4. Legal foundation – do legal remedies work?

Authors: Konstantinos Katevas, Timo Steidle and Max Winter
Academic supervisor: Sebastian Brüggemann

DOI: 10.24989/ocg.v.342.4

4.1. Introduction

With hate speech and fake news on the rise across the internet, politicians are faced with the responsibility to act decisively against their spread.

Several questions need to be answered to approach this issue: How are hate speech and fake news defined in a legal context? What legal remedies or potential policies can be put in place to stifle them? What hurdles and challenges will the government face when enacting these policies? What is the potential impact on universal human rights like the freedom of speech and information?

In the following section, the legal definitions and approaches of individual countries in regards to both hate speech and fake news will be analyzed and compared.

Further, several methods will be evaluated for their expected efficacy in the pursuit of dealing with hate speech and fake news online, while predicting potential short- and long-term effects.

4.2. Fake news

The difference between fake news and hate speech is that hate speech generally harms individuals or members of a specific group, whereas fake news is arguably damaging to society as a whole. This raises problems and questions for governments: whether they should try to regulate fake news with legal restrictions or not. Often those laws directly contradict other basic rights like freedom of speech.

The legal situation regarding the fake news is comparable to the section concerned with hate speech. It can be rather difficult to find common ground, especially on an international level and between different cultures.

The idea of combating fake news is not a subject exclusive to the 21st century. In 1936 the member states of the League of Nations agreed on the “International Convention Concerning the Use of Broadcasting in the Cause of Peace”. One of the main aspects of this agreement was to prohibit the spread of fake news for propaganda purposes. Regarding the period of the agreement, the main concerns were about war propaganda and the related consequences. [4-1]

Articles three to five of the “International Convention Concerning the Use of Broadcasting in the Cause of Peace” mainly address the issues with fake news. Article 3 thereby states, that “any transmission likely to harm good international understanding by incorrect statements shall be rectified at the earliest possible moment”. The following Article 4 elucidates the goal of the agreement by petitioning the participating states to “ensure [...] that stations within their respective territories shall broadcast information concerning international relations the accuracy of which shall be verified”.

To reiterate, following the agreement meant that information needed to be checked and verified before being spread via radio broadcast. If a published statement was found to be incorrect or false, the responsible nation had to ensure that the spread of fake news is stopped and corrected as soon as possible.

While this agreement may be outdated in many aspects, the intentions are comparable with the current state of affairs regarding the spread of fake news.

These days different criminal codes acknowledge fake news as part of the definition of fraud. This means that in many countries, like for example the UK or US, fraud can be committed by spreading fake news if there is a related benefit for the perpetrator or harm for the victim. In German criminal law, the crime of fraud (§ 263 StGB) also requires the perpetrator to "distort or suppress true facts" ("Entstellung oder Unterdrückung wahrer Tatsachen"), which can be described as propagating fake news.[240]

Additionally, the crime of fraud according to the German criminal law also requires immediate disposal of property, which could cause some problems when trying to compare it to the spread of fake news. It might be quite difficult to prove a direct connection between an article that contains fake news and immediate disposal of property.

In addition to the offense of fraud, fake news can potentially fulfill the elements of offenses such as defamation (§ 187 StGB) or incitement of the people (§ 130 StGB), to name a few examples. A closer look on the offense of defamation (§ 187 StGB) shows, that if a perpetrator "asserts or disseminates an untrue fact about another person" ("in Beziehung auf einen anderen eine unwahre Tatsache behauptet oder verbreitet"), or in other words "if someone is spreading fake news about another person", he can be legally punished for that.

This rather small national sample size of legislation already shows the difficulty when trying to combat fake news effectively with the help of legal remedies. There are many different offenses in already existing criminal codes, which can be fulfilled under specific circumstances by spreading fake news.

In the following section, different examples of local legislation regarding fake news will be discussed and further examined. The analysis will focus on the criminal codes of member states of the CoE.

4.2.1. Local legislation regarding fake news (in Europe and other countries)

4.2.1.1. Germany – Netzwerkdurchsetzungsgesetz (NetzDG)

In 2018 the so-called "Netzwerkdurchsetzungsgesetz" (NetzDG) or "Network Enforcement Act" was passed in Germany to stifle different kinds of harmful actions on the internet such as illegal material, hate speech and fake news. The NetzDG applies to social media platforms with at least 2 million members in Germany [4-2].

According to section 3 of the act, social media networks have to implement an "effective and transparent procedure for handling complaints about unlawful content" (NetzDG, paragraph 3) which also includes the likes of fake news. In addition to that, any social media network that reaches this large of an audience has to "remove or block access to content that is manifestly unlawful within 24 hours of receiving the complaint" (NetzDG, paragraph 3).

[240] "Fighting Fake News or Fighting Inconvenient Truths?" - https://verfassungsblog.de/fighting-fake-news-or-fighting-inconvenient-truths/ (Last accessed 26.10.2021).

The responsibility of taking down potentially damaging content and thus supporting the competent authorities is delegated to social media networks. Some of them have already voiced their concerns about including Fake News in the scope of the NetzDG. In a study by Prof. Dr. Marc Liesching from July 2020 Facebook addresses the issue, that the NetzDG only applies if a statement or comment contains criminally relevant contents. Fake news can often not be considered in this context as they fall under freedom of speech [4-3]. Furthermore, this form of prosecution raises the question of whether a company should be eligible to decide if a post or comment is within the legal boundaries or not.

The second problem regarding the NetzDG is the national jurisdiction. On one hand, the scope of this legal regulation is clearly stated in Art. 1 to only social media networks with more than two million registered users in Germany. The scope of the NetzDG is pretty much narrowed down to only the providers of social media networks that meet these requirements.

On the other hand, the problem of limited national jurisdiction is addressed in Art. 4 NetzDG which outlines that "the regulatory offense may be sanctioned even if it is not committed in the Federal Republic of Germany". With this article, the German legislator states that their scope of national jurisdiction can be applied outside of the national territory. This raises concerns regarding the international law principle of territoriality [4-4].

This legal problem is not exclusive to the NetzDG. As already mentioned above, the German criminal code "StGB" also addresses the dissemination of fake news in section 263. The StGB regulates the jurisdiction of crimes committed outside of its territory as follows: "If the participant to an offense committed abroad acted within the territory of the Federal Republic of Germany, German criminal law applies to the participation even if the act is not a criminal offense according to the law of the place of its commission" (§9 StGB). Because almost all websites and contents are accessible in Germany, paragraph 9 StGB is very widely applicable. If executed strictly, this section would lead to an international jurisdiction for the German prosecution authorities when it comes to cybercrime.

4.2.1.2. France – Law against the Manipulation of Information

In comparison to Germany, the French legislation against fake news is focused on a different aspect of the topic. The so-called "Law against the Manipulation of Information", passed in December 2018, is centred around the fight against election misinformation. This legislation was caused by the attempt to interfere with the 2017 presidential election in France. Before the election took place, a coordinated attempt to undermine the presidential candidacy of Emmanuel Macron with the help of a systematic misinformation campaign was started. As a direct result of these actions the "Law against the Manipulation of Information" was enacted to prevent such campaigns in the future.[241]

To respect other rights such as freedom of expression and communication a fake news campaign has to meet certain requirements to be considered a punishable offense in the sense of the "Law against the Manipulation of Information". As such, the digital information has to be objectively false, misleading and threatening to the honesty of an upcoming election. This means that information or news must be considered manifestly false to be punishable. With this large barrier implemented freedom rights can be respected and protected [4-5, p.12].

The "Law against the Manipulation of Information" only addresses a small field regarding the topic of fake news. This further limitation of the material scope may raise questions and concerns about

[241] "Measures to tackle disinformation in selected places" - https://www.legco.gov.hk/research-publications/english/2021in14-measures-to-tackle-disinformation-in-selected-places-20210623-e.pdf (Last accessed 29.11.2021).

the overall usefulness of the legislation. The law only applies to fake news and misinformation that is demonstrably false. However, fake news is often not entirely false but exaggerated, sensationalized or taken severely out of context [4-5, p.13].

On the other hand, this limited scope of the French law guarantees the respect of the freedom rights mentioned above. These strict requirements ensure that only objectively and truly false information is considered illegal while dissenting opinions or statements are still protected from the danger of state-imposed censorship.

4.2.1.3. Global

On a global scale, several countries have addressed the issue of fake news in special sections within their respective criminal codes.

