# Internet Governance and the Freedom of Speech: Analysis of Current Frameworks on Disinformation in the US and EU

By

Maximilian Zangl

Submitted to

Central European University

School of Public Policy

In partial fulfilment of the requirements for the degree of Master of Public Policy

Supervisor: Cameran Ashraf

Vienna, Austria, 2021

**School of Public Policy**

**Author's Declaration**

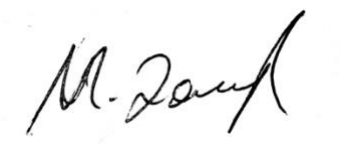
I, the undersigned Maximilian Zangl hereby declare that I am the sole author of this thesis. To the best of my knowledge this thesis contains no material previously published by any other person except where due acknowledgement has been made. This thesis contains no material which has been accepted as part of the requirements of any other academic degree or non-degree program, in English or in any other language.

This is a true copy of the thesis, including final revisions.

Date: 18.06.2021



Name (printed letters): Maximilian Zangl



Signature: ……

## Abstract

The European Union and United States have been subject to various campaigns of disinformation. The attacks originate not only from other states, but also from internet users themselves. Governments are now attempting to tackle disinformation and subsequently reform internet governance. Policy makers are required to respect and adhere to fundamental rights of citizens, when creating regulations. The study analyzes current frameworks and contributes with one of its own. The proposed framework aids to understand the current limitations of internet governance regulating freedom of speech. The results show that, on the one hand, the European Union tends to overregulate speech, while the United States insufficiently addresses the issue. As experienced during the COVID-19 pandemic, disinformation is a threat to society and has the potential to not only impact democratic processes but also public health. Nonetheless, countries have to be careful when implementing regulations and restrictions on the freedom of speech online in order to protect the fundamental rights of citizens in democracy.

# Table of Contents

## Table of Figures

## 1.1 Introduction

In the presidential election of 2016, Russia deployed a campaign to interfere and attack the integrity of the United States democracy. They established a network that included around 4.5 million accounts, disseminating polarizing disinformation (Badawy, Ferrara & Lerman, 2018). Another example is France, where one news outlet spreading disinformation generated over 11 million interactions per month (Fletcher et al., 2018). Also, the United Kingdom fell victim to a false information campaign during the 2016 referendum, that sparked controversy and yet another inquiry into attempting to form policies fighting disinformation (Sabbagh, Harding, & Roth, 2020).

Considering the rapidly growing threat that stems from disinformation, governments have taken over the role of forming and implementing policies. Especially, the European Union has declared itself as the leader against this information war. The Union has been at the forefront of internet governance and setting examples with the General Data Protection Regulation (GDPR). As part of internet governance, disinformation has accelerated to the top of the national security agenda. Nonetheless, governing the internet is not only the obligation of states but includes many more actors from the private sector and civil society. Afterall, the concept of the internet was always set out to regulate itself, but states have stepped in as the authority to govern.

Despite the above mentioned efforts, states have reacted very slowly to the this cyberthreat, but are attempting to tackle the issue. The integrity of democratic processes is directly attacked by the circulation and dissemination of disinformation. The internet is used by billions of people that follow certain protocols and guidelines. These are partly set by organizations and governments but also by internet providers.

Therefore, approaches to regulate disinformation have to be in accordance with the fundamental right of freedom of speech. There is a risk that regulations may restrict this freedom and promote censorship. The US has been slow to react, but is trying to push for transparency in their policies. The approach is to absolutely protect the first amendment online and offline, rather than restricting speech. While this approach doesn't infringe on the freedom of expression, there is no control over disinformation. The European Union has a different philosophy on freedom of speech, which is also reflected in their approach on tackling disinformation online. They shift the liability to the private sector and therefore have implemented a system of governance that is essentially regulated by the social media companies and content providers. The new proposed Digital Service Act (DSA) is accompanied by a bulk of concerns about enabling censorship.

The current literature provides a sufficient overview of the history of internet governance. Also, there are multiple attempts of analyzing the impact of internet policies regarding threats. Generally, disinformation is not seen or treated as a cyber threat. On top of that, there is a clear lack of frameworks that respect the freedom of speech while implementing policies governing the internet. This study tries to explore the current frameworks of the US and the EU and further suggest a possible framework to identify disinformation. First, it defines the important concepts related to the issue; internet governance, disinformation and freedom of speech, then examines the current regulations and proposes a framework on its own. Lastly, the discussion focuses on what policy makers should adhere to when tackling disinformation. The aim is to analyze the shortcomings of the present regulations through the proposed framework, which respects the fundamental right of free speech, and discuss the limitations of internet governance when essentially censoring speech.

## 2 Literature Review

This study tries to explain three major concepts. First, the history of internet governance is examined while establishing a common definition for it. Important works from Mueller and DeNardis are taken into consideration as well as inputs from the European Commission and UN. Further, cyber threats present an important factor in understanding from which standpoint policy makers create regulations for internet governance. Surprisingly, disinformation, which is the focus of the study, frequently misses in cyber threat risk assessments. The phenomena of fake news has been a critical threat to the integrity of democratic process, however policies have failed to effectively address the issue. Concluding, this balancing act between censorship and respecting the fundamental rights of citizens requires an investigation into the definition of freedom of speech.

### 2.1 Internet Governance

The nature of the internet is cyber-physical and therefore disrupts current policy and governing mechanisms. Depending on ideologies or governance structures, the internet differs from country to country. This results in not only government policy making but also includes private actors, i.e. internet and content providers. To understand this, we have to look at the structure of the internet. Generally, the internet has no physical state. What we perceive as the internet is a standardized set of software instructions, also called protocols. These are used to send data over a network to unique addresses. Making use of the protocols, networks are able to communicate with each other and establish a global network. These networks can be owned by private entities, organizations or by the public, which either control and run their own or grant access to external users or nodes (Mueller et al., 2007). These protocols set certain limitations that will be further discussed in the thesis, however, are produced by the private sector, which shifts the distribution of power between governments to private actors and users.

As a result, governments find difficulties to govern and regulate the internet without infringing on the user's freedom.

The term internet governance initially originated from a very narrow set of policy issues that were directed at the global coordination of domain names and addresses (Mueller, 2010, p.9). While previous definitions merely covered a minor part of the internet, there were already attempts of trying to establish governance. The UN installed a working group to define internet governance, find policy issues related to the internet and develop a common understanding of roles and responsibilities of governments and other actors. The definition that was published is:

*"Internet governance is the development and application by Governments, the private sector and civil society, in their respective roles, of shared principles, norms, rules, decision-making procedures, and programs that shape the evolution and use of the Internet"* (WGIG, 2005).

The definition includes the private sector and civil society as well, which are usually not part of the inherent policy making process. However, it still comes from a very classical centralized view on policy making. To keep in mind, in reality internet governance takes place within the interaction of millions of users (Mueller, 2010, p.9). We have to examine a bigger picture and take decentralized policy making into account as well. Without acknowledging the responsibilities of private actors and civil society, applicability and implementation of frameworks are not successful. However, there are still variables and factors to consider that can be governed by governments.

Moving further with the definition of governance we now have to include the role of private actors and civil society within the decision-making process and the influence of them in providing and operating networks. Mueller et al. identify three general areas that are subjective to internet governance. The first one is technical standards. These include network protocols and data formats. The second is resource allocation and assignment which are internet identifiers such as IP addresses and their unique and exclusive use. The last area is human conduct on the internet. While the first two focus on the technical operation of the internet, the later includes much more from cyberattacks to intellectual property infringements, copyrights and the freedom of speech. The three areas display that the core policies have to be in line with private actors and users to ensure freedoms on the internet (2007).

DeNardis provides more insight on how the human component influences protocol making and subsequently the distribution of decision making on the internet (2014, p. 69-71). The approach used to construct the internet was decentralized and leans on the interaction of the users themselves. While this is out of the scope of the study, more insights on internet governance regarding the human component can be found in her book "The Global War for Internet Governance". Ultimately, the definition proposed by the UN does not take the collective approach that is needed to govern the internet into account and misses the inherent nature of it.



