LUISS

Department of Political Science

Bachelor Degree in Politics: Philosophy and Economics

Chair of International Relations

Global gaslighting: how political manipulation through social media shapes the international order

Supervised by

Prof. Raffaele Marchetti

Candidate:

Ludovica Lasconi

Matr. 094552

Academic Year 2022/2023

Index

Introduction

The digital revolution didn't enable people to just post a picture at dinner with their friends to make others jealous, to like a TikTok video on the most effective manifestation techniques, to swipe right in the research of their true love on Tinder, nor even to follow the latest dramas of the Kardashian family. The digital revolution has entrusted each individual to hold a new revolutionary and wide-reaching form of power: the ability to worldwide access, spread and create information of any sort. This contemporary form of power, however, did not only impacted the civil international community at an individual level, but it completely and radically shifted the bedrock of the international order and the pillars of international relations theories. While in the last hundreds of years, IR theorists have always considered states as the principal actors advancing their interests based on military, economic and political power, today everyone has the ability to enter the global political sphere with one simple click from their comfortable couch. Therefore, it is no surprise that Rothkopf in the first stages of this revolution already perceived how the 'rules of the game' were about to change in international relations, predicting how "*the 'realpolitik' of the new era is 'cyberpolitik', in which the actors are no longer just states and raw power can be countered or fortified by information power*" (Rothkopf, 1998). It is the reach of a power born in the complex realm of digital communication, sustained by algorithms, and wielded through the screens that pervade our daily existence. This is the power of information, of narratives, and of perception. However, while the – almost – universal ability to have a share of this power had been considered empowering the international society and fostering democracy and globalization world-wide in the beginning of the 21st century[1], the events of the last decade have demonstrated how this "*spirit of informationalism*" (Castells, 2000) has, indeed, eroded people's trust in the truth, facts, democracy, international institutions, and in their own epistemic autonomy.

This Orwellian-like post-truth world is actually a result of today "*network society*" (Castells, 2000), where since everyone has access to and can exercise such power, everyone is also a subject of it. Therefore, the aim of this thesis is to analyze the drawbacks of this "*tsunami of information*" (Chavalarias, 2022), derived from the latest information revolution, that spread across the international society a sense of helplessness regarding factual knowledge. The possibility to online share disinformation and to politically manipulate social media has, indeed, established worldwide a phenomenon that only recently has gained notoriety: political gaslighting. For this purpose, social media platforms have been exploited by a complex web of actors to share disinformation,

---

[1] Examples of this belief can be found in Daalder & Lindsay (2003), Friedman (1999), Fukuyama (1992), Norris (2001), Tapscott (1996), Shirky (2008), and many others that will be analyzed in the course of the thesis.

undermine political opponents or dissenting voices, and to manipulate public opinion so much that individuals lost both the ability and the interest to investigate the truth, being overwhelmed by an environment where "*lies have become the modern political currency*" (Tesich, 1992).

However, while similar politically manipulation techniques had already been employed domestically before the digital revolution, as I shall analyze in Chapter I, the universal accessibility of the '*hypermedia*' (Deibert, 1997) and its subsequent global interconnectedness have moved this phenomenon from the domestic to the international sphere, creating a complex web of actors trying to influence and re-shape the international order according to their own diversified interests. Therefore, the aim of this thesis is to grasp the complexity of global gaslighting and worldwide online political manipulation, in an attempt to investigate this issue in all its entangled aspects demonstrating the threat it poses on democracy and the international system. Thus, this thesis will be structured as follows.

Firstly, Chapter I will be focused on the analysis of political gaslighting and the reasons why this concept has received much attention only recently. Therefore, I shall first provide a literature review on the concept of power in international relations theories, taking into account realism, liberalism, cultural hegemony and constructivism. Mainly using the last two approaches, according to which "*power is the ability to control the meaning of events*" (Wendt, 1992), I shall examine the theorization of the role of the media in international relations, adopting a necessary holistic ontology to understand why "*changes in modes of communication – the various media by which information is stored and exchanged – have significant implications for the evolution and character of society and politics at a world level*" (Deibert, 1997).

After having asserted how according to some IR scholars whoever controls the media holds the power to shape the relations between the agents acting in the international system, I shall dedicate the following sections to the rise of the concepts of gaslighting and post-truth era, in order to understand them distinctively before delving into their intricate relation. Thus, starting from the theoretical pillars of the post-truth world, I shall demonstrate how Keyes' post-truth politics, in which politicians and public figures can blatantly lie and distort facts with little to no consequence, Harari's thesis on homo-sapiens' preference of power over truth, Frankfurt's distinction between liars and "*bullshitters*", and Latour's constructivist post-modernism fostered a reality characterized by the loss of faith in democracy and in traditional sources of authority and knowledge, paving the way to the spread of political gaslighting.

Then, I shall introduce the ways in which political gaslighting as a phenomenon pervaded today's post-truth system, delving both into its definition and highlighting the similarities with post-truth political manipulation, namely, as expressed by Natascha Rietdijk, the employment of the following

techniques: the introduction of counternarratives, the discrediting of (potential) critics, and the denial of more or less plain fact.

However, in order to explain the magnitude and the peculiarity of such phenomenon, I shall present previous political gaslighting attempts that have been employed throughout history: from Ancient Greece's demagoguery to Nazism's "*Große Lüge*", I shall navigate an historical journey showing how this phenomenon changed based on the characteristic societal conditions of the time in question. This historical overview shall enable me to present the effect that the digital revolution had in amplifying the reach of political gaslighting, shaping the international community in such a way so that this phenomenon could have serious repercussion on worldwide perception of democracy and the international order.

Chapter II shall focus on analyzing the current global political gaslighting in a comprehensive way, untangling the many different aspects that characterize it. In particular it will be structured answering the following questions: how is it spread and how does it threaten democracy and the international order, what is its content, which actors are involved and their specific roles, and what countermeasures have been implemented in order to counteract it.

The modalities in which it is perpetrated on social media shall be investigated according to the techniques expressed on Chavalarias' "*Toxic Data*" – namely through the use of bots and fake accounts, coordinated action networks and sentiment manipulation. As social media continuously "*challenges political and elite-dominated paradigm of media's effect on political behavior*" (Zeitzoff, 2017), I shall provide a thorough analysis on how each of Chavalarias' techniques resonates with Rietdijk's theorization of political gaslighting and investigate further in depth how such methods globally jeopardize democracy and the international order. In doing so, I shall refer to several studies and reports that showed how political manipulation through social media has altered the balance of power in the international system, having a global impact on modern peace, conflict, trade, diplomacy, and a plethora of other international processes, such as influencing the policies and undermining the sovereignty of foreign states.

In examining the content of political gaslighting, I shall take into consideration the pervading phenomenon of digital disinformation, the major setback of today's "*information fatigue*" (Cooper, 2019). In analyzing the contemporary information ecosystem, I shall delve into the distinction between the concepts of information, disinformation, and misinformation, distinguishing Wardle's seven types of MDI and highlighting how these influenced a "*collective cognitive dissonance*" (Rogers and Niederer, 2020) among the international audience. Moreover, I shall also consider how other IR theories have dealt with this issue, focusing on some similar characteristics with E. H. Carr's notion of propaganda, John J. Mearsheimer's typology of lies, and Joseph Nye's

conceptualization of public diplomacy – as La Cour demonstrated with her theorization of digital disinformation as an actual weapon in the international arena.

As I have previously mentioned, the range of actors involved in global gaslighting and digital disinformation are very diversified, ranging from states, transnational relations, such as social movements and transnational terrorism, multinational corporations, and international institutions. In regard to state agents, I shall focus on the widespread phenomenon of online foreign influence, in which states attempt to influence the perception of foreign nations among the public opinion, trying to interfere in their elections or to undermine the functioning of their democratic institutions through digital platforms in order to pursue their own interests. In investigating such phenomenon, I shall refer to several examples and reports, focusing on the comprehensive work of Bradshaw et. al "*Industrialized Disinformation: 2020 Global Inventory of Organized Social Media Manipulation*", in which they distinguished for each country the several techniques employed in such attempts. Concerning non-state actors such as transnational networks, I shall first delve into the intricate subject of global governance, highlighting how in international relations non-state actors have increasingly gained much power in the last decades. Therefore, I shall investigate how disinformation campaigns can be both perpetrated and used to undermine these transnational networks, referring primarily to social movements and transnational terrorism and focusing on their similarities and differences in using and being subjects to such techniques.

As for multinational companies, I shall focus on the Big Tech industry and their potential to act as agents of change as well as stakeholders in global governance structures, having a major impact on the international system also in non-economic laden fields. In particular, I shall refer to Susan Strange (1996) and Edmund Li Sheng (2022) in order to highlight how social media companies have undermined the role of the state in international relations, providing much more direct and responsive platforms that political actors can easily exploit to disseminate tailored messages, target specific demographics, and potentially sway public opinion.

Moreover, the multi-channel relationship between Big Tech companies and governments shall be put under scrutiny, in trying to understand how their different interests often clash due to legislative gaps over social media regulations and the intricate issue regarding the limits of the freedom of speech. Therefore, particular attention shall be dedicated to social media regulations' attempts to counteract online disinformation, which – as highlighted by the European Parliament's Subcommittee on Human Rights - involve three groups of actors, namely digital platforms, legislative and regulatory bodies, and civil society. Thus, I shall delve into the different voluntary content moderation initiatives promoted by the private sector, comparing the strategies of companies such as Facebook, Twitter, Google, and Microsoft. Then, I shall focus on the

institutional actions that have been advanced to raise awareness, prevent and combat this phenomenon: firstly I shall investigate the different approaches that states have employed in facing such matter, then I shall delve into international institutions' initiatives - such as the UN's "Verified" project, UNESCO's Global Media and Information Literacy Week, G7 Rapid Response Mechanism – putting particular emphasis on the commitment that the European Union has shown in counteracting digital disinformation through its plethora of implemented plans.

Finally, Chapter III shall be directed at examining the effectiveness of global political gaslighting and disinformation campaigns in eroding trust in democracy and the international order. For this reason, I shall first analyze the ongoing democratic recess that, according to the Freedom House, the world has been witnessing in the last 17 years. In doing so, I shall compare this phenomenon to Huntington's theory regarding the third democratization wave and his subsequent prevision of the possibility of a democratic recess due to his concept of '*electronic dictatorship*'. In particular, I shall take into consideration the relation between this phenomenon and the spread of social media, as the Freedom House has warned in its last report "*this assault coincided with the rapid uptake of information and communication technologies that have effectively broken many states' media monopolies*" (Freedom House, 2023).

Moreover, after having demonstrated the different effects that social media have on different regime types – as shown by Schleffer and Miller's work (2021) – I shall in dept investigate as a case study the Russian Federation, which in almost every report focused on this matter, has been considered the "*most sophisticated and pioneering actors*" (Bradshaw et. al., 2021) employing these political manipulation techniques through social media, targeting several countries and institutions.

Analyzing Russian "*informatsionnaya voyna*" and its "*full spectrum propaganda*" (Nimmo, 2018) through its several comprehensive disinformation campaigns - like Secondary Infektion, its interference in both 2016 and 2020 US elections, and the digital strategies undertaken regarding its invasion of Ukraine – will demonstrate the actual reach and the threat that this phenomenon poses to the international community.

Lastly, I shall face the major theoretical hole regarding the phenomenon of political gaslighting and digital disinformation, namely the quantitative assessment of its effectiveness. In fact, many attempts have been proposed in trying to provide a satisfactory assessment methodology, as it is highlighted by several institutions and scholars. However, after having analyzed these proposals, I shall conclude that, although presenting significant insights to this problem, neither Nimmo's 'rule of thumb' nor British HMG's RESIST framework provide a satisfactory solution to this assessment issue, leaving the question of the effectiveness' assessment of political gaslighting through social media open to further research.

Chapter I

It is no secret that politicians and prominent figures have tried to manipulate citizens and public opinion since the development of the first complex societies through different means – whether it was through Alcibiades' δημαγωγία, the interpretation of Socrates' trial, the *ars oratoria* in Republican Rome[2], religious and mythological explanations, the pursuit of some ideologies, until the development of the newest forms of propaganda and media control. However, it is strikingly surprising how the concepts of gaslighting and media manipulation have gained such notoriety only in recent times. In this sense, the aim of this section is to demonstrate how the ability to gaslight and manipulate public opinion through media has become the mainstream form of power in international relations. In order to demonstrate how the control of the media has become the primary source of power and to what extent the actors can shape the international system through their manipulation, this chapter will be structured as follows. Firstly, I shall review the concept of power in the main international relations' theories, namely realism, liberalism, cultural hegemony, and constructivism. Then, I shall delineate the ways in which actors are able to change the international system, focusing on the constructivist approach and its theory regarding the media in order to demonstrate that whoever controls the media, has the major power in the international scene. Subsequently, I shall analyze the various techniques of media manipulation, highlighting the concept of gaslighting in the post-truth era so that the widespread phenomenon of political gaslighting could be examined. Therefore, I will provide an overview of the political manipulation techniques that have been employed throughout history, demonstrating how political gaslighting, although its conceptualization is rather new, has always been present in every developed civilization. Thus, I shall conclude by explaining how such manipulation technique has been employed with the technological innovations that characterized the 20th century, bringing us to the question that I will analyze in the next chapter: how can political gaslighting be employed through social media and to what extent can it be effective in the international order.

Section 1.1 – Power in International Relations

International relations have as focus and aim the concept of power, how to gain it, how to keep it, and in which forms power can show itself, since it is power in all its different shapes that determine the relationships among states. Therefore, since it plays a central role in shaping the behavior of states and other international actors, the study of power has been a significant focus of international

---

[2] Cicero in "De Oratore" defines the ways through which an orator can shape the opinions of the listeners ("*ut eos, qui audiant, quocumque incubuerit, possit impellere*" - De Oratore, III, 55), namely the techniques of *probare*, *flectere* and *delectare*.

relations theory, and various approaches have sought to understand the nature and role of power in the international system. Although it is a complex and multifaceted concept that has been defined in various ways by scholars, the most common and all-encompassing definition of power belongs to Dahl, who describes it as the ability to make another actor do what s/he would not have otherwise done or not doing what s/he would have otherwise done (Dahl, 1957)[3]. However, being this a very general definition, it doesn't examine the different forms that power can undertake nor its implications. Therefore, the aim of this section is to analyze the various theorizations of the concept of power and demonstrate the need of understanding it with a multidimensional approach in order to interpret today's international order.

Realism, one of the most prominent and oldest theories, actually theorizes every aspect of international relations in terms of power, considering politics as a fight that has power as a goal and force as an instrument. In this sense, realists start from the universalization of the power struggle, considering the fight for power universal in time and space, and an undeniable empirical fact (Morghentau, 1948)[4]. The entire realist theory has developed based on power in all its aspects. It is the underlying rationale behind the goals that states must pursue, their political tactics, and the relations between states in the international order. Therefore, realists believe that states should strive to maximize their power to ensure security and survival through military force and economic coercion, which they view as the primary sources of power. States employ various political strategies based on their specific goals: preserving power through status quo politics, increasing power through imperialist politics, and showcasing power through prestige politics. Furthermore, the central belief that the global system is anarchic and relies on state behavior in a state of nature enabled the emergence of the balance of power as the mechanism that governs international order. Indeed, "*balance-of-power politics prevail wherever two, and only two requirements are met: that the order be anarchic and that it be populated by units wishing to survive*" (Waltz, 1979). Realists view the balance of power as a mechanism that prevents any single state from gaining excessive power and dominating others, because of the anarchy and perpetual conflicts that define the

---

[3] Specifically, in "*The Concept of Power*", Dahl claims that "*A has power over B to the extent that A can get B to do something that B would not otherwise do*".

[4] In particular, Morghentau in "*Politics Among Nations*" argues "*the struggle for power is universal in time and space and is an undeniable fact of experience. It cannot be denied that throughout historic time, regardless of social, economic and political conditions, states have met each other in contests for power. Even though anthropologists have shown that certain primitive peoples seem to be free from the desire for power, nobody has yet shown how their state of mind can be re-created on a worldwide scale so as to eliminate the struggle for power from the international scene. ... International politics, like all politics, is a struggle for power. Whatever the ultimate aims of international politics, power is always the immediate aim*".

international order. The balance of power is maintained by weaker states forming alliances and counterbalancing the stronger states' power, creating, therefore, a state of deterrence. Hence, the fear of reprisal from other nations prevents states from pursuing aggressive actions, thereby ensuring stability in the global system.

The theory of liberalism emerged in international relations as a contrasting view to power-seeking realism. It interprets power differently, altering its role in the international order. By highlighting institutions, cooperation, individual freedom, human rights, and democracy, liberalism rejects the notion that states must constantly compete and balance power for their survival. For proponents of liberalism, power is not an end in itself or a zero-sum game where one nation's gain means another nation's loss. Instead, power is viewed as a means to promote peace, cooperation, and mutual prosperity among nations. This can be achieved by advocating for democracy, free trade, and, most importantly, international institutions. Power is not just about military and economic strength; it also involves the ability to influence and shape others' behavior through persuasion and cooperation. Liberals stress the significance of non-coercive power, international cooperation, and institutions in addressing shared problems and advancing common interests. Consequently, their agenda sets them apart from realists by encompassing a wider range of topics that extend beyond traditional security concerns, involving also issues of "low politics". So, since liberalism's agenda is no longer focused only on security, liberalists seek to limit the power of states through international law, human rights, and democratic institutions, so that power cannot be concentrated and abused. Therefore, according to liberal theorists, international institutions such as the United Nations, the World Trade Organization and the International Monetary Fund play a crucial role in fostering cooperation among states and ensuring a fair distribution of power. These institutions provide a framework for resolving conflicts, managing economic interdependence, and promoting democratic governance. However, there are critics who contend that the liberal perspective is excessively idealistic and impractical. They argue that it places excessive emphasis on collaboration, adopts ineffective economic policies, and overlooks power dynamics. A prime example of this criticism is highlighted by Robert Cox, who points out the flawed assumption that the interests of powerful states align with those of weaker states. Moreover, liberalism's focus on the role of international institutions does not leave enough space for the newest forms regulating today's international system and for the role of cultural hegemony.

Therefore, to comprehend the various forms of power a state can wield, it is crucial to examine Antonio Gramsci's theory of cultural hegemony. This theory elucidates how power is exercised beyond military or economic means. "*Gramsci, Hegemony, and International Relations: An Essay in Method*" by Robert Cox illustrates how cultural hegemony is relevant to international relations.

Culture dominates how powerful groups in society influence society's values, beliefs, and norms. According to Gramsci, this is accomplished through cultural institutions like education, media, and religion. Dominant groups utilize these institutions to establish a collective perception or a shared comprehension of that which is inherent or customary in society. The dominant group uses this shared understanding to justify their interests and maintain power over subordinate groups. Cox's essay elaborates on Gramsci's concepts, contending that hegemony is not solely about force or coercion, but is also accomplished through the development of "hegemonic ideas" that influence people's perceptions of the world. These dominant ideas are not merely forced from above but are formed in a dynamic struggle between various social groups. And, in fact, Cox, while trying to verify the applicability of Gramsci's theory to international relations, argues that "*world hegemony is describable as a social structure, an economic structure, and a political structure; and it cannot be simply one of these things but must be all three. World hegemony, furthermore, is expressed in universal norms, institutions and mechanisms which lay down general rules of behavior for states and for those forces of civil society that act across national boundaries – rules which support the dominant mode of production*" (Cox, 1983). The theory of cultural hegemony has, therefore, important implications for understanding power in international relations. First, it suggests that dominant states can employ their cultural sway to mold the ideals and convictions of other nations and to further their own interests in the global arena. For example, the United States has been accused of promoting its own cultural values, such as consumerism and individualism, around the world in order to maintain its global dominance. Second, the theory of cultural hegemony suggests that weaker actors can challenge the dominant cultural narrative by promoting alternative values and beliefs. By doing so, weaker actors can create space for themselves in the international system and resist the dominance of more powerful ones. This can be achieved through the figures of the "modern prince", i.e. revolutionary parties, intellectuals, and vanguards able to stimulate new consciousness and political commitment, mobilizing the masses and proposing alternative normative readings of reality through a war of position (spread in civil society) and a war of maneuver (seize the power). The emphasis and the power within these counter-hegemonic movements are fundamental in understanding the power that civil society has in the international system, particularly in light of the fact that anyone can now share their own point of view with a worldwide audience simply using their smartphone. Moreover, Cox's essay further expands on the implications of the theory of cultural hegemony for international relations by arguing that the struggle for hegemony is a fundamental aspect of world order. According to Cox, world order is not simply a matter of the balance of power between states, but also the conflict between social groups to establish a dominant order that reflects their interests.

This aspect regarding changing the dominant narrative and the power of culture and ideas has been widely developed and enlarged by the constructivist approach, which will be the essence of understanding the real nature of power in nowadays' international system. Constructivism, in fact, emphasizes the role of social norms, ideas, and beliefs in shaping state behavior and the international system, considering power not only a function of material capabilities but, mainly, constructed and contested through social interactions. As such, power is also contingent upon the shared understandings and social norms of actors in the international system. According to Alexander Wendt, "*power is the ability to control the meaning of events*" (Wendt, 1992). In other words, power is not a fixed or objective quantity, but is a matter of interpretation and contestation. An important idea for grasping power in constructivism is the influence of identity on state behavior and power distribution in the international system. Wendt argues that "*identities and interests are not given, but constructed through social interaction*" (Wendt, 1999). This means that the way in which states define themselves and their interests is not fixed or predetermined, but it is shaped by their interactions with other actors in the international system. Thus, the distribution of power in the international system is influenced by states' identities and interests, not just their material capabilities. Norms are another important concept for understanding power in constructivism. Norms are shared understandings of appropriate behavior that shape state behavior and constrain the exercise of power in the international system. According to Martha Finnemore and Kathryn Sikkink, two of the most influential scholars on the role of norms in international relations, "*norms are not just 'out there,' waiting to be discovered, but are actively constructed and reinforced through social interaction*" (Finnemore and Sikkink, 1998). This means that norms are not fixed or objective, but are constructed and contested through social interactions. As such, the ability to shape and enforce norms is an important aspect of power in the international system. Upon this conception of power, constructivists have built the theory of the politics of change, which emphasizes the role of social and political mobilization in transforming the existing international order. In fact, due to the flexibility of the narratives in the international system, constructivists look at the critical junctures of history, namely those moments in which different actors have challenged the dominant narrative and promoted some normative changes. The core of the politics of change is the spread into international society the understanding of a central norm, for example, constructivists have analyzed the last four centuries through the promotion and, later, change of these narratives: with the Peace of Westphalia the principle of sovereignty has monopolized the organization of international relations; Wilson's idea of self-determination in the Peace of Versailles has reshaped the international community; the UDHR has promoted worldwide human rights; while since 2005 states seem to have embraced the so-called principle of responsibility to protect (R2P). Therefore,

"*the politics of change is about actors who mobilize to transform, expand, or create norms, rules, and institutions, and who contest those already in place*" (Risse, Ropp, and Sikkink, 1999). In this context, constructivists enlarge the neo-Gramscian conception, by highlighting the fundamental role of social movements in the politics of change. Social movements are collective efforts by individuals or groups to challenge existing norms, rules, and institutions, and to promote alternative visions of the international system. According to Sidney Tarrow, "*social movements can generate new cultural meanings, which provide the basis for new institutional arrangements*" (Tarrow, 1998). Therefore, social movements can have a crucial impact on transforming the global system. For example, the women's movement has challenged traditional gender roles and has succeeded in promoting women's rights and gender equality across the majority of countries. Constructivists have also re-elaborated Gramisci's concept of the "modern princes" by depicting the figure of the norm entrepreneurs, who are individuals or groups who seek to create, promote, or transform social norms in order to advance their political objectives. According to Martha Finnemore, norm entrepreneurs can change the international system by developing, disseminating, and institutionalizing new norms (Finnemore, 1996). Therefore, norm entrepreneurship can play a critical role in transforming the international system by promoting new norms and values that challenge the existing order. For example, the norm of human rights has been promoted by norm entrepreneurs such as Amnesty International and Human Rights Watch, who have monitored state behavior, documented human rights abuses, and advocated for international human rights standards.

Section 1.2 – Media as the catalyst of political change

The several conceptions of power previously analyzed, have delineated four different theories regarding the ways in which actors can shape and change the international order. Realism suggests that the international system is anarchical, with no central authority governing international relations. States constantly compete with each other, resulting in a power balance system where they aim to maintain an equilibrium of relative power, preventing any one state from dominating. Realists alter the international system by pursuing power and security in their own self-interest. This may entail participating in armed conflict, making alliances, and attempting to acquire resources or territories. The change in the global system is thus the change in the dominant power of countries measured by their resources. Contrary to realism, liberalism suggests that cooperation is achievable in the global system and that international organizations and institutions can be utilized to handle conflict and advance collective action. Liberalism posits that actors can alter the international system through dialogue, negotiation, the development of international institutions, and adherence to international norms and laws. Through international organizations like the UN or the IMF, states can collaborate to address common issues like climate change or poverty. The

theory of cultural hegemony, instead, emphasizes how cultural and ideological power shapes the international system. According to this theory, actors can alter the world system by promoting cultural norms and values that become dominant and influence the behavior of others. Cultural dominance can happen when ideas, values, and norms spread through institutions like media, education, and religion. A strong government could spread its culture and beliefs by exporting media and popular culture, causing others in the international system to adopt these values. Additionally, constructivism suggests that the international system is socially constructed, and actors can alter the system through their ideas and actions. This involves forging new identities and norms or questioning current ones. For example, civil society actors can promote new ideas on human rights that may gain wide acceptance and institutionalization, ultimately leading to changes in international law and norms.

In the last two approaches, media plays a crucial role in catalyzing a structural change in the international order. They serve as a platform for modern princes or norm entrepreneurs to share their message. According to these theories, the media can shape social reality and construct the meaning of international events. Therefore, they act as active agents that shape and transform the international system instead of being passive observers. The media narratives play a crucial role in shaping the social construction of reality, by, for example, selecting and presenting the news, and by their engagement with the audience. Bennett and Paletz (1994) contend that media narratives create a symbolic reality that shapes public opinion and policymaking. This reality is not simply a reflection of objective events, but rather a result of the social processes involved in media production, selection, and interpretation. Therefore, Bennett and Paletz developed a social constructionist approach to media studies, emphasizing media's role in constructing the meanings and identities of social actors. They contend that media are more than mere instruments of communication; rather, they are formative forces that shape and uphold social reality. The media shape public discourse on international issues by selecting, framing, and interpreting news events, creating a shared understanding of the world. In particular, in their book "*Taken by Storm: The Media, Public Opinion, and U.S. Foreign Policy in the Gulf War*", the authors provide a comprehensive analysis of the role of the media in shaping public opinion and U.S. foreign policy during the Gulf War, using a combination of qualitative and quantitative research methods in order to demonstrate how media images and narratives influenced public perceptions of the war, which in turn affected political decision-making. In fact, after establishing the historical context of the Gulf War and the media landscape of that time, they contend that the constant news cycle and fierce competition among news outlets to engage audiences led to sensational and shallow coverage of the war. This coverage primarily emphasized the strength and heroism of the American military while

largely disregarding the plight of Iraqi civilians. Bennett and Paletz analyzed public opinion during the war: using survey data to demonstrate how media influenced Americans' attitudes towards the conflict, they showed how support for the war initially started off high but declined as casualties grew and the true costs of the war became more evident. The authors argued that the media had a vital role in shaping public opinion by framing the war in specific ways: for example, they tended to overemphasize the threat of Saddam Hussein's weapons of mass destruction and depicted the conflict as a moral battle between good and evil, an important point that then will be analyzed again while explaining the different gaslighting techniques employed in media coverage. Moreover, the authors have also focused on the relationship between media coverage and political decision-making, with a particular emphasis on the role of the Bush administration in managing the media during the war. The authors assert that the White House adeptly manipulated media coverage to further its political objectives, by employing tactics like assigning journalists to military units and managing information access.

However, media do not only have a direct impact on the social constructions of reality, but they also influence and shape the construction of international norms and identities, which is a central concern of constructivist theory. According to Finnemore and Sikkink (1998), norms are shared understandings of appropriate behavior that guide social action. They argue that media plays a key role in disseminating and institutionalizing international norms. By highlighting instances of normative behavior and moral dilemmas, media can stimulate public debate and mobilize norm entrepreneurs who seek to change international norms and institutions (Finnemore, 1996). Adler and Pouliot (2011) provide a cognitive theory of international institutions, which emphasizes the role of media in shaping the cognitive processes that underpin institutional change. They argue that media narratives create cognitive maps of the international system, which influence the way actors perceive and respond to institutional challenges. Media can generate cognitive dissonance and trigger institutional change by exposing the gap between existing norms and changing realities. Nonetheless, it is the scholar Ronald Deibert who revolutionized the conception of media within international relations. In fact, his aim was to go beyond the state-centric (neo)realist paradigm that had monopolized the realm of international relations, aiming at taking into account the emergence of new actors on the world scene and adapting to the field of IR the findings of media theorists such as McLuhan and Innis in order to formulate his "therapeutic redescription" of the discipline (MacMillan, 2004). He started from the realization that "*the elements of international politics which mainstream rationalist approaches presuppose to be 'natural', 'essential' and 'unchanging' are, in fact, the products of historical contingencies and thus subject to change over time*" (Deibert, 1997). Therefore, such theories are not able to grasp and encompass the multiplicity of actors playing in

the international arena in the 21ˢᵗ century. For this reason, he argues it is necessary to adopt a holistic ontology where all aspects of society are considered as a whole, and related to communication technologies (MacMillan, 2004), building, therefore, his 'media theory'. According to this, while dealing with media and communication fields, the content of the communicative process does not matter, since the focus is on the medium that makes communication possible and the ways in which it is presented. His main argument is that "*changes in modes of communication – the various media by which information is stored and exchanged – have significant implications for the evolution and character of society and politics at a world level*" (Deibert, 1997). Therefore, the emergence of digital media and communication technologies – that he has defined as 'hypermedia'[5] - has transformed the way in which we think about power and politics in the international arena. The proliferation of social media platforms and the rapid spread of information have given rise to a new form of global politics, which is characterized by the increasing importance of digital media in shaping public opinion and influencing policy decisions. Deibert asserts that mass media are no longer confined to conventional mass communication forms, such as newspapers, television, and radio. By rapidly disseminating information and mobilizing people and collectives, digital media has made political engagement and public discussion easier than ever. However, Deibert does not express a draconian judgment on the impact of digital media on international relations, presenting both its positive and negative aspects. On the one hand, the widespread use of social media platforms and other digital technologies has made it easier for individuals and groups to connect and mobilize around shared political goals, being now independent from the "sovereign-territorial" boundaries – an aspect that I shall analyze in more detail in the next chapter. This has led to the emergence of new social movements and political campaigns, which have had a significant impact on global politics. On the other hand, the use of digital media has also posed new challenges for international security and governance. False information and propaganda can undermine democratic institutions and provoke conflicts worldwide through social media platforms, while there are also concerns about civil liberties and privacy when the government and others utilize digital surveillance technologies. The impact of media on international relations in Deibert's study emphasizes the need for academics and policymakers to pay close attention to technological advances in global politics, warning them from underestimating such technologies.

However, the most important point that can be derived from Deibert's theory is that whoever controls the media, whether in an overt or covert manner, holds the real power to shape the relations

---

[5] Deibert defines the 'hypermedia' as "*a complex melding and converging of distinct technologies into a single integrated web of digital-electronic-telecommunications*" which, in conjuction with *social forces, has resulted in "a convergence of both media and instrustries into a single, integrated planetary web of digital electronic communications*" (Deibert, 1997).

between all the actors in the international system. For this reason, it is essential to delineate how this power is exercised through media and how the actors can actually manipulate them in order to shape - according to their own narrative - the world order. It is not a case, in fact, that the concept of 'gaslighting' has firstly monopolized the media field and, later, it has been applied in politics both at a national and international level.

Section 1.3 – The origins of gaslighting as a form of manipulation

Particularly interesting while talking about manipulation is the notion of 'gaslighting', which has witnessed a sudden and rapid development from being a clinical literary term to becoming an everyday common use expression, so much that Merriam-Webster dictionary labeled it the "word of the year 2022". The term 'gaslighting', indeed, increased its lookups by 1,740% in the last year on the famous American dictionary platform, such that the editor Peter Solowski claimed that <<*it's a word that has risen so quickly in the English language, and especially in the last four years, that it actually came as a surprise to me and to many of us* […] *it was a word looked up frequently every single day of the year*>>[6]. For the purpose of this thesis, it is fundamental to analyze in this section the impressive development of the concept of gaslighting and its relevance in contemporary society, particularly in the context of politics and social media.

The first usage of the term was in reference to the 1938 British play "Gas Light", which later became a film in 1944. The term was used to describe the manipulative behavior of the husband who, in the play, used small and subtle actions to make his wife doubt her own sanity. This included turning down the gas lights and denying that he had done so, which made the wife question her own memory and perception of reality. The term 'gaslighting', therefore, initially was established in the clinical field as the <<*psychological manipulation of a person usually over an extended period of time that causes the victim to question the validity of their own thoughts, perception of reality, or memories and typically leads to confusion, loss of confidence and self-esteem, uncertainty of one's emotional or mental stability, and a dependency on the perpetrator*>>[7].

The concept of gaslighting gained greater prominence in the field of psychology in the 1960s and 1970s, particularly in relation to abusive relationships. During this time, researchers and practitioners began to identify the ways in which abusers used manipulative tactics to gain power and control over their victims. Gaslighting was identified as one such tactic in which the *gaslighter*

---

[6] Peter Solowski, editor at large for Merriam-Webster dictionary, explained the concept of gaslighting ad its label as 'word of the year 2022" in an exclusive interview for AP News (https://apnews.com/article/word-of-the-year-2022-80d02a3e0a347e542466571ca9cdb2ef)

[7] According to Merriam Webster, this was the original meaning of 'gaslighting' when it first was used and accepted in the medical community.

tries, consciously or not, to induce in the victim the sense that their reactions, perceptions, memories and/or beliefs are not just mistaken, but rather utterly without grounds, so unfounded as to qualify as crazy (Abramson, 2014). The recognition of gaslighting as a form of abuse was significant and meaningful, as it allowed for greater understanding and support for victims who were experiencing emotional abuse in their relationships. It was so influential that it even became an official part of the British legislation regarding criminal domestic violence law (Sweet, 2019), resulting in the charge of 300 people with this accusation (Mikhailova, 2018).

However, it was only until 2007 that this concept left the psychological and medical sphere in order to conquer a ubiquitous use in the political and media arena.

Section 1.4 – The post-truth era

As previously anticipated, 2007 was a turning point for the development, application, and spread of the concept of gaslighting. In this year, in fact, the American psychoanalyst Robin Stern published the revolutionary book "*The Gaslighting Effect – How to Spot and Survive the Hidden Manipulations Other People use to Control your Life*", which has struck a chord with readers worldwide due to its compelling narrative and insightful analysis of emotional manipulation. One of the key reasons for the resonance of Stern's book is its ability to describe the complex dynamics of gaslighting in a clear and accessible way. By providing a framework for understanding the tactics of gaslighters, Stern helps readers to recognize and name the experiences they may have struggled to articulate. This clarity can be empowering, as it allows readers to feel less alone and more capable of addressing the gaslighting dynamic in their lives. Making the once-elitist concept of gaslighting accessible to the public opinion, not only resulted in a greater awareness of this issue, but it also enabled people and specialists to apply it in different spheres, such as sociology, politics, and communication.

While Sweet provides an exhaustive analysis of gaslighting as a sociological phenomenon, it is Stern herself that presents in the second edition of her book, published in 2018, the ways in which such psychological manipulation dominates the "post-truth" political era (Sweet, 2019). Therefore, in order to provide an exhaustive account of the political role of gaslighting, we cannot omit to define our contemporary society in "post-truth" terms.

The first time such term was used was in the 1992 provocative article "*A Government of Lies*", an article by writer and commentator Steve Tesich. In this article, which results as a scathing critique of the George H.W. Bush administration and its approach to political communication, he analyzes all the previous examples that fostered the "charade of truth" (Tesich, 1992), namely the Watergate syndrome. In order to demonstrate that "*lies have become the modern political currency*" (Tesich, 1992), he shaped his argument on the inverse relation between the citizens' self-esteem and their

knowledge of the truth, to which choice the public responded "*we looked at our government to protect us from the truth*" (Tesich, 1992). Therefore, the Bush government's cynical disregard of the truth not only led to an increased disillusion and disengagement of the public from politics and democratic institutions, but it also paved the way to our current dominant political communication. However, while Tesich foresaw such phenomenon, his definition of post-truth era didn't gain much fame, even if it was recalled by Ralph Keyes' book "*The Post-Truth Era: Dishonesty and Deception in Contemporary Life*" published in 2004. He introduced his work with psychologist Robert Feldman's study, which demonstrated the correlation between threatened self-esteem and the propensity to lie, proving Tesich's thesis. From this study, Keyes elaborated a thought-provoking exploration of the growing trend towards dishonesty and deception in modern society, starting from the definition of "post-truthfulness":

<<*We live in a post-truth era. Post-truthfulness exists in an ethical twilight zone. It allows us to dissemble without considering ourselves dishonest. When our behavior conflicts with our values, what we're most likely to do is reconceive our values. Few of us want to think of ourselves as being unethical, let alone admit that to others, so we devise alternative approaches to morality. Think of them as alt.ethics. This term refers to ethical systems in which dissembling is considered okay, not necessarily wrong, therefore not really "dishonest" in the negative sense of the word*>>.

