SOCIAL MEDIA PLATFORMS INTERMEDIARY LIABILITY FOR USER CONTENT: GERMANY AND THE US – COMPARATIVE ANALYSIS

by Mikheil Rekhviashvili

LLM Final Thesis

SUPERVISOR: Professor Tommaso Soave

Central European University Private University

Quellenstrasse 51-55, 1100 Vienna

Austria

June, 2022

Abstract

Social media platforms (SMPs) are mediums for users to exercise their public rights, but can also be used for unlawful and/or unhealthy purposes such as terrorist recruitment, pornography dissemination, election interference, etc. The question of whether and to what extent SMPs should be liable for their users' activities is more relevant than ever, and states' regulatory responses differ. While the US has opted for a more hands-off approach, Germany enacted a regulation imposing strict content moderation obligations on SMPs.

Drawing from legal and economic sources, this thesis argues that content moderation and the creation of "healthy" content are valuable to all parties concerned – states, SMPs, and users. Moving from this premise, it compares the US Communications Decency Act (CDA) and Germany's Network Enforcement Act (NetzDG) to assess which regime better achieves the goal of promoting "healthy" SMP environments.

My conclusion is that the NetzDG is the way to go. First, under § 230 of the CDA, SMPs' content moderation practices are sometimes skewed. Second, largely privatized content moderation lacks elements of due process while users are exercising their public rights. Third, it lacks transparency. On the other hand, under the NetzDG unlawful content is to be taken down in 24 hours or 7 days. Further, lawful content is to be reinstated; users get to appeal the decisions of SMPs and receive reasoning. Additionally, the NetzDG does not cause collateral censorship. However, the NetzDG still requires improvement – e.g. differentiating between the types of unlawful content and subsequent moderation timeframes.

Table of Contents

Introduction

Relevance of the topic

We live in a time of alternative reality – social media platforms ("SMP") which are created by different and competing commercial entities. Social networks are no longer just a space to connect with other people. On 28th of October Facebook introduced meta[1] - a hyperreal 3D social media space which will not only include user content, but is also intended to be a virtual universe with simulated people and facilities. Major SMPs such as Google, Facebook and Twitter *are* the internet for the majority of the world's population, especially for the emerging markets.[2] Therefore, SMPs are the primary spaces for people to digitally self-realize. Together with many advantages, the trend of shifting almost every aspect of human life to SMPs is characterised by numerous risks both for customers and the business itself. Higher risk of dissemination of unlawful content, fake news, hate speech, violence, terrorism and defamation is the inevitable corollary of the increase of user numbers and content on SMPs. Undoubtedly, this list is not exhaustive.[3]

It follows, therefore, that the law should play a pivotal role and catch up with the development of digital life. Whenever there is a big gap between law and modern technology, states risk facing undesirable consequences. In particular, the delicate question arises of whether and to what extent SMPs should be subject to intermediary legal liability for the harmful content posted by their users. As user content becomes more and more sensitive, personal and disseminated; and SMPs reach such a huge scale, current laws regarding

[1] 'Introducing Meta: A Social Technology Company' <https://about.fb.com/news/2021/10/facebook-company-is-now-meta/> accessed 11 June 2022
[2] Taylor Owen, 'The Case for Platform Governance' (2019) Center for International Governance Innovation 1, 3
[3] Thiago Dias Oliva, 'Content Moderation Technologies: Applying Human Rights Standards to Protect Freedom of Expression' (2020) Volume 20 Issue 4 Human Rights Law Review 607, 607

intermediary liability of SMPs are becoming obsolete. Clearly, a user uploading illicit material on SMPs is personally liable. However, the extent of SMPs intermediary liability is not settled while it is at the same time more relevant than ever. The world is watching closely the impending acquisition of Twitter by Elon Musk.[4] Should he be the only person who determines the discourse on Twitter and its intermediary liability? Regulators are trying to find a balance between the interests of SMPs and customers. Subsequently, some of them are trying to approach these platforms with heavily regulated laws such as the NetzDG[5] whereas others have opted for loose ones – § 230 of the CDA.[6]

Central research question

The present thesis compares two opposite approaches – that of the United States' and Germany – to SMP intermediary liability, the modes and purposes of content moderation, and the interests of the parties concerned. The goal is to determine how a pertinent legislative approach to these issues should look. Approach can be labelled as pertinent if the interests of all parties concerned – the public, SMP businesses and users - are taken into consideration and well-balanced. It is a general perception that less legal liability is better for business, and that companies do not like to be the subject of legal limitations. But is this true for SMPs in today's world? There is no straightforward answer. No liability means less action and responsibility from moderators, potentially leading to a toxic and uncontrolled digital environment. In turn, this might prove detrimental or profitable for the business itself depending on the circumstances. On the one hand, customers may shift to healthier platforms, while states will have incentives to step in and adopt more severe measures to mitigate the consequences, thus imposing costly and unexpected burdens on businesses. On the other hand, heavy regulations

[4] Clare Duffy, 'Elon Musk to Buy Twitter In $44 Billion Deal' (2022) CNN Business <https://edition.cnn.com/2022/04/25/tech/elon-musk-twitter-sale-agreement/index.html> accessed 11 June 2022
[5] Network Enforcement Act 2017 (NetzDG)
[6] Section 230 of the Communications Act of 1934 (§ 230 of the CDA)

might lead to a censored social media environment. This affects not only business but also the users and their freedom of expression. Business will typically try to pass the burden onto the customers in order to avoid legal liability. Additionally, social media laws have repercussions on the price of stocks, since almost all the major SMP companies are publicly listed entities. Consequently, the existence of heavy or light regulations are indicative parameters for investors.

Throughout this thesis I put a particular emphasis on two jurisdictions – the United States and Germany. The reason for choosing the abovementioned countries is that they are both major global markets. The United States is the cradle of SMPs and one of the most advanced countries in the world in this regard. Further, US relevant law – Section 230 of the Communications Decency Act of 1996 (CDA) – is generally deemed to be out of date, and there have been several attempts to repeal[7] and amend[8] the law. Germany, for its part, is the major economy of the European Union. I intentionally do not focus on the EU at large, because the EU Directive[9] is quite generic. Germany also seemed most appropriate because it enacted legislation in 2017 - the Network Enforcement Act ("NetzDG") - which is quite restrictive and burdensome as compared to the relatively lax CDA. The NetzDG has been a subject to a lot of criticism and commendation, making the German and the US regulations on SMPs stand out as two polar models in the West, which may influence other nations' efforts in the future. I believe that the contrast between these two leading jurisdictions and their comparison leads to compelling observations and viable solutions with regard to the proper regulation of SMPs' intermediary liability. It is worth mentioning that these countries are different not only in this

[7] eg S.2972 - A bill to repeal section 230 of the Communications Act of 1934 <https://www.congress.gov/bill/117th-congress/senate-bill/2972> accessed 11 June 2022, S.3538 - EARN IT Act of 2022 <https://www.congress.gov/bill/117th-congress/senate-bill/3538> accessed 11 June 2022

[8] eg S.4066 - PACT Act <https://www.congress.gov/bill/116th-congress/senate-bill/4066> accessed 11 June 2022, H.R.4027 - Stop the Censorship Act <https://www.congress.gov/bill/116th-congress/house-bill/4027#:~:text=Introduced%20in%20House%20(07%2F25%2F2019)&text=This%20bill%20limits%20a%20social,unlawful%20rather%20than%20merely%20objectionable> accessed 11 June 2022

[9] Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market (E-Commerce Directive)

regard but in the general framework of their legal systems, as the United States is a common law country and Germany a civil law one. Lastly, this paper might also prove to be useful for legislative purposes in my homeland – Georgia – which does not currently have any relevant law concerning SMPs' content and its subsequent intermediary liability.

Research methodology

The methodology of the research is a comparative analysis of the US and Germany's applicable jurisprudence and laws, namely the US Communications Decency Act of 1996 and the Network Enforcement Act of 2017. Intermediary liability of SMPs is a broad category, so the present research will focus on the relevant norms of the NetzDG and the CDA which shape the intermediary liability of SMPs in the abovementioned nations. In addition to the legal documents, the research touches upon the internal processes, standards, bodies, guidelines, announcements and public hearings of SMPs and public authorities. While these are "unofficial" sources, they form an important part of the regulatory regime of SMPs' intermediary liability and can therefore not be discounted. The comparison of German and US legal regimes is followed by an analysis, and a subsequent conclusion outlining what a proper legislative approach should look like and what elements should it include.

Road map

The thesis is divided into 3 chapters.

Chapter 1 is devoted to the general concept of SMPs' intermediary liability for user content. It defines and clarifies the scope of intermediary liability and indicates the relevant field of laws concerning user content. The first chapter briefly describes the different approaches for regulating SMPs' content and their subsequent intermediary liability, and illustrates the internal bodies and regulations of SMPs that are not purely positive law. In the

final section of the chapter, I discuss the primary purposes of SMPs' intermediary liability regulation and the interests of the parties concerned. The chapter concludes with the notion that the interests of the state and SMPs are not far from each other in maintaining "healthy" content.

Chapter 2 goes into the depth of the relevant provisions of § 230 of the CDA and the NetzDG which shape intermediary liability of SMPs. In the first part of the chapter, I discuss the applicability of those regulations to SMPs. The second section is devoted to substance, namely what kinds of liability and immunity are granted or deprived under the norms and respective jurisprudence. The chapter concludes by briefly comparing US and German legal regimes delineating general characteristics and differences, and their respective pros and cons.

After having established the primary purpose of SMP's regulation in chapter 1 and analyzing two major different legal regimes in chapter 2, chapter 3 explores which approach is the superior methodology – German or US. In this chapter I examine how a sensible legislative approach should be engineered and the relevant issues that should be considered.

The chapters are followed by a brief conclusion which sums up the research findings.

Chapter 1. SMP's intermediary liability

1.1 Definition

User content responsibility in the context of SMP's is quite a broad term. It refers to the legal relationships which might arise from the third-party user content posted and shared on SMPs. It contains not only positive law, but also "platform-made law" which is created by the private actors themselves. ""Social media law" is not a monolith. It is, instead, an amalgam of constitutional law, employment law, criminal law, administrative law, intellectual property law, tort law, and even the rules of discovery as applied to the ubiquitous and relatively recent presence and use of online and mobile platforms for sharing and creating content."[10] Given their wide reach, social media laws have significant implications for business and the public. Relevant laws set forth the responsibility of SMPs for user content, and therefore determine the behavior of SMPs. Users, without a doubt, are individually liable for their content. However, the scope of liability of the platforms on which they communicate is still under debate, and more relevant than ever. Throughout this thesis, accent is put on intermediary liability of SMPs for the content provided by other parties – namely, their users.

In terms of intermediary liability, several parallels could be drawn - including with Internet Service Providers, telecommunication companies, libraries, bookstores, newspapers, magazines, and the like. Under US law, Internet Service Providers are seen as mere conduits - therefore, they are exempted from intermediary liability unless they have any editorial or participatory function in connection with the dissemination of defamatory (unlawful) content.[11]

[10] Christopher Escobedo Hart, 'Social Media Law: Significant Developments' (2016) 72 Bus Law 235, 235

[11] See eg *Lunney v Prodigy Servs Co* [1998] 250 A.D.2d 230 (Supreme Court of New York, Appellate Division, Second Department).

The same standard applies to telephone companies.[12] And even if both these types of companies are considered publishers because of exercising editorial or participatory functions, they are protected from liability by common-law qualified privilege unless they know that the transmitted message is defamatory (unlawful).[13] Libraries and bookstores are secondary publishers: distributors which are sometimes distinguished from mere conduits.[14] Lastly, newspapers and magazines are primary publishers as they exercise editorial function and control. Traditionally, publishers are held liable for the published content in common law.[15] On the other hand, distributors are not liable for the third-party content unless they have knowledge of the unlawful content.

