FIGHTING HATE SPEECH AND TERRORIST PROPAGANDA ON SOCIAL MEDIA IN GERMANY: ‘LESSONS LEARNED’ AFTER ONE YEAR OF THE NETZDG LAW

This paper, part of the Legal Perspectives on Tech Series, was commissioned in conjunction with the Congressional Counterterrorism Caucus

ALEXANDER RITZMANN
SEPTEMBER 2019
About the Program on Extremism

The Program on Extremism at George Washington University provides analysis on issues related to violent and non-violent extremism. The Program spearheads innovative and thoughtful academic inquiry, producing empirical work that strengthens extremism research as a distinct field of study. The Program aims to develop pragmatic policy solutions that resonate with policymakers, civic leaders, and the general public.

About the Author

Alexander Ritzmann is a member of the Steering Committee of the European Commission’s Radicalisation Awareness Network (RAN) and Co-Chair of its Communication and Narratives (C&N) Working Group.

Ritzmann is also a Senior Research Fellow at the Brandenburg Institute for Society and Security (BIGS) in Potsdam, Germany, where he heads the project “Integration at Eye-level”, funded by the German federal government as part of the program “Living Democracy”. In addition, Alexander Ritzmann worked with the European Foundation for Democracy in Brussels, Belgium, in various capacities. From September 2012 to December 2015, he acted as Senior Advisor for the MENA Region and Project Manager for the German Development Cooperation (GIZ) at their Cairo offices. Previously, he was DAAD-Fellow at the American Institute for Contemporary German Studies (AICGS), Johns Hopkins University, in Washington, DC.

The views expressed in this paper are solely those of the author, and not necessarily those of the Program on Extremism or the George Washington University.
Introduction

Over the course of 2015 and 2016, Germany had accepted around one million refugees, most from the Middle East. A right-wing anti-migrant backlash ensued, leading to a dramatic increase of crimes often committed by Germans who had no prior affiliation with extremist right-wing groups. On social media, hate speech proliferated, targeting both refugees and government officials who were deemed responsible for Germany’s open immigration policy. At the same time, the online propaganda and recruitment efforts of the so called “Islamic State” (IS) were at their peak and several IS-claimed terrorist attacks were committed within the European Union and Germany.

In 2016 and 2017, the German federal government initiated an investigation into online activities that violated Article 130 (incitement to hatred and Holocaust denial) and Article 86a (use of symbols from unconstitutional organizations) of the penal code, and violations against the Youth Protection Act. The organization mandated with the investigation reported 200 pieces of content per tested social media company (SMC). Facebook removed 39%, YouTube 90% and Twitter 1%. Looking solely at content removed within 24 hours of being flagged, the rates fell to 31% for Facebook, 82% for YouTube and 0% for Twitter.¹

Realizing that social media failed their own community standards and did not police their networks effectively in respect to illegal activities, Germany’s Network Enforcement Act, or NetzDG law, was introduced on May 16, 2017 and passed several weeks later. The short period for deliberation in the Bundestag was criticized heavily. Federal elections were held in September that year so it appeared that for the governing coalition of the Christian Democrats (CDU) and Social Democrats (SPD), speed was more relevant than deliberation. The law was passed virtually unanimously among the CDU and SPD. The Free Democrats (FDP) and the party “The Left” voted against, the Greens abstained.
Prior to the drafting of the law there were attempts by the Federal Ministry of Justice and Consumer Protection to discuss with and convince the major SMCs to substantially increase their self-policing efforts through “regulated self-regulation.”

NetzDG came into binding effect January 2018. Now, 16 months later, some findings, conclusions and lessons can be drawn. Germany’s NetzDG represents a relevant test case for combatting hate and terrorist propaganda on the internet by law.

Before going into the details of the NetzDG, one of the key questions often raised in this context is: Why is it not a priority for SMCs like Facebook/Instagram, Twitter, Google/YouTube, Pinterest or Soundcloud to proactively and robustly police their platforms regarding all content that is in violation of their own community standards or the laws of the countries they do business in?