He follows his analysis by exploring the post-truth's "ethical twilight zone", by delineating a new category in the truth-lie spectrum, "*a third category of ambiguous statements that are not exactly the truth but fall just short of a lie. Enhanced truth it might be called. Neo-truth. Soft truth. Faux truth. Truth lite*" (Keyes, 2004). In today post-truth world, Keyes emphasizes its corresponding post-truth politics, in which politicians and public figures can blatantly lie and distort facts with little to no consequence. Keyes suggests that this phenomenon is due in part to the changing nature of media and the increasing fragmentation of society, which has allowed individuals to consume only the news that confirms their existing beliefs, while ignoring opposing viewpoints.

However, only in 2016 the term "post-truth" skyrocketed in its usage so much that it was chosen by Oxford Dictionary as "word of the year", defining it as an adjective 'relating to or denoting circumstances in which objective facts are less influential in shaping public opinion than appeals to emotion and personal belief'[8]. Several events contributed to the spread of the aforementioned concept in this specific year, raising awareness regarding the issue of factual accuracy and its role in public discourse. For example, the proliferation of social media platforms both helped and contested the use of misleading or false claims in the political campaigns of the Brexit referendum and the US presidential election. Moreover, not only did the spread of social media enable people to spread

---

[8] https://languages.oup.com/word-of-the-year/2016/

news and information rapidly and without fact-checking, but it also enhanced the increasing polarization of political views and the rejection of expert opinions. As a result, the term "post-truth" became a popular way to describe the current state of public discourse, in which emotions and personal beliefs often take precedence over facts and evidence.

Yael Brahms in his article "*Philosophy of Post-Truth*" provides a better understanding of Oxford Dictionaries' definition of the term, offering a thought-provoking analysis of the post-truth era and its implications for philosophy, politics, and society as a whole. In fact, after analyzing philosophical and research work regarding the phenomenon of post-truth – among which Harari's thesis on homo-sapiens preference of power over truth, Frankfurt's distinction between liars and "bullshitters", and Latour's constructivist post-modernism – Brahms delineates the four characteristics that fostered the rise and spread of this reality. According to the author, these "peak waves" are: "*the information explosion and disruptive technology; the dwindling of faith in institutions and in the 'truth tellers'; postmodernist ideas, which seeped into such fields as science and national security and laid the foundation for substandard discourse about truth; and the bitter political battles over the criteria of truth*" (Brahms, 2020).

Therefore, after suggesting that post-truth poses a challenge to traditional philosophical views on truth, as well as to our ability to make informed decisions and participate in democratic processes, Brahms argues that this loss of faith in traditional sources of authority and knowledge - such as science, expertise, and the media – led to a proliferation of conspiracy theories, fake news, and alternative facts due to our common relying on social media, personal anecdotes, and emotional appeals to shape our beliefs and opinions.

Only the acknowledgement of these several phenomena enables us to understand the dynamics and the nature of contemporary post-truth world, which, for the purpose of this paper, will permit the implication that political gaslighting has in today society. It is, therefore, necessary for the purpose of this study to adopt Brahms' definition of post-truth as "*a term denoting circumstances in which our ability to clarify the reality in order to understand it and in order to function within it on the basis of facts is weakening as a result of high-intensity interference by four peak waves: the information explosion and disruptive technology; the dwindling of faith in institutions and in 'truth tellers'; undermining postmodernist ideas; and bitter political battles*" (Brahms, 2020).

Section 1.5 Political Gaslighting in the post-truth era

After having explained the meaning, nature and implication of gaslighting as a manipulative technique, it is only with the spread of the concept of "post-truth politics" that political gaslighting emerged. Due to the interconnection of these two notions, it is no surprise that they both emerged in the public discourse at the same time and strictly linked to each other. As previously illustrated, in

essence post-truth politics refers to a political discourse that relies on appeals to emotions and personal beliefs, rather than facts and evidence, to shape public opinion, while gaslighting is a form of psychological manipulation in which the perpetrator seeks to sow seeds of doubt in the victim's mind, causing them to question their own memory, perception, or sanity. The relationship between gaslighting and post-truth politics is complex, as they are both mechanisms that serve to undermine truth and erode people's trust in objective sources of information and in their epistemic autonomy. In a post-truth political landscape, politicians may employ gaslighting tactics to delegitimize opposing viewpoints and bolster their own narratives, creating a situation where truth becomes relative and subjective. It is, therefore, in such context that scholars and journalists have begun to define and apply the concept of "political gaslighting".

One of the first uses of gaslighting in a political context was Lauren Duca's article "*Donald Trump is gaslighting America*"[9], an op-ed published on Teen Vogue right after his victory in the 2016 elections. She provides several examples demonstrating how "*Trump won the presidency by gaslight*" (Duca, 2016), such as undermining the CIA's trustworthiness, repeating lies, false claims, and conspiracy theories through social media, and by disseminating disinformation with no evidence backing his allegations. Duca's article highlights the insidious nature of political gaslighting, which had become the central strategy of the Trump administration in its efforts to maintain power and control, and its potential to erode the foundations of democracy. In fact, she argues that "*at the hands of Trump, facts have become interchangeable with opinions, blinding us into arguing amongst ourselves as our very reality is called into question*" (Duca, 2016). By attacking the media, discrediting scientific evidence, and labeling any dissenting voices as "fake news", Trump has *de facto* created a world hostile to critical thinking and reasoned debate, in which truth and facts are no longer relevant, and the only thing that matters is the narrative that he chooses to promote.

Duca's article inspired many other journalists, scholars, and political scientists to analyze and explore the techniques and the effects of political gaslighting, shading a completely different light on current political tactics and the behavior of the electorate. In this context, one of the most influential works consists in the book "*Gaslighting America: why we love it when Trump lies to us*" by Amanda Carpenter, which results in a thorough and compelling exploration of the use of gaslighting in American politics, and a call to action to resist this manipulation and defend the truth. Carpenter, in fact, not only breaks down Trump's gaslighting techniques, assisted by the rise of social media platforms, but she also provides an exhaustive account regarding the ways in which

---

[9] https://www.teenvogue.com/story/donald-trump-is-gaslighting-america

gaslighting had been used in American politics throughout history, from the Red Scare to McCarthyism to the birther movement.

Inspired by Duca's article, Carpenter's book and previous studies, Farah Latif in her article "*Political Gaslighting in the Climate Change Discourse Surrounding 2016 Elections*" provides an all-encompassing definition of political gaslighting and analyzes the practical effects of such phenomenon. In fact, after accounting the works of many scholars – among which Welch, Ahern, Davis & Ernst, Fox, Kovacs, Stosny, Carpenter and Duca – on the basis of such literature review across many fields, she defines political gaslighting as follows: "*Gaslighting is a political strategy that utilizes deceptive and manipulative use of information, which destabilizes and disorients public opinion on political issues concerning the public. Several types of discourse and actions may constitute political gaslighting, such as lying, hypocrisy, misleading by mischaracterizing or rejecting facts, use of logical fallacies, and minimizing or deliberate silence on the seriousness of issues that are important to constituents of a politician*" (Latif, 2020). Therefore, by arguing the alteration of public perceptions and the disorientation of public opinion, she intelligently chose to analyze political gaslighting in terms of climate change due to its scientific evidence[10]. The article, therefore, cites numerous examples of political gaslighting in the context of climate change discourse. One example is the way in which politicians and media outlets have used language that downplays the severity of climate change, such as referring to it as "global warming" instead of "climate change". This tactic creates confusion among the public and allows politicians to deny the reality of the issue. Another example of political gaslighting cited by the author is the way in which politicians and media outlets have used selective facts and data to support their arguments. By cherry-picking data that supports their views and ignoring data that contradicts them, politicians can create a false sense of security among the public. Additionally, after having demonstrated the threat that political gaslighting is to democracy, Latif introduced another important implication of such phenomenon, namely – what we shall define - the "duplicity of the victims". Political gaslighters not only induce a sense of confusion and disorientation in the public that results in their inability to make sound judgments based on their observations and beliefs, but they also undermine the credibility of the ones holding opposing views (Latif, 2020) – in this context indeed scientists. Regarding this second type of victims, Latif claims, therefore, that political gaslighting is a tactic used with the purpose of character assassination of the opponents and that it might be a side effect of character attacks (Latif, 2020).

---

[10] Latif argued that she chose to uncover the evidence of political gaslighting in politicians because "*science is clear on climate change with a 97% consensus from climate scientists that man-made climate change exists*".

Another author that expands this duplicity of the victims, by shifting the spotlight from the gaslighters to the gaslightees, is Natascha Rietdijk in her article "*Post-truth Politics and Collective Gaslighting*", which consists of a colossal insight into political gaslighting's techniques and their strict link with the post-truth world. Rietdijk, actually, is the first author that provides a comprehensive and factual demonstration of the link between gaslighting and post-truth politics through their main common implication, namely the undermining of the victims' epistemic self-trust and epistemic autonomy, introducing new far-reaching psychological threats in addition to the ones addressed to knowledge and democracy. Therefore, she elaborates her demonstration on three techniques that are common both in post-truth politics' manipulation and in psychological gaslighting, namely the introduction of counternarratives, the discrediting of (potential) critics, and the denial of more or less plain fact (Rietdijk, 2021). Her argument, indeed, revolves around the fact that "*post-truth politics is not just an expression of disregard for truth or an origin of false beliefs; it can also undercut our own faith in our capacity to distinguish between truth and falsehood, between reliable and unreliable sources*" (Rietdijk, 2021), which subvert the public's epistemic autonomy, just as Gregory – from the play "*Gas light*" - does to his wife Paula. The author manages to meticulously apply individual psychological phenomena into the broader sphere of politics by creating a juxtaposition between gaslightees' epistemically, emotionally or prudentially predisposition to believe their gaslighters to the trust and dependency that the public has towards politicians due to their power and respect given by their public offices. In this sense, both the gaslighters' aim and the politicians' purpose is not the undermining of the victims' self-trust or epistemic autonomy, which results in a highly problematic side effect, but it is the consolidation of their power and control over them. In order to achieve such goal, both the gaslighters and the politicians perform the same techniques. The first technique that she analyzes is the introduction of counternarratives, which, by challenging dominant existing story, aims at distracting, disorienting and confusing people so that they do not know what to believe anymore (Rietdijk, 2021). The second category regards the discrediting of critics, which attempts at deceiving people's trustworthiness of anyone that holds an opposing opinion from the gaslighters. Here, Rietdijk highlights a difference between one-on-one gaslighting and post-truth rhetoric: while in the former case the gaslightees are often themselves the target of the discrediting, in the latter arises the aforementioned "duplicity of the victims", where the critics targeted, such as newspaper, scientists or institutions, do not correspond to the public. Lastly, the denial of more or less plain fact not only enables gaslighters to contradict themselves by confusing and questioning people' own understanding and experience through ambiguous communication, but it also dismisses politicians' responsibility for the expression of socially unacceptable attitudes – like sexism, racism, and

fascism (Rietdijk, 2021). Since "*each of these strategies of deception, disorientation, isolation and confusion used in post-truth politics resembles those used in traditional, one-on-one gaslighting*" (Rietdijk, 2021), the author explains the newer and more complex phenomenon that she has named "collective gaslighting", which is what we currently witness and are subjects to. Her work has been fundamental in analyzing, recognizing and resisting the threat to democracy, knowledge and, mainly, our epistemic autonomy that political gaslighting in our post-truth world poses to our society. By exposing how politicians can manipulate public opinion and create a false reality that serves their interests through disinformation and propaganda, the author sheds light on the dangers of a political climate in which truth is malleable and subjective and highlights the importance of maintaining a critical perspective and seeking out reliable sources of information in order to combat the effects of collective gaslighting.

However, many scholars have questioned whether such phenomenon has risen with the spread of social media or whether it has been a political tactic that has been in use before the latest technological revolution. For the purpose of this paper, therefore, it is essential to overview political manipulation techniques employed before the introduction of social media platforms and to understand whether collective gaslighting is correlated to the trust that the public has in social media.

Section 1.6 – Political Manipulation before the Technological Revolution

Although political gaslighting has been usually linked to a dystopian Orwellian-like society[11], influenced by the Second World War experience of totalitarianism and authoritarianism and the advent of technological devices, political manipulation – among which techniques we can witness also some gaslighting techniques - has been a pervasive feature of human societies for centuries,

---

[11] George Orwell in his book "*1984*" describes in detail a dystopian future society in which the government, headed by Big Brother, controls every aspect of people's lives through several political manipulation – but mainly gaslighting – techniques in order to maintain power over Oceania's citizens. The author, in 1949, depicts all the political gaslighting techniques that have become prevalent in our society, such as the concept of "doublethink", which entails the ability to hold two contradictory beliefs simultaneously and believe them both to be true. In this way the Party is able to control people's perceptions and understanding of reality and manipulate them into believing whatever they want them to believe – which results much similar to the aforementioned post-truth gaslighting method of counternarratives. Another characteristic common both to Orwellian society and current political gaslighting, is the rewriting of history: the protagonist himself works for the "Ministry of Truth", which job revolved around deleting and rewriting historical documents in line with the history that the Party had designed ("*The party told you to reject the evidence of your eyes and ears. It was their final, most essential command*"). This technique aimed at controlling people's understanding of the past and present, and ultimately their future, can be easily compared to Rietdijk's denial of more or less plain facts.

long before the advent of modern technology. Through various means, therefore, political leaders have sought to influence public opinion and gain power, often resorting to tactics that are deceitful, manipulative, and even violent. It is therefore fundamental to spot those historical political gaslighting examples in order to understand the ways in which social media have influenced the spread of such manipulation techniques.

In his book "*Tribes, Festivals and Processions; Civic Ceremonial and Political Manipulation in Archaic Greece*", W.R. Connor examines the role of civic ceremonial in political manipulation in ancient Greece during the Archaic period. Connor argues that festivals, processions, and other civic ceremonial events were essential tools for political manipulation, as they provided a means of reinforcing social and political hierarchies and legitimizing the power of the ruling elite. By analyzing the symbolism and ritual practices associated with these events, Connor demonstrates how the ruling class used them to construct a narrative of legitimacy and to reinforce their own power and authority. Fundamental was the role of religion, which, through the interpretation of the μύθοι or the Πυθία's oracles in Delphi, was exploited by the ruling elite in order to be associated with the gods and goddesses of the pantheon, so that they could portray themselves as divinely ordained and legitimize their rule. Another example of mythology's exploitation in Ancient Greece as a form of political manipulation is Alcibiades' δημαγωγία. According to the Merriam-Webster Dictionary's definition, a demagogue is "*a leader who makes use of popular prejudices and false claims and promises in order to gain power*"[12], as indicated by Aristophanes in the "*Knights*"[13], a demagogue is an unscrupulous master of slippery rhetoric who, for his own ends, plays the crowd like a cheap fiddle (Bartlett, 2020). Plutarch in his work "*Alcibiades*" describes Alcibiades as a demagogue, due to his known extravagant lifestyle and his willingness to do whatever it took to achieve his goals, including using deceit and manipulation to sway public opinion[14], by playing both sides of the political spectrum and switching alliances depending on what was more beneficial for him. His figure did indeed present both the characteristics of the gaslighter, such as his charismatic character and his ability to gain the trust of the δήμος no matter what, and of the post-truth politicians. He convinced the Athenians to launch a disastrous military campaign against Syracuse through Rietdijk's techniques of discrediting the critics – namely the other military leaders – and the creation of a counternarrative based on the need of such victory and the interpretation of

---

[12] https://www.merriam-webster.com/dictionary/demagogue
[13] "*demagoguery no longer belongs to a man acquainted with the things of the Muses or to one whose ways are upright / But to somebody unlearned and loathsome*" (Aristophanes in the "*Knights*")
[14] "*But the weakness which his tempters played upon most of all was his love of distinction (philotimia) and his desire for fame (philodoxia)*" (Plutarch in "*Alcibiades*")

an oracle. Moreover, when the campaign went awry and the Athenians suffered a crushing defeat, Alcibiades was quick to shift blame onto others, convincing the people of Athens that the failure was due to the incompetence of other military leaders, rather than his own poor judgment. This allowed him to maintain his influence and continue to hold political power. The Syracuse's defeat, in addition to his widely known rhetoric based on the distortion of the truth and misrepresentation of facts in order to gain support for his policies and proposals, proved his political gaslighting without the use of social media, but through the exploitation of Ancient Greece's political features *sui generis*.

Even in Ancient Rome, political leaders employed various forms of gaslighting to gain and maintain power, creating confusion and disorienting the public from reality. In particular, Roman politics witnessed more episodes of political gaslighting, due to the high competition and corruption characterizing the power positions and the hierarchical feature of the system which left the control of the government to a few wealthy and powerful families. One example of political gaslighting in Ancient Rome was the propaganda used by Julius Caesar during his campaigns. Caesar used manipulation tactics such as exaggerating his achievements and spreading false rumors about his opponents to sway the public's opinion in his favor. He also manipulated the truth about his reasons for crossing the Rubicon, claiming that he had no other choice but to take up arms against the Senate. We can see his political gaslighting techniques both in the myth that he made of his persona and in the books he wrote, namely "*De Bello Gallico*" and "*De Bello Civili*". In particular, by writing these highly notorious works, he highlights the same manipulation that post-truth politicians apply to recent forms of media, namely presenting his own narrative and discrediting the opponents in order to provide a perception of reality different from the actual facts that enabled him to manipulate the public opinion[15]. Also Augustus employed the same manipulative techniques of his predecessor. Firstly, he created a cult of his personality through an intelligent duplicity: if on the one hand he gained the favor of the public by presenting himself as a normal man – through the address of *princeps*, i.e. first among the others – on the other, he established his divine origins by the epithet *divi fili*, through the "divinization" of Caesar. Secondly, he was "*the supreme master of*

---

[15] Regarding his crossing of the Rubicon, he in fact writes: "*Quibus rebus cognitis Caesar apud milites contionatur. Omnium temporum iniurias inimicorum in se commemorat; a quibus deductum ac depravatum Pompeium queritur invidia atque obtrectatione laudis suae, cuius ipse honori et dignitati semper faverit adiutorque fuerit. Novum in re publica introductum exemplum queritur, ut tribunicia intercessio armis notaretur atque opprimeretur, quae superioribus annis armis esset restituta. Sullam nudata omnibus rebus tribunicia potestate tamen intercessionem liberam reliquisse. Pompeium, qui amissa restituisse vldeatur, dona etiam, quae ante habuerint, ademisse*".

the art of propaganda in the entire history of the West"[16], establishing an unprecedented manipulation of the literary elite, which consisted in a *unicum* throughout the entire history. Through the manipulation of the literature and of the political scene in Rome through his propaganda, he was able to promote the concept of *Pax Augustana* and establishing himself as the exclusive promoter of peace after the long political turmoil that Roman citizens had witnessed for the previous decades. Moreover, he repeated one Caesar-specific political gaslighting technique: while Caesar committed an illegal act by crossing the Rubicon to highlight the opponents' corruption and incompetence and then later justifying himself in *De Bello Civili*, Augustus did approximately the same thing when he publicly read Marcus Antonius' private testament in the Senate in order to depict him as oriental and to move the public opinion against him. However, the most evident political gaslighting example in Ancient Rome was the Great Fire of Rome and the reaction of Emperor Nero. The different narratives that had followed the *incendium magnum Romae*, whether it was Nero's accusation towards the Christians as a pretext to persecute them or Cassius Dio, Tacitus and Suetonius' blaming Nero himself, are the peak of ancient political gaslighting so that until today no one knows the truth about the events of that fire. In this episode we can see the efficiency of all aforementioned Ratdijk's techniques combined together and employed by the different figures involved, resulting in a complete confusion and disorientation of the public opinion and of the posterity, making everyone unable to make a judgment and attacking people's epistemic autonomy.

During the Middle Ages those in positions of power frequently employed political gaslighting to control their subjects. Monarchs and nobles utilized various tactics to manipulate their subjects, such as spreading false rumors, inventing threats, and employing propaganda to solidify their power. In order to justify the Church's military campaigns and maintain its authority, the Catholic propaganda during the Crusades depicted Muslims as a threat to Christianity. The Church also relied on fear and intimidation during the Inquisition to oversee the populace and uphold its authority, often accusing individuals of heresy and using torture to extract false confessions. The primary tactic of political deceit involved discrediting opponents through disseminating false accusations via propaganda. During the Investiture Controversy in the 11th and 12th centuries, the Holy Roman Emperor and the pope vied for control over the authority to appoint bishops. Both sides engaged in manipulation and propaganda to sway public opinion. The papacy accused the empire of engaging in simony (the sale of church positions) while the empire accused the papacy of corruption and misuse of power.

---

[16] According to an interview of the historian Tom Holland on Censorship magazine

The Renaissance presents a form of political gaslighting which is the result of two previous and already-tested manipulation techniques, namely the Medieval presentation of counternarratives and discrediting of the opponents through the spreading of rumors and the Augustan cultural propaganda. This was, in fact, the reflection of the Renaissance period, a time characterized by a rebirth of art, culture, and humanism, but also of political upheaval and power struggles mainly in Italy. In regard to the Medieval tradition, during the Renaissance political gaslighting often took the form of conspiracies and accusations of treachery, as seen in the case of the Pazzi Conspiracy in 1478, in which members of the Pazzi family attempted to assassinate the ruling Medici family of Florence. However, the most important political tool in this period consisted in the commission of artistic projects, recalling the phenomenon of Maecenas' patronage from the Augustan Age, through which the ruling families could manipulate the public perception of their power and authority. According to C. Jean Campbell in her book "*The Commonwealth of Nature: Art and Poetic Community in the Age of Dante*", during the Renaissance, rulers commissioned artworks that portrayed themselves as strong and just leaders, while depicting their enemies as weak and inferior. These depictions were often used to justify political decisions and actions, while also demonizing opponents. For example, the portrait of Lorenzo de' Medici painted by Agnolo Bronzino, depicts the humanist as a wise and just ruler, surrounded by classical symbols of learning and philosophy, even though in reality he was known for his ruthless tactics and suppression of political opponents. Also the commission of Michelangelo's David by the Republic of Florence is a clear example of political manipulation through artistic means. The statue was intended to symbolize the strength and bravery of the Florentine people, but it was also used as a political tool to assert the Republic's dominance over neighboring city-states. The role of art in the Renaissance politics was analyzed by Caroline Elam in her work "*Art and Diplomacy in Renaissance Florence*", which depicted art as a political mean both towards securing peace and with the aim of reinforcing the power and supremacy of the ruling families of the city-states. She claims that this interconnection between art and politics becomes impossible to ignore with Lorenzo de Medici's "*politique de prestige artistique*"[17], which made difficult to disentangle whether his artistic advice and commissions had to do with his own dynastic aspiration and political agenda or with the reputation and political end for the city of Florence (Elam, 1988).

The invention and proliferation of newspapers marked a pivotal moment in political manipulation. Prior to this development political elites relied on face-to-face communication as well as printed materials like pamphlets and posters to spread their message. However, with newspapers entering the scene they gained access to a broader audience enabling them to influence public opinion on a

---

[17] Elam here is quoting André Chastel's definition of Lorenzo de Medici political strategy.

larger scale and control political discourse. In his study "*The American Newspaper: A Study in Social Psychology*" Wilcox highlights how newspapers greatly impacted politics in late 19th early 20th century United States. He argues that newspapers possess significant power in shaping public opinion beyond just information dissemination. Instead, they have the ability to shape people's perspectives on specific issues by selectively reporting on them employing language that promotes certain viewpoints, and editorializing in favor of particular policies or politicians. Wilcox cites the case of William Jennings Bryan as an illustrative example to support his assertion. Bryan, a populist statesman who campaigned for president multiple times during the early 20th century, adeptly utilized newspapers to connect with working-class readership. The media coverage often emphasized Bryan's populist messaging and portrayed him as a champion of ordinary citizens' causes. Another demonstration of newspapers' influence on public opinion comes from the French Revolution, wherein periodicals like *L'Ami du peuple* and *Le Père Duchesne* played a vital role in radicalizing revolutionaries by endorsing violent rhetoric and attacking political foes. Corroborating this view, Popkins' article "*The Press and the French Revolution after Two Hundred Years*" suggests that newspapers frequently sensationalized events or presented them in ways supporting their own political motives through incendiary language designed to evoke strong emotions. Popkin contends that these tactics proved effective due to high levels of literacy among the French populace who heavily relied on newspapers for information dissemination. Consequently, understanding the historical diffusion of newspapers becomes crucial when contemplating modern instances of political gaslighting - techniques that have evolved alongside advances in mass media innovations throughout centuries.

Section 1.7 – From the Big lie to the big network

However, no form of political manipulation throughout history can be even remotely compared to the totalitarian gaslighting employed by the Nazi regime. We can easily argue that a whole branch of post-WW2 scholars has focused their careers on trying to identify the techniques and the efficiency of this form of manipulation that shaped an entire population into believing in the Nazi ideology.

The most effective gaslighting technique was the so-called "Big lie", or "*Große Lüge*" in German. Hitler himself in *"Mein Kampf"* defines the concept of the big lie, even though he described it as a form of propaganda used by the Jews and not as a manipulation technique that his regime would later impose. "*All this was inspired by the principle - which is quite true in itself - that in the big lie there is always a certain force of credibility; because the broad masses of a nation are always more easily corrupted in the deeper strata of their emotional nature than consciously or voluntarily; and thus in the primitive simplicity of their minds they more readily fall victims to the big lie than the*

small lie, since they themselves often tell small lies in little matters but would be ashamed to resort to large-scale falsehoods. It would never come into their heads to fabricate colossal untruths, and they would not believe that others could have the impudence to distort the truth so infamously. Even though the facts which prove this to be so may be brought clearly to their minds, they will still doubt and waver and will continue to think that there may be some other explanation. For the grossly impudent lie always leaves traces behind it, even after it has been nailed down, a fact which is known to all expert liars in this world and to all who conspire together in the art of lying. These people know only too well how to use falsehood for the basest purposes."[18]

Even – allegedly - Goebbels, the infamous Nazi Propaganda Minister, tried to elaborate on this by explaining that "*if you tell a lie big enough and keep repeating it, people will eventually come to believe it. The lie can be maintained only for such time as the State can shield the people from the political, economic and/or military consequences of the lie. It thus becomes vitally important for the State to use all of its powers to repress dissent, for the truth is the mortal enemy of the lie, and thus by extension, the truth is the greatest enemy of the State.*"[19]

The Nazi regime employed this method extensively demonstrating its efficacy towards the masses. The famous philosopher Hannah Arendt, actually, in her masterpiece "*The Origins of Totalitarianism*" explored the role that this technique had in shaping the masses' perception of the Nazi regime, defining it as the most significant propaganda tool employed by the Nazis. Many examples reveal how such gaslighting tactic worked. One example is the portraits of the Jews as a subhuman race responsible for Germany's economic woes and other social problems as opposed to the Nazis portrayed as defenders of traditional German values and culture in their relentless propaganda campaigns. Moreover, the use of the big lie was not limited to propaganda campaigns: it was also used to justify military aggression and conquest, like the 1939 invasion of Poland based on the principle of the *Volksgemeinschaft*, and to promote official governmental policies, for example by denying the existence of the concentration camps claiming that they were merely re-educational centers for political dissidents. However, Hannah Arendt provides an additional psychological factor, which really foresaw and anticipated nowadays post-truth world, namely that such a technique can be effective only in a society where individuals had lost faith in objective truth and were willing to accept whatever narrative was presented to them.[20]

---

[18] According to James Murphy translation, page 134
(https://www.sourcewatch.org/index.php/Big_lie)
[19] It is not sure whether this quote is attributable to Goebbels himself, but for a long time it was considered as such by the literature (https://www.jewishvirtuallibrary.org/joseph-goebbels-on-the-quot-big-lie-quot)
[20] In her 1951 "*The Origins of Totalitarianism*" she stated: "*in an ever-changing, incomprehensible world the masses had reached the point where they would, at the same time, believe everything and*

The concept of the big lie, as I shall argue, can be considered not only the pinnacle of political gaslighting ever employed - a fact made evident by numerous accounts highlighting the perceptions of the German people - but also the most comprehensive one, due to the integration of all Rietdijk's techniques under one main concept. The Nazis skillfully employed scapegoating, singling out Jews, homosexuals, communists, and other marginalized groups as responsible for Germanys myriad problems. This deliberate strategy instilled fear and distrust towards these minorities while simultaneously fostering an unwavering faith in the regime from the masses. Furthermore, spreading lies and misinformation through various channels - predominantly propaganda - was a method consistently used to sow confusion and cast doubt on objective truth. These measures were adopted with a singular purpose in mind: to perpetuate their control over power and manipulate public opinion. Added to this was their reliance on censorship and the menacing threat of violence as means to maintain dominance over public discourse and suppress any opposition that dared to challenge their authority.

Joseph Goebbels emerges as a mastermind in terms of utilizing propaganda to shape public perception. As elucidated by Kershaw in his book "*The Hitler Myth: Image and Reality in the Third Reich*", Goebbels seized every available medium or method possible to disseminate Nazi ideology, for example, notably grand-scale rallies were employed as a powerful tool for fostering a sense of unity and strength among Germans. These meticulously choreographed events featured expertly crafted speeches along with meticulously calculated displays of power designed to invoke both awe and loyalty among those present. Within this framework relentless messaging repetition through slogans played an integral part in reinforcing party doctrine. One particularly renowned example is "*Ein Volk, ein Reich, ein Führer*" ("One People, One Nation, One Leader"), which saw extensive usage within Nazi propaganda campaigns aimed at promoting an image of unified Germany under Nazi rule. Other slogans like "*Arbeit macht frei*" ("Work sets you free") were strategically employed to legitimize forced labor and concentration camps. The objective underlying this propaganda extended beyond seeking acceptance or spreading Nazi ideology alone; it also aimed at constructing a mythical figure around Hitler for consolidating and maintaining his grip on power through elaborate visual displays and extensive media coverage, as pointed out by Kershaw. In Kershaws' analysis, he also delves into the psychological and sociological factors that made Hitler

---

nothing, think that everything was possible and nothing was true… The totalitarian mass leaders based their propaganda on the correct psychological assumption that, under such conditions, one could make people believe the most fantastic statements one day, and trust that if the next day they were given irrefutable proof of their falsehood, they would take refuge in cynicism; instead of deserting the leaders who had lied to them, they would protest that they had known all along that the statement was a lie and would admire the leaders for their superior tactical cleverness."

such an effective propaganda tool - particularly his personal charisma and ability to resonate with the masses.

David Welch, another scholar shedding light on Nazi propaganda technique, in his book "*Propaganda and the German Cinema*", focuses specifically on their utilization of media. Welch argues that Nazis employed the film industry as a means to communicate their ideology and disseminate their message to broad audiences. Their effectiveness relied on a combination of censorship, subsidies, and coercion ensuring that films reflected Nazi ideology and served as effective tools for propaganda. Moreover, Welch posits that cinema proved exceptionally impactful due to its ability to manipulate emotions while fostering unity and identification with Nazi ideals. By utilizing film to cultivate collective identity while reinforcing Nazi ideology, the regime managed to retain support from German citizens while justifying acts of aggression and persecution. One prominent technique employed by Nazis in their films involved portraying Germany as a victim of foreign aggression. Exemplifying this type of propaganda is Leni Riefenstahls' renowned film "*Triumph des Willens*" ("Triumph of the Will"), which depicts the 1934 Nazi Party rally in Nuremberg; showcasing both Germanys' military prowess and citizens' unwavering loyalty towards Hitlers' leadership. Riefenstahl expertly used camera angles and editing techniques to create an imagery presenting Germany as a strong nation proud under Hitler's guidance - so much so that Susan Sontag defined it as "*a film whose very conception negates the possibility of the filmmaker's having an aesthetic conception independent of propaganda*"[21]. It is important to acknowledge that Hitler's mystification received significant support from the film industry at that time - movies like "*Hitlerjunge Quex*" relentlessly portrayed him as the epitome of discipline and loyalty effectively symbolizing traditional German values.

Although the profound impact of 20th century totalitarianism raised people's awareness about such propaganda techniques, it became evident during the Cold War that political gaslighting was an intrinsic part of mutual discrediting and misinformation campaigns between the United States and the USSR. Both nations employed diverse tactics to demonize one another: while the U.S. utilized propaganda films to label the Soviet Union as a totalitarian regime posing a threat to democracy and freedom, the Soviet Union presented themselves as advocates for socialism and championed working-class rights, simultaneously portraying America as an oppressive imperialist power. In addition to these conventional propaganda techniques employed by totalitarian regimes earlier, political gaslighting achieved a significant triumph during this period through scandals involving the U.S administration. The Watergate scandal serves as a prime example of this phenomenon with

---

[21] As reported by Ala Sennett in "*Film Propaganda: Triumph of the Will as a Case Study*" (2014)

Nixon and his administration employing various gaslighting strategies to manipulate public perception in their favor. One prominent tactic was their consistent denial of any wrongdoing, dismissing allegations as nothing more than politically biased witch hunts. Furthermore, they launched attacks on the credibility of press organizations and individuals who uncovered evidence of malpractice, seeking to taint their reputation by branding them as biased or unpatriotic. Additionally, contradictory public statements made by Nixon himself and his aides further contributed to confusion among citizens regarding the facts surrounding this case. Discrediting witnesses possessing incriminating information was another favored approach for Nixon's team - either through character assassination or undermining their testimonies by questioning reliability. All these tactics were aimed at creating doubt and uncertainty among citizens in order to obstruct any form of accountability for their actions. In advancing these political gaslighting tactics, Nixon's administration relied heavily on various media channels: leaking fake news stories to manipulate press coverage, launching direct attacks on journalists themselves to discredit their work, or strategically crafting the president's image through staged press conferences and meticulously planned interviews for television broadcasts.

As I shall argue, all the aforementioned political gaslighting techniques, which were applied to public debates, propaganda, newspaper, press, television etc. have led the way to the employment of these tactics in the context of social media, which, however, have a much broader spectrum of audience.

The technological revolution has transformed society in many ways, and one of the most significant changes has been the rise of social media. This new form of communication has brought about significant changes in society, particularly in terms of how people interact, share information, and engage with political issues. Scholars have closely examined the impact of social media on political participation, the dissemination of political messages, and the effectiveness of political communication. The widespread adoption of platforms like Facebook, Twitter, Instagram, and YouTube has disrupted traditional political processes and structures, in such a way that politicians and political organizations have utilized social media to revolutionize the nature of political communication. In "*Social Theory after the Internet: Media, Technology, and Globalization*" Ralph Schroeder analyzes how the digital revolution has transformed both mass media and politics focusing on the shift of political communication from traditional media to the network system provided by social media. He emphasizes that traditional media outlets like newspapers and television had limited space for political discourse due to hierarchical structures maintained by journalists and editors. On the contrary "*digital media are an autonomous subsystem, a transmission belt between citizens and elites in the political process*" (Schroeder, 2018), which

decentralization allows anyone with internet access to create content and reach a global audience. The ability for politicians to instantly broadcast their message worldwide has fundamentally changed how they communicate with citizens on a global scale while bypassing traditional media channels. A key aspect that affects how political elites use social media is a pivotal dual distinction: on one hand, it allows political leaders to reach a wider audience and directly engage with citizens, making it easier for them to spread their agendas and narratives, on the other, they need to face the direct responsiveness to their agendas and the issue of the counterpublics, who challenge their vision via media (Schroeder, 2018). This means that while politicians can manipulate public opinion through this direct relationship, they are also susceptible to influence from the public. Thanks to social media platforms citizens can connect with others who hold similar beliefs and come together around issues that matter to them, for example through hashtags, viral videos, and online petitions grassroots movements have been able to mobilize and exert pressure on the political elite.

However, this is not the only implication that the global reach of social media has shown throughout these years, indeed, the pivotal point is how foreign actors can influence politics both on a domestic and an international level. In this resides the main difference between political gaslighting before the technological revolution and after the advent of social media: while all the other forms of media manipulation I have analyzed throughout history had always had a domestic audience – even though their consequences often affected the dynamics of the international system - the global reach of social media enabled actors to manipulate the public opinion transcending the "sovereign-territorial" boundaries. For this reason, the rise of social media shall be the main field to investigate the ways through which political gaslighting can influence international relations, where single politicians and actors can share their own narrative not only within their own country, but with the whole world, re-shaping, therefore, the international order.