Canada used to have a legal section that addressed the spreading of false news in their criminal code. According to Criminal Code Section 181, if a perpetrator were to publish a statement or news that is known to be "false and that causes or is likely to cause injury or mischief to a public interest", they can be punished with imprisonment for up to two years.[242]

This section was repealed in 2019 because it interfered with the constitutional right of freedom of expression and was thus deemed unconstitutional.

This decision was made, after a neo-Nazi, who had published antisemitic literature, was prosecuted under section 181 of the Canadian criminal code. The Supreme Court of Canada ruled, that there is a fine line between the truth and falsehood, that cannot be defined by law. This then led to the repeal of section 181.[243]

Another prime example of legislation against fake news is the Anti-Fake News Act (AFNA), created by the Malaysian parliament in 2018. This ordinance punishes actions like the creation, publication, and distribution of fake news, "with intent to cause, or which is likely to cause fear or alarm to the public, or any section of the public" (AFNA, section 4). The AFNA was repealed in December of 2019 because of a new coalition after the national elections of that year.

This law is under heavy criticism because it was re-enacted in 2021 via a Proclamation of Emergency to fight the COVID-19 pandemic, whilst some believe the ordinance was promulgated to restrict media reports of the pandemic and thereby constrain aspects like free speech and press freedom.[244]

In comparison to the others, South Korea took quite a different approach. To combat fake news in their respective state, the ruling party wanted to change the "Press Arbitration Act" so that publishing fake news or false information could be punished. The bill specifically targets media outlets such as newspapers, magazines or TV and radio channels. If a media outlet publishes fake news, the

[242] Criminal Code (R.S.C., 1985, c. C-46)" https://laws-lois.justice.gc.ca/eng/acts/c-46/section-181-20030101.html (Last accessed 26.10.2021).
[243] "Is Canadian Law Better Equipped to Handle Disinformation?" - https://www.lawfareblog.com/canadian-law-better-equipped-handle-disinformation (Last accessed 27.10.2021).
[244] "The rebirth of Malaysia's fake news law – and what the NetzDG has to do with it" - https://verfassungsblog.de/malaysia-fake-news (Last accessed 27.10.2021).

retribution will be five times the estimated damage caused by false information. This damage will be calculated according to the social influence, as well as the total amount of sales or views.[245]

Instead of battling fake news on social media like Germany or in the context of presidential elections like France, the South Korean focus is more narrowed down towards "classical and established" media.

Within this approach, one big advantage can be determined: the prosecution of this bill. As outlined in chapter 3 – Technical Foundations, there are many options for internet users to conceal themselves and hinder or impede the law enforcement authorities. Another problem, especially regarding multi-national participants can be the jurisdiction in the context of state sovereignty. These difficulties are circumvented by the South Korean bill. Because the remedy focuses on media companies that are situated within their state, the law enforcement is similar to any other crime and does not have the issues associated with the prosecution of online crimes.

This bill still faces a lot of criticism by the government's opposition and international journalism organisations for encroaching on the "freedom of the press". As a result, the law was not passed as intended and the vote has been delayed until 2022.[246]

The differences between the main examples of national legislation against fake news, Germany, France and South Korea are listed in the following table to add clarity.

[245] "The Trouble With South Korea's 'Fake News' Law - " https://thediplomat.com/2021/08/the-trouble-with-south-koreas-fake-news-law (Last accessed 20.12.2021).
[246] "How South Korea Is Attempting to Tackle Fake News" - https://thediplomat.com/2021/11/how-south-korea-is-attempting-to-tackle-fake-news (Last accessed 20.12.2021).

	Germany	France	South Korea
Legislation	Network Enforcement Act	Law against the Manipulation of Information	Press Arbitration Act
Scope / Objectives	Preventing illegal content online (like fake news, hate speech, etc.)	Countering pre-election disinformation campaigns	Preventing media outlets from spreading false information
Parties regulated	Social network platforms (with over 2 million users in Germany)	Online platforms (exceeding a certain amount of distinct French users)	Media outlets (TV, radio, newspapers)
Regulatory tools	Social network platforms are required to remove illegal content within 24 hours	Online platforms have to provide a mechanism to report fake news, judges are responsible for deciding about the content before a general election	Courts can issue heavy fines on media outlets, if they are found guilty of spreading fake news
Criticism / Issues	Social network platforms have to decide in the first place whether a post is illegal or not; leading to the danger of overblocking	Fake news has to be determined objectively false within a short timeframe; in practice, this can be quite difficult to decide	Possible infringement of basic rights, such as the Freedom of Speech may occur

Table 1. National legislation against fake news

4.2.1.4. Conclusion

As shown with these brief examples, legislators in different countries all over the world have been trying to quash fake news with the help of regulations and laws. Their approaches might differ, but the development in legislation against the harm of fake news shows, that many countries are well aware of this problem and actively try to prevent the negative impact via legal remedies.

A common criticism against such laws is their encroaching on the freedoms of speech, expression and the press. The national legislators acknowledge these concerns and try to balance the fight against fake news with the preservation of individual and societal rights.

Further problems and legal issues, like the national jurisdiction, the applicability of laws online, as well as the principle of proportionality in the context of freedom-rights versus restrictive legal remedies, will be discussed and explained in section 2.2.3 of this chapter.

4.2.2. Legal Prerequisites

Local legislation concerning fake news can vary significantly, as showcased in section 1.1. As such, there are no general or international prerequisites, which would make something qualify as fake news and be punishable all over the world.

In the following sections, some examples of legal prerequisites will be showcased and discussed whilst referring to the previously outlined national laws from section 1.1.

4.2.2.1. A common definition is desirable to be able to take legal action against fake news

As explained in the beginning chapter of this book, an all-encompassing and universally correct definition of the term fake news simply does not exist. In different societies or legal and political systems, the conception of what exactly is fake news may vary in considerable ways. Therefore, legal prosecution especially across national borders is pretty much impossible.

A common definition of fake news may solve this issue or is at least desirable for governments to be able to suppress fake news effectively using legal remedies. Because most criminally relevant actions regarding fake news are often taken via the internet, a common definition is even more important due to the global reach of the medium.

Another reason that speaks to the importance of a clear-cut definition of fake news is the legal principle of “nulla poena sine lege” (“no penalty without law”). According to this principle, a clearly defined criminal offense is needed to enforce a prosecution. It is applied when there are no clear definitions or assessments for a criminal offense [4-6]. If "nulla poena sine lege” was neglected, judgments in fake news cases would be easy to challenge.

Even though a common definition would be ideal for legislators to be able to fight fake news with legal remedies, the controversies around the restrictions a law like that would pose to the freedom of the press, showcase that such a common definition is unlikely to find purchase in the near future.

Governments should therefore try to find other ways to counteract fake news on a smaller scale, whilst still working on a global solution.

4.2.2.2. Different treatments depending on the perpetrator

Is the legal treatment of fake news different, when it originates from an “official” news source, or even a government institution when compared to a private blogger who is spreading disinformation?

This question is not only considered as an ethical or political problem but can be examined in a legal context. First of all, it is important to determine whether a distinction between a private person/company or the public authorities is even necessary for a legal examination. From a societal point of view, this difference is a substantial part. If a person only speaks for themselves and shares fake news on his or her private blog, the amount of damage from this action is most likely much smaller than if a government organisation is spreading false information. The public authority would arguably have a higher responsibility to be truthful. The shown examples of national legislation further above do not differentiate between those two groups of perpetrators. This further emphasizes the thesis that the distinction is not relevant in a legal context and more of an ethical aspect.

A recent example of fake news published by a public authority is the case of the former Austrian chancellor Sebastian Kurz, who has allegedly spread disinformation as part of his political campaign. Before further explaining these actions, it is important to state, that by the time of writing this paper those actions are still under investigation and the presumption of innocence is still entitled.

The former Austrian chancellor and his allied colleagues are suspected of paying tabloids to publish fake opinion polls and overly positive media coverage before the national elections of 2016. The financing was allegedly funded by state money.[247]

Even though not all details of this incident are screened as of now, it already shows the difficulties and problems that the connection of false information and public authorities can cause. Whether a legal remedy against such behaviour would work is up for debate.

4.2.2.3. Are those regulations applicable to the internet?

At first glance, the internet might appear as a legal environment with a completely different set of rules or even as a place without any form of regulation. Whereas this might be a somewhat common point of view, it generally does not apply in a legal sense. Many laws that apply offline are equally applicable online. It is mandatory to state that this might vary regarding different national or local legislation.

Additionally, in this context, the question of fake news being a cybercrime or not can be raised. As with many other terminologies regarding this topic, there is also no clear-cut definition of what can be considered a cybercrime. One definition of cybercrime was published by Nir Kshetri in 2010, which states that "[...] a cybercrime is defined as a criminal activity in which computers or computer networks are the principal means of committing an offense or violation of laws, rules, or regulation" [4-7].

