Platforms (VLOPs) and Very Large Online Search Engines (VLOSEs), adherence to the Code may serve as evidence of their efforts to limit systemic risks, particularly in the context of the independent audits foreseen under Article 37.[454]

Nevertheless, some concerns have emerged regarding the real-world impact of the Code, especially in relation to certain fact-checking obligations that some industry actors perceive as excessive[455] or not necessarily the most effective means of tackling disinformation.[456] The success of the Code, therefore, will likely depend on the extent to which platforms perceive its commitments not only as a matter of voluntary engagement but also as a strategic element in their broader compliance

[451] KING, O'KEEFFEE, ENGLISH, *EU's Code of Conduct on Disinformation Integrated into DSA*, cit., p.1.
[452] *Ibidem*.
[453] *Ibidem*.
[454] *Ibidem*.
[455] NAUTA, *You can't fact check disinformation away. What can you do?*, Free Press Unlimited, 2021, pp. 1–2.
[456] KING *et al.*, *EU's Code of Conduct on Disinformation Integrated into DSA*, cit., p. 1.

with EU regulation. In a context where disinformation merges with threats to public safety, radicalization, and terrorism, implementing the Code under the DSA framework marks a major advancement in the EU's ongoing strategy to protect its information space from domestic and foreign interference.

6. Conclusion

The EU's contribution to countering MDM in the context of terrorism showcases a comprehensive and coordinated institutional strategy, primarily expressed through the EU Internet Forum and additional regulatory measures. The EU's advantage is its ability to unite a wide range of stakeholders and promote public-private partnerships that improve the removal of terrorist material online while respecting fundamental rights. The introduction of harmonized legislation, especially Regulation (EU) 2021/784 targeting the spread of terrorist content online, clearly exemplifies the Union's legal assertiveness by establishing mandatory removal timelines (within one hour) and enhancing transparency through reporting requirements. Moreover, programs like the RAN and Europol's Internet Referral Unit provide operational support and facilitate knowledge sharing, aiding in both detection and prevention efforts. In fact, the Union has effectively enabled the removal of significant amounts of terrorist content; e.g., the detection of extremist content on more than 300 platforms thanks to the Internet Referral Unit.

Nevertheless, limitations persist. The voluntary nature of certain initiatives, the reliance on private platforms' cooperation, and the inherent challenge of balancing security imperatives with freedom of expression pose enduring constraints. Fragmentation risks due to uneven national implementations and differing interpretations of fundamental rights further complicate enforcement consistency. Besides, while technical removal capacities have advanced, preventive strategies addressing the root causes of radicalization in the digital sphere remain less consolidated.

CHAPTER IV

Euro-Atlantic Security Organizations' role against MDM and Terrorism

1. Introduction

This chapter will focus on the two major international organizations between the North Atlantic and Europe, being the Organization for Security and Co-operation in Europe (hereafter OSCE) and the North Atlantic Treaty Organization (NATO). The choice of the international organization is not fortuitous: they are both focused on the security aspect at an international level, however, operating in different spheres. The first one is a military and political defense alliance, the second one is a political and security organization. Interestingly enough, they both could be considered international organizations of a regional type, as NATO only allows for countries from the North American and European continents, and OSCE is considered the "world's largest regional organization" with at its core the European continent.

Furthermore, another reason for this choice of international organization lies in the fact that a comprehensive analysis of all international organizations cannot be undertaken within the scope of a single thesis. So, to counter this practical limitation, this thesis will be focusing on a specific institutional area that enables a clearer and more comprehensive understanding of how the fight against terrorism is addressed across multiple dimensions. In this case, the European continent is central. In fact, out of the 27 members in the European Union, 23 are also in NATO.[457] Besides, ever since 2006, all EU member states have also been part of the OSCE.[458] Hence why these two international organizations would not only complete the overview of this specific area but also analyze their contribution in different areas than the EU and the ONU, being military and security.

2. Organization for Security and Co-operation in Europe

The Organization for Security and Co-operation in Europe (OSCE) is quite an important actor in the fight against online MDM in counterterrorism efforts. The OSCE seeks to reconcile the imperatives of national security with the fundamental principles of human rights, particularly the protection of freedom of expression, which is done through a multi-dimensional approach. This capacity for balance is demonstrated in its array of projects, all of which shape policy, train, build partnerships, and create societal resilience around the dissemination of terrorist content and harmful narratives. The OSCE's notion of societal resilience implies that security, media integrity, and human rights are closely linked concepts, and seeks to position itself as a key actor of influence on the development of counter-disinformation in OSCE states.

The OSCE's comprehensive approach also revolves around engaging the broadest range of societal actors. It recognizes that counter-disinformation must be informed by a multitude of perspectives, including policymakers, educators, media professionals, and civil society.

A great example of this approach is the 2023 training initiative in Bosnia and Herzegovina, which equipped stakeholders with the necessary tools to combat digital information disorder while steadfastly upholding the principles of free speech.[459] Similarly, the INFORMED project, which still represents a significant contribution of the OSCE's counter-MDM work, places a strong

[457] European Council, *EU-NATO cooperation*, Consilium Europa, 2024 (updated), p. 1.

[458] OSCE, *The European Union*, Organization for Security and Co-operation in Europe Website, p. 1.

[459] OSCE Secretariat Transnational Threats Department, *Preventing and countering violent extremism online through media and information literacy focus of OSCE training programme*, Organization for Security and Co-operation in Europe Website, 2023, p. 1.

emphasis on media literacy to build a societal defense against extremist narratives.[460] This project is particularly interesting due to its inclusion of local voices and gender-sensitive considerations, which fit with the OSCE's commitment to acknowledging these concerns and utilizing inclusiveness in its security-based frameworks. The OSCE also advocates for public-private partnerships in the global effort to combat the exploitation of digital and digital platforms by terrorists while seeking safeguards to ensure regulatory actions do not infringe upon fundamental freedoms.

The initiatives that will be discussed in this chapter include the 2017 Joint Declaration on "Fake News, Disinformation, and Propaganda," the 2024 policy report "Fostering Media Freedom Literacy across the OSCE Region", the "Policy Manual: Spotlight on Artificial Intelligence and Freedom of Expression" (2022), the "Communiqué on Propaganda in Times of Conflict" (2014), the report "Beyond Fake News – Advancing Media and Information Literacy for an Informed Society" (2025), and the INFORMED project (2023–2028). The selection of these documents is based on their comprehensive coverage of both normative and operational dimensions relevant to countering the manipulation of information in a security context; Besides, they represent the main philosophy of action of the organization. They reflect the OSCE's long-standing normative commitment to freedom of expression while also showing the evolving policy tools addressing emerging digital threats.

2.1 Copenhagen Document (1990)

The Copenhagen document, adopted in 1990 by the participating states of what was then the Conference on Security and Cooperation in Europe (CSCE), now known as the OSCE, serves as a national normative framework. Its principles, even if crafted before the digital age, can still remain applicable to contemporary issues such as MDM and the information manipulation related to counterterrorism. Although the document does not directly tackle MDM nor mention it – understandable given that the phenomenon wasn't very well documented at the times since its

[460] OSCE Secretariat Transnational Threats Department, *INFORMED: Information and Media Literacy in Preventing Violent Extremism – Human rights-based and gender-sensitive approaches to addressing the digital information disorder*, Organization for Security and Co-operation in Europe Website, 2023, p. 1.

relevance came with the rise of social media and internet news – it still articulates democratic commitments that act as essential normative anchors for the organization.[461] Especially considering key focuses which include ensuring free and fair elections,[462] media freedom,[463] political plurality,[464] and accountable governance,[465] all of which are critical for protecting democratic societies from the disruptive impacts of information manipulation, whether in the context of counterterrorism or other. This document, ratified by the participating States, is a good starting point for analyzing the OSCE's contribution as it contains these principles, which all members agree upon.

Firstly, the Copenhagen Document's provisions on electoral integrity (both Articles 5.1 and Articles 7.1-7.6) highlights the importance of informed participation in democratic processes. The commitment to transparency in electoral procedures, equal opportunity for political contenders, and unrestricted access to public discourse is fundamentally linked to the quality of information accessible to citizens. By promoting strong electoral frameworks rooted in openness and fairness, the logic is that it helps to cultivate an environment that, while indirectly, effectively counters the damaging influence of false or misleading narratives, and especially those that aim to undermine democratic institutions or foster extremist ideologies. Article 5.4 is also worth mentioning as it establishes the principle of a clear division between the state and the media. Though not explicitly addressing MDM, this provision does, however, advocate for media independence as a structural safeguard against the concentration of narrative power and thereby prohibits state or non-state entities from monopolizing public discourse for propaganda Especially now with the digital sphere, a diverse and independent media landscape is quite important as a frontline defense against the widespread distribution of MDM. This is done, following the provisions of the Copenhagen document, by promoting competition among perspectives, supporting investigative journalism, and allowing for public scrutiny of both governmental and non-governmental information sources. Furthermore, Article 5.2 focuses on the need for accountable governance, meaning that public

[461] Conference on Security and Co-operation in Europe, *Document of the Copenhagen Meeting of the Conference on the Human Dimension of the CSCE*, Copenhagen, 1990, p. 2.
[462] *Ivi.*, Article 5(1) – 6, pp. 3 – 5.
[463] *Ivi.*, Article 26, p. 16.
[464] *Ivi.*, Article 3, p. 3.
[465] *Ivi.*, Article 5 (2), p. 3.

institutions must respond to citizens and operate under the rule of law: this provision shows that the ratifying states strive for a strengthened democratic system. This, consequently, doubles down on the fight against MDM, which flourishes in environments lacking transparency or accountability.

As such, by setting clear expectations for institutional openness and public oversight, the Copenhagen document fosters trust-building between state and society, which is an essential element in combating MDM, particularly when altered narratives attempt to manipulate public cynicism, fear, or alienation. However, like the other previous analyzed framework of international organizations, the Copenhagen document also contains an article about freedom of thought and its potential limitation, contained in Article 9 (4):

> «*Everyone will have the right to freedom of thought, conscience and religion […] The exercise of these rights may be subject only to such restrictions as are prescribed by law and are consistent with international standards.* »[466]

The ratifying states, along with the OSCE, thus recognize that this freedom is only limited by existing specific laws, which respect international standards as well. This goes with the line of thought shared by other international organizations. Hence, it seems that all international organizations so far share this acknowledgement of the freedom of thought while still recognizing that it can be limited by specific cases and without breaking any *ius cogens* or international standards. Besides, the very fact that all these provisions imply that there can't be any prescription without law is itself a respect of international criminal law's principle of *nullum crimem sine lege*.

While the framework seems to be shared between these international organizations as they contain similar principles, how does the OSCE contribute to the fight against MDM and terrorism?

[466] *Ivi.*, Article 9 (4), p. 8.

2.2 Representative on Freedom of the Media (RFOM)

In this framework provided by the OSCE, the Representative on Freedom of the Media (RFOM) may be one of its most important institutional player in navigating the complex interplay of counterterrorism efforts and freedom of expression. The RFOM's role is complex, as it involves monitoring violations of media freedoms,[467] supporting the establishment of laws to protect journalists,[468] and facilitating guidance for going after hate, all within the framework of protecting legitimate dissent. A key component of the RFOM's work also focuses on hate speech mitigation,[469] with an emphasis on educational programs and raising awareness designed to counteract intolerance, but not engaging in counter-productive censorship, as well. The RFOM is also active in advancing media plurality[470] by promoting the establishment and ongoing support of independent public service media to help counter misinformation. This includes consistently advocating for neutral editorial independence[471] and adequate resourcing for public broadcasters to allow them to provide facts and perspectives that counter MDM, especially with their frequent expert meetings.[472]

Several important initiatives highlight the OSCE's proactive approach to the link between MDM and counterterrorism. A notable document is the 2017 Joint Declaration on "Fake News, Disinformation, and Propaganda," co-authored by the RFOM: this declaration outlines the essential principles aimed at combating disinformation, while ensuring free expression. In fact, this is a prime example of the role of RFOM as an advisory body for national policy setting by

[467] OSCE Representative on Freedom of the Media, *Media freedom on the Internet*, Organization for Security and Co-operation in Europe Website, p. 1.

