A HOOVER INSTITUTION ESSAY

# Roles and Responsibilities of Information Intermediaries

FIGHTING MISINFORMATION AS A TEST CASE FOR A HUMAN RIGHTS–RESPECTING GOVERNANCE OF SOCIAL MEDIA PLATFORMS

**WOLFGANG SCHULZ**

Aegis Series Paper No. 1904



## Introduction

The pendulum is swinging back. While many countries have initially opted to give online platforms a safe harbor to eschew liability for third-party content, we are now witnessing the reverse trend to weaken that protection. In Europe, this includes the creation of regulatory regimes specifically dealing with social media platforms and other information intermediaries as well as soft law instruments.[1] This is to a large extent driven by the (felt) need to deal with phenomena like "misinformation," which might have—among other negative effects—the potential to erode the integrity of elections and undermine the very basis of our democracies.[2] Those intermediaries seem to be the only actors capable of dealing with the scale and the speed of the problem.[3] Furthermore, the debate is driven by a growing awareness of their (potential) influence on public opinion-making.[4]

Platforms' control of the chokepoints of online speech make their content moderation practices especially dangerous for freedom of speech and freedom of information.[5] Speech can be made imperceptible before there can be any debate about the content. Overblocking of content is always a possibility, and it can go unnoticed. Regulatory attempts like the German Network Enforcement Act (NetzDG), which place responsibility on big platforms to restrict speech on states' behalf, are therefore criticized even by those who feel the need to act against harmful content.[6]

It is, however, not fair to blame policy makers alone for the persistent lack of an adequate and balanced approach to intermediary governance. Even in academic research, the functions of intermediaries are not yet fully understood, nor are the potential effects of the various forms of regulation.

Against this background, this article will discuss the structural challenges posed by online misinformation compared to traditional offline misinformation; the conceptual

difficulties of using information intermediaries as a basis for rational governance; recent European policy approaches vis-à-vis communication on platforms and their shortcomings; and suggestions from a European perspective for a path towards a human rights–respecting governance of information intermediaries.

## Misinformation as a Test Case

**Defining misinformation** Following the definition of the *Oxford English Dictionary* (*OED*),[7] *misinformation* denotes "wrong or misleading information." *Disinformation* is also wrong information but unlike *misinformation,* it is a known falsehood, whose active diffusion is intentional. The *OED* defines *disinformation* as "the dissemination of deliberately false information" and refers specifically to wrong information supplied by governments. The term "fake news" is rarely used in academic research anymore since it is imprecise and since the reference to "news" can be understood to cover only false information in the shape of traditional media reporting.

Of course, when we speak about misinformation, we have to bear in mind that reality is a mediated construction.[8] Apart from verifiable scientific facts, most of what we think of as true is (only) discursively constructed. Our understanding of social reality depends to a large extent on dominant discourses and not exclusively on "hard facts."[9] In other words, there is no single "truth," considering that different people see the world in different ways. This fluid and context-dependent state of truth puts us into a quandary in the context of online misinformation.

There is indeed a risk in constructing a universe where any attempt to objectively define truth is destined to fail and where social reality is formed by hegemonic discourses and power structures.[10] Nevertheless, we should not lose sight of the value of evidence and facts, even when we are discussing nonscientific social reality. We cannot allow truth to be reconstructed on dubious and potentially harmful grounds, as attempted by the current US president and other "post-truth" politicians, who spin misleading narratives to shore up voter support. When we talk about truth and misinformation, our baseline should therefore be, to the extent possible and depending on the nature of the statement in question, a correspondence with fact-based reality.[11]

This article uses misinformation on social media platforms as a test case, but the arguments developed should inform the broader debate about creating an adequate governance framework for information intermediaries.

**What makes online misinformation different?** Misinformation is by no means a new phenomenon. Long before the digital era, tabloids and other media formats spread falsehoods, construed as propaganda, gossip, or conspiracy theories. Still, the ubiquity of the internet has enabled countless new forms of expression that were unavailable before. This transformation has fundamentally altered the ways in which information is produced and received. Statements of individual speakers can now be amplified to reach millions of users across the globe.

Alongside numerous opportunities, this development has also spurred the spreading of harmful content. Misinformation is particularly virulent. Conspiracy theories and false allegations spread around like-minded groups on social media and poison the digital environment.[12] In addition, targeted online disinformation campaigns by bad actors sow mistrust and increase polarization in society.[13] Such manipulation of the public opinion has the potential for destabilizing democratic processes.[14]

The challenges posed by misinformation are exacerbated by the specific characteristics of online speech. The first aspect is acceleration. While some content on a social media platform might never be noticed by anyone apart from its author, other content can go viral in no time. Therefore, governance mechanisms need to work swiftly, since it is difficult to locate and deal with misinformation after it has been widely distributed and copied.[15]

Volume is another obvious factor. The sheer quantity of information has become a structural challenge.[16] This means that proposals to have national courts—or any private adjudicatory body, such as Facebook's new oversight board[17]—adjudicate all content-related conflicts among users or between user and platform are simply not feasible.[18]

Additionally, it can be difficult to find out the real identity of speakers online, which is necessary to file a claim and, at times, even necessary to check the veracity of a statement. Even though some platforms have a clear name policy, anonymity is still common online and a central characteristic of internet communication. The exact source of online misinformation is also often unclear, so attribution is generally fuzzy.[19]

The attribution problem is accompanied by weak social control. The operability of legal concepts like privacy and its protection depends largely on internalization of the relevant norms by the individual. Social norms and social control—not law

enforcement—are essential. Those mechanisms are distorted online, especially when the speakers have not met each other in "real life." Research indicates that when people communicate anonymously, they are more uninhibited—and ultimately more hurtful and dangerous.[20]

Furthermore, online platforms often allow users or even the general public to observe communication among groups or between individuals. This has enabled the business models of rating platforms, but it also creates challenges: observers often do not know the context or other crucial information that is necessary to understand the speakers' motivations. Many of the problems with online communication stem from the fact that the talk at the "regulars' tables" is now visible to everyone, with all its exaggerations, polemics, and stereotypes.[21]