While the EU is utilizing the proposed definition by the UN, they have focused on and included various areas of internet governance. The Singe Digital Market encompasses any policy making regarding the internet and rests on the same principles the European Single Market is built on. Free movement of persons, services and capital under common and fair conditions for individuals and businesses (European Commission, 2020). The EU has been

trying to involve all stakeholders and has complemented existing frameworks and the definition set by the UN in the Declaration by the Committee of Ministers on Internet governance principles. Emphasizing human rights, democracy and the rule of law, they also focus on multistakeholder governance (Council of Europe, 2011). This amendment to the declaration includes that all public policy related to the internet should enable full participation of all stakeholders in every country.

To conclude, we arrive at a definition that is still broad but now includes all aspects that are concerned with internet governance:

*"Internet governance is collective decision-making by owners, operators, developers, and users of the networks connected by Internet protocols to establish policies, rules, and dispute resolution procedures about technical standards, resource allocations, and/or the conduct of people engaged in global internetworking activities"* (Mueller et al., 2007).

It offers a decentralized approach including all stakeholders in the policy and decision making process to safeguard the freedoms of all actors.

Before going further, we have to make the distinction between cyberspace and the internet. The literature does not provide a sufficient definition to differentiate the two terms. As discussed earlier, the internet consists of networks that communicate with each other. Independently of who is providing the service, users have access to it. It provides a distinct form of communication, transfer of information and other services. The easiest example for utilizing the internet is writing and sending an email. Cyberspace on the other hand, is a much broader subject. There are networks that have no access to the internet, hence they are not

communicating with other networks. Further, the term offline provides a sufficient example that includes activities in cyberspace, however not on the internet. Devices can be used without the internet (offline) and still stay in the realm of the cyberspace. One example that aids to understand the difference between cyberspace and the internet is Stuxnet. The virus was designed to attack the computer systems that were controlling the speed of centrifuges within the Iranian nuclear powerplant Natanz. The computers were air-gapped from the internet and therefore not reachable except physically. Therefore, the attackers had to plant viruses that spread over flash drives. Attacking five outside companies that were believed to be in close contact with the powerplant, until one flash drive reached its target (Zetter, 2014). The virus was not connected to the internet, but was still active within the realm of cyberspace.

## 2.2 Cyber threat

To define cyber threats, there is a necessity to also examine the definition of cyber security. Security was defined by Wolfers in an objective sense as the absence of threats to acquired values (1952, p. 485). To relate this to cyber space the definition requires additions and clarifications due to the fluidity of the realm. First, the domain does not have clear boarders or frontiers. This results in unsure lines of defense, while attacks and threats can virtually happen anywhere. Secondly, due to the structure of the internet, networks are numerous and connected to various types of other infrastructures. Plus, cyber space provides almost infinite points of entries. Thirdly, the impact of threats or attacks can be locally but spread throughout networks (European Organization for Security, 2010). Cyber security does not necessarily take place on the internet and can take on physical forms as well. This is where the distinction between cyber space and the internet becomes progressively important.

Threats to the cyber space are also categorized as threats to the internet, which are specific to the structure and nature of the internet. However, we have to keep in mind, that even though the threat or attack happened somewhere not connected to the internet, it still has the capabilities to spread to other networks and therefore advance on the internet. The usage of the internet leads to a more complex assessment of security within cyber space. Of course, both terms are overlapping and relate to each other. The challenges with internet security are demanding an require further examination. Software and protocols are designed by humans and accessible to essentially anybody who is on the internet. Humans make mistakes and once an attack has pierced through, networks that are connected can also be infiltrated. This causes internet security to be even more exhausting. On top of that, the previously mentioned additions and clarifications of the security of cyber space are also applicable for the internet as it is part of the cyber space.

Establishing cyber and internet security helps us to understand how to define cyber threats. The term cyber threats will be used for all threats regarding the cyber space and the internet as they can be subject to both.

Threat is defined within the European Commission's proposal for a cybersecurity taxonomy as a potential cause of an unwanted incident, which may result in harm to a system or organization (2019). The definition is also in accordance with the International Organization for Standardization and therefore uniform with other actors. While broad, it complements the above mentioned points about cyber/internet security, but also needs more clarification and an improved scope. However, before assessing advancements to the definition we have to examine the different categories of cyber threats that are currently considered in the literature.

First, we have to make the distinction between attackers, which are perceived as threats, and tools of the process. Both are categories of cyber threats, however the discrepancy provides more insight on how to manage threats and later on construct collective policies that are effective.

Attackers can be categorized in different groups. While "hackers" are perceived as the major source of attacks, the term seems too broad and doesn't capture the motivations behind the attackers. In total, there are five classifications of groups. First, advance persistent threats (ATP) are one of the most dominant drivers in cyber-attacks and warfare. The term exemplifies state guided attacks and is focused on digital spying and espionage. The next group is organized crime, which are mostly concentrated on some sort of economic benefit. Insider threats can cause a lot of damage, however motivations can range dramatically. From employees, who are displeased and want to hurt the company or organization to financial motivations. Another category are hacktivists. Their motivations range from political views to religious or cultural beliefs and other ideologies. The last group of attackers are script kiddies or noobs. They use whatever tools they have available online to attack targets. They are important to mention because of the sheer amount activity they produce (Andress & Winterfield 2014, p. 28-30). Categorizing attackers is important to provide an overview of motivations and actors which are part of cyber space and the internet. Moreover, it presents options and objectives to optimize internet governance once actors are defined and categorized. Of course all of the above mentioned types of attackers can overlap and converge in different motivations.



Attacks require tools and techniques that are directed to harm, observe or infiltrate a specific target. The major steps to an attack are recon, attack and the exploit. Tools and techniques can also be classified in four different areas. Of course, this field is very dynamic

and fluid and continuously produces new technologies and machinery, however the main categories are reconnaissance, attack, exploit and social engineering (SE) (Andress & Winterfield 2014, p. 24).

So far we have established and defined internet governance and have taken threats and security into account. Furthermore, the examination of cyber threats demanded an investigation into cyber security, which also includes the security of the internet. The literature shows, that there have been efforts made to define threats and categorize groups of attackers as well. However, this study is focused on the limitations of internet governance in regards to free speech. Most of the current threats in the literature point to an effort to assert harm through gaining some sort of asset, whether this is information, data, knowledge etc. Yet, there is one component missing which requires further investigation into the definition of threats which is disinformation. Considering that there have been efforts to interfere in elections, i.e. the 2016 US presidential elections or the UK Brexit referendum, through disinformation it grants further research (Sabbagh, Harding, & Roth, 2020). The issue at hand is, how to control deliberate acts of disinformation in a decentralized governing system without infringing on fundamental human rights on freedom of speech.

### 2.2.1 Disinformation (and misinformation) as a Cyber Threat

We have already established the definition of what a cyber threat is, however to incorporate disinformation as one, there are more aspects to examine. Disinformation is not mentioned in most of the literature on cyberthreats, due to the nature of it. As discussed earlier, the sources of threats don't include a group that might fall into the category of spreading disinformation. This also becomes apparent when we examine the steps to an attack; recon,

attack and the exploit (Andress & Winterfield 2014, p. 24). Disinformation is left out because there is no clear perpetrator, steps, procedures or protocols that are followed.

First, the European Commission understands disinformation as, "verifiably false or misleading information that is created, presented and disseminated for economic gain or to intentionally deceive the public, and may cause public harm" (2018b). Important to understand is that the definition of harm in this case builds on the previously discussed one, however broadens in regards to public harm. Public harm are threats to democratic processes as well as to public goods such as citizens' health, environment or security (European Commission, 2018b). This provides a clearer picture of what disinformation is and how to categorize it. Nonetheless, we have to make the distinction between disinformation and misinformation as well. Misinformation is "verifiably false information that is spread without the intention to mislead, and often shared because the user believes it to be true" (European Commission, 2021c). The difference between the two is immensely important. Both require different handling, nonetheless have the potential to cause private and public harm. Although, disinformation is associated with being a threat online, it has not been taken into consideration by the literature. There have been efforts by the European Union to address the issue, however, rethinking id required.

We begin with the IT-Grundschutz catalogue of Germany. Generally, harm has to be done to assets, knowledge, objects or health. They go further and provide a definition for harm in the realm of cyber space and especially on the internet. A potential cause of an unwanted incident that impairs or jeopardizes the availability, integrity, or confidentiality of information which can inflict harm to users and owners of the data or information (Federal Office for

Information Security, 2018). While this definition is still broad, it fulfills the requirements to include disinformation as a cyber threat.