This pretty narrowed-down definition only focuses on crimes like fraud, forgery, data manipulation and other similar actions. In recent years other criminal offenses in the online spectrum have emerged and are no longer part of this definition. Actions like online stalking or bullying are noteworthy examples of such [4-8].

So, regarding Kshetri's definition, spreading fake news would not be considered a cybercrime. However, there are opposing opinions who are including fake news in the terminology of cybercrime. For example, Robert Smith and Mark Perry state in their article about "Fake News and the Convention on Cybercrime", that spreading fake news on social media can be considered a cybercrime [4-9].

Even though these statements support different opinions they have something in common, both of them only focus on the online aspect of fake news. In these articles, publishing or spreading of fake news in other forms of media, such as TV, radio or newspapers, are not taken into consideration. Other than hate speech, which is committed most of the time in an online environment, fake news is not necessarily a cybercrime. This marks a clear distinction in a legal context between the likes of hate speech in comparison to fake news.

Fake news can be published online, but are also oftentimes spread via other forms of media. Therefore, the term cybercrime is not perfectly suited for fake news, as the criminal acts are equally happening offline and online. Regarding fake news, the term cybercrime may be applicable for some aspects of it but certainly not the whole topic.

[247] "Austria: First arrest in Kurz corruption probe – reports" - https://www.dw.com/en/austria-first-arrest-in-kurz-corruption-probe-reports/a-59483916 (Last accessed 30.12.2021).

Besides the question regarding the applicability of laws online and the terminology of cybercrime, another urgent problem is the legal jurisdiction, particularly concerning legal cases across national borders. The global nature of the internet might lead to conflicts between different national laws and moral concepts. Crimes like the spread of fake news introduce a new possibility: a criminal action can be committed in one state whilst unfolding its effect in a different state. One action might be legal in the state where it is taken but be illegal in the state where its effects take place. This undermines national sovereignty with the principles of national legislation and criminal prosecution. As already stated in section 1.2.2 the legal jurisdiction can be a crucial aspect for example whilst debating between contradicting legal regulations.

In the EU there is a specific regulation for organisations regarding this exact problem. The so-called country of origin principle (with a few exceptions) was introduced by the e-commerce directive (2000/31/EC) to address this issue [4-10].

Article 3 of the directive states that "Each Member State shall ensure that the information society services provided by a service provider established on its territory comply with the national provisions applicable in the Member state in question which fall within the coordinated field" [4-11].

This means that in the EU the establishment of an organisation is the decisive reason on which national laws apply to them. So, if for example, an organisation located in member state A spreads fake news regarding a topic on their website, hosted by a server in member state B, the legal regulations of member state A are applicable and not the national laws of state B. The jurisdiction is thereby regulated clearly by article 3 of this direction. In a case of fake news spreading within the EU the jurisdiction resides with the member states in which the organisation that caused the incident is established.

As already stated, this regulation only applies to organisations or companies and not to private individuals. Furthermore, the directive is inapplicable to organisations established outside of the EU and its member states. Therefore, organisations that want to cause harm by spreading fake news can circumvent the regulation quite easily by simply setting up their head office in another country outside of the EU.

Another difficulty, which has been briefly mentioned before in this chapter, is the principle of proportionality regarding legal remedies introduced by constitutional states. This is one of the main reasons why many examples from section 1.1 "Local legislation regarding fake news (in Europe and other countries)" struggled with their respective legislation. Legal regulations against fake news almost always come with a more or less extensive restriction on other fundamental rights, such as the freedom of speech or the freedom of information. While the principle of proportionality applies to the internet as well as to offline circumstances, about the freedom of speech, the internet is oftentimes seen as an "unregulated place", where anybody can share their thoughts freely. Thereby making a regulation to combat fake news in an online environment is a difficult task for any legislator because arguments like censorship almost immediately arise.

Whereas these principles are present in all constitutional states, the reach or possible limitation of those freedom rights are different in their specific manifestations. When comparing the European country Germany to the United States of America, this difference becomes much more prevalent. In Germany, for example, the freedom of speech is restricted by section 130 of the criminal code, where expressions like approving or denying "an act committed under the rule of National Socialism" (§ 130 StGB), are legally punished.

On the other hand, in the United States of America, attempts to combat fake news with the help of legal restrictions contradict the First Amendment where “freedom of expression by prohibiting Congress from restricting the press or the rights of individuals to speak freely” is guaranteed by the constitution. This showcases, that different countries place different emphasis and boundaries on their citizen’s fundamental rights.

Described by Dr. Ron Paul as a “war on Free Speech” the proposed legal regulation against fake news is seen from a different perspective, with a much bigger focus on freedom rights and a distinct refusal of any sort of government restriction [4-12].

So as a conclusion, it is important to state, that applying legal regulations to the internet is very much possible and laws are often either applicable online as well as offline, or even specifically meant to combat crimes on the internet. The difficulties hereby are usually less of a legal problem like a missing criminal offense in the applicable laws and more of practical execution.

In many cases, the prosecution of internet crimes is too complex and disproportionate to the seriousness of the offense. Also, the legal jurisdiction is an important aspect when trying to combat fake news online, as the perpetrator, potential victims, the server, and the internet platform or social media network can all be located in different countries with a variety of potentially applicable laws.

4.2.3. Hate speech

4.2.3.1. The legal definition of hate speech

4.2.3.1.1. Definition of hate speech as determined above

“Hate speech is to be understood as the advocacy, promotion or incitement in any form of denigration, hatred or disparagement of any person or group of persons, as well as any harassment, insult, negative stereotyping, stigmatization or threat to such person or group of persons, and the justification of any of the foregoing on the grounds of 'race', color, descent, national or ethnic origin, age, disability, language, religion or belief, age, disability, language, religion or belief, sex, gender identity, sexual orientation and other personal characteristics or status, as well as the form of public denial, trivialization, justification or approval of genocide, crimes against humanity or war crimes found by courts of law, and the glorification of persons convicted of committing such crimes.”

That is the definition of hate speech from the Council of Europe as it was shown in the first chapter. But just as there is no clear and universal default definition, there is also no internationally legally recognized definition, what is considered hateful is disputed [4-13].

Therefore, each country has its own definition and its own way of dealing with hate speech, as the following examples will show:

4.2.3.1.2. Different legislations

4.2.3.1.2.1. Germany

In Germany, criminal offenses are punished under the German Criminal Code (Strafgesetzbuch). For the offense of hate speech, several paragraphs can be used. For example, the sections "§ 130 Volksverhetzung" or "§ 185 Beleidigung" come into question. Hate speech is also defined in § 130

section 1 No. 2 of the German Criminal Code. This is understood to mean a disruption of public peace by inciting hatred, violence or arbitrariness towards groups, population groups or individuals based on nationality, religion or ethnic origin. This also includes an attack on human dignity in which an individual is maliciously insulted, defamed or despised, based on his or her membership in a particular group or part of the population [4-14].

As already mentioned in chapter 1.1 Fake news - local legislation, the NetzDG came into force on September 1st, 2017, to combat not only fake news but also other cybercrimes such as hate speech [4-15].

It is intended to ensure that criminal content is deleted in social networks within a reasonable time frame. Similar to fake news, in the case of hate speech, difficulties arise when assessing the content, or the network provider does not adhere to the given deadlines for deletion. The Federal Office of Justice acts as a supervisory authority and can impose fines in the event of non-action.[248]

4.2.3.1.2.2. France

In France, criminal offenses are punished according to the French criminal code “code penal” Hate speech is punishable there as "discours de haine” [4-16].

The proposal for a law to combat hateful content on the Internet is intended to increase the obligation to remove hateful comments by making criminal reactions more efficient and preventing dissemination. Furthermore, the responsibilities should be clarified and the platform operators should be asked to cooperate more. Proposition de loi visant à lutter contre les contenus haineux sur internet, “Avia Law” is based on the German NetzDG and also has a definition of hate speech [4-17].

4.2.3.1.2.3. Austria

In Austria, hate speech, if it is made public and accessible to many people, is defined as incitement to hatred under § 283 of the Austrian Criminal Code. This paragraph also contains a definition that is very similar to the German definition. [4-18]

4.2.3.1.2.4. Interim conclusion (member states of the EU)

As can be seen in the examples shown above, the European countries all have a similar point of view regarding hate speech. Nevertheless, there is still a need for action, as an example from 2015 shows. In a report published in 2015 by the European Commission against Racism and Intolerance (ECRI), incitement to racism was only punishable in Estonia when the victim's health, life or property were threatened [4-19].