[468] OSCE Representative on Freedom of the Media, *Safety of Journalists*, Organization for Security and Co-operation in Europe Website, p. 1.

[469] OSCE Representative on Freedom of the Media, *Hate speech*, Organization for Security and Co-operation in Europe Website, p. 1.

[470] OSCE Representative on Freedom of the Media, *Media pluralism*, Organization for Security and Co-operation in Europe Website, p. 1.

[471] OSCE Chairpersonship et al., *Free and independent media are vital for strong democracies and our common security OSCE leaders say*, Organization for Security and Co-operation in Europe Website, 2025, p. 1.

[472] OSCE Representative on Freedom of the Media, *Sixth Expert Meeting: The role of public service media in countering disinformation*, Organization for Security and Co-operation in Europe Website, 2022, p. 1.

making recommendations to OSCE participating states so they can formulate their own policies on counter-MDM initiatives that do not violate international free speech legal standards.

It states that «disinformation and propaganda are often designed and implemented so as to mislead a population, as well as to interfere with the public's right to know and the right of individuals to seek and receive, as well as to impart, information and ideas of all kinds, regardless of frontiers, protected under international legal guarantees of the rights to freedom of expression and to hold opinions ».[473] Hence, this joint declaration explicitly considers that information manipulation in propagandistic use can actually hinder people's rights and freedoms: so, it's not only important to fight while trying not to respect these rights and freedoms, but countering MDM in the first place already helps to ensure them. Besides, it's not only the right to know, seek and receive, or freedom of expression, the declaration also notes that «some forms of disinformation and propaganda may harm individual reputations and privacy, or incite to violence, discrimination or hostility against identifiable groups in society ».[474]

Furthermore, what is appreciated in this joint declaration is the fact that disinformation is immediately defined, although incomplete, since it lacks other forms of MDM. However, what they don't forget to mention is that propaganda and disinformation can be from either state and non-state sources. For example, it explicitly condemns state-sponsored disinformation efforts and warns against broadly criminalizing false statements:[475] instead of the latter, it recommends a strategy that focuses on fostering independent media[476] and fact-checking services.[477] This document perfectly embraces what seems to be the OSCE's philosophy by stressing the importance of quality journalism and media outlets in the declaration's general principles, which truly reinforces the idea that the organization is committed to improving journalistic standards as a way to combat harmful narratives. Besides, this Joint Declaration represents a way to not only connect better with other international organizations, but to raise the flag on the issue so they can help against the spread of MDM. In this case, this document is shared with RFOM, the African

[473] UN Special Rapporteur on Freedom of Opinion and Expression et al., *Joint Declaration of Freedom of Expression and "Fake News", Disinformation and Propaganda*, Organization for Security and Co-operation in Europe Website, 2017, p. 1.
[474] *Ibidem*.
[475] *Ibidem*.
[476] *Ivi.*, p. 3 – 4.
[477] *Ivi.*, p. 4.

Commission on Human and Peoples' Rights, the Organization of American States, and the UNHR Office of the High Commissioner. As such, we can consider this document a comprehensive effort between all these international organizations, which also illustrates how OSCE, and more specifically the RFOM, operates.

Another important policy instrument in this regard is the "Fostering Media Freedom Literacy across the OSCE region",[478] which reinforces the organization's commitment to protecting media freedom within a broader counterterrorism framework. It was commissioned in 2024 by the RFOM in order to aid policymaking for a more effective way to counter the problem. Structurally, the report offers a comprehensive framework: it opens with an introduction and methodological approach, followed by a detailed overview of the legal and normative obligations stemming from key international organizations (notably the OSCE RFOM, the European Union, and the Council of Europe) before presenting a survey of relevant programmes and interventions. The report concludes with a series of policy recommendations alongside illustrative media literacy case studies.

What is interesting in this document is definitely the section detailing the obligations specific to each organization, in particular highlighting some RFOM documents such as the "Policy Manual: Spotlight on Artificial Intelligence and Freedom of Expression" (2022) and the "Communiqué on propaganda in times of conflict" (2014). The first document supporting digital literacy so that individuals themselves are more equipped with the proper knowledge and know-how for a responsible media consumption,[479] the latter stressing the importance of both freedom of media and its plurality as the best counters to propaganda and extremism.[480] As such, the recommendations this document proposes are: (1) adopt a "Cradle to Grave" approach to media literacy;[481] (2) maximize impact via a multi-stakeholder approach;[482] (3) empower citizens with

[478] CHAPMAN, ROKŠA-ZUBČEVIĆ, *Fostering Media Freedom Literacy across the OSCE region*, Office of the Organization for Security and Co-operation in Europe Representative on Freedom of the Media, 2024, pp. 1 – 82.
[479] *Ivi.*, p. 19.
[480] *Ibidem*.
[481] *Ivi.*, pp. 48 – 52.
[482] *Ivi.*, pp. 52 – 55.

timely and relevant information;[483] (4) insist on transparency from all stakeholders;[484] (5) prominence of public interest content;[485] (6) create an evidence base to develop effective interventions;[486] (7) guarantee availability and access to quality content;[487] (8) raise awareness and understand of media concentration and media pluralism.[488] While all the points are extremely valuable, the most intriguing one for the content of this thesis is point 5 "Prominence of Public Interest Content" which the document defines as «journalistic work and media content which deals with issues of interest and relevance to citizens and communities ».[489] In particular, the role of the different stakeholders with their different recommendations such as for media regulatory authorities to work with media platforms in order that the platform itself doesn't favor the flux of clickbait or harmful content such as misinformation and disinformation, and instead, prioritize fact-checked information.[490] Besides, the document also mentions that online platforms should work closely with fact-checking organizations to counter MDM content.[491]

All of these recommendations are in line with the approach of the OSCE, but also the UN with its Global Principles for Information Integrity and the EU with its 2022 Strengthened Code of Practice on Disinformation. This proves that, as analyzed so far in this thesis that, despite these international organizations not formally agreeing on a shared strategy, they seem to agree on similar strategies to counter MDM and terrorism while respecting human rights and freedoms.

2.3 Recent OSCE initiatives in the fight of MDM

Finally, while these projects are not the most recent, it doesn't mean that OSCE isn't active in this area. On the contrary, the 20th March 2025, a report was released called "Beyond Fake News - Advancing media and information literacy for an informed society", which was part of the OSCE

[483] *Ivi.*, pp. 55 – 57.
[484] *Ivi.*, pp. 57 – 58.
[485] *Ivi.*, pp. 58 – 61.
[486] *Ivi.*, pp. 61 – 63.
[487] *Ivi.*, pp. 63 – 64.
[488] *Ivi.*, pp. 64 – 66.
[489] *Ivi.*, p. 58.
[490] *Ivi.*, p. 60.
[491] *Ibidem*.

Mission in Kosovo. Not only is it one of the most recent reports at the time of writing this thesis, but it is exactly the theme of this thesis. In fact, this report contains a multitude of policy briefs related to Media and Information Literacy (MIL) but also MDM and other media challenges in the digital sphere. The choice of this project to be set in Kosovo probably stems from the fact that, like the report mentions, Kosovo is one of the lowest-ranking countries in media literacy in Europe.[492]

Gërguri, a lecturer expert on disinformation and political communication, in his policy brief contained in this official OSCE report, defines three different types of information manipulation: disinformation; misleading information; misinformation. However, the given definitions are different from what this thesis and many other sources define them as. According to Gërguri misleading information is «unintentional dissemination of incorrect information »[493] and misinformation is «the use of true information to cause harm ».[494] As a reminder, this thesis considers that the latter is a false information but with no malintent, and that malinformation is an information based on reality but manipulated or removed from its context with the intent to deceive and cause harm. Nevertheless, both this thesis and the policy brief by Gërguri share the same definition of disinformation, being «the deliberate spread of false information for manipulation ».[495]

With that being said, in the policy brief 6 “The role of artificial intelligence in increasing accuracy and speed in the fight against disinformation”, still authored by Gërguri, he notes that there are already some programs and platforms that are training AI to aid in the combat against MDM, with each their peculiarity.[496] ClaimBuster, for example, is to verify claims online,[497] while GPTzero is an AI detector for anything that is written.[498] These are only examples, as many more companies are working on using AI as a tool against MDM, and this proves that even if MDM techniques are changing with deepfakes, for example, the defense mechanism also may change by introducing the same tools. However, Gërguri reminds us of the current limitations of AI tools, being that the

[492] GËRGURI, *Strengthening Media Literacy for a More Resilient Society Against Information Disorder (Policy brief 1*), Organization for Security and Co-operation in Europe Website, 2025, p. 7.
[493] *Ibidem*.
[494] *Ivi.*, p. 8.
[495] *Ivi.*, p. 7.
[496] *Ivi.*, pp. 27 – 28.
[497] *Ibidem*.
[498] *Ibidem*.

algorithm’s training is still circumstantial based on the data brought to it:[499] the input can significantly limit the output. Besides, Gërguri also points out that the AI’s data is more trained in some languages than others, hence why the first recommendation, in the context of the mission in Kosovo, is a support for the Albanian language.[500]

While it is evident that the OSCE has made considerable efforts, all of the OSCE's work is still contextual in nature and therefore requires thoughtful scrutiny. What stands out as a strength of the OSCE is its ability to link security concerns as well as human rights concerns, like for example the gender-sensitive application of the INFORMED project (2023-2028)[501] and the RFOM's push towards a more balanced regulatory framework. In addition, the OSCE has played a significant role in setting norms in relation to global conversations on counter-MDM initiatives, and its documents, as seen with the 2017 Joint Declaration, which serve as an example of meaningful barriers to regulatory overreach in counterterrorism. Nevertheless, there could be implementation challenges, especially with respect to the inconsistent promotion of media literacy initiatives among the participating states. Additionally, the conflicts between counterterrorism intents and free speech rights still represent a significant challenge and a contradiction in those areas where restrictive media laws have been enacted that violate OSCE principled frameworks.

Overall, the OSCE's engagement with disinformation in the context of counter-terrorism policies demonstrates a complex relationship between policy development, capacity-building, and a rights-driven agenda. By utilizing its normative frameworks and institutional resources, the organization seeks to mitigate terrorist propaganda proliferation while strengthening democratic resilience. This dual mission, however, requires ongoing adjustment to address the evolving challenges posed by digital information chaos in a rapidly changing geopolitical environment.

[499] *Ibidem*.
[500] *Ibidem*.
[501] OSCE Secretariat, *INFORMED: Information and Media Literacy in Preventing Violent Extremism*, cit., p.1.

3. North Atlantic Treaty Organization (NATO)

3.1 What does counterterrorism and disinformation have to do with NATO?

NATO, which has primarily focused on collective defense and military deterrence, is increasingly aware of the evolving threats from terrorism and now also MDM. As will be analyzed later in this chapter, it recognizes the connection between these threats as a crucial challenge to the security of its allies: in today's digital age, extremist ideologies have gained greater visibility, allowing terrorist groups to use online platforms for radicalization, recruitment, and spreading propaganda.

As such, this new situation requires NATO to readjust its strategic focus, as outlined in the Strategic Concept which was adopted at the Madrid Summit on 29 June 2022, evolving the Alliance's commitment to three main tasks: being «deterrence and defence, crisis prevention and management, and cooperative security ».[502] Each of these areas is not only related to combating terrorism and countering MDM but crucial areas for them, thereby expanding NATO's role beyond conventional military operations.