So which one is the most suitable for SMPs? That is the one of the questions of this thesis. Unlike newspapers and primary publishers, SMPs do not control and pre-edit the content to be published by their users. However, they still have the best position and tools to effectively moderate content, be it *ex ante* or *ex post*.[16] Applicable laws such as § 230 of the CDA concern not only SMPs and users but also investors and the investment environment. As Facebook CEO Mark Zuckerberg testified before the United States Senate Committee on Commerce, Science, and Transportation, without § 230 of the CDA "platforms would likely censor more content to avoid legal risk and would be less likely to invest in technologies that enable people to express themselves in new ways."[17] The extent and far-reaching nature of state regulation or non-

[12] See eg *Anderson v New York Tel Co* [1974] 35 N.Y.2d 746 (Court of Appeals of New York).
[13] ibid
[14] Hilary Young and Emily Laidlaw, 'Internet Intermediary Liability in Defamation: Proposals for Statutory Reform' (2017) 1, 5
[15] Agnieszka McPeak, 'Platform Immunity Redefined' (2021) 62 Wm. & Mary L. Rev 1557, 1566
[16] Here under *ex ante* I mean the process of human and automated moderation which occurs before a post is published on SMP and respectively, *ex post* here refers to the moderation after a post is published
[17] Mark Zuckerberg, 'Testimony of Mark Zuckerberg Facebook, Inc' (2020) Hearing Before the United States Senate Committee on Commerce, Science, and Transportation 1, 1 <https://www.commerce.senate.gov/services/files/E017B34E-F87F-4127-88A7-2C32B6BC3810> accessed 11 June 2022

regulation requires a very subtle touch from both legislative branch and SMPs in order to strike a fair balance between the interests of business on one hand and customers on the other.

1.2 US and German Approaches

Speaking of a legislative approach from a state to user content regulation, I put emphasis on two approaches – the self-regulatory US style and the more stringent, government-driven German style. In the United States, SMPs are, under § 230 of the CDA, granted immunity from the liability of certain actions of their users. According to the CDA, interactive computer service providers are not "the publishers or speakers of any information provided by another information content provider".[18] Therefore, in the US, user content liability regulation and platform moderation is delegated to private actors. Consequently, in the US model, platforms are not obligated to take illegal content down. They are also not liable for removing the content which they do not want to appear.[19] However, although § 230 immunity applies to most torts concerning publishing including defamation, it does not influence criminal law, intellectual property law, consistent state law, communications privacy law and sex trafficking law.[20]

On the other hand, there is the NetzDG which is quite burdensome for SMPs. It obliges SMPs to take illegal content down within set timelines. Additionally, it imposes on the providers of social network reporting obligations and sets forth the penalties in case of non-compliance. Under the NetzDG, decisions regarding the lawfulness of content are made by SMPs themselves. Therefore, the German state delegates a quasi-judicial duty to the SMP, namely the assessment of the lawfulness of content, which is a quite complex and vague term.[21]

[18] Maayan Perel, 'Enjoining Non-Liable Platforms' (2020) 34 Harv. J. Law & Tec 1, 3
[19] Patrick Zurth, 'The German NetzDG as Role Model or Cautionary Tale? Implications for the Debate on Social Media Liability' (2021) 31 Fordham Intell Prop Media & Ent LJ 1084,1093
[20] § 230 (e) of the CDA
[21] Amélie Heldt, 'Let's Meet Halfway: Sharing New Responsibilities in a Digital Age' (2019) 9 Journal of Information Policy 336, 342

However, SMPs' intermediary liability for the user content is not set forth in the NetzDG. In this regard relevant provisions can be found in the Telemedia Act articles 7-10, which correspond to the E-commerce Directive articles 12-14. In Germany, liability for the user content comes with the knowledge under Störerhaftung doctrine.[22] This corresponds to the distributor liability in the United States, according to which an intermediary is liable when it has acquired knowledge or should have known about unlawful content. Therefore, the NetzDG can be seen as a set of rules which prescribes more concrete actions of enforcement, reporting obligations and penalties in case of non-compliance.

1.3 Beyond official legislation

In addition to positive national legislative instruments, such as the § 230 of the CDA and the NetzDG, soft laws and self-regulation bear mention in this analysis. By self-regulation, I mean the internal rules and standards of SMPs, such as Facebook's Community Standards.[23] In some cases, self-regulations have a more far-reaching impact on society than state laws in that concrete issue. SMPs have reached such a scale that the legal market demanded "privatization of law and judiciary." A result of this was the creation of the Facebook Oversight Board which adjudicates complaints with regard to the user content according to its values, community standards and international law in its decisions[24]. This creates a type of private legal order within the company - nonetheless its decisions are "binding" upon the users from different states and legal regimes. Since the § 230 of the CDA was enacted many years previously and it gives the private actors space for self-regulation, the creation of the Oversight

[22] See Sandra Schmitz, Christian M. Berndt, 'The German Act on Improving Law Enforcement on Social Networks (NetzDG): A Blunt Sword?' (2018) 1, 9
[23] Facebook Community Standards <https://transparency.fb.com/policies/community-standards/> accessed 11 June 2022
[24] See eg *Case Decision 2020-006-FB-FBR* [2021] (The Oversight Board) and *Case decision 2021-016-FB-FBR* [2022] (The Oversight Board).

Board can be seen as way "to stave off actual government regulation".[25] Further, the Oversight Board could well be seen as a scapegoat which will take the blame for controversial decisions instead of Facebook.[26] In this particular case business might be trying to fill the § 230 of the CDA's legislative intentional vacuum by creating its own "legislation" and order. The Oversight Board itself can be seen as a body akin to the Supreme Court of the United States in the sense that it selects cases to be adjudicated and its decisions are final within private mechanisms.[27] However, Facebook and other SMPs still have to adjust to national legislations such as the NetzDG in Germany. Whereas § 230 of the CDA gives freedom to SMPs to self-regulate and create bodies, the NetzDG "forces" them to adjudicate whether the content is obviously illegal or just illegal according to German criminal law.[28] Be that as it may, "the advent of the internet environment has prompted parallel consolidation of power in the hands of private intermediaries, demonstrating an increasing tendency towards the privatization of traditionally public functions."[29] This is true for the US model under § 230 of the CDA, but also not far from truth for German model as the NetzDG compels SMPs to make criminal law assessments in restricted timelines.

1.4 History and reasons for adopting the CDA and the NetzDG

There is a significant time gap between the enacting of § 230 of the CDA and the subsequent enacting of NetzDG 21 years later. Generally, legislation is the answer to contemporary challenges, and these challenges shift in form and function as time goes on. It is important to take a look at the situation of defamation, free speech and technological

[25] Kate Klonick, 'The Facebook Oversight Board: Creating an Independent Institution to Adjudicate Online Free Expression' (2020) 129 Yale L.J. 2418, 2488
[26] ibid 2426.
[27] ibid 2476.
[28] NetzDG, s 3.2 and 3.3.
[29] Luca Belli, Pedro Augusto Francisco and Nicolo Zingales, 'Law of the Land or Law of the Platform? Beware of the Privatisation of Regulation and Police' (2017) 41, 59

development at the time § 230 of the CDA was proposed. Before enacting § 230 of the CDA, the Supreme Court of New York rendered its decision in *Stratton Oakmont v. Prodigy Servs. Co.*, where the court equated online service providers to publishers.[30] By contrast, in *Cubby, Inc. v. Compuserve, Inc*, the District Court for the Southern District of New York reasoned that the defendant, a computerized database owner, was not a publisher but a distributor.[31] These two cases were the main drivers for the adoption of § 230 of the CDA[32], as the Congress wanted to provide relief for internet service providers for third party user content. The legislation was a reflection of the United States' approach to free speech. The § 230 of the CDA's primary purpose was to promote and foster free speech and technology. However, it must be taken into account that at this time no platforms with the influence and scale of the likes of Facebook, and Twitter existed - platforms which are used by millions of people globally, with information disseminated throughout the world in seconds. Therefore, the advancement of SMPs has not been without accompanying risks, because of the potential for this nearly instantaneous spread of illicit information.

By comparison, the NetzDG, went into force in 2018, imposing strict rules and liabilities on SMPs. In this regard, the NetzDG is almost the opposite approach to § 230 of the CDA. In contrast to the situation of § 230 of the CDA's enactment in 1996, this time the goal of the legislation was to tackle the issue of online hate speech and fake news.[33] Proliferation of illegal content and hate speech on SMPs in Germany is mostly ascribed to the 2015 European migrant crisis in Germany, when more than 1 million Syrian refugees were accepted into the

[30] *Stratton Oakmont v. Prodigy Servs.* Co [1995] 1995 N.Y. Misc. LEXIS 229 (Supreme Court of New York, Nassau County).
[31] *Cubby, Inc. v. Compuserve, Inc.* [1991] 776 F. Supp. 135 (United States District Court for the Southern District of New York).
[32] Vanessa S. Browne-Barbour, 'Losing Their License to Libel: Revisiting § 230 Immunity' (2015) 30(2) Berkeley Technology Law Journal 1505, 1519
[33] Sebastian Schwiddessen, Birgit Clark, Thomas Defaux, John Groom, 'Germany's Network Enforcement Act - closing the net on fake news?' (2018) 40(8) E.I.P.R. 539, 539

country.[34] This was followed by the spread of hate speech against refugees and the government, and subsequent investigations of unlawful online activities.[35] Consequently, the German legislator provided quite burdensome and heavy legislation to counter the problem.

Therefore, on the one hand we have a "private legislation" and "private supreme court" of Facebook which stems from § 230 of the CDA and on the other, heavy administrative state-regulation – the NetzDG. The main questions that arise then are: What is the direction of these laws? What is the primary aim – healthy content? Is it primary for the business or for the state?

1.5 Is healthy content the primary purpose? And whose interest does it serve?

In order to ascertain which legal approach is most suitable, the primary purpose of the regulations of SMPs in terms of user content needs to be determined. When § 230 of the CDA was enacted in 1996, the legislator sought to exempt intermediaries from liability for user content. Distilled to its essence, the goal was to make it easier for interactive computer service providers to conduct business and promote free speech. In the case of § 230 of the CDA, the interests of the state and business were largely aligned, since § 230 of the CDA provided complete freedom and exemption from liability for interactive computer service providers. Naturally, Big Tech companies such as Facebook, Google and Twitter still prefer § 230 of the CDA as it is today. Clear evidence of this comes from the hearing on "Does Section 230's Sweeping Immunity Enable Big Tech Bad Behavior?" before the US Senate Committee on Commerce, Science, and Transportation. According to the hearing[36] and statements[37], these

[34] William Echikson and Olivia Knodt, 'Germany's NetzDG: A Key Test for Combatting Online Hate' (2018) No. 2018/09 CEPS Research Reports 1, 2

[35] Alexander Ritzmann, Hans-Jakob Schindle and Marco Macori, 'NetzDG 2.0 – Recommendations for the amendment of the Network Enforcement Act (NetzDG) and Investigation into the actual blocking and removal processes of YouTube, Facebook and Instagram' (2020) CEP Policy Paper 1, 1

[36] *Does Section 230's Sweeping Immunity Enable Big Tech Bad Behavior?* (2020) <https://www.commerce.senate.gov/2020/10/does-section-230-s-sweeping-immunity-enable-big-tech-bad-behavior> accessed 11 June 2022

[37] Jack Dorsey, 'Testimony of Jack Dorsey Chief Executive Officer Twitter, Inc.' *Does Section 230's Sweeping Immunity Enable Big Tech Bad Behavior?* (2020) <https://www.commerce.senate.gov/services/files/7A232503-B194-4865-A86B-708465B2E5E2> accessed 11 June 2022, Sundar Pichai, 'Written Testimony of Sundar

companies see § 230 of the CDA as a cornerstone of business viability for SMPs and they support the preservation of this. However - at least for the states and public - priorities have shifted since 1996. Today, in combination with state regulations, SMPs are also trying to self-regulate. Yet is arguable whether the interests of SMPs and states are still mutually agreeable in the contemporary climate.