However, in the European Union, the definition, the elements of the crime and the opinion that hate speech should be legally prosecuted are largely the same. Other cultures may have completely different moral values and views. An example of this is the USA, which has a very different view on hate speech.

4.2.3.1.2.5. United States of America

In principle, it must be mentioned that constitutional protection is very pronounced in the United States of America. Among other things, this is due to the fact that the American Constitution is very

[248] “Hasskriminalität in sozialen Netzwerken bekämpfen” - https://www.bundesjustizamt.de/DE/Themen/Buergerdienste/NetzDG/NetzDG_node.html (Last accessed 12.01.2022).

old compared to most other countries. In addition, the American courts grant almost absolute protection to freedom of expression, which also distinguishes them from the courts of other countries. This is based on the fact that there is a strong distrust of government. People believe in a competition of opinions and are not supposed to distinguish between good or bad opinions [4-20].

A famous United States Supreme Court ruling has defined the limits of freedom of speech. In the 1969 case of Brandenburg v. Ohio, a Ku Klux Klan leader (Clarence Brandenburg) called for the possibility of revenge against “Jews” and “Negroes” and announced a march on the United States Congress. Brandenburg was initially convicted for this speech, but the Supreme Court overturned the conviction, calling Brandenburg's actions “imminent lawless action”, which cannot be punished under the Constitution [4-21].

4.2.3.1.3. Conclusion / Problem

Aiming to avoid serious differences in the handling of hate speech, as shown by the example of the USA and European countries above, international rules must be established. This would facilitate prosecuting and punishing hate speech crimes internationally in a uniform way. These should best be established by major international associations or organizations to include as many different countries as possible. In the first instance, this should be incorporated into international law by the United Nations.

4.2.3.1.3.1. International law / United Nations

Under international law, no provision would prohibit hate speech at the moment. This is even though hate speech has a big impact on many areas of activity from the United Nations, such as the protection of the population and the fight against violence, racism and discrimination. Taking into account the international Human rights norms and standards and the right to freedom of expression, the United Nations has adopted a UN Strategy and a Plan of Action on Hate Speech. These grant the United Nations the necessary resources to take action against hate speech. The United Nations tactic is to undermine the causes of hate speech and to find effective responses to the impact of hate speech on society.

Also, incitement to violence, discrimination or hostility is nevertheless prohibited. This is a special form of hate speech, which is particularly harmful because of the risk of misdeeds or even terrorism [4-22].

4.2.3.1.3.2. European Union

Another supranational organization is the European Union, which, with its many bodies and member states, also has the possibility of drawing up supranational rules. Because of its smaller size compared to the United Nations, it also has a greater chance of ensuring that the rules it draws up are accepted by all and that compliance can be better monitored. So far, there is no legal regulation on hate speech at the European level. However, there are various approaches from European bodies such as the Council of the European Union or the European Commission to take action against hate speech. One possibility is “soft-laws”, which prescribe certain things just like real laws, but these do not have to be adhered to and non-compliance is not punished [4-23].

European Code of conduct

The European Commission, Facebook, YouTube, Microsoft and Twitter signed the Code of conduct (CoC) on countering illegal hate speech online on May 31, 2016. Instagram, Dailymotion, Snapchat

and Goggle+ were added later. Jeuxvideo.com joined in January 2019, also TikTok in September 2020 and LinkedIn in June 2021 [4-24].

The Code of Conduct obliges the companies to check reports of hate speech within 24 hours, delete illegal content or block users. As reported by the Commission, IT companies investigate 89% of reported cases within 24 hours, of which 72% are deleted due to illegal content [4-25].

Thus, the Code of Conduct is the significant instrument for self-regulation of illegal hate speech on the Internet [4-26, p.53].

Guideline-Audiovisual Media Services

The European Audiovisual Media Services Directive (AVMS) was also amended to combat hate speech. The regulations, which previously applied only to broadcasters, now also cover video-sharing and video-on-demand platforms such as Netflix, YouTube and Facebook. As a result, video platform operators are required to create easy-to-understand and use mechanisms through which videos containing hate speech or glorifying violence can be reported and, after subsequent review, deleted by the operators.[249]

European Commission

Combating hate speech and hate crime is also a priority for the President of the European Commission, Ursula von der Leyen. On February 23, 2021, the European Commission published a proposal to declare hate speech an EU crime. At the moment, the European Commission is working on an initiative that should lead to a Council decision against hate speech. If this Council decision is taken, the European Commission will then have the power to propose substantive legislation. This would make it possible to standardize definitions and penalties for hate speech.[250]

4.2.3.1.3.3. Council of Europe

Minimum rules for the definition of criminal offenses and sanctions are necessary and decided to align laws and regulations for the implementation of a common policy of member states. According to Art. 83 section 1 TFEU, the European Council may, through a legislative procedure, lay down guidelines for minimum rules on criminal offenses. However, a special need and a cross-border dimension are required for this. The directives apply to many areas of crime, including cybercrime. If there are concerns about a directive, under section 1 or 2 regarding its compatibility with the respective criminal laws in the countries, a member of the Council of the EC can refer this to the ER for consideration and request a suspension and deliberation on it. A decision will be made within 4 months.

There are often differing opinions when it comes to establishing guidelines, but if at least nine Member States reach an agreement about cross-border cooperation, they manifest their decision to the European Parliament, the European Council and the European Commission within the defined 4 months. The conclusion is deemed to have been granted [4-27].

[249] "EU-Richtlinie für audiovisuelle Mediendienste" - https://www.medienkorrespondenz.de/politik/artikel/eu-richtlinie-fuer-audiovisuelle-mediendienste-umsetzung-bis-septembernbsp2020.html (Last accessed 04.01.2022).
[250] "Commission: Hate Crime Should Become an EU Crime" - https://eucrim.eu/news/commission-hate-crime-should-become-an-eu-crime/ (Last accessed 04.01.2022).

European Commission against Racism and Intolerance

Besides these legal attempts, the European Union has established some bodies that are only there to protect human rights such as the European Commission against Racism and Intolerance (ECRI), which specializes in combating discrimination and racism. This is closely networked with the Equal Treatment Bodies of the Länder and thus monitors them. The ECRI monitors Member States by analysing the circumstances and the actual state of affairs. When problems arise, the ECRI put forward proposals and makes recommendations. The equality bodies are independent authorities that combat racism and discrimination at the national level. In addition, relations are maintained with international organizations, such as the United Nations [4-28].

European Court of Human Rights

Another body of the European Union is the European Court of Human Rights (ECHR), which is a supranational judicial body that ensures that member states respect the human rights set out in the Convention on Human Rights. All 47 member states are also members of the Council of Europe. If the ECHR would develop a common practice/definition per jurisdiction, where the important parameters are defined, the 47 member states with 47 different criminal codes would not have 47 ways the courts apply them. For hate speech to be punishable uniformly, it is important to define the parameters. Does the post need to be public and have a specific reach or is an insult through a private message enough to be considered hate speech? Does it have to be specifically directed and received by the targeted person, or is it enough if the hate speech is posted in a private forum, which excludes the targeted group?[251]

Summing up, there are several enactments and rules concerning the handling of hate speech, but a common European guiding principle is missing. The present definitions of hate speech, no matter nation or international, differ about the modalities, the impacts as well as the consequences.

A guidance note issued by the "High-Level Group", which was launched by the European Commission in 2016, refers merely to the observance of case law in the responsible member states. Combating illegal online content and disinformation cannot be solved nationally. It must be tackled by the Member States together. Basic approaches already exist, for example, "Tackling online disinformation: a European approach," but need to be expanded. The focus should be on promoting tolerance and plurality in public institutions and in positions with political, media and financial responsibilities. The policy must represent values and set standards as well as communicate improved, wide-ranging and coherent solutions [4-29, pp.21,53].

4.2.4. Are those regulations applicable to the internet?

Hypothetically, if there was a common legal definition for hate speech, could regulations against it be applied to the internet?

As already stated in section 1.2.3 of this chapter, the question of whether legal remedies are applicable online or not can often be quite difficult to answer. In the following segments, these problems are going to be further addressed and discussed to emphasize the explanations given in the section about fake news prior.

[251] "What is the European Convention on Human Rights?" - https://www.amnesty.org.uk/what-is-the-european-convention-on-human-rights (Last accessed 20.01.2022).

4.2.4.1. Determining the perpetrator

As someone can be using the profile of someone else, either by hacking into their account or simply using their computer to post something through their profile it is nearly impossible to be able to determine a poster's identity with 100% certainty. This legally makes determining the perpetrator without reasonable doubt complicated. If hate speech was posted through someone's personal computer on their account and they live alone, it's difficult to argue that they weren't the perpetrator.