While people initially worried about the internet being a legal vacuum, it has become the opposite. Jurisdictional uncertainties are a major challenge, which can hinder effective law enforcement online.[22] For instance, content that is criminal in Germany can be safely hosted on servers in the United States, where such material is protected by the First Amendment.[23]

It is now a popular saying that the internet never forgets. That is true in the sense that there is not yet a way to ensure that a piece of information is deleted from all servers, wherever they are located.[24] The so-called right to be forgotten only provides a data subject with a claim against a certain controller.[25] While it can help make information about the subject harder to find in searches, it would be a mistake to think that the information can really be expunged.[26]

In sum, these factors further exacerbate the problems posed by misinformation.[27] The actors who can address these challenges in the most efficient way are the internet intermediaries, whose technological capacities to restrain harmful content are in many ways superior to those of states. This is certainly one reason why several European governments have introduced regulation of intermediaries targeting misinformation, and many others are considering doing so. But before we turn to intermediaries, it seems worthwhile to look at the role of the traditional media in the context of misinformation.

**The role of the media** Traditional journalistic media still plays a crucial role in informing the public and tackling misinformation.[28] Research has established that journalistic

media—even though it is itself changing and blurring traditional boundaries—still fulfills a specific role for the self-observation of societies. Benkler, Faris, and Roberts have demonstrated that the role of the media in the US elections of 2016 was significantly greater than individual attempts at misinformation.[29] Ideally, the news industry can mitigate the effects of misinformation by providing an antidote to fake stories and trolling. Strong professional journalism can correct falsehoods and build public trust, although this is becoming increasingly difficult in today's polarized societies.[30]

Given this urgent need for high-quality journalism, the current state of the news media is worrisome. Traditional media formats face overwhelming competition in the advertisement market from nonjournalistic online players, who have nearly dried out their revenue streams. Depleted newsrooms and the dwindling of local newspapers and TV stations are the consequence.[31] To survive in the contest for peoples' attention, traditional media actors are increasingly producing sensationalist and click-bait content, which in turn undermines trust in the news media.[32]

In addition, the constraints of time-consuming professional reporting collide with the speed of online misinformation. The window of opportunity for dealing with misinformation, i.e., to correct falsehoods or debunk rumors before significant social effects can occur, is closing ever more rapidly. Consider the difference: In the midst of the 1800 US presidential election campaigns of Thomas Jefferson and John Adams, the rumor spread that Jefferson had suddenly died. News reports that clarified the situation traveled slowly then, but so did the rumors. Today, the pace has picked up dramatically. Unlike in the times of Adams and Jefferson, online misinformation can spread virally in a matter of hours. By the time professional journalists have picked up the issue, it is often too late, and the damage is done.

## Understanding Internet Intermediaries

In order to address misinformation in a meaningful way and to design adequate regulatory responses, an understanding of internet intermediaries is crucial. They have become extraordinarily powerful institutions in the networked digital sphere. Performing a crucial function as the internet's "middlemen," they facilitate online interaction and dissemination of information between different parties by offering a variety of services.[33]

Various activities and business models fall under the broad category of internet intermediaries. Most of them carry out several functions in parallel. Notably,

intermediaries may moderate and rank content, using algorithms designed to increase profits. This allows them to influence users' access to information online in ways comparable to those of traditional media.[34] While their importance is obvious, there are structural difficulties in conceptualizing intermediaries for governance purposes. Their ambiguous nature challenges traditional governance techniques like defining a type of service and attaching rights and obligations to the so-defined entities. This also leaves room for political "fights about narratives": for example, while lobbyists for publishing houses tend to frame intermediaries as distribution platforms for their own journalistic content, the platform companies themselves tend to portray their services as technical mere tools for users (or at least want politicians to see them that way). This section considers these conceptual challenges for regulators.

**No necessary link between services and communicative practices** Even with traditional media, there is no necessary link between the type of media and the communicative practices of the audience. You can be educated by comedy shows and you can be entertained by news shows. When it comes to intermediaries, the link between their intended use and their actual use is further attenuated. Take search engines as an example. You can use search tools to find a website whose URL you cannot remember, but you can also "ask" a search engine a question like, "What has the president done today?" You may even use it as a spell-checker and for other purposes that the search engine provider might not have foreseen. This is an observation with relevance not limited to media studies, given that the ability of intermediaries to influence public opinion depends on and varies according to those communicative practices. It is therefore crucial that policy makers assess the risks stemming from platforms' various affordances when they design regulation.

**Hybrid nature of services** Another obstacle in designing adequate regulatory solutions is that many information intermediaries are hybrids. For example, one aspect of Facebook's activity is merely providing a platform for hosting and sharing content that users upload. Simultaneously, another aspect of its activity—the newsfeed—can be seen as an information service that at least partly fulfills functions comparable to traditional news media. It has even become an important means of distribution for media content. This complicates the application of traditional concepts of regulation, since one might have to simultaneously treat the different subservices differently: whereas the curation of content could be conceptualized as a form of editorial control under media law, for instance, this approach would be pointless when it comes to regulating other, content-neutral functions.

**Differentiated functions for public communication** The recent Reuters Digital News Survey reveals the importance of Instagram for news consumption in Germany, to take an example. Instagram is the social medium most frequently used to get access to news by online users in the age group of 18–24.[35] If one looks at older target groups, however, traditional television is still by far the most important source of news, and social media plays a secondary role.[36] Considering this diversity of usage patterns, which is unlikely to change in the foreseeable future, it is difficult to develop a generalized understanding of the role of intermediaries and to develop regulatory concepts. By way of example, emphasizing intermediaries' role in informing the public and obligating them to prioritize trusted news sources may have an impact with respect to younger users, but would be of limited use for older user demographics.

**Algorithmic sorting and ranking are not fully understood** A further challenge arises from the fact that many functions of intermediaries are based on algorithmic calculations, which sort or select information. These algorithms have various effects, such as content amplification. The current discussion about errors related to decision making by and in connection with algorithmic decision systems shows the range of problems. We still lack definitive knowledge about how errors or biases in the training data for machine learning systems impact the algorithms themselves, and, subsequently, human decision making.