In many instances, it is unclear were the information is coming from, especially with social media. This is where the human component can help to explain more. Humans are both functioning as the actor and the target. Nonetheless, there have been artificial intelligence (AI) programs developed that are capable of creating and disseminating disinformation. However, humans are ultimately involved and directing such programs to intentionally spread "fake news" content. Further, disinformation is extremely hard to differentiate from misinformation. However, the distinction is important to combat both. As mentioned earlier disinformation and misinformation require different responses (Caramancion, 2020). Disinformation comes from a place of intent. There may be different factors involved on why an actor wants to spread malicious information, but its nevertheless dangerous. Especially, the damage done to the public can take on immense forms. Misinformation poses an even greater challenge to internet governance. It is spread without intent and the user truly beliefs in it, which categorizes it as free speech. There are many different angles to consider, i.e. how much of the information is actually false, what is the reach of the misinformation within the social media cycle, what are consequences of this information, what is the degree of harm it poses, etc. Nonetheless, both have the potential to threaten assets, knowledge, objects or health and the democratic process. Therefore, we have to include especially disinformation as a cyber threat.



In combination with the above mentioned points, disinformation has to be incorporated as a cyber threat and therefore be part of the internet governance. However, as this issue is very delicate and sensitive, there are other factors that have to be taken into account. Regulating disinformation is complex and might threaten fundamental human rights, i.e. the freedom of

speech. The distinction between disinformation and misinformation aids to understand the threat and define the scope. The IT-Grundschutz catalogue contributes with an improved definition of cyber threats that includes the disinformation.

## 2.3 Freedom of Speech

Freedom of expression and speech is a fundamental right within the European Union, however, there are grave differences on how freedom of speech is governed compared to the US. In order to compare and show limitations of governance, we have to assess the dissimilarities and scope both. Ultimately, this will have a grave impact on how disinformation will be assessed and governed.

### 2.3.1 Freedom of Speech European Union

Freedom of speech and expression are both covered by the European Union's Charter for Fundamental Rights (Article 11) and the European Convention on Human Rights (Article 10) (FRA, 2000; ECHR, 1953). The European Court of Human Rights (ECHR) has listed two major reasons for the importance of freedom of speech, first, it is an essential foundation of a democratic society and its progress and second, it is critical for the development of every man and women. Further, the freedom of speech is vital for open discussions and opinion forming. The political debate often enjoys a greater degree of protection than other forms. This is important to ensure the fundamental rights of Europeans, develop new ideas and to exercise criticism (Nieuwenhuis, 2000).

Nonetheless, there are certain limitations laid out to freedom of speech, by the Article 17 of the European Convention on Human Rights. It states that:

"*nothing […] may be interpreted as implying for any State, group or person any right to engage in any activity or perform any act aimed at the destruction of any of the rights and freedoms set forth herein or at their limitation to a greater extent than is provided for in the Convention"* (ECHR, 1953).

The reason for this article is the prevention of anti-democratic movements expressing themselves. This is mostly, due to the history that Europe has faced in the past of fascism and communist repression of the East. Therefore, the constitution of European countries, i.e. Germany, has the ability to impose limitations on the freedom of speech. The law has been set out to honor and protect human dignity and the assumptions on which democracy is founded upon. Hence, there are certain elements that protect democracy from ideas that are opposing and contemptuous of human dignity (Nieuwenhuis, 2000).

The above mentioned points provide the basis for the ECHR, which establishes the grounds for limiting free speech in case law. The court has founded the definition of "pressing social need" and a proportional restriction which evaluates speech that threatens or attacks human dignity. This includes hate speech and glorification of violence which may be subject to restrictions. Further, the utterness of racial speech is prohibited due to the infringements or threats to the rights of the minority groups (Nieuwenhuis, 2000).

The underlying approach and philosophy of freedom of speech within the European Union lies in the tolerance and respect of others. Freedom of expression is granted as long as it does not threaten democracy or the fundamental rights of others.

### 2.3.2 Freedom of Speech United States of America

In the US the first amendment of the constitution is the protector of the freedom of expression. The Supreme Court has been immensely clear and strict in protecting the freedom of speech. The philosophy behind the court's decision often relies on "marketplace of ideas". The assumption is that superior ideas will prevail over bad ones. John Stuart Mill originally founded his ideas on the basis that no one alone knows the truth and the competition of ideas is the preeminent solution in separating the falsehoods from fact (Schultz & Hudson, 2017).

Moreover, the Supreme Court has defined two imperative reasons why the first amendment in multiple instances. The freedom of speech is not only an individual liberty but also a good unto itself and it is essential for a common quest for the truth and the vitality of society (Nieuwenhuis, 2000). Compared to the European interpretation of freedom of expression, there are similarities to be found. The prevalent drive for progress and quest for truth within a society is highly regarded in both interpretations and illustrates the importance of opinion-forming and therefore political debate.

The greatest difference between the two approaches is that there are almost no limitations to free speech in the US. In the eyes of the US Supreme Court, the "marketplace of ideas" provides a natural selection process for ideas. Therefore, speech that might be offense to minority groups or even go against democracy are not illegal, however, in theory will also not survive. Arriving from the assumption that in a democracy, the people have the highest power and therefore every political idea should be considered. Following the logic, the US democratic debate could potentially lead to the abolishment of democratic freedoms. However, the "marketplace of ideas" would conclude that comparable thoughts, ideas and debate will not survive (Nieuwenhuis, 2000).

Both approaches are founded in similar ideas and philosophies. The US has a stricter interpretation of limitations on free speech than the European Union. The only limitation that can be found in are instances, where speech incited violence. Even then, the courts have been very careful in prohibiting the freedom of expression. In the case of disinformation internet governance and therefore also regulating free speech, the two approaches have different effects. In the following section, the different frameworks and regulations for combating disinformation will be provided.

# 3 Frameworks on Disinformation

Internet governance has to be a collective effort by all stake holders involved. The decision-making process on specific policies not only calls for a multistakeholder model but also requires private actors and users to be the enforcement mechanism. The cyber and in our case especially internet realm is constantly evolving and requires flexible, sustainable and fast adaptations to procedures, rules and operation. Further, threats on the internet in our society are emerging and developing as fast as the internet is changing. The inclusion of disinformation as a threat is vital to assess threats correctly and treat them accordingly. Nonetheless, individual rights have to be respected while accomplishing that. Especially, the enigmatic relationship between disinformation and freedom of speech, demands governments to be careful in governing the internet.

In order to assess the limitations of the governance of freedom of speech and disinformation on the internet, we have to examine the existing frameworks. The internet is governed on multiple different levels, the top level is for the most part governed by the International Corporation for Assigned Names and Numbers (ICANN). The next level are either regions or the national level itself. In our case, we will examine the European Union and the United States of America. Due to the nature of ICANN and their multistakeholder model of governance, it is interesting to see how such an approach applies in real life. The investigation into the European Union and the United States delivers two approaches to internet governance and specifically to freedom of speech. Both actors have completely different philosophies and approaches on regulating the internet. Privacy laws are taken immensely serious within Europe and frameworks like the GDPR have solidified the Union as one of the pioneers in that area. On the other hand, the US as the inventor of the internet itself, has been very strict in implementing policy that might restrict the users and the companies' ability to

operate. The acumination of the internet, social media and unrestricted speech has led to the emergence of disinformation and misinformation. Both, the US and the European Union have been hit by this phenomenon. While the latter has been attempting to formulate policies targeting disinformation, the former is very content due to the possible infringements of the freedom of speech. We will examine three internet governance frameworks on disinformation, first, ICANN then the European Union and lastly, the US.

## 3.1 Global Level

On the global level, there is only one authority that could have a major crucial impact on content and therefore disinformation. ICANN is a self-proclaimed multistakeholder governance organization that essentially has the power to access and regulate top level domains (TDL).

### 3.1.1 Internet Corporation for Assigned Names and Numbers (ICANN)

ICANN has been an integral part of internet governance. It was established in 1998 and has been independent from governments. Its main purposes are to organize and manage the Domain Name System (DNS), allocate and assign the internet's unique identifiers and provide a platform for a multistakeholder governance system. The aim of the system is to ensure the safety and stability of the internet, while incorporate anybody who wants to participate in internet governance. The structure of ICANN is built around the board of directors who ultimately vote on different policies. However, there are community driven support organizations that contribute all recommendations (ICANN, 2013).