However, if it was a PC at work and they happened to not lock their screen when for example getting a cup of coffee, the perpetrator could have been several people who had access to that PC. In the German legal system, the company owns the account and is thereby liable to third parties under private law for any actions taken via this account.[252]

4.2.4.2. Legal jurisdiction

In addition to the many different definitions, the preconditions for committing a crime, the way the courts prosecute these crimes, and the problem that these rules cannot be perfectly applied to the Internet, there is another key issue. In the case of crimes committed over the Internet, it is also not clear at the first moment which court has jurisdiction at all. The territorial principle that assigned jurisdiction in the past no longer works in today's world with modern technology. When clarifying jurisdiction, several factors must be taken into account, such as the location of the perpetrator, the nationality of the perpetrator, the location of the victim or the nationality of the victim. Also to be included is the portal or platform through which the crime was committed. Thus, the laws of the country in which the company headquarters or servers are located may also become relevant.[253]

As the internet is accessible all over the world, determining which country's laws apply to a situation can be complicated. This situation can even get more complicated depending on which of the locations are within the EU or outside the EU. There are several factors to consider, for example, if person A posted hate speech against person B on a popular platform, that is accessible from the EU Member States, which of the following is relevant to pinpoint which countries' jurisdiction applies to the case.

- Person A's nationality/country of origin/ citizenship?
- Person A's location when making the post?
- Person B's nationality/country of origin/ citizenship?
- The location where the information is accessible?

[252] "Use of company internet connections and e-mail accounts as well as mobile phones and notebooks" - https://www.anwalt.de/rechtstipps/nutzung-betrieblicher-internetanschluesse-und-e-mail-accounts-sowie-von-mobiltelefonen-und-notebooks_062654.html [Last accessed: 01.21.2022].

[253] "Territorial principle of Germany" - https://www.juraforum.de/lexikon/territorialprinzip (Last accessed: 01-21-2022).

Victim/Perpetrator	Same Country	Different Country within CoE	Country outside CoE
Country A (CoE)	Laws of country A fully applicable	Application of Country A's laws limited	Likely no application of Country A's laws. In rare cases bilateral treaties
Country B (non-CoE)	Laws of country B fully applicable	Likely no application without bilateral treaties	Likely no application without bilateral treaties

Table 2. Sorting out legal jurisdictions

If this would not make the clarification of jurisdiction difficult enough, it must also be taken into account that each country has its own special rules for law enforcement.

Example: In Germany, the legal jurisdiction is determined in §§ 5 and following the German Penal Code (StGB). The so-called "Handlungs- / Erfolgsort-Prinzip" contains rules on jurisdiction. For websites, this means that the availability of the content in Germany is sufficient for the German Penal Code to apply, thus practically everywhere. Also, some crimes can be generally persecuted by law even if they were committed outside Germany, for example, §130 section 2 No. 1 StGB "Volksverhetzung", which can also be applied to hate speech [4-30].

Server location

In addition to the nationality of the perpetrator/victim, the server or business location may also be relevant. The Canadian Court of Justice has faced this issue before. The case considered whether a Romanian website that had no servers or business locations in Canada was subject to Canadian laws. The content was passed from the website, which was taken from the Canadian database of court decisions CanLII.org. The original website ensured that personal information provided by litigants could not be found by search engines. This was not done by the Romanian website. The Canadian court then found jurisdiction. The reason given was that although the business location and the server location were in Romania, there was nevertheless a sufficient connection to Canada, as Canadians were affected.[254]

A similar point of view was shown at the conference "Law, Borders and Speech", which was held at Stanford. At this conference, the importance of the server location in clarifying jurisdiction was discussed in particular. From Silicon Valley, the opinion was expressed that the server location should be the decisive factor in determining jurisdiction. The U.S. Court of Appeals for the Second Circuit takes a similar view, mentioning this in the comments to the decision in the lawsuit "14-2985 In the Matter of a Warrant to Search a Certain E-Mail Account Controlled and Maintained by Microsoft Corporation" with Microsoft.[255]

[254] "Server location not definitive in determining jurisdiction over foreign defendant" - https://www.lexology.com/library/detail.aspx?g=9ad16ad9-c363-4e00-a293-c61f19f9fcf6 (Last accessed: 01-20-2022).
[255] "Server Location, Jurisdiction, and Server Location Requirements" - https://blog.ericgoldman.org/archives/2016/12/server-location-jurisdiction-and-server-location-requirements-guest-blog-post.htm (Last accessed 15.01.2022).

Bilateral treaties

Enforcement of court judgments in many countries, including the US, depends on the principles of comity, reciprocity and res judicata and ultimately on the internal laws of each country. Bilateral or multilateral treaties and agreements can be made between countries to regulate legal issues and the recognition of judgments and their enforcement. There is no agreement in this regard with the United States. Reasons for this seem to be, among other things, the high sums imposed by US courts in connection with liability claims. Many countries consider the fines to be too high. There are also different opinions regarding extraterritorial jurisdiction, which prevent joint agreements with the US. However, in most countries, foreign judgments not providing for damages can generally be enforced if the following conditions have been met:

- the court, that made the judgement was authorized to judge and had jurisdiction in the designated case;
- the defendant was informed of all relevant facts about the case;
- the process was not influenced by fraud;
- the judgment was compatible with the public order of the country.

If a judicial decree does not require compensation for damages, after approval by the domestic local court, the judgement can in many cases be enforced, despite differences in the procedures of the countries.[256]

4.2.5. Hate speech vs. freedom of expression and freedom of religion

In the second chapter, it was explained that the right to freedom of expression is a fundamental human right, which was included in the Universal Declaration of Human Rights. In Germany, it is enshrined in the Basic Law of Germany (Grundgesetz). Art. 5 section 1 Grundgesetz grants everyone the right to freely express and disseminate their opinions in speech, writing and images. In the well-known Lüth decision, the Federal Constitutional Court describes this fundamental right as a fundamental element of democratic state order and a direct expression of human personality [4-31].

However, unlimited freedom of expression is not granted, so in certain cases, restrictions are applied. For example, statements that incite, encourage or justify hatred based on intolerance. Certain statements thus fall into the category of hate speech and are therefore no longer protected by the fundamental right to freedom of expression [4-29, p.16].

Moreover, the Federal Republic of Germany guarantees according to Art. 4 sections 1 and 2 Grundgesetz, everyone the right to confess a religion, to join it or to change religious affiliation, as well as the right not to confess any religion or to leave a religious community.[257]

Freedom of religion, also guaranteed by Art. 18 ICCPR (International Covenant on Civil and Political Rights) often leads to discussions and is considered controversial because it affects other fundamental

[256] ”Enforcement of Judgments” - https://travel.state.gov/content/travel/en/legal/travel-legal-considerations/internl-judicial-asst/Enforcement-of-Judges.html (Last accessed 10.01.2022).
[257] ”Religious Constitutional Law” - https://www.bmi.bund.de/DE/themen/heimat-integration/staat-und-religion/religionsverfassungsrecht/religionsverfassungsrecht-node.html (Last accessed 23.01.2022).

rights. Freedom of religion can be restricted too, to protect fundamental rights and freedoms, health and morals, and in the event of violations of public order and security.[258]

Concerning freedom of expression and freedom of religion, each case must be carefully considered. In regards to different speech restrictions, there are distinctive case-laws applied by the European Court of Human Rights, but there is still no precise specification about hate speech. So, there is a need for clarification and guidelines [4-26, p.34].

The answer to the question of whether freedom of expression also applies to use on the Internet is yes. However, it is not permitted, and therefore no longer covered by freedom of expression, if the personal rights of persons or groups of persons are thereby violated. This includes insults or falsehoods about them. It is also not allowed to incite violence on the Internet and it is forbidden to post pictures with certain symbols, for example, the swastika. All statements classified as hate speech are punishable, whether on the Internet or in real life [4-32].

Thus, everyone should be aware that statements made in the online world can have serious consequences. Some comments are posted in a small or closed group within social media, but others are posted on a public, mass-accessible platform that reaches a global audience. It seems that social networks are communicating with an ever-increasing number of people. All statements and content can be called on the Internet at any time and can therefore also spread uncontrolled. All statements once posted are available in the social networks until they are deleted, in contrast to verbal statements. Quick deletions are difficult but necessary to minimize the potential damage of some utterances.