In sum, policy makers and scholars are faced with a twofold challenge: On the one hand, online misinformation presents substantial problems, which I have discussed in the previous section. On the other hand, any attempt to address these challenges through platform regulation is additionally burdened with the structural difficulties in conceptualizing internet intermediaries.

## European Approaches to the Governance of Internet Intermediaries

In this section, I will discuss recent regulatory approaches by European policy makers to tackle online misinformation on platforms. This analysis focuses on the governance *of* platforms, that is, the regulation of platforms by states, as opposed to the governance *by* platforms, that is, platforms applying their own standards to the communication of their users. Of course, these two components are intertwined, and governance of intermediaries might lead to new forms of governance by intermediaries.

### *Reasons for Safe Harbor Rules*

The initial approach to platform regulation by many states was to shield internet intermediaries from liability—an approach that may now seem counterintuitive. Given intermediaries' importance for online communication, policy makers recognized early on that holding intermediaries liable for illegal activities by third parties might significantly inhibit the free flow of information. They feared that platforms might overblock content in order to avoid liability.[37] In response, key jurisdictions—the United States and the European Union—adopted so-called safe harbor rules that have become a cornerstone of today's platform economy.[38]

In the European Union, the safe harbors are enshrined in the e-Commerce Directive,[39] which shields intermediaries from liability for third-party content. The motivation behind the safe harbor clause in the e-Commerce Directive was not so much to protect freedom of speech as it was to allow the European IT sector to grow without having to fear incalculable liability risks. Platforms such as Facebook or YouTube are granted immunity under Article 14 of the e-Commerce Directive if they meet specific conditions. To benefit from the safe harbor protection, they need to expeditiously remove or block access to unlawful content once they become aware of it. Article 14 of the e-Commerce Directive has led to the development of notice and takedown procedures, but it does not regulate those procedures in detail.

Compared to its US equivalent, Section 230 of the Communications Decency Act (CDA), the immunity afforded by the e-Commerce Directive is more limited, since the former shields intermediaries from liability even if they have positive knowledge of unlawful content on their platforms. Unlike Section 230 of the CDA, the e-Commerce Directive also only insulates platforms from monetary liability; it does not affect court injunctions to take down content. Many countries have adopted similar regimes, although they differ in their scope of application and in how much immunity they grant to information intermediaries.

### *Towards Intermediary Responsibility?*

In recent years, regulators in the European Union have called the liability protection afforded by the e-Commerce Directive into question.[40] Given the proliferation of misinformation and hate speech online, European policy makers have increasingly pressured intermediaries to assist them in combating harmful speech.[41] They have also

started to reduce the immunity granted to platforms in specific sectors, for instance by adopting regulation in the copyright sector.

Intermediaries are tempting targets for government regulation because they control the important chokepoints of online information. Moreover, their capacity to restrain illegal activities on their services is in many ways superior to that of states. They can block access to certain content or eliminate misconduct by suspending the accounts of wrongdoers, regardless of their anonymity.[42] Unlike most states, big platforms also have powerful content-recognition technologies at their disposal to identify infringing content before it spreads around the web.[43]

Platform regulation therefore allows governments to, at least indirectly, set the rules for online speech. This raises hard questions as to how much responsibility should be placed on platforms to implement speech control on behalf of states. On the one hand, delegating the enforcement of rules set by states to intermediaries is seen as the most efficient, if not only, way of maintaining control over online speech. On the other hand, enlisting private entities as "proxy censors" creates new problems, especially from a human-rights perspective.[44]

### *New European Approaches to Tackle Misinformation*

Unfortunately, there is no easy fix to misinformation online. Calls for government regulation are understandable, but they can create new problems. As I show in this section, any attempt to curtail misinformation carries risks of unintended collateral effects, especially regarding the right to freedom of expression.

In the following, I discuss two recent regulatory approaches for addressing misinformation, which represent two broader trends in European intermediary regulation: First, I will briefly examine the EU Code of Practice against Disinformation as an example of using "soft law" instruments to achieve public-policy objectives. Second, I will focus on the German NetzDG as an example of a novel law specifically tailored for intermediaries.

**"Soft law" approach: The EU Code of Practice against Disinformation** Faced with growing concerns about the impact of online misinformation on European election campaigns and democracy in general,[45] the EU Commission negotiated a voluntary Code of Practice against Disinformation with large platform companies, including

Google, Facebook, and Twitter, and representatives from the advertising industry in September 2018.[46] The signatories agreed to implement a wide range of commitments to fight disinformation that is "created, presented and disseminated for economic gain or to intentionally deceive the public" and that "may cause public harm." Depending on the signatories' various modes of operation, the commitments include measures to increase transparency in political advertising, efforts to reduce the revenues of commercial distributors of misinformation, and the suspension of fake accounts.

After first reports by the signatories were submitted in early 2019 to document their implementation of the Code, the EU Commission welcomed the progresses made.[47] Yet it urged the platform companies to "develop a more systematic approach to enable a proper and regular monitoring and assessment, on the basis of appropriate performance data."[48] The commission also announced a comprehensive assessment of the code's initial twelve-month period, stating that, "should the results prove unsatisfactory, the Commission may propose further actions, including of a regulatory nature."[49]

The Code of Practice was no small feat. The commission successfully encouraged stakeholders from the private sector to agree, for the first time and on a voluntary basis, to reduce the spread of online disinformation.[50] At the same time, human-rights activists have voiced concerns over this "soft law" approach, arguing that it pressures platforms into removing content without meaningful safeguards such as an appeal system or review of the takedown decisions made under the Code.[51]

**New laws for intermediaries: The NetzDG as an example** Germany's NetzDG, short for Network Enforcement Act, is an instructive example of the second regulatory approach—specific laws governing information intermediaries. Even though the act does not address misinformation directly, it was designed in part to solve the problem of misinformation online.