The mission statements of major SMCs highlight that they aim at “making the world more open and connected”,2 to “give everyone the power to create and share ideas and information instantly without barriers,”3 and to “organize the world's information and make it universally accessible and useful.”4 Since most SMCs are for-profit enterprises, driven and measured by their ability to increase profits and the value of the company, conflicts between different interests and objectives on how to prioritize resources and investments are inevitable.

For SMCs to justify the allocation of significant resources towards a more effective self-policing (compliance) regime, the incentives of doing so, or the costs of not doing so, must be of higher priority than investing those resources into the growth of the core business. In most cases this means selling community member data to the highest eligible bidder.

Incentives to allocate additional resources towards effective self-policing might include:

- avoiding legislation that might interfere with the business model at large
- preventing reputational damage due to unfavorable media reporting
- reducing the amount of law suits and possible fines
The concept of “regulated self-regulation,” which appears to work with different industries, seems not to have led to the desired results for many governments and policy makers on the EU level and with EU member states. NetzDG and recent regulation on the European Union level is evidence of this.5

Implementing NetzDG: Procedures, Objectives, and Limitations

Who and what is subjected to the law?

NetzDG does not create new categories of illegal content, but provides a binding structure for SMCs for an effective compliance system. In German legal tradition, the contribution to or the support of illegal actions (by providing a platform for example) results in the obligation to remove them. This is called “Störerhaftung,” which also applies to SMCs.67As a result, SMCs need to remove illegal content once they are aware of it.

The law aims at closing the gap between the online and offline world by bundling and enforcing 22 statutes in the online space that already existed in the German criminal code and to hold large social media platforms responsible for their effective enforcement. NetzDG applies to all for-profit social media platforms with at least two million registered users in Germany who receive at least 100 complaints regarding (supposedly) illegal content per year. Direct messaging services like WhatsApp, Telegram and Signal and media outlets are exempt.

The table below, taken from the Facebook NetzDG transparency report for 2018, lists most of the 22 statues and shows the number reported violations.8
This additional, or in Facebook’s case, enhanced compliance system comes at a cost. In 2018, Facebook had a total of 63 employees working on NetzDG related complaints, including two in-house lawyers. In total, Facebook employs 30,000 experts worldwide who work on security and community standards related issues.

The relatively low number of complaints could be a result of Facebook’s decision to create a separate NetzDG reporting form which might be perceived as too complicated by many users.

YouTube’s transparency reports for 2018 show that approximately 465,800 complaints by users based in Germany were received. According to YouTube, 25% of those complaints actually violated the community rules or German laws, resulting in the
removal or blocking of that content. 94% of the reported content that violated community rules or German laws was removed within 24 hours. Most cases related to YouTube’s own rules.¹²

Twitter, like YouTube, made it quite easy to report complaints under NetzDG and received more than 490,000 complaints, accepting only about 10% as being justified. Incitement to hatred, insult and defamation were the most reported illegal acts. Twitter employs a staff of 50 to analyse NetzDG complaints.¹³

**What is the Law Requiring from SMCs?**

SMCs are obliged to set up an “effective and transparent” complaint mechanism. The internal design of this mechanism, however, is up to the company. SMCs must also produce a report every six months on how they have handled complaints, following guidelines laid out in the law. The platforms must designate a domestic point of contact to receive information requests from German law enforcement, to which they must respond within 48 hours of receipt. SMCs have the additional option of setting up a self-regulation body, to whom they can outsource the decision on content, which would have to be approved by the Federal Office of Justice, the agency responsible for the oversight of the implementation of the law (BfJ).¹⁴ A proposal for such a self-regulation body is currently under supervision by the BfJ.

SMCs must delete or block “manifestly” illegal content within 24 hours. What “manifestly” means exactly is part of the controversy. SMCs have up to a week to decide on all other complaints. The seven-day period can be extended if the case remains unclear. The SMCs can contact the user who filed the complaint. If the content is “manifestly unlawful,” platforms must remove it within 24 hours after it being reported. If content is illegal in Germany but acceptable according to the SMC’s community standards, the SMC might only block access to that content from within Germany.
What are Possible Penalties?