Chapter II

As we have seen in the previous chapter, the rise of social media, and, mainly, their global reach enabled every actor to manipulate public opinion transcending the "sovereign-territorial" boundaries, being able to influence an international audience. This has severe consequences for the international order, due to its facility to shape the narratives and norms that lie behind the international system. Therefore, through the use of social media, there are an infinite number of possibilities regarding both the identity of the actors motivated to manipulate social media, the identity of the actors subject to such manipulation, and the ways in which such manipulation occurs. In this possibility lies the pivotal change between traditional and social media, which is the basis for understanding how such phenomenon impacts the international order and its complexity due to the inconceivable high number of actors involved. Referring to Zeitzoff's work analyzing social media influence in inter-state conflicts, traditional communication models were founded on a simple distinction: tactical communication referred to the transmission of information elite-to-elite, while mass communication involved the ways in which the political elites spread information to the masses. However, while each traditional type of communication considered the masses as mere passive receivers of elite communication (Scheufele, 2000), social media continuously "*challenges political and elite-dominated paradigm of media's effect on political behavior*" by allowing users to create and share "*their own content and directly respond to elites and leaders*" (Zeitzoff, 2017). The high complexity in analyzing the impact on the international order of political gaslighting through social media is a result of both the extensive reach of the latter and of the difficulty in its detection and assessment. Therefore, it is necessary to try to untangle this phenomenon through a multi-faceted approach, separately investigating its different aspects.

Therefore, I shall try to give a comprehensive overview of this phenomenon, in order to understand its effectiveness and to assess the threat it poses on the international order. For the purpose of this thesis, I shall firstly examine the ways in which political gaslighting is employed through social media. Hence in this section I shall explain the techniques of social media manipulation used for achieving political goals – namely the use of bots and fake accounts, coordinated action networks, and sentiment manipulation - and their impact on democracy and the international order. After having delineated "how" such manipulation is employed, another section shall be entirely devoted to the phenomenon of digital disinformation and fake news through social networks, unveiling the content of political gaslighting carried on through social media and focusing on its understanding under an international relation lens and its distinction from misinformation. Secondly, I shall take into account the main actors involved in political manipulation on these platforms, starting from distinguishing the most common and stereotypic scenarios: 1) political elites or state agents that aim

at manipulating foreign domestic politics; 2) political elites or state agents that target international institutions or transnational relations entities; 3) political elites or state agents that intend to influence the international public opinion; 4) transnational relations entities that aim at influencing foreign political actors; and 5) transnational relations entities that want to shape the international public opinion. Thus, in order to grasp it in the most straightforward and unequivocal manner, I will distinguish political manipulation through social media both perpetrated and endured by these two main actors, namely state agents and transnational relations entities, focusing in the latter on international social movements and transnational terrorism. I shall consider them as both subjects and objects to social media manipulation because while on one hand they have a proactive role in trying to shape the narratives according to them, other actors may employ the same techniques to undermine their manipulation. However, in order to assess in a thorough form such phenomenon, I shall also mention and explore the multi-channel relationships between governments and private companies owning social networks platforms, investigating how the economic field is involved with global political gaslighting. In particular, I shall examine the complex scenario due to the overlapping of interests involving multinational companies, state agents and international institutions, focusing on the Big Tech's attempts and responsibility to regulate their own platforms. Lastly, in order to assess the impact of political gaslighting through social media manipulation, I shall delve into the attention and awareness that states and international institutions have devoted to such phenomenon, highlighting how these agents have tried to prevent and combat it.

Section 2.1 – "Toxic Data": how political manipulation is employed through social media

In Chapter I, I have analyzed the three forms of political gaslighting identified by Natascha Rietdijk, namely the introduction of counternarratives, the discrediting of critics, and the denial of more or less plain facts, which are all strategies that are constantly employed by political actors in the social media realm. However, Rietdijk's work remains on a theoretical and psychological level, meaning that she elaborates her theory without analyzing the means through which the manipulation techniques specified are employed. In this context, the work of David Chavalarias provides a thorough insight into the various ways in which political manipulation can be employed and spread through social media. He starts his work by acknowledging the global dimensions of social networks, which "*have thus become a gigantic marketplace of influence, where millions of messages and reactions are exchanged every second, and where we are all both targets and potential vectors of influence*" (Chavalarias, 2022), making every citizen an actor that is both subject and object to media influence. Due to this intrinsic feature of social media, namely collective accessibility, he argues that "*today, more than ever, our societies are under the threat of*

a tsunami of information that is overwhelming our cognitive capacities and our democratic institutions" (Chavalarias, 2022), demonstrating how such political manipulation constitutes a threat for democracy and critical thinking. In order to address this issue, Chavalarias assumes that due to the fact that social media networks have made it easier for politicians to connect with their constituents, mobilize support, and disseminate their messages, this increased connectivity and access enabled them to become a powerful tool for manipulating public opinion. Therefore, he explores three forms of political manipulation through social media, focusing mainly on politicians as subjects and demonstrating how easily and effectively they can influence public opinion.

The first form of social network manipulation discussed by Chavalarias is the use of bots and fake accounts, which "*are automatic programs or users that behave like human beings, with the ability to post, like, or comment on social media, as well as interact with other users or groups. Their purpose is to spread propaganda or fake news, hijack trending topics, create echo chambers, and manipulate public opinion*" (Chavalarias, 2022). Hence, these are automated accounts that can post and interact with other accounts without human intervention. Bots can amplify political messages or support groups, creating the illusion of widespread support. Fake accounts, in contrast, spread propaganda, false information, or harass political opponents. Bots and fake accounts are frequently employed to manipulate social media algorithms, resulting in increased visibility and engagement for specific content, as Chavalarias observes. One consequence of the use of bots and fake accounts is the distortion of public opinion. Political actors can manipulate public opinion by spreading false or misleading information on social media, creating a false perception of widespread support or opposition for a certain issue or candidate. This can produce a chain reaction, as actual users of social media are more inclined to share and interact with content that seems popular or widely embraced. The use of bots and fake accounts, therefore, can lead to amplifying extreme or fringe views, creating the impression that they are more widely held than they actually are. This is even more evident and effective since it exploits social media's algorithms themselves: being hyper-personalized, they propose to the user content that is specifically customized for their interests and opinions, therefore bots and fake accounts reinforce their views by showing wider support. In fact, bots and fake accounts can distort public opinion by artificially boosting the followers, likes, and shares of a specific account or post. This can create a misleading perception of a political actor or party's popularity or influence, further distorting the public's understanding of political reality. In "*Algorithms, bots, and political communication in the US 2016 election: The challenge of automated political communication for election law and administration*", Howard, Woolley, and Calo contend that these algorithms can create the appearance of a vast online following for a politician who is, in fact, an outsider with no real-world support (Howard et al., 2018). Furthermore,

bots and false accounts contribute to spreading disinformation and propaganda, eroding trust in democratic institutions. These consequences are prevalent in all the forms mentioned by Chavalarias and will be analyzed later.

The second form of social network manipulation discussed in "Toxic Data" involves coordinated action networks. These are groups of individuals who work together to promote a particular message or agenda on social media. Since coordinated action networks can be used to spread propaganda, attack opponents, or mobilize supporters, Chavalarias observes that these typically consist of individuals who feel strongly about a specific issue. However, they can also be manipulated by politicians or other groups to advance their own agenda. In fact, Chavalarias, in an interview on CNRS News about disinformation strategies employed on Twitter by climate skeptics, explained how such minority groups "*manage to persuade a significant proportion of the population to believe in pseudo-facts that run contrary to the scientific consensus and to people's perceived experience year after year*"[22]. He argued that their minority status is compensated by a strong presence online, enabling them to wage a "communication warfare" that effectively sows doubt. As evidence for such claims, in fact, he highlights how climate change denialists are much more active on social networks than those who defend the scientific consensus on climate change. Therefore, when coordinated action networks amplify certain messages and silence others, an unbalanced public discourse arises that, in extreme cases, can lead to the creation of echo chambers, where individuals are only exposed to information that supports their existing beliefs and opinions. This phenomenon has many intricate and multifaceted outcomes, like the rise of public opinion polarization, where individuals become more extreme in their views and less receptive to engaging with those who possess differing opinions. However, the primary consequence is, as previously emphasized, the exploitation of individuals by political actors who aim to manipulate public opinion. By targeting content to specific groups of individuals, in fact, politicians and interest groups can amplify their messages and create an illusion of consensus, even if that consensus does not actually exist. This can have a corroding impact on democratic processes, particularly if misinformation or propaganda is being spread. The consequences worsen when this phenomenon is applied globally through social media. Individuals, protected from opposing views and only exposed to information that supports their existing beliefs, become resistant to compromise and collaboration with those who hold different beliefs, resulting in discriminatory behaviors.

The work "Toxic Data" delves into the third form of social network manipulation, known as sentiment manipulation, which refers to the practice of artificially shaping the emotional reactions of individuals on social media platforms for political purposes. This technique relies on the idea that

---

[22] https://news.cnrs.fr/articles/investigating-climate-sceptics-disinformation-strategy-on-twitter

individuals are more likely to act on emotions rather than rational thought. As a result, political actors can manipulate public opinion by influencing people's emotional responses on social media platforms. Sentiment manipulation can be achieved by using language, imagery, and other forms of content that evoke particular emotions through selectively emphasizing events or issues or manipulating conversation tones. For example, Chavalarias notes that political actors can use social media to amplify negative emotions such as fear or anger towards a particular group or issue, aiming to rally support for their cause, or to diminish them such as downplaying the significance of a scandal or adverse event. Sentiment manipulation, like coordinated action networks, can polarize public opinion, causing people to become more entrenched in their emotional responses and less willing to engage in rational debate or compromise. Since sentiment manipulation amplifies emotions, hindering rational discourse, people become more likely to be swayed by their emotional reactions than the facts. Moreover, sentiment manipulation can also contribute to the spread of conspiracy theories and disinformation, as people become more receptive to emotional appeals that lack factual evidence. This is particularly problematic in the context of international relations, where the spread of false or misleading information poses a significant problem for peace and stability. For example, sentiment manipulation campaigns have fueled conflicts between ethnic and religious groups, while also undermining the legitimacy of democratic institutions. One instance of this technique occurred during the 2018 Italian general elections, where the far-right party, the Lega, used social media to disseminate propaganda and manipulate public opinion. The party's leader, Matteo Salvini, particularly adept at using social media platforms such as Facebook and Twitter to spread his message, targeted specific groups such as young males and those with an interest in Italian football. Hence, the party employed the language and symbols linked to football to appeal to this demographic, crafting memes and other visual content that resonated with their interests which resulted in 17.4% of the votes in the general elections.

Every form of social media manipulation described by Chavalarias encompasses the three forms of political gaslighting depicted by Rietdijik: both the use of bots and fake accounts, coordinated action networks, and sentiment manipulation can be easily employed to spread disinformation, discredit political opponents, and distort narratives that do not align with objective reality. However, while the purposes of these techniques may seem different, all of them constitute a massive threat to worldwide democracy and the international order.

Section 2.2 – Social media manipulation as a threat to democracy and the international order

In regard to the threat to worldwide democracy, political manipulation employed through social media is one of the main risks of our generation, as it undermines the public's ability to make

informed decisions and creates an environment in which disinformation and polarization thrive. Social media platforms have, in fact, become an appealing target for political actors who want to manipulate public opinion due to their broad reach and low cost. As previously anticipated, bots and fake accounts are commonly used in social media to manipulate information and control the narrative by amplifying selected messages while silencing others. An analysis of organized social media manipulation by the Oxford Internet Institute, "*Industrialized Disinformation: 2020 Global Inventory of Organized Social Media Manipulation*"[23], found that 76 out of 81 countries surveyed between 2019 and 2020 suspected political parties or governments of manipulating public opinion via social media. In the case of the 2016 U.S. presidential election, it was found that Russian operatives used social media to create false accounts and amplify divisive messages, in an effort to sow chaos and confusion. As well as creating fraudulent accounts and pages, the Russian Internet Research Agency (IRA) orchestrated demonstrations and gatherings using these strategies in order to "*spread distrust toward the candidates and the political system in general*" as part of the 'Translator Project' (Graff, 2018[24]). As a result, informed decisions were hindered by rampant disinformation during the election. Regarding coordinated action networks, they can be utilized to generate the illusion of widespread endorsement for a specific candidate or issue. This can sway public opinion and create the perception of a broad-based movement. The report highlighted the impact that coordinated action networks had in the case of the Brexit referendum, which both the Leave and Remain campaign employed to influence public opinion. The Leave campaign made use of these networks to create the illusion of widespread support for their cause, while the Remain campaign manipulated them to instill fear and urgency in voters' minds. Another way in which social media manipulation poses a threat to democracy relates to the formation of echo chambers. These confine individuals within a bubble where they encounter information and viewpoints that align with their existing beliefs. Consequently, alongside the sentiment manipulation, this phenomenon can engender polarization and informed electorate's susceptibility to emotional appeals and misinformation instead of rational discourse. The exacerbation of this issue lies in social media algorithms that prioritize content based on user preferences and engagement patterns leading to a feedback loop that reinforces existing biases and restricts exposure to perspectives. The implications of these three forms of social media manipulation for democracy in international relations are profound. Disinformation campaigns breed discord and erode trust in democratic

---

[23] https://www.ox.ac.uk/news/2021-01-13-social-media-manipulation-political-actors-industrial-scale-problem-oxford-report

[24] Here the Report cites the article "*Inside the Mueller Indictment: A Russian Novel of Intrigue*" written by Garrett M. Graff (https://www.wired.com/story/inside-the-mueller-indictment-a-russian-novel-of-intrigue/)

institutions, while echo chambers hinder voters' ability to make well informed decisions based on reliable information. Moreover, bots and other automated accounts distort discourse making difficult to distinguish genuine public opinion from artificially generated content. In culmination all these manipulative tactics foster an environment inclined to authoritarianism and demagoguery, enabling actors to exploit social media platforms for public opinion manipulation and power consolidation.

In order to assess the threat that political manipulation through the use of social media poses on the international order, Bradshaw and Howard in their article "*The Global Organization of Social Media Disinformation Campaigns*" examine how low barriers to entry and reduced communication costs afforded by technology have altered the balance of power in the international system, through "soft power" and persuasion, framing and agenda setting, ideological hegemony, symbolic power, or sharp power to achieve desired outcomes (Bradshaw and Howard, 2018). Referring to the Computational Propaganda Project's 2017[25], the authors classified the countries surveyed by their regime type, focusing on the modal actors that employed such techniques, the level of formal organization of these cyber-troops, their level of capacity, and their modal targets. Particularly interesting, they found out that while in democracies the main actors were political parties with domestic targets, in authoritarian states the primary modal actors were the governments themselves targeting both domestic and foreign audiences. While these findings will be analyzed further in depth in the next chapter, Bradshaw and Howard's work is fundamental to understand the global impact of political manipulation through social media, since such phenomenon has consequences on modern peace, conflict, trade, diplomacy, and a myriad of other international processes (Bradshaw and Howard, 2018). Therefore, they argue that due to the proliferation of examples of state actors using and abusing social media to achieve political goals, it is necessary to adopt a global perspective of such phenomenon in order to understand the new nature of cyber power and computational propaganda as essential tools for cyberwarfare. Moreover, the threat to the international order is strictly linked to the one posed on the democratic process and institutions in other countries. Disinformation campaigns, often sponsored by foreign actors, aim to spread false or misleading information to influence the opinions and decisions of foreign populations. Bradshaw and Howard argue that these campaigns are often designed to exploit existing societal divisions and amplify fringe voices, creating a climate of polarization and distrust that can undermine the credibility of democratic institutions (Bradshaw and Howard, 2018). The impact of these campaigns was evident in the use of fake news and propaganda in India to incite and exploit communal

---

[25] https://demtech.oii.ox.ac.uk/wp-content/uploads/sites/12/2017/06/Casestudies-ExecutiveSummary.pdf

tensions, where "*manipulated content, disinformation, and misinformation plague the online environment*" (Freedom House, 2019). For example, in 2018, a series of fake news, stories and doctored videos circulated on WhatsApp, claiming that child kidnappers were on the loose in several Indian states. This led to several incidents of mob violence, with innocent people being attacked and even killed by vigilante mobs, inciting Hindu-Muslim tensions. Furthermore, foreign governments may use social media manipulation to exacerbate existing conflicts both within a foreign country and between foreign nations, spreading disinformation and propaganda. State-sponsored actors, such as Russia and China, have been accused of using such techniques to manipulate public opinion in other countries. For example, in the aftermath of the Skripal poisoning case in the UK, the UK government found a 4,000% increase in the spread of propaganda from Russia-based accounts since the attack - many of which were identified as automated accounts - to discredit the UK government and undermine the international response to the attack. The use of social media manipulation in international relations also has implications for the stability of the global order. In some cases, disinformation campaigns can be used to influence the policies and decisions of foreign governments, undermining the sovereignty of the targeted state. As previously mentioned, the most striking the case was the 2016 U.S. presidential election, where Russian state-sponsored actors reached 126 million people on Facebook, 20 million on Instagram, 1.4 million on Twitter, and uploaded 1,000 videos to YouTube (DiResta et al., 2018). This type of interference is not limited to the U.S. or the Russian context, as other countries have also experienced attempts by foreign actors to interfere in their political processes. For instance, a Singaporean Government Agency – namely the SG101 – engaged in a project to raise awareness against foreign interference and hostile information campaign after Singapore was a victim of "*an artificial impression to netizens of the opposition to Singapore's position*" (Senior Minister of State for Law Edwin Tong in an interview on 12 February 2019[26]) during their maritime and airspace disputes with Malaysia. Furthermore, it is important to acknowledge the threat of cyber-attacks, which possess the ability to destabilize political processes, pilfer confidential data, and incite disorder. Cyber-attacks become significantly more precarious when aimed at infrastructures, like energy grids or communication systems as the ramifications can extend beyond immediate repercussions. A poignant illustration of manipulation is exemplified by the cyber-attack on the Ukrainian power grid in 2015. This audacious breach resulted in power failures and a severe disruption of communication systems further exacerbating tensions, between Ukraine and Russia.

---

[26] https://elevenmyanmar.com/news/singapore-parliament-was-told-about-curious-spike-in-online-comments-critical-of-singapore

After having reviewed the modalities and consequences of social media manipulation for democracy and the international order, I shall investigate what is the content of the previously analyzed techniques. In this sense, I shall proceed by disentangling the common denominator of social media manipulation, namely the possibility to share disinformation.

Section 2.3 – Digital disinformation: what is the content of political gaslighting through social media

Imagine being one of the 293 million followers on Instagram of the worldwide famous singer Justin Bieber, connecting to one of his live streams, and seeing him carelessly adjusting his hat. Now, the majority of people wouldn't even notice such a common gesture, so how did it explode virally on every social network reopening one of the biggest conspiracy theories of the last decade that involved the US Democratic party, including the ex-presidential candidate Hilary Clinton? Allegedly, among the myriad comments during Bieber's live, one user challenged the singer to touch his hat if he was a child sex trafficking sufferer, making the "Pizzagate scandal" viral again and its followers more fervent than ever. But how can a conspiracy theory involving the most prominent political figures of the US, the Turkey government and a pizzeria in Washington DC become so spread that ended up in a shooting and on the frontline of major news media outlets? During the 2016 US elections campaign, some emails of John Podesta, Clinton's campaign manager, were leaked, among which one talked about a possible fundraising event at the pizzeria Comet Ping Pong. This "evidence" then started to spread on the controversial website 4Chan, where many began to consider this pizzeria as the headquarters of a Satanic global child-trafficking ring controlled by the Democratic Party. All it took was a false document posted on this website that the theory rapidly crushed a far-right extremist section of Reddit, and then every other social media platform. Thanks to the high involvement of Turkish President Tayyip Erdogan's supporters in promoting and virally spreading such news in order to accuse the US politicians of hypocrisy, on December 4, 2016, Edgar Maddison Welch burst into the pizzeria and started shooting, luckily not harming anyone. However, this episode did not succeed in stopping people from sharing such fake news, including also Michael Flynn Jr., the son and chief of staff of Trump's former national security advisor Ret. Lt. Gen. Michael Flynn, and many other online followers.

Among all the techniques of social media manipulation and political gaslighting analyzed in the previous sections, the *fil rouge* at the core of their effectiveness is the possibility to share any kind of information rapidly and easily, thus resulting in what many have defined as "*tsunami of information*" (Chavalarias, 2022), "*information pollution*" (firstly coined by Jakob Nielsen in

2003), "*streams of information*" (Wardle & Derakhshan, 2017), "*information overload*"[27], "*information fatigue*" (Cooper, 2019), and many other terms. However, regardless the different terminology and nuances, all these notions refer to the same main concept, namely, as investigated by Wardle and Derakhshan[28], an "*outcome of information revolution, wherein people are supplied with contaminated information, which is of less importance, irrelevant, unreliable and unauthentic, which lacks exactness and precision, which always has an adverse effect on society at large*" (Pandita, 2014). And, as explained in Section 1.3, 1.4, and mainly 1.5, its effects are the same of gaslighting in the post-truth era: people, due to the insane amount of information without fact-checking to which they are exposed constantly, have lost the ability or/and the interest to investigate the truthfulness of the facts, becoming passive receivers to whatever they want to believe. Therefore, it is not surprising that this perpetual stream of information, combined with people's unwillingness to spend time to fact-checking, their heuristic bio-psychological approach, and the intention of many actors to manipulate public opinion for their own goals, leads to an "*information war*", as defined by Claire Wardle in "*Fake News. It's Complicated*". Therefore, information, or – better - the manipulation of information, not only consists in the content of political gaslighting, but it also becomes one of the main sources of power in the international landscape. However, before delving into its links with political gaslighting and international relations, it is necessary to delineate some aspects of its terminology that, due to the common usage of these words, are often misconceived.

Firstly, we need to point out that disinformation, misinformation and fake news are all different aspects of the information ecosystem, that, although they do share "falseness" in common, have some distinct characteristics. While misinformation is generally defined as "*the inadvertent sharing of false information*" (Wardle, 2017), the terminology for disinformation and fake news was subject to different studies and interpretations. For example, in order to address the lack of consensus regarding this issue, Edson C. Tandoc Jr., Zheng Wei Lim and Richard Ling examined how the term "fake news" has been used in the academic field, delineating different typologies of fake news based on its previous application. However, although they extensively provide an insight into the multifaceted aspects of fake news, they do not postulate a synoptic definition of this notion. Therefore, I shall refer to Glenn Anderau's work "*Defining Fake News*", in which, after having overviewed several existing philosophical definitions on this issue, he manages to comprehensively resolve the question. He defines fake news as "*misleading information intentionally published and*

---

[27] It firstly became popular in the 1960s within sociology and psychology's scholars, only later it was applied to the context of social media

[28] They do not provide a concise definition of this phenomenon, but, after having highlighted the need for a specific terminology, they extensively explored all the aspects involved in this notion

presented as news which has the function of deliberately misleading its recipients about its status as news" (Anderau, 2021), including both lying and bullshitting in a Frankfurtian sense[29]. Moreover, he reinstates Fallis's definition of disinformation as "*misleading information that has the function of misleading someone*" (Fallis, 2015), distinguishing its main three properties, namely that it is a type of information, that is misleading and "*likely to create false beliefs*", and that it is nonaccidental (Fallis, 2015).

However, since neither Anderau nor Fallis managed to provide a full-fledged categorization of this phenomenon, I shall consider the distinction of MDI[30] depicted by Wardle, even though the "satire or parody" kind lies on the border of Fallis' definition. In her article, in fact, she distinguishes 7 different types of MDI based on the degree of the intention to deceive (Figure 1):

1. Satire or parody: this type involves intentionally humorous or satirical content that may be mistaken for genuine news. Satire acts as a way to comment on and criticize, but its potential to mislead requires the audience to think critically and discern the difference[31].

2. False connections: they involve the linking of genuine content to unrelated events or entities, creating misleading associations. By taking advantage of existing trust or familiarity, false connections manipulate perceptions, generating false narratives.

3. Misleading content: it includes inaccuracies or distorted information that can mislead audiences. It often involves selective factual presentation, contextual omissions, or cherry-picked evidence, manipulating perceptions and distorting public understanding.

4. False context: it entails manipulating content by placing it in a false or misleading context. It may involve misappropriating images or quotes to support a distorted narrative or mislead audiences about the intended meaning.

---

[29] In his philosophical work "*On Bullshit*", Harry Frankfurt investigates the difference between liars and bullshitters and its effects in contemporary society, highlighting that "*someone who lies and someone who tells the truth are playing on opposite sides, so to speak, in the same game. Each responds to the facts as he understands them, although the response of the one is guided by the authority of the truth, while the response of the other defies that authority and refuses to meet its demands. The bullshitter ignores these demands altogether. He does not reject the authority of the truth, as the liar does, and oppose himself to it. He pays no attention to it at all. By virtue of this, bullshit is a greater enemy of the truth than lies are*" (Frankfurt, 2005).

[30] With "MDI" I refer to the subgroup of the information ecosystem that encompasses all the forms of mis- and dis- information.

[31] For example, the newspaper "*China's People's Daily*" republished an article from the satirical journal "*The Onion*", in which the authors asserted that North Korea's Kim Jong Un was voted 2012's sexiest man alive.

5. Imposter content: it involves making representations where individuals or entities impersonate others, mimic sources or create fake accounts. The goal of content is to deceive audiences by leveraging credibility or authority associated with reliable sources.

6. Manipulated content: it refers to the alteration of genuine information, such as photos or videos, through editing or digital manipulation techniques. The intention behind this is to deceive viewers by presenting manipulated content as genuine.

7. Fabricated content: this can range from purposefully created news stories to doctored images or videos, exploiting digital tools to engender false narratives and mislead audiences.

Figure 1: Types of MDI
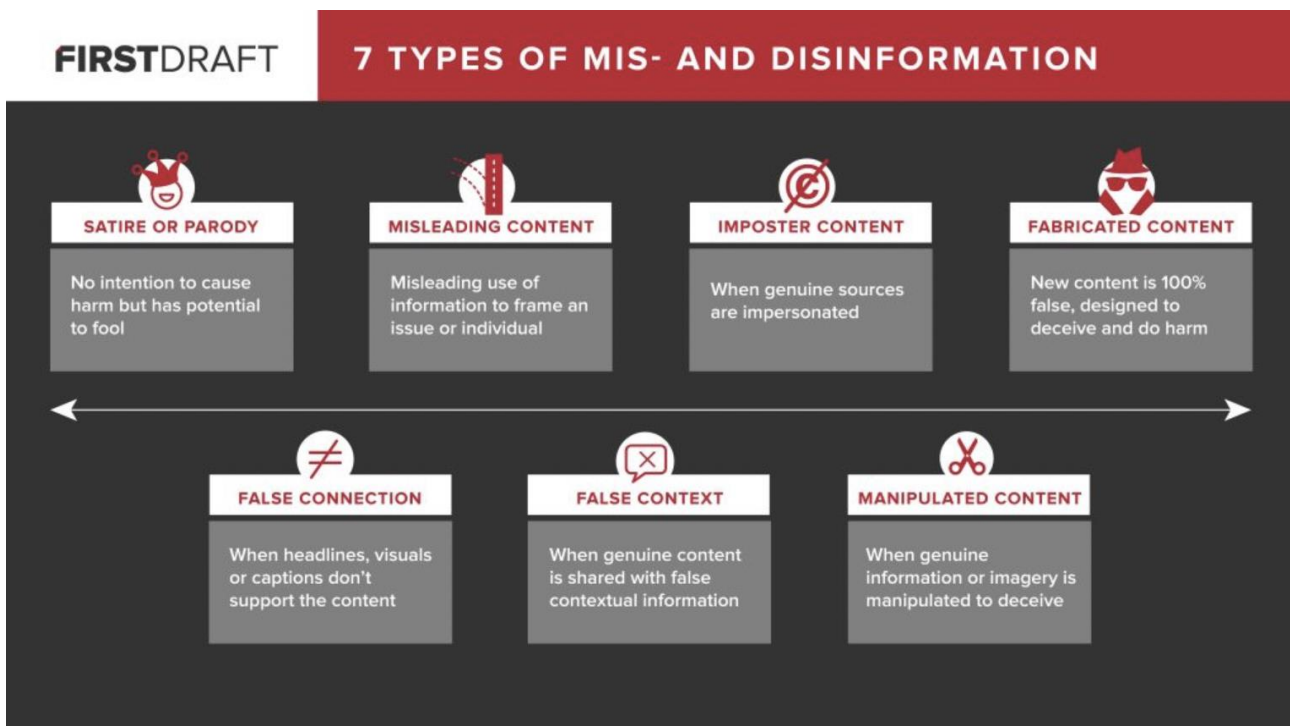
As the description of Wardle's seven types of MDI suggested, they do also have a fundamental visual component, that can easily enhance their effectiveness in misleading the audience. Highlighting the role of visual content in the MDI system, Nausicaa Renner explained that "*the fake news conversation has taken place in the realm of words, but that's missing a big part of the story. Much of the content that circulates on Facebook are images, often memes. They're not attached to an article, and there's often no way to trace their source. And while Facebook's algorithm is notoriously elusive, it seems to favor images and video over text. As such, images have the potential to reach more readers than articles — whether fake, real, non-partisan or hyper-partisan*" (Renner, 2017). In fact, due to our neurological predisposition to process images way faster than words, any visual subject has an essential impact on spreading whatever message through social media. For this reason, Wardle and Derakhshan inserted this component while delineating the four characteristics that make a message more likely to be consumed, processed, and shared widely – namely the

provocation of an emotional response, the presence of a powerful visual component, the address of a strong narrative, and an extensive repetition.

Now that I have examined the content of social media manipulation, it is necessary for the purpose of this thesis to understand digital disinformation in the context of political gaslighting and international relations. As we have previously examined, "*post-truth politics is not just an expression of disregard for truth or an origin of false beliefs; it can also undercut our own faith in our capacity to distinguish between truth and falsehood, between reliable and unreliable sources*" (Rietdijk, 2021), therefore the analogy between political gaslighting and digital disinformation springs to mind. The introduction of counternarratives, the discrediting of (potential) critics, and the denial of more or less plain facts (Rietdijk, 2021) cover all the implications of digital disinformation that we have analyzed previously in this section. The introduction of counternarratives is exactly what lies at the core of the aforementioned information fatigue and information overload, referring to the effects of being constantly exposed to a plethora of information shared online. As the information ecosystem becomes more saturated with contradictory narratives, individuals face a cognitive overload that hampers their ability to discern truth from fiction, half-truths, or manipulated facts. The constant stream of information on media platforms further intensifies this phenomenon, bombarding people with an unending amount of content. In this state of cognitive saturation, individuals may resort to heuristic thinking[32] – a cognitive shortcut that leads them to accept information that aligns with their preexisting beliefs while dismissing anything that contradicts their perspectives. Political gaslighters capitalize on this psychological tendency, skillfully crafting counternarratives that resonate with their target audience. The introduction of counternarratives, therefore, exploiting the pre-existing tsunami of information, creates a sense of epistemic chaos and undermines the foundations of objective truth. The continuous flow of conflicting information not only desensitizes individuals to the importance of fact-checking but also erodes their trust in established sources of information. For this reason, the saturation of counternarratives in the digital realm undermines people's ability to distinguish reliable sources from untrustworthy ones, blurring the lines between reputable journalism and disinformation and leaving individuals uncertain about whom or what to believe. The erosion of the public's epistemic autonomy enables the gaslighters to share disinformation also with the aim of discrediting their opponents and denying plain facts, shaping a "*collective cognitive dissonance*" (Rogers and Niederer, 2020) among the international audience.

---

[32] Based on Daniel Kahneman's system of thinking, heuristics is the fast-thinking pathway that is unconscious, emotional, quick, instinctive, automatic, and that it requires little to no effort.

In investigating digital disinformation within the context of international relations, La Cour insightfully explores the notion and the uses of international digital disinformation through three theoretical lenses from international relations theory – namely E. H. Carr's notion of propaganda, John J. Mearsheimer's typology of lies, and Joseph Nye's conceptualization of public diplomacy (La Cour, 2020). The author, in fact, in the article "*Theorising digital disinformation in international relations*" brilliantly demonstrates how digital disinformation becomes an actual weapon in the international arena, since "*when planted by foreign powers with malign intent, false information becomes disinformation – and thus a tool employed in a global power struggle*" (La Cour, 2020). Therefore, in order to elaborate on this, she establishes three main proto-types of disinformation perpetrated by state actors, assigning to each one an actual example, which are:

1. Disinformation story: it is "*one singular false (news) story placed in a foreign information sphere*" (La Cour, 2020). An example of a disinformation story is the so-called "Lisa Affair", which involved the publication on Russian Channel One of a German-Russian girl's story raped by a group of Arab immigrants. Although the news turned out to be false, it caused Anti-Muslim mass demonstrations and political unrest. This Russian disinformation was allegedly intended to influence the domestic debate, decrease support for the government's immigration policies, and undermine Chancellor Merkel due to her perspective Ukraine.

2. Disinformation campaign: it is a "*coherent campaign spreading multiple false stories in a foreign country linked to a particular event*" with the aim of creating "*confusion about the circumstances and facts surrounding the given event*" (La Cour, 2020). An example of a disinformation campaign regards pro-democracy Hong Kong protests against a Chinese legislative proposal that would have undermined Hong Kong's judicial independence. Chinese agents, therefore, spread on the social media several narratives and images discrediting the protestors, like claiming that they were funded by the CIA or that they were members of the ISIS.

3. Disinformation operation: it is "*a long-term effort to systematically deceive a foreign public*" (La Cour, 2020), in which inauthentic profiles created by foreign officers pretend to be real people, "*producing various forms of 'organic content' and engaging with foreign citizens with the purpose of influencing them*" (La Cour, 2020). An example of a disinformation operation is the Endless Mayfly, an Iran-linked network that through inauthentic personas and "ephemeral disinformation"[33] spread fake news in order to

---

[33] It refers to the technique through which, "*after achieving a degree of amplification, Endless Mayfly operators deleted the inauthentic articles and redirected the links to the legitimate news*

undermine Saudi Arabia, the United States, and Israel. They were seeking to exacerbate wadge issues, like Gulf states' cooperation with Israel, and diplomatic crises, like 2017 Qatar diplomatic crisis, posing as legitimate news outlets and directly approaching journalists and politicians.

In her literature review, La Cour explores digital disinformation in international relations referring to the role of emotions in IR, the holistic approach of information warfare, and the field of strategic narratives. Although she does point out how these different fields in IR mention and include international digital disinformation in their studies, they fail in disentangling "*the motives behind the use of disinformation as a distinct tool separate from other tools*" (La Cour, 2020). Therefore, she uses the theories of E. H. Carr, John J. Mearsheimer and Joseph S. Nye to grasp a sense of what digital disinformation consists in international relations. Indeed, international digital disinformation does share some features with Carr's notion of propaganda, as illustrated in his book "*The Twenty Years' Crisis*". For example, both concepts can be employed in various forms, they involve a sort of emotional appeal or 'spiritual element' (Carr, 1939), they have risen because of "*the broadening of the basis of politics, which has vastly increased the number of those whose opinion is politically important*" (Carr, 1940), and they both disregard the truth in their targeting the publics. However, within this context, Carr's notion of propaganda cannot fully incorporate modern international digital disinformation because the former becomes ineffective without a 'national home' (Carr, 1939), which is why nowadays it is particularly difficult to trace the perpetrators of the latter. Moreover, digital disinformation may seem related to the notion of "inter-state lying" depicted by Mearsheimer in "*Why Leaders Lie: The Truth About Lying in International Politics*". Nonetheless, differently from inter-state lying, not only does digital disinformation not focus on military capabilities usually during wartime or crises, but it also shares some features with the other two forms of deception presented by Mearsheimer[34]. Finally, La Cour highlights the similarities between digital disinformation and Nye's juxtaposition of public diplomacy and propaganda, mainly for their objective to increase one nation's soft power by appealing to an abroad public. However, modern international digital disinformation differs from both concepts: while public diplomacy "*projects positive stories about the home country in a foreign context – not negative*

---

sites that they had impersonated. References to the false content would continue to exist online, however, further creating the appearance of a legitimate story, while obscuring its origins" (Lim et. al., 2019).

[34] Mearsheimer identifies three forms of deception, namely lying, spinning – which involves "a person telling a story emphasizes certain facts and links them together in ways that play to his advantage, while, at the same time, downplaying or ignoring inconvenient facts" (Mearsheimer, 2011) – and concealment – which entails the "withholding of information that might undermine or weaken one's position" (Mearsheimer, 2011).