This new Strategic Concept replaces the 2010 one in order to be able to respond to a drastically altered geopolitical landscape, particularly following Russia's full-scale invasion of Ukraine in February 2022.[503] This updated version reaffirms NATO's three core tasks (deterrence and defence, crisis prevention and management, and cooperative security). Not surprisingly, considering the history of NATO, it explicitly identifies the Russian Federation as the most significant and direct threat to Allied security.

However, for the first time, the Concept addresses China's systemic challenges to the rules-based international order and the security interests of the Alliance.[504] Most importantly, though, it integrates hybrid threats, cyber resilience, and the weaponization of energy as critical domains of

[502] NATO, *NATO 2022 Strategic Concept*, Madrid, 2022, p. 1.

[503] NATO Defense College Research Division, *NATO's New Strategic Concept*, edited by Tardy T., in «NDC Research Paper 25», NATO Defence College, 2022, pp. 1 – 2.

[504] NATO, *NATO 2022 Strategic Concept*, cit., p. 5.

concern.[505] Though not structured with a fixed expiration, the Strategic Concept is a guiding reference until replaced or revised by consensus.[506] Its implementation can be observed through NATO operations, doctrinal adjustments or even enhanced forward defence postures, particularly along NATO's eastern flank. As of 2024, regular progress has been monitored via the NATO Secretary General's Annual Reports and ministerial communiqués, affirming its operationalization across both military and political dimensions.

Terrorism is explicitly identified in the 2022 Strategic Concept as:

> «*Terrorism, in all its forms and manifestations, is the most direct asymmetric threat to the security of our citizens and to international peace and prosperity. Terrorist organisations seek to attack or inspire attacks against Allies. They have expanded their networks, enhanced their capabilities and invested in new technologies to improve their reach and lethality. Non-state armed groups, including transnational terrorist networks and state supported actors, continue to exploit conflict and weak governance to recruit, mobilise and expand their foothold.* »[507]

This perspective frames the evolving nature of terrorist threats as organizations expand their operational reach, adopt new technologies, and refine their use of information warfare strategies. MDM are crucial in this scenario since they enable terrorist groups to manipulate public sentiment, deepen societal rifts, and exploit vulnerabilities in democratic systems. The online dissemination of extremist content serves multiple objectives, such as radicalizing susceptible individuals, justifying terrorism through misleading information, and eroding public trust in national and global

[505] *Ivi.*, p. 3.
[506] NATO, *Consensus decision-making at NATO*, North Atlantic Treaty Organization Website, (updated) 2023, p. 1.
[507] NATO, *NATO 2022 Strategic Concept*, cit., p. 4.

entities. These elements not only heighten the terrorist threat but also exacerbate instability among NATO member states, presenting security challenges that require a unified and comprehensive response. As such, NATO's recognition of the strategic role of MDM in modern security threats is evident in its emphasis on hybrid warfare and the tactics employed by both state and non-state actors to undermine democratic resilience. The Strategic Concept explicitly warns that:

> «*Authoritarian actors challenge our interests, values and democratic way of life [...] They interfere in our democratic processes and institutions and target the security of our citizens through hybrid tactics, both directly and through proxies. They conduct malicious activities in cyberspace and space, promote disinformation campaigns, instrumentalise migration, manipulate energy supplies and employ economic coercion* ».[508]

This acknowledgment recognizes that disinformation is more than just propaganda; it poses a real security threat. Such disinformation includes intentional attempts to disrupt societies, sway public opinion, and undermine unity among transatlantic partners: and thus, we can say that the link between terrorism and disinformation is firmly within NATO's security responsibilities. This means that NATO must work together to build resilience, improve responses to misinformation, and assist member countries in dealing with the effects of information manipulation. By taking this approach, NATO not only upholds its traditional security roles but also evolves to meet contemporary threats, ensuring it stays relevant in a world where information has become a battleground.

To complement the analysis of NATO's actions in the fight against disinformation in the context of counterterrorism, it's important to look at other initiatives and organs such as the actions of Center of Excellence Defense Against Terrorism and the Strategic Communications Center of

[508] *Ivi.*, p. 3.

Excellence to complete the hybrid threats framework while also planning out the sphere of action of the organization. Nevertheless, the following parts will also focus on documents such as the 2023 “Joint Declaration on EU-NATO Cooperation” or the “Hybrid Threats and Hybrid Warfare Reference Curriculum” to capture the posture of the organization on the matter. Finally, the analysis draws on the contributions of experts such as James Pamment, whose research on information influence operations has informed NATO StratCom reports and related policy discussions.

3.2 NATO's Centre of Excellence

NATO's Centers of Excellence (COEs) are a key part of NATO's focus on research, training, and knowledge-sharing in various security and defence areas: COEs exist to provide NATO member states with expert analysis and strategic guidance.[509] Besides, they also act as hubs for specialized expertise that operate outside and independently of NATO's official command structure. However, only two COEs are pertinent to the content of this thesis, and hence, that concentrate on various security-related topics specifically address problems of terrorism and MDM: the Center of Excellence Defense Against Terrorism (COE-DAT) and the Strategic Communications Center of Excellence (StratCom COE). Their distinct roles demonstrate NATO's recognition that, in order to address the increasingly complex nature of hybrid warfare, contemporary security challenges call for not only military strength but also intellectual and analytical capabilities.

Located in Riga, Latvia, the StratCom COE has a critical function to help NATO advance its mission of understanding strategic communications,[510] specifically in relation to information warfare, including propaganda, and as we’ll see, also MDM. Its research agenda focuses on analyzing how state and non-state actors manipulate public perception through digital platforms with communication techniques, with a special focus on countering the influence of hostile

[509] NATO, *Centres of Excellence*, North Atlantic Treaty Organization Website, (updated) 2025, p. 1 – 2.
[510] *Ibidem*.

narratives which can threaten security or NATO activities and missions.[511] MDM campaigns, often designed to weaken trust in democratic institutions and sow discord within societies, are among the primary concerns addressed by the Centre. By studying the mechanisms through which adversaries exploit the information space, the StratCom COE provides NATO and its member states with the necessary tools and knowledge which aid them to develop effective countermeasures to enhance societal resilience.

Its publications, including "Social Media in Operation - A Counter-terrorism Perspective" (2017), examine the intersection of digital communications and terrorist recruitment strategies, shedding light on how extremist organizations utilize social media platforms to disseminate ideology, mobilize support, and coordinate operations.[512] But it doesn't stop with this insightful publication; in fact, there are more StratCom documents that were published over the years, which focus on different aspects of MDM. As early as December 2021, they already released two, rather analytic and technical, publications about AI and its role in both disseminating disinformation[513] but also how AI can be used as a tool against it;[514] subjects which were barely emerging at the time and still are contemporary to our times. Furthermore, they continue their research with an analysis of case studies of different countries and their overall framework and policies against the "infodemic" which started with COVID-19.[515] This document is noteworthy for its reports of tools and innovative solutions to increase societal resistance against the threat of MDM in general. However, another relevant publication by James Pamment, a lecturer and expert in hybrid threats and information influence operations who has done multiple reports for NATO StratCom and other important international organizations. His report for StratCom called "A Capability Definition and Assessment Framework for Countering Disinformation, Information Influence, and Foreign Interference" is an analysis for defining and assessing capabilities in order to contrast

[511] NATO Strategic Communications Center of Excellence, *About Strategic Communications*, NATO Strategic Communications Center of Excellence, p. 1.
[512] KANDEMIR et al., *Social Media in Operation - A Counter-terrorism Perspective,* Workshop report from NATO Centre of Excellence Defence Against Terrorism and NATO Strategic Communications Center of Excellence, 2017, pp. 21 – 22.
[513] JURŠĖNAS et al., *The Double-Edged Sword of AI: Enabler of Disinformation*, NATO Strategic Communications Center of Excellence, 2021, p. 4.
[514] JURŠĖNAS et al., *The Role of AI in the Battle Against Disinformation*, NATO Strategic Communications Center of Excellence, 2022, p. 4.
[515] HASSAIN, *Disinformation in Democracies: Improving Societal Resilience to Disinformation*, NATO Strategic Communications Center of Excellence, 2022, p. 4.

disinformation and its related phenomena.[516] In fact, the report does mention that «few have attempted to systematically define what those countermeasures are, and how they could be placed within a single, coherent capability assessment framework »[517] and so this research offers itself as literature to fill this gap. It proposes a "pragmatic toolset" for the capability assessment framework – constituted by Objectives, Indicators, Process maturity, and Risk assessment[518] – which enables Pamment to propose countering measures for different phenomena, notably counter-disinformation. This report is extremely valuable for policymakers in NATO countries to propose better and more realistic solutions to counter MDM. But the StratCom doesn't stop at publishing researches. In 2015, faced with a major disinformation incident, the High Representative for Foreign Affairs initiated the establishment of the East StratCom Task Force, while the European Commission complemented this effort by creating a High-Level Expert Group dedicated to addressing the spread of fake news.[519]

In addition to emphasizing the information aspect of security, the COE-DAT in Ankara, Turkey, is also contributing to NATO's counterterrorism objectives directly through research, educational programs, and developing doctrines.[520] The Centre is responsible for identifying the changing terrorist threats which has positioned the Centre to be a key asset in providing subject matter expertise related to terrorism and recommending effective practices for prevention, response, and mitigation. The Centre has theoretical and operational responsibilities in NATO Counter Terrorism, which allows NATO to be agile when confronting emerging terrorist threats.[521] The COE-DAT not only expands the level of collaboration amongst member nations, but they also offer advanced training programs,[522] which enhances the Alliance capabilities to counter all facets of terrorism.

[516] PAMMENT, *A Capability Definition and Assessment Framework for Countering Disinformation, Information Influence, and Foreign Interference*, NATO Strategic Communications Center of Excellence, 2022, pp. 4 – 31.

[517] *Ivi.*, p. 4.

[518] *Ivi.*, p. 13.

[519] NATO, *Hybrid Threats and Hybrid Warfare Reference Curriculum*, North Atlantic Treaty Organization Website, 2024, p. 54.

[520] NATO Center of Excellence Against Terrorism, *Functions and Activities*, NATO Center of Excellence Against Terrorism Website, (updated) 2023, p. 1.

[521] *Ibidem*.

[522] NATO Center of Excellence Against Terrorism, *Mobile Education and Training Activities*, NATO Center of Excellence Against Terrorism Website, (updated) 2023, p. 1.

The COE-DAT outlines several key activities on its official website: (1) helping NATO members and partners harmonize their efforts through standardization and combined initiatives to better anticipate, prevent, and respond to terrorism;[523] (2) promoting a comprehensive approach that incorporates the shared knowledge and capabilities of all contributing nations;[524] (3) providing expert advice on issues related to defence against terrorism and counterterrorism;[525] (4) supporting the development of NATO concepts, doctrines, and related documents in the field;[526] (5) offering education and training programs to both civilian and military personnel, from senior leaders to staff-level teams;[527] (6) sending mobile training teams to deliver specialized instruction where needed;[528] (7) assisting in the planning and execution of terrorism-related exercises and experiments, including the creation of realistic scenarios;[529] (8) managing a central repository of information, analysis, and lessons learned in close cooperation with the JALLC;[530] (9) organizing academic and professional events such as courses, seminars, conferences, and workshops;[531] (10) conducting specialized workshops and meetings focused on concept and doctrine development;[532] (11) participating actively in NATO's committees, boards, and working groups related to counterterrorism;[533] (12) carrying out evaluations to extract lessons learned from counterterrorism initiatives;[534] (13) engaging in academic research and related projects;[535] (14) producing and disseminating publications focused on defence against terrorism and counterterrorism topics:[536] (15) undertaking other specific tasks requested by the Steering Committee;[537] (16) and adapting its activities as necessary to meet emerging demands within the scope of its mandate.[538] The activities carried out by the COE-DAT are particularly relevant to understanding NATO's

[523] NATO Centre of Excellence Defence Against Terrorism, *Functions and Activities*, cit. p. 1.

[524] *Ibidem*.

[525] *Ibidem*.