Like any other business, SMPs are primarily motivated by economic incentives and profits. Social media is an attention economy which basically means that it generates income from user engagement.[38] Accordingly, in order to raise revenue and profit, SMPs are trying to engage as many users as possible. In this regard, it is argued that the harmful posts which are causing radicalization and negative emotions such as fear, shock, surprise and disgust generate a large number of shares and engagement on SMPs and, consequently are responsible for significant profit.[39] Similarly, artificial intelligence of the major SMPs "are competing for views by promoting completely false information - a race to the bottom by emphasizing engagement."[40] Therefore, it is doubtful that it is in SMPs' economic interests to be regulated externally and create purely "healthy" content. However, besides economic incentives, SMPs are motivated by long-term viability, public image and good citizenship.[41] Nevertheless, I believe that ultimately, all the mentioned motives are connected to and derived from economic and financial well-being. It can be argued though, that stemming from the same economic incentive, sometimes it is beneficial for business to say no to today's revenues in order to avoid

---

Pichai, Chief Executive Officer, Alphabet Inc.' *Does Section 230's Sweeping Immunity Enable Big Tech Bad Behavior?* (2020) <https://www.commerce.senate.gov/services/files/894758B6-F55E-471D-904F-480FCD9E4D98> accessed 11 June 2022, Mark Zuckerberg, 'Testimony of Mark Zuckerberg Facebook, Inc.' *Does Section 230's Sweeping Immunity Enable Big Tech Bad Behavior?* (2020) <https://www.commerce.senate.gov/services/files/E017B34E-F87F-4127-88A7-2C32B6BC3810> accessed 11 June 2022

[38] Owen (n2) 3.

[39] Nina I. Brown, 'Regulatory Goldilocks: Finding the Just and Right Fit for Content Moderation on Social Platforms' (2021) 8 Tex. A&M L. Rev, 451, 482 citing Martin Jones, 'Emotional Engagement Is the Key to Viral Content Marketing' Cox Blue <https://perma.cc/54DH-DWLM?view-mode=server-side&type=image> accessed 11 June 2022

[40] Owen (n2) 4.

[41] Frank Fagan, 'Systemic Social Media Regulation' (2018) 16 Duke L & Tech Rev 393, 396

tomorrow's heavy costs, such as burdensome regulations from the state. Prominent SMPs such as Facebook and Twitter have started shifting towards creating a healthier environment.[42] As Twitter CEO Jack Dorsey put it "people from around the world come together on Twitter in an open and free exchange of ideas. We want to make sure conversations on Twitter are healthy and that people feel safe to express their points of view".[43] In contrast to 1996, when unrestricted speech was advocated, nowadays emphasis is more on "healthy speech" which resembles the European approach.[44] The abovementioned strategy of avoidance of future costs might be the reason for this tendency of SMPs strategic shift to creating "healthy" content. A further risk is that if the content is abusive and contains illegal material, part of the customer base will become dissatisfied with the service and shift to other platforms.

Translated into real world terms, however, the risk of loss of user base is not that threatening in case of large SMPs such as Facebook and Google. Although it has not been yet confirmed by the US courts, it would not be an exaggeration to say that some of the Big Tech players, especially Google, are close to monopolies. For instance, in the online advertising market Facebook and Google account for half of the US market sales.[45] Consequently, they have substantial pricing power in that area.[46] Unlike the business model of most industries, in the case of SMPs, "marginal cost of serving additional consumers is essentially zero".[47] This means that SMPs can add as many users as they want without additional production or any associated cost. SMPs such as Facebook, YouTube and Twitter have capitalised on this by

[42] Tim Wu, 'Will Artificial Intelligence Eat the Law? The Rise of Hybrid Social-Ordering Systems' (2019) Vol. 119 No. 7 Columbia Law Review 2001, 2009 citing eg Facebook Community Standards <https://perma.cc/D27N-XJEY> accessed 11 June 2022, YouTube Help Hate Speech Policy <https://perma.cc/AZD2-VH4V> accessed 11 June 2022, and Facebook Community Standards Objectionable Content <https://perma.cc/9TMH-R2HG> accessed 11 June 2022

[43] Dorsey (n37) 1.

[44] Zurth (n19) 1098 citing Wu (n42) 2009-10.

[45] Marc Jarsulic, 'Addressing the Competitive Harms of Opaque Online Surveillance and Recommendation Algorithms' (2022) The Antitrust Bulletin 100, 101

[46] ibid 108.

[47] Robert H. Frank, 'The Economic Case for Regulating Social Media' The New York Times <https://www.nytimes.com/2021/02/11/business/social-media-facebook-regulation.html> accessed 11 June 2022

developing a design which psychologically encourages users' addiction.[48] "The technologies we use have turned into compulsions, if not full-fledged addictions. It's the impulse to check a message notification. It's the pull to visit YouTube, Facebook, or Twitter for just a few minutes, only to find yourself still tapping and scrolling an hour later. It's the urge you likely feel throughout your day but hardly notice."[49] All these economic and psychological factors of major SMPs' business model render it very difficult for users to leave the platform or to shift to another one.

Undoubtedly though, there are some levels of toxicity where it becomes intolerable for users to stay on the platform and which may motivate many users and advertisers to leave it.[50] If the content is sufficiently abusive, the risk of losing customers might be realized. According to the Pew Research Center survey, 27% of 4,248 US adults abstained from posting online after witnessing the harassment of others and 13% ceased using the service.[51] Other pieces of research are available which corroborate these results.[52] This again indicates that maintaining "healthy" content is in the interests of the business as well, in order to ensure maximum engagement. In summation, although strict content moderation and a high number of content take-downs is not desirable for SMPs, synergies might be found between state and business interest.[53]

Since the major SMPs have reached scales previously thought unimaginable, states cannot afford the proliferation of illegal content, especially when it comes to terrorism or imminent threat. In this regard, their primary focus is to compel SMPs to ensure a healthy environment. As discussed above, this objective - with little conflict of interest - is also primary

[48] Not with the exact medical sense of this word
[49] Nir Eyal, *Hooked: How to Build Habit-Forming Products* (Portfolio Penguin 2014)
[50] Perel (n18) 25.
[51] Zurth (n19) 1131 citing Maeve Duggan, 'Online Harassment 2017' (2017) Pew Research Center <https://www.pewresearch.org/internet/2017/07/11/online-harassment-2017/> < https://perma.cc/H935-DKKE> and Danielle Keats Citron and Benjamin Wittes, 'The Internet Will Not Break: Denying Bad Samaritans § 230 Immunity' (2017) 86 FORDHAM L. REV. 401, 410 420
[52] See ibid 1132.
[53] Heldt (n21) 337.

for SMPs. Basically, the financial model of an SMP is foundational on the number of users and their engagement. Therefore, SMPs primary purpose is to attract and retain as many users as possible. This objective could not effectively be attained in an environment where a user is constantly anticipating harassment or abuse. "Take down too much content and you lose not only the opportunity for interaction, but also the potential trust of users. Likewise, keeping up all content on a site risks making users uncomfortable and losing page views and revenue".[54] Therefore, content moderation is not only valuable to the state and the public in general, but also to SMPs themselves. This might be one of the occasions where the economic incentives of the business and the interests of state and public are aligned. Although all the interested parties want to play the same game and strive for "healthy" content, they might not agree on the rules of the game. The devil is in the details and these details, once generalized, create the character of the legal approach - be it be self-regulation or external laws. Usually, business is against heavy regulation because of their association with difficulties and costs of compliance. However, the main question here is - is this one of those instances? In order to find the most suitable method for achieving "healthiness" of SMP's content, I look into two opposing and characteristic approaches, namely § 230 of the CDA and the NetzDG, and from this comparison derive the most suitable approach.

The primary determination to be made here is what is a "healthy" platform? What level should it reach to be deemed "unhealthy"? And which is the proper regulatory approach to attain "healthiness"? The US approach, the German approach, or neither?

[54] Kate Klonick, 'The New Governors: The People, Rules, and Processes Governing Online Speech' (2018) 131 Harv. L. Rev. 1598, 1627

Chapter 2. Comparative analysis of the NetzDG and § 230 of the CDA provisions shaping intermediary liability

Having examined the interests of the parties and the primary purpose of SMP intermediary liability regulation, the next step of the analysis is to explore and assess the NetzDG and § 230 of the CDA in order to ascertain what sensible regulation should look like and what elements can be borrowed therefrom [NetzDG and § 230 of CDA]. In this chapter I examine the applicability of § 230 of the CDA and NetzDG to SMPs and relevant provisions thereof which shape the intermediary liability of SMPs, in a comparative context. The chapter concludes with a brief comparative summary of § 230 of the CDA and the NetzDG, delineating the main characteristics and differences.

2.1 Applicability to SMPs

Before going into the substance of the norms, it should be determined what the legal qualifications of SMPs under section § 230 of the CDA and the NetzDG are, and whether they are covered by these legislations. According to § 230 (c)(1) of the CDA, "no provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider". Under this framework, SMPs are qualified as providers of interactive computer service, as they enable computer access by multiple users to a computer server. Information content provider itself means "any person or entity that is responsible, in whole or in part, for the creation or development of information provided through the Internet or any other interactive computer service".[55] Therefore, users are

[55] § 230 (f)(3) of the CDA

information content providers, as they create or develop the content which is provided through SMPs. Sometimes it is difficult to discern and determine whether a platform (interactive computer service provider) itself becomes information content provider. For example, this can happen by editing lawful content and thus rendering it unlawful, providing questionnaires requiring unlawful answers[56], actively soliciting unlawful content, or other such cases.[57] In a case where they become information content provider, the immunity under § 230 of the CDA is lost.[58] In *Fraley v. Facebook*, the court stated that Facebook is an interactive computer service provider under § 230 of the CDA.[59] However, it also reasoned that in the context of the plaintiffs' claims, Facebook also qualified as an information content provider, thus losing its immunity under § 230 of the CDA. Consequently, under § 230 of the CDA, SMPs are providers of an interactive computer service and enjoy the immunity thereby granted. However, as illustrated by *Fraley v. Facebook*, they might also qualify as an information content provider and consequently lose their immunity.

A similar provision of the NetzDG is the section 1.(1), according to which "NetzDG applies to telemedia service providers which, for profit-making purposes, operate internet platforms which are designed to enable users to share any content with other users or to make such content available to the public (social networks)." Its text continues: "platforms offering journalistic or editorial content, the responsibility for which lies with the service provider itself, shall not constitute social networks within the meaning of this Act. The same shall apply to platforms which are designed to enable individual communication or the dissemination of

[56] See eg *Fair Hous. Council v. Roommate.com, LLC* [2012] 666 F.3d 1216 (United States Court of Appeals for the Ninth Circuit).
[57] See Edward Fenno and Christina Humphries, 'Protection under CDA § 230 and Responsibility for "Development" of Third-Party Content' (2011) 28 Comm Law
[58] Michelle Jee, 'New Technology Merits New Interpretation: An Analysis of the Breadth of CDA Section 230 Immunity' (2012) 13 Hous Bus & Tax LJ 178, 185 citing § 230(f)(2)-(3) of the CDA and *Zeran v. Am. Online, Inc.* [1997] 129 F.3d 327 [332]-[34] (United States Court of Appeals for the Fourth Circuit).
[59] *Fraley v. Facebook, Inc.* [2011] 830 F. Supp. 2d 785 (United States District Court for the Northern District of California, San Jose Division).

specific content." The norm has two sides – positive and negative. On the one hand, it covers all the telemedia service providers which enable users to share content or make it public. On the other hand, it excludes individualized communication services such as messaging and email applications.[60] Furthermore, content which is subject to editing, such as news websites, are also excluded.[61] The wording of this provision clearly reveals that the statute is specifically aimed at SMPs. A key difference from the US jurisprudence, though, is that the NetzDG is silent about what the journalistic or editorial content is and what type and degree of activity would equate to editorial or journalistic content on the part of SMPs. In the end all major SMPs are covered by the NetzDG, as all qualifying criteria are met and they [SMPs] are not editing the content. In sum, both § 230 of the CDA and the NetzDG apply to major SMPs.