(false) information about various issues in those foreign countries" (La Cour, 2020), Nye's propaganda presents the same discrepancy of Carr's concept of propaganda. In conclusion, La Cour argues that, although these previous IR notions help to understand digital disinformation in international relations, the fact that they do not take into consideration that "*the credibility of the perpetrators of disinformation may remain intact despite their use of disinformation*" (La Cour, 2020) requires additional research focused on this phenomenon, which "*appears to be ever more present in modern-day international politics than ever before*" (La Cour, 2020).

Thus, after having disentangled the content of political gaslighting through social media and its characteristics within the realm of international relations, in order to assess to what extent can these techniques succeed it is necessary to investigate the various global actors involved. Previous research has mainly focused on media manipulation *per se*, without focusing on a comparative analysis of the ability that different actors have in influencing international relations through social media manipulation. Therefore, I shall proceed by decomposing this phenomenon, analyzing several actors that can influence the world order through social media and focusing on the interplay between political elites and social movements. This is pivotal in trying to untangle the extent to which political elites can influence the public opinion and to which they are subject to other actors influencing the public opinion. As for the aim of this paper, I shall start by investigating the phenomenon of online foreign influence, namely the attempt of a foreign actor to manipulate the public opinion of the domestic audience of another country.

Section 2.4 Online Foreign Influence

As previously anticipated, the global reach of social media also raises concerns about the impact of foreign actors on domestic politics. The ability to target populations with customized messages has been exploited by political actors, including foreign governments, to manipulate political outcomes in other countries. This has created concerns about the fairness of elections and the potential for media to be used as a tool for interference and manipulation. This is an aspect to consider: unlike traditional media, where citizens were primarily influenced by their own domestic political actors, the expansive reach of the new digital media allows for the dissemination of particular narratives on a global scale. This is likely to challenge the established system that has been in place since the end of World War II. The methods employed by foreign actors to manipulate domestic politics through social media are similar to the techniques of political gaslighting discussed earlier. Disinformation campaigns can be highly effective when targeted at groups, such as swing voters during elections. The promotion of divisive views can create a polarized climate that is susceptible to exploitation and bots and trolls can be utilized to create an illusion of widespread support or opposition, towards certain viewpoints or individuals. The potential consequences of this scenario are beyond

comprehension both at the international level and within domestic politics. It has the ability to erode trust in the elections and the overall democratic system, thus undermining its legitimacy. Additionally, it can fuel existing divisions leading to instability and conflict while also promoting extreme ideologies, posing a challenge to standing narratives that have been established for decades.

Regarding this issue, in their paper "*Recent Trends in Online Foreign Influence Efforts*", Martin, Shapiro and Nedashkovskaya analyze the characteristics of foreign influence campaigns and the measures governments can take to mitigate their impact. They define online foreign influence as "*efforts by foreign actors to shape public opinion, interfere in elections, or otherwise undermine the functioning of democratic institutions through the use of social media, online propaganda, and other forms of digital communication*" (Martin et al., 2019). These efforts can take many forms, including the creation of fake social media accounts, the dissemination of false or misleading information, and the use of bots to amplify certain messages. In particular, the authors focus on the issue of foreign influence efforts (FIEs) defined as "*coordinated campaigns by one state to impact one or more specific aspects of politics in another state through media channels, including social media, by producing content designed to appear indigenous to the target state*" (Martin et al., 2019). Their work is fundamental in analyzing the patterns and trends in foreign online influence campaigns from 2013 to 2018. They used a wide database, composed of hundreds of news articles, government reports, and previous academic research, in which they identified 53 FIEs that respected their criteria – namely their definition – that targeted at least 24 countries. However, they noted that their account is much lower than the real number of FIEs, due to their difficulty to detect, the limited range of languages of the sources[35], and the identification of 40 more influence efforts that met some of their criteria but not all of them. In order to analyze the trends of this global phenomenon, for each FIE, the authors have identified: the political goal, the attacking party, the platform, the sources, the strategy, the topic, the approach, and the tactics. Table 1[36], summarizes all their findings:

---

[35] Their work took into consideration only the sources found in English, Arabic, French, Spanish, and Russian
[36] The Table has been taken from their study "*Recent Trends in Online Foreign Influence Efforts*"

Table 1: Summary statistics

Variable	Freq.	Variable	Freq.
First sighting		Last sighting	
2013	1	2013	0
2014	9	2014	1
2015	10	2015	1
2016	15	2016	7
2017	12	2017	12
2018	6	2018	32

Variable	Mean	Variable	Mean
Actor		Platform	
Astroturf	0.08	Email	0.08
Company	0.40	Facebook	0.57
Cyber espionage group	0.08	Fake websites	0.15
Foreign government	0.17	Google	0.17
Intelligence Agency	0.19	Instagram	0.21
Real NGO	0.04	Line	0.02
Wealthy individual	0.06	News outlets	0.64
Unknown	0.26	Other media	0.15
		Reddit	0.09
Strategy		Twitter	0.83
Defame	0.64	Whatsapp	0.04
Persuade	0.57	Wikipedia	0.02
Polarize	0.15	Youtube	0.26
Shift agenda	0.06		
Undermine institutions	0.06	Tactic	
		Bot	0.53
Approach		Fake account	0.55
Amplify	0.77	Hashtag hijacking	0.17
Create	0.98	Other tactics	0.23
Distort	0.72	Steal information	0.13
		Troll	0.85

Foreign influence efforts (FIEs) are defined as coordinated campaigns by one state to impact politics in another state through media channels, including social, in a manner which involves producing content that appears indigenous to the target state. In total, 53 FIEs. Each category is not mutually exclusive.

As Russian trolls initiated a campaign to delegitimize Ukraine in the Polish Internet space in 2013, the authors identified the first FIE in their data. Regarding their timeline, they found out that 70% of cases started between 2015 and 2017, and the efforts took an average of 2.2 years to complete. As of the end of 2018, a number of FIEs were still active, such as Russia supporting materials that criticized the Belarusian government and attempting to lower Ukrainians' support for the Donbas conflict. As for the nature of the attackers, the authors discovered that the most frequent actors in FIEs were private firms (47%), media outlets (39%), and intelligence agencies (22%). However, in one-fourth of FIEs, the relevant individuals could not be clearly identified in the media coverage due to insufficient detail. In regard to strategy, attempts to damage the reputation of individuals or institutions, known as "defamation", is the most frequently employed tactic (65%). Persuasion, which is defined as "*trying to move the average citizen to one side of an issue*" (Martin et al., 2019), is the second most prominent tactic (55%). It's noteworthy that just 15% of FIEs attempted to "polarize" ideas by pushing them to the extremes on one or more issues. Moreover, they found much less heterogeneity while identifying the approaches used: in three out of every five instances, the three techniques – "amplify," "create," and "distort" - are used together in one operation. In

99% of the cases, new content was created, existing information was amplified in 78% of the cases, and objectively verifiable facts were distorted in 73% of the cases. Other results showed that among the 24 targeted countries, the U.S. was the main object of FIEs scoring 38%, followed by 9% for the U.K., 6% Germany; Australia, France, Netherlands, and Ukraine 4% each; Austria, Belarus, Brazil, Canada, Finland, Israel, Italy, Lithuania, Poland, Spain, South Africa, South Saudi, Sweden, Taiwan, and Yemen were each targeted once. Additionally, particularly interesting is their research identifying the attacking countries, as shown by Figure 2.

Figure 2: Origins of attacks
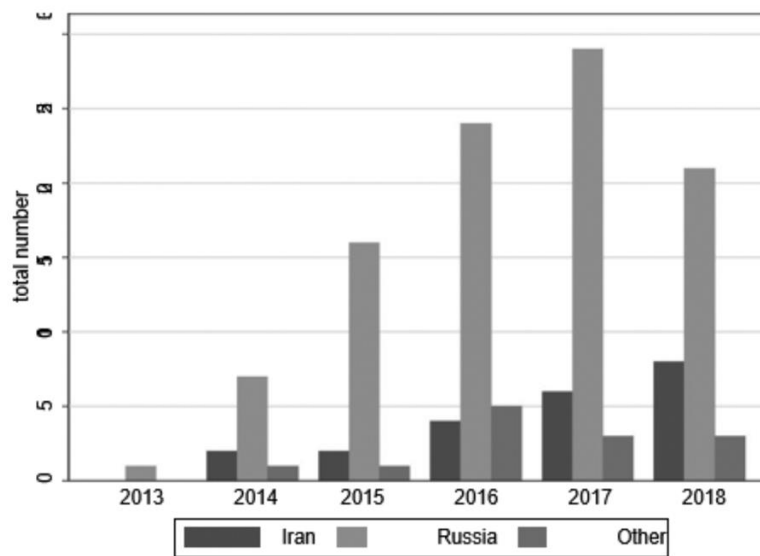
Figure 2 shows that the majority of FIEs launches to date have come from Russia. The authors estimated that by 2017, Russia had participated in 29 different global initiatives. Iran participated in two instances in 2014 and 2015, but its activities increased gradually through 2018, targeting now eight other countries. Throughout the time of their analysis, Saudi Arabia, China, and Iran all started FIEs. Several studies have confirmed these conclusions. Vilmer et al. (2018), for instance, claim that European authorities ascribe 80% of influence attempts to Russia, with the other 20% coming from China, Iran, and the non-state entities, like the ISIS. The scholars also focused specifically on all the FIEs started by Russia, both for their tactics and for their political goals, however, such analysis shall be explored further in detail in the next chapter.

Their findings suggest that foreign actors are increasingly using social media platforms to spread disinformation and influence political outcomes in other countries, becoming increasingly sophisticated and challenging to spot. In fact, since many campaigns use organic and artificial techniques, such as the employment of bots and organized networks of accounts, to spread their message, this makes it more difficult for social media sites and governments to recognize and block

these initiatives. Moreover, the authors point out that geopolitical objectives often serve as the driving force behind campaigns, such as influencing public opinion in a target nation or endorsing a certain political candidate or party, but they also include that some campaigns might be driven by monetary gain or other reasons, such as ideological convictions or personal grudges. In addition to this, a significant discovery made by this study concerns the involvement of non-traditional political actors are also in foreign influence campaigns, as non-state actors like terrorist groups use social media to disseminate propaganda and enlist new members. This emphasizes the necessity for a comprehensive strategy that involves both public and private sector actors to counteract attempts by foreign governments to influence cyberspace.

However, the most comprehensive study that investigated such phenomenon, analyzing it country by country, was "*Industrialized Disinformation: 2020 Global Inventory of Organized Social Media Manipulation*" by Samantha Bradshaw, Ualan Campbell-Smith, Amelie Henle, Antonella Perini, Sivanne Shalev, Hannah Bailey & Philip N. Howard. By analyzing more than 1300 sources, the authors managed to disentangle disinformation campaigns through social media manipulation in 79 countries, highlighting for each one the organizational form, including the actors involved[37], the strategies, tools and techniques employed[38], and their organizational capacity and resources[39]. Although they do not provide a comparative analysis among the different countries analyzed, their work is extremely insightful in exposing evidence-based country-specific examples, which helps the public grasp the reality of such phenomenon and understand the challenges posed by it and its impact on a global scale. For example, they shed light on how the United Arab Emirates spent more than USD 12 million on lobbying and PR in order to counter online allegations of their human rights abuses (Bradshaw et. al., 2021), showing how social media manipulation can be used to silence dissenting voices and suppress free speech, particularly in authoritarian regimes. They also showed how disinformation campaigns perpetrated by foreign governments aim at manipulating the electoral process, for instance Wang Liqiang revealed that he was instructed by the Chinese government to interfere in Taiwan's midterm elections in 2018 and the presidential elections January 2020. However, among the plethora of examples provided by the report, the Russian account seems to be - along with China and Iran - the country that conducted the most online foreign influence operation due to its "global reach", resulting in attacks targeting many EU states,

---

[37] The actors were classified among the categories "government agencies", "politicians & parties", "private contractors", "civil society organizations", and "citizens and influencers".

[38] They assessed this section by classifying the evidence into "account types", "messaging and valence", "content and communication strategies", and "platforms".

[39] Their measure accounted for the "team size", "resources spent in USD", "activity levels", "coordination", and "capacity measure".

the U.S., the U.K., Syria, African countries, and mostly Ukraine. With a budget of USD 35 million for Project Lakhta – an operation aimed at inflaming US political debates – and their diversified target, however, the Russian example shall be discussed further in detail in the next chapter. However, the report does not explicitly assess when a foreign disinformation campaign is effective, since its efficacy depends on a variety of factors, such as the goals of the campaign, the target audience, and the broader political context. However, an example of their effectiveness is provided by Rogers and Niederer in "*The Politics of Social Media Manipulation*", where they demonstrated how Russian campaigners organized and promoted through "front groups" and "faux hashtags publics" several demonstrations in the U.S., as shown in Table 2.

Table 2 Overview of 2016 fake rallies planned and promoted, as listed in the US indictment of 13 Russian nationals concerning foreign election interference

2016 fake rallies planned and promoted		
Date	*Fake rally*	*Location*
25 June	March for Trump	New York
9 July	Support Hillary. Save American Muslims	Washington, D.C.
23 July	Down with Hillary	New York
20 Aug.	Florida goes Trump	Several Florida cities
2 Oct.	Miners for Trump	Several Pennsylvania cities
12 Nov.	Show your support for President-Elect Donald Trump	New York
12 Nov.	Trump is NOT my president	New York
19 Nov.	Charlotte against Trump	Charlotte, N.C.

Source: Parlapiano and Lee (2018)

Therefore, it is evident how foreign social media manipulation can actually be effective in sowing discord, undermining trust in institutions, and influencing the public opinion through the aforementioned techniques. Thus, it is no surprise, as Bradshaw and Howard argued, that before the advent of social media "*most public concern and academic inquiries on cyber power have been preoccupied with the 'hard power' capabilities that affect both the cyber and real-world domains, such as cyber-crime and data theft, attacks that damage critical infrastructure, or online surveillance. Over the past five years, international affairs have become replete with examples of how governments have leveraged social media to manipulate public opinion, assigning personnel and financial resources to disinformation and propaganda campaigns online. These campaigns focus less on the ability to obtain desired outcomes through coercion or material resources, but instead employ what scholars have referred to as 'soft power' and persuasion, framing and agenda setting, ideological hegemony, symbolic power, or sharp power to achieve desired outcomes. This emerging domain of inquiry should be treated as central to the study of international studies as it*

provides an opportunity to both expand and refine the concept and exercise of cyber power, and acts as a countervailing example of state power in the digital age" (Bradshaw & Howard, 2018). However, as highlighted by the previous most comprehensive reports, foreign governments and state agencies are not the only actors that are involved in political gaslighting through social media manipulation. On an international level, in fact, "*even non-state actors are stoking political tensions through the spread of misinformation online – which are - often aimed at the existing liberal order and the institutions that buttress it, portend potential geopolitical upheavals*" (Frederick, 2019). For this reason, the next section shall be focused on this phenomenon perpetrated by social movements and other non-state actors that aim at manipulating the public opinion through social media for political purposes.

Section 2.5 – Non-state actors: transnational relations

As previously anticipated, the modern technological development and the new international dynamics deeply impacted the realm of international relations and the world order. While the conventional state-centric world order, deeply rooted in the historical Westphalian system, prominently underscored the preeminence of nation-states as principal actors in international affairs, the rise of liberalism and interdependence fostered the establishment of today global governance. Since the conceptualization and the definition itself of global governance has been subject to different studies that highlighted its multifaceted features[40], for the purpose of this thesis I shall consider one of the most comprehensive ones, which depicts global governance as "*the complex of formal and informal institutions, mechanisms, relationships, and processes between and among states, markets, citizens, and organizations—both intergovernmental and nongovernmental— through which collective interests are articulated, rights and obligations are established, and differences are mediated*" (Thakur & Van Langenhove 2006). In fact, the seminal concept of "complex interdependence", as propounded by Robert Keohane and Joseph Nye, laid the bedrock for comprehending the augmented interconnectedness among states and non-state actors within the global arena. Their cogent argument illuminated the disentanglement of states from being the sole protagonists, as non-state actors, including international organizations, NGOs, and multinational corporations, gained ascendancy in shaping the international system. In this context, Anne-Marie Slaughter's work "*A New World Order*" provides insights into how non-state actors have significantly shaped the current global landscape. These actors, which go beyond national borders,

---

[40] Some examples are Rittberg's short definition (2002), Risse's focus on the role of collective goods (2012), Castells highlighting its task in conflict-resolution (2005), Weiss & Wilkinson's multidimensional approach (2014), Slaughter's focus on the centrality of non-state actors (2004), and many others

including networks, civil society organizations, and private entities, have played a significant role in establishing norms, laws, and policies both at an international and domestic level. For example, transnational networks, made up of corporations and advocacy groups, have influence in addressing and pressing global challenges such as human rights and environmental preservation. While civil society organizations, ranging from grassroots social movements to non-governmental organizations (NGOs), have empowered individuals to engage in global affairs by advocating for social change and promoting democratic accountability. Moreover, she argues that the emergence of "governmental networks" showcases the rise of intermestic's partnerships that, through the cooperation between domestic and international agents, promote collaborative problem-solving on global common issues, like sustainable development and humanitarian aid.

The role of non-state actors in shaping the international arena was also emphasized by Risse-Kappen's work "*Bringing Transnational Relations Back In*", in which he shifts the focus from the states to transnational advocacy networks, civil society organizations, and transgovernmental relations[41]. In disregarding the traditional state-centric view, the author shows how these non-state actors go beyond national boundaries to form connections across borders working together to achieve shared goals and changing how global governance works. In this context, he argues that "*transnational and transgovernmental issue networks seem to have a major impact on the global diffusion of values, norms, and ideas in such diverse issue-areas as human rights, international security, or the global environment*" (Risse-Kappen, 1995). In fact, by exploring how non-state actors form identities – which will also be the basis of his later work on collective identity – he reveals how the interconnection between transnational actors and domestic structures determines their role in shaping the international order. Moreover, not only does his emphasis on non-state actors as influential agents encourage scholars and policymakers to reconsider current global governance structures and promote dialogue and collaboration among diverse participants, but it also raises awareness of its risks. In fact, Risse-Kappen warns the international community of the ethical flexibility of these agents, by arguing: "*there is no reason to assume that transnational relations regularly promote "good" causes. Transnational terrorism poses a serious threat to internal stability in many countries, while some scholars have identified Islamic fundamentalism – another transnational social movement – as a major source of future inter-state conflicts*" (Risse-Kappen, 1995). The ethically unbiased view expressed in this work represents the approach that is necessary to employ in this section, in which epistemically "good" transnational movements will be

---

[41] Risse-Kappen defines 'transnational relations' as "*regular interactions across national boundaries when at least one actor is a non-state agent or does not operate on behalf of a national government or an intergovernmental organization*" (Risse-Kappen, 1995).

considered alongside "bad" transnational movements and other non-state actors. In fact, in order to understand the impact that such non-state actors have in today's world system, we shall apply Risse-Kappen's ethical unbiasedness to Keck and Sikkink's work "*Activists beyond Borders: Advocacy Networks in International Politics*". Although they focus on transnational advocacy networks[42], which are considered epistemically positive, with the rise of social media their impact is very similar to the one of any other transnational movement – whether it promotes positive or negative values. In fact, as argued by the authors, transnational advocacy networks amplify suppressed voices in an international arena. However, "*the multiplication of voices is imperfect and selective, but in a world where the voices of states have predominated, networks open channels for bringing alternative visions and information into international debate*" (Keck & Sekkink, 1998). The reason for their effectiveness lies on two components, namely the acceptance of the aforementioned 'politics of change' and their 'non-classic' power. As I have analyzed in the previous chapter, global governance and the importance of non-state actors in the international system can happen only if its values are open to change and to be challenged. As for their power, it consists in the sharing of information. In particular, "*besides sharing of information, groups in networks create categories or frames within which to generate and organize information on which to base their campaigns*" and "*their ability to generate information quickly and accurately, and deploy it effectively, is their most valuable currency*" (Keck & Sikkink, 1998). By sharing information, these agents are able to raise awareness across the globe about a particular issue, presenting counternarratives to the established order with the aim to gain international support. However, as I have analyzed in section 2.3, information per se cannot be considered "inherently good", due to the fact that the information ecosystem includes also the mis- and dis-information sphere. Therefore, as the other side of the coin, other transnational actors can show the same characteristics of transnational advocacy networks, have the same impact, and share their message worldwide but for their "malignant" political agendas. An example of this is transnational terrorism, which it is indeed a form of non-state actor that operates across national boundaries, composed by people who promote their shared ideas and values to change the international arena. In this case, however, they manage to gain support transnationally by sharing disinformation and by upholding some counternarratives that not only differ from the existent mainstream narrative, but that also do not reflect Western liberal values. Thus, in order to understand the role of non-state actors in

---

[42] The authors define them 'transnational advocacy networks' as a form of non-state actors' organization that operates across national borders and that "*are organized to promote causes, principled ideas, and norms*" in order to change international politics, consisting of various actors, like non-governmental organizations (NGOs), activists, scholars, and other civil society groups, bound together by shared values and goals (Keck & Sikkink, 1998).

perpetrating political gaslighting through social media, it is necessary to approach this issue from an ethical unbiasedness, considering that in terms of organizational features and media capability "positive" or "negative" transnational movements are quite similar. It should be noted that transnational relations like social movements do differ from others like transnational terrorism[43], mainly in their different political aims and the tactics employed in their attempt to change international politics – primarily regarding the use of violence. However, for the aim of this thesis, they shall be considered as different components of the larger group of transnational relations, since the ways in which they use social media to gain support and share their narratives is analogous. In this sense, the work Beck's work "*The Contribution of Social Movement Theory to Understand Terrorism*", in which he analyzes terrorism under the tripartite model of social movement theory[44] and emphasizes for both phenomena their link with transnational relations. Concerning resources mobilization, Beck shows how "*many terrorist groups seem to be structured like modern social movement organizations – a highly professionalized core that directs and manages attacks, assembles resources, and provides overall leadership to a broader base of supporters*" (Beck, 2008). He also demonstrated their similarities regarding the structure of political opportunities, since both phenomena can arise only if the overall political or social conditions are ripe for successful and sustained contention, and the rhetorical and symbolic side of political contention, since both seek to mobilize support by framing their cause in a way that resonates with potential supporters.

Now that I have demonstrated how both social movements and transnational terrorism are both included in the category of transnational relations, the focus of this section shall be on their use of social media manipulation and disinformation campaigns, in order to show how these are both subjects and objects of international gaslighting and whether it involves a threat to democracy and the international order. In fact, in the realm of contentious politics - namely "*the use of disruptive tactics by groups and individuals (challengers and incumbents) to change or maintain the status quo*" (Zeitzoff, 2017) – the characteristics of social media that I have explored in the previous sections allow for transnational movements to gain support and mobilize masses by facilitating connections among activists, lowering coordination and communication costs, getting feedback on protest platforms, and raising funds (Zeitzoff, 2017). In fact, transnational social movements like

---

[43] Although the definition of 'terrorism' has been subject to several academic debates over the last decades, in which many have proposed different theories and school of thoughts without reaching an arbitrary interpretation, most of them convey on the following features of terrorism – namely the presence of violence or threat of violence, its unconventional targets (e.g. civilians), and its political goals.

[44] The author recalls McAdam's 3-fold framework of social movement's theoretical perspective that involves mobilizing resources, political opportunities, and framing (Beck, 2008).

Black Lives Matter, Hong Kong's Umbrella Movement, or the Arab Spring would have not achieved the same international resonance without the use of social media. Nor would have ISIS and other Salafi-jihadist terrorist groups been able to recruit such followers and provide them tips on how to cross into Syria and stage the attacks (Zeitzoff, 2017). For this reason, An Xiao Mina's book "*Memes to Movements: How the World's Most Viral Media Is Changing Social Protest and Power*" highlights how non-state actors, encompassing grassroots activists, civil society groups, and transnational networks, utilize the potency of social media platforms to catalyze and shape global socio-political dynamics through the use of memes, videos, and other viral content. Although much of Mina's work focuses on how social media platforms serve as powerful tools for communication, coordination, and dissemination of information, enabling non-state actors to swiftly garner support, raise awareness, and hold states and institutions accountable for their actions, she does acknowledge their threat. In fact, she recognizes that social media-driven activism is filled with complexities that need a thorough examination to understand its implications, one of these being the spread of mis- and disinformation due to the quick and unfiltered dissemination of inaccurate or manipulated content. As previously stated, this flood of misleading information not only undermines the credibility of movements but also creates an environment of uncertainty making it difficult to separate fact from fiction. Moreover, as examined with political gaslighting, Mina highlights polarization as another aspect where these very platforms, meant to amplify voices and encourage dialogue, unintentionally deepen divisions: social media-driven activism runs the risk of reinforcing echo chambers and hindering discussions ultimately impeding our ability to foster nuanced understanding and collaborative solutions. It isn't surprising that oppositional movements or groups spread disinformation in order to discredit one another: since the beginning of politics parties or figures have tried to undermine their opposing counterpart by sharing fake news or challenging their creditability, as I have thoroughly examined in Chapter I. A modern example of this may be the attack to Capitol Hill that happened on January 6, 2021, where many Trump supporters identified some antifa groups as the assaulters, while Anti-Trump groups shared edited images and videos to picture a much more devastating episode. However, in the realm of international relations, such disinformation campaigns become much more relevant when they leave the domestic sphere of opposing national groups through the involvement of governments and state-actors, that try to undermine such movements or that are discredited by them. In this sense, also Mina's analysis encompasses the intertwining between social movements, the advent of social media and the state-actors, focusing mainly on state surveillance looming over media-driven activism. She argues that the ubiquity of social media platforms, while on one hand empowers individuals and movements to gain worldwide support, on the other it makes them vulnerable to the surveillance of governments

and authoritative entities. This extensive monitoring, often justified in the name of security and public order creates a dynamic between the potential of online activism and the infringement on civil liberties. An example of this is represented by Muammar Gaddafi's government, who, just before the outbreak of the Lybian civil war, sent threatening text messages to every citizen in order to demobilize people who had organized nonviolent protests against the government. Therefore, being these pervasive surveillance mechanisms much more easy to employ in the context of social media platforms, they become much more effective in discouraging people from participating in discussions and expressing dissent creating an atmosphere where self-censorship prevails, due to fear of consequences. This echoes effects seen in traditional forms of surveillance and has the potential to undermine the very essence of a vibrant and democratic public sphere that social media platforms initially aimed to foster, posing also a threat to individuals' right to privacy. An example of the severity of such state surveillance may be represented by the Egyptian government which used all kinds of social media platforms to arrest journalists or activists who opposed – privately or publicly- the regime, even restoring to dating apps to track down and later arrest and torture members of the LGBTQ+ community[45]. Another digital technique employed by (authoritative) governments to oppose such movements is a complete shutdown of social media and the internet, as happened in 2011 when President Hosni Mubarak blocked all digital platforms in order to stop the protests against his government (Hassanpour, 2014).

Many more studies have been conducted on how transnational disinformation campaigns have actually undermined social movements' stances, public images, and support. For example, as Erica Chenoweth has argued in the Wiener Conference Call[46], *"the IRA had set up an account called Blacktivist on Facebook, which meant to simulate black racial justice activists in the United States, but actually was trying to basically divide Black Lives Matter and make it look more radical in the public mind. They also set up an Instagram account called Woke Black that discouraged people from voting for Hillary Clinton in the 2016 election, spreading stories about her and trying to discourage black voters from turning out and supporting anyone in the polls. And then using the Black Lives Matter hashtag to try to depict legitimate solidarity while also using divisive messages into their media operations. There's also the issue of deep fakes, which has become even more concerning with facial recognition technologies, which many people worry about enabling opponents of peaceful activists using to essentially spread completely irrefutable fake stories about them. So basically, all of these things can undermine the unity of movements, the ability to attract large and diverse segments of society. It can undermine their discipline. It can increase targeted*

---

[45] https://www.bbc.com/news/world-middle-east-64390817
[46] https://www.hks.harvard.edu/wiener-conference-calls/erica-chenoweth#transcript

repression, and it can reduce their ability to expand their basis of support" (Chenoweth, 2020).
However, it is important to highlight that disinformation campaigns against transnational social movements are not only perpetrated by authoritarian state-actors or domestic opposing groups, but also by democratic governments, in order to shift public opinion accordingly to their own political agendas. One notable instance is when ex-President of the United States, Donald Trump, reacted to the rise of Greta Thunberg's FridayForFuture movement regarding climate change. I've previously discussed how this matter has been a prime subject for various social media manipulation methods, leading to political gaslighting tactics such as discrediting scientists and rejecting scientific outcomes. However, Trump's disinformation campaigns on this issue not only did it impact U.S. domestic policies, but also aimed at sowing doubt in an international audience. In several speeches and tweets, former US President Donald Trump has dismissed the issue of global warming, considering it a 'hoax' "*created by and for the Chinese in order to make U.S. manufacturing non-competitive*"[47] and asserting some absurd claims with no scientific basis whatsoever – for example that the sound of wind turbines caused cancer. Not only did this lead to 346 anti-science actions taken by the federal government during Trump administration[48], but it also led to scientific censorship and self-censorship[49], allowing him to carry forward his economic-oriented agenda and dismiss that transparency characteristic of democratic regimes. The ex-President used social media also to discredit the activist Greta Thunberg and 'her' social movement by adopting some classic gaslighting tactics, as is evident in a tweet Trump posted as a response to her nomination as Time's Person of the Year for 2019, in which he wrote: "*So ridiculous. Greta must work on her Anger Management problem, then go to a good old fashioned movie with a friend! Chill Greta, Chill!*". In this context, it becomes apparent how political gaslighting perpetrated by state actors impacts social movements through the spread of disinformation or discrediting of political figures through social media.

However, disinformation campaigns do not impact social movements unilaterally. Simon Norin, for example, in "*Cooperation With Disinformation Campaigns as a Social Movement Strategy*" explores how disinformation campaigns can foster social movements, providing as an example the mobilization against the Swedish Social Service. By starting from previous literature - such as

---

[47] https://twitter.com/realdonaldtrump/status/265895292191248385?lang=en

[48] These data were revealed by the Silencing Science Tracker, a joint project of the Climate Science Legal Defense Fund and Columbia Law School's Sabin Center for Climate Change Law.

[49] During his administration, there were 154 documented instances of federal government censorship and 19 documented instances of scientists' self-censorship (Webb & Kurtz, 2022), due to "*fear of retaliation, belief that reporting would make no difference, perceived suppression or interference by Agency leadership, and belief that politics and policy outweigh science*" (https://www.epa.gov/sites/default/files/2020-05/documents/_epaoig_20200520-20-p-0173.pdf).

Jankowicz (2020), Bennet & Livingston (2018), Earl et. al. (2021) – Norin delves into the intricate interplay between disinformation campaigns, social movements, and their shared reliance on grievances as catalysts for altering political behavior. On one hand, disinformation campaigns, strategically aimed at eroding societal resilience, disseminate false narratives concerning existing grievances, thereby fostering diminished trust within societal groups and institutions. This orchestrated deception yields tangible shifts in political conduct, ranging from escalated protest activities to heightened polarization of discourse. Concurrently, social movements materialize as responses to collective grievances that acquire recognition as salient issues necessitating concerted engagement. Therefore, since both disinformation campaigns and social movements pivot upon the central tenet of grievances in order to change political behaviors, he argues that "*in some cases the two phenomena are interlinked by disinformation campaigns exploiting and amplifying a grievance and thus initiating the creation of a social movement*" (Norin, 2022). Thus, disinformation campaigns, thanks to their ability to amplify grievances, may serve as 'catalyst' for the formation of social movements (Norin, 2022) – some examples, as cited by the author, are the Velvet Revolution in Czechoslovakia, the anti-vaccination movement, and the bronze soldier movement in Estonia. Moreover, disinformation campaigns through social media can foster social movements internationally in another way – namely getting support by exposing them. In critical moments of civic unrest, many governments – whether democratic or authoritarian – tried to, directly or indirectly, share disinformation or misleading information through social media in order to discredit such movements and contain the ongoing turmoil. Many cases previously mentioned fall in this context. However, social media, while on one hand facilitated state actors to gaslight the public opinion against these social movements, on the other they enabled the activists to expose such manipulation, which ended up in discrediting even more the governments at issue. In fact, this dynamic interplay underscores the profound transformative potential that orchestrated disinformation can wield in mobilizing and sustaining collective action on a global scale. Through the strategic dissemination of false narratives, fabricated content, and emotionally resonant messaging, disinformation campaigns permeate the virtual spaces of social media platforms, engendering a ripple effect that resonates far beyond individual screens. The catalytic role of disinformation in fostering and supporting social movements is palpable, exemplified by instances such as the Arab Spring and the "Black Lives Matter" movement. In the former, governments and state actors deliberately tried to block, monitor or spread disinformation on social media platforms like Facebook or Twitter in order to monopolize their countries' news outlets globally and put an end to the protests. However, these attempts actually backfired on them. Thanks to the multifaceted and decentralized features of social media, the privately owned submarine cable running to Europe,

and the involvement of the group of hackers and activists Anonymous (Howard & Hussain, 2013), the activists managed not only to deceive governments' attempts at blocking and arresting them, but they also exposed them to a domestic and global audience, raising awareness on the disinformation campaigns and illicit actions perpetrated by the state actors. This blatant exposure catalyzed mass mobilization, galvanizing public sentiment and coalescing diverse segments of society around a shared cause, leading to seismic political shifts across regions. The same phenomenon happened with the 'Black Lives Movement', in which unveiling the state actors and opposing groups' sharing of fake news and edited visual content on Instagram, Facebook, Twitter, and Youtube resulted in challenging public outrage, spotlighting systemic injustices, and prompting individuals to engage in collective actions. Therefore, this complicated dynamic highlights how both governments and state actors, as well as social movements and other transnational relations, are on one hand perpetrators of political gaslighting through the dissemination of disinformation through social media, and on the other objects of such phenomenon. In this sense, political gaslighting can undermine the involved actors – whether they are the perpetrators or the targets – as well as it can benefit and support them.

Shifting the focus on transnational terrorism in particular, political gaslighting through social media manipulation has the potential to wield far-reaching consequences, spanning from the recruitment and radicalization of individuals to the perpetration of violent acts. The interconnected nature of social media platforms serves as a breeding ground for the propagation of fabricated narratives and inflammatory rhetoric, which can effectively exploit vulnerabilities, amplify grievances, and galvanize potential terrorists across borders. In fact, although it does constitute 'the other side of the coin' of social movement in this context, it is necessary to analyze some specific characteristics of this type of transnational relations, mainly due to the involvement of a violent component. Instances abound wherein political gaslighting, facilitated through social media manipulation, has, in fact, left an indelible imprint on the contours of transnational terrorism. The rise of extremist ideologies, propagated and amplified through digital channels, showcases the profound impact of manipulated narratives on individuals susceptible to radicalization. By distorting historical events, manipulating perceptions of geopolitical conflicts, and amplifying grievances, political gaslighting fuels a sense of victimization and fosters a distorted worldview that resonates with marginalized populations susceptible to extremist ideologies. Additionally, the deployment of misinformation campaigns by state and non-state actors alike can manipulate public discourse, divert attention from underlying causes, and manipulate the narrative surrounding transnational terrorism, thereby exacerbating its complexities. Furthermore, the ramifications of political gaslighting extend beyond mere rhetoric, as its strategic manipulation can facilitate the coordination and execution of terrorist acts. False

information disseminated through social media channels can manipulate potential recruits into carrying out attacks, harnessing their conviction for violent means. This underscores the potential for disinformation to transcend the digital realm, materializing in tangible acts of terror that can destabilize nations, inflict harm on innocent lives, and disrupt the global order. The testimony of author and former nonresident fellow in the Project on U.S. Relations with the Islamic World at Brookings, J.M. Berger, in the U.S. Senate hearing "*Jihad 2.0: social media in the next evolution of terrorist recruitment*" provides a detailed account on how the ISIS employed such strategy worldwide through social media. The group, in fact, effectively utilized social media platforms to propagate distorted narratives and manipulate perceptions through the strategic dissemination of videos showcasing acts of brutality and leveraging emotionally charged imagery, so that they could attract recruits and foster a sense of belonging among vulnerable individuals (Berger, 2015). In particular, after having addressed some alarming numbers regarding ISIS use of social media[50], he depicted three major components to ISIS social media campaign: "*The first is disseminating propaganda to generate support for the group and attract potential recruits and supporters locally and abroad. The second is disseminating propaganda designed to manipulate its enemies' perceptions and political reactions. While some of this material purports to demoralize and deter potential enemies from taking action, its real intent is often to inflame animosity and engage foreign countries in a wider regional war. Some of this propaganda also aims to undermine the unity of the coalition opposing ISIS. Its terrorist actions are synchronized with this goal. The third major component is recruitment. Here, the broad strokes of ISIS's highly visible propaganda campaign give way to a host of smaller, individualized activities*" (Berger, 2015). It is due to this type of propaganda and this use of social media that ISIS has expanded internationally, drawing and counting members all over the world. Similarly, the Rohingya crisis witnessed the strategic dissemination of misinformation on platforms like Facebook, exacerbating tensions and fueling violence against the Rohingya Muslim minority in Myanmar. In regard to this, Aung San Suu Kyi, the de-facto leader of Myanmar, asserted that the crisis in Rakhine state is being misrepresented by a "huge iceberg of misinformation"[51] due to the extensive dissemination of fake news and altered visual imagery. In particular, in this case, Facebook had a "determining role", according to the UN, since the anti-Rohingya propaganda campaign went undetected on the platform, on which it incited murders, rapes, and the largest forced human migration in recent history[52]. Moreover, misleading images claiming to depict violence in Myanmar and being shared on social media sowed distrust

---

[50] He claimed that since "*ISIS is a large organization [it] can afford to have 2,000 people who tweet 150 times every day*" (Berger, 2015).