[526] *Ibidem*.

[527] *Ibidem*.

[528] *Ibidem*.

[529] *Ibidem*.

[530] *Ibidem*.

[531] *Ibidem*.

[532] *Ibidem*.

[533] *Ibidem*.

[534] *Ibidem*.

[535] *Ibidem*.

[536] *Ibidem*.

[537] *Ibidem*.

[538] *Ibidem*.

multi-dimensional approach to counterterrorism. In particular, COE-DAT's missions related to standardization and interoperability (points 1–2), expertise provision (point 3), doctrine and concept development (points 4, 10), education and training (points 5–6, 9), and lessons learned analysis (points 8, 12) demonstrate the Alliance's emphasis on both operational readiness and strategic adaptation. In the centralization of knowledge production, support for coordinated responses, and human resource development at civilian and military levels, the COE-DAT embodies NATO's desire to approach terrorism not simply as a reactionary outcome but to establish longer-term resilience. Consequently, enumerating COE-DAT's functions is not only appropriate, it is also warranted, as it demonstrates the range of the NATO counter terrorism toolkit and the institutional depth mobilized in response to changing terrorist threats.

Finally, the Centres of Excellence illustrate NATO's commitment to tackling security challenges with a balance of foresight and execution. As efforts are undertaken to influence policy and operational considerations, they strengthen and support NATO's overall efforts to counter terrorism and the destabilizing effect of MDM. As the security environment continues to change, the data and observations produced by these collective organizations will remain assets as NATO faces the unprecedented and complex challenges of a modern world. The COEs make sure that NATO and its member states aren't just reacting to terrorism, hybrid threats and MDM, but also are able to proactively build the necessary knowledge, prepare forces, and create strategies to prevent and counter threats.

However, not all scholars were satisfied on the intensity of NATO's contribution to the fight against MDM as Kristine Berzina et al. state in their report (2019): «NATO's ambition in countering disinformation is relatively modest and the issue is largely portrayed as a regional rather than alliance-wide concern ».[539] Nevertheless, the situation has evolved since 2019, and NATO has developed its strategy over time, as will be discussed later in this chapter.

[539] BERZINA et al., *European Policy Blueprint for Countering Authoritarian Interference in Democracies*, German Marshall Fund of the United States, 2019, p. 44.

3.3 Counter Hybrid Threats Strategy

From what analyzed so far, NATO addresses MDM within the broader context of hybrid threats, which reflect the complex and varied nature of today's security issues. These hybrid threats, as a reminder of the EU definition, emerge from a mix of both conventional and unconventional methods, involving also military power with cyber tactics, economic pressure, and advanced information warfare. In fact, NATO also seems to have a similar working definition of hybrid warfare, being the following:

> «*Hybrid warfare is the creative use of hard, soft, and smart power by malign state or nonstate actors to achieve war-like objectives and political goals. Malign acts include a broad spectrum of military and nonmilitary instruments of coercive power beyond the conventionally conceived multidomain battlespace. Hybrid warfare encompasses politics, diplomacy, information, the economy, technology, the military and society, as well as dimensions like culture, psychology, legitimacy and morale. The coordinated performance of these malign acts occur both overtly and covertly in the ambiguous grey zones of blurred interfaces: between war and peace, friend and foe, internal and external relations, civil and military, and state and nonstate actors, as well as in fields of responsibilities generally below the threshold of war or as an accompaniment to more regular armed conflict.* »[540]

So, often executed by a mix by state and non-state of these threats aim to exploit societal weaknesses, damage democratic systems, and diminish faith in both national and international governance. MDM is a central component of this strategic playbook, making information warfare

[540] NATO, *Hybrid Threats and Hybrid Warfare Reference Curriculum*, 2024, p. 16.

an integral aspect of contemporary conflict. As such, NATO, much like other international organizations, has a multi-pronged approach with the aim to lessen the effects of MDM while enhancing the overall resilience and strengthening the international organization and its member states' capacity to deal with a more challenging information environment.

One of NATO's main efforts in this regard appears to be the continuing development of situational awareness, through its many COEs for example, which are crucial for identifying, tracking, and challenging MDM campaigns before they escalate. This would require continuous monitoring of the information environment using advanced technologies and systems that promote intelligence-sharing to identify coordinated influence operations. By studying patterns of information manipulation, NATO can proactively identify MDM threats, reducing their impact and safeguarding public trust. Recognizing and attributing these operations is critical, as it allows member states to formulate realistic and appropriate responses. This is also seen in the 2022 Strategic Concept which explicitly mentions that the organization and its members should «prepare for, deter, and defend against the coercive use of political, economic, energy, information and other hybrid tactics by state and non-state actors ».[541]

As such, beyond merely identifying MDM, NATO focuses additionally on the creation and sharing of counter-narratives aimed at disputing falsehoods. The goal extends beyond simply refuting MDM; as analyzed so far, it seems to seek to proactively influence the information environment to ensure that credible, evidence-based messages reach audiences ahead of adversarial narratives.

Achieving this shows a deep understanding of how various audiences process and interact with information, alongside the skill to communicate effectively across different societal and cultural backgrounds; this is why strategic communications become essential in bolstering NATO's credibility while countering harmful influences. Also, building resilience is one of guiding principles of NATO which recognizes that the best defense against hybrid threats[542] is to not only provide institutional defenses when they arise but to also enable societies to appraise the

[541] NATO, *NATO 2022 Strategic Concept*, cit., p. 7.
[542] NATO, *Resilience, civil preparedness and Article 3*, North Atlantic Treaty Organization, p. 1.

information they receive critically. This prevention model thus would work in addition with NATO's reactive elements to ensure any response to information threats are recognized as embedded within a greater culture of resilience rather than only based on crisis management.

But what exactly does NATO do concretely against hybrid threats? The document "Hybrid Threats and Hybrid Warfare Reference Curriculum" explains it quite well, but concisely the contribution of the organization is the following:

> «*NATO has invested in its ability to prepare for, deter and defend against the full spectrum of hybrid threats. It has expanded its tool box while recognising that primary responsibility for responding to hybrid attacks lies with the targeted nation. NATO has adopted an actor specific approach to countering hybrid threats by developing tailored comprehensive (civil and military) preventive and responses options for Allies to consider in countering specific threats. Tools include a deployable Counter Hybrid Support Team, consultations under Article 4 of the Washington Treaty and military activities all of which aim to pose strategic dilemmas for potential adversaries. NATO doctrine highlights that hybrid operations against the Alliance could reach the level of an armed attack and could lead to the invocation of Article 5 by the North Atlantic Council*. »[543]

Thus, as hybrid threats continue to evolve, NATO's approach to information warfare seems to remain dynamic, adapting to new challenges while reinforcing core principles of democratic resilience and collective security. NATO's aim is to defend the integrity of the information environment against intentional intrusions from individuals or organizations that seek to exploit the information for their strategic value through intelligence-driven situational awareness systems,

[543] NATO, *Hybrid Threats and Hybrid Warfare Reference Curriculum*, cit., p. 54.

proactive strategic communications, societal capacity building and international partnerships. In this process, the international organization reestablishes its position as a defender of not only territorial security but also the defense of the informational environment upon which democratic governance is contingent. But as will be seen in the next part, NATO does not work alone.

3.4 NATO-EU relationship

The strategic partnership between NATO and the EU is based on a shared commitment to ensuring security, stability and democratic values. They have a partnership that has matured into a willingness to collaborate in addressing multidimensional threats that require comprehensive and coordinated solutions. This includes terrorism and MDM, which pose significant challenges for open societies and democratic principles.

In fact, NATO explicitly mentions, in the “Hybrid Threats and Hybrid Warfare Reference Curriculum” that «NATO cooperates closely with partners and with the European Union ».[544] It is a common saying that communication is key in any relations, and political dialogue in a partnership is the same; here it enables both organizations to align strategic objectives, exchange perspectives on emerging security threats, and coordinate policy responses. As a matter of fact, NATO and the EU regularly consult and collaborate to strengthen resilience, increase situational awareness, and develop effective countermeasures against threats related to terrorism and hybrid threats.[545] Regular consultations at various levels ensure a flow of information is maintained and NATO's military and defense capacities are complementary to the extensive security and governance work of the EU.

Intelligence-sharing and the harmonization of policies across the EU’s 27 member states are especially relevant in the fight against terrorism as a coordinated, comprehensive response helps mitigate threats to our common citizens and disrupts extremist networks. In parallel, the exchange of information on MDM campaigns allows both organizations to identify and expose malign influence operations designed to manipulate public opinion and weaken democratic institutions.

[544] *Ibidem*.
[545] European Council, *EU-NATO cooperation*, cit., p. 1.

Besides, also the previously mentioned document from NATO recognizes that EU has similar aims: «the EU recognizes that responding to hybrid threats is a national issue but aims to support its partners and to coordinate actions with both member states and NATO. Its emphasis is on growing societal, economic, and political resiliency at the national and EU level; however, EU members still perceive that it lacks a top-level political commitment to responding seriously to these threats ».[546] This comment illustrates that while the EU has made important strides towards resilience-building, it seems that the EU member states identify gaps in strong political leadership in relation to hybrid threats and hybrid warfare. NATO's emphasis on this shortcoming probably is to prove of the complementary nature of the NATO-EU partnership: by providing operational support, strategic coordination, and deterrence capabilities, NATO helps to address the political and operational gaps left by the EU's approach, thereby reinforcing the collective security framework against hybrid challenges in the EU region.

There is a vast literature discussing NATO-EU partnership, such as "Disinformation Campaigns: Battling Misinformation for Resilience in Hybrid Threats Model" by William Steingartner et al., who are all scholars expert in cybersecurity and information security. Their research presents a dual perspective on the evolving security landscape, applying emphasis on corporate cybersecurity governance but also the strategic imperatives of EU-NATO cooperation;[547] in fact, they draw upon findings from the EEAS, which highlight the strategic necessity of strengthened EU-NATO cooperation in a volatile geopolitical context.[548] They also stress that, given the overlapping membership of EU and NATO states, a robust partnership is indispensable for safeguarding European security.[549] Both organizations are encouraged to leverage their complementary capabilities through structured collaboration in areas such as hybrid threats, cyberdefense, cybersecurity, and defense industry research.[550] This cooperation enhances resilience by improving information-sharing, synchronized planning, vulnerability disclosure, and joint responses to cybersecurity crises.[551] Additionally, it fosters responsible state behavior and contributes to

[546] NATO, *Hybrid Threats and Hybrid Warfare Reference Curriculum*, cit., p. 54.

[547] STEINGARTNER et al., *Disinformation Campaigns: Battling Misinformation for Resilience in Hybrid Threats Model*, Acta Polytechnica Hungarica, 2024, pp. 523 – 524.

[548] *Ibidem*.

[549] *Ivi.*, p.524.

[550] *Ibidem*.