*Genesis of the act* After an influx of refugees in 2015 created a massive backlash from right-wing populists, the German government was concerned that hate speech and misinformation could influence the 2017 Bundestag election campaign.[52] One instructive example of harmful misinformation that received much attention in Germany involved a selfie that a Syrian refugee took with Angela Merkel. Right-wing groups used the photo to falsely assert that the refugee was involved in terrorist

activities. That fake news went viral. Some of the posts that used the word "terrorist" relating to the refugee were removed by social media platforms as harmful content. These takedowns were not only aimed at right-wing trolls, however. Equally affected were users who referenced the fake news in order to critically reflect on hate speech.

Germany's Ministry of Justice initially encouraged big platforms to voluntarily improve their complaints mechanisms. After reviewing the measures taken, the German government was not satisfied with the platforms' effort at self-regulation. According to the Ministry of Justice, Facebook in particular reacted too slowly to complaints about content; Twitter likewise generally had low response rates.[53]

Against this background, the government hurriedly introduced a draft bill to impose binding rules on platforms. Although experts roundly criticized its approach, the ruling coalition pushed the NetzDG through parliament.[54] It came into force on October 1, 2017, and has been fully applicable since January 1, 2018.[55]

*The NetzDG's regulatory concept* The NetzDG covers social network providers; to exempt start-ups and small and mid-size enterprises (SMEs), the act imposes some obligations only on platforms with at least two million users. The main provision of the NetzDG stipulates that large providers with more than two million users must maintain an effective and transparent procedure for handling complaints about unlawful content. They have to remove or block access to content that appears to be "manifestly unlawful" within twenty-four hours after a complaint has been filed. Other unlawful content has to be taken down within seven days. This review period may exceed seven days in case more time is required to reduce "overblocking," or when providers refer the decision to an independent co-regulatory body.[56] Social network providers that receive more than one hundred complaints per year also have to produce transparency reports every six months.

The act does not include new definitions of hate speech or misinformation. Instead, it refers to existing definitions under the German criminal code. Content is considered illegal under the NetzDG if it falls within the scope of an exhaustive list of speech-related criminal offenses. Several of these offenses are aimed at protecting public safety (such as incitement to hatred), while others are aimed at safeguarding individual rights (such as slander and defamation).[57] While earlier drafts included the obligation

to properly handle each individual case, the final act only requires a functioning complaints system.[58] The Federal Office of Justice, an administrative body directly subordinated to the Ministry of Justice, oversees compliance with the NetzDG. If providers systemically fail in their obligations under the NetzDG, they face a fine of up to 50 million euros. So far, no fines have been imposed.

*Basic critique* On the whole, the NetzDG is an example of good intentions falling short. Its vague wording has been sharply criticized; for instance, there is uncertainty as to what constitutes "manifestly unlawful" content.[59]

There is also good reason to argue that the NetzDG conflicts with the e-Commerce Directive.[60] Proponents of the act maintain that there is no contradiction between the instruments, since the NetzDG merely enforces existing obligations to remove unlawful content upon notification.[61] However, as previously mentioned, Article 14 of the e-Commerce Directive does not impose tight time frames, whereas the NetzDG does. Under the directive, platforms are only required to act expeditiously. The time frames imposed by the NetzDG might therefore counteract the e-Commerce Directive's objective to harmonize cross-border services,[62] since they lead to divergent procedures to establish intermediary liability across the European Union.[63]

In addition, the ramped-up efforts of big platforms to tackle harmful content have caused users with extreme views to migrate to smaller platforms that lack the resources to moderate content that major players such as Facebook or YouTube have.[64] In effect, misinformation still circulates around the web, but seems partly to have moved to smaller platforms that are not subject to the NetzDG.

Another point of criticism is that NetzDG enforcement is not detached from politics. As mentioned already, the Federal Office of Justice supervises the NetzDG's implementation and directly reports to the minister of justice. That institutional setup is especially notable in Germany, where state-controlled media evokes memories of the Nazi dictatorship and the German Democratic Republic. Therefore, the Federal Constitutional Court regards the independence of media regulation from state interference as an eminently important principle.[65]

These are only the broad lines of criticism, which explain why many scholars and activists maintain that the law was poorly crafted. On a more general level, the

NetzDG's many weaknesses indicate that targeting intermediaries to control speech may not be the silver bullet after all.

*Debate on overblocking and recent developments* The NetzDG has initiated a heated debate as to whether the law encourages overblocking. Critics argue that the NetzDG creates a strong incentive to systematically take content down rather than leave it up.[66] There is indeed a good case to be made that this regulatory approach leads to the excessive removal of content, given that the law forces platforms to make decisions about the lawfulness of content within tight time frames, under threat of substantial fines.

This argument is lent further weight by the fact that the NetzDG obligates intermediaries to make highly context-dependent decisions at scale. To provide some perspective, YouTube reportedly received more than 160,000 NetzDG-related complaints between July and December 2018 (the number of pieces of content flagged under the company's private content policies is likely much higher).[67] Moreover, intermediaries can only assess the information available on the platform in order to understand the context. They lack the adversarial evidentiary process that courts have at their disposal.

There is still insufficient data to empirically substantiate the concerns regarding overblocking, however. As of now, the extent of overblocking—if any—and a potential chilling effect is hard to measure. Although big platforms have released three rounds of biannual NetzDG transparency reports so far, the reported numbers are inconclusive regarding overblocking, since they only reveal the number of complaints and the actions taken. In addition, there appears to be a significant divergence between the number of NetzDG complaints received by Facebook, Twitter, and YouTube. While Twitter[68] and YouTube[69] reported 264,818 and 214,827 complaints, respectively, from January to June 2018, Facebook only counted 886 complaints during the same period[70] (the number dropped even further to just 500 complaints for the period from July to December 2018[71]). This is mostly due to the fact that platforms appear to prioritize their own private content policies over the NetzDG to varying degrees.[72] Facebook especially makes it more difficult for users to file a NetzDG complaint than a complaint under its private framework.[73] Since content removed under these policies does not fall within the scope of the NetzDG transparency obligations, the effect of the NetzDG is not really visible.[74]

To overcome the lack of available empirical data,[75] platforms need to open their data troves to external researchers to facilitate a more informed discussion about the actual effects of harmful content and the impact of content regulation. What has been described as Facebook's recent "Glasnost moment"[76] is a welcome development in this regard.[77] Meanwhile, the NetzDG has been imitated elsewhere, including in nondemocratic countries. Most notably, Russia passed anti–fake news legislation modeled on the NetzDG, with explicit reference to the German law in the *travaux préparatoires*.[78]

## A European Perspective on Human Rights and Platform Regulation

As discussed above, the recent regulatory attempts to deal with misinformation in Europe raise grave concerns given their broad scope, their restriction of public and individual communications, and their lack of procedural safeguards. While such cases of (enforced) private censorship are not commonly framed as a human-rights issue in the United States, these risks have human-rights implications from the perspective of European fundamental-rights doctrine.