2.2 Substance

According to the § 230 (c)(1) of CDA, "no provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider." According to the following article (c)(2), "no provider or user of an interactive computer service shall be held liable on account of any action voluntarily taken in good faith to restrict access to or availability of material that the provider or user considers to be obscene, lewd, lascivious, filthy, excessively violent, harassing, or otherwise objectionable, whether or not such material is constitutionally protected". Hence, under (c)(1) SMPs are given special status and they are not treated as publishers or speakers of any information provided by users.[62] Consequently, SMPs are not under the obligation to restrict or take down illegal content, and they are consequently exempt from the intermediary liability for the content. Further, under (c)(2)(A) of the CDA, SMPs are exempt from civil liability

[60] Stefan Theil, 'The German Netzdg: A Risk Worth Taking?' (2018) Verfassungsblog on Matters Constitutional <https://verfassungsblog.de/the-german-netzdg-a-risk-worth-taking/> accessed 11 June 2022
[61] ibid
[62] Browne-Barbour (n32) 1523.

whenever they restrict, remove or limit access to the content they consider "to be obscene, lewd, lascivious, filthy, excessively violent, harassing, or otherwise objectionable, whether or not such material is constitutionally protected".[63] Hence, SMPs are allowed, without incurring civil liability, to take down the content which they deem inappropriate whether or not it is constitutionally protected. "This is a double-pronged protection for moderation: it gives moderators immunity both for the content they moderate and the content they miss".[64]

Because of this bilateral immunity, content moderation and SMP's subsequent liability for the user content is largely delegated to SMPs. From the text of the § 230 of the CDA it is clear that SMPs are not deemed either publisher or speaker. However, it is disputed whether or not SMPs face distributor liability under § 230 of the CDA. In the renowned case *Zeran v. Am. Online*[65] the fourth circuit rejected the defendant's distributor liability reasoning that "interpreting § 230 to leave distributor liability in effect would defeat the two primary purposes of the statute and would certainly "lessen the scope plainly intended" by Congress' use of the term "publisher"".[66] Basically, the court interpreted distributor as the publisher for the purposes of § 230 of the CDA and thus exempted the defendant from the liability even upon notice of illegal content. *Zeran* established "a national standard for the interpretation of § 230 of the CDA".[67] This suit was followed by the Supreme Court of California in *Barrett (Stephen) v. Rosenthal (Ilena)*[68]. Here the Court of Appeal established that the § 230 immunity did not cover common law distributor liability. However, the Supreme Court of California reversed and stated that Congress did not intend to leave distributors out of the § 230 immunity. It concluded that § 230 of the CDA "does not permit Internet service providers or users to be sued as

[63] § 230 (c)(2)(A) of the CDA
[64] James Grimmelmann, 'The Virtues of Moderation' (2015) 17 Yale Journal of Law & Technology 42, 103
[65] *Zeran v. Am. Online, Inc.* [1997] 129 F.3d 327 (United States Court of Appeals for the Fourth Circuit).
[66] ibid [334].
[67] Amanda Bennis, 'Realism About Remedies and the Need for a CDA Takedown: A Comparative Analysis of § 230 of the CDA and the U.K. Defamation Act 2013' (2015) 27 2 Florida Journal of International Law 297, 311
[68] *Barrett v. Rosenthal* [2006] 40 Cal. 4th 33 (Supreme Court of California).

distributors."[69] Several federal courts of the US went on to interpret § 230 this way.[70] But this approach and interpretation of § 230 of the CDA was also criticized by many authors.[71] Further, not every court followed the same suit.[72] As a result, under the § 230 there is no consistent approach nor coherent case law regarding the form of liability of interactive computer service providers. Undoubtedly, they do not face publisher's or speaker's liability. As mentioned above, according to the majority of the US federal courts they do not face distributor liability either. However, this approach is subject to a lot of criticism and there were a considerable number of endeavors to repeal, amend and/or edit § 230 of the CDA, both on an academic and legislative level.[73] Both the jurisprudence and academia show that § 230 of the CDA lacks consistency and does not meet modern challenges.[74] Nevertheless, § 230 of the CDA is still there and content moderation is largely delegated to SMPs thereunder. As outlined earlier, the response of this hands-off approach is the creation of the Facebook Oversight Board which adjudicates cases under Facebook's Community Standards.

Conversely, Germany opted for more hands-on, administrative regulation – the NetzDG which imposes strict rules on SMPs regarding user content. For the present comparative analysis to be comprehensive comparing only the NetzDG to § 230 of the CDA is not sufficient. The issue of liability for SMPs is also covered by the E-commerce Directive and the German Telemedia Act. Specifically, articles 7-10 of the Telemedia Act are a reflection of articles 12-14 of the E-commerce Directive. As previously stated, according to the E-commerce Directive

[69] ibid [63].
[70] See eg *Green v. Am. Online (AOL)* [2003] 318 F.3d 465 (United States Court of Appeals for the Third Circuit) and *Dowbenko v. Google Inc.* [2014] 582 Fed. Appx. 801 (United States Court of Appeals for the Eleventh Circuit).
[71] Browne-Barbour (n32) 1531-33.
[72] See eg *Fair Hous. Council v. Roommates.com, LLC* [2008] 521 F.3d 1157 (United States Court of Appeals for the Ninth Circuit).
[73] See eg H.R.874 - AOC Act <https://www.congress.gov/bill/117th-congress/house-bill/874/text?q=%7B%22search%22%3A%5B%22section+230%22%5D%7D&r=1&s=3> accessed 11 June 2022, H.R.492 - Biased Algorithm Deterrence Act of 2019 <https://www.congress.gov/bill/116th-congress/house-bill/492> accessed 11 June 2022, S.797 - PACT Act <https://www.congress.gov/bill/117th-congress/senate-bill/797/text> accessed 11 June 2022
[74] McPeak (n15) 1581.

and the Telemedia Act, SMPs are compelled to block or remove access to unlawful content expeditiously upon knowledge.[75] Thus, even before NetzDG, SMPs faced distributor liability upon knowledge. "The NetzDG does not introduce a new liability regime nor does it render previously legal speech illegal."[76] The NetzDG can be seen as an extension of the Telemedia Act and distributor liability, as it specifies and prescribes concrete timelines for the review of the content and subsequent actions. Under the NetzDG, unlawfulness is to be assessed under the relevant German Criminal Code provisions referred to by the NezDG.[77] According to the NetzDG, SMPs shall review and remove or block the content which is manifestly unlawful within 24 hours of receiving a complaint.[78] This norm was subject to criticism because of freedom of speech considerations – namely the potential over-blocking of content and additionally, privatization of judiciary function to assess the content under criminal law in restricted timelines.[79] However, as we will see below in both Facebook's and Google's NetzDG transparency reports that [over-blocking] did not turn out to be the case. The same obligation of reviewing unlawful content, except that which is manifestly unlawful, is to be performed immediately but within 7 days of receiving complaint.[80] This 7-day timeline also triggered some scepticism as it is sometimes tricky even for the courts to assess and differentiate whether a speech is a fact or an opinion, and thus make a proper judgement under German criminal law.[81] It was argued that this could lead to taking down almost all contested content and thus violate human rights.[82] Thus, in addition to previously existed distributor liability, the NetzDG imposes on SMPs strict timelines for content moderation. Further, it obliges SMPs to moderate

[75] Berndt, Schmitz, (n22) 11.
[76] ibid 16.
[77] NetzDG, s 1.(3).
[78] ibid s 3.(2).2.
[79] Amélie Heldt, 'Reading between the lines and the numbers: an analysis of the first NetzDG reports' (2019) 8(2) Internet Policy Review 1, 4
[80] NetzDG, s 3.(2).3.
[81] Wolfgang Schulz, 'Comments on the Draft for an Act improving Law Enforcement on Social Networks (NetzDG)' (2017) 1, 3
[82] ibid

the content under local, German criminal law. Therefore, the NetzDG compels SMPs to embrace the role of public courts and adjudicate content under German criminal law. This privatization of judicial tasks is also economically detrimental to SMPs since it requires an additional workforce for the adjudication under local law.[83]

Despite all concerns and criticism NetzDG reports illustrate a different picture. I invoke here the most recent Facebook and YouTube NetzDG transparency reports. Within the period between January 1, 2021 and June 30, 2021 out of 77,671 NetzDG reports only 11,699 resulted in blocking or removal[84], an average of roughly 15%. Moreover, in the interval between July 1, 2021 and December 31, 2021 out of 115,085 NetzDG complaints 17.791 resulted in blocking or removal[85] which is roughly also 15%. To see the bigger picture reports from other platforms are also relevant. According to YouTube's NetzDG transparency report, approximately 83% of content was retained because it did not violate neither Community Guidelines nor criminal statutes referred in the NetzDG.[86] However, the most meaningful data in this regard is the number of removals which happened explicitly because of the NetzDG. In case of YouTube the number between the period July, 2021 and December, 2021 was 1404 out of 43,847[87] which is roughly 3%. In the interval between January and June, 2021 this number was 571 of 48,157[88] which is roughly 1%. In case of Facebook, in the interval between January and July 2021, this number was 1,092 out of 11,699[89] which is approximately 9%, and in the

[83] Thomas Kasakowskij, Julia Fürst, Jan Fischer, Kaja J. Fietkiewicz, 'Network enforcement as denunciation endorsement? A critical study on legal enforcement in social media' (2020) 46 Telematics and Informatics 1, 11
[84] 'Netzdg Transparency Report' (Facebook 2021) <https://about.fb.com/de/wp-content/uploads/sites/10/2021/07/Facebook-NetzDG-Transparency-Report-July-2021.pdf> accessed 11 June 2022
[85] 'Netzdg Transparency Report' (Facebook 2022) <https://scontent-vie1-1.xx.fbcdn.net/v/t39.8562-6/272780755_501288795008908_3397613968114653452_n.pdf?_nc_cat=107&ccb=1-7&_nc_sid=ae5e01&_nc_ohc=J-vhq4r8w7EAX-h2wIV&_nc_ht=scontent-vie1-1.xx&oh=00_AT-r-z9Qd_U-_aipp0msRDMwqAHPFYm2OF-Iylg4lRwcVA&oe=62AA71D3> accessed 11 June 2022
[86] 'Removals Under The Network Enforcement Law' (Google 2021) <https://transparencyreport.google.com/netzdg/youtube?items_by_submitter=period:2019H2&lu=turnaround_time_by_reason_chart&turnaround_time_by_submitter=period:2021H1&turnaround_time_by_reason_chart=period:2018H2> accessed 11 June 2022
[87] ibid
[88] ibid
[89] (n 84)

period between July and December 2021 this number was 1,082 out of 17,730[90] which is around 6%. This data illustrates that the content taken down explicitly due to the NetzDG and respective relevant German criminal law provisions was a very little part of overall takedowns. Therefore, fears of curbing freedom of expression excessively did not realize in practice, most probably because of SMPs' already existent content moderation rules and policies.