[51] https://www.bbc.com/news/world-asia-41170570

[52] https://www.nytimes.com/2018/10/15/technology/myanmar-facebook-genocide.html

and undermined the work of human rights groups. Therefore, as real attacks had escalated, graphic images, in some cases relating to atrocities elsewhere, were being posted alongside claims of violence both by the Rohingya people and against them, leaving the public opinion with the impossibility to unveil the truth. Collectively, these instances underscore the tangible and intricate repercussions of political gaslighting through social media manipulation on the landscape of transnational terrorism.

As we have analyzed the impact that political gaslighting has on transnational relations, focusing on the role of social movements and transnational terrorism, there is also an additional kind of non-state actor that is deeply involved in the international dynamics of social media manipulation worldwide, namely multinational companies. The MNCs, which, thanks to globalization and the rise of global governance, have gained much prominence in the realm of international relations from a political and economic point of view, are also a pivotal point in understanding to what extent the phenomenon in question can be perpetrated and contrasted.

Section 2.6 – MNCs and Big Tech and the issues for their regulation
In the context of the rise of global governance, one key non-state actor in delineating the dynamics of the international order are multinational corporations (MNCs), which role in international relations has been debated since the decades following World War II and the rise of the so-called '*Pax Americana*', even though their political importance expanded over the US influence. This intricate interplay between multinational corporations and international relations has been thoughtfully elucidated by Joseph S. Nye in his seminal work "*Multinational Corporations in World Politics*", in which he underscores the multifaceted role that these corporations play in shaping the global landscape, with implications that extend beyond mere economic dimensions. Understanding his argument that the transnational activity of MNCs transcends the classical 'military-security determinism' (Nye, 1974), one can discern the significant influence that multinational corporations wield in international affairs, traversing economic, political, and social realms. Nye's examination of the power dynamics within multinational corporations illuminates their potential to act as agents of change and as stakeholders in global governance structures, highlighting both their direct and indirect effects on the international system through private foreign policy and by influencing the agenda-setting of the nations. The presence of multinational corporations in trade negotiations, exemplified by the negotiations surrounding the Trans-Pacific Partnership (TPP), exemplifies their capacity to influence trade policies, regulations, and investment provisions. In particular, the comprehensive TPP agreement added new market entry plus market presence provisions on services, public procurement, intellectual property rights, investment protection, labour and environmental standards, digital economy, customs

and trade facilitation, small and medium-sized enterprises, monetary policy, competition, anti-corruption and transparency, regulatory coherence and cooperation (Fergusson & Williams, 2016). This underscores Nye's assertion that multinational corporations are not passive actors; rather, they actively shape international economic interactions through their strategies and operations. Through relentless lobbying, coalition-building, and advocacy, in fact, multinational corporations navigate the complexities of international economic interactions, actively molding the outcomes to align with their interests. Moreover, Nye's exposition illuminates how multinational corporations transcend the traditional role of profit-making entities by engaging in corporate social responsibility initiatives and partnerships that extend beyond economic realms. In fact, not only did MNCs transformed substantially international trade, including new trade policy questions on the intergovernmental agenda, but they also "*have unintentionally affected the agenda of interstate relations by stimulating other social groups to press for particular governmental policies*" (Nye, 1974). By collaborating with governments, international organizations, and civil society, these corporations play an instrumental role in addressing global challenges, such as environmental sustainability, public health, and human rights. One example of the key role they played in the international order is exemplified by the Paris Agreement on climate change, which witnessed corporations engaging in collaborative initiatives to reduce carbon footprints and advance sustainable practices, thereby signifying their participation in shaping international efforts towards a more ecologically responsible world. Additionally, the extractive industries provide insight into the intricate nexus between multinational corporations and global governance. The activities of oil and mining companies in resource-rich nations underscore the complexities of balancing economic development with environmental protection and social welfare. In this context, a coalition of civil society, governments, NGOs, and MNCs advanced the multistakeholder voluntary principles of behavior "Extractive Industry Transparency Initiative" (EITI), which aimed at eliminating the problems of resource-curse by ensuring good governance through transparency and accountability in the management of resource rents (Kostald & Wiig, 2008). Such proactive engagement underscores their agency and their recognition of the responsibilities that accompany their transnational operations. In fact, since "*governments are faced with trade-offs between their objectives of welfare and autonomy, even when government controls constrain and diminish the direct corporate role in world politics, they may simultaneously increase the indirect importance of multinational corporations as an instrument or agenda item in intergovernmental politics*" (Nye, 1974).

However, while in the decades of the rise of the 'complex interdependence' the international political role of MNCs was more intended as an indirect influence on the agenda setting or as a

byproduct of the economic transnational dynamics of the world order, with the advent of tech giants, such as internet and social media companies, their impact on non-economically-laden fields of the international landscape has become much more direct. As Li Sheng argued, "*recently, and especially over the latest decade, references to the power of technology companies have become common in studies of international relations. The cross-border influence of the so-called 'Big Tech' companies has increased in the domains of industry, commerce, telecommunications, education, entertainment, culture, and media with the expansion of the Internet, which has diffused information and communications technology throughout societies and across many sectors of the economy. The COVID-19 pandemic has left people even more dependent on the services provided by these technology giants in a range of areas, including news, education, work, food, and health. As their influence has grown, Big Tech firms have attained more power to intervene in the economy and are playing increasingly significant roles in domestic politics and international relations*" (Sheng, 2022). In fact, while she underscores the international economic centrality of these companies – for example, with their complex supply chains and global manufacturing operations, they substantially and directly influence international trade networks and the global markets, fostering interdependence and reshaping traditional notions of economic sovereignty - she delves into the political, social, and ethical of Big Tech's global presence, "*challenging and transcending authority, [which is] essential for scientific and technological innovation*" (Sheng, 2022). In order to demonstrate how the information revolution driven by the latest big tech companies has engendered a reconfiguration of power dynamics, leading to a perceptible reduction in the hegemonic authority of states both domestically and internationally, Sheng uses Susan Strange's arguments presented in her work "*The Retreat of the state: The Diffusion of Power in the World Economy*". Susan Strange's prescient observation, highlighting the substantial surge in individuals' access to vast troves of information, serves as a starting point for a nuanced exploration of how the ascendancy of Big Tech corporations has significantly reshaped the role of states and governments within the realm of international relations. As the digital age unfolds, governments worldwide find themselves navigating uncharted waters, wherein the once-unchallenged primacy of their authority is gradually eroded. Her discerning analysis, in fact, underscores a pivotal transformation in the dynamics of state authority. In particular, the influence of large governments and their instrumental organs, such as central banks and intelligence agencies, confronts a stark reduction in power, which emanates from the empowerment of non-state actors, an effect both relative and absolute. The crux of this transformation lies in the rapidity with which information is diffused, enabling nonstate entities to assert their influence in real-time. This velocity of response, exemplified by the agility of markets that react within seconds, stands in stark contrast to the inherently hierarchical and

deliberate decision-making process of governments. As Strange underscores, the imperative for states to convey a unified stance necessitates a time-consuming hierarchy-driven approach, rendering them unable to match the swiftness of non-state actors. Moreover, the interplay between domestic policies and foreign affairs amplifies this shift, rendering traditional state-centric paradigms increasingly intertwined. Governments, driven by the need to safeguard their interests, are compelled to circumvent other nation-states and engage directly with foreign citizens. This dynamic not only redefines the contours of diplomatic interaction but also underscores the potent emergence of transnational civil society as a consequential international actor (Strange, 1996). The global landscape, therefore, witnesses the convergence of state and non-state actors, each wielding distinct influence within an intricate web of international relations. Furthermore, the recalibration of roles transcends the boundaries of politics and extends into the economic domain. Global business enterprises, propelled by the advent of Big Tech, have assumed roles once held exclusively by governments. These corporations, with their expansive reach and operational fluidity, influence economic policies, global trade, and the socio-economic welfare of nations. The shift is particularly evident in the realm of digital governance, as technology giants like Google and Facebook[53], by design or consequence, contribute to the formation of norms and standards that inherently encroach upon state regulatory frameworks (Strange, 1996). Subsequently, Sheng focuses on Big Tech's impact on the international order mainly from a political and social perspective, advancing the diminishing role of the states. Therefore, the author argues that one prominent facet of their political impact lies in their capacity to shape and manipulate public discourse, thereby exerting influence over electoral processes and democratic outcomes. The nexus between social media platforms and political communication has emerged as a potent tool for political actors to disseminate tailored messages, target specific demographics, and potentially sway public opinion. The infamous Cambridge Analytica scandal, a watershed moment, vividly illustrates the potential manipulation of user data for micro-targeted political messaging, thereby underscoring Big Tech's role in political campaigns. This newfound ability to directly engage with voters and bypass traditional media channels has transformed the dynamics of electoral processes, raising vital questions about the ethics and transparency of these digital interventions. Moreover, from a socio-political point of view, Sheng argues that Big Tech's political influence transcends national borders, manifesting in its role as a conduit for transnational interactions and collaborations. The digital platforms offered by these companies facilitate cross-border dialogues, grassroots movements, and advocacy

---

[53] It needs to be noted that when 'Facebook' refers to a social media platform it indicates the social 'Facebook', while it indicates 'Meta' when referring to the big tech company. Throughout the paper, therefore, the term 'Facebook' will be used to indicate both the social media platform and the tech company affiliated in order to maintain coherence with the academic literature

campaigns that challenge traditional notions of state-centric diplomacy. Examples abound, such as Twitter serving as a catalyst for global social movements, from the Arab Spring to climate activism, as I have analyzed in the previous section. This borderless influence extends to the realm of geopolitical dynamics, where technology giants can find themselves embroiled in international conflicts due to their market reach and digital presence.

However, she introduces and delves into one of the most complex and controversial questions regarding both the impact on the international system of the Big Tech and the possibility for any agent to manipulate social media, namely the multi-channel relationships between governments and private companies owning social networks platforms and the possible regulations of such media. This is a result both of their global reach and their different nature from other private companies: while other firms sell goods or provide services, digital platforms are more capable of creating value by exploiting data, which allows them to operate as critical market makers (Mayer-Schonberger & Ramge, 2018). The key problematic factor is, thus, that their distinctive nature lies in their role as self-contained markets, characterized by data-rich environments that actively shape consumer preferences (Mayer-Schinberger & Ramge, 2018). Therefore, the speculations over their regulations do not only involve users' privacy and data collecting, but expand on a much broader discourse entailing the right to freedom of expression, democracy, and the complicated relationship between governments, international law, and multinational companies that are not merely considered for their economic value. As highlighted by Susskind, "*the basic case for legislative intervention is, in fact, non-partisan. It's simply that, as more and more of our discourse migrates online, social media platforms are increasingly trusted to draw the borders of free expression. They order, filter and present the world's information. They set rules about what may be said and who may say it. They approve, and ban, speakers and ideas. And when they do these things, they necessarily apply their own rules, principles, biases, and philosophies*" (Susskind, 2022). The overlapping of regulatory efforts with the right to freedom of expression underscores the complexities inherent in governing these platforms. In fact, the intricate relationship between these platforms and governments spans a spectrum from cooperation to conflict, reflecting the multifarious ways in which information dissemination, censorship, and freedom of expression intersect. While the right to freely express opinions and access information is a cornerstone of democratic societies, it becomes a formidable task to define the boundaries of permissible speech within the online realm. The challenge lies in discerning between legitimate expression and content that incites violence, promotes hate speech, or disseminates misinformation, raising critical questions about the extent to which regulatory interventions can coexist with the principles of open discourse. For this reason, legal frameworks such as the International Covenant on Civil and

Political Rights (ICCPR) Article 19, which guarantees the right to freedom of expression, have been invoked in attempts to regulate these platforms. In particular, Article 19 states:

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 reputations of others;

   (b) For the protection of national security or of public order (ordre public), or of public health or morals.

As governments grapple with the ramifications of unfettered digital communication, the application of ICCPR Article 19 exemplifies this intricate web of challenges. While this article enshrines the fundamental right to express and receive information, its interpretation within the realm of social media platforms has been subject to debates. Instances of hate speech, incitement to violence, and dissemination of disinformation on these platforms have spurred discussions about the limits of free expression. Striking a balance between upholding democratic values and preventing harm has led to nuanced debates, exemplifying the complexities of regulating social media. First of all, the first point to address is whether social media companies are subjected to any sort of law. Within the realm of international law, the regulations that govern the conduct of social media companies in managing their platforms carry profound implications for the safeguarding of human rights, particularly in the context of freedom of expression, privacy, and discrimination. As conduits of information and communication, these platforms wield substantial influence over how individuals express, access, and engage with information. However, as private companies, they are not bound by human rights law, unless human rights standards are translated into national regulation. Workers' rights, child, environmental, journalist, and press freedom protection are a few examples of areas where such translation has taken place. However, there is still no regulation that specifies the role and responsibility of tech giants, whose influence often transcends that of states in shaping individual speech, public discourse, discrimination, and privacy, underscoring a crucial gap in the existing legal framework. The Leibniz Institute for Media Research's[54] investigation into the development process of Facebook's Standards, namely the development of content-related policies,

---

[54] Kettemann, M. C., & Schulz, W. (2020). *Setting rules for 2.7 billion: A (first) Look Into Facebook's Norm-making System ; Results of a Pilot Study.*

shed light on the intricate mechanisms at play. Employing a defined, multi-step process in drafting and refining its Standards, Facebook's consultations with stakeholder groups evoke a non-systematic yet representative-oriented approach. The influence of economic interests on these standards remains ambiguous, with the potential impact on individual rights being a subject of contention. In fact, while from an economic standpoint "*their goal– to make Facebook a more attractive platform where everyone's "voice" can be heard – is supported by the economically sensible goal to be used more broadly*" (Kettemann & Schulz, 2020), they pointed out that "*the "primary focus of our engagement is civil society organizations, activist groups, and thought leaders, in such areas as digital and civil rights, anti-discrimination, free speech, and human rights*" (Kettemann & Schulz, 2020). Therefore, Facebook, like other big tech companies, finds in a complex position trying to manage different interests from all sorts of fields, such as their economic profits, human rights law, and national legislation. As concludes the study, "*Facebook has been constructing a prima facie autonomous and private normative order for public communication that seeks to reconcile interests within that order and is conceived largely without reference to state law or international human rights standards*" (Kettemann & Schulz, 2020) – or, in the words of one Stakeholder Engagement team member, they "*are making the rules up*". Within this context, the application of ICCPR Article 19, which upholds freedom of expression, poses nuanced challenges. While traditionally applicable to state actors, the evolution of social media platforms prompts a reassessment of its reach. The UN Guiding Principles on Business and Human Rights (UNGPs) establish that companies possess human rights responsibilities, offering a framework relevant to various business enterprises, including social media giants. While not binding law, Facebook's voluntary commitment to the UNGPs further underscores its acknowledgment of these principles. However, the complexity of their regulation transcends the mere adhesion to human rights standards, since in the digital public sphere it becomes a compromise between the obligation to protect freedom of expression (ensuring legal content remains online), to enforce the boundaries of freedom of expression (removing unlawful content), to adhere to national legislation, and to simultaneously carry on the company's economic interests. An example to understand the complexity of this 'web of challenges' is Turkey's amendment of the Internet Law (Law no. 5651), which places a regulatory burden on social media companies, requiring them to appoint local representatives in Turkey, store user data within the country, and swiftly respond to removal requests. While ostensibly aimed at safeguarding against content that is deemed offensive or harmful, these measures also raise concerns regarding potential infringements on freedom of expression and the privacy of users – mainly considering that Turkey has been found guilty in multiple occasions in front of the ECHR of unlawful or disproportionate use of content blocking

measures[55]. Therefore, social media companies found themselves having to choose either to respect international human rights standards – being the right to free speech guaranteed under the Universal Declaration of Human Rights (UDHR), the International Covenant on Civil and Political Rights (ICCPR), and the European Convention on Human Rights (ECHR) –risking to exit the Turkish market, or to comply to the national legislation, jeopardizing free speech in the country. Another issue in their regulation is their responsibility in removing unlawful content. A paradigmatic example is Germany's legislative measures such as the Network Enforcement Law (NetzDG), which compels rapid content removal under the threat of substantial penalties. While the reasons behind this legislation do comply with human rights law, its enforcement raises several questions of accountability. Being big tech companies responsible to remove in the fastest way possible unlawful content, the rapidity of their action may compromise freedom of expression, since delegating such profound decisions to private entities within compressed timeframes risks overregulation, potentially leading to the unintended removal of legitimate content. Furthermore, the absence of rigorous safeguards characteristic of state-based processes, including independent judicial review, oversight, and complaint mechanisms, underscores the exigency for a more robust legal framework to ensure a balanced protection of freedom of expression. The duty to uphold this balance rests squarely with the state, even when outsourcing decisions regarding unlawful content to private entities. The extant legal framework often falls short of this responsibility, warranting a more comprehensive discourse on content regulation. Moreover, in terms of preserving lawful content online, social media companies enjoy substantial autonomy in defining and enforcing their terms of service and community guidelines, even in cases involving protected speech under human rights law. Responding to this challenge, the UN Special Rapporteur on Freedom of Expression[56] has advocated for the alignment of company content moderation practices with international freedom of expression standards. This calls for adherence to the principles of legality, legitimacy, and necessity that traditionally bind states when curtailing freedom of expression. The articulation of clear and specific company rules that mirror these principles is imperative, ensuring predictability and transparency for users while simultaneously safeguarding fundamental rights. Therefore, the juxtaposition of these principles against the complex operations of social media platforms raises fundamental questions about the mechanisms for ensuring adherence to human rights norms and other international standards in the context of social media regulation. As digital

---

[55] European Court of Human Rights, Ahmet Yıldırım v Turkey, no. 3111/10, 18 December 2012; Akdeniz v Turkey (dec.), No 20877/10, 11 March 2014; Cengiz & Others v Turkey, Nos 48226/10 and 14027/11, 1 December 2015.
[56] https://documents-dds-ny.un.org/doc/UNDOC/GEN/G18/096/72/PDF/G1809672.pdf?OpenElement

spaces increasingly serve as central platforms for communication, commerce, and development, the challenges posed by the application of human rights principles to private entities demand a thoughtful and comprehensive approach. Balancing the autonomy of corporations with the safeguarding of individual rights remains an imperative for the evolving landscape of international law.

Additionally, I shall argue that there is also another sphere in which social media companies' interests and governments' agendas intertwine and overlap – namely when technological advancement and data exploitation are involved. In this sense, governments play a dual role in their interaction with social media companies, simultaneously supporting their growth while also exploiting the opportunities they provide. This complex relationship reflects the intricate interplay between technology, governance, and international relations. As Carolyn Abbot underscores, governments are driven by a compelling need to foster technological advancement, recognizing its pivotal role in propelling economies toward high-tech knowledge-based models and ensuring global competitiveness. This imperative often prompts governments to exercise caution in imposing stringent regulations on social media companies, as overregulation could stifle innovation and hinder the development of emerging technologies that hold immense economic promise (Abbot, 2012). However, this support for technological development also underscores a delicate balance, where governments must grapple with challenges arising from the rapid evolution of technology and the informational asymmetries between state regulatory bodies and private enterprises. The intricate nature of new technologies makes it complicated to formulate precise and effective regulatory frameworks that accommodate their fluidity. As private entities amass significant expertise and knowledge, traditional state-centered regulatory approaches can falter in delivering optimal outcomes. In light of these complexities, the utilization of "soft-law" mechanisms, as proposed by Abbot, emerges as a pragmatic strategy for 'bridging the gap' between governments and the evolving technology landscape. By fostering collaboration and information-sharing between governments, industry stakeholders, and interested parties, soft-law mechanisms can facilitate the timely exchange of knowledge and insights necessary for effective regulation. This approach mitigates the risk of regulatory disconnection from rapidly evolving technological advancements, ensuring that regulatory controls remain relevant and adaptable. However, this governments' strategic alignment with social media companies extends to a realm of exploitation, whereby states harness the influence and reach of these platforms to advance their own agendas. Governments have been known to utilize social media platforms for propaganda, misinformation campaigns, and surveillance purposes, thereby capitalizing on the platforms' ability to disseminate information to vast audiences in real-time. In particular, a research made by European Parliamentary Research

Service[57] highlighted five main risks of governments' exploitation of social media, namely "surveillance, personalisation, disinformation, moderation and microtargeting" (Dumbrava, 2021). Specifically, it focuses on the ways in which state agents gather, exploit, and use data provided by social media companies in order to further their own agendas, causing profound implications for democracy and governance. One notable dimension is the surveillance of social media platforms for intelligence and law enforcement purposes. Governments have increasingly employed sophisticated algorithms to monitor public sentiment, track individuals, and anticipate potential threats. The revelations surrounding the Prism program, disclosed by Edward Snowden, revealed how intelligence agencies, including the United States National Security Agency, have tapped into social media data streams to gather information on citizens, illustrating the extensive reach and potential misuse of such surveillance. Thus, while this surveillance is necessary, as I have explained before, in order to prevent threats and criminal activities, it provides also personal data of social media users that are not related to such schemes. Therefore, in addition to the previously analyzed disinformation campaigns spread by state-actors thanks to the use of such data, governments have exploited the personalized nature of social media to tailor political messaging and influence public opinion. By leveraging microtargeting techniques, governments can deliver tailored content to specific demographic groups, exacerbating filter bubbles and echo chambers. The Cambridge Analytica scandal exemplifies this trend, wherein personal data harvested from Facebook was allegedly used to target and manipulate voters during the 2016 U.S. presidential election and the Brexit referendum. Microtargeting[58] demonstrated to be highly more effective than traditional target advertising. As Dumbrava explained in her analysis, "*targeted advertising, including for political purposes, is nothing new. Political campaigners have for decades used demographic data (such as age, education, employment, or residence) to refine and focus electoral strategies (e.g. voter segmentation). The novelty of political microtargeting is the use of greater and broader types of data than conventional advertising, including data on online behaviour (e.g. purchase history, browsing history, social media likes and shares). This allows targeters to identify micro-groups of people who share certain characteristics or inclinations that make them more likely to respond to a specific message. Psychological profiles may also be used to capture personal inclinations and*

---

[57] "Key social media risks to democracy: risks from surveillance, personalisation, disinformation, moderation and microtargeting"

[58] According to T. Dobber, R. 'O Fathaigh and F. J. Zuiderveen Borgesius, microtargeting describes advertising strategies that use data and predictive modelling techniques to disseminate highly personalised messages to influence individual behaviour, while political microtargeting refers to a subset of such strategies aiming to influence citizens' political opinions and behaviour, i.e., to persuade or dissuade, inform or confuse, and mobilise or demobilize voters.

predispositions. While appealing to people's unconscious fears and yearnings is not a novel advertising tactic, microtargeting provides far more effective ways to identify these predispositions and to reach people who have them. Lastly, unlike traditional advertising, which is able to adapt its messages to none other but the general demographic groups, microtargeting allows messages to be tailored to much smaller and homogenous groups. For example, this can be done by using 'A/B testing' techniques, which imply sending out slightly different versions of the same message to different population segments to test patterns in their responses" (Dumbrava, 2021). The results of this microtargeting technique are, therefore, political manipulation and the distortion of the electoral process, which consequences, both for the democratic order and the international system, have been analyzed in the previous sections.

Thus, the issue of social media regulation in order to control the spread of disinformation or other manipulation techniques is highly complex and difficult to untangle, due to the different interests of the various actors involved. On the one hand, big tech companies want to increase their profits and their use in most countries, they have an international obligation to respect human rights law even when they are not translated in national legislation. While on the other, governments are keen to leave social media regulation to the big tech firms, they also expect to use data provided by these platforms not only to prevent threats but also to foster their own political agenda. Moreover, the question becomes more complex when dealing with the boundaries of such regulation, establishing the right standards through which the right to free speech and express opinion is not attacked by removing content suggesting misleading information or offensive comments. Therefore, in such a bewildering situation, I shall try to answer the questions 'what has been done?' and 'what could be done?'.

Section 2.7 – What has been done and what could be done
In the previous section, I have analyzed the issues and the interests' overlap that characterize the establishment of social media regulation. Spotting and removing disinformation shared on social media, therefore, has become one of the main challenges of today's post-truth world, which still has not been addressed adequately, resulting in eroding trust in democratic institutions and deeply affecting the international order. As Schiffrin exposed in 2017 this issue in "*Disinformation and Democracy: The Internet Transformed Protest But Did Not Improve Democracy*", a definitive solution, six years later, still has to be found. As she argued, "*in many cases, countries with laws against hate speech and incitement will need to find democratic ways to enforce them online so that the fight against dis- information does not become an excuse for corporate and government censorship. Asking big tech companies to deal with the problem on their own opens the way to corporate censorship, free expression advocates have consistently warned. On the other hand, it is*

important not to let technology companies use free speech as an excuse not to take action"
(Schiffrin, 2017). In fact, as many have argued, disinformation through social media is a relatively
new phenomenon, of which governments and other agents became aware in the last six or seven
years. As pointed out by Schiffrin, actually, only with the scandal of Cambridge Analytica and the
different 2016 elections, the narrative and the perception of social media as a catalyst of a
democratic promoter of social activism – due to their role in the Arab Spring – completely shifted,
revealing its 'dark side'. The advent of the COVID-19 pandemic and the Russian invasion of
Ukraine contributed to highlighting and spreading the phenomenon of political disinformation
through social media, urging the agents involved to take action in order to counter-measure it. An
example illustrating the gravity and the impact of the spread of disinformation online previous to
the countermeasures adopted after the 2016 'boom' of fake news, is represented by Vosoughi, Roy,
and Aral's study "*The spread of true and false news online*", in which they delve into the intriguing
dynamics of information propagation across social media platforms. Focusing on Twitter as a
microcosm of online information diffusion, the researchers examine the contrasting trajectories of
true and false news stories, unearthing significant implications for the understanding of
disinformation dissemination. The authors employ a large dataset comprising rumors and news
stories shared on Twitter over an eleven-year period (from 2006 to 2017), meticulously tracking the
virality of both true and false narratives. Notably, their approach encompasses a comprehensive
array of factors, including the source, content, structure, and dynamics of retweets. Through
rigorous quantitative analyses involving 126,000 rumor cascades[59] verified by six independent fact-
checking organizations, Vosoughi et al. ascertain the extent to which false information spreads
more rapidly and extensively than its true counterparts, thereby prompting profound inquiries into
the mechanisms underlying such disparities. Figure 3 summarizes their findings, where (A) is an
example rumor cascade collected illustrating its depth, size, maximum breadth, and structural
virality over time; (B) is the complementary cumulative distribution functions of true, false, and
mixed cascades, measuring the percentages of rumors that exhibit a given number of cascades; (C)
is the quarterly counts of all true, false, and mixed rumor cascades that diffused on Twitter between
2006 and 2017, annotated with example rumors in each category; (D) is the complementary
cumulative distribution functions of true, false and mixed political cascades; (E) is the quarterly
counts of all true, false, and mixed political rumor cascades that diffused on Twitter between 2006

---

[59] As explained by the authors, "*a rumor cascade begins on Twitter when a user makes an assertion about a topic in a tweet, which could include written text, photos, or links to articles online. Others then propagate the rumor by retweeting it. A rumor's diffusion process can be characterized as having one or more cascades, which we define as instances of a rumor-spreading pattern that exhibit an unbroken retweet chain with a common, singular origin*" (Vosoughi et. al., 2018).

and 2017, annotated with example rumors in each category; (F) is a histogram of the total number of rumor cascades analyzed across the seven most frequent topical categories.

Figure 3: Twitter rumor cascades 2006-2017
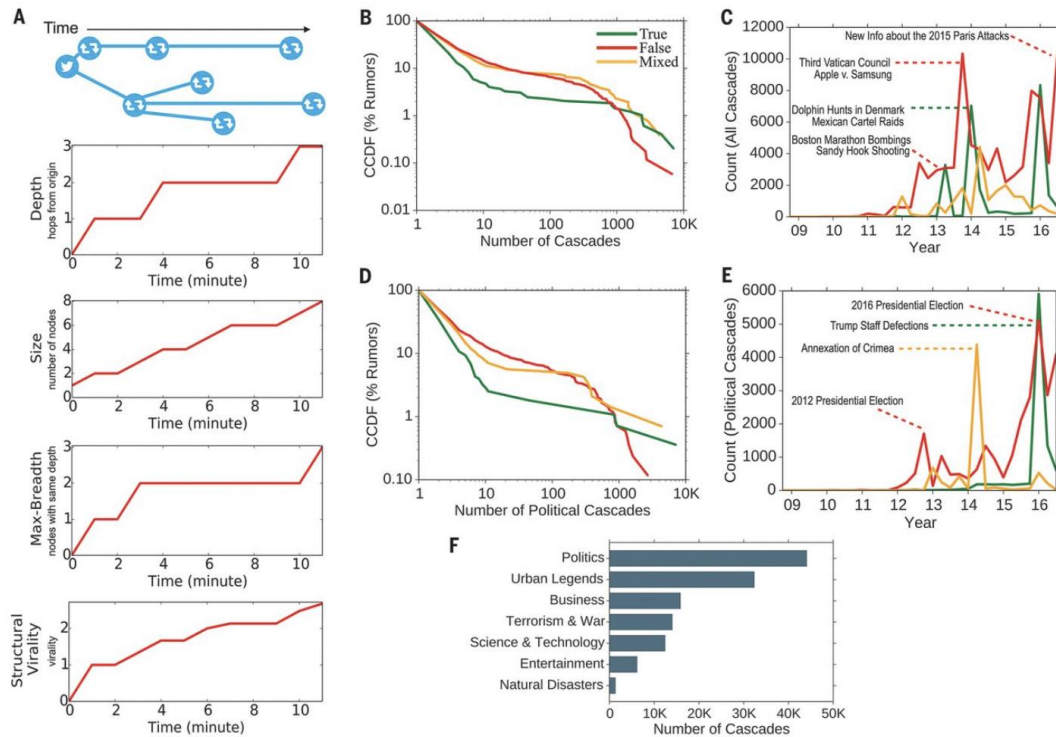
Key findings from the study underscore the notable speed and reach with which false news diffuses, outpacing the propagation of accurate information. This puzzling observation contrasts prevailing theories that posit truthful content to be more viral, highlighting the counterintuitive nature of information flow on social media platforms. In particular, they found that not only is false news shared more rapidly and by a greater number of users, but they also highlighted how political rumor cascades are much more prominent on Twitter, significantly exceeding all the other categories. The study also delves into the characteristics of users propagating false information, revealing that they possess fewer followers, follow fewer accounts, and are less active in comparison to users disseminating true information, underscoring the role of user engagement and behavior in shaping the dissemination of disinformation. Moreover, after dissecting the psychology behind individuals' engagement with false news stories, the authors also explore the temporal dynamics of information spread, discerning that false narratives experience rapid spikes in activity, often due to concentrated bursts of retweets, in contrast to the steadier and more sustained pace of true news stories. The implications of these findings have proven profound, underscoring the multifaceted challenges posed by disinformation in the digital age. The study not only challenged conventional wisdom surrounding information diffusion but also highlighted the need for an enhanced understanding of

the cognitive, psychological, and social factors that contribute to the virality of false news. The research offers critical insights into the interplay between human behavior, platform dynamics, and the proliferation of disinformation, furthering the discourse on strategies for mitigating the harmful impact of false news spread on social media on society. Thanks to this study, other research and investigations, governments and civil society became increasingly aware of the threat of political disinformation through social media platforms, starting to develop strategies aimed at counteracting such phenomenon.

In the study requested by the European Parliament's Subcommittee on Human Rights, "*The impact of disinformation on democratic processes and human rights in the world*", the authors distinguish three groups of actors responsible for providing a response to this issue, namely the private sector (digital platforms), legislative and regulatory bodies, and civil society (Colomina et. al., 2021). It is therefore needed to analyze the involvement of these actors in their attempt to regulate the spread of political disinformation, starting from the social media companies themselves. In response to the rampant spread of disinformation across digital platforms, social media companies have taken notable steps to regulate and mitigate the dissemination of false content. With governments initially slow to address the issue of digital election interference, tech companies have emerged as pivotal actors in combating the spread of disinformation. These self-imposed measures vary in their design and focus, but they undeniably contribute to curbing the dissemination of political disinformation on their platforms. Such voluntary content moderation practices are underpinned by the legal rights that social media platforms possess as publishers, as defined by their terms-of-service agreements between users and platforms, which hold significant influence over digital speech and serve as key documents governing online discourse in today's world. According to Metzger, policies responses can be distinguished in two groups based on their purposes: the ones aimed at 'reducing exposure', which involves the removal of false content and its visibility 'downranking' algorithmically, and the ones aimed at 'reducing belief', which main strategy consists in labeling (Metzger, 2019). For instance, Twitter's decision to ban all political ads in November 2019 aimed to counter the spread of disinformation during election campaigns, and in February 2020, the platform implemented a partial ban on manipulated visual material, including deepfakes. After Elon Musk's acquisition of Twitter, their different policies depend on the "potential offline harm"[60] of the contents shared, determining whether the platform will adopt a 'reducing exposure' or a 'reducing belief' strategy. In particular, the platform informs that if offline consequences can be immediate or severe, they will remove the content or limit its amplification, while in other cases they will "*inform and contextualize by sharing timely information or credible content from third-party sources*". The

---

[60] https://help.twitter.com/en/resources/addressing-misleading-info

latter will be carried on through several techniques, such as 'labeling content', 'prompting when engaging with a misleading tweet', 'creating Twitter moments', 'launching prebunks', and by getting users' feedback. Moreover, Twitter provides a really transparent insight for its users regarding its strategies to combat mis- and disinformation, presenting a table[61] describing and synthesizing the criteria and the correspondent policies employed (Table 3).

Table 3: Twitter's criteria and strategies to combat online mis- and disinformation

Is the media significantly and deceptively altered or fabricated?	Is the media shared in a deceptive manner?	Is the content likely to impact public safety or cause serious harm?	
✓	✕	✕	Content **may** be labeled
✓	✕	✓	Content is **likely** to be labeled, or **may** be removed.
✓	✓	✕	Content is **likely** to be labeled.
✓	✓	✓	Content is **very likely** to be removed.

Meanwhile, YouTube reinforced its guidelines on misleading election-related content before the 2020 US elections, focusing on the removal of technically doctored or manipulated content. However, compared to Twitter, Youtube – a Google platform – provides a less clear and open policy regarding the fight to mis- and disinformation. In particular, in its guidelines[62], the regulation focuses more on the content that shall not be shared – distinguishing several scenarios such as 'misinformation policy', 'elections misinformation policy', 'COVID-19 medical misinformation policy', and 'vaccine misinformation policy' – and on users' reports of such content. Moreover, unlike Twitter, Youtube does not apply different strategies based on the content reported, but it sticks to its '3 strikes system', after which the account that has been spreading mis- and disinformation will be terminated. However, Google has promoted various partnerships in its fight against disinformation. Notably, after an IPSOS survey[63] showing that "*more than a quarter of the Central European and Baltic population is strongly affected by disinformation*" due to the war in Ukraine, Google participated in the strengthened EU Code of Practice on Disinformation – which will be analyzed later – and built a long-term partnership with the Civic Resilience Initiative, the Baltic Center for Media Excellence and the Charles University to fight online disinformation.