In theory, the policy implementation operates mainly through the Generic Names Supporting Organization (GNSO). The board of directors has to approve the policy and the

Global Domains Division (GDD) forms implementation strategies according to the global guidelines. After completing this process an Implementation Review Team (IRT) is installed. Both the GDD and IRT design the policy and seek public comment. The policy is then officially implemented and continuously overlooked by the GDD (ICANN, 2021).

Even though, the organization is supporting and claiming to have a bottom-up approach, the ICANN has drawbacks in governing the internet. Multistakeholder governance has to derive from a place within the community. ICANN is powerful as it controls and regulates the DNS and therefore can regulate the domain name industry as a whole. This puts them in the position of a monopoly, however influenced through outside political and economic forces. Through the control of the unique identifiers it ultimately has the power to change and influence user behaviors (Mueller, 2010, p.227).

### 3.1.2 Internet Assigned Numbers Authority (IANA)

One of the most powerful organizations regarding internet governance that is incorporated within the ICANN is IANA. The authority's obligations include the coordination of the internet protocols, administration of the DNS root zone management, the allocation of internet numbering resources and other tasks related to the management of top-level domains (Komaitis, 2014).



There have been multiple issues raised with ICANN and the IANA in regards to the multistakeholder governance. First, the US government wanted to create a system that was operated by a non-profit private organization, however did not detach themselves from the organization until 2016. 18 years after establishing the white papers, which set up the governance system the US government officially stepped aside. Before that, the US National

Telecommunications and Information Administration (NTIA) was heavily involved in supervising all policy decisions (Farrell, 2016). This shift is immensely important for internet governance, because the US essentially had dominance over the biggest telecommunications network in the world. First, the US government had the power to influence decisions about the root zone management, second it was able to supervise and regulate the behavior of ICANN and IANA and keep all legal matters within the organization under US law (Mueller, 2014).

The sheer importance of the governance comes to light when we examine the power that lies within the DNS root zone. It allocates domains to unique identifiers. Thus, any domain on the internet could potentially be controlled by these two organizations. These domains, are commonly known under internet addresses, i.e. google.com. The websites contain text, videos, pictures, etc. and therefore are an integral part of providing content. In simple terms, whoever controls and governs the DNS root zone is also able to control content. Therefore, governments, the private sector and the public have a keen interest on who governs it. Therefore, the functions of the ICANN and IANA become inherently policy related (Bradshaw & DeNardis, 2016). To provide an example of the influence a possible infringement of free speech, is to block all websites that have a TLD of .edu. Although, an extreme example but millions of university websites could potentially be shut off. This is important within the context of internet governance as governments also have the ability to block such websites and therefore block content.



### 3.1.3 Domain Abuse Activity Reporting (DAAR)

The DAAR was established by the ICANN to report any abuse of domains. The focus lies on TLDs, however the organization encourages domains that are country specific to take part in the program as well. Essentially, the DAAR is a platform for studying and reporting

domain name registration and security threats to them. Domains are often abused for spreading malware, however also have the ability to spread disinformation. The focus of the DAAR and ICANN is on phishing, malware, botnet command-and-control and spam (ICANN, 2017). ICANN could have the potential to focus on disinformation as well, due its abilities to help organize a multistakeholder model to combat it. The organization has a great influence and access to the different identifiers of the internet, domains and IP addresses, making them an influential part in internet governance.

## 3.2 Regional and National Level

The country or for the EU the Union level, has the power to regulate their own content. Depending on how a country regulates TDLs or other content, users are influenced. For example, in the case of Saudi Arabia, filtering and blocking websites works, amongst other things, through blocking TDLs (Zittrain & Edelman, 2002). Even though, internet providers are private, governments still have the power to regulate and block content. The European Union attempts in internet governance has had upsides, i.e. the GDPR, and now they are embarking on a new project. Regulating and limiting disinformation as much as possible. The different approaches within the US and the EU derive from the distinct philosophies of implementing free speech.



### 3.2.1 European Union Framework on Disinformation

The EU's plan to counter disinformation is laid out in several communications with several institutions. The important factor to take into consideration is, that most of the frameworks or regulations are mere suggestions and don't have to be followed by member states. The Union plans on implementing a new framework that would be a regulation and therefore obligatory for member states to follow, the DSA. The act is designed to change the

current status quo of internet governance, especially the freedom of speech online. Nonetheless, it has not been implemented yet, but will be subject of debate within the discussion session of the thesis.

### 3.2.2 Code of Practice on Disinformation

The Code of Practice on Disinformation was introduced in 2018. This was an unprecedented agreement with the biggest technology companies and the European Union to set standards to fight disinformation and regulate the advertising industry. However, it has to be noted that these practices are on a voluntary basis and companies are self-regulated. Companies like Facebook, Google and Twitter have signed the code and submitted themselves to deliver an annual self-assessment report which includes the efforts taken and shows progress made. The European Union was clearly contented about the progress which the Code of Practice delivered. Moreover, the dedicated Monitoring Programme helped to understand and examine specific actions taken by platforms to counter false information concerning the COVID-19 crisis. Overall, the code contributed a framework for a structured dialogue between policy makers and the platforms and warranted more transparency and accountability of the their policies countering disinformation. The Code of Practice will also be part of the DSA (European Commission, 2021a).

### 3.2.3 The Action Plan on Disinformation

The Action Plan was develop in order to protect democratic systems and processes within the European Union from disinformation in regards to elections. It builds on the Code of Practice, which is focused on disinformation and advertising, and tries to provoke a coordinated response from member states. It builds on four pillars:

(a) improving the capabilities of Union institutions to detect, analyze and expose disinformation;

(b) strengthening coordinated and joint responses to disinformation;

(c) mobilizing private sector to tackle disinformation;

(d) raising awareness and improving societal resilience.

Based on each pillar there are certain tasks and actions required. The Action Plan includes all parts of governments to work together, this ranges from cybersecurity to data protection and electoral to law and media authorities.

(a) improving the capabilities of Union institutions to detect, analyze and expose disinformation;

To improve capabilities the European Commission has allocated funds and is reinforcing the Strategic Communication Task Forces of the European External Action Service, the Union Delegations and the EU Hybrid Fusion Cell by affording additional staff and expertise on detecting and analyzing threats. Focus of this pillar is the strengthening of communication between member states and stakeholders involved and detect, analyze and expose threats.

(b) strengthening coordinated and joint responses to disinformation;

Ultimately, this pillar is aimed at a quick response from the European Union and member states. Therefore, a Rapid Alert System will be set up to provide alerts on disinformation campaigns. The emphasis lies on installing technological tools to flag attacks and stop them before they are able to take off. In order to strengthen cooperation each member state has to set up a designated contact person within the strategic communications department in order to have a coordinated and effective defense. This pillar also encourage further intelligence sharing and communication among member states.

(c) mobilising private sector to tackle disinformation;

The private sector plays a vital role in tackling disinformation. This pillar is built on the currently existing Code of Practice on Disinformation. While online platforms are signing up, the European Commission also encourages advertisers and the advertising industry to adhere to the Code and the Action Plan. Furthermore, the Commission offers assistance to the private sector in adhering to the rules, i.e. the annual report.

(d) raising awareness and improving societal resilience.

The last pillar aims to address civil society. It encourages discussions, forums and other methods of spreading awareness. Member states are called to engage with media, online platforms and communication technology providers to raise awareness and increase transparency. Nonetheless, independent organizations, i.e. fact-checkers are needed in order to fight disinformation or at least mitigate the negative impact. There is also an emphasis on the independent work of media in order to expose and counter disinformation, which is an integral part of the democratic processes.

The Action Plan is also accompanied with the "Tackling Online Disinformation: a European Approach". The Communication outlines five clusters of action for private and public stakeholders.

1. Scrutiny of ad placements
2. Political advertising and issue-based advertising
3. Integrity of services
4. Empowering consumers
5. Empowering the research community

In combination with the Code of Practice on Disinformation the five clusters provide different actions to take part in. Private companies who are signatories of the Code are allowed to have different approaches, however have to include them in their annual report, which has to be shared to all member states. This mechanism is profoundly for the policy process of the European Union. Different kind of actions can fill or expose policy gaps that have to be filled in order to combat disinformation (European Commission, 2021a).