Fines can only be imposed for “systematic” breaches of the law, meaning that if companies don’t create and execute an effective moderation and complaints system, they will be held liable. Platforms that fail to comply risk fines of up to €50 million. If the SMCs make an honest mistake in judgment, or overlook an item by error, it faces no liability.

Currently, the Federal Office for Justice has five active investigations into claims that SMCs did not implement an effective moderation and complaints system. One case has been taken to the courts. Two additional SMCs did not publish the transparency reports.¹⁵

NetzDG: A Controversial Law

NetzDG, during its parliamentary readings and after its passing, received criticism from many sides. Its supporters from civil society, for example the Lesbian and Gay Federation in Germany (LSVD) and the Central Council of Jews in Germany, claimed that the list of criminal offences included in the law was incomplete, that the 24-hour limit for deleting “manifestly illegal content” is too long, that the law should apply to all tech companies, not only those with more than two million users and that only individuals with a verified online profile should be allowed to post.¹⁶ Other civil society organisations disagreed and put a focus on the potential limiting of free speech, highlighting that anonymity allows whistle-blowers and other critics to speak without fear of retribution.¹⁷

NetzDG was also criticized for “privatizing” law enforcement, which became one of the key points of controversy. Coming back to the earlier mentioned “Störerhaftung,” it is irrelevant if illegal activities take place during “Oktoberfest” in Munich, in a public square or on social media platforms. The legal entity providing the space where the illegal activity takes place has to intervene as soon as they are (made) aware that illegal activities are taking place.¹⁸
Unimpressed by this legal position, the Federal Association for Information Technology (Bitkom), which has several SMCs amongst its members, claimed the law shifts the responsibility for tackling illegal content away from public authorities and courts to private companies. The Association of the Internet Industry (eco), which has a similar membership, agreed, alleging that the state was abdicating its responsibility.

“Journalist organisations such as Reporters without Borders argued that courts, not social media platforms, should decide on the legality of content. This point is obviously correct in rule-of-law based state. The question here might rather be, who is in charge of creating “facts” and who and who can those decisions be challenged.”

Some of these arguments seem to ignore the point that private companies have the right (and as mentioned the obligation) to police their social media platforms in accordance to their community rules and within the existing legal framework of the country. Like in any other private association, house rules apply.

It also appears that some critics consider large SMCs no longer private entities but rather public spaces. “The legal status of intermediaries like social media services under freedom of speech protection is still unclear. They clearly enjoy freedom of speech protection for their own statements on the platform, but whether the provision of the platform as such and specific functions provided for the users are protected as well is heavily debated. The central question is whether an establishment of a complaint system as required by the NetzDG is encroachment on the fundamental rights of the provider. There is a limited number of types of content where publication is illegal under any circumstances. In any other case, the protection of freedom of speech requires a context-sensitive determination of the meaning of the act of speech. That is especially true for possible infringements of personal rights. Under German constitutional law, there is a complex balancing to be performed when reporting about a person without consent. It is unlikely to encounter any “obvious” case in this field.”
Another line of criticism was that NetzDG would lead to “over-blocking,” meaning that tech companies would take down legal content that could potentially be illegal just to avoid possible legal troubles.

“Since no punishment exists for blocking or deleting legal content in Germany, the platforms would push the delete button to avoid the potential heavy fines. David Kaye, United Nations Special Rapporteur on freedom of opinion and expression, raised further concerns: he argued that decisions about the legitimacy of content would in many cases require an in-depth assessment of the context of speech, something social media companies would not be able to provide. Human Rights Watch and other international critics also opposed NetzDG because, according to them, it would set a precedent for governments around the world to restrict online speech.”

While these concerns seemed plausible at the time of the inception of the law, 16 months after its implementation, the published transparency reports by the SMCs show that NetzDG played a minor role in their decision making. Relating to the earlier point about the core business model of SMCs, over-blocking would result in deleting data (content or user profiles) that could be turned into profits for the company. Hence, over-blocking is against the core interests of the companies, unless penalties or other costs (like reputational risks) are higher than the expected gains.

Another aspect relevant to the discussion about “over-blocking” is that NetzDG does not mandate the SMCs to hear the user posting the content before its removal or blocking. Also, there is no “put back” mechanism in place to reinstate removed or blocked content that was taken down by mistake. Users have to go through the regular court system, which takes several months if not years and costs them between €5000 and €15000.
Expectations and reality:

What can be learned from the German experience?