However, there are still some concerns about the NetzDG. One of these is that the NetzDG does not make gradations of, and differentiation between, for example, insulting content and content propagating terrorist groups.[91] Prioritization in the NetzDG, and consequently in content moderation would enhance the level of moderation.[92] Another important issue to be taken into consideration is the rights of the parties concerned. When the state decides to heavily intervene in the private sphere with hands-on regulation such as the NetzDG, it should fairly provide remedies for all the parties. In this regard, one significant drawback of the NetzDG was the absence of appeal mechanism for the users whose content was taken down.[93] Or, the other side of the coin would be the fact that the NetzDG did not prohibit taking down lawful content at all. Speech of users could be freely censored by SMPs in case of a complaint because they [users] were left without effective remedies. Although users still had the possibility to submit the matter to the public courts, the costs of litigation in case of one public post on SMP would in most cases outweigh its benefits. Therefore, these users found themselves with no potent mechanism to redress. "This was seen to create a lopsided incentive structure that would motivate companies to take down more than necessary, resulting in a restriction on freedom of expression and undermining liberal democracy".[94] However, after

[90] (n 85)
[91] Rebecca Zipursky, 'Nuts About NETZ: The Network Enforcement Act and Freedom of Expression' (2019) 42 Fordham Int'l L.J. 1325, 1351
[92] ibid
[93] ibid 1352.
[94] Laura Fichtner, 'Moderating the Regulators/Regulating the Moderators: NetzDG and online content moderation in Germany' (2021) (3rd Weizenbaum Conference: Democracy in Flux – Order, Dynamics and Voices in Digital Public Spheres, Berlin) 1, 2 <https://doi.org/10.34669/wi.cp/3.5> accessed 11 June 2022

recent amendments, the NetzDG entitles users to appeal decisions of SMPs regarding removal or blocking of the content.[95] Complainants now also enjoy the right to appeal a decision of SMP regarding the retention of content.[96] Thus, the amended NetzDG equally equips both parties with the right to appeal, which deserves commendation. By implementing the appeal mechanism, the NetzDG introduced due process elements in content moderation.[97] What is more, under the NetzDG SMPs shall provide a reasoning for their decisions to the complainant and the respondent user.[98] However, it does not specify how deep and extensive this reasoning should be and therefore gives some flexibility to SMPs. This approach is reasonable, since the mandatory appellation mechanism is supposedly already a heavy burden on SMPs. In this regard transparency reports, which would include statistics abouts appeals and results, are not yet available - as according to the NetzDG, prior version of the NetzDG and respectively former section 2 - report requirements are applicable to reports provided up to December 31, 2021.[99]

One of the main characteristics of the NetzDG is its thorough reporting obligations for SMPs.[100] The NetzDG obliges SMPs which receive more than 100 complaints per calendar year to provide half-early reports on the handling of complaints about unlawful content.[101] Paragraph 2, which has been amended, delineates all the required information to be included in reports. Although, there is no report available yet according to the new version of paragraph 2, required information seems to be comprehensive because it includes not only general information about complaint handling but also very specific and indicative data.[102] Utility of this data is twofold. On the one hand, it answers human rights' concerns as it brings

[95] NetzDG, s 3b.(1).
[96] ibid
[97] Sonja Solomun, Maryna Polataiko, Helen A. Hayes, 'Platform Responsibility and Regulation in Canada: Considerations on Transparency, Legislative Clarity, and Design' (2021) 34 Harv. J.L. & Tech. Dig. 1, 16
[98] NetzDG, s 3.(2).5.a).
[99] ibid s 6.(3).
[100] Schwiddessen, Clark, Defaux, Groom, (n 33) 3.
[101] NetzDG, s 2.(1).
[102] See ibid s 2.(2).

transparency and informs users about which content is taken down and why, what the numbers of complaints and taken-down content are, what the decision-making criteria are and examination procedures. On the other hand, certain data under section 2 informs the state about the NetzDG's implication on SMPs' business viability. For example, in addition to data and numbers mentioned above, section 2 includes the data about the units responsible for processing complaints, trainings and support for the personnel responsible for processing complaints, number of complaints where an external body was consulted, etc. All this data informs the state indirectly what the compliance costs for SMPs are.

Finally, under the NetzDG, SMPs are obliged to pay fines in case they do not meet reporting requirements, complaints' handling requirements including organizational requirements, crime reporting requirements, or if they fail to name a person authorized to receive service in the Federal Republic of Germany, or in case such person fails to respond to request for information from public authorities.[103] Basically, under the NetzDG any violation of obligations thereof can be subject to a fine. Although these fines are severe and might reach up to 50 million Euros, they should be imposed only in case of systematic noncompliance according to legislative materials.[104] One should not expect efficacy of a hands-on approach such as the NetzDG without proper tools in the hands of the state. These administrative fines are mechanisms to ensure compliance on the part of SMPs because such financial harm deters SMPs from breaching NetzDG norms. One noteworthy novelty in this regard is supervisor administrative authority - the Federal Office of Justice which monitors compliance with NetzDG provisions.[105] If this body determines that the SMP has violated or is violating NetzDG provisions, it shall take necessary measures against the SMP - in particular, it may compel the SMP to cease infringement.[106] Authority and function of this body, however, can be

[103] See ibid s 4.
[104] Schwiddessen, Clark, Defaux, Groom, (n 33) 5.
[105] NetzDG, s 4a.(1).
[106] ibid s 4a.(2).

characterized as too vague. It is not obvious what its capacity is for monitoring the activities of SMPs. Under the NetzDG, two of the main obligations of SMPs are content moderation and reporting. With the latter there is arguably no pressing need for such a body, since reports are already provided twice a year and they are publicly available. In terms of content moderation, according to the legislative material the Federal Office of Justice will be able to approach SMPs in a forward-looking manner on the basis of its supervisory function (Section 4a.(1). NetzDG) with the aim of eliminating possible violations.[107] This also begs the question – how is the Federal Office of Justice able to eliminate or assist in eliminating violations? Content moderation happens *ex ante* and *ex post*. The former is usually an automated process and happens through the algorithms of SMPs before the content is published. *Ex post* moderation happens after the content is already published and it is flagged by the user. In such a case human moderators evaluate the content in the light of applicable rules. In this process it is hardly conceivable what the role of the Federal Office of Justice can be. Such a norm which does not specifically prescribe the capacity of public authority and its designation can unjustifiably broaden intermediary liability of SMPs and it raises concerns about legal certainty. Content moderation is a very specific and peculiar activity which is why even German legislation delegates this duty to SMPs, though content adjudication happens under German criminal law. Every effective tool for content moderation is in the hands of SMPs. They are the first to detect illegal content and they are the first to take it down. Therefore, it is very hard to justify the role of an administrative body and process in content moderation.

2.3 Comparison in a nutshell

[107] Carsten Müller, Florian Post, Stephan Brandner, Roman Müller-Böhm, Niema Movassat, Manuela Rottmann, 'Beschlussempfehlung und Bericht des Ausschusses für Recht und Verbraucherschutz (6. Ausschuss) zu dem Gesetzentwurf der Bundesregierung Entwurf eines Gesetzes zur Änderung des Netzwerkdurchsetzungsgesetzes' (Deutscher Bundestag 2021) 1, 14 <https://dserver.bundestag.de/btd/19/293/1929392.pdf> accessed 11 June 2022

Although a string of US court cases is pointing in a different direction, under § 230 of the CDA and US leading case law SMPs enjoy blanket immunity - they do not face even distributor liability upon notice of illegal content. Conversely, under German legislation SMPs face distributor liability, which is enforced through strict and concrete requirements and timelines. Further to this, in the US SMPs are allowed to take down any content whether or not it is constitutionally protected.[108] Meanwhile Germany takes a different approach – the NetzDG entitles users whose content was taken down to appeal the decision thereabout.[109] This implicitly means that it is not allowed to remove lawful content since on the appeal SMPs have the obligation to reinstate such [lawful] content. This inference is supported by German jurisprudence, according to which Facebook did not have the right to delete content which was legal according to German law.[110] As mentioned above, under § 230 of the CDA content moderation is largely delegated to SMPs. This means that the US approach offers centralized moderation since it is carried out according to SMPs uniform, internal rules[111] and decisions made according to them are binding globally.[112] Conversely, moderation under the NetzDG is local since it is carried out locally in Germany and decisions made according to the NetzDG is binding in Germany meaning that the unlawful content is blocked within Germany.[113] Undoubtedly, from a business viability perspective, uniform and centralized moderation is preferable as SMPs do not have to adapt to local rules and regulations, which means additional costs for business. On the other hand, in large communities such as Facebook and Twitter decentralized and localized moderation might be more efficient, because maintaining norms and their legitimacy in a smaller and fragmented community is easier.[114] It also takes cultural

[108] § 230 (c)(2)(A) of the CDA
[109] NetzDG, s 3b.(1).
[110] Echikson and Knodt (n 43) 9.
[111] eg Facebook Community Standards <https://transparency.fb.com/policies/community-standards/> accessed 11 June 2022
[112] Grimmelmann (n 64) 55.
[113] ibid
[114] ibid 73.

and legalistic differences between communities and nations into account. In today's world order absolutely centralized moderation seems impossible due to the fact that the states are trying to impose their own legal regimes and preferences on SMPs.[115] One important consideration when evaluating moderation is the issue of transparency. Under the NetzDG moderation is necessarily transparent as it requires SMPs to provide reasoning for their decisions to the user and complainant.[116] Through the biannual reporting obligations on the part of SMPs, transparency is further enhanced. However, under § 230 of the CDA the reality is different, because SMPs are not legally obliged to provide reasoning and justification for the decisions made upon content or transparency reports. In sum, the NetzDG and § 230 of the CDA are poles apart as the latter gives freedom to SMPs whereas the former imposes strict requirements. § 230 of the CDA was enacted to promote development of interactive computer services, technology and free speech. Under § 230 of the CDA, realization of free speech rights of users is largely dependent on SMPs. On the other hand, as seen in the earlier quoted statistics, the NetzDG did not cause over-blocking and collateral censorship. What is interesting in terms of free speech is that in a sense the NetzDG is more promotive of free speech as it compels SMPs to retain and reinstate legal content, and to provide reasoning for decisions.

[115] See eg Austrian Federal Act on measures to protect users on communication platforms 2021 (Communication Platforms Act)
[116] NetzDG, s 3.(2).5.a).

Chapter 3. Healthy content – which way to go?

Given the concerns highlighted with respect to both § 230 of the CDA and the NetzDG, this Chapter offers some views on the core tenets of a sensible SMP intermediary liability regulation. First, it presents an analysis on whether external, state laws are needed. Second, it lays out the relevant elements and issues which should be covered and taken into consideration by such regulation. Lastly, this chapter is followed by the thesis' final conclusion.

3.1 Need for external regulation?

Fundamentally, a "healthy" SMP content shall have two aspects. It should guarantee freedom of expression for users on the one hand, while avoiding toxic content on the other. The previous chapters have outlined how this model is closer to European freedom of expression rather than American. Further definition of a "healthy" content is not necessary here, as the central question that remains is not what, but: How? In order to find the most pertinent approach for "healthy" content several aspects should be taken into consideration. Primary among these is whether we need an external state regulation, or should the issue be left to the private actors as in the US? It has already been discussed how creating "healthy" content is the primary purpose for the public and state. The same applies to the business sector but, albeit with some reservations, as "unhealthy" posts are often more pervasive and trigger massive user engagement.