### 3.2.4 European Digital Media Observatory (EDMO)

EDMO is a hub that provides an opportunity for fact-checkers, academics or other stakeholder to connect and work with each other. It serves as an aide to coordinate and fight against disinformation. There are five general activities that EDMO adheres to. First, mapping fact-checker organizations across Europe and tries to interconnect them. There is a substantial amount of funding within the hub, with more than €11 million at disposal. Second, is supporting research on disinformation. This is aimed to prompt scientific research on the topic and provide regular updates on the issue. Third, building a public portal that provides a data base to media practitioners, teachers, academics and citizens to further increase awareness. Fourth, the EDMO wants to establish access for researchers to acquire data on media platforms that is secure and private in order to understand and study disinformation. Last but not least, support the public authorities in monitoring media platforms on existing policies (European Commission, 2021b).

The observatory is completely free of any political institution. It is governed through an advisory board that consists of different academic institutions, which also have the power to decide over the strategy and direction of the EDMO. It was called to life with the Action Plan and targets four specific areas on disinformation:

1. improving detection
2. coordinating responses
3. working with online platforms and industry
4. raising awareness and empowering citizens to respond to disinformation online (European Commission, 2021b).

### 3.2.5 East Strategic Communication Task Force of the European External Action Service

The Task force was installed within the boundaries of the European External Action Service (EEAS). The diplomatic service of the European Union, aids the foreign affairs chief to implement, operate and carry out foreign and security policy. It functions as communication hub for foreign and defense ministries of member states and other parties like the European Commission, the European Parliament and the Council, but also is in close contact with the United Nations and other international organizations (EEAS, 2019a).

In 2015, the European Council pressed to tackle the ongoing issue of disinformation. The EEAS installed the East StratCom Task force, which has gone on to detect and catalogue over 4,500 cases of pro-Russian disinformation. In the course of the Action Plan, the task force was commissioned to establish a system, where disinformation could be identified and countered quicker and more efficiently. The Rapid Alert System (RAS) was developed (EEAS, 2019a).

The RAS features a digital platform with 28 national contact points. Member states are able to upload and feed data into the system in order to gather data, alert disinformation campaigns, coordinate a response and discuss general practices. The EEAS, the Commission

and the task force closely work together to raise public awareness about activities, flag content and empower researchers (EEAS, 2019b).

### 3.2.6 US Framework on Online Content

The United States of America have a different approach to freedom of speech on the internet than the European Union. Of course, there are concerns rising on how to tackle disinformation and misinformation online, however, the importance of free speech anywhere is upheld by the first amendment of the constitution and by the section 230 of the Communications Decency Act (CDA).

The CDA was established to protect minors from indecent content online. While, many parts of the Act were deemed unconstitutional section 230 remained and has developed as one of the most important laws for free speech on the internet. The goal was to allow websites to regulate themselves by removing indecent content on their own (Magee, 2020). The section also allows tech companies to regulate their own content. For example, some platforms chose to ban hate speech, these moderation rules are also in line with the section 230 (Newton, 2020).

One of the most important landmark cases that underlined the section 230 as the most important law to protect free speech online is the supreme court decision on Reno v. ACLU. The American Civil Liberties Union was contesting the CDA, which included other forms of regulations on content and freedom of speech online. The ruling stated:

*"We agree with the District Court's conclusion that the CDA places an unacceptably heavy burden on protected speech, and that the defenses do not constitute the sort of "narrow*

*tailoring" that will save an otherwise patently invalid unconstitutional provision (Reno v. ACLU, 1997).”*

Although, the most of the proposed laws and regulations within the CDA failed, the section 230 remains and the importance of the Supreme Court Ruling empowers the current law. On the other hand, this protected and boosted the freedom speech online and solidified the internet as a the “marketplace of ideas” (Abrams v. United States, 1919). The “marketplace of ideas” has been invoked by the Supreme Court multiple times to counter censorship and therefore has become a powerful analogy that also supports freedom of speech on the internet (Schultz & Hudson, 2017).

In 2021, policy makers in the US have been making an effort to prevent disinformation online, especially reforming section 230 and reforming political ads online. Two bills are currently under way and have been passed to the Senate. The SAFE TECH act is aimed at reforming section 230. Although, the proposed law would only target paid content, concerns over free speech infringements prevail (Hatmaker, 2021). The second bill, the Honest Ads act is directly linked to political advertisement. The bill will close a loop hole and make political online advertisement illegal for foreign nationals. Further, it would improve transparency by expanding disclosure rules and social media platforms would be compulsory to provide a list of political advertisement information, i.e. target audience and payment information (Lau, 2020).



The overall theme of the reforms, targets the transparency of content rather than limiting it. Of course, community guidelines set by websites or social media companies still

have to be respected by users, however, content as long as its transparent where it comes from (if it is paid for or political) still faces no regulation.

## 3.3 Framework on Limitations of Governance

This study will use this framework to examine the above mentioned laws, regulations and cases to pinpoint the limitations in governing not only disinformation but also freedom of speech online. The approach created originates from an US based interpretation of freedom of speech. As discussed earlier, the US has a vastly different approach to regulating freedom of expression online through the "marketplace of ideas". For the European Union, one has to be more careful. The limitations of the policy makers are more complex due to the existence of hate speech. Nonetheless, we will examine test the model against the processes of the European Union especially the proportionality test of the European Court of Human Rights.

| | | |
|---|---|---|
| 1. Victim | Society/Public | Individual |
| 2. Motivation/Intent | Intent | No-Intent |
| 3. Conduct through content | Not-Protected | Protected |
| 4. Assessment of Harm | Harm | No Harm |

*Figure 1: Proposed Framework to Assess Disinformation*

Disinformation has distinctive categories that have to be met. Liability is only applicable if the victim is the society/public, there was a clear intent to deceive, the content is not protected under freedom of speech and there has been harm done. If any of these categories is not fulfilled, it is not disinformation.

### 3.3.1 Victims

The victim of disinformation has to be the society or the public. This is essential to the nature of the intend of disinformation. If the intent is to only deceive one or a limited number of persons with wrong information, this would simply be lying. That is why it is important, that the general public is the target group of disinformation.

### 3.3.2 Motivation or Intent

Intent is arguably the most important component in assessing, whether or not disinformation was conducted. Furthermore, this is also the toughest part to proof. The intentions and motivations have to be to deceive the public in order to have some sort of gain from it. This could be economic gain or to intentionally deceive the public, and may cause public harm.

### 3.3.3 Conduct through Content

The conduct has to take place through content online, this includes, posts, pictures, videos, websites etc. Important for the content is freedom of expression. Of course, in combination with the intent, freedom of speech becomes more limited and has to fall under non-protected speech. Therefore, information that is accurate, cannot fall under the category of disinformation. Even though there is an intent to harm the public, it requires the distinction between the interest of the public and public interest. The latter is in the interest of the wellbeing of the general public rather than the former, which is only what might be appealing to know for the public.

### 3.3.4 Assessment of Harm

Harm to the general public needs to be done in order to qualify under disinformation. Other scenarios where there is harm to individuals, limited groups or even no harm does not grant the classification of disinformation. To assess the harm of certain content there is a test required.

In the case of Schenck v. United States (1919) the Supreme Court established the Clear and Present Danger Test. It states that any printed or spoken word may not be subject to restraint or subsequent punishment unless the expression establishes a clear and present danger of bringing about a substantial evil. There are two requirements that the test has to fulfill. First, the expression has to pose a threat that a substantive evil might follow and second, the threat has to be real and imminent. (Cornell Law School, 2020). Even though, Schenk lost the case, the courts had a structure to rule freedom of expression by. Justice Holmes was pressing the issue that the circumstances and time are vital factors in determining, whether speech is protected or not (Cornell Law School, n.d.-b).

The case Brandenburg v. Ohio, the Supreme Court adjusted the Clear and Present Danger Test. The test determines whether speech may be prohibited if it is advocating force or crime. Therefore, two parts have to be satisfied, first, the speech is "directed to inciting or producing imminent lawless action," and the speech is "likely to incite or produce such action" (Cornell Law School (c), n.d.-a).