The right to freedom of expression is protected in Europe under Article 10[79] of the European Convention on Human Rights (ECHR).[80] In addition to a focus on state interferences, similar to the First Amendment state-action doctrine, it also imposes positive obligations on states to protect citizens' freedom of expression vis-à-vis private entities.[81]

**General threats to freedom of speech** Against this normative background, the NetzDG in particular does not strike a good balance between its objective of countering harmful content and protecting online expression. Considering that the act is likely to make platforms remove lawful content, there is a good case to be made that it infringes on freedom of speech. Moreover, besides curtailing lawful speech, overblocking likely has an overall chilling effect on users' (future) exercise of their freedom of expression.[82]

**Doctrinal difficulties** There remain many open questions regarding the level of protection afforded by ECHR Article 10 in the context of misinformation. In particular, the private nature of intermediaries and the fact that regulation such as the NetzDG does not directly target users complicate the analysis, given that under classical liberal doctrine, only state actors can interfere with fundamental rights.[83]

This complexity is further increased by the multitude of actors whose communication freedoms are potentially affected: (1) the "victim" smeared by misinformation on the platform, who may also be the complainant; (2) the provider of the social media platform; (3) the creator of illegal content that is taken down; (4) the author of lawful content that is (wrongly) taken down; (5) the intended recipients of the removed content.[84]

Among these actors, the fundamental-rights protection of intermediaries is particularly unclear. It remains an open question whether regulatory efforts to address misinformation might also infringe on platform companies' own right to freedom of speech (besides the right to do business).[85] Whereas they clearly enjoy protection for their own statements, it is heavily debated whether the provision of a platform as such and of its specific functions are protected as well. This is especially relevant to the curation of social media content, such as Facebook's news feed, since the editorial decisions taken in this context might also constitute protected speech.[86] Further, intermediaries enable or at least facilitate the communication of others and may indirectly fall within the scope of free-speech guarantees for this reason. An adequate fundamental-rights analysis would have to differentiate between the various functions performed by intermediaries described above.

**Potential justification** A last aspect I would like to stress in the context of the NetzDG is how potential interferences with fundamental rights could be justified under ECHR Article 10. Under this framework, interferences can be justified under certain conditions. Their legality is contingent on a proportionality test, which takes into account whether the legislation "is necessary in a democratic society." In most cases, the condition is met if coercion by the state is necessary to prevent harm to others.[87] This harm principle, tracing back to John Stuart Mill, could justify content regulation if it prevents users from being subjected to harmful speech.

There is reason to be skeptical about the existence of such a justification regarding the NetzDG. It is noteworthy that the official explanatory memorandum for the NetzDG does not even directly refer to the harm principle. It instead refers to the necessity to maintain a culture of political debate. This is understandable in light of the far-right populist movements at the time of the NetzDG's inception. Nevertheless, a desire to protect the civility of political debate is insufficient for limiting fundamental rights; it instead turns the NetzDG into an instrument of a purely moral nature. At least as long

as there are no grave structural risks for a free and open political debate, it is not the role of the state to govern how public opinion is formed in society.[88]

## Toward a Human Rights–Respecting Approach

Going forward, it is crucial that we develop regulatory approaches to misinformation that respect freedom of speech.

**Council of Europe's Standards on the Roles and Responsibilities of Intermediaries as a starting point**[89] Apart from the work of the UN Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression,[90] the Council of Europe's Standards on the Roles and Responsibilities of Intermediaries[91] is the most comprehensive attempt to create a human rights–oriented framework in the European context.

The Council of Europe (COE) acknowledged the possibilities for communication and information access opened by information intermediaries but also highlighted their role in spreading potentially harmful speech. The COE recommended a functional approach that takes the above-mentioned hybrid nature of many intermediary services into account.

Under the COE framework, it is primarily the obligation of states to make sure that laws, regulations, and policies applicable to internet intermediaries effectively safeguard the human rights and fundamental freedoms of users. At the same time and in line with the UN Guiding Principles on Business and Human Rights,[92] internet intermediaries have the responsibility to respect the internationally recognized human rights of their users and of third parties affected by their activities. States and intermediaries therefore need to cooperate in safeguarding freedom of expression.

As for regulation by the states, the COE reaffirms that any limitation of human rights requires a clear legal basis, and that the process of enacting legislation applicable to internet intermediaries should be transparent and inclusive. This poses challenges to systems of coregulation, which—as the EU Code of Practice against Disinformation instructively shows—are often predicated on nonbinding commitments by private companies. Shifting the responsibility to the companies cannot release the state from its fundamental-rights obligation, however; at least in the case of coregulation, an interference with freedom of expression by intermediaries may be attributed to the state.

In order to ensure that content regulation meets the conditions foreseen in ECHR Article 10, state authorities should therefore carefully evaluate the prospective (unintended) impact on freedom of expression and should opt for the least intrusive means. Such a human-rights impact assessment would force regulators to preemptively address adverse consequences. State authorities should also ensure that effective redress mechanisms are made available and adhere to applicable procedural safeguards.