4.3. Possible legal approaches

This section will explore possible approaches to combat both hate speech and fake news. By analysing each approach and determining its short- and long-term effects, the goal is to find a recommendation for the procedure.

4.3.1. Platform Liability

This approach makes any platform that is accessible from within the European Union liable for the content that is on their platform. It aims to shift the responsibility of keeping a website clear of hate speech and fake news to the platform owner themselves by relying on their economic interest to have access to the European userbase. A similar strategy was utilized for the German Netzwerkdurchsetzungsgesetz as showcased in 1.1 of this chapter.

The perceived advantage of such a method is that large platforms are easier to address and sue when compared to individual users. One of the prerequisites to apply such a measure is clearly defining what constitutes hate speech and fake news respectively, as the platform owners themselves would have to determine what content they would have to take down.

However, several problems may arise when using this strategy:

Large platforms host huge amounts of content, with additional massive amounts being uploaded daily. For example, on the social media platform Twitter the daily upload is on average 500 million posts

[258] ”Freedom of religion and freedom of speech” - https://menschenrechte-durchsetzen.dgvn.de/menschenrechte/politische-buergerliche-rechte/religionsfreiheit-und-meinungsfreiheit/ (Last accessed 23.01.2022).

which translate to 200 billion posts per year.[259] As such, the expectation of content being checked manually is not realistic. A combination of automatic filters, as well as a reporting system, would be required, that allows visitors of the platform to report each other's content in case of violation.

Automatic filters using algorithms or artificial intelligence have come a long way in the past decade and are actively in use to take down posts that infringe the copyright or that depict child sexual abuse. However, they haven't been sufficiently developed to be able to discern the intent of a post [4-33]. Whether something qualifies as hate speech or remains within the boundaries of fundamental rights such as the freedom of expression, can be controversial [4-34].

A well-known example of courts being in disagreement over a matter of hate speech was the case of the German politician Renate Künast, who was the target of a large mass of insults online because of a remark she made regarding sexual abuse towards children. The first instance court ruled that the insults directed at her were not hate speech and that, as a politician speaking about a delicate topic, she was supposed to withstand harsher forms of criticism. Künast appealed against this ruling and the case went to higher courts that ended up revising the decision several times, each court disagreeing with the previous ruling [4-35, 4-36].

As such, writing a general predetermination into a program aiming to distinguish between something being hate speech or protected speech is unlikely to produce sufficiently accurate results – especially when dealing with 47 different legal systems in the 47 member states of the Council of Europe.

Depending on the severity of the sanctions toward the platform owner, they may choose to use filters that follow the philosophy of "rather safe than sorry". Doing so will result in a lot more content being flagged and removed than intended by the policy writers, resulting in so-called "overblocking" [4-37, 4-38].

Smaller companies and platforms, that do not have access to advanced filtering technology, would struggle to keep up with the Silicon Valley tech giants, forcing them to rely on their technical solutions and thus creating additional market entry barriers [4-39].

Platform owners may also choose to simply not make their content accessible for Europe-based users, as a way to avoid liability issues – similar to what happened when the GDPR was introduced.[260] If enough platforms react in this fashion, citizens of member states may end up in digital isolation from the rest of the global userbase. That could, in turn, result in a sizeable public backlash.[261]

In short – making the platform liable for the content on it would arguably be an effective way for governments to restrict undesirable content. However, the de-facto delegation of complex judiciary tasks to private companies without any financial compensation is likely to have external effects. It could lead to considerable downsides for internet users such as the restrictions being heavier than intended by the legislators. It would make the owners of social media platforms the arbiters of what falls under freedom of speech and what needs to be taken down ebased on either fake news or hate speech. It would also shift the liability towards the platform owners and generate additional costs for them.

[259] "Internet Live Stats", Available at https://www.internetlivestats.com/twitter-statistics/ (Last accessed: 25 January 2022).
[260] SENTANCE, REBECCA, GDPR: Which websites are blocking visitors from the EU?, 2018, Available at https://econsultancy.com/gdpr-which-websites-are-blocking-visitors-from-the-eu-2/ (Last accessed 28 January 2022).
[261] SOUTH, JEFF, More than 1,000 U.S. news sites are still unavailable in Europe, two months after GDPR took effect, 2018, Available at https://www.niemanlab.org/2018/08/more-than-1000-u-s-news-sites-are-still-unavailable-in-europe-two-months-after-gdpr-took-effect/ (Last accessed 28 January 2022).

4.3.2. Blocking Access

A straightforward way to deal with platforms not abiding by European laws, regulations, or standards would be blocking access to them for users within the member states.

To successfully set this up, a list of standards would need to be created that are to be followed by websites wanting access to the European market. Those standards would have to be attainable and maintainable. Within the context of hate speech and fake news, however, this might prove more difficult. Hate speech in particular is most often produced in social media by the platform's users - rarely by the platform owners themselves. As such, it may be difficult to curate the posted content, especially for larger platforms for the reasons outlined under 3.1 of this chapter.

A blanket-blocking of (often foreign) websites and platforms might lead to geopolitical consequences. Similar requirements of following foreign standards may be imposed on European-based platforms by other countries in the long term, thus potentially disassembling the internet along national lines due to cultural differences.

Therefore, a possible public backlash needs to be taken into account in this case, as it constitutes a heavy restriction on the freedoms of European citizens. The preventative attempts may be viewed as exclusion and censorship and might even increase the interest in the banned sites, thus causing a so-called Streisand effect [4-40].

Finally, blocking access to certain websites based on a user's physical location or country of origin might very well be futile. As outlined in chapter 3, such a restriction can be easily circumvented by utilizing VPN clients and proxies. Successfully stifling access to the targeted websites seems unlikely as a result.

4.3.3. Liability of the Individual

An opposing approach to making platforms liable for content that was posted on them is to attempt to prosecute the individual who uploaded the illegal content in the first place.

This avoids having to use the platform as a scapegoat or legal arbiter as well as establishing direct consequences for transgressions in matters of hate speech and fake news. Additionally, it doesn't move the responsibility of prosecution away from the judicial branch.

This method may also face several issues, that can be divided into three categories: Detection, Identification and Prosecution.

4.3.3.1. Detection

Detection involves recognizing illegal content as such and commencing the legal process against the responsible individual.

As illustrated under 3.1 the sheer mass of data that gets posted onto social media daily makes a manual checking of everything nearly impossible. A system like that would require either automated checking of content before it is uploaded or an integrated report system for other users to utilize. Both of these would have to be implemented into all social media platforms. Furthermore, the intricacies of distinguishing hate speech and fake news from the legal and protected speech are difficult to program into an algorithm. If the platform is not the one responsible for checking whether a post is illegal,

then government or non-government agencies would have to be established to do so. Even with assuming that common definitions and outlines are accepted internationally, the number of media to check would likely be overwhelming. To reduce the workload the responsible agency could focus on content flagged by users, though even that is unlikely to make the number manageable.

4.3.3.2. Identification

Assuming a piece of uploaded content was found to be illegal and the poster needs to be held accountable, it would be required for the responsible individual to be identified without a doubt. Such an identification can be problematic through the internet for several reasons: On the one hand, users on social media platforms tend to make use of pseudonyms, rather than using their real names – and on the other hand, a perpetrator could be using someone else's profile, account or computer to mask their identity. While one could attempt to trace a user's IP address back to them, there are several ways for them to avoid being successfully tracked, as discussed in Chapter 3.

Several politicians and influential political figures have instead called for an enforced deanonymization of social media platforms with the hopes of increasing accountability on the internet.[262] [4-41] This would mean that a person could only register into social media (or other websites that allow users to interact with each other) using their real first and last name, possibly in addition to other personal data that can be used to identify them. The proponents of this method aim to improve internet culture by taking away a user's ability to conceal themselves.

Experts in the field are concerned that such a mandate could bring about a lot of undesirable side effects [4-42]. A common example is that an employee could no longer criticize their company or superiors without fearing retaliation.[263] Marginalized groups may face difficulties when wanting to express their opinions and voices – leading to a chilling effect that stifles freedoms of speech and free expression [4-43]. Additionally, users having to display their real names and possibly additional personal data could make it easier for them to get targeted by doxing – meaning that their private address and other sensitive information could be leaked into the internet, making them easy targets for retaliation, stalking or other crimes.

Studies on online culture have shown, that changes in tone are negligible between users using pseudonyms or their real names [4-44, 4-45, 4-46]. A user who wants to spread hateful speech will do so regardless of whether their real name is shown or not.