The two tests offer an improved sense of how to measure harm and when to prohibit free speech. Although, the use of both is targeted at speech that incites violence, force or crime, there are key takeaways to relate to disinformation. In order to assess harm and gravity of the

disseminated disinformation, context, time and circumstances have to be taken into consideration. On top of that, it has to be directed to incite imminent harm to the public and likely to incite or produce actions damaging the integrity of democratic processes and/or the public well-being. To conclude, each case will require an individual assessment of harm that takes all the above mentioned points into consideration, in order to arrive at a decision.

# 4 Discussion

The aim of this discussion is to examine the current frameworks and their potential effectiveness on current users. The central question is, whether frameworks are infringing on the freedom of speech and what are the limitations of internet governance in the regard of effectively tackling disinformation. The proposed framework in section 3.3. will be tested against the current framework, and also compared and amended to the EU's regulations and the DSA.

## 4.1 Framework and the US

The US has fallen victim to multiple disinformation campaigns over the course of history. However, there seems to be an spike in the false information spread. The rise of social media and the shift of roles within the dissemination of information to the individual, anybody can actively promote, share and create information (Chang, Mukherjee, & Coppel, 2020). The US has made itself vulnerable to conspiracy theories, fake news and other forms of disinformation by ignoring to set ground rules for information exchange online. Of course, it is not the only country to do so, however, in as the world hegemon and the current leaders of the free world, it has struggled to contain false information online. The current frameworks for the US are not aiming at regulating disinformation. Section 230 is protecting the content moderation of social media platforms and their content. The proposed bills, i.e the HONEST ads act, is only aimed to close the gap on foreign influence on political advertisements and to increase transparency. The first amendment is rarely touched within the legal frameworks, unless there is a clear incitement of violence. The proposed framework could address issues and combat disinformation.

To showcase the operability of the proposed framework the example of the heavily discussed school shooting, the Sandy Hook Massacre in 2012, was chosen. On December, 14, 2012, a gunman stormed into an elementary school and brutally killed 20 first-graders, before taking his own life. This sparked a nationwide conversation about gun control where reform ultimately failed. After the tragic incident, stories on social media started to circulate. The narrative of the stories were that the shooting was staged, fake and that the grieving parents were actors. The front runner in disseminating disinformation was Alex Jones (Berman, 2017). The face of Infowars, a self-proclaimed news source, opened a bombarding of fake news stories. Some of his social media accounts and content has been banned from platforms themselves, however he remains immensely popular. Up until the ban of Infowars, the website had a daily average of 1.4 million visitors and views on their content. This number dropped to 750,000 after the ban (Nicas, 2018). However, the reach of Alex Jones and his conspiracy theories is still enormously influential. According to a Fairleigh Dickinson poll, around 22% of Americans thought that the 2012 shooting was faked, in order to increase gun control (2016).

To test this case against the proposed framework in the methodology section we have to go through the four categories. First, the victims of the disseminated false information are not only the families, however, also the public as a whole. The immense size of traffic that Infowars produces alone on average each day, provides a substantial reach of the public. To keep in mind, there are also other victims, i.e. the family members of the diseased. If they would be the only harmed group, it would not be a disinformation case. The victims' families could sue for defamation (which they did) (Collins, 2021). However, because of the extensive outreach and the following countrywide discussion and debate about the incident, leaves no doubt that the general public was the victim of the spread disinformation. Further, the motivation or intent of the media outlet was politically motivated. Alex Jones, who has been

known as a close friend of Trump and a far right conspiracy theorist. He has been advocating guns fiercely on his show (Anti-Defamation League, 2021). The intent to deceive the public was to downplay the shooting and persuade them to stop gun reforms by forceful spreading disinformation about the tragedy. On top of that, it can be argued that the more outrageous the information is, the more views it will generate. Therefore providing Infowars with economic gain as well. The third category is conduct through content. This is where protected and non-protected speech comes into play. As of right now, we have established that his speech has the target audience of the general public and the intent to deceive them for his own political and economic gain. In combination with the first two classifications we can establish that the conduct of spreading disinformation via his own website and social media channels, was not protected. Furthermore, taken all the above mentioned points into consideration, we have to examine circumstances and timing as well. 2012 was the reelection year were Barack Obama won his second term in office. A rather decisive win by the Democrats was met by an outraged political right. During this time, other disinformation campaigns started to gain ground as well, i.e. President Obama was a Muslim or that he wasn't born in the United States. Undoubtedly, Infowars tried to delegitimize the current leadership and tried to damage the integrity of democratic processes. Especially on a topic like gun control, where the voters partisan gap is at 59 percentage points. Higher than any other issue on the political agenda (Pew Research Center, 2019).



While this is a clear case of spreading disinformation under the proposed framework, there are further investigations required. In the case of the US, free speech is arguably the most important right. The limitations with the proposed framework is judging whether the disinformation can fall under the category of non-protective speech if intent and harm is proven.

## 4.2 Framework of the European Union

For the European Union the philosophy in regulating disinformation is shaped through the approach taken for regulating free speech. Although, there have been efforts made to protect fundamental freedoms, there are still questions to be answered. The framework on disinformation presented in this study are not regulations but communications to follow on a voluntary basis. The four pillars of the Action Plan on Disinformation are very broad, however provide a starting point. Generally, the call for the involvement of the private sector appears to be a good idea as well. Nonetheless, the role of the civil society is overlooked. In the further discussion, we will examine proposed regulations within the DSA that underline this relationship. Furthermore, the frameworks, including the RAS are there to regulate processes not content per se. The new push to the DSA will change that.

The aim of the DSA is to change the rules for targeting, monitoring and moderating harmful illegal content online the liability of online providers for third party content, quality assurance obligations of third party suppliers and the protection of users' fundamental rights online (Gerritzen et al., 2020). While the EU is hoping to change the governance of the internet and set an example, there are major concerns regarding free speech. The greatest concern regarding the regulation of free speech is the push to let platforms enforce censorship. There is a clear incentive for social media companies to overregulate content and therefore limiting free speech. First, the intermediary liability is upheld, which makes the platform or content provider liable for not engaging in active content moderation. Therefore, companies are seen as playing an active role in the dissemination of illegally deemed content. Even though, the Article 6 of the DSA is seen as the a comparison to section 230 as a “good faith” clause and protects service providers from liability, there is still unclarity what is meant by “active role” (Schmon, 2021).

Article 14 of DSA shifts the responsibility of creating "dissemination to the public" to the social platform provider. As large social platform providers are hosting millions or even billions (Facebook) of users, they are forced to implement smart algorithms to detect dissemination. The term used is "endpoint content blocking", which consists of several techniques like content scanning, URL categorization, IP address blocking, and DNS interception. To detect dissemination, only filters within a social media platform (content scanning) can be used.

These filters have to be installed by the platform provider. Very little is known about the effectiveness, scope, or side effects of platform blocking technologies. While the major social platforms universally block certain types of content (such as malware and pornographic material) and provide customized content feeds to their users, information on policy induced blockages is not public. Content filters are required to include image recognition techniques, as users can post written content in picture format (Schmon, 2021).

As DSA regulations are fuzzy and do not specifically describe illegal content, there is no standard definition of rules across these platforms to filter dissemination. The only time illegal content is defined is in the Commission's recommendation on measures to effectively tackle illegal content. However, the explanation is very vague, "information relating to terrorism, child sexual abuse, illegal hate speech or infringements of consumer protection laws, which can undermine the trust of their users and damage their business models" (2018a). Depending on company philosophy, the rules are set tighter or looser, but in general to avoid any liabilities, social media platforms tend to over block content. In addition, the platform provider has to justify the reason why the content is blocked to be compliant with Article 3 of DSA (Schmon, 2021). This implies, that the provider needs to understand the intention of the

user, who has posted the content, as not every user is acting with bad intentions to safeguard freedom of speech.

On the other hand, the platform provider can be held liable, if he doesn't set appropriate actions to block users who are posting disseminating content. Facebook users upload approximately 250 million images per day (Smith, 2019), which all have to be scanned for content. As compliance to DSA represents a sheer impossible task, providers tend to manage content restrictions through their own terms of service. Also, this move has been restricted by the DSA in Article 12 ("Providers of intermediary services shall act in a diligent, objective and proportionate manner in applying and enforcing the restrictions") (Schmon, 2021). Through this article, provider's hands are tied, leaving them alone to find a delicate balance between preserving freedom of speech and held liable for distributing illegal content.