**Differentiated approach to take context into account** As the example of the NetzDG shows, there is reason to worry that content regulation disproportionally curtails protected speech, since it forces intermediaries to make highly context-sensitive decisions within tight time frames and based on insufficient available information. The heavy reliance on (semi-)automated flagging systems further exacerbates the problem. Despite all the progress in the field of automated content recognition and intelligent speech processing, these systems do not (yet) have the cognitive capabilities of human reviewers. They are still error-prone and cannot reliably detect irony or other double meanings.[93]

A human rights–oriented approach to intermediary governance therefore needs to take both this state of the art of automated content moderation and the complex nature of misinformation into account.[94] Of course, some content should still be taken down as quickly as possible, despite fundamental-rights concerns. These should be limited to severe cases, such as misinformation that may cause an imminent risk of grave harm, for instance where online incitement based on false allegations may spill over into the offline world and mobilize violent mobs against vulnerable groups.

In all other cases of less grave misconduct, it is necessary to establish a procedure that allows for a sufficient consideration of individual pieces of content, including by human content moderators. Context is obviously crucial here. The same material may either be lawful or unlawful depending on the respective circumstances. For instance, a tweet that violates the law may be repeated elsewhere as satire or as a critical discussion. The latter can only be recognized as a lawful expression after close inspection of the speaker's intention.[95]

Moreover, we need to differentiate between content decisions that the intermediary is able to make based on the information available on the platform alone and other content decisions that require further input. The latter could require cooperation with

external fact-checkers or a procedure that allows the parties involved to be heard. For instance, there is evidence that the platforms have experienced more difficulty in reacting to alleged acts of insult, slander, and defamation than responding to other criminal offences under the NetzDG, such as incitement to hatred. This might be because the former category of offences is more context sensitive, and deciding related disputes is often only possible after hearing both sides. Furthermore, there is a thin line between those acts and offensive but legal speech.

**Additional measures** Considering the complexity of the challenges posed by misinformation, an adequate governance approach will require a wide range of complementary measures. In the following, I outline some of them (naturally, all these proposals have costs and benefits that require further analysis to examine their advantages and potential shortcomings):[96]

- With regards to misinformation, a system of fact-checking and labeling of questionable content might be part of the solution. We also need more cooperation with platforms to be able to research the effects of fact-checking efforts. This could be done in a much more granular way than it is done now.

- There should be user-friendly tools to report content that users think is misleading. This especially pertains to digital election campaigning.

- Further, we need systemic support for people who fall victim to misinformation or hate speech. In particular, journalists, who are often massively harassed online, need better protection.

- Platforms should also be encouraged to design and implement instruments of dispute resolution, so that conflicts can be solved between the parties themselves.

- Users need to be able to distinguish between quality online journalism and material coming from dubious sources. This could involve online tools developed by self-regulatory bodies that assist users in that regard.

- Platforms should also be more transparent about how they curate content, so users have a better understanding of why they are exposed to certain information and can appraise information more critically.

- Another component of a comprehensive approach is to enhance knowledge about counterspeech, for instance by teaching techniques on how bystanders can be

mobilized to de-escalate in an online conflict. Extensive research about various approaches and their effects already exists.[97] Put into practice, it could help to limit dangerous speech.

- Traditional law enforcement needs to be brought up to speed in the digital era, so perpetrators and victims alike do not feel that speech-related crimes go unpunished online. Content regulation may complement criminal prosecution, but it must not replace it.

- At the same time, intermediaries should minimize financial incentives for those who seek to profit from misinformation.

- Finally, fixing the problem of misinformation and hate speech goes hand in hand with strengthening civil society and the media. Given the opaque business models of many platforms, we need a high level of scrutiny from journalists and civil society organizations. A healthy information ecosystem requires professional journalism with sufficient budgets and civil society actors committed to the public interest.

## Conclusion

Unfortunately, discussions about an appropriate regulatory response to misinformation often center on issues such as politics or the overt influence of big (US) tech companies, which are only tangentially related to the actual problem. This is unfortunate, since what is really at stake is the unmanipulated formation of public opinion and the freedom of online speech. This does not mean that curbing companies' unchecked power and challenging "post-truth" politicians are not legitimate concerns. In order to rationally dissect the issue of misinformation, however, we need a more nuanced discussion that differentiates between different types of content (including the basis of their context dependency), respects the proportionality principle, and seeks to develop a clear and transparent separation of the responsibility of states and of platform providers.

## Acknowledgments

The author wants to thank Alexander Pirang for many valuable suggestions and extensive support.

## NOTES

1 See Jack Balkin, "Free Speech Is a Triangle," *Columbia Law Review* 118, no. 7 (November 2018): 2015*ff.*; David Kaye, "Report of the Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression," Human Rights Council, United Nations General Assembly, April 6, 2018, 12*ff.*

2 See Wolfgang Schulz, "Regulating Intermediaries to Protect Privacy Online: The Case of the German NetzDG," in *Personality and Data Protection Rights on the Internet*, ed. Marion Albers and Ingo Sarlet (New York: Springer, forthcoming), 5.

3 See Jack Balkin, "Free Speech in the Algorithmic Society: Big Data, Private Governance, and New School Speech Regulation," *UC Davis Law Review* 51, no. 3 (February 2018): 1175.

4 See Schulz, "Regulating Intermediaries," 11.

5 See Jack Balkin, "Old-School/New-School Speech Regulation," *Harvard Law Review* 127, no. 8 (June 2014): 2297.

6 See, e.g., Bernd Holznagel, "Legal Review of the Draft Law on Better Law Enforcement in Social Networks," Organization for Security and Co-operation in Europe, May 2017, https://www.osce.org/fom/333541.

7 *Oxford English Dictionary*, http://www.oed.com.

8 See Nick Couldry and Andreas Hepp, *The Mediated Construction of Reality* (Cambridge: Polity, 2016).

9 See Anastasia Degliaouri, "Discursive Construction of Truth, Ideology and the Emergence of Post-Truth Narratives in Contemporary Political Communication," *International Journal of Media, Culture and Politics* 14, no. 3 (September 1, 2018): 302.

10 See Degliaouri, "Discursive Construction," 303*ff.*

11 See Degliaouri, "Discursive Construction,"302*ff.*

12 See Cass R. Sunstein, *#Republic: Divided Democracy in the Age of Social Media* (Princeton, NJ: Princeton University Press, 2017).