---

[61] https://blog.twitter.com/en_us/topics/company/2020/new-approach-to-synthetic-and-manipulated-media

[62] https://support.google.com/youtube/answer/10834785?hl=en&ref_topic=10833358&sjid=17688554870839399945-EU

[63] https://www.ipsos.com/cs-cz/more-quarter-central-european-and-baltic-population-strongly-affected-disinformation

Additionally, recognizing the gravity of deepfakes, Microsoft and Facebook initiated in 2019 the Deepfake Detection Challenge, inviting global participation to develop innovative technologies for detecting AI-generated videos and raffling US $1 million. Facebook, facing significant scrutiny, has actively navigated its role in tackling disinformation. While refusing to ban political ads based on free speech protections, the platform announced in January 2020 that it would remove deepfakes and manipulated video content that might mislead viewers. Ahead of the 2020 US election, Facebook has also heightened its efforts against disinformation by increasing transparency, labeling content flagged as false, and relying on both human fact-checkers and technological solutions for content identification. The platform's advocacy for government regulation, as evidenced by its white paper in February 2020, underscores the need for balanced safety, freedom of expression, and other values. Facebook's strategies, in fact, highly resemble Twitter's policies, adopting different procedures based on the content at issue[64]. In particular, having "*built the largest global fact-checking network of any platform and having contributed more than $100 million to programs supporting our fact-checking efforts since 2016*"[65], Facebook bases its disinformation countermeasures on independent fact-checking agents, which categorize the content at issue as 'true', 'altered', 'partially false', 'lacking context', 'satire', and 'true'. And, depending on this categorization, Facebook will determine whether to adopt a 'reducing exposure' or 'reducing belief' strategy – much like Twitter's policies. In the report "*Disinformation, Deepfakes & Democracy: The European response to election interference in the digital age*" for the Alliance of Democracies, Waldemarsson summarizes a broad picture of voluntary restrictions by social media platforms (Table 4).

Table 4: voluntary restrictions by social media companies

	Facebook	Twitter	YouTube/Google	Reddit	LinkedIn
Allows deepfakes?	No	No	No	No, unless satirical	N/A
Allows political advertising	Yes	No	Yes	Yes, on federal level and only for US elections	No
Allows micro-targeted political advertising	Yes	N/A	Yes, but only by age, gender or zip code	Yes	N/A
Allows untrue content in political ads?	Yes, but some fact-checking for content not posted by candidates or parties	N/A	Yes, excepts ads that could undermine participation or trust in electoral processes	No	N/A

However, the voluntary content moderation landscape isn't without complexities. Definitions and interpretations vary, leading to differences in practical application. Despite these efforts, the

[64] https://transparency.fb.com/it-it/policies/community-standards/misinformation/
[65] https://www.facebook.com/formedia/blog/third-party-fact-checking-industry-investments

efficacy of such measures remains a subject of scrutiny. Furthermore, while these voluntary commitments reflect the willingness of social media platforms to address disinformation, they are not legally binding. The unique business model of these platforms, where increased traffic translates to higher revenues, underscores the necessity of regulatory oversight to prevent undue influence on content moderation. Though some states have initiated regulatory attempts, significant challenges remain in harmonizing regulatory frameworks across jurisdictions and ensuring effective enforcement mechanisms. Thus, while social media companies have demonstrated commitment to combat disinformation, the broader regulatory landscape continues to evolve in response to the intricate challenges posed by digital information dynamics.

Therefore, it is necessary to address such matter also to legislative and regulatory bodies. The surge in disinformation on social media platforms has prompted legislative and regulatory bodies at both national and international levels to address the pressing need for effective countermeasures. The complexities of disinformation, intertwined with challenges related to freedom of expression and technological advancements, have led to a diverse range of approaches aimed at mitigating its impact, both from a national and an international perspective. In this regard, the Poynter Institute's study[66] provides a guide for existing attempts of global anti-disinformation actions, analyzing each country taken into consideration. Figure 4 summarizes their findings across countries.

Figure 4: global anti-disinformation actions
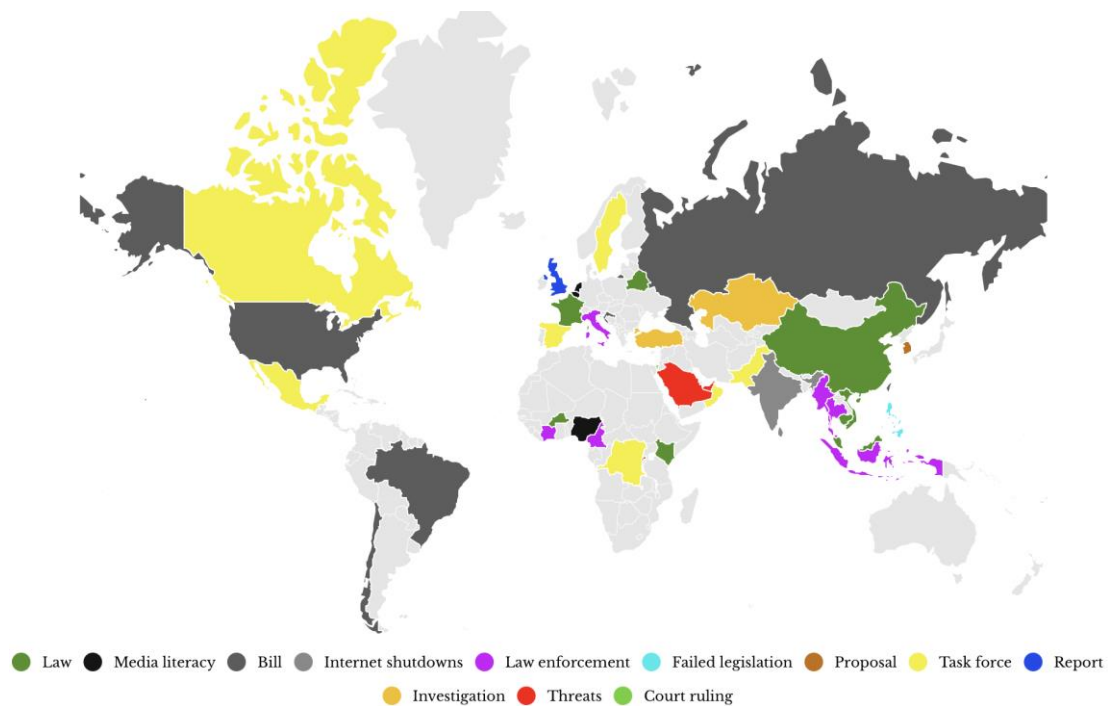
This study maps diverse responses, encompassing legislative amendments, regulations, and initiatives by different nations to counter disinformation, underscoring the multifaceted and

---

[66] https://www.poynter.org/ifcn/anti-misinformation-actions/

evolving nature of legislative and regulatory approaches to address disinformation's pervasive influence. Moreover, by analyzing country-by-country governmental actions, the study provides valuable insights into understanding the more effective ways to prevent and combat disinformation and into highlighting which country exploits such regulation for undermining the right of free speech. One example of the latter situation has been provided – among other countries as well - by China, which has "some of the strictest laws in the world when it comes to misinformation", according to the report. In fact, while some countries focus on transparency and accountability, China's approach to disinformation is shaped by its political context, emphasizing social stability and state control over information. The government seeks to curb the potential harm caused by false information by implementing a combination of legal measures, technology-driven solutions, and strict censorship mechanisms. China's legal framework encompasses various laws targeting the dissemination of false information that could disrupt social order or public safety. The Criminal Law, for instance, criminalizes the spread of rumors or false information that may lead to social panic, empowering authorities to take action against individuals and entities responsible for spreading disinformation. Additionally, China's response to disinformation involves leveraging advanced technologies like AI and machine learning, which are harnessed to identify and remove false information swiftly. Real-time analysis of content patterns and keywords allows for the efficient detection of potentially harmful content while flanked by a unique digital landscape characterized by robust content censorship and control mechanisms. The "Great Firewall" is, in fact, a complex system that filters and blocks content deemed politically sensitive and extends to social media platforms, where the government can promptly remove disinformation and suspend accounts that spread false information. In addition, a law was passed in 2017 that mandates social media platforms to only share and link to news articles from registered news media sources. Furthermore, in 2018, authorities began requiring microblogging sites to identify and dispute rumors on their platforms. Therefore, while its regulatory approach has yielded some success in curbing the spread of disinformation, it raises concerns about freedom of expression. The stringent censorship mechanisms risk stifling legitimate voices and stifling open discourse and the government's oversight of tech companies sparks debates about privacy and surveillance.

Some governments, instead, have chosen to penalize those who disseminate false information, recognizing the potential harm to public discourse and societal stability. For instance, Chile passed legislation in 2019 imposing penalties on politicians and individuals for "the dissemination, promotion or financing of false news".

Transparency and accountability mechanisms have also gained prominence in the battle against disinformation. Countries like Canada and France have sought to enhance the transparency of

political advertising on social media platforms. Through legislation, these countries have mandated that tech platforms create repositories for political advertisements, thereby allowing citizens to better understand the origin and intent of such content. In France, additional measures empower the broadcasting agency to suspend or terminate broadcasters influenced by foreign states if they disseminate false information that threatens electoral integrity. Building upon an 1881 law, in fact, in November 2018, later amended in 2019, a law passed in France to control the spread of disinformation through media before the elections. In particular, this legislative decree includes three primary provisions. Firstly, a judge has been granted the authority to take necessary measures to prevent the dissemination of false information prior to an election. Secondly, social media platforms, including Facebook, Twitter, and YouTube, must disclose the identity of those who have financed sponsored content or campaign advertisements, as well as their associated costs. This requirement is similar to the Honest Ads Act that applies to television and radio in the United States. Lastly, the Higher Audiovisual Council (CSA), which regulates broadcasting, has been bestowed with enhanced administrative and executive powers to ensure that platforms adhere to the legislation. The CSA will also provide periodic reports on the efficacy of the measures implemented by the platforms. In addition, the CSA possesses the unilateral authority to revoke the broadcasting privileges of television and radio outlets operating in France that promote misinformation and are "under the control or influence of a foreign state".

Another prominent approach involves empowering regulatory authorities to oversee and enforce content moderation standards. For instance, Germany, as previously mentioned, introduced the Network Enforcement Act (NetzDG) in 2017, which mandates social media platforms to swiftly remove illegal hate speech and disinformation and which failure to comply can result in hefty fines. However, while this approach aims to hold platforms accountable, it also raises concerns about potential over-removal of content and stifling freedom of expression.

Nonetheless, the global nature of social media requires coordinated international efforts to mitigate the spread of false information that can influence public opinion, elections, and societal stability. Such provisions promoting transparency, accountability, and responsible content dissemination have been advanced and supported by several international organizations and institutions. The UN has recognized the urgency of tackling disinformation to uphold democratic values and human rights. "The Age of Digital Interdependence"[67] report underscores the profound impact of digital technology on societies and emphasizes the need for collective global responses. In this regard, in fact, the report highlights how "*around the world, many people are increasingly – and rightly – worried that our growing reliance on digital technologies has created new ways for individuals,*"

---

[67] https://www.un.org/en/pdfs/DigitalCooperation-report-for%20web.pdf

companies and governments to intentionally cause harm or to act irresponsibly. Virtually every day brings new stories about hatred being spread on social media, invasion of privacy by businesses and governments, cyber-attacks using weaponised digital technologies or states violating the rights of political opponents". Moreover, one of the UN's notable initiatives is the "Verified" campaign that, launched in response to the COVID-19 pandemic, leverages a public-private partnership to enlist trusted individuals as messengers who disseminate accurate information to counter the spread of misleading content. By enlisting individuals as reliable sources, the UN managed to combat false narratives that could have exacerbated the impact of the pandemic. Moreover, also UNESCO recognizes the crucial role of media and information literacy in countering disinformation. The organization promotes media literacy education to empower individuals with critical thinking skills necessary to discern reliable sources from false information. UNESCO's Global Media and Information Literacy Week[68] serves as a platform to advance media literacy awareness and best practices worldwide. By fostering media literacy education, UNESCO aims to equip individuals with the tools to navigate the complex landscape of online information and distinguish between credible and misleading content, which "*is undermining the very foundations of our societies and democracies and putting lives at risk through the propagation of fake cures, the fueling of vaccine conspiracy theories, or the spread of racism and hate speech*". Additionally, the G7 has taken proactive measures to address disinformation and threats to democracy through the establishment of the Rapid Response Mechanism (RRM)[69]. This mechanism, initiated by Canada, aims to coordinate responses to evolving threats to democratic societies, including disinformation campaigns. The RRM underscores the significance of information sharing, best practices, and cooperation to combat the spread of false narratives. By focusing on collaboration among G7 countries, the RRM enhances the collective ability to counter malign influence and protect the integrity of democratic discourse, by sponsoring various institutional partnerships to combat disinformation, foreign interference and foster economic and research security.

The European Union response to the spread of online political disinformation has been widely studied and analyzed, due to its proactive commitment to combating this phenomenon by implementing legislative and regulatory measures. Since this concerted effort involves various key actors, strategies, and initiatives, for the purpose of this thesis I shall not delve into the legal details of each action, as done in the study "*The impact of disinformation on democratic processes and human rights in the world*", but I will provide a general overview of the several EU attempts in

---

[68] https://www.unesco.org/en/weeks/media-information-literacy
[69] https://www.international.gc.ca/transparency-transparence/rapid-response-mechanism-mecanisme-reponse-rapide/index.aspx?lang=eng#a3_1

facing this issue. Its rapid and urgent response is a result both of the awareness raised by Russia's interference with the US elections and the annexation of Crimea and of the fear of online manipulation in the then-upcoming 2019 European Parliament elections. Central to the EU endeavors in combating disinformation is the European External Action Service (EEAS), which has been entrusted with the critical task of countering disinformation. The EEAS spearheads the EU response to foreign disinformation campaigns and is responsible for implementing significant measures such as the Action Plan Against Disinformation[70] and the establishment of the Rapid Alert System (RAS)[71], underlining the EU commitment to addressing disinformation threats in a comprehensive and coordinated manner. The RAS, for instance, acts as a coordination mechanism between the EU and its member states, facilitating the real-time sharing of information and enhancing the EU capability to promptly expose disinformation campaigns. Moreover, acknowledging the significance of accurate fact-checking and in-depth research, the EU established the European Digital Media Observatory (EDMO)[72], a collaborative initiative that serves as a hub for tracking and analyzing disinformation campaigns across Europe. By fostering partnerships between fact-checkers, researchers, and media organizations, EDMO strengthens the region's capacity to identify and counteract false narratives effectively. At the forefront of these efforts is the East StratCom Task Force (ESCTF), an essential unit created indeed in response to Russia's disinformation endeavors aimed at destabilizing the EU political processes. Initially focused on addressing Russian disinformation, ESCTF has expanded its scope to proactively counter disinformation across different regions. This expansion reflects the EU acknowledgment of the broader challenge posed by disinformation campaigns and its commitment to fostering societal resilience against such threats[73]. The EU response to disinformation, however, extends beyond monitoring and identification to encompass strategic actions that fortify democratic processes. For instance, the EU has established a Hybrid Fusion Cell within the EU Intelligence and Situation Centre (EU INTCEN)[74] to analyze unconventional threats, including information campaigns and radicalization efforts. This approach demonstrates the EU recognition of the multifaceted nature of disinformation and its dedication to safeguarding its political and societal resilience. In addition to its collaboration with international partners such as the G-7 Rapid Response Mechanism and NATO, The EU has also devised specific initiatives to address disinformation within its borders. In autumn 2018, the European Commission introduced the aforementioned Action Plan Against

---

[70] https://www.eeas.europa.eu/sites/default/files/action_plan_against_disinformation.pdf

[71] https://ec.europa.eu/commission/presscorner/detail/en/ip_20_1270

[72] https://edmo.eu/

[73] https://www.eeas.europa.eu/eeas/2022-report-eeas-activities-counter-fimi_en

[74] https://www.statewatch.org/media/documents/news/2016/may/eu-intcen-factsheet.pdf

Disinformation, concurrently entering into agreements with major social media companies to adhere to a Code of Practice on Disinformation. While laudable for its aim to promote self-regulation within the tech industry, this Code has also faced criticism for its potential impact on freedom of speech. In this regard, however, The EU has adopted guidelines to safeguard freedom of expression online and offline, thereby enhancing its involvement in preventing violations affecting freedom of opinion and expression. The EU has also initiated roadmaps for engaging with civil society in external relations, allowing for structured dialogues and cooperation that contribute to the EU overarching strategy. The Action Plan for Human Rights and Democracy for 2020-2024[75] serves as a comprehensive roadmap that underscores the EU commitment to countering disinformation and promoting democratic resilience. This ambitious plan not only highlights the importance of countering disinformation and extremist content but also emphasizes the reinforcement of online media literacy and digital skills, underscoring the EU dedication to quality journalism, investigative reporting, and fact-checking as vital components of its strategy. The EU commitment to global human rights leadership is evident in its engagement with civil society, parliamentary collaborations, and cross-border initiatives. The EU influence extends beyond its borders as it collaborates with like-minded countries to push for a UN Convention on Universal Digital (Human) Rights. Through its special committee on Foreign Interference in all Democratic Processes[76], the European Parliament plays a crucial role in advocating for global human rights and coordinating efforts to counter disinformation. Additionally, from a legislative point of view, the EU has solidified its commitment to addressing disinformation through legislative measures, exemplified by the introduction of the Digital Services Act and Digital Markets Act. The former establishes clear rules for online platforms, including obligations to combat disinformation, ensure content transparency, and enhance accountability in content moderation. While the latter focuses on promoting competition in the digital sphere, thereby minimizing the potential for undue influence on the spread of disinformation.

Having showed how states and transnational institutions have raised awareness and tried to combat political gaslighting and disinformation campaigns through social media, in order to assess whether these tactics are effective I shall analyze the Russian example, being it globally the leading state actor employing political gaslighting through social media directed at various foreign entities.

---

[75]https://www.eeas.europa.eu/sites/default/files/eu_action_plan_on_human_rights_and_democracy_2020-2024.pdf

[76] https://www.europarl.europa.eu/doceo/document/A-9-2023-0187_EN.html

Chapter III

Having delved into the nature, actors involved, techniques employed, and effects of political gaslighting advanced through social media, the analysis of the countermeasures that states, international institutions and tech companies have shown the efforts for these agents to worldwide counteract and raise awareness about the threat of such phenomenon. However, the extent of its impact requires further analysis, due to its multi-faceted nature and complexity to detect. For this reason, in this chapter firstly I shall examine how the advent of social media and the perpetration of disinformation campaigns has contributed to the worldwide democratic recess. Using Huntington's theory regarding the reasons behind the democratic recess that followed the first two democratization waves and his anticipating concept of 'electronic dicatorship', I shall demonstrate how the mistrust in the media and their freedom has had a fundamental role in the erosion of democracy worldwide, as highlighted by the former Freedom House report. In particular, I shall focus on the effects that social media have on different regime types, explaining with Schleffer and Miller's study, why they have a weakening effect on strong democratic regimes, an intensifying effect on strong authoritarian regimes, a radicalizing effect on weak democratic regimes, and a destabilizing effect on weak authoritarian regimes (Schleffer & Miller, 2021). Then, I shall take into consideration the case of the Russian Federation, being it globally the first country involved in online foreign influence efforts and disinformation campaigns. By analyzing Russian employment and modernization of Cold War-era 'smart power' in their operations – such as 'Secondary Infektion', their interference in the 2016 US elections, and the spread of their narrative following the invasion of Ukraine – I shall expose how political gaslighting is perpetrated by an authoritarian regime so powerful. Lastly, I shall investigate how the effectiveness of such phenomenon can be measured, concluding that today's attempts are far from satisfying in assessing this factor, requiring further research due to the threat that in poses on democracy and the international order.

Section 3.1 – Social media and the erosion of democracy

The landscape of global politics has been punctuated by a series of seismic shifts, often driven by the ebb and flow of democratic governance. As the third wave of democratization swept across continents, buoyed by aspirations of greater political freedom and civic participation, a prevailing optimism emerged about the prospects of enduring democratic institutions. However, the last decades have witnessed an unforeseen recalibration of this trajectory. The once-steadfast course of democracies has become increasingly intricate, marked by nuances that evoke contemplation and concern alike.

Samuel Huntington's "*The third wave: Democratization in the Late Twentieth Century*" highlighted how between 1974 and 1990 more than 30 countries transitioned to democracies, constituting,

therefore, what he defined the "third wave of democratization". Exploring the reasons behind this phenomenon, Huntington also managed to explain how this transition was in compliance with the then-world order and the following international system: the rise of democracies worldwide has reflected the transition from the Cold War bipolar world order to the "pax americana" that followed the collapse of the Soviet Union. And, in fact, the United States, which emerged as the hegemonic power after the end of the Cold War, managed to maintain its role as the main worldwide democratization promoter and to further the third wave of democratization from different points of view. Apart from the cultural perspective, in which capitalism and liberal democracies emerged globally as the winners against any other regime, affecting international public opinion, the US further established its "policeman of the world" role both politically, with Bill Clinton's global integration, economically, sponsoring international free trade with the creation of the NAFTA and the WTO, and militarily, with the Gulf War and the war in Yugoslavia. Moreover, this unipolar moment remained unchallenged for the following 10 years mainly for the lack of any strategic rivals: while Russia had been defeated and left in a very unstable situation that led to its attempt to bandwagoning on the West with its admission to the Council of Europe and the G7, China had not established in the international order as a super-power. However, while this third wave of democratization unfolded, engendering a wave of optimism about the global trajectory toward democratic governance, and the emergence of democracies in Latin America, Eastern Europe, and parts of Asia was celebrated as a harbinger of stability and progress, the following decades revealed that democratization, while transformative, was far from linear. The oscillation between democratic consolidation and democratic backsliding has become an intricate theme that requires insightful investigation. Referring to Huntington's analysis regarding the factors that contributed to transitioning away from democracy during the first and second reverse waves, he highlighted seven different elements:

1. the weakness of democratic values among key elite groups and the general public;
2. severe economic setbacks, which intensified social conflict and enhanced the popularity of remedies that could be imposed only by authoritarian governments;
3. social and political polarization, often produced by leftist governments seeking the rapid introduction of major social and economic reforms;
4. the determination of conservative middle-class and upper-class groups to exclude populist and leftist movements and lower-class groups from political power;
5. the breakdown of law and order resulting from terrorism or insurgency;
6. intervention or conquest by a nondemocratic foreign power;

7. "reverse snowballing" triggered by the collapse or overthrow of democratic systems in other countries.

Actually, Huntington's theory almost exactly – except for the political orientations of the challengers - predicted how the '90s world order established by the US hegemony was actually contested, leading to 17 consecutive years of democratic recess, according to the Freedom House. The factors elucidated by Huntington in his argument, such as the fragility of democratic values among influential elites and the general public, have been accentuated by the erosion of trust in democratic institutions fueled by the widespread dissemination of disinformation through social media, leading to the distortion of public discourse, diminishing the foundations of democratic deliberation and decision-making. Thus, it is not surprising how in the last 18 years, according to the Pew Research Center, public trust in the US government witnessed a progressing decrease, reaching its historic low in 2019, where only 17% of US citizens claimed that "they trust the government to do what is right just about always or most of the time"[77]. Table 5, in fact, represents the percentage of US citizens who claimed to trust the government.

Table 5: US citizens' trust in the government 1958-2022
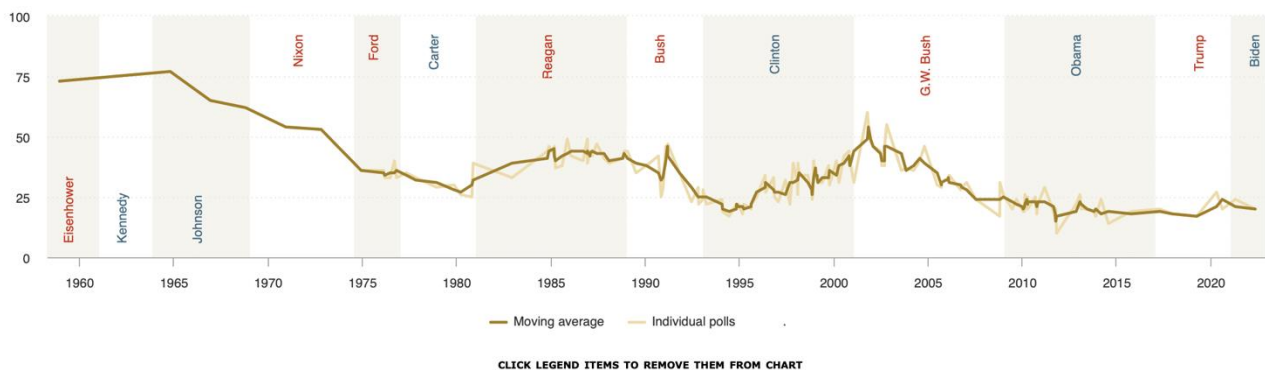
PEW RESEARCH CENTER

Therefore, these data become quite impressive compared to Statista's research[78] showing by country the share of the population who trust their government worldwide. In particular, this report revealed that the countries with the most trust in the government are all considered "not free" by Freedom House – namely China, the United Arab Emirates, and Saudi Arabia, respectively having the trust of 89%, 86%, and 83% of their population – followed by other three nations defined "partly free" – namely Singapore, Indonesia, and India which all scored 76%. However, how these

---

[77] https://www.pewresearch.org/politics/2022/06/06/public-trust-in-government-1958-2022/
[78] https://www.statista.com/statistics/1362804/trust-government-world/

data reflect the relationship between trust in democracy and the media outlet and the overthrowing of the US hegemonic system is still a subject under academic inquiry.

The severe economic setbacks that Huntington highlighted as a potential catalyst for democratic backsliding have materialized in various instances, further strengthening his theoretical underpinnings. For instance, the global financial crisis of 2008 triggered economic disparities, fostering social conflicts that were exploited by authoritarian regimes as an impetus to consolidate power. The rising appeal of remedies advocated by autocratic governments during times of economic turmoil aligns with Huntington's assertion that the allure of authoritarian governance intensifies in periods of socioeconomic distress. Moreover, while the West deteriorated due to the economic crisis originated by the Lehman Brothers' bankruptcy, the rest of the world – namely non-western liberal democracies – increased intensively their power both from an economic and political perspective, exploiting such a dreadful situation. The recognition of the G20, which included also non-western countries, the 2008 Olympic Games held in China, the first BRIC meeting taken in 2009, and the New Silk Road as a response to the interregional trade blocks – namely the TPP and the TTIP – promoted as to exclude China, were only some of the various events that drastically challenged the US hegemony after the 2008 economic crisis. This has led to a striking economic growth rate increase for China, which overcame the US in terms of PPP (Table 6), although it did not surmount the US in terms of GDP (Table 7).

Table 6: China and US economy in terms of PPP (World Bank)
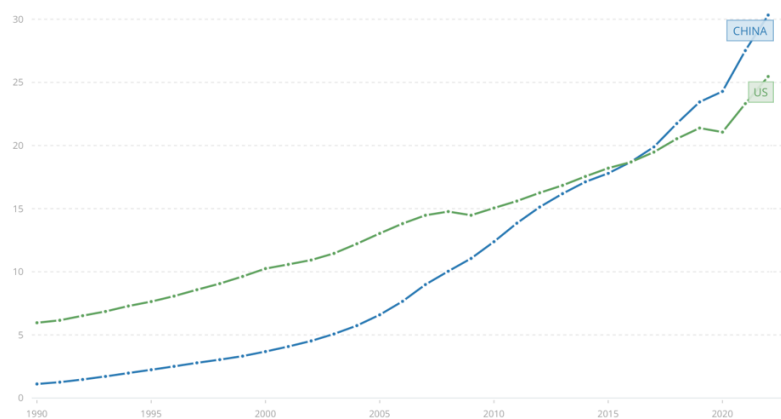
Table 7: China and US economy in terms of GDP (World Bank)
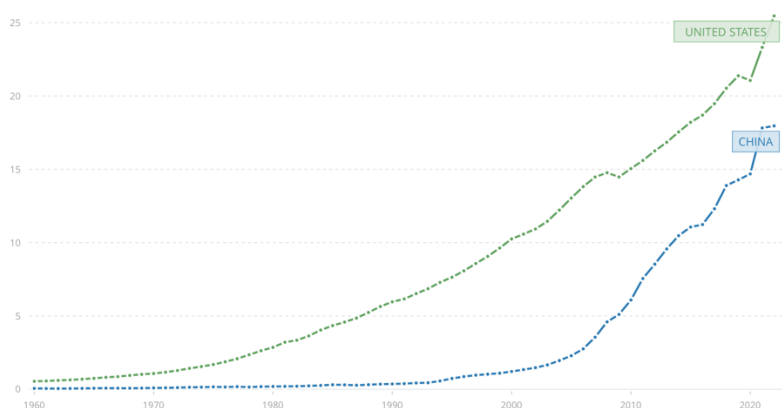
As Huntington observed, in fact, "*if China develops economically under authoritarian rule in the coming decades and expands its influence and control in East Asia, democratic regimes in the region will be significantly weakened*" (Huntington, 1992).

His predictions also resonate in the context of social and political polarization, which has been exacerbated by governments that swiftly introduce sweeping reforms. The 2008 economic crisis wasn't merely an economic phenomenon, since it brought a moment of political fatigue in the West not only due to the increasing power of external rivals, but also due to the challenges from within. Populist, nationalist, and anti-establishment movements of different forces rose in all Western countries – for instance France's Front National, Italy's Movimento 5 Stelle and Lega Nord, the Brexit in the UK, the election of Trump in the US, etc… - causing a "democratic recession", characterized by democratic backsliding, populist surges, and challenges to the liberal norms that underpin these systems (Diamond, 2015). Additionally, the determination of conservative segments of society to thwart these populist movements mirrors Huntington's prognosis of power struggles within democratic contexts.

Also the phenomenon of law and order breakdown due to insurgency or terrorism, as delineated by the author, is reflected in contemporary international relations. Instances of political violence and the ensuing security challenges have provided authoritarian regimes with pretexts to curtail civil liberties in the name of maintaining stability. The 9/11 terrorist attack on the Twin Towers and the Pentagon was actually one of the first events that contested the US unipolar moment after the USSR collapse. Not only did it consist in the first offensive on US soil after Pearl Harbor, but it constituted also a symbolic dramatic attack aimed at undermining the US both from an economic, targeting the Twin Towers in the financial center of New York, and a political perspective, addressing the Pentagon as the political center of Washington DC. Therefore, having caused a national trauma perceived worldwide as an attack on all Western liberal values, it heavily influenced and shaped US foreign politics and the global rise of terrorism's concerns and fear.

Moreover, Huntington's anticipation of the impact of foreign intervention or conquest on democratic systems is exemplified in modern geopolitical scenarios. The interference of nondemocratic foreign powers in the affairs of sovereign states has often subverted democratic processes, aligning with his prediction of external influences on democratic governance. The annexation of Crimea by Russia in 2014 serves as a stark illustration of how external powers can exploit vulnerabilities within democracies to undermine their sovereignty, a threat that has been enhanced by Russia's declaration of war in Ukraine.

Lastly, the concept of "reverse snowballing", wherein the collapse of democratic systems in one country triggers a domino effect, has manifested through the contagion of illiberal practices across regions. The election of far-right leaders in Western democracies, paralleling the ascent of illiberal regimes in Eastern Europe, substantiates Huntington's prognosis of democracy'' demise begetting similar trends regionally and globally.

However, Huntington elaborated a very visionary possibility while describing the factors that are likely to foster a transition from democracy to authoritarianism, warning that "*all of these forms of authoritarianism have existed in the past. It is not beyond the wit of humans to devise new ones in the future. One possibility might be a technocratic 'electronic dictatorship', in which authoritarian rule is made possible and legitimated by the regime's ability to manipulate information, the media, and sophisticated means of communication*" (Huntington, 1992). This prospect strikingly and accurately resonates with the latest Freedom in the World report[79] that highlighted how "*of all the indicators that Freedom in the World uses to assess political rights and civil liberties, freedom of the media and freedom of personal expression have declined the most precipitously since 2005. This assault coincided with the rapid uptake of information and communication technologies that have effectively broken many states' media monopolies*" (Freedom House, 2023). The intricate interplay between the erosion of democratic values, particularly the freedom of expression and the media - as portrayed in the Freedom House report of 2023 stressing how the number of countries and territories that have a score of 0 out of 4 on the media freedom indicator has increased from 14 to 33 during the 17 years of global democratic decline - and the broader trajectory of democratic recession within the international order forms a compelling narrative that resonates profoundly with scholars of international relations. The paramount role of freedom of expression, characterized as a fundamental component of democracy, is underscored by both the Freedom House report and existing scholarly discourse. As the report indicates, this foundational element has borne the brunt of sustained attacks across the globe for nearly two decades. Of notable significance is the concurrent rise of information and communication technologies, which, while shattering traditional

---

[79] https://freedomhouse.org/sites/default/files/2023-03/FIW_World_2023_DigtalPDF.pdf

media monopolies and fostering the democratization of information dissemination, has paradoxically provided authoritarian regimes with new avenues to curtail freedom of expression, as the Chinese example illustrated[80]. This evolving dynamic elucidates a critical aspect of the decline in the third wave of democratization. The pervasive infiltration of new communication technologies into various strata of society, as discussed in international relations literature, has facilitated a transformation in the way information is accessed, consumed, and shared. However, this change has also enabled autocratic governments to mount counteroffensives against independent journalism and dissenting voices. The report's documentation of the expansion of laws targeting the spread of false information, particularly within the context of Russia's media crackdown after the invasion of Ukraine, resonates with Schleffer and Miller's study on the political effects of social media. Both the report and scholarly discourse on international relations elucidate the disconcerting reality that technological advancements can be harnessed to erode democratic safeguards.

In particular, Schleffer and Miller's research offers outstanding insights into the relationship between today's democratic decline and the role of social networks, demonstrating the political effects that these platforms have on different regime types. As a result, they managed to thoroughly explain why American social media platforms have a different impact on democracies and authoritarian regimes, concluding that "*there are four different effects that social media can have: it can have a weakening effect on strong democratic regimes, an intensifying effect on strong authoritarian regimes, a radicalizing effect on weak democratic regimes, and a destabilizing effect on weak authoritarian regimes*" (Schleffer & Miller, 2021). Therefore, after having distinguished the parameters used for defining democratic and authoritarian regimes and their strength[81], they analyzed how the use of social media by three main political actors – namely domestic opposition (dissidents or populist candidates), external forces (other countries or multinational corporations), and the governing regime – determine their effects on the countries in consideration. Table 8 summarizes their findings and the four different political effects that social media have.

---

[80] According to the report, in fact, Chinese "*residents' activities are invasively monitored by public security cameras, urban grid managers, and automated systems that detect suspicious and banned behavior, including innocuous expressions of ethnic and religious identity*" (Freedom House, 2023)
[81] They took into consideration for their research democracies, defined as countries that scored over 74 in the "freedom score", and authoritarian regimes, defined as states that scored under 25 on the same scale. Additionally, they used three different indexes for determining whether a state was to be considered "strong" or "weak", namely the Fund for Peace's Fragile States Index, the World Bank data on Gross National Income, and their total military expenditures.

Table 8: the four types of political effects that social media can have (Schleffer & Miller, 2021)

	Democratic Regime	Authoritarian Regime
Strong State	Weakening effect	Intensifying effect
Weak State	Radicalizing effect	Destabilizing effect

This multi-dimensional approach captures the essence of social media's multifaceted impact on states, ultimately shaping their political trajectories through a complex interplay of state capacity and regime characteristics. In weak authoritarian states, like Egypt in 2011, the destabilization effect emerges as a potent dynamic, wherein social media platforms convert into instruments of empowerment for domestic opposition. In fact, due to their very nature, these platforms enable like-minded individuals to connect with each other, fostering the cooperation and organization of grassroots movements opposing the government's tyranny. Moreover, dissidents are encouraged by Facebook's 'filter bubbles' that, by creating a façade of more support for the cause in question, may incite more people to oppose the authoritarian regime. However, this destabilizing effect is only possible in weak states thanks to their lack of capacity to suppress such protests and maintain the internal order. Contrarily to the weak ones, in strong authoritarian regimes, like Russia and China, social media have an intensifying effect, by enabling government agents to bolster their surveillance on citizens, facilitating the suppression of civil rights and the subversion of domestic opposition. Not only do social media foster the establishment of "digital dictatorships" (Schleffer & Miller, 2021) - as predicted by Huntington with the concept of 'electronic dictatorship' – but they also enable them to exercise their power on other states. In particular, in regard to other liberal democracies, authoritarian regimes employ these platforms to influence democratic elections through the spreading of disinformation, while they cheaply sell their surveillance restrictive practices to weak authoritarian states in order to avoid their destabilizing effect. The research also underscores the radicalizing effect of social media platforms in weak democratic regimes, where populist and anti-establishment candidates and external actors – like Russia – exploit them as guarantors of free speech to disrupt democratic elections. This is possible due to the lack of a suitable system of checks and balances that, not being able to restrict and combat this negative use of social networks, enables such actors to impose an illiberal or authoritarian regime, as happened in Brazil with Jair Bolsonaro. Meanwhile, although strong democratic states do have the capability to appeal to a sufficient system of checks and balances, they are not able to adequately counter the exploitation of these platforms by domestic populist forces or external actors because of their own democratic norms. Therefore, these regime types result weakened by such use of social media,

which fosters the rise of populism and distrust in liberal democratic values, as resulted in the UK with the Brexit referendum and in the US with the election of Trump.