As a consequence of this, I argue that the matter should not be fully delegated to the business. The state should have some instant method of remedy in case self-regulation is not properly functioning. Since the scale of major SMPs are global, dissemination of the content might cause irreparable harm for example in the case of child pornography. SMPs' content

moderation policies and rules have a huge impact on the rights to free speech, public discourse and information flow.[117]

Moreover, since § 230 of the CDA was enacted in 1996, SMPs and interactive computer service providers today currently need no further assistance and promotion from the state as they have become large-scale and powerful. "The Internet is no longer a child of the 90s' needing congressional helicopter-parenting."[118] Nowadays, the state is the runner-up to Big Tech and should try to catch up urgently with its speedy development. A further concern is that self-regulatory laws might not be sufficient "in quality" meaning not providing proper and equal guarantees and remedies for the users.[119] In order for these arguments not to stay general and abstract, whenever there are examples and empirical data, it is better to use them. In this regard, such an example is § 230 of the CDA and Facebook's self-regulation. Whenever SMP breaches self-imposed rules, it is not going to face any kind of external and effective legal liability. According to internal self-imposed rules, SMPs are not entitled to derogate from them. However, in some cases they [SMPs] might behave in a way that contravenes internal rules which is illustrative of ineffectiveness thereof. SMPs might unfairly derogate from internal rules, for example, in favor of powerful people.[120] Kate Klonick invokes two such examples[121]. First, in 2016, Tom Egeland, a Norwegian author, posted a photo[122] of a little Vietnamese girl running naked following a napalm attack and it violated Facebook's terms of service, most

[117] Hannah Bloch-Wehba, 'Automation in Moderation' (2020) 53 Texas A&M University School of Law Legal Studies Research Paper No. 20-33 41, 51
[118] Alex S. Rifkind, 'Dyroff v. Ultimate Software Group, Inc.: A Reminder of the Broad Scope of § 230 Immunity' (2021) 51 Golden Gate U. L. Rev. 49, 68
[119] Karanjot Gill, 'Regulating Platforms' Invisible Hand: Content Moderation Policies and Processes' (2021) 21 Wake Forest J. Bus. & Intell. Prop. L. 171, 209
[120] Kate Klonick, 'The New Governors: The People, Rules, and Processes Governing Online Speech' (2018) 131 Harv. L. Rev. 1598, 1654-55
[121] ibid
[122] ibid 1654 citing Kate Klonick, 'Facebook Under Pressure' (2016) SLATE <https://slate.com/technology/2016/09/facebook-erred-by-taking-down-the-napalm-girl-photo-what-happens-next.html> <https://perma.cc/6A4U-UYC5> accessed 12 June 2022

likely because of nudity, according to the author.[123] The photo was taken down and the account was suspended.[124] CEO of Norwegian newspaper – Espen Egil Hansen also had the same photo taken down.[125] The same happened to Norwegian Prime Minister Erna Solberg.[126] Following this, Espen Egil Hansen published a letter to Zuckerberg on the front page of a newspaper urging Facebook to avoid censorship.[127] Shortly after, the photo was reinstated and Facebook's COO apologized.[128] In a different such scenario, Facebook maintained announcements of then-presidential candidate Donald Trump which contravened the internal rules.[129]

These examples illustrate that the behavior of SMPs sometimes might be flawed in favor of some persons or circumstances. This is exceptionally risky, especially when we are speaking about one of the biggest spaces for discourse where people are exercising their public rights. SMPs also have an impact on security as they have been used by terrorists to recruit volunteers, livestream attacks and sow propaganda.[130] Therefore, leaving the bulk of the responsibility to SMPs is not the solution since, as examples illustrate, SMPs might not treat its users equally. This type of discriminatory action should be absent from the platforms which are deemed to be the among the biggest means of expression and communication. Consequently, externally binding state laws are necessary. Which begs the questions: how

[123] Ibid 1654 citing Kjetil Malkenes Hovland and Deepa Seetharaman, 'Facebook Backs Down on Censoring 'Napalm Girl' Photo' (2016) WALL ST. J. < http://on.wsj.com/2bYZtNR> accessed 12 June 2022
[124] ibid
[125] ibid citing Espen Egil Hansen, 'Dear Mark. I am writing this to inform you that I shall not comply with your requirement to remove this picture' (2016) Aftenposten <https://www.aftenposten.no/meninger/kommentar/i/G892Q/Dear-Mark-I-am-writing-this-to-informyou-that-I-shall-not-comply-with-your-requirement-to-remove-this-picture> <https://perma.cc/49QW-EDUT> accessed 12 June 2022
[126] Klonick (n 120) 1654 citing Hovland and Seetharaman (n 123)
[127] Klonick (n 120) 1654 citing Hansen (n125)
[128] Klonick (n 120) 1654 citing Claire Zillman, 'Sheryl Sandberg Apologizes for Facebook's "Napalm Girl" Incident' (2016) Time <http://time.com/4489370/sheryl-sandberg-napalm-girl-apology> <https://perma.cc/Z7N4-WA2P> accessed 12 June 2022
[129] Klonick (n 120) 1655 citing Deepa Seetharaman 'Facebook Employees Pushed to Remove Trump's Posts as Hate Speech' (2016) WALL ST. J. <http://on.wsj.com/2ePTsoh> accessed 12 June 2022
[130] Brittany Doyle, 'Self-Regulation Is No Regulation--The Case for Government Oversight of Social Media Platforms' (2022) 32 Int'l & Comp. L. Rev. 97, 105 citing Md Sazzad Hossain, 'Social Media and Terrorism: Threats and Challenges to the Modern Era' (2018) 22 S. Asian Surv. 136

extensive these laws should be? Should they be procedural or substantive, or both? I argue that respective laws should be both procedural and substantive.

3.2 Relevant issues to be covered

Having established that state intervention is needed, the discussion turns to the question – which issues should be covered by state laws in order to provide "healthy" content? First and foremost, I believe laws should be clear about what kind of liability is imposed on SMPs for user content. The law imposing certain types of liability on SMPs should not be engineered so as to provide incentives for SMPs to simply take down the content when there is a doubt whether it falls under free speech, in order to avoid legal liability. Hence, the law (whether state or internal) should differentiate between certain types of unlawful content. If SMPs have too short a timeline for removing doubtful and disputable content, it might put them in the position to take the content down to minimize legal risk. In such a case, legislation will have a chilling effect on freedom of expression and cause collateral censorship. However, when the content contains an imminent threat or is obviously illegal, maintaining it might lead to risk of severe consequences. In such a scenario, restricted deadlines are justified. SMPs should be given proper and proportional timelines so that the necessary processes can be duly followed.

The second important consideration in terms of external regulation is its role and designation - specifically, what the interplay between external laws and SMPs' self-regulation should be to better attain the objective of creating "healthy content". Does the existence of external laws necessarily mean applying them in every individual case? Would it be reasonable to compel SMPs to make criminal law "judgements" every time content is reported? As I argue below, external laws should be granted the role of minimum standard. This model ensures that effective self-regulation and internal rules are not made redundant.

However, this alone is not enough. Internal rules and decisions should be trusted by the customer base. In order to make them trustworthy, the process, rules and reported statistics have to be transparent. Yet, transparency is not always in business's interest. Therefore, we can reason that together with self-regulation SMPs need a little nudge from the state. As previously mentioned, sometimes SMPs might prefer to retain "not wholly healthy" content which triggers different kind of negative emotions or act unfairly in favor of powerful people – another reason why a model relying entirely on self-regulation will not suffice. A final important consideration is the enforceability of norms. Since the most prominent SMPs are based in the United States, states should take into account whether the law would be enforceable in the US.

3.2.1 Form of liability

This thesis does not intend to provide a full-fledged legislative proposal nor indicate the structure of the law concerning SMPs' intermediary liability, but only to suggest the essentials of a sensible regulatory approach. However, I believe there are questions which have to be answered with certainty in regard to the aim of creating "healthy" content. It would not be an exaggeration to say that the first and most prominent such issue is a form of liability. In the United States, platforms are exempted from liability under § 230 of the CDA, whereas in Germany they face distributor (Störer) liability. In Germany, the NetzDG, read together with the Telemedia Act, imposes distributor liability on SMPs, since once they receive complaints, they are obliged to check and remove or block the content, or face penalty and possible tort liability. In the first instance, I believe that SMPs should not face publisher's liability in any case, since it will undermine the business model of every SMP. Unlike newspapers, one of the main amenities of these platforms [SMPs] is the opportunity given to the users to react and express opinions instantly, without pre-editing from SMPs. Undoubtedly, with publisher's liability this feature will be compromised, and user engagement will decrease dramatically,

thus decreasing income for business sharply. Costs will also increase because of the resources needed to pre-edit every post. SMPs will have a strong incentive not to allow content to avoid liability, having a dampening effect on freedom of speech and encouraging collateral censorship. Although publisher's liability is exceptionally burdensome, absolute exemption from liability can be equally bothersome, since consumers will be dependent on private actors to enforce their public rights.[131] As § 230 of the CDA experience illustrates above, absolute exemption might in some cases lead to discriminatory results, especially in cases where powerful public figures are involved.

Concerningly, under § 230 of the CDA, users are deprived of effective remedies and due process elements. In *Caraccioli v. Facebook,* an anonymous person created a Facebook account with plaintiff's name, and uploaded videos and pictures of him sexually pleasuring himself.[132] Facebook did not delete the account, stating that it did not infringe Facebook Community Standards. The Plaintiff claimed that Facebook was liable, as it reviewed the account and did nothing to take action. However, the court reasoned that "liability based on that sort of vicarious responsibility, however, is exactly what § 230(c) seeks to avoid".[133]

In light of these facts I argue that the most suitable form of liability for creating "healthy" content is German liability upon notice

Liability upon the notice mitigates the abovementioned risks, and consequently SMPs will not be able to retain the unlawful post once it is reported. However, simply providing that the SMPs face liability upon knowledge or notice, is not enough to ensure compliance. It is of vital importance to ensure efficient enforcement of distributor liability – in this instance, by efficiency, I mean timely action on the part of the SMP. To exclude the SMP's conflict of interest regarding a possibly delayed action against a harmful post, specific timelines and

[131] Klonick (n 120) 1668.
[132] *Caraccioli v. Facebook, Inc.* [2016] 167 F. Supp. 3d 1056 (United States District Court for the Northern District of California, San Jose Division).
[133] ibid [1066].

procedural rules should be provided by the law. The procedural provisions should be external as well, because SMPs' conflict of interest is likely to be mirrored in internal regulations or possibly in practice. Such enforcement rules should additionally take different types of harmful content into account. In the case of imminent threat or danger, prompt action is desirable. For instance, giving the same deadlines for reviewing terrorism related content as for defamatory content is not justified. Liability upon notice is not only a guarantee for "healthy" content, but also a means for protecting users' rights. But to hammer home the point, in order for this protection to be effective, concrete and external procedural rules are needed.

Content also cannot be deemed "healthy" if freedom of expression is not guaranteed and the conditions to encourage collateral censorship exist. Therefore, when considering specific, external timelines and liabilities, an important issue to be considered is SMPs' incentive to just simply remove the content when its legality is disputable. There was an opinion that removing § 230 of the CDA protection and penalizing platforms does not work in a way to ensure content safety, but rather it induces platforms to remove the content to minimize the risks.[134] However, this opinion was not supported by the SMPs' transparency reports analyzed above.