13 See Cherilyn Ireton, "Truth, Trust and Journalism: Why It Matters," in *Journalism, "Fake News" & Disinformation. Handbook for Journalism Education and Training*, ed. Cherilyn Ireton and Julie Posetti, UNESCO Series on Journalism Education (Paris: United Nations Educational, Scientific and Cultural Organization, 2018), 32.

14 See "Disinformation and 'Fake News': Final Report," report of the Digital, Culture, Media and Sport Committee, House of Commons, United Kingdom, Session 2017–2019, February 18, 2019, 68*ff.*, https://publications.parliament.uk/pa/cm201719/cmselect/cmcumeds/1791/1791.pdf

15 See Schulz, "Regulating Intermediaries," 2.

16 Schulz, "Regulating Intermediaries," 2.

17 Mark Zuckerberg, "A Blueprint for Content Governance and Enforcement," Facebook, November 15, 2018, https://www.facebook.com/notes/mark-zuckerberg/a-blueprint-for-content-governance-and-enforcement/10156443129621634.

18 Schulz, "Regulating Intermediaries," 2.

19 See Schulz, "Regulating Intermediaries," 3.

20 See Russell Haines, Jill Hough, Lan Cao, and Douglas Haines, "Anonymity in Computer-Mediated Communication: More Contrarian Ideas with Less Influence," *Group Decision and Negotiation* 23, no. 4 (July 2014): 765–86.

21 See Schulz, "Regulating Intermediaries," 2.

22 Schulz, "Regulating Intermediaries," 3. Extensive work on jurisdictional problems is done by https://www.internetjurisdiction.net.

23 See Rebecca MacKinnon, Elonnai Hickok, Allon Bar, and Hai-in Lim, "Fostering Freedom Online: The Role of Internet Intermediaries," UNESCO Series on Internet Freedom (Paris: United Nations Educational, Scientific and Cultural Organization, 2014), 34; although there are strict requirements under the German constitution regarding any limit on freedom of speech enshrined in Article 5(1) Basic Law, there is more scope for speech regulation than in the United States.

24 Schulz, "Regulating Intermediaries," 2.

25 See Google Spain SL v. AEPD, C-131/12, Court of Justice of the European Union (CJEU, 2014); the CJEU also limited the deletion obligation to the results in European search queries, leaving open the possibility of still finding the results when using, for example, Google.

26 Schulz, "Regulating Intermediaries," 2*ff.*

27 See Nathaniel Persily, *The Internet's Challenge to Democracy: Framing the Problem and Assessing Reforms*, Kofi Annan Foundation, 2019.

28 See Ireton and Posetti, *Journalism.*

29 Yochai Benkler, Robert Faris, and Hal Roberts, *Network Propaganda: Manipulation, Disinformation, and Radicalization in American Politics* (New York: Oxford University Press, 2018).

30 See Ireton and Posetti, *Journalism.*

31 See Julie Posetti, "News Industry Transformation: Digital Technology, Social Platforms and the Spread of Misinformation and Disinformation," in Ireton and Posetti, *Journalism*, 57*ff.*

32 See Ireton and Posetti, *Journalism*, 10*ff.*, 18; see generally Jill Lepore, "Does Journalism Have a Future?," *New Yorker*, January 28, 2019, https://www.newyorker.com/magazine/2019/01/28/does-journalism-have-a-future.

33 See Schulz, "Regulating Intermediaries," 13; see also Alexandra Kuczerawy, "Private Enforcement of Public Policies: Freedom of Expression in the Era of Online Gatekeeping" (PhD thesis, KU Leuven, 2018), 39*ff.*

34 See Schulz, "Regulating Intermediaries," 13.

35 Sascha Hölig and Uwe Hasebrink, "Reuters Institute Digital News Report 2019: Results for Germany" (working paper no. 47, Leibniz Institute for Media Research, Hans Bredow Institute, June 2019), 7, https://www.leibniz-hbi.de/uploads/media/default/cms/media/os943xm_AP47_RDNR19_Deutschland.pdf.

36 See Hölig and Hasebrink, "Reuters Institute," 7.

37 On the genesis of of the Communications Decency Act, Section 230, see Kate Klonick, "The New Governors: The People, Rules, and Processes Governing Online Speech," *Harvard Law Review* 131, no. 6 (April 2018): 1605*ff.*

38 See Jeff Kosseff, *The Twenty-Six Words That Created the Internet* (Ithaca, NY: Cornell University Press, 2019).

39 European Parliament, Council of the European Union, "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 ['Directive on electronic commerce']," https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX%3A32000L0031.

40 See Giancarlo F. Frosio, "From Horizontal to Vertical: An Intermediary Liability Earthquake in Europe," *Journal of Intellectual Property Law & Practice* 12, no. 7 (July 2017): 565–75.

41 See, for instance, the European Commission, "Commission Recommendation of 1.3.2018 on measures to effectively tackle illegal content online (C(2018) 1177 final)," March 1, 2018, https://ec.europa.eu/digital-single-market/en/news/commission-recommendation-measures-effectively-tackle-illegal-content-online.

42 See Balkin, "Free Speech in the Algorithmic Society," 1182.

43 Balkin, "Free Speech Is a Triangle," 2019.

44 See Kuczerawy, "Private Enforcement," 41*ff.*; Seth Kreimer, "Censorship by Proxy: The First Amendment, Internet Intermediaries, and the Problem of the Weakest Link," *University of Pennsylvania Law Review* 155, no. 1 (November 2006): 11–101.

45 In a survey conducted in February 2018, 85 percent of respondents described misinformation as a problem in their countries; see European Commission, "Final Results of the Eurobarometer on fake news and online disinformation," March 12, 2018, https://ec.europa.eu/digital-single-market/en/news/final-results-eurobarometer-fake-news-and-online-disinformation.

46 European Commission, Code of Practice on Disinformation, September 26, 2018, https://ec.europa.eu/digital-single-market/en/news/code-practice-disinformation; see also Olga Robinson, Alistair Coleman, and Shayan Sardarizadeh, "A Report of Anti-Disinformation Initiatives," Oxford Internet Institute, August 2019, 3*ff.*, https://comprop.oii.ox.ac.uk/wp-content/uploads/sites/93/2019/08/A-Report-of-Anti-Disinformation-Initiatives.