[551] *Ibidem*.

sustaining cyber deterrence.[552] The EU's Strategic Compass further underlines the need to bolster collective strength by integrating cyber resilience measures – including supply chain risk management, diplomatic engagement, and civilian-military cooperation – with the EU-NATO partnership serving as a key component of this strategic framework.[553]

Apart from NATO-EU initiatives, similar activity by its member states and other institutions is important because it supports the conditions of security architecture. The Helsinki-based European Centre of Excellence for Countering Hybrid Threats, for example, reflects the overall resolve against hybrid warfare,[554] such as MDM.[555] While not a NATO body, the Centre serves as a cooperative arrangement for EU member states, NATO Allies, and partner countries to conduct research, share knowledge, and develop policy in the area of hybrid threats. Though not exclusively, the Centre's contribution is pertinent to the case of disinformation because it has the personnel, talent, and organizational capacity to provide research and analysis of how hostile actors have exploited the information environment, and to build and develop resilience in society against exploitation by hybrid threats.[556]

Recent developments in NATO-EU cooperation highlight the growing importance of this partnership to deal with hybrid threats. The 2018 Brussels Summit formalized joint NATO-EU action and made a commitment to greater unity in cybersecurity and countering disinformation, doubling down on the statement that «the European Union remains a unique and essential partner for NATO ».[557] In particular, NATO and the EU have developed tangible collaboration across several key domains, notably in countering hybrid threats, operational cooperation – including maritime security – cyberdefense, capability development, industrial research, joint exercises, and capacity-building initiatives.[558] Both organizations underlined the necessity of full mutual

[552] *Ibidem*.
[553] *Ibidem*.
[554] European Centre of Excellence for Countering Hybrid Threats, *Our work*, European Centre of Excellence for Countering Hybrid Threats Website, pp. 1 – 2.
[555] European Centre of Excellence for Countering Hybrid Threats, *Hybrid threats as a concept*, European Centre of Excellence for Countering Hybrid Threats Website, pp. 1 – 2.
[556] European Centre of Excellence for Countering Hybrid Threats, *Research and analysis*, European Centre of Excellence for Countering Hybrid Threats Website, pp. 1 – 2.
[557] NATO, *Brussels Summit Declaration*, 2018, p. 21.
[558] *Ibidem.*

openness, transparency, complementarity, and respect for their distinct mandates and decision-making autonomy. The implementation of the common set of 74 proposals was highlighted as a critical instrument to maintain coherence and probably to avoid unnecessary duplication of efforts.[559] NATO further welcomed the strengthening of European defense, noting that coherent, complementary, and consistent capabilities would reinforce the Alliance's collective security and contribute to transatlantic burden-sharing. Political dialogue between the organizations remains essential to advancing these objectives,[560] and NATO encouraged further steps to facilitate third-state participation in EU initiatives, where suitable.[561]

The 2023 Joint Declaration on EU-NATO Cooperation was envisioned in this same spirit and reiterated the need for enhanced partnership, especially regarding cybersecurity, hybrid threats,[562] and information warfare.[563] In making this commitment, both NATO and EU expressed the belief that it's better not to combat these threats alone, and thus, that a coordinated action will better serve their transatlantic community in defending security and democracy.

The growing importance of this NATO–EU cooperation is also confirmed by the literature, which confirms that this partnership is increasingly regarded as indispensable in addressing contemporary security threats, notably in the interconnected domains of counterterrorism and hybrid threats. Dominika Roslon, lecturer at the Lviv Polytechnic National University in the Department of Political Science and International Relations, with two students, published a research paper in 2023 discussing exactly this. They state that both organizations have demonstrated a clear political commitment towards deepening cooperation.[564] However, they also argue that both the EU and NATO have «for too long » fallen short in systematically analyzing developments in Russia and its neighboring post-Soviet states, leaving vulnerabilities that are readily exploited by Russian

[559] *Ibidem*.
[560] *Ibidem*.
[561] *Ibidem*.
[562] NATO, *Joint Declaration on EU-NATO Cooperation by the President of the European Council, the President of the European Commission, and the Secretary General of the North Atlantic Treaty Organization*, 2023, p. 1.
[563] *Ibidem*.
[564] ROSLON, KRUZHKOVA, SYVULKA, *EU and NATO Strategy to Counter and Prevent Russian Propaganda*, Scientific Journal Regional Studies, 2024, p. 101.

disinformation campaigns.[565] They argue that achieving long-term resilience requires extensive institutional reforms to strengthen stability in areas where Russian propaganda resonates.[566] While recognizing the ongoing discussion about how assertively to counter Moscow's disinformation, and considering the potential retaliation, they strongly support a more robust and visible strategy.[567]

To operationalize this stance, they propose a set of pragmatic policy measures such as expanding and adequately financing the EU's East StratCom Task Force and establishing a corresponding North StratCom unit to enhance analytical capacities in the Baltic region.[568] Another interesting suggestion is the creation of national disinformation working groups within both the EU and NATO, with the ambition of fostering cross-border knowledge exchange and harmonizing best practices.[569] Joint meetings between these institutional fora are viewed as a mechanism to deepen transatlantic coherence.[570]

Finally, as NATO and the EU navigate an increasingly complex security environment, their partnership continues to evolve, adapting to emerging threats and refining their joint response mechanisms. By integrating political dialogue, intelligence-sharing, capacity-building, and joint operational initiatives, both organizations contribute to a comprehensive security architecture capable of addressing the multidimensional challenges posed by terrorism and disinformation. In doing so, they reaffirm their role not only as defenders of territorial integrity but also as custodians of the democratic principles that underpin the stability of the Euro-Atlantic region.

[565] *Ivi.*, p. 104.
[566] *Ibidem*.
[567] *Ibidem*.
[568] *Ivi.*, pp. 104–105.
[569] *Ivi.*, p. 105.
[570] *Ibidem*.

4. Conclusion

This chapter provides a comparative analysis of the OSCE and NATO that demonstrates that both organizations make valuable but distinct contributions to the fight against MDM in the context of counterterrorism. They might share a foundational commitment to safeguarding security and democratic values, but their approaches reflect their respective mandates, institutional capacities, and normative orientations.

The OSCE's primary strength is its ability to connect security needs with a rights-based framework. Through initiatives led by the RFOM – including the 2017 Joint Declaration on "Fake News, Disinformation, and Propaganda," the INFORMED project, and various media literacy reports – the OSCE emphasizes media freedom, pluralism, and public resilience. As analyzed in this chapter, the focus on media literacy, public interest content, and the inclusive engagement of civil society reflects a normative strategy that positions independent journalism and societal awareness as essential counterweights to both disinformation and extremist narratives. Furthermore, the organization demonstrates a significant capacity to engage diverse stakeholders while grounding its interventions in human rights standards, a defining aspect of its contributions. However, there are a few problems. Firstly, the OSCE's interventions are inherently contextual and often require careful adaptation to diverse regional settings which complicates uniform implementation; such as reports focusing only on specific areas, like for example Kosovo policy brief. Secondly, its initiatives demand continual adjustment to keep pace with the rapidly evolving challenges posed by digital information disorder in a shifting geopolitical landscape. Especially since they don't have a proper main document setting out the framework for these subjects, as is the case for the EU with its Code of Conduct on Disinformation. Finally, while the OSCE acknowledges the intersection between disinformation and terrorism, its efforts predominantly focus on combating MDM in general terms, with comparatively fewer initiatives explicitly addressing the nexus between MDM and terrorism, aside from selected examples such as the INFORMED project.

Conversely, NATO employs a security-driven, operationally robust approach that embeds counter-disinformation efforts within its broader frameworks for addressing hybrid threats and counterterrorism. Its CoEs — notably StratCom COE and COE-DAT — deliver strategic research, doctrinal development, and practical training to bolster member states' resilience and situational awareness. NATO's Counter Hybrid Threats Strategy, along with its focus on intelligence-sharing, strategic communications, and proactive defense, underscores its capability to develop institutional and military capacities to tackle complex, multi-domain threats. Additionally, NATO's strong collaboration with the EU through the 2023 Joint Declaration enhances its ability for transatlantic burden-sharing and coordinated operational responses. Nevertheless, certain limitations, for NATO as well, must be acknowledged; its conceptual framing tends to situate MDM primarily within the broader category of hybrid threats, rather than recognizing disinformation as an autonomous and distinct security challenge in its own right. This perspective may risk underestimating the standalone dangers posed by MDM, particularly in contexts where disinformation operates independently of coordinated hybrid campaigns. Furthermore, NATO's focus on state-centric security strategies and military preparedness may underemphasize the protection of freedom of expression and media pluralism, especially in sensitive information environments.

However, overall, the OSCE contributes predominantly through normative frameworks and societal capacity-building, while NATO focuses in operational resilience and strategic deterrence.

Final Considerations

Summary: 1. Brief Summary of main points and answer to the thesis's research question. – 2. Comments and recommendations.

1. Brief Summary of main points and answer to the thesis's research question

This thesis has examined the extent to which international organizations contribute to addressing the complex and evolving challenge of online disinformation, misinformation, and malinformation (MDM) within counterterrorism frameworks. The research provided a comparative analysis of the UN, EU, NATO and the OSCE. It is clear that while each of these international organizations have each developed instruments and initiatives to counter MDM, their contributions remain varied in scope, depth, and operational capacity. The analysis has shown that international organizations are increasingly recognizing the threat of MDM to international and global security, particularly given its role in radicalization, propaganda, and operational coordination by terrorist actors. However, their methods and functions are not identical because of varying mandates, legal authorities and institutional structures.

While the UN itself does not have an enforcement mechanism, it is still a key actor in its coordinating and normative activities. From the UN Strategy and Plan of Action on Hate Speech, to the Verified Initiative to the Global Counter-Terrorism Strategy, the UN provides a global legal and political framework, focusing on prevention, human rights, and interagency cooperation. Still, its effectiveness is often weakened by its intergovernmental nature and the political will of Member States. The EU stands out for its normative and regulatory leadership. Instruments such as the Digital Services Act (DSA), the EU Internet Forum, and the Code of Practice on Disinformation represent an advanced framework for tackling terrorist content and platform accountability. The EU is, nonetheless, limited in some respects by its treaties and the reliance on Member State implementation, which affects uniform application. NATO, traditionally focused on collective defense, has made notable conceptual progress in recognizing the threat of MDM as part of hybrid warfare. Its Centers of Excellence and strategic doctrines integrate MDM as a component of emerging security threats. However, NATO's actions remain limited to coordination and awareness, given its lack of jurisdiction over internal legal frameworks or content moderation. While the OSCE is developing a rights-based response to MDM, particularly through the work of the Representative on Freedom of the Media, there are some limitations; the organization's consensus-driven decisions with a broad and diverse membership, its tools are mostly soft and non-binding, relying heavily on persuasion and dialogue.

In conclusion, this thesis finds that while all four organizations recognize MDM as a relevant factor in modern terrorism, their responses differ substantially in scope and emphasis. The UN offers normative leadership and multilateral platforms, the EU delivers binding regulations and operational tools, the OSCE strengthens rights-based resilience, and NATO ensures security-oriented preparedness. The research question of this thesis was "to what extent do international organizations contribute to addressing the challenge of online disinformation, misinformation, and malinformation in counterterrorism efforts?". Now, after all these analyses, it's possible to answer that international organizations provide essential normative, regulatory, and cooperative frameworks, but face significant operational, political, and jurisdictional constraints. While these organizations succeed in raising awareness, encouraging best practices, and fostering collaboration, the implementation gap, the voluntary nature of many measures, and divergences between member states' priorities hinder the full realization of their potential.

2. Comments and recommendations

This thesis has aimed to offer a comprehensive account of how international organizations engage with the growing phenomenon of MDM in the context of counterterrorism. But while this thesis tried to be as comprehensive as possible, it is still subject to the inherent limitations of a research project conducted within a limited timeframe and by a sole researcher. Thus, while it has focused on institutional frameworks and legal-political responses, several important avenues for further research deserve deeper scholarly attention:

First, future research should explore empirical assessments of effectiveness. While the current study primarily addresses normative frameworks and strategic instruments, the actual effects of these initiatives on diminishing terrorist propaganda and online radicalization are still under-explored. Gathering quantitative data on the amount of MDM content removed, the speed of takedowns, or the outreach of media literacy campaigns could greatly improve our comprehension of whether these actions yield significant results or remain merely symbolic. Second, it is essential to examine the long-term societal implications of MDM countermeasures, especially regarding civil liberties. Although this thesis highlights the conflict between security and freedom of expression, future research should perform specific legal and sociopolitical studies to evaluate whether and how certain counter-disinformation methods have undermined trust in institutions, restricted journalistic freedom, or caused unintended effects like suppressing dissent in vulnerable democracies. Third, upcoming studies could undertake comparative regional analyses beyond the Euro-Atlantic space. This thesis concentrated on the UN, EU, NATO, and OSCE due to their established relevance in European security. However, regional organizations in Africa, Latin America, or Southeast Asia might present different conceptual frameworks and counter-MDM strategies influenced by diverse political cultures, technological landscapes, and legal systems. Incorporating these viewpoints would enhance the global perspective on multilateral approaches to disinformation and terrorism. Fourth, the rise of generative AI, synthetic media, and algorithmic bias necessitates a reevaluation of current frameworks. Researchers should explore how international law can respond to the blending of truth and deception fostered by deepfakes and automated disinformation campaigns. Additionally, addressing the governance of AI-driven recommendation systems, which frequently promote radical content, is imperative in both academic and policy discussions. Lastly, there is an urgent need for further theoretical contemplation on the ontological status of MDM within international law. Unlike terrorism, which has dominated international legal discourse for decades, MDM remains conceptually scattered and often regarded as a subordinate concern. Future legal scholarship should investigate whether MDM could (or should) be considered an independent category of international threat and the implications this might have for doctrines such as sovereignty, intervention, and state accountability.