Instead, data from the transparency reports indicate that the content taken down explicitly because of NetzDG was a negligible fraction of overall removals. Therefore, the assumption that the NetzDG or external liability necessarily leads to collateral censorship does not stand up to scrutiny. Furthermore, these numbers indirectly show that there is a very small difference between the interests of state and business to ensure "healthiness" of content. As mentioned above, to avoid SMPs defaulting to removing content, deadlines should be adjusted according to the nature of content. A relevant example of this is the NetzDG which provides different timelines for content removal. Namely, it gives platforms 24 hours, from the receipt

[134] Brown (n 39) 475.

of complaint, to block or remove manifestly unlawful content.[135] Regarding unlawful content, platforms have to act immediately, within 7 day from receiving complaint.[136]

However, the NetzDG has one noteworthy drawback. For instance, two of the crimes referred to by the NetzDG are forming terrorist organizations and use of symbols of unconstitutional organizations. Even if the content containing these two crimes are manifestly unlawful, is it warranted for SMPs to be given the same timeframe of 24 hours to remove both kind of content? I disagree, because allowing terrorism related content[137] for 24 hours can be devastating and cause irreparable harm. Hence, the NetzDG needs a more sophisticated framework to differentiate between certain types of crimes and application of respective moderation timelines.

As regards substantive law, I believe that the law should be specific as possible about what type of content is subject to removal. Although this means empowering platforms with the right to interpret criminal or administrative norms, this process is arguably inevitable and the NetzDG is an example of this.

3.2.2 The role of external regulation

In today's world free speech is a triangle comprised of speakers, states and private actors.[138] Hence, we need to accept that some aspects of the judiciary will be and in fact already is privatized and find the best interplay between private and public. In the case of SMPs' internal regulations and state law, a balanced example of this interplay would be, I believe, granting the external law the function of the net and minimum standard, because of economical and efficiency arguments.

[135] NetzDG, s 3.(2).2.
[136] ibid s 3.(2).3.
[137] See eg *Force v. Facebook, Inc.* [2019] 934 F.3d 53 (United States Court of Appeals for the Second Circuit).
[138] Jack M. Balkin, 'Free Speech is a Triangle' (2018) 118 Colum L. Rev 2011, 2055

Analysis of the NetzDG showed that the existence of external laws does not necessarily mean applying them in all cases. As Facebook's NetzDG transparency reports[139] show, complaints under the NetzDG are first checked in the light of internal rules and only if they do not infringe them is German criminal law called into play.

The state's objective to remove and avoid dissemination of the content, which is unlawful according to German criminal laws, is attained through the NetzDG model because if the content is not filtered according to internal laws, there is an additional level of moderation against German criminal laws. For SMPs, this model is cost and labor saving as no SMP wants to embrace the role of public courts and adjudicate content under different state laws, since it is an additional cost to the business. Thus, in this way, the state retains leverage over business but does not heavily interfere. Presumably, German legislator could not foresee this practice of checking content first under internal rules and then under German criminal laws in case of compliance with internal rules. However, this model should be retained and SMPs should not be deprived of possibility to moderate under internal rules as it works the most efficiently for all concerned parties. For the sake of legal clarity, it would be preferable to insert such provisions in the NetzDG entitling SMPs to moderate under internal rules at first and then under local criminal laws.

The absence of internal rules and an effective grievance mechanism can also be detrimental for all interested parties. Although courts have more expertise and training than SMPs in adjudicating the lawfulness of content, they [courts] are not well-positioned for doing it in most cases, because court proceedings are time-consuming and the cost of a dispute often outweighs the social value thereof.[140] And finally, the absence of the internal mechanism is obviously troublesome for the consumer, as instead of simply clicking on the button to report

[139] (n 84) and (n 85)
[140] Eric Goldman, 'Content Moderation Remedies' (2021) 28 Michigan Technology Law Review 1, 52

the content, they will have to apply to the court and pay additional fees. As effective internal rules are beneficial for every concerned party, it should be prompted by external laws. Under the NetzDG SMPs are given this possibility. This model inherently promotes self-regulation. Consequently, SMPs have a strong incentive to come up with an effective and "high quality" internal rules and grievance system in order to avoid further examination of state laws in each case. However, one can also raise the question - why is adjudication under internal rules not as costly? The answer to the question is that the internal rules are global and centralized and they apply to every user around the globe. If these rules are sufficiently sophisticated, there is no more need for compliance with external laws. However, external laws and fines are still there in any case if the SMPs choose to abuse self-regulation. In summation, subsidiarity of external laws encourages and fosters effective self-regulation, which is beneficial for all the parties concerned.

3.2.3 Users' trust and jurisdiction

The proposed model, in which the content is reviewed under state law subsidiarily after internal adjudication, would be effective only if internal regulations and procedures have trust among SMP's users. Otherwise, users would just see internal regulations as an obstacle on the way to require evaluation internally under the state law, or even externally by courts. One legislative way for this is to provide reporting obligation in the law. A practical example of this is the NetzDG which includes transparency-reporting obligations.[141] Undoubtedly, reports should contain meaningful information. Another non-legislative way to attain users' trust is the launching of public campaigns to provide the public with pertinent information. However, exploration of non-legislative tools does not fall within the scope of this research.

[141] Danya He, 'Governing Hate Content Online: How the *Rechtsstaat* Shaped the Policy Discourse on the NetzDG in Germany' (2020) 14 International Journal of Communication 3746, 3748

Any state can enact legislation regulating user content liability issues. One crucial question to consider here though, is whether these legislations would be enforceable. The majority of the most influential SMPs such as Google, Facebook and Twitter are based in the United States. Therefore, if another country wants to enforce a decision made according to its legislation, it has to be enforceable in the United States. Problematically, not all decisions are enforceable in the United States because of the First Amendment. Although neither Yahoo! nor Google are exactly SMPs, the following cases are still relevant to illustrate the issues concerning jurisdiction. In the renowned proceedings *Yahoo! v. LICRA and UEJF*[142], Yahoo sought a declaratory judgement to declare judgements, rendered by the High Court of Paris, neither recognizable nor enforceable in the United States on the basis that they violated the First Amendment.[143] According to the orders of the High Court of Paris, Yahoo! was held liable for the content on its auction website which contained Nazism-related material.[144] District Court for the Northern District of California granted motion for summary judgement in favor of Yahoo! and stated that orders were not enforceable, because they chilled the First Amendment.[145] Another relevant case in this regard is *Google LLC v. Equustek Sols. Inc.*[146] The factual background of this case is as follows: The defendants, Equteek, had a dispute with persons associated with Datalink, a computer hardware distributor and seller company.[147] Equteek alleged that Datalink, together with a former Equsteek engineer incorporated Equsteek's trade secrets in Datalink's products.[148] Equteek prevailed and was granted Canadian

[142] Mentioned proceedings include several different court cases both in France and the US, however, here I put an accent on one of them in order to illustrate jurisdictional issues which might arise from the foreign content regulation in terms of free speech
[143] *Yahoo!, Inc. v. La Ligue Contre Le Racisme et L'Antisemitisme* [2001] 169 F. Supp. 2d 1181 (United States District Court for the Northern District of California, San Jose Division)
[144] ibid
[145] ibid
[146] *Google LLC v. Equustek Sols. Inc.* [2017] 2017 U.S. Dist. LEXIS 182194 (United States District Court for the Northern District of California, San Jose Division)
[147] ibid
[148] ibid

court orders against Datalink[149]. The defendants did not comply and fled Canada. Subsequently, Equustek asked Google to take down Datalink's material from search results. Google initially refused, but when afterwards the Canadian court granted injunctive relief to Equsteek against Datalink, Google acquiesced and blocked Datalink websites from Canadian search results only.[150] Equsteek then pursued a court order compelling Google to remove mentioned materials worldwide, which was granted by the trial court and then confirmed by the Court of Appeal and Supreme Court of Canada.[151] In this case before the United States District Court for the Northern District of California, Plaintiff Google LLC sought to prevent enforcement of a Canadian court order compelling Google to delist search results not only in Canada but worldwide.[152] Here the court based its reasoning on § 230 of the CDA and stated that all the elements of it were satisfied. First, Google was the provider of an interactive computer service. Second, the provider of information was not Google but Datalink. Third, and the most importantly, the court concluded that the enforcement of the Canadian court order would mean imposing the publisher's or speaker's liability on Google which would be incompatible with § 230 of the CDA. Subsequently, the court granted motion to Plaintiff Google LLC for preliminary injunctive relief.

These cases once again illustrate that the enforceability, and therefore the effectiveness, of the norms regarding user content liability is partly dependent on the compliance with § 230 of the CDA and US free speech principles in case of major SMPs. The German and EU[153] distributor liability of SMPs would likely be incompatible with US leading case law regarding § 230 of the CDA. However, in practice, this does not mean that Facebook is not complying with the NetzDG and EU state national laws. Enforceability of the EU national laws is not the

[149] ibid
[150] ibid
[151] ibid
[152] ibid
[153] E-commerce Directive

primary factor for SMPs' compliance. German and EU markets are precious for SMPs revenue models and user base, and therefore they have incentive to comply with the NetzDG and EU legislations.

Conclusion

In this thesis, I analyzed SMP intermediary liability laws of two major yet very different jurisdictions – Germany and the US. The topic of SMPs intermediary liability is dynamic and states' response differs.

In the US, SMP intermediary liability is still regulated under § 230 of the CDA which was enacted in 1996. Under its immunity rules, content moderation and consequent intermediary liability are largely delegated to SMPs. In response, SMPs are trying to self-regulate, as exemplified by the fact that Facebook created an Oversight Board for content adjudication. Reasons for this may vary: signaling that they do not need external state regulation, avoiding blame for controversial decisions, creating "healthy" content and a host of other operational circumstances. However, as demonstrated in this thesis, moderation under § 230 of the CDA has not been without its shortcomings. In some cases, SMPs moderated unfairly or in a discriminatory fashion. Further, sometimes SMPs prefer to retain "unhealthy" content as it disseminates faster and engages lot of users.

By contrast, Germany decided to interfere more substantively in SMPs' content moderation and imposed liability and notice take-down obligations on SMPs. By enacting the NetzDG, Germany sought to curb proliferated hate speech and fake news on SMPs. Under the NetzDG, SMPs have the duty to check the content under German criminal law once it is reported by the user. If the content is manifestly unlawful it has to be taken down in 24 hours, if it is simply unlawful – in 7 days. Thus, SMPs have been *de facto* vested with some of the authority of public bodies. Contrary to popular belief, the NetzDG did not cause collateral censorship and over-blocking, as evidenced by the statistics presented in their self-reporting.

Since the enactment of § 230 of the CDA in 1996, the priorities of regulatory authorities and SMPs have shifted. SMPs have reached a previously unimaginable scale and their user content disseminates within seconds around the globe. Content which triggers negative emotions and polarization sometimes travels even faster. States cannot afford to allow dissemination of unlawful content, especially when it contains imminent and real threats. States' primary interest is to maintain a "healthy" content environment. SMPs, too, inherently have the same interest of maintaining "healthy" content, since toxic environment deters users from engagement and therefore cause a loss of income. Consequently, it can be argued that the priorities of SMPs and states are largely aligned. As discussed in the thesis, although both SMPs and State want to create and maintain "healthy" content, as mentioned, § 230 of the CDA in many respects does not meet modern requirements of due process, security and "healthy" content.

In order to eliminate and mitigate these or future risks, state external laws such as the NetzDG are necessary and can constitute a model going forward. As the NetzDG reports show, content is not necessarily checked under German criminal provisions referred to by the NetzDG and only a small portion of content is taken down explicitly of the NetzDG. However, it has to be present if some kind of risk arises. Therefore, the NetzDG implicitly allows SMPs to assess the content in the light of internal rules at first and if it is not infringing those rules, it is checked against German relevant criminal laws.

Such flexibility is valuable to all the parties concerned. After recent amendments to the NetzDG, it has become more sophisticated as it equally entitles the complainant and the user whose content was taken down to appeal the decision and receive reasoning. However, the NetzDG is not a *panacea* and still needs further development. In particular, it needs to make distinctions between certain types of crimes and apply respective content moderation timelines in order to be most effective for its stated purpose.