In the case of South Korea, a country that enacted a real-name policy on social media back in 2007, malicious comments on internet forums decreased only by 0.9%. At the same time, hackers were able to take advantage of the citizen's personal information being stored in the website databases for identification purposes. A single cyber-attack leaked the personal information of over 35 million Koreans – which amounted to more than half of South Korea's national population at the time [4-47].

---

[262] WITTENHORST, TILMAN, Gegen Hetze im Netz: Schäuble fordert Klarnamen-Pflicht, 2019, Available at https://www.heise.de/newsticker/meldung/Gegen-Hetze-im-Netz-Schaeuble-fordert-Klarnamen-Pflicht-4425451.html (Last accessed 28 January 2022).

[263] KELBERER, ULRICH, "Klarnamenpflicht im Netz vertreibt nicht den Hass, sondern unsere Freiheit", 2020, Available at https://www.focus.de/digital/internet/gastbeitrag-von-ulrich-kelber-eine-klarnamenpflicht-im-netz-vertreibt-nicht-den-hass-sondern-unsere-freiheit_id_11614881.html (Last accessed 29 January 2022).

Finally, the South Korean Constitutional Court declared the real-name policy unconstitutional in 2012 and it was abolished as a result.[264]

A recent ruling by Germany's federal court lines up with this analysis, as they support their citizen's rights to make use of pseudonyms – preventing Facebook from demanding real names from users that have been on the platform for longer than four years.[265]

Identification of Users on the internet has become more reliable over time with the help of algorithms, meta-data and machine learning,[266] but it remains difficult and unreliable when an individual knows how to conceal themselves.

4.3.3.3. Criminal Prosecution

Even when a post has been flagged and determined as either hate speech or fake news and the user who is responsible for that post has been positively identified, that is still no guarantee for prosecution or any consequences for that individual.

The internet is globally accessible, but perpetrators are not. As showcased in 2.2.3. of this chapter, a state's jurisdiction will rarely allow them to prosecute criminals beyond their borders. Those cases are rare for murderers, war criminals and other forms of wanted individuals.[267] [4-48] Expecting to be able to extradite someone who may or may not have posted hate speech on the internet seems optimistic at best.

Another thing to be considered is that the damage of hate speech campaigns or the spread of fake news is inflicted quickly. It has been shown that falsehoods are several times more likely to be shared on social media when compared to facts [4-49]. Thus, even if a smear campaign is shown to be false, the damage has already been done.

A prosecution across national borders, even if successful, would potentially be a years-long process. With the ubiquity of hate speech across social media, processing the cases would drain sizeable resources both in time and money.

In conclusion, attempting to prosecute for hate speech and fake news on an individual level would use vast amounts of resources and face numerous difficulties in the process – all for likely disappointing results.

4.4. Conclusion

There are numerous definitions and approaches to fake news and hate speech across different countries. This difference will potentially make it difficult to agree on an internationally uniform approach to resolving these problems.

---

[264] KYUNGHYANG, SHINMUN, Internet "Real Name" Law Violates the Constitution, Of Course, 2012, Available at http://english.khan.co.kr/khan_art_view.html?artid=201208241354087&code=790101 (Last Accessed 29 January 2022).
[265] BUDRAS CORINNA, Facebook muss Pseudonyme auf seiner Plattform dulden, 2022, Available at https://www.faz.net/aktuell/wirtschaft/digitec/facebook-muss-pseudonyme-auf-seiner-plattform-dulden-17756980.html (Last Accessed 29 January 2022).
[266] STOKEL-WALKER, Chris, Twitter's vast metadata haul is a privacy nightmare for users, 2018, Available at https://www.wired.co.uk/article/twitter-metadata-user-privacy (Last Accessed 29 January 2022).
[267] Ibid.

Legal remedies are hindered by the internet's global nature and the respective nation's lack of jurisdiction outside of its borders. A perpetrator outside of a legal authority's sphere of influence faces next to no effective consequences.

Resorting to perceived easy solutions, such as delegating the prosecution to the platforms themselves or blanket-blocking the access to them, leads to numerous unintended side effects. These include but are not limited to digital isolation, a heavy restriction on the freedom of expression and the citizens side-stepping the measures altogether.

Therefore, existing and intended legal remedies on a national legislative basis do not provide a feasible solution to combat fake news and hate speech on a global scale. It might be preferable to spend the resources on extensive education with digital media in conjunction with more steps towards political transparency. Open government initiatives might prove to be a better path towards mitigating the spread of fake news.

The next chapter will offer some insight into these initiatives and analyze their potential benefits.

References Chapter 4

[4-1] UN GENERAL ASSEMBLY, International Convention concerning the Use of Broadcasting in the Cause of Peace (Geneva, 1936) Available at https://www.refworld.org/docid/3b00f0838.html (Last accessed 25 January 2022).

[4-2] KALSNES, BENTE, Fake News, Kristiania University College, 2018, Available at https://doi.org/10.1093/acrefore/9780190228613.013.809 (Last accessed 05 November 2021).

[4-3] LIESCHING, MARC, Das NetzDG in der praktischen Anwendung, 2021, P. 75, Available at https://library.oapen.org/handle/20.500.12657/48794 (Last accessed 20 January 2022).

[4-4] CLAUSSEN, VICTOR, Fighting Hate Speech and Fake News. The Network Enforcement Act (NetzDG) in Germany in the context of European legislation, in: media laws 03/2018, Available at https://www.medialaws.eu/wp-content/uploads/2019/05/6.-Claussen.pdf (Last accessed 24 November 2021).

[4-5] COUZIGOU, IRENE, The French Legislation Against Digital Information Manipulation in Electoral Campaigns: A Scope Limited by Freedom of Expression, in Election Law Journal Vol. 20, No. 1, 2021. p. 112, Available at https://doi.org/10.1089/elj.2021.0001 (Last accessed 29 November 2021).

[4-6] SANZ-CABALLERO, SUSANA, The Principle of Nulla Poena Sine Lege Revisited: The Retrospective Application of Criminal Law in the Eyes of the European Court of Human Rights, in European Journal of International Law, Vol. 28, 2017, P. 788, Available at https://doi.org/10.1093/ejil/chx049 (Last accessed 19 January 2022).

[4-7] KSHETRI, NIR, The Global Cybercrime Industry, Springer, 2010, P. 143.

[4-8] BUSSMANN, KAI-D., Organisationen als Opfer, in BKA Reihe Polizei und Forschung, Viktimologie Deutschland, 2015, P. 395, Available at https://www.bka.de/SharedDocs Downloads/DE/Publikationen/Publikationsreihen/PolizeiUndForschung1_47_1_ViktimisierungsbefragungenInDeutschland.html (Last accessed 30 December 2021).

[4-9] SMITH, R., PERRY, M., Fake News and the Convention on Cybercrime, in Athens Journal of Law-Volume 7, 2021, P. 336, Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3878059 (Last accessed 10 January 2022).

[4-10] KOTOWSKI, MATEUSZ, The Country of Origin Principle and the Applicable Law for Obligations Related to the Benefit of Information Society Services, in Adam Mickiewicz University Law Review, 11, 161-183, Available at https://doi.org/10.14746/ppuam.2020.11.09 (Last accessed 30 October 2021).

[4-11] EUROPEAN PARLAMENT; Directive 2000/31/EC, 2000, Art. 3, Available at http://data.europa.eu/eli/dir/2000/31/oj (Last accessed 28 December 2021).

[4-12] PAUL, RON, War on "Fake News" Part of War on Free Speech, Available at https://mises.org/wire/war-fake-news-part-war-free-speech (Last accessed 20 January 2022).

[4-13] UNITED NATIONS, Strategy and plan of action on hate speech, 2019, p.1f., Available at https://www.un.org/en/genocideprevention/hate-speech-strategy.shtml (Last accessed 30 December 2021).

[4-14] FEDERAL OFFICE OF JUSTICE, German Criminal Code, 2019, § 130, Available at https://www.gesetze-im-internet.de/englisch_stgb/ (Last accessed 25 January 2022).

[4-15] FEDERAL OFFICE OF JUSTICE, Law to improve law enforcement in social networks - Netzwerkdurchsetzungsgesetz – NetzDG, 2017, Available at https://www.gesetze-im-internet.de/netzdg/BJNR335210017.html (Last accessed 13 January 2022).

[4-16] EUROPEAN COURT OF HUMAN RIGHTS, Press Unit, January 2022, p. 1., Available at https://www.echr.coe.int/documents/fs_hate_speech_fra.pdf (Last accessed 13 January 2022).

[4-17] FRENCH SENATE, Fight against hate on the internet, September 2021, Available at https://www.senat.fr/dossier-legislatif/ppl18-645.html (Last accessed 06 January 2022).