## 4.3 Final Discussion of the Limitations of Internet Governance and Freedom of Speech

Within the Unites States internet governance regarding free speech has been lenient. The protection of the first amendment is a priority to policy makers and courts. However, disinformation has taken its toll on the country and political system as seen in the 2016 presidential election. "Information warfare" and social media campaigns targeting both sides including staging events within the United States boarders effectively favored Trump and disparaged Clinton (Mueller, R., 2019). Showing that the democratic process have been harmed and interfered with. The current frameworks set in place are not sufficient to tackle misinformation online and need to revised under certain conditions.

On the other hand, the European Union, with a different philosophy, has attempted to tackle disinformation online. The current frameworks set in place, are aimed at raising transparency and involving the private sector. However, there has not been a substantial regulation implemented for all member states. Social media platforms and content providers are voluntarily participating in programs that help to understand disinformation, but fail to regulate without infringing on fundamental freedoms. The proposed DSA shifts the liability and censorship power to the private sector, which raises significant concerns for the freedom of speech of users. Essentially, platforms are incentivized to overregulate content in order to avoid fines by the EU. The vagueness of the DSA is concerning and has to be revised.

Taking the previous discussion into account, there are multiple limitations of internet governance regarding free speech. These are the shortcomings of the current frameworks in place that require further revision.

Governments are required to be very specific on their policies. The policy making process is still heavily influenced by authorities. In order to have good policies regarding the governance of the internet to protect the integrity of democracies while respecting fundamental rights, laws and frameworks cannot be vague. Especially, the DSA allows an considerable amount of interpretation to private companies. In combination with possible fines and penalties this poses an outcome of overregulating content. The balance between protecting fundamental rights and enforcing censorship is immensely delicate and has to be respected and taken into consideration.

Policies must not exist to restrict ideas and beliefs. This is immensely important when the distinction between misinformation and disinformation is not clear. Even though the

differentiation is made, the line between legal and illegal conduct has to be clear. Here, the intention emerges as the deciding factor. If there is no intent to deceive for political or economic gain, rather to disseminate one's own belief, it cannot be censored. Nonetheless, the speech has to either be protected under the first amendment or under Article 10 of the European Convention on Human Rights. Decisions for restrictions have to consider many different factors in order to limit ideas and beliefs. Respecting the two different philosophies of the US and the EU, decisions have to be conducted under immense due diligence and care. Governments and online platforms have to actively work against censorship, rather than constraining ideas and beliefs.

The proposed framework could potentially provide a solution to the problem and aid in assessing whether or not governments are able to restrict, prohibit or censor speech. All four categories have to be met by the information in order to classify as disinformation. While the framework looks simple it presents a fairly complex and thorough analysis. Further, it would somewhat standardize disinformation and could set an example to future cases. The application of the framework within the US, would tackle the issue of non-protective speech on the internet. Only when the victim is the public and the intent and harm is proven, there can be a limitation of speech. This provides a high threshold, that has to be proven, however has a clear outcome. The current laws in place, are not sufficient to tackle disinformation. Multiple examples, most importantly the 2016 presidential election, have shown that the US has fallen victim to disinformation campaigns, stemming from outside and inside its borders.



The European Union tends to overregulate content and has established a dangerous internet governance system, that puts the regulation of freedom of speech online in the hands of the biggest social media companies in the world. Especially the above mentioned limitations

on vagueness and restrictions on ideas and beliefs apply to the DSA. Although the proposed framework would be more difficult to implement, due to the rigorous and thorough process of the ECHR, there are upsides that have to be taken into account. With the framework, censorship and content moderation would be harder for the European Union and therefore providing adequate protection to of fundamental rights. It would also shift the responsibility of censorship to the ECHR, which would guarantee a correct assessment and decision.

# 5 Conclusion

Internet governance will incrementally become more and more important. There are many facets and aspect that have to be included with the field, i.e. the inclusion of civil society within the policy making process. Nonetheless, this study examined cyber threats, focusing on disinformation, and the existing frameworks in place from the EU and the US. Although, there are multiple levels to internet governance, the country level, or in the case of the EU, the regional level, has the most impact on the users.

Fundamental rights have to be respected when creating and implementing new policies. Especially, the fight against disinformation has an impact on the integrity of democratic processes and requires smart and well thought through policies. In an attempt to provide a framework for governing disinformation online, we established a threshold for courts to decide whether information has been disseminated illegally. Although, the framework applies to the US, the ECHR puts certain limitation on the freedom of speech, which potentially could influence it.

As experienced during the COVID-19 pandemic, disinformation is a major threat to society and has the potential to not only impact democratic processes but also public health. Nonetheless, countries have to be careful when implementing regulations and restrictions on the freedom of speech online in order to protect the fundamental rights of citizens in democracy.

# References

*Abrams v. United States* (250 US 616 (1919)). (1919, November). https://www.oyez.org/cases/1900-1940/250us616

Andress, J. & Winterfeld, S. (2014). Cyber Threatscape. *Cyber Warfare*, 19–34. https://doi.org/10.1016/b978-0-12-416672-1.00002-7

Anti-Defamation League. (2021). Alex Jones: Five Things to Know. *Anti-Defamation League*. https://www.adl.org/resources/backgrounders/alex-jones-five-things-to-know

Badawy, A., Ferrara, E. & Lerman, K. (2018). Analyzing the Digital Traces of Political Manipulation: The 2016 Russian Interference Twitter Campaign. *2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*. Published. https://doi.org/10.1109/asonam.2018.8508646

Berman, N. (2017). The victims of fake news. *Columbia Journalism Review*. https://www.cjr.org/special_report/fake-news-pizzagate-seth-rich-newtown-sandy-hook.php

Bradshaw, S. & DeNardis, L. (2016). The politicization of the Internet's Domain Name System: Implications for Internet security, universality, and freedom. *New Media & Society*, *20*(1), 332–350. https://doi.org/10.1177/1461444816662932

Caramancion, K. M. (2020). An Exploration of Disinformation as a Cybersecurity Threat. *2020 3rd International Conference on Information and Computer Technologies (ICICT)*. Published. https://doi.org/10.1109/icict50521.2020.00076

Chang, L. Y. C., Mukherjee, S. & Coppel, N. (2020). We Are All Victims: Questionable Content and Collective Victimisation in the Digital Age. *Asian Journal of Criminology*, *16*(1), 37–50. https://doi.org/10.1007/s11417-020-09331-2

Collins, D. (2021, 5. April). High court nixes Alex Jones' appeal in Newtown shooting case. *AP NEWS*. https://apnews.com/article/connecticut-shootings-lawsuits-alex-jones-school-shootings-59360449ed878c5cdfcbb550291c90a2

Cornell Law School. (n. D.-a). *Brandenburg test*. LII / Legal Information Institute. Abgerufen am 18. Juni 2021, von https://www.law.cornell.edu/wex/brandenburg_test

Cornell Law School. (n. D.-b). *Clear and Present Danger*. LII / Legal Information Institute. Abgerufen am 18. Juni 2021, von https://www.law.cornell.edu/constitution-conan/amendment-1/clear-and-present-danger

Cornell Law School. (2020, May). *Clear and Present Danger*. LII / Legal Information Institute. https://www.law.cornell.edu/wex/clear_and_present_danger

Council of Europe. (2011, 21. September). *Declaration by the Committee of Ministers on Internet Governance Principles*. www.coe.int. https://search.coe.int/cm/Pages/result_details.aspx?ObjectID=09000016805cc2f6

DeNardis, L. (2014). *The Global War for Internet Governance*. Oxford University Press, USA.