Bibliography

AALTOLA, Mika, *Democracy's Eleventh Hour: Safeguarding Democratic Elections Against Cyber-Enabled Autocratic Meddling*, in «Fiia Briefing Paper(226)», The Finnish Institute of International Affairs, 2017, pp. 3 – 4.

Advokatfirman Segerström, *Legal advice – The Swedish Care of Young Persons Act (LVU*), Advokatsegerstrom.

AMMAR, Jamil, *Disinformation: The Jihadists' New Religion*, edited by ARCOS R., CHIRU I., IVAN C., *Routledge Handbook of Disinformation and National Security*, Routledge, Abingdon, 17 November 2023, p. 112–116–118.

ARCOS, Rubén, ARRIBAS, Cristina M., *Anticipatory Approaches to Disinformation, Warning and Supporting Technologies*, edited by ARCOS R., CHIRU I., IVAN C., *Routledge Handbook of Disinformation and National Security*, Routledge, Abingdon, 17 November 2023, p. 401.

ARQUILLA, John et al., *Russian Strategic Intentions: A Strategic Multilayer Assessment (SMA) White Paper*, in «SMA White Papers», NSI Inc., May 2019, pp. 70 – 71.

Article 19, *Social Media 4 Peace: A handbook to support freedom of expression*, Article 19 Website, 24 August 2023, p. 1.

ASMOLOV, Gregory, *The Effects of Participatory Propaganda: From Socialization to Internalization of Conflicts*, in «Journal of Design and Science», issue 6, 7 August 2019, pp. 6–15.

BAYER, Judith et al., *The fight against disinformation and the right to freedom of expression*, Think Tank European Parliament, 5 July 2021, pp. 19–20.

BERZINA, Kristine et al., *European Policy Blueprint for Countering Authoritarian Interference in Democracies*, German Marshall Fund of the United States, no. 18, Washington, 25 June 2019, p.44.

BJOLA, Corneliu, MANOR Ilan, *The Use and Abuse of History by Russian Embassies on Twitter*, edited by ARCOS R., CHIRU I., IVAN C., *Routledge Handbook of Disinformation and National Security*, Routledge, Abingdon, 17 November 2023, pp. 154 – 157.

BONTCHEVA, Kalina et al., *Balancing Act: Countering Digital Disinformation While Respecting Freedom of Expression*, United Nations Educational, Scientific and Cultural Organization, Paris, 2020, p. 25.

BORISOVICH-ORISHEV, Aleksandr et al., *The UN role in combating international terrorism: achievements and challenges*, in «Revista de investigaciones Universidad del Quindio», vol. 34, no. S2, 22 August 2022, p. 269.

CASSESE, Antonio, *Terrorism is also disrupting some crucial legal categories of international law*, in «European journal of international law », vol. 12, no. 5, 2001, p. 995.

CASSESE, Antonio, *The Multifaceted Criminal Notion of Terrorism in International Law*, in «Journal of International Criminal Justice», vol. 4, no. 5, 2006, p. 957.

CHAPMAN, Martina, ROKŠA-ZUBČEVIĆ, Asja, *Fostering Media Freedom Literacy across the OSCE region*, Office of the Organization for Security and Co-operation in Europe Representative on Freedom of the Media, April 2024, pp. 1 – 82.

CHESNEY, Robert, CITRON, Danielle K., *Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security*, in «California Law Review», University of California Berkeley School of Law, vol. 107, no. 6, 2019, p. 1785.

CHILUWA, Innocent, *The Discourse of Terror Threats: Assessing Online Written Threats by Nigerian Terrorist Groups*, in «Studies in Conflict and Terrorism», Taylor & Francis Online, vol. 40, issue 4, 2016, pp. 318 – 338.

CHIRU, Irena, BULUC, Ruxandra, *An Ethical Understanding of Military Strategic Communication, Public Relations, And Persuasion*, edited by ARCOS R., CHIRU I., IVAN C., *Routledge Handbook of Disinformation and National Security*, Routledge, Abingdon, 17 November 2023, p. 264.

CIALDINI, Robert B., *Influence: The Psychology of Persuasion (Revised Edition)*, Collins Business, New York, 2006, pp. 115 – 116.

COATS, Dan, *Transcript: Dan Coats Warns The Lights Are 'Blinking Red' On Russian Cyberattacks*, NPR, 18 July 2018.

Conference on Security and Co-operation in Europe, *Document of the Copenhagen Meeting of the Conference on the Human Dimension of the CSCE*, Copenhagen, 29 June 1990, pp.2–16.

CONROY, Meghan and AL-DAYEL, Nadia, *Identity Construction through Discourse: A Case Study of ISIS's No Respite Video*, in «Studies in Conflict and Terrorism», Taylor & Francis Online, vol. 45, issue 12, 7 April 2020, p. 13.

CORBU, Nicoleta et al., *Fake news going viral: The mediating effect of negative emotions*, in «Media Literacy and Academic Research», Univerzita sv. Cyrila a Metoda v Trnave, Fakulta masmediálnej komunikácie, issue 2, 2021, pp. 58–87.

Council of Europe Committee of Ministers, *Guidelines on Human Rights and the Fight Against Terrorism*, in «Policy & Guidance», Strasbourg, July 2002, Preamble(a), p. 3.

Council of Europe, *A Convention to protect your rights and liberties*, Council of Europe Official Website.

Council of the European Union, *EU measures to prevent radicalisation*, *Consilium Europa*, 2022, p. 1.

Council of the European Union, *Strategic Compass for Security and Defence*, Brussels, 21 March 2022, pp.5–53.

Council of the European Union, *The European Union Counter-Terrorism Strategy*, *Consilium Europa,* 30 November 2005, pp. 3–6.

DE LONDRAS, Fiona, *Terrorism as an international crime*, edited by SCHABAS, W.A., BERNAZ, N., *Routledge Handbook of International Criminal Law*, Routledge, Oxfordshire, 2010, pp. 175 – 176.

European Centre of Excellence for Countering Hybrid Threats, *Hybrid threats as a concept*, European Centre of Excellence for Countering Hybrid Threats Website, pp. 1 – 2.

European Centre of Excellence for Countering Hybrid Threats, *Our work*, European Centre of Excellence for Countering Hybrid Threats Website, pp. 1 – 2.

European Centre of Excellence for Countering Hybrid Threats, *Research and analysis*, European Centre of Excellence for Countering Hybrid Threats Website, pp. 1 – 2.

European Commission, *A Counter-Terrorism Agenda for the EU: Anticipate, Prevent, Protect, Respond (COM 2020/795)*, EUR-Lex, 9 December 2020, p. 3.

European Commission, *Commission sends preliminary findings to X for breach of the Digital Services Act*, European Commission Website, Brussels, 12 July 2024, pp. 1–2.

European Commission, *Communication from the Commission to the European Parliament and the Council on the Seventh Progress Report on the implementation of the EU Security Union Strategy*, 15 May 2024, pp. 4–9.

European Commission, *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions Supporting the Prevention of Radicalisation Leading to Violent Extremism* (COM/2016), Eur-Lex, 17 June 2016, p. 2–14.

European Commission, *Disinformation: EU assesses the Code of Practice and publishes platform reports on coronavirus related disinformation*, European Commission Website, 10 September 2020, pp. 1–2.

European Commission, *Code of Practice on Disinformation*, European Commission Website, 26 September 2018, p. 1.

European Commission, *EU Security Union Strategy: connecting the dots in a new security Ecosystem*, European Commission Website, 24 July 2020, pp 1–9.

European Commission, *European Commission Guidance on Strengthening the Code of Practice on Disinformation*, in *Communication to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions*, Brussels, 26 May 2021, pp. 1 – 24.

European Commission, *Fighting Terrorism Online: Internet Forum pushes for automatic detection of terrorist propaganda*, European Commission Website, 6 December 2017, pp. 1–2.

European Commission, *Guidance on Strengthening the Code of Practice on Disinformation*, *Policy and Legislation*, European Commission Website, 26 May 2021, p. 1.

European Commission, *Tackling coronavirus disinformation*, in «Strategy and Policy», European Commission Website, 2020.

European Commission, *Tackling online disinformation*, in «Policies», Digital Strategy, 2024, pp. 1–2.

European Commission, *The 2022 Code of Practice on Disinformation*, European Commission Website, updated on 13 February 2025, pp. 1–4.

European Commission, *The Code of Conduct on Disinformation*, European Commission Website, 13 February 2025, p.1.

European Commission, *The Strengthened Code of Practice on Disinformation*, 16 June 2022, pp. 1–40.

European Commission: Directorate-General for Communications Networks, Content and Technology, *A multi-dimensional approach to disinformation: report of the independent High level Group on fake news and online disinformation*, Publications Office of the European Union, 2018, p. 3.

European Commission's Directorate-General for Defence Industry and Space, *Hybrid Threats*, Directorate-General for Defence Industry and Space, p. 1.

European Council, *EU-NATO cooperation*, Consilium Europa, updated in 2024, p. 1.

European e-Justice, *Protecting fundamental rights within the European Union*, European e-Justice Portal, p. 2.

European External Action Service, *Annual Progress Report on the Implementation of the Strategic Compass for Security and Defence*, European Union External Action, March 2024, p. 8.

European Union, *EU countries*, European Union Official Website, p. 1.

European Union, *The Schuman Declaration*, European Union Official Website, 1950, p. 1

EUvsDisinfo, *'To Challenge Russia's Ongoing Disinformation Campaigns': Eight Years of EUvsDisinfo*, in «*News and Analysis*», EUvsDisinfo, 5 July 2023, p. 1.

FINO, Audrey, *A critique of the UN Strategy and Guidance on 'Hate Speech': Some Legal Considerations*, edited by FORTIN K. et al., *Netherlands Quarterly of Human Rights*, Sage Journals, vol. 41, issue 4, 2023, p. 191–207.

Foreign, Commonwealth & Development Office, *Promoting interreligious and intercultural dialogue and tolerance in countering hate speech: UK statement at the UN General Assembly*, Government of UK official website, 26 July 2023.

FRAU-MEIGS, Divina, *Media and information literacy*, in «Background paper prepared for the 2023 Global education monitoring report: Technology in education», UNESCODOC Digital Library, 2023, p. 4.

FRY, James D., *Terrorism as a Crime Against Humanity and Genocide: The Backdoor to Universal Jurisdiction*, in «UCLA Journal of International Law and Foreign Affairs», vol. 7, no. 1, 2002, pp. 198 – 198.

GËRGURI, Dren, *Strengthening Media Literacy for a More Resilient Society Against Information Disorder (Policy brief 1*), in «Beyond Fake News – Advancing media and information literacy for an informed society», Organization for Security and Co-operation in Europe Mission in Kosovo report, 20 March 2025, pp. 7–28.

GIANDOMENICO, Jessica et al., *Disinformation landscape in Sweden (V2)*, EUDisinfoLab, 2 April 2025, pp. 3– 4.