The BfJ recently stated that it deems NetzDG suitable to increase the effectiveness of actions against illegal activities online. However, the agency also sees potential for the optimization of the law, particularly regarding the reporting mechanisms, the complaint management and the content and structure of the transparency reports. The BfJ is not an independent body but a federal agency that is being directed by the federal Ministry of Justice and Consumer Protection, a fact that also has been criticized.

The BfJ expected to be handling 25,000 complaints per year regarding “over-blocking” as well as regarding cases of supposedly illegal content that has not been taken down. In 2018, the total number of complaints the agency received was below 1,000. The practicability of the reporting mechanisms for users seem to have had an effect on the relatively low number of filed complaints, particularly on Facebook. While Facebook installed a separate reporting mechanism for NetzDG related complaints, YouTube and Twitter have integrated this into their existing system.

NetzDG aimed at a more effective execution of existing laws on social media platforms with more than two million users. So far, it remains unclear if this objective has been achieved. The requirements and proscribed structure of the mandatory transparency reports seems not to be sufficient to generate robust empirical data to support a clear verdict about NetzDG’s success. Amendments to the law could aim at designating a report structure that can provide the necessary empirical data.

In addition, it is hard to predict and measure the full effects of NetzDG simply by focussing narrowly on the number of complaints. This tells little about any potential larger effects of this law on Germany’s information ecosystem or political discourse. “German politicians drew lessons from history to try to protect democracy by curtailing free speech. In the long run, however, they must be careful not to undermine the freedoms embedded in the political system that they seek to protect.”
process that led to NetzDG left some key issues undetermined, like impact assessments on human rights and the right of free speech in particular.

According to the German Association of Judges (DRB), Twitter, Facebook and other SMCs continue to cooperate very little with security authorities in the fight against illegal content online. The expectation that the Network Enforcement Act (NetzDG) would lead to the improved prosecution of criminal offenses, such as incitement to hatred and insult, never came to fruition, according to the Managing Director of the DRB Sven Rebehn. He criticized that "the local offices that social media platforms were required to implement rarely share the names or e-mail addresses of suspects" calls for a "legal obligation of the networks to release a user's data on suspected crime, as is already required for telecommunication providers."28

**Lessons to be learned for the NetzDG experience so far:**

1. **Establish clear reporting and enforcement standards**
The three major social network platforms have each come up with their own approaches towards reporting formulas, making it difficult in some cases for users to flag NetzDG violations. The BfJ highlights the importance of easily recognisable directly accessible and easily understandable reporting mechanism.29

2. **Establish a clearing house for disputed content**
NetzDG offers little recourse to users who believe their content has been wrongfully deleted. The need for a reinstatement procedure ("put back"), where users could appeal to the SMC to have their content reinstated, is obvious. An independent oversight body would be suited best for such a task to avoid mixing business interests and politics on the sensitive issue of free speech.

3. **Implement re-upload filters for designated illegal content**
What often happens after illegal content, in particular extremist propaganda, is being removed, is that the same content is being uploaded several times again. This can lead to the paradoxical situation that an SMC can claim they deleted 99,9 % of illegal extremist content, yet extremists are still able to spread their illegal content. A study by the Counter Extremism Project (CEP) from July 2018 showed that a selection of ISIS videos
YouTube were removed mostly within one hour after detection, yet 91% of them were uploaded more than once. As a result, these videos received 163,391 views. Hence, effective re-upload filters (robust hashing algorithms) need to be deployed to avoid the continuous re-upload of already designated illegal material.
References

7 This legal concept is also part of Section 10 of the Telemedia Act, which is the German implementation of the EU E-Commerce Directive.
9 Ibid.
14 Act to Improve Enforcement of the Law in Social Networks (NetzDG).
FIGHTING HATE SPEECH AND TERRORIST PROPAGANDA ON SOCIAL MEDIA IN GERMANY

https://www.ceps.eu/ceps-publications/germanys-netzdg-key-test-combating-online-hate/