Thus, after having delved into the intricate web that links the 17-years decline of democracy and the different political effects that social media have on different regime types, it is necessary to analyze in depth the cases of Russia, in order to demonstrate the effect that political gaslighting through social media manipulation has worldwife and to assess its effectiveness.

Section 3.2 – The case of Russia

After the scandal of Cambridge Analytica and influence operations "*burst into global public consciousness*" (Gleicher et. al.), Facebook published a report[82] aimed at identifying and countering these operations, which are "*foreign and domestic run by governments, commercial entities, politicians, and conspiracy and fringe political groups*" (Gleicher et. al.). In particular, among the 50 countries found as perpetrators of such operations, Russia was declared the state where most of CIB networks – defined as "*any coordinated network of accounts, pages and groups that centrally relies on fake accounts to mislead Facebook and people using our services about who is behind it and what they are doing*" (Gleicher et. al.) – had origin (27 out of the 150 detected). Therefore, it is no surprise that "*actors operating on behalf of the Federation of Russia are often cited as the most sophisticated and pioneering actors to engage in the manipulation of social media*" (Bradshaw et. al., 2021), nor that Russian propaganda has been defined a "firehose of falsehood" (Paul & Matthews, 2016). In particular, several scholars have claimed how Russian propaganda has its roots in a strategic recalibration of the techniques employed both by the Soviet Union and by the US before the end of the Cold War. Taking the lessons from the US use of "smart power" – defined as a "*smart combination of soft and hard power*" (Nye, 2009) – in defeating the USSR, today's Russia learned how to modernize Soviet propaganda inserting its traditional information warfare into its current hybrid warfare and translating it into the digital realm. In particular, the information warfare perpetrated by the USSR was labeled as 'active measures', a set of "*overt and covert techniques for influencing events and behavior in foreign countries*" that included the spread of disinformation, the support of front organizations, the employment of agents of influence, the diffusion of fake news in non-Soviet media and of forgeries (Cull et. al., 2017). However, in placing such tactics in a broader 'hybrid warfare', a multifaceted concept which definition is still the object of an ongoing academic

---

[82] https://about.fb.com/wp-content/uploads/2021/05/IO-Threat-Report-May-20-2021.pdf

debate[83], the lack of an iron curtain, the end of an ideological war, and the unrestricted reach of new media led to today's post-truth politics. As highlighted by Cull et. al. in their report "*Soviet Subversion, Disinformation and Propaganda: How the West Fought Against it*", while during the Cold War both parties spread likely-true disinformation in order to win a rational debate, today's world completely disregards such rationality, directing such tactics for mere opportunistic and inconsistent purposes aimed at sowing confusion among the public. As they put it: "*Today's Russia is not trying to prove that it is on a path to a greater future, and so it can dispense with facts too. When Putin annexed Crimea, he told the international press there were no Russian soldiers there, even though he knew that everyone knew there were. When Russian warplanes bombed civilian targets in Syria, Putin claimed to be attacking ISIS. In doing so, he was not telling a truthful-seeming lie, but showing disrespect for the very idea of 'facts'. There is something alluring in this: facts, after all, are unpleasant things, reminders of limitations. The very outrageousness of the Kremlin's lies is part of their attraction. Donald Trump's appeal might be partly about this too. Neither Putin nor Trump are attempting to establish a factual, rational narrative*" (Cull et. al., 2017). For this reason, the Russian term "*informatsionnaya voyna*" (information warfare) is used in a holistic sense that crosses many fields, such as computer network operations, electronic warfare, psychological operations, and information operations with the aim of dominating the global information landscape (Connell & Vogler, 2017). Indeed, due to the nature of its authoritarian regime, Russian state agents were able to employ a double-edge information ecosystem that enabled them to gaslight their domestic audience through surveillance and state-controlled media strategies and the international audience through the spread of false narratives and disinformation campaigns aimed at sowing chaos and confusion in the targeted foreign countries. It is therefore quite useful to understand how the Russian (dis)information ecosystem allowed "Putin to gaslight us all" (Baba, 2022).

Therefore, before delving into the most striking cases of Russia's foreign influence operations through social media manipulation, it is necessary to provide a general overview of Russian computational propaganda, which refers to the "*use of algorithms, automation, and human curation to purposefully distribute misleading information over social media networks*" (Woolley & Howard, 2018). Firstly, Russia's computational propaganda presents an intricate organizational structure, where different agents, such as state institutions, intelligence agencies, and media outlets are strictly intertwined into a web of coordinated disinformation. According to the Freedom House,

---

[83] For the purpose of this paper, I will refer to Bilal's definition of 'hybrid warfare' as a "*combination of conventional and unconventional strategies, methods, and tactics in contemporary warfare as well as the psychological or information-related aspects of modern conflicts*" (Bilal, 2021)

loyalist security forces, a controlled judiciary, and a suppressed media environment, all serve as conduits for Russian media manipulation (Freedom House, 2019), laying the foundations for such a multifaceted organizational structure. In regard to state institutions, many strategies have been employed in order to facilitate the spread and support of such propaganda, such as the Karachay-Cherkess Republic seeking to manipulate social media through public tenders, Moscow employing pro-government troll farms, paid activists to flood negative press articles with hundreds of supportive comments, and government employees in various cities being directed to endorse political posts. However, at the heart of these activities is the Internet Research Agency (IRA), an intelligence agency led by Yevgeny Prigozhin, a figure really close to Putin also known as 'Putin's chef'. Emerging in 2013 and later attracting global attention, the IRA fosters fake accounts to disseminate pro-Putin and anti-Western content, having increasingly shifted its targets from domestic political manipulation to FIEs. This troll farm, operating similarly to a marketing agency, boasts specialized departments for graphics, data analysis, search engine optimization, and financing (Boyd-Barrett, 2019). Additionally, in the realm of military intelligence, the Main Intelligence Directorate (GRU) takes center stage, being also implicated in the 2016 US presidential election interference, with Special Counsel Robert Mueller attributing operations to GRU Unit 26165 and Unit 74455, as discovered by the Mueller report. The GRU's operations combined hacking with a network of fake social media personas, grassroots movements and journalists, and dedicated websites (Nimmo, 2018). While smaller in scale compared to the IRA, the GRU's 2016 presidential election interference aimed at mobilizing African American opinions and supporting Russian military endeavors, fostering US political division within the civil society. Moreover, as previously anticipated, due to its nature of an authoritarian regime, Russia has also the power to censor and control its own domestic audience through media. In fact, Russia's media landscape stands completely state-controlled, with state-run or Kremlin-linked entities dominating such field, like RT – once Russia Today – that is registered as a state-owned Autonomous Non-commercial Organization with the Russian Ministry of Justice and is almost entirely funded by the state budget (Nimmo, 2018). Bradshaw et. al. in their report summarizes Russian computational propaganda's organizational form as illustrated by Table 9:

Table 9: Russian computational propaganda's organizational form

Initial Report	Government Agencies	Politicians & Parties	Private Contractors	Civil Society Organizations	Citizens & Influencers
2012	GRU (Units 26165, 74455. 54777), RT, Sputnik, Public Sector workers	President Vladimir Putin	Internet Research Agency (Yevgeny Prigozhin)	Nashi	Evidence Found

Moreover, as highlighted by Cunningham, other state agencies, which originally were not involved in foreign influence operations, like the Federal Security Service (FSB), the Foreign Intelligence Service (SVR), are strictly linked with additional networks of disinformation, through their affiliations to Advanced Persistent Threats (APTs)[84], cyber networks which create and coordinate influence and interference operations. As illustrated by Cowman et. al., Figure 5 shows the major foreign influence networks associated with Russian intelligence services:

Figure 5: major foreign influence networks associated with Russian intelligence services
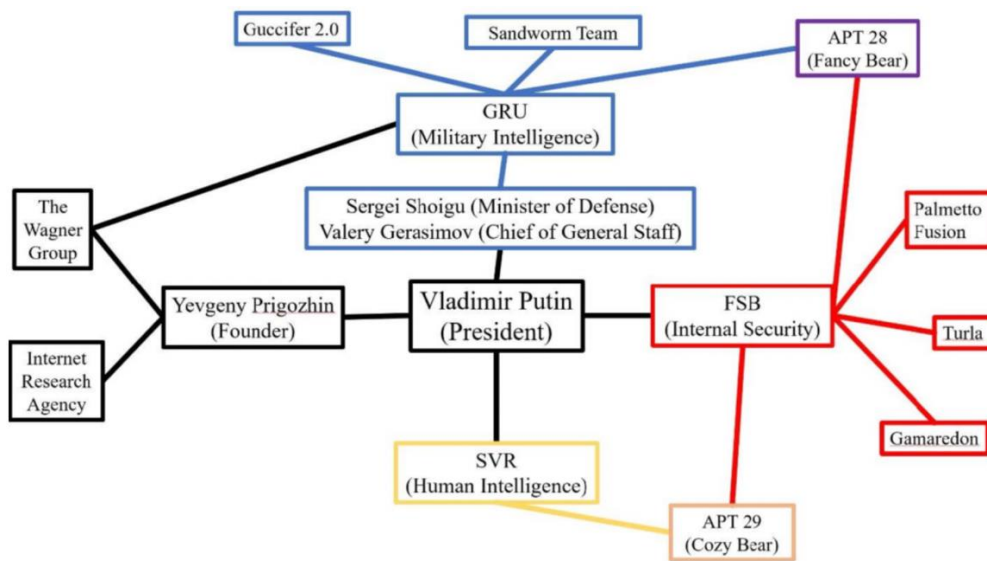
Also their strategies present a multifaceted and holistic approach, being employed in all kinds of online platforms. For this reason, Russian "Full Spectrum Propaganda" – as defined by Nimmo - revolves around the principles of "dismiss, distort, distract, dismay" (Nimmo, 2018). This approach relies on a fusion of both official and fringe media outlets alongside social media operations, seamlessly combining attributed and non-attributed elements in order to create a spectrum of complexity and plausible deniability. As highlighted by Bradshaw et. al., underlying this approach is 'narrative laundering', where a story originates in fringe online communities, gaining legitimacy through repeated citations across various media sources and amplified via fictitious social media accounts. Russian 'full spectrum propaganda' infiltrates almost every significant media, social media, and technology platform, ranging from SMS text messages to Ukrainian soldiers, to

---

[84] Among these, APT28 and APT29 were determined two of the most prolific APTs in Western election interference, beings responsible for the interference in 2016 US elections, the attacks against the German Parliament, and a campaign against Immanuel Macron (Cowman et. al., 2023).

inauthentic activity on Russian service VK among Russian-speaking populations in several countries, and to the use of less widespread online platforms – such as Vine, Pinterest, SoundCloud, Pokémon Go, Tumblr, Reddit, Google (Google+, Gmail, Voice, Ads), Meetup, Medium, Gab and PayPal – to influence the 2016 US elections (DiResta et al., 2018). A cornerstone of Russia's information warfare is the 'hack and leak' strategy, reflecting their comprehensive engagement in the information domain. This approach integrates both offensive cyber capabilities and content strategies, setting it apart from Western tactics, which involves hacking sensitive information or fabricating fake leaked documents and subsequently disseminating them across traditional and social media channels. The GRU's involvement in hacking and leaking emails from the Democratic Party, as well as targeting the World Anti-Doping Agency, exemplifies this which evolved in the 2019 UK general election, where leaked documents were amplified by the Labour Party itself. Additionally, in the complicated task of assessing Russia's computational propaganda's organizational capacity and resources, Mueller in his report found in 2018 at least thirteen Russian nationals and three organizations for "conspiracy" to illegally influence the US presidential campaign (Mueller, 2019). In particular, the IRA alone – thus not taking into consideration the network of affiliates of this agency – was claimed to have a budget of US$1.25 million monthly (Lee, 2018) and have circa 300-400 employees paid higher than average salaries, more or less around 40,000 rubles a month (Graff, 2018). While 'Project Lakhta', which mantra was "information warfare against the United States" (Osborne, 2018), had a budget of more than US$35 million between January 2016 and June 2018 (Osborne, 2018). While Table 10 by Bradshaw et. al. summarizes these findings, these operations constitute only the surface of the capacity of Russian computational propaganda, which loose network makes it almost impossible to have an accurate account.
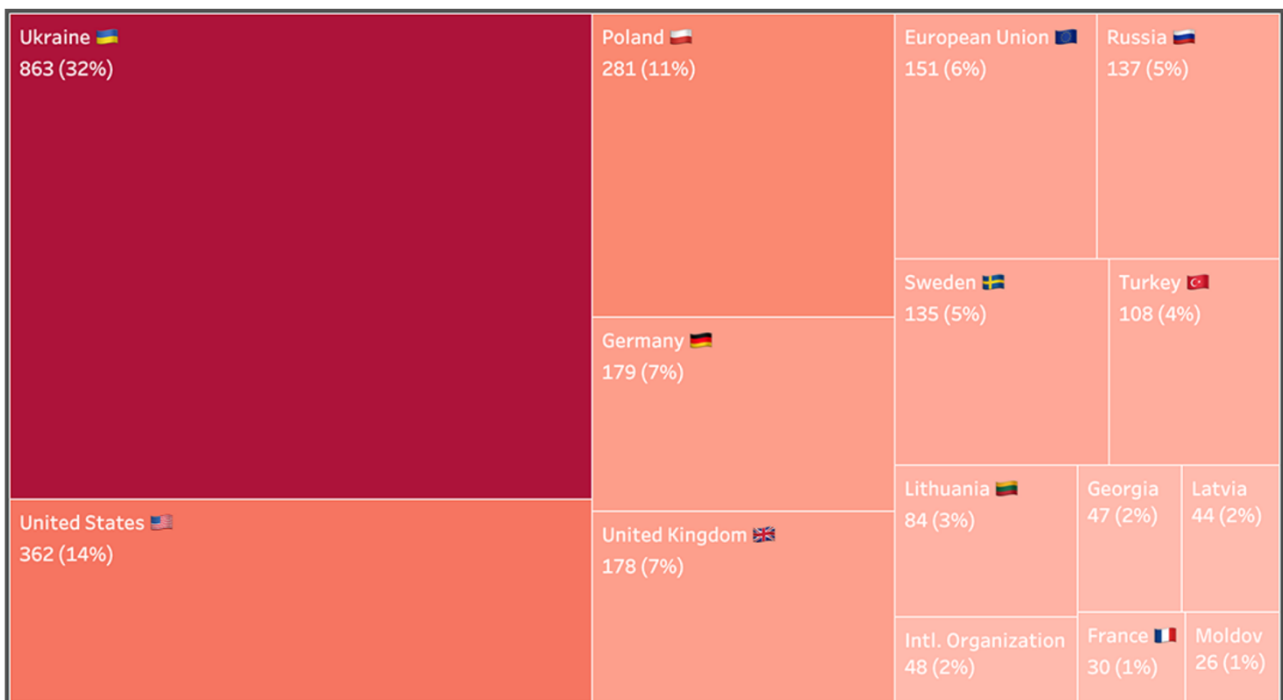
Table 10: Russia's computational propaganda capacity and resources

Team Size	Resources (USD) Spent	Activity Levels	Coordination	Capacity Measure
Multiple teams, IRA (300-1,000)	$35 million for Project Lakhta	High	High	High

Russia's foreign influence operations through social media is such a far-reaching and multifaceted phenomenon that it is almost impossible to detect each operation perpetrated by so many different actors, due to the complexity of its networks, the agents involved, and the cyber tactics employed to hide Russian links. However, several computational propaganda efforts have been exposed, targeting countries all over the world, especially European states, the US, Canada, Australia, Africa, and Central Asia for different purposes. In particular, the 'Secondary Infektion' operation is a prime

example of Russia's evolved tactics in the digital age, where the spread of false narratives and disinformation is used to create confusion, sow discord, and undermine trust in established institutions. This long-running multifaceted Russian (dis)information operation embraced several disinformation and political gaslighting techniques spread through a striking amount of social media platforms targeting countries all over the world. Started in 2014 and carried on until 2020, it was named 'Secondary Infektion' after the East German Stasi operation that the Russian KGB conducted in the 1980s with the aim of convincing the international community that the US had created the HIV in a bio-lab and then spread it worldwide. As Soviet 'Operation Infektion', today's operation involves the spread of forged documents, fake 'leaks', conspiracy theories, made-up stories, and many other disinformation contents aimed at exploiting existing doubts and grievances within target audiences, attacking foreign political figures, eroding trust in democratic institutions, undermining Western countries, polarizing public opinion, and promoting pro-Kremlin rhetoric. Also the targets were as diversified as the content, trying to disrupt 2016 US elections, Turkey's diplomatic dispute with Russia between 2016 and 2018, 2017 French elections, and many more, although Ukraine always remained the main target. According to a report made by Graphika[85], figure 6 represents all the countries that have been targeted by 'Secondary operation', delineating also the number of stories found for each country:

**Figure 6: countries and number of stories targeted by 'Secondary Operation'**

---

[85] https://www.courthousenews.com/wp-content/uploads/2020/06/secondary-infektion-report.pdf

Despite its diversified targets and contents – which ranged from posting articles on the Armenian genocide to accusations of sexual violence committed by Muslim refugees and depicting Angela Merkel as an alcoholic – Nimmo et. al. in their report managed to divide the content into nine categories according to the most prominent theme of each article:

1. Ukraine as a failed or unreliable state;
2. US and NATO aggression or interference in other countries;
3. European divisions and weakness;
4. Elections, especially in the United States, United Kingdom, and France;
5. Migration and Islam;
6. Russia's doping scandals in various sports competitions;
7. Turkey as an aggressive, destabilizing power;
8. Defending Russia and its government;
9. Insulting Kremlin critics, including Aleksei Navalny and Angela Merkel;

In particular, Table 11 shows the number of contents shared for each abovementioned category:

Table 11: number of contents shared by themes
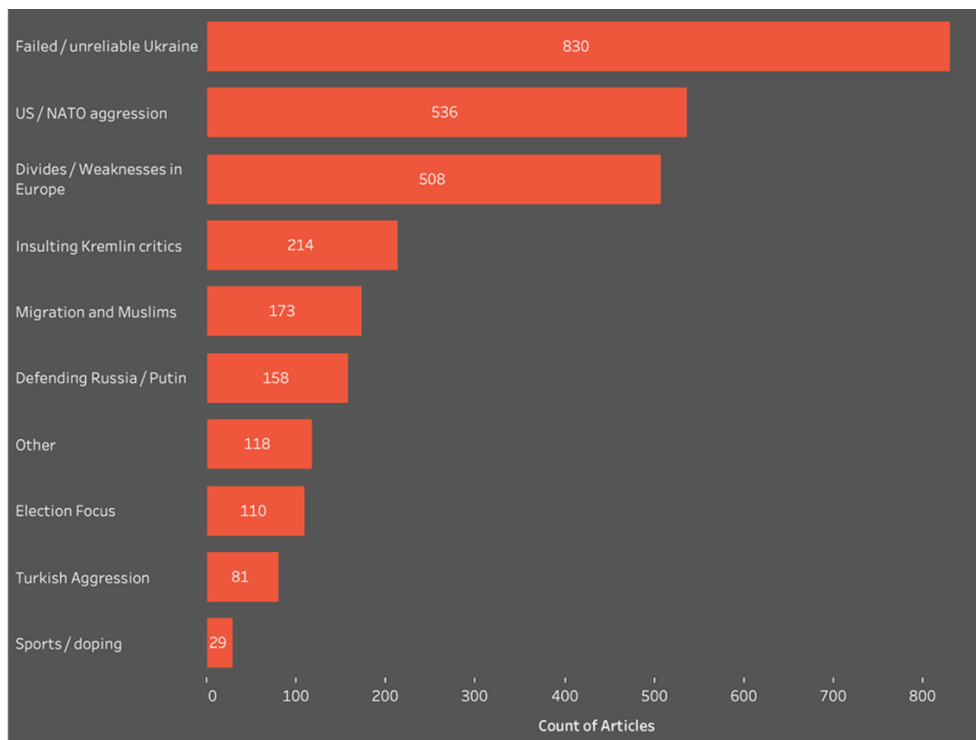
The operations that targeted French President Immanuel Macron are one of the many examples that can be mentioned to understand how 'Secondary Infektion' conducted its disinformation campaigns aimed at gaslighting the French audience by discrediting the figure of Macron. During the 2017

elections, according to Jean-Baptiste Jeangene-Vilmer, Director of the French Institute for Strategic Research (IRSEM), "*there was a coordinated attempt to undermine Macron's candidacy, through a classic three-dimensional information operation: (1) a disinformation campaign consisting of rumors, fake news, and even forged documents; (2) a hack targeting the computers of his campaign staff; (3) a leak—15 GB of stolen data, including 21,075 emails, released on Friday, May 5, 2017— just two days before the second and final round of the presidential election*"[86] (Jeangene-Vilmer, 2019). Additionally at the same time, 'Secondary Infektion' spread three different narratives aimed at undermining Macron's legitimacy and image during the spring of 2017, primarily targeting the issue of immigration, particularly pertaining to Muslims. Employing a well-crafted narrative, Secondary Infektion initiated a story titled "Emmanuel Macron: the new hope of migrants in Europe?"[87], which, being disseminated across platforms, alleged that Macron had criticized Angela Merkel for not taking in migrants, questioning whether this stance would lead to an immigrant influx at French borders. Continuing its agenda, Secondary Infektion then planted a story asserting that Macron prioritized political support from migrants over French national security. This claim, based on a screenshot of a purported "official document", was riddled with linguistic inaccuracies that betrayed its fabricated nature. But, as the final round of the election neared, the operation shifted focus to the "Macron-leaks", painting Macron as an "*embodied synthesis of industrial and banking PR technologies and pan-European red-tape lobbyists*" and a "*synthetic media homunculus that matches all criteria of a negative political selection*"[88]. Although these stories did not gain the necessary virality to have a relevant impact on the population, this narrative showed how Russian agents aimed at casting doubt on Macron's legitimacy through disinformation spread on social media.

On the contrary, another Russian foreign influence campaign perpetrated through social media, which had not only a broader range of contents but also a much more far-reaching impact on public opinion, was the IRA interference with the US 2016 elections. Beginning this campaign on Twitter in 2012, since 2015 "*they reached 126 million people on Facebook, at least 20 million users on Instagram, 1.4 million users on Twitter, and uploaded over 1,000 videos to YouTube*" sharing circa 61,500 unique Facebook posts across 81 Pages, 116,000 Instagram posts across 133 accounts, and 10.4 million tweets across 3841 Twitter accounts (DiResta et. al., 2018). In this context, the New Knowledge report "*The Tactics & Tropes of the Internet Research Agency*", was the first document

---

[86] https://www.atlanticcouncil.org/wp-content/uploads/2019/06/The_Macron_Leaks_Operation-A_Post-Mortem.pdf
[87] http://web.archive.org/web/20200418160806/https://medium.com/@DreuxVachon/emmanuel-macron-a-new-hope-of-migrants-in-europe-b6d07e0a111b
[88] https://archive.is/mHxb6

that thoroughly analyzed this phenomenon, using legal documents and social media platforms' provided data – although the authors admitted how the latter may have been slightly manipulated so that the companies could be held less accountable. However, they managed to address such a difficult issue to detect, examining the various techniques and targets employed in this disinformation campaign. With the aim of spreading distrust toward US candidates and the democratic political system in general (Graff, 2018), IRA agents stole US citizens' social security numbers and identity documents to create fraudulent bank accounts in order to purchase political advertisements, enabling them to significantly spread hashtags like #Trump2016 and #Hillary4Prison and use only 20 Facebook pages to dominate 99% of engagement during the elections (Bradshaw et. al., 2020). However, this IRA interference wasn't only concentrated on promoting Trump's presidency and undermining the figure of Hilary Clinton through the previously analyzed techniques of Russian disinformation campaigns – like the spread of conspiracy theories like the Pizzagate or the lack of Obama's birth certificate as shown in figure 7 – but its themes were accurately chosen on the basis of their targeted section of the population.

Figure 7: an example of a conspiracy theory post shared by IRA that went viral during 2016 US elections (DiResta et. al., 2018)

I can't believe this. Our so-called President, Kenyan illegal bastard Barack Hussein Obama encourages illegal aliens to vote – because as you know law breaking comes naturally to both Democrats and Illegal aliens! "This is not a surprise at all to me" - says Obama. What the hell are you talking about Barry?? This is a CRIME, a VOTER FRAUD! And you just saying that you're not surprised?? Are you encouraging this? You're illegal. You cannot vote. And the President of the United States is saying, 'Don't worry, no one will be spying on you, or catching you.' Why President Obama says so? Apparently because he himself is illegal and cares nothing for this country!! When maybe whether you're for the president, against the president, whether you're pro-immigration reform, anti-immigration reform – you are ignoring the fact that you've been questioned about illegal voting, which you can't do. Why? Because you're not a citizen of this country! Isn't that true, Mr. Obama? I cannot stand the level of corruption that surrounds this administration. Obama and Hillary have committed treasonous acts against our country. They both belong behind bars, not in the White House! Do you agree?

In fact, the IRA exploited the microtargeting, echo chambers, and algorithms features of social media platforms to manipulate in distinctive ways the Black, Left-leaning, and Right-leaning groups over the same timeframe (DiResta et. al., 2018). For those aligned with Right-leaning ideologies, the IRA elaborated a strategy based on the sharing of provocatory content that would have caused extreme anger and suspicion in order to spark the flames of motivation among potential voters. Darkly veiled in conspiracy theories, references to voter fraud, and alarming talk of illegal participation in the election, the IRA's orchestrated posts sought to stoke the fires of rebellion— should the election be 'stolen' from them by the Democrats. In this context, therefore, the IRA's scheme laid in generating an intense emotional reaction driven by anger that would galvanize the masses towards the ballot boxes. Differently, the IRA managed the Black-targeted content with a calculated aloofness towards the election, producing content that focused on the societal alienation

and unending struggles against police brutality that pervaded the African-American experience. Their narrative, a tapestry of defiance and resilience, took on an unexpected turn as the election's horizon neared, intertwining the themes of suppression with voter suppression narratives — persuading the Black community to abstain from voting, to retreat and boycott, painting a landscape where America stood as a forsaken land for its Black citizens. In this case, the IRA managed to infiltrate the electoral behaviors of citizens by just appealing to emotions, but they captured and exploited also the intricate nuances of US sociopolitical dynamics. And then, those swayed by Left-leaning ideologies were targeted with content tinged with anti-establishment attitudes. By sharing posts focused on identity and pride, they targeted communities such as Native Americans, LGBT+, and Muslims, trying to impact and influence their marginalized voices. In fact, beyond these appeals, a broader narrative emerged: one urging the challenging of the status quo or, at least, shifting their votes to candidates other than Clinton. In general, the IRA goals were to engender intense emotional responses of anger and engagement for those most likely to support then-candidate Donald Trump, and to generate disillusionment and disengagement on the Left-leaning and Black communities (DiResta et. al., 2018). By selectively adjusting and spreading disinformation based on the sociopolitical conditions of their different targets, the IRA managed to spread and influence more effectively their support for former president Donald Trump, creating a political environment imbued with fake news regarding his opponents, amplified political divisions, and distrust toward the US political system – therefore, filled with political gaslighting. In DiResta et. al. words, "*throughout its multi-year effort, the Internet Research Agency exploited divisions in our society by leveraging vulnerabilities in our information ecosystem. They exploited social unrest and human cognitive biases. The divisive propaganda Russia used to influence American thought and steer conversations for over three years wasn't always objectively false. The content designed to reinforce in-group dynamics would likely have offended outsiders who saw it, but the vast majority wasn't hate speech. Much of it wasn't even particularly objectionable. But it was absolutely intended to reinforce tribalism, to polarize and divide, and to normalize points of view strategically advantageous to the Russian government on everything from social issues to political candidates. It was designed to exploit societal fractures, blur the lines between reality and fiction, erode our trust in media entities and the information environment, in government, in each other, and in democracy itself. This campaign pursued all of those objectives with innovative skill, scope, and precision*" (DiResta et. al., 2018).

Coming from this experience, the US 2020 elections were expected to be a target for further Russian foreign online influence efforts and, although many social media platforms tried to prevent such phenomenon through the policies I have analyzed in the previous chapter, the report "Foreign

Threats to the 2020 US Federal Elections"[89] by the US National Intelligence Council has found that "*President Putin and the Russian state authorized and conducted influence operations against the 2020 US presidential election aimed at denigrating President Biden and the Democratic Party, supporting former President Trump, undermining public confidence in the electoral process, and exacerbating sociopolitical divisions in the US*" (N.I.C., 2021). In particular, although Russia's online foreign influence efforts were not on the same scale as the ones perpetrated in 2016 US elections, the National Intelligence Council highlighted that "*throughout the election cycle, Russia's online influence actors sought to affect US public perceptions of the candidates, as well as advance Moscow's long-standing goals of undermining confidence in US election processes and increasing sociopolitical divisions among the American people*" (N.I.C., 2021), using some of the strategy previously employed in the former US elections – such as spreading misleading or unsubstantiated allegations against President Biden, amplifying mistrust in the electoral process by denigrating mail-in ballots, and deflecting their influence on 2016 elections to Ukrainian government.

A completely different case, however, has been the political gaslighting that the Kremlin tried to perpetrate across the global public opinion regarding Russia's invasion of Ukraine. From spreading Putin's speeches on Youtube in which he shared his own narrative regarding the 'special military operation' in Ukraine - claiming it was "*in accordance to the Article 51 Chapter VII of the UN Charter*" in order to "*protect people who have been subjected to abuse and genocide by the regime in Kyiv for eight years [...] bringing to justice those who committed numerous bloody crimes against civilians, including citizens of the Russian Federation*"[90] (Putin, 2022) – to sharing on TikTok manufactured videos of alleged real-time battles (Kern & Sott, 2022), Russian invasion of Ukraine has been the *hapax* of online disinformation and political gaslighting[91]. Not only did the Kremlin tried to globally spread a narrative depicting Russia as a knight in shining armor finding itself "forced to help" the victims of atrocities supported by Western politicians and, in particular, NATO interests (Putin, 2022) – therefore, deflecting the responsibility to take actions on the "*allegedly winners of the Cold War*" – but they also created and spread throughout all social media conspiracy theories aimed at undermining and depicting their opponents as the real offenders and villains – for instance, by claiming that the Ukrainian president Volodymyr Zelenskyy was an agent of the "New World Order" or that the United States had funded bioweapons research in Ukraine

---

[89] https://int.nyt.com/data/documenttools/2021-intelligence-community-election-interference-assessment/abd0346ebdd93e1e/full.pdf

[90] http://en.kremlin.ru/events/president/news/67843

[91] An example of the impact of this phenomenon has been US Political Counselor John Kelley who accused Russia of "gaslighting", after Moscow had summoned on March 14, 2023, the UN Security Council over 'Russophobia' in an attempt to further justify the invasion of Ukraine among western countries (Massoud, 2023).

(Thompson, 2022). A cue of Russian government's use of social media to spread pro-Russia disinformation is shown by their Twitter operation circa the time of Ukraine's invasion: 75 official Russian government accounts – having 7,366,622 followers, being re-tweeted 35.9 million times, receiving 29.8 million likes and 4 million replies – only during the week of the invasion posted 1,157 tweets, the majority being false narratives on Ukraine (Thompson, 2022). Table 12 shows the daily tweet volume from January 1, 2022, to March 5, 2022, distinguishing true and false tweets about Ukraine.

Table 12: daily tweet volume in 2022 of Russian government accounts
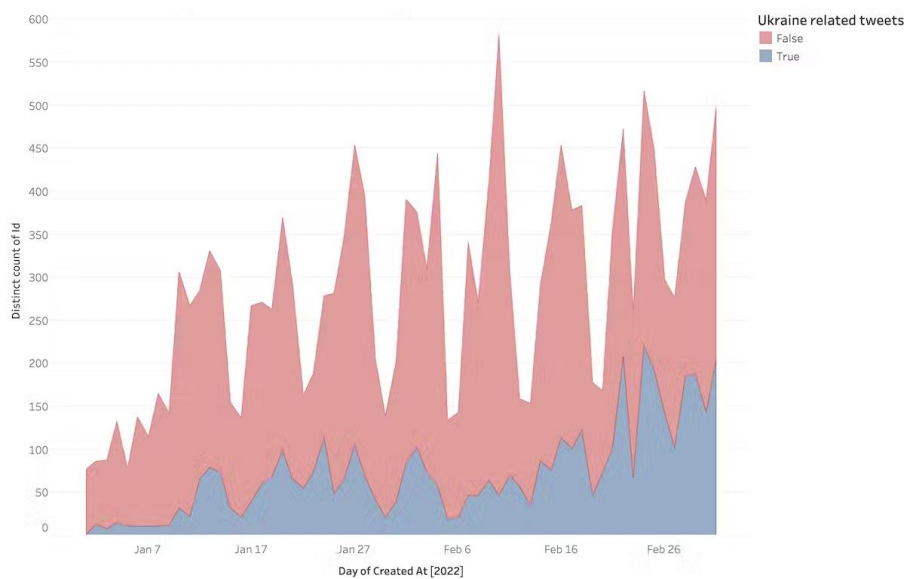
In particular, the disinformation narratives that were mostly spread through Twitter regarded the delegitimization Ukraine as a sovereign state, sowing doubt and mistruths about the Ukraine government and neo-Nazi infiltration, the spread of "whataboutisms"[92] that downplay the Ukraine invasion by drawing attention to alleged war crimes by other countries, and the dissemination of conspiracy theories about Ukraine/US bioweapons research (Thompson, 2022). In addition to this, the NewsGuard's Russia-Ukraine Disinformation Tracking Center has found 395 websites – up to 25 August 2023 – spreading disinformation and pro-Russian propaganda according to at least 130 narratives about the Russian-Ukraine conflict. Through a multi-layered strategy, therefore, their report shows how government-funded and operated websites use digital platforms such as YouTube, Facebook, Twitter, Telegram and TikTok to introduce, amplify, and spread false and

---

[92] According to the Merriam-Webster dictionary 'whataboutisms' refers to "*the act or practice of responding to an accusation of wrongdoing by claiming that an offense committed by another is similar or worse*" (https://www.merriam-webster.com/dictionary/whataboutism)

distorted narratives across the world (*Russia-Ukraine Disinformation Tracking Center - NewsGuard*, 2023). As shown by this report, some of these myths include:

- Russia not stealing grain from Ukraine or blocking shipments, as the West has charged;
- The stage of the massacre of civilians in Bucha;
- Videos showing the "Ghost of Kyiv" shooting down Russian planes;
- The support of Ukrainian authorities to Nazism;
- The genocide of Russian-speaking residents in Donbas;
- The US network of bioweapons labs in Eastern Europe.

If this attempt is enough to sow chaos and confusion among an international audience, part of which lives with access to the internet and the possibility to fact-check the news they encounter online, it is unimaginable the effects that such manipulation has on Russian citizens, who are also embodied in an environment characterized by intense anti-war sentiment's monitoring and heavy censorship[93]. Therefore, the only question that remains unanswered is the following: how effective are such tactics? Are disinformation campaigns and political gaslighting through social media able to shape international or foreign public opinions? In the next section, I will try to provide an attempt in assessing the effectiveness of such a phenomenon that completely embraced the global online space.

Section 3.3 – How much is it effective?