GIANNOPOULOS, Georgios, SMITH, Hanna, THEOCHARIDOU, Marianthi, *The Landscape of Hybrid Threats: A conceptual model*, Publications Office of the European Union, Luxembourg, 2021, p. 32.

GLOBSEC, *Policy Brief: United Nations Code of Conduct for Information Integrity on Digital Platforms*, GLOBSEC Centre for Democracy & Resilience, 18 December 2023, pp. 2–6.

GUTERRES, António, *Foreword by the Secretary-General United Nations Strategy and Plan of Action on Hate Speech*, in «United Nations Strategy and Plan of Action on Hate Speech: Detailed Guidance on Implementation for United Nations Field Presences», United Nations Digital Library, September 2020, p.3.

GUTERRES, António*, Information Integrity on Digital Platforms*, in «Our Common Agenda», Policy brief 8, June 2023, p. 5–16.

HASSAIN, Jon, *Disinformation in Democracies: Improving Societal Resilience to Disinformation*, North Atlantic Treaty Organization Strategic Communications Center of Excellence, Riga, 26 April 2022, p. 4.

HATTOTUWA, Sanjana et al., *High-Level Panel on Digital Cooperation: Reflections and Recommendations from the ICT4Peace Foundation*, ICT4Peace Publishing, Geneva, 30 November 2018, p. 14.

HOSAKA, Sanshiro, *Cold War Active Measure*s, edited by ARCOS R., CHIRU I., IVAN C., *Routledge Handbook of Disinformation and National Security*, Routledge, Abingdon, 17 November 2023, p. 45.

ICT4Peace, *More support needed for smaller technology platforms to counter terrorist content*, in «Policy Research», ICT4Peace Foundation, 25 November 2018, p. 1.

ICT4Peace, *UN and ICT4Peace engage with private sector on responding to terrorist use of ICT*, in «Activities», ICT4Peace Foundation, 12 April 2016, p. 1.

ICT4Peace, *UN Security Council recognises ICT4Peace's work with the United Nations*, in «Activities», ICT4Peace Foundation, 8 January 2018, p. 1.

INGELEVIČ-CITAK, Milena, PRZYSZLAK, Zuzanna, *Jihadist, Far-Right And Far-Left Terrorism in Cyberspace –Same Threat and Same Countermeasures?*, in «International Comparative Jurisprudence», Mykolas Romeris University, vol. 18, no. 2, December 2020, p. 159.

IVAN, Cristina, *Protective Factors Against Disinformation*, edited by ARCOS R., CHIRU I., IVAN C., *Routledge Handbook of Disinformation and National Security*, Routledge, Abingdon, 17 November 2023, p. 295.

JURŠĖNAS, Alfonsas et al., *The Double-Edged Sword of AI: Enabler of Disinformation*, North Atlantic Treaty Organization Strategic Communications Center of Excellence, Riga, 21 December 2021, p.4.

JURŠĖNAS, Alfonsas et al., *The Role of AI in the Battle Against Disinformation*, North Atlantic Treaty Organization Strategic Communications Center of Excellence, Riga, 28 February 2022, p.4.

KANDEMIR, Berfin et al., *Social Media in Operation - A Counter-terrorism Perspective,* Workshop report from NATO Centre of Excellence Defence Against Terrorism and NATO Strategic Communications Center of Excellence, 27 September 2017, pp. 21–22.

KHAN, Irene, *Disinformation and freedom of opinion and expression (A/HRC/47/25)*, in «Report of the Special Procedure of the Human Rights Council United Nations», United Nations Digital Library, 13 April 2021, p. 4.

KING, Stephen, O'KEEFFEE, Eoghan, ENGLISH, Rosalyn, *EU's Code of Conduct on Disinformation Integrated into DSA, Tech Law Blog*, 24 March 2025, p. 1.

KIVIMÄKI, Veli P*., Open-Source Information for Intelligence Purposes: the Challenge of Disinformation*, edited by ARCOS R., CHIRU I., IVAN C., *Routledge Handbook of Disinformation and National Security*, Routledge, Abingdon, 17 November 2023, p. 283–287.

KOUFA, Kalliopi K., *Terrorism and Human Rights: Progress Report by Special Rapporteur Kalliopi K. Koufa*, in «Report of the Special Rapporteur on the Promotion and Protection of Human Rights while Countering Terrorism», United Nations Digital Library, 2001, p. 11.

KRZAK, Andrzej, *Operational Disinformation of Soviet Counterintelligence during the Cold War*, in «International Journal of Intelligence and CounterIntelligence», Routledge, vol. 35, issue 2, 2022, pp. 2 – 3.

KUMAR, Abhijeet, *UN unveils 'Global Principles' to combat online misinformation, hate speech*, in «World News», Business Standard, 25 June 2024, p. 1.

LAKOMY, Miron, *Recruitment and Incitement to Violence in the Islamic State's Online Propaganda: Comparative Analysis of Dabiq and Rumiyah*, in «Studies in Conflict and Terrorism», Taylor & Francis Online, vol. 44, issue 7, 7 February 2019, p.577.

LUDDEN, Vanessa et al., *Evaluation of impact and effectiveness of counter- and alternative campaigns stemming from the CSEP programme aiming at preventing radicalisation leading to violent extremism and terrorism (Final Report)*, Publications Office of the European Union, Luxembourg, October 2022, pp. 152–179.

LUKITO, Josephine, *Digital Disinformation, Electoral Interference and Systemic Distrust*, edited by ARCOS R., CHIRU I., IVAN C., *Routledge Handbook of Disinformation and National Security*, Routledge, Abingdon, 17 November 2023, p.123.

Medicine Lodge Cresset. *Issue of 13th February 1887*, Medicine Lodge, 1887, p. 3.

MEHRAN, Weeda et al., *Deep Analysis of Taliban Videos: Differential Use of Multimodal, Visual and Sonic Forms across Strategic Themes*, in «Studies in Conflict and Terrorism», Taylor & Francis Online, 11 January 2021, pp. 1–21.

MILTON, Daniel, *Truth and lies in the Caliphate: The use of deception in Islamic State Propaganda*, edited by MALTBY, S. et al, *Media, War & Conflict*, vol. 15, no. 2, Sage Journals, London, 2020, p. 231.

MOHAMMED, Amina J., *Strengthening Regional Cooperation and Institution Building to Address the Evolving Threat of Terrorism in Africa*, in «United Nations Statements», Abuja, 2024, p. 1.

NAUTA, Myrthe, *You can't fact check disinformation away. What can you do?*, in «Access to Independent Information», Free Press Unlimited, 23 July 2021, pp. 1–2.

NENADIĆ, Iva, *EC's Guidance to Strengthening the Code of Practice on Disinformation: A Mis-take with Mis-information?*, in «Discussion series on Media Policy and Journalism», Centre for Media Pluralism and Media Freedom, 22 October 2021, pp. 1 – 2.

NÍ AOLÁIN, Fionnuala, *Visit to Kazakhstan: Report of the Special Rapporteur on the promotion and protection of human rights and fundamental freedoms while countering terrorism*, in «Report of the Special Rapporteur on the Promotion and Protection of Human Rights while Countering Terrorism», UN Digital Library, 22 January 2020, p. 9.

NIMMO, Ben et al., *Secondary Infektion*, in «Graphika Report», Graphika, 2020, p. 4.

North Atlantic Treaty Organization Center of Excellence Against Terrorism, *Functions and Activities*, North Atlantic Treaty Organization Center of Excellence Against Terrorism Website, updated in 2023, p.1.

North Atlantic Treaty Organization Center of Excellence Against Terrorism, *Mobile Education and Training Activities*, North Atlantic Treaty Organization Center of Excellence Against Terrorism Website, updated in 2023, p.1

North Atlantic Treaty Organization Defense College Research Division, *NATO's New Strategic Concept*, edited by Tardy T., in «NDC Research Paper 25 », NATO Defence College, 14 September 2022, pp. 1 – 2.

North Atlantic Treaty Organization Strategic Communications Center of Excellence, *About Strategic Communications*, NATO Strategic Communications Center of Excellence, p.1.

North Atlantic Treaty Organization, *Brussels Summit Declaration*, issued by the Heads of State and Government participating in the meeting of the North Atlantic Council, Brussels, 11 July 2018, p. 21.

North Atlantic Treaty Organization, *Centres of Excellence*, North Atlantic Treaty Organization Website, updated in 2025, p. 1–2.

North Atlantic Treaty Organization, *Consensus decision-making at NATO*, North Atlantic Treaty Organization Website, updated in 2023, p. 1.

North Atlantic Treaty Organization, *Hybrid Threats and Hybrid Warfare Reference Curriculum*, North Atlantic Treaty Organization Website, June 2024, p.54.

North Atlantic Treaty Organization, *Hybrid Threats and Hybrid Warfare Reference Curriculum*, North Atlantic Treaty Organization Headquarters, Brussels, June 2024, pp. 16–54.

North Atlantic Treaty Organization, *Joint Declaration on EU-NATO Cooperation by the President of the European Council, the President of the European Commission, and the Secretary General of the North Atlantic Treaty Organization*, Brussels, 10 January 2023, p. 1.

North Atlantic Treaty Organization, *NATO 2022 Strategic Concept*, adopted by Heads of State and Government at the NATO Summit, Madrid, 29 June 2022, pp. 1–7.

North Atlantic Treaty Organization, *Resilience, civil preparedness and Article 3*, North Atlantic Treaty Organization Center of Excellence Against Terrorism Website, p. 1.

Ó FATHAIGH, Ronan, *Brzeziński v. Poland: Fine over 'false' information during election campaign violated Article 10*, *Strasbourg Observers*, 8 August 2019, pp. 2 – 3.

Organization for Security and Co-operation in Europe Chairpersonship et al., *Free and independent media are vital for strong democracies and our common security OSCE leaders say*, Organization for Security and Co-operation in Europe Website, 17 March 2025, p. 1.

Organization for Security and Co-operation in Europe Representative on Freedom of the Media, *Media freedom on the Internet*, Organization for Security and Co-operation in Europe Website, p. 1.

Organization for Security and Co-operation in Europe Representative on Freedom of the Media, *Safety of Journalists*, Organization for Security and Co-operation in Europe Website, p. 1.

Organization for Security and Co-operation in Europe Representative on Freedom of the Media, *Hate speech*, Organization for Security and Co-operation in Europe Website, p. 1.

Organization for Security and Co-operation in Europe Representative on Freedom of the Media, *Media pluralism*, Organization for Security and Co-operation in Europe Website, p. 1.

Organization for Security and Co-operation in Europe Representative on Freedom of the Media, *Sixth Expert Meeting: The role of public service media in countering disinformation*, Organization for Security and Co-operation in Europe Website, 2022, p. 1.

Organization for Security and Co-operation in Europe Secretariat Transnational Threats Department, *Preventing and countering violent extremism online through media and information literacy focus of OSCE training programme*, Organization for Security and Co-operation in Europe Website, 2 November 2023, p. 1.

Organization for Security and Co-operation in Europe Secretariat Transnational Threats Department, *INFORMED: Information and Media Literacy in Preventing Violent Extremism – Human rights-based and gender-sensitive approaches to addressing the digital information disorder*, Organization for Security and Co-operation in Europe Website, 2023, p. 1.

Organization for Security and Co-operation in Europe, *The European Union*, Organization for Security and Co-operation in Europe Website, p. 1.

ORSINI, Alessandro, *La radicalisation des terroristes de vocation*, in « Commentaire », n°156, 2016, pp. 783-790.

ORSINI, Alessandro, *Poverty, ideology and terrorism: The STAM bond*, in «Studies in Conflict and Terrorism», 2012, vol. 35, issue 10, pp. 665-692.

Oxford University Press, *Misinformation*, Oxford English Dictionary, 2024.