ECHR. (1953). *European Convention on Human Rights* (Rome, 4.XI.1950). https://www.echr.coe.int/documents/convention_eng.pdf

EEAS. (2019a, March 11). *Countering disinformation*. EEAS - European External Action Service - European Commission. https://eeas.europa.eu/topics/countering-disinformation/59411/countering-disinformation_en

EEAS. (2019b, March 12). *Rapid Alert System*. EEAS - European External Action Service - European Commission. https://eeas.europa.eu/sites/default/files/ras_factsheet_march_2019_0.pdf

European Commission. (2018a). *COMMISSION RECOMMENDATION of 1.3.2018 on measures to effectively tackle illegal content online* (C(2018) 1177 final).

https://digital-strategy.ec.europa.eu/en/library/commission-recommendation-measures-effectively-tackle-illegal-content-online

European Commission. (2018b, December). *JOINT COMMUNICATION TO THE EUROPEAN PARLIAMENT, THE EUROPEAN COUNCIL, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS* (JOIN(2018) 36 final). https://eeas.europa.eu/sites/default/files/action_plan_against_disinformation.pdf

European Commission. (2019, 5. December). *European Cybersecurity Taxonomy*. EU Science Hub - European Commission. https://ec.europa.eu/jrc/en/science-update/european-cybersecurity-taxonomy

European Commission. (2020). *Shaping the Digital Single Market*. https://ec.europa.eu/. https://ec.europa.eu/digital-single-market/en/shaping-digital-single-market

European Commission. (2021a). *Code of Practice on Disinformation*. Shaping Europe's Digital Future. https://digital-strategy.ec.europa.eu/en/policies/code-practice-disinformation

European Commission. (2021b). *European Digital Media Observatory*. Shaping Europe's Digital Future. https://digital-strategy.ec.europa.eu/en/policies/european-digital-media-observatory

European Commission. (2021c). *Online disinformation*. Ec.Europa.Eu. https://digital-strategy.ec.europa.eu/en/policies/online-disinformation

European Organization for Security. (2010, July). *Towards a concerted EU approach to cyber security* (Version 5.0-draft July 23 rd 2010). http://www.eos-eu.com/Files/Cuber-policy-docs/EOS_ICTWG_CyberSec_v5%200.pdf

Fairleigh Dickinson University. (2016, 4. May). *Fairleigh Dickinson University's PublicMind Poll Finds Trump Supporters More Conspiracy-Minded Than Other Republicans*. https://view2.fdu.edu/publicmind/2016/160504/

Farrell, M. (2016, 14. March). Quietly, symbolically, US control of the internet was just ended. *The Guardian*. https://www.theguardian.com/technology/2016/mar/14/icann-internet-control-domain-names-iana

Federal Office for Information Security. (2018). *IT-Grundschutz-Kompendium*. Beltz Verlag.

Fletcher, R., Cornia, A., Graves, L. & Nielsen, R. K. (2018, 1. Februar). *Measuring the reach of "fake news" and online disinformation in Europe*. Reuters Institute Digital News Report. https://www.digitalnewsreport.org/publications/2018/measuring-reach-fake-news-online-disinformation-europe/

FRA. (2000). *Article 11 - Freedom of expression and information*. European Union Agency for Fundamental Rights. https://fra.europa.eu/en/eu-charter/article/11-freedom-expression-and-information

Gerritzen, F., Van Dyck, P., de Vries, Y., Wolters Ruckert, N., Taelman, E., Ruers, C. & van der Leeuw-Veiksha, A. (2020, 16. December). *The Digital Services Act package is here*. Allen & Overy. https://www.allenovery.com/en-gb/global/news-and-insights/publications/the-digital-services-act-package-is-here

Hatmaker, T. (2021, 5. February). *TechCrunch is now a part of Verizon Media*. Tech Crunch. https://techcrunch.com/2021/02/05/safe-tech-act-section-230-warner/

ICANN. (2013). *Beginner's Guide to Participating in ICANN*. https://www.icann.org/en/system/files/files/participating-08nov13-en.pdf

ICANN. (2017). *Frequently Asked Questions: ICANN's Domain Abuse Activity Reporting (DAAR) Project*. https://www.icann.org/. https://www.icann.org/octo-ssr/daar-faqs

ICANN. (2021, March). *Implementing Policy at ICANN*. https://www.icann.org. https://www.icann.org/policy/implementation

Komaitis, B. K. (2014, 10. October). *Understanding the IANA Functions*. Internet Society. https://www.internetsociety.org/blog/2014/10/understanding-the-iana-functions/

Lau, T. (2020, 17. January). *The Honest Ads Act Explained*. Brennan Center for Justice. https://www.brennancenter.org/our-work/research-reports/honest-ads-act-explained

Magee, A. F. (2020, 29. October). *Back Against the Wall: Are Section 230's Days Numbered?* Wake Forest Law Review. http://www.wakeforestlawreview.com/2020/10/back-against-the-wall-are-section-230s-days-numbered/

Mueller, M. (2010). *Networks and States*. Amsterdam University Press.

Mueller, M. (2014). Detaching Internet Governance from the State: Globalizing the IANA. *Georgetown Journal of International Affairs*, 35–44. http://www.jstor.org/stable/43773647

Mueller, M., Mathiason, J. & Klein, H. (2007). The Internet and Global Governance: Principles and Norms for a New Regime. *Global Governance: A Review of Multilateralism and International Organizations*, *13*(2), 237–254. https://doi.org/10.1163/19426720-01302007

Newton, C. (2020, 29. December). Section 230: everything you need to know about the law protecting internet speech. *The Verge*. https://www.theverge.com/21273768/section-230-explained-internet-speech-law-definition-guide-free-moderation

Nicas, J. (2018, 5. September). Alex Jones Said Bans Would Strengthen Him. He Was Wrong. *The New York Times*. https://www.nytimes.com/2018/09/04/technology/alex-jones-infowars-bans-traffic.html

Nieuwenhuis, A. (2000). Freedom of Speech: USA vs Germany and Europe. *Netherlands Quarterly of Human Rights*, *18*(2), 195–214. https://doi.org/10.1177/092405190001800203

Pew Research Center. (2019, 17. December). In a Politically Polarized Era, Sharp Divides in Both Partisan Coalitions. *Pew Research Center - U.S. Politics & Policy*. https://www.pewresearch.org/politics/2019/12/17/in-a-politically-polarized-era-sharp-divides-in-both-partisan-coalitions/

*Reno v. ACLU* (521 U.S. 844). (1997, June). https://supreme.justia.com/cases/federal/us/521/844/

Sabbagh, D., Harding, L. & Roth, A. (2020, 22. July). Russia report reveals UK government failed to investigate Kremlin interference. *The Guardian*. https://www.theguardian.com/world/2020/jul/21/russia-report-reveals-uk-government-failed-to-address-kremlin-interference-scottish-referendum-brexit

*Schenck v. United States* (249 US 47 (1919)). (1919, March). https://www.oyez.org/cases/1900-1940/249us47

Schmon, C. (2021, 18. March). Twitter, Trump, and Tough Decisions: EU Freedom of Expression and the. *Electronic Frontier Foundation*. https://www.eff.org/de/deeplinks/2021/03/twitter-trump-and-tough-decisions-eu-freedom-expression-and-digital-services-act

Schultz, D. & Hudson, D. L. (2017, June). *Marketplace of Ideas*. The First Amendment Encyclopedia. https://www.mtsu.edu/first-amendment/article/999/marketplace-of-ideas

Smith, K. (2019, 1. June). 53 Incredible Facebook Statistics and Facts. *Brandwatch*. https://www.brandwatch.com/blog/facebook-statistics/

Special Counsel Robert S. Mueller, III. (2019, March). *Report On The Investigation Into Russian Interference In The 2016 Presidential Election* (Submitted Pursuant to 28 C.F.R. § 600.8(c)). https://www.justice.gov/archives/sco/file/1373816/download

Wolfers, A. (1952). „National Security" as an Ambiguous Symbol. *Political Science Quarterly*, *67*(4), 481. https://doi.org/10.2307/2145138

Working Group on Internet Governance. (2005). *Report of the Working Group on Internet Governance* (05.41622). https://www.wgig.org/docs/WGIGREPORT.pdf

Zetter, K. (2014, 3. November). An Unprecedented Look at Stuxnet, the World's First Digital Weapon. *Wired*. https://www.wired.com/2014/11/countdown-to-zero-day-stuxnet/

Zittrain, J. & Edelman, B. (2002). Documentation of Internet Filtering in Saudi Arabia. *cyber.harvard.edu*. Published. https://cyber.harvard.edu/filtering/saudiarabia/