In the previous sections and chapters I have already examined the impact and effects that political gaslighting has both on a domestic level – for example, the inadequate US COVID-19 response due to the 'Trumpian gaslighting'[94] – on an international point of view and in respect to democracy. However, determining the effectiveness of such operations, like the online foreign influence efforts made by Russia during the US elections, is almost impossible to assess quantitatively, for which reason in the academic community there is a huge gap regarding this attempt, where no one has been able to find an accurate measurement. Theoretically speaking, in fact, the best way to assess whether political gaslighting and online FIEs had been effective would be a quantitative analysis

---

[93] Striking examples of this can be represented by Roskomnadzor's censorship, enacting that media organizations could only publish information from official government media outlets on the war, the restricted access to the Internet for Russian citizens, the arrest of several journalists and citizens for their reporting or public comments, etc…

[94] Trump firstly claimed that the pandemic was the "new hoax" of the Democratic Party, then he miscredited the death toll in January 2021 claiming the deaths were not real, and kept denying the realities of the COVID-19 emergency leading to poor policies to face the pandemic (Wild, 2023).

case-specific based on a plethora of gathered data. For instance, in the case of Russia's interference in 2016 US elections, firstly one should have gathered data about the voting preferences of individuals, representing all the different targets that they addressed, before their exposure to such disinformation content. Second, one should be able to estimate the nature and the amount of IRA's manipulated content they have been exposed to, distinguishing it from non-Russian online domestic disinformation campaigns. Then, one should analyze their current voting behaviors and compare them with the previous ones, in order to assess whether these operations had been effective in changing their perception or behavior. It is evident that this constitutes an impossible task for several different reasons, first of all lack of data. As the Director of Investigations of Graphika, Ben Nimmo, has claimed "*unless you know the baseline and unless you actually have a very reliable way of judging what people were thinking before and after, you can't measure the change that went on*"[95] (Nimmo, 2022). Moreover, this system entails an additional obstacle, namely the lack of ability to establish a causal relation between these factors. Let's suppose that we do have the possibility to analyze all the necessary data previously mentioned in order to assess a change in voters' perception and behavior, we would still not be able to determine whether this shift has happened due to a specific foreign disinformation campaign since such operations "*don't exist in a vacuum, they're surrounded by all the real world events and all the other considerations that people go through in their daily lives*" (Nimmo, 2022). Therefore, such change is affected by all the other factors that influence individual decision-making, among which disinformation campaigns would result at most as just a contributor. For instance, as Ellehuus has argued, after the aforementioned case of the Skripal poisoning in the UK and months of an aggressive Russian information operation, despite the perceptions of Russia's guilt in the attack decreased from 65 percent to 55 percent, it is impossible to determine whether this shift of opinion was caused by this operation or by the government's inconsistent messaging on the facts of the case (Ellehuus, 2020). Another quantitative method that one could consider in assessing the effectiveness of this phenomenon may consist in analyzing the engagement statistics on social media: the more a content is shared, liked, or reposted the more it is likely to be believed and the more impact it can have on people's opinion. However, even this technique presents two insurmountable obstacles – the first being the so-called 'online marketing fake engagement', as highlighted by Nimmo, and the second being, again, the impossibility of establishing a causal relationship. In regard to the former, in fact, as it has been witnessed in the operation 'Secondary Infektion' and the Chinese 'Spamouflage Dragon', likes and shares can be easily bought online, so that while a post or a campaign may seem

---

[95] https://www.futurelearn.com/info/courses/combating-disinformation-and-fake-news/0/steps/157095

effective, in reality, the only accounts that engaged with that content were fake accounts (Nimmo, 2022). The latter, instead, involves the fact that whether an individual likes or shares an online content on social media by any means does it mean that s/he will actually believe in it – that is why, despite the statistics showing how much fake news creates more engagement on social media, this does not reflect at all the credibility of them, since people may share it for a variety of reasons, one of which being, for instance, satire. Therefore, we can conclude that "*there are problems with using the numbers as basic metrics*" (Nimmo, 2022). Thus, all we have left is a 'rule of thumb', as Nimmo declared, in assessing the effect that a disinformation campaign can have, revolving around its intent, consumption, and spread, focusing in particular on the targeted communities, the platforms on which it is shared and whether it is amplified by mainstream media and public figures. For example, the fact that Russian troll farm posts during 2016 US elections had been quoted by traditional media or publicized in politicians' election campaigns has enabled this disinformation campaign to become much more effective, being no more limited to a particular group and being 'legitimized' by public figures. As described by Ellehuus, this 'rule of thumb' has been employed in the British HMG's RESIST framework, where these structured analysis techniques provide insights in the potential effectiveness of a disinformation campaign. In fact, the RESIST process involves firstly the recognition of the goals of such operation – whether they consist of economic gains, discrediting experts and leaders, increasing polarization, undermining national security and wellbeing – and the likelihood that they will be achieved – for example, whether it poses a significant risk to the general public in terms of national security, public safety, public health, or the ability to have a fair and balanced debate (Ellehuus, 2020). This assessment, backed up by an analysis of the reach and modalities in which this disinformation operation will spread online, enables individuals to understand the possible impact of specific disinformation and online FIEs. However, it can be considered far from an accurate and comprehensive method that could be applied indiscriminately regardless of the circumstances, for which reason the threat of disinformation is so difficult to assess and counteract, requiring further inquiry and studies aimed at developing an infallible system in this field. In regard to political gaslighting specifically, in addition to all the factors previously mentioned, one should also implement and inquire the relationship of at least these elements, namely the trust that people have in social media and information shared on them, whether they feel they have access to free and reliable information, their media literacy, whether they are aware of this political manipulation, and the trust they have on political public figures, international institution and democracy itself. Until a system is developed that integrates all of these factors and accurately enables people to investigate their relationship, not only do disinformation and political gaslighting through social media remain a

visible threat worldwide for democracy and the international order, but also we are still not able to determine the damage that this phenomenon has already caused and will cause.

Conclusion

In the pervasive sphere of international relations, where the dynamics of power, diplomacy, and statecraft have traditionally shaped the course of history, the entire international community finds itself riding the wave of a digital revolution that has fundamentally altered the features and principles of global politics. Aiming at analyzing and unraveling the – now commonly known - complex phenomenon of political gaslighting and digital disinformation, this thesis has delved into an intricate exploration of power, deception, and manipulation in the age of social media and international online interconnectedness. As I conclude this academic journey, this paper shall provide some illuminating insights while raising more doubts and uncertainties that still characterize the study of global political gaslighting and international digital disinformation. This research began by revisiting the fundamental question "how does political manipulation through social media shape the international order?", highlighting the ways in which such manipulation transformed into the phenomenon of political gaslighting and the threats it poses on both democracy and the international system. In pursuit of an answer, I have employed a multidimensional approach across various fields of international relations, delving into different IR theories, psychological dynamics exploited by political actors, sociological media theories, cyber-political influence, and democratic recessions. The aim of such a multidimensional approach was to dissect this question layer by layer, trying to untangle any aspect that contributed to its spread and effectiveness, in order to provide a more comprehensive basis for future research focused on finding more sophisticated and effective methods to counteract this phenomenon.

For this reason, I started this thesis looking at and analyzing the heart of international relations theories, where I have given a general overview of the foundational concepts of realism, liberalism, cultural hegemony, and constructivism, focusing primarily on their different understanding and interpretation of the concept of power. Among these theories, I have uncovered a compelling narrative logically resonating with the topic under discussion, namely the rise of media as a primary source of power in the contemporary international order. The intertwining of cultural hegemony and constructivism has indeed unveiled and explained the intricate web of influence through which media, both traditional and – especially – digital, interweaves with global politics, influencing public perceptions, shaping identities, challenging international societal norms, and molding worldviews.

With this understanding, I have provided a theoretical framework that was necessary in order to conscientiously explore further the socio-political developments that the digital age has brought into our daily lives, reaching the unsettling conclusion that political gaslighting, a manipulation technique with deep historical roots, has evolved into a menacing force in today's post-truth world.

Through the lens of international relations, I managed to generally explain and highlight the threat that today's political gaslighting, now reaching a global audience, poses to democracy and the international order. The precision-engineered dissemination of false narratives, the strategic erosion of trust, and the weaponization of misinformation have ushered in an era where truth itself is under siege.

Then, the increasing awareness of political gaslight has led me to navigate the shadowy sphere of digital disinformation, the main downside of the ongoing information revolution. In the second chapter, I have delved into the intricate modalities through which this phenomenon is perpetrated and the various forms it can employ on social media platforms. Therefore, it became evident how digital disinformation campaigns are not homogenous endeavors, but they consist of intricate webs involving a multitude of actors: foreign governments, state agents, transnational networks, multinational corporations, big tech companies, international institutions, and society as a whole are all entangled in it, sometimes knowingly and at times unwittingly. Thus, I have examined how each of these actors is involved in spreading or counteracting this phenomenon, highlighting the monumental challenges that democratic societies and international bodies face in their fight against online disinformation. While striving to safeguard democratic principles and the international order and facing the challenges regarding the economic interests of the big tech corporations and the legislative gaps in international law, they find themselves repeatedly trying to find a compromise in the delicate sphere of freedom of expression and information flow, cautiously avoiding the correspondent downfalls.

In the last chapter, I have focused on the democratic recess that has characterized the last 17 years, emphasizing the Freedom House's link between this phenomenon and the development and management of the media. Building upon several reports and Samuel Huntington's notion of a "third wave of democratization", I have discerned the disturbing patterns of democratic decline that have swept across nations, increasingly going towards an "electronic dictatorship". As an example of the effectiveness and reach of digital disinformation, I have thoroughly examined the initiatives employed by the Russian Federation, which has asserted itself as the main actor perpetrating global disinformation campaigns both domestically and, above all, internationally. Their full-spectrum, sophisticated and multifaceted approach, in fact, served to the entire international community as a stark reminder of the capability that this phenomenon has in eroding the foundations of democracy and people's epistemic autonomy.

Yet, despite the comprehensive exploration and understanding of political gaslighting and digital disinformation campaigns, the question of their assessment lingers like a phantom in the dark, leaving the international community with a paradoxical reality. The answer to the question

regarding the modalities to measure the effectiveness of global political gaslighting or international disinformation campaigns though social media remains elusive – much like the phenomenon itself. Indeed, existing methodologies and metrics do not sufficiently manage to capture and assess all the factors involved in this complex interplay of psychology, sociology, technology, and geopolitics. I shall recognize that the very elusiveness of an effective assessment method underscores the insidious nature of these manipulative techniques, for which the global community, academics, policymakers, and digital platforms themselves continue to grapple with this perplexing challenge. Therefore, I shall conclude this thesis with a call to action – a call for continued scholarly inquiry, interdisciplinary collaboration, and international cooperation. However, I shall also highlight another aspect that has been underestimated both in my thesis and by all the other actors involved: the international community. As I have stated at the beginning of this thesis, people are now empowered. They are not merely passive receivers of information or governmental decisions. People have the power to actively take part in this tsunami of information, which was the starting point for phenomena like global political gaslighting to spread in the first place through social media. Perpetrators of this manipulation have exploited this universal power with which everyone has been entrusted: since – in many cases – they could not take away that power, they have managed to make people 'feel' powerless. It is now time to give back to people the awareness of their power and their confidence in their ability to make informed decisions and to access truthful information. Only with the involvement of the public – the gaslightees – political gaslighting efforts and disinformation campaigns will become mere memes to make fun of. Therefore, my future research will be focused on finding the most effective way in which the international public community could regain their trust in facts, truthful information, democracy, the international order, and their own epistemic autonomy in order to end this pervasive shadow of political manipulation.

Bibliography

Abbot, C. (2012). Bridging the Gap - Non-state Actors and the Challenges of Regulating New Technology. *Journal of Law and Society*, *39*(3), 329–358. https://doi.org/10.1111/j.1467-6478.2012.00588.x

Abramson, Kate (2014). *Turning up the lights on gaslighting*. Philosophical Perspectives 28 (1):1-30.

Address by the President of the Russian Federation. (2022, February 24). President of Russia. http://en.kremlin.ru/events/president/news/67843

Adler, E., & Pouliot, V. (2011). *International practices*. Cambridge University Press.

Anderau, G. (2021). Defining Fake News. *Journal of Philosophy*, *35*(3), 197–215. https://doi.org/10.1515/krt-2021-0019

Arendt, H. (1968). *The origins of totalitarianism*.

Baba, A. (2022, February 16). Putin is gaslighting us all. *The Independent*. https://www.independent.co.uk/voices/putin-gaslighting-invasion-ukraine-biden-russia-b2016847.html

Bartlett, R. C. (2020). *Against Demagogues: What Aristophanes Can Teach Us about the Perils of Populism and the Fate of Democracy, New Translations of the Acharnians and the Knights*. University of California Press.

BBC News. (2017, September 6). Rohingya crisis: Suu Kyi says "fake news helping terrorists." *BBC News*. https://www.bbc.com/news/world-asia-41170570

Beck, C.J. (2008). The Contribution of Social Movement Theory to Understand Terrorism. *Sociology Compass*, *2/5*, 1565–1581. https://doi.org/10.1111/j.1751-9020.2008.00148.x

Bell, P. (2022, June 6). *Public Trust in Government: 1958-2022 | Pew Research Center*. Pew Research Center - U.S. Politics & Policy. https://www.pewresearch.org/politics/2022/06/06/public-trust-in-government-1958-2022/

Bennett, W. L., & Paletz, D. L. (1994). *Taken by storm: The Media, Public Opinion, and U.S. Foreign Policy in the Gulf War*. University of Chicago Press.

Berger, J.M. (2015). *Jihad 2.0: social media in the next evolution of terrorist recruitment*. Washington, DC: US Senate Committee on Homeland Security & Governmental Affairs

Bilal, A. (2021). NATO Review - Hybrid Warfare – New threats, complexity, and 'Trust' as the antidote. *NATO Review*. https://www.nato.int/docu/review/articles/2021/11/30/hybrid-warfare-new-threats-complexity-and-trust-as-the-antidote/index.html

Boyd-Barrett, O. (2019). Fake news and 'RussiaGate' discourses: Propaganda in the post-truth era. *Journalism*, *20*(1), 87–91. https://doi.org/10.1177/1464884918806735

Bradshaw, S., & Howard, P. (2017). Troops, Trolls and Troublemakers: A Global Inventory of Organized Social Media Manipulation. In Computational Propaganda Research Project (pp. 1–37). Oxford Internet Institute.

Bradshaw, S., Bailey, H., & Philip N. Howard. (2021). "*Industrialized Disinformation: 2020 Global Inventory of Organized Social Media Manipulation.*" Oxford, UK: Programme on Democracy & Technology. demtech.oii.ox.ac.uk.

Brahms, Y. (2020). *Philosophy of Post-Truth*. Institute for National Security Studies. http://www.jstor.org/stable/resrep23537

Caesar, G. (2020). *De Bello Civili: The Civil Wars*.

Caesar, J. (2015). *Commentarii de Bello Gallico*. Andesite Press.

Carpenter, A. (2018). *Gaslighting America: Why We Love It When Trump Lies to Us*. HarperCollins.

Castells, M. (2000). *The Rise of The Network Society: The Information Age: Economy, Society and Culture*. Wiley-Blackwell.

Castells, M. (2005). Global governance and global politics. *PS: Political Science and Politics*, *38*(1), 9–16.

Chavalarias, D. (2022). *Toxic data: comment les réseaux manipulent nos opinions*.

Cicero, M. T. (2011). Cicero: *De oratore* Book III. In Cambridge University Press eBooks. https://doi.org/10.1017/cbo9780511778292
CNRS News. (n.d.). CNRS News. https://news.cnrs.fr/articles/investigating-climate-sceptics-disinformation-strategy-on-twitter

Connell, M., & Vogler, S. (2017). *Russia's approach to cyber warfare*.

Connor, W. R. (1987). Tribes, Festivals and Processions; Civic Ceremonial and Political Manipulation in Archaic Greece. *The Journal of Hellenic Studies*, *107*, 40–50. https://doi.org/10.2307/630068

Cooper, C. (2019). Deciding the Facebook question. *Issues in Science and Technology*. https://issues.org/deciding-the-facebook-question/

Council, N. I. (2021). *Foreign threats to the 2020 US federal elections*.

Cowman, C., Hernandez, A., & Singh, J. (2023). Russian and Chinese influence actors and operations against the American electorate. *GDIL | Global Disinformation Lab at UT Austin*. https://gdil.org/russian-and-chinese-influence-actors-and-operations-against-the-american-electorate/

Cox, R. W. (1996). *Gramsci, Hegemony, and International Relations: An Essay in Method (1983).*

Cull, N. J., Gatov, V., Pomerantsev, P., Applebaum, A., & Shawcross, A. (2017). *Soviet Subversion, Disinformation and Propaganda: How the West Fought Against it*. LSE.

Cunningham, C. (2020). *A Russian Federation Information Warfare Primer*. University of Washington, The Henry M. Jackson School of International Studies.

Daalder, I. H., & Lindsay, J. M. (2003). *The Globalization of Politics: American Foreign Policy for a New Century.* Brookings. https://www.brookings.edu/articles/the-globalization-of-politics-american-foreign-policy-for-a-new-century/

Dahl, R. A. (1957). *The concept of power*.

Dalpino, C. E. (2001). *Does Globalization Promote Democracy?: An early assessment* Brookings. https://www.brookings.edu/articles/does-globalization-promote-democracy-an-early-assessment/

Deibert, R. J. (1997). *Review: The Global Political Economy of Communications*. International Journal, 52(2), 384–385. https://doi.org/10.1177/002070209705200217

Diamond, L., & Plattner, M. F. (2015). *Democracy in decline?* JHU Press.

DiResta, R., Shaffer, K., Ruppel, B., Sullivan, D., Matney, R., Fox, R., Albright, J., & Johnson, B. (2018). *The Tactics & Tropes of the Internet Research Agency*. New Knowledge. https://disinformationreport.blob.core.windows.net/disinformation- report/NewKnowledge-Disinformation-Report-Whitepaper.pdf

DiResta, R., Shaffer, K., Ruppel, B., Sullivan, D., Matney, R., Fox, R., Albright, J., & Johnson, B. (2018). *The Tactics & Tropes of the Internet Research Agency*. New Knowledge. https://disinformationreport.blob.core.windows.net/disinformation- report/NewKnowledge-Disinformation-Report-Whitepaper.pdf

Disinformation. (n.d.). https://transparency.fb.com/it-it/policies/community-standards/misinformation/

Dobber, T., 'O Fathaigh, R., & Zuiderveen Borgesius, F.J., (2019). The regulation of online political micro-targeting in Europe, *Internet Policy Review.* 8(4).

Duca, L. (2016, December 10). *Donald trump is gaslighting america*. Teen Vogue. Retrieved February 11, 2023, from https://www.teenvogue.com/story/donald-trump-is-gaslighting-america.

Elam, C. (1988). Art and Diplomacy in Renaissance Florence. RSA Journal, 136(5387), 813–826. http://www.jstor.org/stable/41377296

Ellehuus, R. (2020). Measuring Impact. In *Mind the Gaps: Assessing Russian Influence in the United Kingdom* (pp. 24–26). Center for Strategic and International Studies (CSIS). http://www.jstor.org/stable/resrep25326.7

Envtl. Prot. Agency Office of Inspector General. Further Efforts Needed to Uphold Scientific Integrity Policy at EPA. https://www.epa.gov/sites/production/files/2020-05/documents/_epaoig_20200520-20-p-0173.pdf [https://perma.cc/H8PZ-CA92]. Published May 20, 2020.

Fallis, D. (2015). What Is Disinformation? *Library Trends, 63*, 401 - 426.

Fergusson, I. F., & Williams, B. R. (2016). *The trans-Pacific partnership (TPP): Key provisions and issues for congress* (CRS Report No. R4489).

Finnemore, M. (1996). *National interests in international society*. Cornell University Press.

Finnemore, M. (1996b). *National interests in international society*. Cornell University Press.

Finnemore, M., & Sikkink, K. (1998). *International Norm Dynamics and Political Change*. International Organization, *52*(4), 887–917. http://www.jstor.org/stable/2601361

Frankfurt, H. G. (2005). *On bullshit*.

Frederick, K. (2019). *The New War of Ideas: Counterterrorism Lessons for the Digital Disinformation Fight*. Center for a New American Security. http://www.jstor.org/stable/resrep20399

Freedom House. (2023). *Freedom in the World 2023*. https://freedomhouse.org/sites/default/files/2023-03/FIW_World_2023_DigtalPDF.pdf

Friedman, T. L. (1999). *The Lexus and the olive Tree*. http://cfile25.uf.tistory.com/attach/152723244AE92F6E1358BA

Fukuyama, F. (1992). *End of History and the Last man*.

Gleicher, N., Franklin, M., Agranovich, D., Nimmo, B., Belogolova, O., & Torrey, M. (2021). *Threat Report: The State of Influence Operations 2017-2020*. Facebook.

Global Media and Information Literacy Week 2023. (2023). *UNESCO*. https://www.unesco.org/en/weeks/media-information-literacy

Graff, G. (2018). Inside the Mueller Indictment: A Russian Novel of Intrigue. *Wired*. https://www.wired.com/story/inside-the-mueller-indictment-a-russian-novel-of- intrigue

Harding, C. (2009). [Review of *The Commonwealth of Nature: Art and Poetic Community in the Age of Dante*, by C. J. Campbell]. Renaissance Quarterly, *62*(3), 869–871. https://doi.org/10.1086/647365

Hassanpour, N. (2014). Media disruption and revolutionary unrest: Evidence from Mubarak's Quasi-Experiment. *Political Communication*, *31*(1), 1–24. https://doi.org/10.1080/10584609.2012.737439

Hitler, Adolf (21 March 1939). *Mein Kampf*. Translated by Murphy, James. Archived from the original on 24 July 2008 – Project Gutenberg.

How we address misinformation on Twitter. (n.d.). https://help.twitter.com/en/resources/addressing-misleading-info

Howard, P. N., & Bradshaw, S. J. (2018). The global organization of social media disinformation campaigns. *Journal of International Affairs*, *71*. https://ora.ox.ac.uk/objects/uuid:67cd8a98-8b3c-45c0-b07f-5e25b25ea67a

Howard, P. N., Woolley, S., & Calo, R. (2018). Algorithms, bots, and political communication in the US 2016 election: The challenge of automated political communication for election law and administration. *Journal of Information Technology & Politics*, *15*(2), 81–93. https://doi.org/10.1080/19331681.2018.1448735 https://medium.com/dfrlab/trolltracker-russias-other-troll-team-4efd2f73f9b5

Huntington, S. P. (1991). *The third wave: Democratization in the Late Twentieth Century*.

Hussain, M., & Howard, P. N. (2013). Democracy's fourth wave?: digital media and the Arab Spring. *Choice Reviews Online*, *51*(04), 51–2322. https://doi.org/10.5860/choice.51-2322

Jackson, R., & Sørensen, G. (2012). *Introduction to International Relations. Theories and Approaches* (fifth ed.). Oxford: Oxford University Press.

Joseph Goebbels On the "Big Lie." (n.d.). Copyright 2023. https://www.jewishvirtuallibrary.org/joseph-goebbels-on-the-quot-big-lie-quo

Keck, M. E., & Sikkink, K. (1998). *Activists beyond borders: Advocacy Networks in International Politics*. Cornell University Press.

Kern, R., & Scott, M. (2022, February 24). *Social media platforms on the defensive as Russian-based disinformation about Ukraine spreads*. POLITICO. https://www.politico.com/news/2022/02/24/social-media-platforms-russia-ukraine-disinformation-00011559

Kershaw, I., & Kershaw, P. O. M. H. I. (1987). *The `Hitler myth': Image and Reality in the Third Reich*. Oxford : Clarendon Press ; New York : Oxford University Press.

Kettemann, M. C., & Schulz, W. (2020). *Setting rules for 2.7 billion: A (first) Look Into Facebook's Norm-making System ; Results of a Pilot Study*

Keyes, R. (2004). *The post-Truth era: Dishonesty and Deception in Contemporary Life*. Macmillan.

La Cour, C. H. (2020). Theorising digital disinformation in international relations. *International politics*, Vol. 57, No. 4, pp. 704-723   - https://hdl.handle.net/1814/70162

Latif, F. (2020). Political Gaslighting in the Climate Change Discourse Surrounding the 2016 Election. In R. X. Browning (Ed.), *President Trump's First Term: The Year in C-SPAN Archives Research, Volume 5* (pp. 39–66). Purdue University Press. https://doi.org/10.2307/j.ctvh9w11h.8

Lavorgna, A. (2021). Information Pollution as Social Harm: Investigating the digital drift of medical misinformation in a time of crisis. In *Emerald Publishing Limited eBooks*. https://doi.org/10.1108/9781800715219

Lee, D. (2018). The tactics of a Russian troll farm. *BBC News*. https://www.bbc.com/news/technology-43093390

MacMillan, A. (2004). *Towards a critical theory of the media in international relations*. Simon Fraser University. Retrieved February 10, 2023, from https://core.ac.uk/download/pdf/56372022.pdf.

Martin, D. A., Shapiro, J. N., & Nedashkovskaya, M. (2019). Recent Trends in Online Foreign Influence Efforts. *Journal of Information Warfare*, *18*(3), 15–48. https://www.jstor.org/stable/26894680

Massoud, A. (2023, March 14). US accuses Russia of gaslighting with its claims of Russophobia. *The National*. https://www.thenationalnews.com/world/us-news/2023/03/14/us-accuses-russia-of-gaslighting-with-its-claims-of-russophobia/

Mayer-Schonberger, V., & Ramge, T. (2018). *Reinventing capitalism in the age of big data*. Hachette UK.

Metzger, M. (2019, November 15). *Effectiveness of Responses to Synthetic and Manipulated Media on Social Media Platforms*. Retrieved from Carnegie Endowment for International Peace: https://carnegieendowment.org/2019/11/15/legal-ethical-and-efficacy-dimensions-of- managing-synthetic-and-manipulated-media-pub-80439#effectiveness

Mikhailova, A. (2018, May 23). Theresa may pledges to tighten the law on "gaslighting" abuse. *The Telegraph*. Retrieved February 10, 2023, from https://www.telegraph.co.uk/politics/2018/05/23/theresa-may-pledges-tighten-law-gaslighting-abuse/.

Mina, A. X. (2019). *Memes to movements: How the World's Most Viral Media Is Changing Social Protest and Power*. Beacon Press.

Misinformation policies - YouTube Help. (n.d.). https://support.google.com/youtube/answer/10834785?hl=en&ref_topic=10833358&sjid=17688554870839399945-EU

Morgenthau, H. J. (1948). *Politics among nations: The Struggle for Power and Peace*. New York : A.A. Knopf.

Mozur, P. (2018, October 15). A genocide incited on Facebook, with posts from Myanmar's military. *The New York Times*. https://www.nytimes.com/2018/10/15/technology/myanmar-facebook-genocide.html

Mueller, R. (2019). *Report on the Investigation Into Russian Interference in the 2016 Presidential Election* (U.S. Department of Justice).

Nimmo, B. (2018). #TrollTracker: Russia's Other Troll Team. *DFRLab*.

Nimmo, B. (2022, October 25). *Measuring Impact of Disinformation Campaigns* [Video]. FutureLearn. Retrieved August 25, 2023, from https://www.futurelearn.com/info/courses/combating-disinformation-and-fake-news/0/steps/157095

Norin, S. (2023). Cooperation With Disinformation Campaigns as a Social Movement Strategy : A case study on the connection between the disinformation campaign and the mobilization against the Swedish Social Services (Dissertation). Retrieved from https://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-494904

Nye, J. S. (1974). Multinational Corporations in World Politics. *Foreign Affairs*, *53*(1), 153–175. https://doi.org/10.2307/20039497

Nye, J. S. (2009). Get Smart: Combining Hard and Soft Power. *Foreign Affairs*, *88*(4), 160–163. http://www.jstor.org/stable/20699631

Norris, P., & Norris, M. L. I. C. P. P. (2001). *Digital divide: Civic Engagement, Information Poverty, and the Internet Worldwide*. Cambridge University Press.

Orwell, G. (1949). *1984* George Orwell Hardback: Hardcover Book.

Osborne, C. (2018). Project Lakhta: Russian national charged with US election meddling. *ZDNET*. https://www.zdnet.com/article/russian-national-charged-with-us-election-meddling/

Pandita, R. (2014). Information pollution, a mounting threat: Internet a major causality. *Journal of Information Science Theory and Practice*, *2*(4), 49–60. https://doi.org/10.1633/jistap.2014.2.4.4

Paul, C., & Matthews, M. (2016). The Russian "Firehose of Falsehood" Propaganda Model: Why It Might Work and Options to Counter It. *RAND Corporation*. https://www.rand.org/pubs/perspectives/PE198.html

Popkin, J. D. (1990). The Press and the French Revolution after Two Hundred Years. *French Historical Studies*, *16*(3), 664. https://doi.org/10.2307/286493

Rapid Response Mechanism Canada: Global Affairs Canada. (2023). GAC. https://www.international.gc.ca/transparency-transparence/rapid-response-mechanism-mecanisme-reponse-rapide/index.aspx?lang=eng#a3_1

Renner, N. (2017) Memes trump articles on Breitbart's Facebook page, *Columbia Journalism Review*. https://www.cjr.org/tow_center/memes-trump-articles-on-breitbarts- facebook-page.php

Rietdijk, N. (2021). *Post-truth Politics and Collective Gaslighting*. Episteme*, 1-17. doi:10.1017/epi.2021.24

Rinehart, A. (2017, February 16). *Fake news. It's complicated.* First Draft. https://firstdraftnews.org/articles/fake-news-complicated/

Risse-Kappen, T. (Ed.). (1995). *Bringing Transnational Relations Back In: Non-State Actors, Domestic Structures and International Institutions* (Cambridge Studies in International Relations). Cambridge: Cambridge University Press. doi:10.1017/CBO9780511598760

Risse, T. (2012). Governance in areas of limited statehood. In D. Levi-Faur (Ed.), *Oxford handbook of governance* (pp. 716–729). Oxford, UK: Oxford University Press.

Risse, T., Risse-Kappen, T., Ropp, S. C., & Sikkink, K. (1999). *The power of human rights: International Norms and Domestic Change*. Cambridge University Press.

Rittberger, V. (2002). *Global governance and the United Nations system*. New York, NY: United Nations University

Rogers, R., & Niederer, S. (Eds.). (2020). *The Politics of Social Media Manipulation*. Amsterdam University Press. https://doi.org/10.2307/j.ctv1b0fvs5

Roll Call. (2019). Mueller report: Russia hacked state databases and voting machine companies. *Roll Call*. https://rollcall.com/2019/04/22/mueller-report-russia-hacked-state-databases-and-voting-machine-companies/

Rothkopf, D. J. (1998). Cyberpolitik: The Changing Nature of Power in the Information Age. *Journal of International Affairs*, *51*(2), 325–359. http://www.jstor.org/stable/24357498
Russia-Ukraine Disinformation Tracking Center - NewsGuard. (2023, August 25). NewsGuard. https://www.newsguardtech.com/special-reports/russian-disinformation-tracking-center/

Schiffrin, A. (2017). Disinformation and Democracy: the Internet Transformed Protest but Did Not Improve Democracy. *Journal of International Affairs*, *71*(1), 117–126. https://www.jstor.org/stable/26494367

Schleffer, G., & Miller, B. (2021). The political effects of social media platforms on different regime types. *Texas National Security Review*. https://tnsr.org/2021/07/the-political-effects-of-social-media-platforms-on-different-regime-types/#_ftn1

Schroeder, R. (2018). *Social Theory after the Internet: Media, Technology, and Globalization*. UCL Press. https://doi.org/10.2307/j.ctt20krxdr

Sebastian, M., & Bruney, G. (2021). Years After Being Debunked, Interest in Pizzagate Is Rising—Again. *Esquire*. https://www.esquire.com/news-politics/news/a51268/what-is-pizzagate/

Sennett, A. (2014). Film Propaganda: Triumph of the Will as a Case Study. *Framework: The Journal of Cinema and Media*, *55*(1), 45. https://doi.org/10.13110/framework.55.1.0045

Sheng, E. L. (2022). *Big tech firms and international relations: The Role of the Nation-State in New Forms of Power*. Springer Nature.

Shihab-Eldin, B. A. (2023, January 30). How Egyptian police hunt LGBT people on dating apps. *BBC News*. https://www.bbc.com/news/world-middle-east-64390817

Shirky, C. (2008). *Here comes everybody: The Power of Organizing Without Organizations*. Penguin.

Slaughter, A. M. (2004). A new world order. Princeton University Press.
Social media manipulation by political actors an industrial scale. (2021, January 13). https://www.ox.ac.uk/news/2021-01-13-social-media-manipulation-political-actors-industrial-scale-problem-oxford-report

Social Movements in the Age of Fake News with Erica Chenoweth. (n.d.). Harvard Kennedy School. https://www.hks.harvard.edu/wiener-conference-calls/erica-chenoweth

Stathis, S. W. (1983). Nixon, Watergate, and American Foreign Policy. *Presidential Studies Quarterly*, *13*(1), 129–147. http://www.jstor.org/stable/27547895

Stern, R. (2007). *The gaslight effect: How to Spot and Survive the Hidden Manipulation Others Use to Control Your Life*. Harmony.

Strange, S., & Strange, S. O. I. R. S. (1996). *The retreat of the state: The Diffusion of Power in the World Economy*. Cambridge University Press

Susskind, J. (2022, July 22). We can regulate social media without censorship. Here's how. *Time*. https://time.com/6199565/regulate-social-media-platform-reduce-risks/

Sweet, P. L. (2019). *The sociology of gaslighting*. American Sociological Review, *84*(5), 851–875. https://doi.org/10.1177/0003122419874843

Tandoc, E. C., Lim, Z. J., & Ling, R. (2017). Defining "Fake news." *Digital Journalism*, *6*(2), 137–153. https://doi.org/10.1080/21670811.2017.1360143

Tapscott, D. (1996). *The digital economy: Promise and Peril in the Age of Networked Intelligence*. McGraw-Hill Companies.

Tarrow, S. (1998). *Power in movement: Social Movements and Contentious Politics*. Cambridge University Press.

Tesich, S. (1992, January 6). *A government of lies*. The Nation. https://www.thefreelibrary.com/A+government+of+lies.-a011665982

Thakur, R., & Van Langenhove, L. (2006). Enhancing global governance through regional integration. *Global Governance*, *12*(3), 233–240.

The Poynter Institute. (2019, August 14). *A guide to anti-misinformation actions around the world - Poynter*. Poynter. https://www.poynter.org/ifcn/anti-misinformation-actions/

Thompson, J. and T. Graham (2022), "Russian Government Accounts Are Using a Twitter Loophole to Spread Disinformation", *The Conversation*, http://theconversation.com/russian-government-accounts-are-using-a-twitter-loophole-to-spread-disinformation-178001

Thompson, J. D. (2022, March 15). *Russian government accounts are using a Twitter loophole to spread disinformation*. The Conversation. https://theconversation.com/russian-government-accounts-are-using-a-twitter-loophole-to-spread-disinformation-178001

UN Secretary-General's High-level Panel on Digital Cooperation. (2019). *The Age of Digital Interdependence*. United Nations.

UN Special Rapporteur on Freedom of Expression. (2018). *Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression*. https://documents-dds-ny.un.org/doc/UNDOC/GEN/G18/096/72/PDF/G1809672.pdf?OpenElement

Vasu, N., Ang, B., Teo, T.-A., Jayakumar, S., Faizal, M., & Ahuja, J. (2018). Unpacking Fake News. In *FAKE NEWS: NATIONAL SECURITY IN THE POST-TRUTH ERA* (pp. 5–9). S. Rajaratnam School of International Studies. http://www.jstor.org/stable/resrep17648.5

Verdegem, S. (2010). *Plutarch's Life of Alcibiades: Story, Text and Moralism*. Leuven University Press.

Vilmer, J.-B. J., Escorcia, A., Guillaume, M. & Herrera, J. (2018), *Information manipulation: A challenge for our democracies*, Technical report, Policy Planning Staff (CAPS) of the Ministry for Europe and Foreign Affairs and the Institute for Strategic Research (IRSEM) of the Ministry for the Armed Forces.

Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380), 1146-1151. DOI:10.1126/science.aap9559

Waldemarsson, C. (2020). Disinformation, Deepfakes & Democracy: The European response to election interference in the digital age. In *Alliance of Democracies*. Alliance of democracies.

Waltz, K. N. (2010). *Theory of international politics*.

Wardle, C., & Derakhshan, H. (2017). *Information disorder: Toward an Interdisciplinary Framework for Research and Policymaking*.

Webb, R., & Kurtz, L. (2022). Politics v. science: How President Trump's war on science impacted public health and environmental regulation. In *Elsevier eBooks* (pp. 65–80). https://doi.org/10.1016/bs.pmbts.2021.11.006

Welch, D. (1983). *Propaganda and the German cinema, 1933-1945*. Oxford University Press, USA.

Wendt, A. (1992). *Anarchy is what States Make of it: The Social Construction of Power Politics*. International Organization, *46*(2), 391–425. http://www.jstor.org/stable/2706858

Wendt, A. (1999). *Social Theory of international politics*. Cambridge University Press.

Wilcox, D. F. (1900). The American Newspaper: A Study in Social Psychology. *The Annals of the American Academy of Political and Social Science*, *16*, 56–92. http://www.jstor.org/stable/1009441

Wild, S. (2023, January 12). *How political gaslighting undermines public policy; a case study of South Africa, Brazil and the United States*. Global Voices. https://globalvoices.org/2023/01/12/how-political-gaslighting-undermines-public-policy-a-case-study-of-south-africa-brazil-and-the-united-states/

Woolley, S., & Howard, P. (2017). Computational propaganda worldwide: Executive summary. *Oxford Internet Institute*. https://ora.ox.ac.uk/objects/uuid:d6157461-aefd-48ff-a9a9-2d93222a9bfd

Worland, J. (2019, July 9). Donald Trump called climate change a hoax. Now he's awkwardly boasting about fighting it. *Time*. https://time.com/5622374/donald-trump-climate-change-hoax-event/

Yerlikaya, T., & Aslan, S. T. (2020). Social Media and Fake News in the Post-Truth Era: The Manipulation of Politics in the Election Process. *Insight Turkey*, *22*(2), 177–196. https://www.jstor.org/stable/26918129

Zeitzoff, T. (2017). How Social Media Is Changing Conflict. *The Journal of Conflict Resolution*, *61*(9), 1970–1991. http://www.jstor.org/stable/26363973