The inquiry of SMP intermediary liability will continue in the future and require careful observation of the interplay between positive law and self-regulation, free speech principles, states' preferences and political ideologies. Meanwhile, these two Western jurisdictions model different paths - time will tell which, if either, will prevail.

Bibliography

Case Law

Anderson v New York Tel Co [1974] 35 N.Y.2d 746 (Court of Appeals of New York).

Barrett v. Rosenthal [2006] 40 Cal. 4th 33 (Supreme Court of California).

Caraccioli v. Facebook, Inc. [2016] 167 F. Supp. 3d 1056 (United States District Court for the Northern District of California, San Jose Division).

Case Decision 2020-006-FB-FBR [2021] (The Oversight Board).

Case decision 2021-016-FB-FBR [2022] (The Oversight Board).

Cubby, Inc. v. Compuserve, Inc. [1991] 776 F. Supp. 135 (United States District Court for the Southern District of New York).

Dowbenko v. Google Inc. [2014] 582 Fed. Appx. 801 (United States Court of Appeals for the Eleventh Circuit).

Fair Hous. Council v. Roommate.com, LLC [2012] 666 F.3d 1216 (United States Court of Appeals for the Ninth Circuit).

Fair Hous. Council v. Roommates.com, LLC [2008] 521 F.3d 1157 (United States Court of Appeals for the Ninth Circuit).

Force v. Facebook, Inc. [2019] 934 F.3d 53 (United States Court of Appeals for the Second Circuit).

Fraley v. Facebook, Inc. [2011] 830 F. Supp. 2d 785 (United States District Court for the Northern District of California, San Jose Division).

Google LLC v. Equustek Sols. Inc. [2017] 2017 U.S. Dist. LEXIS 182194 (United States District Court for the Northern District of California, San Jose Division).

Green v. Am. Online (AOL) [2003] 318 F.3d 465 (United States Court of Appeals for the Third Circuit).

Lunney v Prodigy Servs Co [1998] 250 A.D.2d 230 (Supreme Court of New York, Appellate Division, Second Department).

Stratton Oakmont v. Prodigy Servs. Co [1995] 1995 N.Y. Misc. LEXIS 229 (Supreme Court of New York, Nassau County).

Yahoo!, Inc. v. La Ligue Contre Le Racisme et L'Antisemitisme [2001] 169 F. Supp. 2d 1181 (United States District Court for the Northern District of California, San Jose Division).

Zeran v. Am. Online, Inc. [1997] 129 F.3d 327 (United States Court of Appeals for the Fourth Circuit).

Law

Austrian Federal Act on measures to protect users on communication platforms 2021 (Communication Platforms Act)

Communications Decency Act of 1996

Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market (E-Commerce Directive)

Network Enforcement Act 2017 (NetzDG)

Books

Eyal N, *Hooked: How to Build Habit-Forming Products* (Portfolio Penguin 2014)

Mangan D. and Gillies L, *The Legal Challenges of Social Media* (Edward Elgar Publishing 2017)

Articles

Balkin J, 'Free Speech is a Triangle' (2018) 118 Colum L. Rev 2011

Belli L, Francisco P and Zingales N, 'Law of the Land or Law of the Platform? Beware of the Privatisation of Regulation and Police' (2017) 41

Bennis A, 'Realism About Remedies and the Need for a CDA Takedown: A Comparative Analysis of § 230 of the CDA and the U.K. Defamation Act 2013' (2015) 27 2 Florida Journal of International Law 297

Bloch-Wehba H, 'Automation in Moderation' (2020) 53 Texas A&M University School of Law Legal Studies Research Paper No. 20-33 41

Brown N, 'Regulatory Goldilocks: Finding the Just and Right Fit for Content Moderation on Social Platforms' (2021) 8 Tex. A&M L. Rev, 451

Browne-Barbour V, 'Losing Their License to Libel: Revisiting § 230 Immunity' (2015) 30(2) Berkeley Technology Law Journal 1505

Doyle B, 'Self-Regulation Is No Regulation--The Case for Government Oversight of Social Media Platforms' (2022) 32 Int'l & Comp. L. Rev. 97

Echikson W and Knodt O, ‘Germany’s NetzDG: A Key Test For Combatting Online Hate’ (2018) No. 2018/09 CEPS Research Reports 1

Fagan F, 'Systemic Social Media Regulation' (2018) 16 Duke L & Tech Rev 393

Fenno E and Humphries C, 'Protection under CDA § 230 and Responsibility for “Development” of Third-Party Content' (2011) 28 Comm Law 1

Fichtner L, ‘Moderating The Regulators/Regulating The Moderators: Netzdg And Online Content Moderation In Germany’ (3rd Weizenbaum Conference: Democracy in Flux – Order, Dynamics and Voices in Digital Public Spheres, Berlin, June 2021) 1

Gill K, ‘Regulating Platforms’ Invisible Hand: Content Moderation Policies and Processes’ (2021) 21 Wake Forest J. Bus. & Intell. Prop. L. 171

Goldman E, ‘Content Moderation Remedies’ (2021) 28 Michigan Technology Law Review 1

Grimmelmann J, ‘The Virtues of Moderation’ (2015) 17 Yale Journal of Law & Technology 42

Hart C, 'Social Media Law: Significant Developments' (2016) 72 Bus Law 235

He D, ‘Governing Hate Content Online: How the *Rechtsstaat* Shaped the Policy Discourse on the NetzDG in Germany’ (2020) 14 International Journal of Communication 3746

Heldt A, ‘Let's Meet Halfway: Sharing New Responsibilities in a Digital Age’ (2019) 9 Journal of Information Policy 336

—— ‘Reading between the lines and the numbers: an analysis of the first NetzDG reports’ (2019) 8(2) Internet Policy Review 1

Jarsulic M, ‘Addressing the Competitive Harms of Opaque Online Surveillance and Recommendation Algorithms’ (2022) The Antitrust Bulletin 100

Jee M, ‘New Technology Merits New Interpretation: An Analysis of the Breadth of CDA Section 230 Immunity’ (2012) 13 Hous Bus & Tax LJ 178

Kasakowskij T, Fürst J, Fischer J, Fietkiewicz K, ‘Network enforcement as denunciation endorsement? A critical study on legal enforcement in social media’ (2020) 46 Telematics and Informatics 1

Klonick K, ‘The New Governors: The People, Rules, and Processes Governing Online Speech’ (2018) 131 Harv. L. Rev. 1598

—— ‘The Facebook Oversight Board: Creating an Independent Institution to Adjudicate Online Free Expression’ (2020) 129 Yale L.J. 2418

McPeak A, ‘Platform Immunity Redefined’ (2021) 62 Wm. & Mary L. Rev 1557

Müller C, Post F, Brandner S, Müller-Böhm R, Movassat N, Rottmann R, 'Beschlussempfehlung Und Bericht Des Ausschusses Für Recht Und Verbraucherschutz (6. Ausschuss) Zu Dem Gesetzentwurf Der Bundesregierung Entwurf Eines Gesetzes Zur Änderung Des Netzwerkdurchsetzungsgesetzes' (Deutscher Bundestag 2021) 1

Oliva T, 'Content Moderation Technologies: Applying Human Rights Standards to Protect Freedom of Expression' (2020) Volume 20 Issue 4 Human Rights Law Review 607

Owen T, 'The Case for Platform Governance' (2019) Center for International Governance Innovation 1

Perel M, 'Enjoining Non-Liable Platforms' (2020) 34 Harv. J. Law & Tec 1

Rifkind A, 'Dyroff v. Ultimate Software Group, Inc.: A Reminder of the Broad Scope of § 230 Immunity' (2021) 51 Golden Gate U. L. Rev. 49

Ritzmann A, Schindle H and Macori M, 'NetzDG 2.0 – Recommendations for the amendment of the Network Enforcement Act (NetzDG) and Investigation into the actual blocking and removal processes of YouTube, Facebook and Instagram' (2020) CEP Policy Paper 1

Schmitz S, Berndt C, 'The German Act on Improving Law Enforcement on Social Networks (NetzDG): A Blunt Sword?' (2018) 1

Schulz W, 'Comments on the Draft for an Act improving Law Enforcement on Social Networks (NetzDG)' (2017) 1

Schwiddessen S, Clark B, Defaux T, Groom J, 'Germany's Network Enforcement Act - closing the net on fake news?' (2018) 40(8) E.I.P.R. 539

Solomun S, Polataiko M, Hayes H, 'Platform Responsibility and Regulation in Canada: Considerations on Transparency, Legislative Clarity, and Design' (2021) 34 Harv. J.L. & Tech. Dig. 1

Wu T, 'Will Artificial Intelligence Eat the Law? The Rise of Hybrid Social-Ordering Systems' (2019) Vol. 119 No. 7 Columbia Law Review 2001

Young H and Laidlaw E, 'Internet Intermediary Liability in Defamation: Proposals for Statutory Reform' (2017) 1

Zipursky R, 'Nuts About NETZ: The Network Enforcement Act and Freedom of Expression' (2019) 42 Fordham Int'l L.J. 1325

Zurth P, 'The German NetzDG as Role Model or Cautionary Tale? Implications for the Debate on Social Media Liability' (2021) 31 Fordham Intell Prop Media & Ent LJ 1084

Online articles

Duffy C, 'Elon Musk To Buy Twitter In $44 Billion Deal' (2022) CNN Business <https://edition.cnn.com/2022/04/25/tech/elon-musk-twitter-sale-agreement/index.html> accessed 11 June 2022

Frank R, 'The Economic Case For Regulating Social Media' The New York Times <https://www.nytimes.com/2021/02/11/business/social-media-facebook-regulation.html> accessed 11 June 2022

Theil S, 'The German Netzdg: A Risk Worth Taking?' (2018) Verfassungsblog on Matters Constitutional <https://verfassungsblog.de/the-german-netzdg-a-risk-worth-taking/> accessed 11 June 2022

Other resources

Dorsey J, 'Testimony of Jack Dorsey Chief Executive Officer Twitter, Inc.', *Does Section 230's Sweeping Immunity Enable Big Tech Bad Behavior?* (2020) Hearing Before the United States Senate Committee on Commerce, Science, and Transportation <https://www.commerce.senate.gov/services/files/7A232503-B194-4865-A86B-708465B2E5E2> accessed 11 June 2022

H.R.4027 - Stop the Censorship Act <https://www.congress.gov/bill/116th-congress/house-bill/4027#:~:text=Introduced%20in%20House%20(07%2F25%2F2019)&text=This%20bill%20limits%20a%20social,unlawful%20rather%20than%20merely%20objectionable> accessed 11 June 2022

Pichai S, 'Written Testimony of Sundar Pichai, Chief Executive Officer, Alphabet Inc.', *Does Section 230's Sweeping Immunity Enable Big Tech Bad Behavior?* (2020) Hearing Before the United States Senate Committee on Commerce, Science, and Transportation <https://www.commerce.senate.gov/services/files/894758B6-F55E-471D-904F-480FCD9E4D98> accessed 11 June 2022

S.2972 - A bill to repeal section 230 of the Communications Act of 1934 <https://www.congress.gov/bill/117th-congress/senate-bill/2972> accessed 11 June 2022

S.3538 - EARN IT Act of 2022 <https://www.congress.gov/bill/117th-congress/senate-bill/3538> accessed 11 June 2022

S.4066 - PACT Act <https://www.congress.gov/bill/116th-congress/senate-bill/4066> accessed 11 June 2022

Zuckerberg M, 'Testimony of Mark Zuckerberg Facebook, Inc.', *Does Section 230's Sweeping Immunity Enable Big Tech Bad Behavior?* (2020) Hearing Before the United States Senate Committee on Commerce, Science, and Transportation <https://www.commerce.senate.gov/services/files/E017B34E-F87F-4127-88A7-2C32B6BC3810> accessed 11 June 2022