PALMERTZ, Björn, PAMMENT James, *Asymmetrical Conflict in the Information Domain—The Case Of Russia*, edited by ARCOS R., CHIRU I., IVAN C., *Routledge Handbook of Disinformation and National Security*, Routledge, Abingdon, 17 November 2023, p. 95.

PAMMENT James, *A Capability Definition and Assessment Framework for Countering Disinformation*, *Information Influence, and Foreign Interference*, North Atlantic Treaty Organization Strategic Communications Center of Excellence, Riga, 5 December 2022, pp. 4–31.

PHERSON, Randolph H., LABRINY, Deanna, DIORIO, Abby, *Historical Disinformation Practices: Learning from the Russians*, edited by ARCOS R., CHIRU I., IVAN C., *Routledge Handbook of Disinformation and National Security*, Routledge, Abingdon, 17 November 2023, p. 60–78.

REINHOLD, Steven, *Good Faith in International Law*, UCL Journal of Law and Jurisprudence, vol. 2, issue 1, 24 May 2013, pp. 40–63.

RICHARDS, Julian, *The Use of Discourse Analysis in Propaganda Detection and Understanding*, edited by ARCOS R., CHIRU I., IVAN C., *Routledge Handbook of Disinformation and National Security*, Routledge, Abingdon, 17 November 2023, p. 387–391.

RID, Thomas, *Active Measures: The Secret History of Disinformation and Political Warfare*, Profile Books, New York, 2020, p. 7.

ROSLON, D.T., KRUZHKOVA, E. M., SYVULKA, U. I.*, EU and NATO Strategy to Counter and Prevent Russian Propaganda*, in «Scientific Journal Regional Studies», Uzhhorod National University, issue 36, Uzhhorod, 2024, pp. 101–105.

RUSSELL, Stuart J. and NORVIG, Peter, *Artificial Intelligence, A Modern Approach (Third Edition*), Pearson, London, 1 January 2014, p. 694.

SALING, Lauren L. et al., *No One Is Immune to Misinformation: An Investigation of Misinformation Sharing by Subscribers to a Fact-Checking Newsletter*, PloS One, vol. 16, no. 8, 10 August 2021, pp. 1–12.

SAUL, Ben, *Defining Terrorism in International Law*, in «GlobaLex», New York School of Law, December 2021, pp. 3–4.

SAUL, Ben, *Defining Terrorism in International Law*, in «Oxford monographs in international law», Oxford University Press, Oxford, 2006, p. 1 – 232.

SAUL, Ben, *Terrorism in Customary International Law*, in «Oxford Monographs in International Law», Oxford University Press, Oxford, 2008, pp. 251–253.

SHATTOCK, Ethan, *Should the ECtHR Invoke Article 17 for Disinformation Cases?*, *EJIL:Talk!* (Blog of the European Journal of International Law), 26 March 2021, pp. 4 – 5.

STEINGARTNER, William et al., *Disinformation Campaigns: Battling Misinformation for Resilience in Hybrid Threats Model*, Acta Polytechnica Hungarica, vol. 21, no. 10, January 2024, pp. 523 – 524.

Strategic Communications, *Information Integrity and Countering Foreign Information Manipulation & Interference (FIMI)*, European Union External Action Website, 14 March 2025.

Tech against Terrorism, *Position Paper: Content Personalisation and the online dissemination of terrorist and violent extremist content*, Tech against Terrorism, 17 February 2021, p. 1.

The Guardian, *Rival Disinformation Campaigns Targeted African Users, Facebook Says*, in «News Central African Republic + Africa», The Guardian, 15 December 2020.

Thomson Reuters, *Direct applicability (EU)*, Thomson Reuters Practical Law, p. 1.

TUDORACHE, Adrian, *A Perception Management Take on Propaganda as Political Warfare*, edited by ARCOS R., CHIRU I., IVAN C., *Routledge Handbook of Disinformation and National Security*, Routledge, Abingdon, 17 November 2023, p.135.

United Nations Department of Global Communications, *'Verified' initiative aims to flood digital space with facts amid COVID-19 crisis*, United Nations Website, 2020, p.1–4.

United Nations Educational, Scientific and Cultural Organization, *Evaluation of the project "Social Media 4 Peace"*, United Nations Global Marketplace, 19 March 2024, p. 1.

United Nations General Assembly, *Activities of the United Nations system in implementing the United Nations Global Counter-Terrorism Strategy*, in «Report of the Secretary-General», United Nations Website, New York, 2 February 2023, p. 3–7.

United Nations General Assembly, *As Terrorists Exploit Online Platforms, Sixth Committee Delegates Urge Holistic, Global Collaboration with Governments, Civil Society, Internet Providers*, United Nations Press Releases, 8 October 2019, pp. 1 – 3.

United Nations General Assembly, *Resolution 75/309: Promoting interreligious and intercultural dialogue and tolerance in countering hate speech*, United Nations Digital Library, New York, 21 July 2021, p. 2–5.

United Nations General Assembly, *The United Nations Global Counter-Terrorism Strategy (Resolution 60/288)*, United Nations Website, New York, 20 September 2006, p. 1–9.

United Nations General Assembly, *The United Nations Global Counter-Terrorism Strategy: eighth review (Resolution 77/298)*, in «Resolutions and Decisions», New York, 22 June 2023, p. 2–12.

United Nations Office of Counter-Terrorism, *United Nations Global Counter-Terrorism Strategy*, United Nations Global Counter-Terrorism Website, 2023, p. 1.

United Nations Secretary General, *Countering disinformation for the promotion and protection of human rights and fundamental freedoms (A/ 77/287)*, in «Report of the Secretary-General», United Nations Digital Library, 12 August 2022, p. 2.

United Nations Security Council, *(Annex) Comprehensive International Framework to Counter Terrorist Narratives (S/2017/375)*, in «Letter dated 26 April 2017 from the Chair of the Security Council Committee Established pursuant to Resolution 1373 (2001) concerning Counter-Terrorism addressed to the President of the Security Council», 28 April 2017, p. 5.

United Nations Special Rapporteur on Freedom of Opinion and Expression et al., *Joint Declaration of Freedom of Expression and "Fake News", Disinformation and Propaganda,* Organization for Security and Co-operation in Europe Website, 3 March 2017, pp. 1–4.

United Nations, *Global Issues*, *United Nations Website, p.1.*

United Nations, *NoToHate Factsheets for the United Nations Strategy and Plan of Action on Hate Speech*, United Nations Website, 2023, p. 1–3.

United Nations, *Secretary-General Launches United Nations Strategy and Plan of Action against Hate Speech, Designating Special Adviser on Genocide Prevention as Focal Point*, United Nations Press Releases, 2019.

United Nations, *United Nations Global Principles For Information Integrity: Recommendations for Multi-stakeholder Action*, United Nations Website, 2024, p. 3–18.

United Nations, *United Nations Strategy and Plan of Action on Hate Speech: Detailed Guidance on Implementation for United Nations Field Presences*, United Nations Digital Library, September 2020, p.3–50.

United States Department of Treasury, *Treasury Escalates Sanctions Against the Russian Government's Attempts to Influence U.S. Elections*, in «Press Releases», U.S. Department of Treasury, 15 April 2021.

ÜNVER, Hamid Akin, ERTAN, Arhan S., *The Strategic Logic of Digital Disinformation: Offence, Defence And Deterrence in Information Warfare*, edited by ARCOS R., CHIRU I., IVAN C., *Routledge Handbook of Disinformation and National Security*, Routledge, Abingdon, 17 November 2023, p. 192–202.

VENEMA, Agnes E., *Deepfake Disinformation: How Digital Deception and Synthetic Media Threaten National Security*, edited by ARCOS R., CHIRU I., IVAN C., *Routledge Handbook of Disinformation and National Security*, Routledge, Abingdon, 17 November 2023, p. 175–186.

Verified, *Our Mission*, Verified, p. 1.

WALZER, Michael, *Just and Unjust Wars: A Moral Argument with Historical Illustrations*, 5th edition, Basic Books, New York, August 11 2006, p. 323 – 325.

WARDLE, Claire, DERAKHSHAN Hossein, *Information Disorder: Toward an interdisciplinary framework for research and policy making*, in «Report», Council of Europe, 27 September 2017, p. 5.

WILLIAMS, Heather J., BLUM, Ilana, *Defining Second Generation Open Source Intelligence for the Defense Enterprise*, RAND Corporation, California, 17 May 2018, p.40.

Regulatory acts and case law

Council of Europe, *European Convention on Human Rights (as amended by Protocols No. 15)*, in «Council of Europe Treaty Series», entered into force on 4 November 1950, 1 August 2021, Art. 10, p. 12.

Council of the European Union and European Parliament, *Regulation (EU) 2022/2065 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act)*, in «Official Journal of the European Union», Article 1(1), 19 October 2022, pp. 20–78.

European Court of Human Rights, *Pavel Ivanov v. Russia*, Application no. 35222/04, 20 February 2007.

European Court of Human Rights, *Roj TV A/S v. Denmark*, Application no. 24683/14, 24 May 2018.

European Parliament and Council of the European Union, *Directive (EU) 2017/541 of the European Parliament and of the Council of 15 March 2017 [on combating terrorism and replacing Council Framework Decision 2002/475/JHA and amending Council Decision 2005/671/JHA]*, in «Official Journal of the European Union», 31 March 2017, p. 4.

European Union, *Charter of Fundamental Rights of the European Union* in «Official Journal of the European Union», entered into force on 1 December 2009, 26 October 2012, Article 11(1), p. 11–21.

European Union, *Consolidated version of the Treaty on the European Union*, in «Official Journal of the European Union», entered into force on 1 December 2009, 26 October 2012, Art. 5(1), p. 4–7.

European Union, *Consolidated version of the Treaty on the Functioning of the European Union*, in «Official Journal of the European Union», entered into force on 1 December 2009, 26 October 2012, Art. 83, pp. 34 – 102.

European Union, *Consumer Protection*, in «Summaries of EU legislation», EUR-Lex, 2023, p. 1.

European Union, *Digital Services Act*, in «Summaries of EU legislation», EUR-Lex, 2022, p. 1.

International Court of Justice, *Reparation for injuries suffered in the service of the Nations*, Overview of the Case, in «ICJ Reports 174», 11 April 1949, p. 1., par. 1 – 24.

International Court of Justice, *Reparation for injuries suffered in the service of the Nations*, Advisory Opinion of 11 April 1949, in «ICJ Reports 174», 11 April 1949, p. 177–184.

International Law Commission, *Vienna Convention on the Law of Treaties*, in «Treaty Series», Vienna, Art. 26, 22 May 1969, p. 11.

League of Arab States, *The Arab Convention For The Suppression of Terrorism*, Translated from Arabic by the United Nations English translation service (Unofficial translation) on 29 May 2000, United Nations Office on Drugs and Crime Website, Cairo, 22 April 1998, Preamble, p. 1.

League of Nations, *Convention for the Prevention and Punishment of Terrorism*, in «Treaty Collection», League of Nations Archives, Geneva, 16 November 1937, Article 1, p. 6.

United Nations General Assembly, *International Covenant on Civil and Political Rights*, in United Nations, «Treaty Series», New York, Art. 19(3), 16 December 1966, p. 11–24.

United Nations Security Council, *Resolution 1373 on threats to international peace and security caused by terrorist acts*, in «Security Council Resolution», New York, 28 September 2001, pp. 1–3.

United Nations Security Council, *Resolution 2354 on implementation of the Comprehensive International Framework to Counter Terrorist Narratives*, in «Security Council Resolution», New York, 24 May 2017, p.1–2.

United Nations Security Council, *Security Council resolution 2178 on threats to international peace and security caused by foreign terrorist fighters*, in «Security Council Resolutions», 24 September 2014, p. 1.

United Nations Special Tribunal for Lebanon (Appeals Chamber), *Interlocutory Decision on the Applicable Law: Terrorism, Conspiracy, Homicide, Perpetration, Cumulative Charging*, in «Case Law», 16 February 2011, p. 57–73.