[4-18] GERMAN BUNDESTAG, Scientific Services, Regulating hate speech and fake news on social media networks through selected countries, p. 9., 2019, Available at https://www.bundestag.de/resource/blob/662048/190949149266f3df2e27a0f098a53026/WD-10-059-19-pdf-data.pdf (Last accessed 30 December 2021).

[4-19] COUNCIL OF EUROPE, Reports of the Anti-Racism Commission on Estonia, Austria and the Czech Republic, 2015, Available at https://go.coe.int/K94ds (Last accessed 30 December 2021).

[4-20] BRUGGER, WINFRIED, Ban on or protection of hate speech - some observations based on German and American law, 2019, Available at https://journals.tulane.edu/teclf/article/view/1662 (Last accessed 20 January 2022).

[4-21] SUPREME COURT OF THE UNITED STATES, U.S. Reports: Brandenburg v. Ohio, 395 US 444 (1969), Available at https://www.loc.gov/item/usrep395444/ (Last accessed 19 January 2022).

[4-22] UNITED NATIONS, Strategy and plan of action on hate speech, 2019, p.2., Available at https://www.un.org/en/genocideprevention/hate-speech-strategy.shtml (Last accessed 30 December 2021).

[4-23] EUROPEAN PARLIAMENT, Policy Department - Citizens rights and constitutional affairs, p. 126, 2015, Available at https://www.europarl.europa.eu/RegData/etudes/STUD/2015/536460/IPOL_STU(2015)536460_EN.pdf (Last accessed 17 January 2022).

[4-24] EUROPEAN COMMISSION, The EU Code of conduct on countering illegal hate speech online, 2019, Available at, https://ec.europa.eu/info/policies/justice-and-fundamental-rights/combatting-discrimination/racism-and-xenophobia/eu-code-conduct-countering-illegal-hate-speech-online_en (Last accessed 14 January 2022).

[4-25] GERMAN BUNDESTAG, Regulating hate speech and fake news on social media networks through selected countries, p. 11, 2019, Available at

https://www.bundestag.de/resource/blob/662048/190949149266f3df2e27a0f098a53026/WD-10-059-19-pdf-data.pdf (Last accessed 30 December 2021).

[4-26] EUROPEAN PARLIAMENT, Hate speech and hate crime in the EU and the evaluation of online content regulation approaches, p. 53, 2020, Available at https://www.europarl.europa.eu/RegData/etudes/STUD/2020/655135/IPOL_STU%282020%29655135_EN.pdf (Last accessed 10 January 2022).

[4-27] EUROPEAN UNION, Consolidated version of the Treaty on the Functioning of the European Union – Part Three, Article 83 (ex Article 31 TEU), 2008, Available at https://eur-lex.europa.eu/eli/treaty/tfeu_2008/art_83/oj (Last accessed 07 January 2022).

[4-28] COUNCIL OF EUROPE, European Commission against Racism and Intolerance ECRI, Available at https://rm.coe.int/leaflet-ecri-2019/168094b101 (Last accessed 08 January 2022).

[4-29] EUROPEAN UNIVERSITY INSTITUTE, Handbook on Techniques of Judicial Interaction in the Application of the EU Charter: Freedom of expression and countering hate speech, pp. 21, 53, Available at https://cjc.eui.eu/wp-content/uploads/2020/05/eNACT_Handbook_Freedom-of-expression-compresso.pdf (Last accessed 08 January 2022).

[4-30] FEDERAL OFFICE OF JUSTICE, German Criminal Code, 2019, § 5, Available at https://www.gesetze-im-internet.de/englisch_stgb/ (Last accessed 08 January 2022).

[4-31] FEDERAL CONSTITUTIONAL COURT GERMANY, Decision of January 15, 1958 –1 BvR 400/51, Available at https://www.bundesverfassungsgericht.de/SharedDocs/Entscheidungen/DE/1958/01/rs19580115_1bvr040051.html (Last accessed 09 January 2022).

[4-32] BERLIN STATE CENTER FOR POLITICAL EDUCATION, Hate speech and fake news - questions and answers, p. 9, 2018, Available at https://www.amadeu-antonio-stiftung.de/w/files/pdfs/hate_speech_fake_news.pdf (Last accessed 20 January 2022).

[4-33] EUROPEAN PARLIAMENT, The impact of algorithms for online content filtering or moderation, p.44, 2020, Available at https://www.europarl.europa.eu/RegData/etudes/STUD/2020/657101/IPOL_STU(2020)657101_EN.pdf (Last accessed 26 January 2022).

[4-34] HOUSE OF LORDS LIBRARY, Freedom of speech: Challenges and the role of public, private and civil society sectors in upholding rights, 2021, Available at https://lordslibrary.parliament.uk/freedom-of-speech-challenges-and-the-role-of-public-private-and-civil-society-sectors-in-upholding-rights/ (Last accessed 26 January 2022).

[4-35] KORNMEIER, CLAUDIA, Gericht mit Kehrtwende im Fall Künast, 2020, Available at https://www.tagesschau.de/inland/kuenast-beleidigung-103.html (Last accessed 27 January 2022).

[4-36] KÖVER, CHRIS, Interview zum Fall Künast - Dieses Urteil ist ein gutes Zeichen, 2020, Available at https://netzpolitik.org/2020/dieses-urteil-ist-ein-gutes-zeichen/ (Last accessed 27 January 2022).

[4-37] HELDT, A. P., Intelligente Upload-Filter: Bedrohung für die Meinungsfreiheit, 2018, Available at https://www.ssoar.info/ssoar/bitstream/handle/document/57609/ssoar-2018-heldt-Intelligente_Upload-Filter_Bedrohung_fur_die.pdf (Last accessed 28 January 2022)

[4-38] DREXL JOSEF, Bedrohung der Meinungsvielfalt durch Algorithmen, ZUM 2017, 529

[4-39] SPOERRI, THOMAS, On Upload-Filters and other Competitive Advantages for Big Tech Companies under Article 17 of the Directive on Copyright in the Digital Single Market, 10 (2019) JIPITEC 173 para 1., Available at https://www.jipitec.eu/issues/jipitec-10-2-2019/4914 (Last accessed 27 January 2022).

[4-40] CACCIOTTOLO, MARIO, The Streisand Effect: When censorship backfires, 2012, Available at https://www.bbc.com/news/uk-18458567 (Last accessed 28 January 2022).

[4-41] SCOTT, JENNIFER, Can Online Safety Bill tackle social media abuse of MPs?, 2021, Available at https://www.bbc.com/news/uk-politics-58958244 (Last accessed 28 January 2022).

[4-42] NEUERER, DIETMAR, Digitalverbände und Datenschützer warnen vor Klarnamenpflicht, 2019, Available at https://www.handelsblatt.com/politik/deutschland/cdu-vorstoss-digitalverbaende-und-datenschuetzer-warnen-vor-klarnamenpflicht/24446344.html?ticket=ST-3333381-QQCgAfzLxmTFvGCebieC-ap3 (Last accessed 29 January 2022).

[4-43] SCHWANDER, TIMO, Das digitale Vermummungsverbot – eine irreführende Analogie, in ZRP 52/7 (2019), p. 207-208.

[4-44] HAARKÖTTER, HEKTOR, Anonymität im partizipativen Journalismus, in: Zöllner, Anonymität und Transparenz in der digitalen Gesellschaft, 2015, p.133-149.

[4-45] CARAGLIANO, DAVID, Real Names and Responsible Speech: The Cases of South Korea, China, and Facebook, 2013, Available at https://www.yalejournal.org/publications/real-names-and-responsible-speech-the-cases-of-south-korea-china-and-facebook (Last accessed 29 January 2022).

[4-46] THE CHOSUNILBO, Real-Name Online Registration to Be Scrapped, 2011, available at http://english.chosun.com/site/data/html_dir/2011/12/30/2011123001526.html (Last accessed 29 January 2022).

[4-47] KIM, KATE JEE-HYUNG, Lessons Learned from South Korea's Real-Name Policy, 2012, available at http://www.koreaittimes.com/news/articleView.html?idxno=19361 (Last accessed 29 January 2022).

[4-48] SALUZZO, STEFANO, EU Law and Extradition Agreements of Member States: The Petruhhin Case, 2017, Available at https://www.europeanpapers.eu/en/europeanforum/eu-law-and-extradition-agreement-of-member-states-the-petruhhin-case (Last accessed 29 January 2022).

[4-49] VOSOUGHI, SOROUSH, The spread of true and false news online, 2018, Available at https://www.science.org/doi/10.1126/science.aap9559 (Last accessed 29 January 2022).