47 European Commission, "First results of the EU Code of Practice against disinformation," January 29, 2019, https://ec.europa.eu/digital-single-market/en/news/first-results-eu-code-practice-against-disinformation.

48 European Commission, "Code of Practice against disinformation: Commission calls on signatories to intensify their efforts," January 29, 2019, https://ec.europa.eu/commission/news/code-practice-against-disinformation-2019-jan-29_en.

49 European Commission, "Code of Practice against disinformation: Commission recognizes platforms' efforts ahead of the European elections," May 17, 2019, https://europa.eu/rapid/press-release_STATEMENT-19-2570_en.htm.

50 See Annina Claesson, "Coming Together to Fight Fake News: Lessons from the European Approach to Disinformation," *New Perspectives on Foreign Policy* 17 (Spring 2019): 15.

51 See, e.g., Laura Blanco, "Tackling Disinformation: Proposed EU Code of Practice Should Not Lead to Politically Biased Censorship," *Center of Democracy and Technology* (blog), August 21, 2018, https://cdt.org/blog/tackling-disinformation-proposed-eu-code-of-practice-should-not-lead-to-politically-biased-censorship.

52 See Schulz, "Regulating Intermediaries," 4.

53 German Ministry of Justice, "Löschung von strafbaren Hasskommentaren durch soziale Netzwerke weiterhin nicht ausreichend" (press statement, March 14, 2017), https://www.bmjv.de/SharedDocs/Pressemitteilungen/DE/2017/03142017_Monitoring_SozialeNetzwerke.html.

54 See Schulz, "Regulating Intermediaries," 4.

55 *Bundesgesetzblatt Jahrgang Teil I* 16 (September 7, 2017): 3352.

56 That might be a newly created entity or an established one. The concept of an independent self-regulatory body was inserted into the bill during the legislative process, after experts proposed the successful model of "regulated self-regulation," i.e., coregulation in German minor protection law as an alternative approach. This model allows providers to form a self-regulatory body, which takes over a monitoring and sanctioning function. The public regulatory authority's role is to supervise the scheme's

implementation. In the context of the NetzDG, however, the lawmakers only included the notion of an independent self-regulatory body in the bill but did not build any further on the lessons learned from minor protection law. As of September 2019, a system of coregulation has not yet been implemented by the industry. On the concept, see Schulz, "Regulating Intermediaries," 5.

57 See Schulz, "Regulating Intermediaries," 5.

58 See Schulz, "Regulating Intermediaries," 5.

59 See Article 19, "Germany: The Act to Improve Enforcement of the Law in Social Networks" (legal analysis, August 2017), 19.

60 See Gerald Spindler, "Internet Intermediary Liability Reloaded: The New German Act on Responsibility of Social Networks and Its (In-) Compatibility with European Law," *Journal of Intellectual Property, Information Technology and E-Commerce Law* 8, no. 2 (2017): 166–79.

61 See Thomas Wischmeyer, "'What Is Illegal Offline Is Also Illegal Online'—The German Network Enforcement Act 2017," in *Fundamental Rights Protection Online: The Future Regulation of Internet Intermediaries*, ed. Bilyana Petkova and Tuomas Ojanen (Cheltenham, UK: Edward Elgar, 2019), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3256498.

62 See European Parliament, Council of the European Union, "Directive 2000/31/EC," Article 5.

63 See Schulz, "Regulating Intermediaries," 7.

64 Some small platforms, such as Gab, even make a point about not policing content at all.

65 See Schulz, "Regulating Intermediaries," 10.

66 See Holznagel, "Legal Review of the Draft Law," 23*ff.*

67 Google, "Entfernungen von Inhalten nach dem Netzwerkdurchsetzungsgesetz" (NetzDG transparency report), https://transparencyreport.google.com/netzdg/youtube?hl=de.

68 Twitter, "Netzwerkdurchsetzungsbericht Januar – Juni 2018" (NetzDG transparency report), https://cdn.cms-twdigitalassets.com/content/dam/transparency-twitter/data/download-netzdg-report/netzdg-jan-jun-2018.pdf.

69 Google, "Entfernungen"

70 Facebook, NetzDG Transparency Report, July 2018, https://fbnewsroomus.files.wordpress.com/2018/07/facebook_netzdg_july_2018_english-1.pdf.

71 Facebook, NetzDG Transparency Report, January 2019, https://fbnewsroomus.files.wordpress.com/2019/01/facebook_netzdg_january_2019_english71.pdf.

72 See Amélie Pia Heldt, "Reading between the Lines and the Numbers: An Analysis of the First NetzDG Reports," *Internet Policy Review*, 8, no. 2 (June 12, 2019): 8*ff.*, https://doi.org/10.14763/2019.2.1398.

73 In response, the Federal Ministry of Justice fined Facebook 2 million euros in July 2019, arguing that the company underreported complaints (after an appeal by Facebook, a final decision is still pending); see Thomas Escritt, "Germany Fines Facebook for Under-Reporting Complaints," Reuters, July 2, 2019, https://www.reuters.com/article/us-facebook-germany-fine/germany-fines-facebook-for-under-reporting-complaints-idUSKCN1TX1IC?il=0.

74 See Kirsten Gollatz, Martin Riedl, and Jens Pohlmann, "Removals of online hate speech in numbers," *Digital Society* (blog), August 9, 2018, https://www.hiig.de/en/removals-of-online-hate-speech-numbers.

75 See Benkler, Faris, and Roberts, *Network Propaganda*, 384*ff.*; see also Daphne Keller, "Empirical Evidence of 'Over-Removal' by Internet Companies under Intermediary Liability Laws," *Center for Internet*

*and Society, Stanford Law School* (blog), October 12, 2015, http://cyberlaw.stanford.edu/blog/2015/10/empirical-evidence-over-removal-internet-companies-under-intermediary-liability-laws.

76 Timothy Garton Ash, Robert Gorwa, and Danaë Metaxa, "GLASNOST! Nine Ways Facebook Can Make Itself a Better Forum for Free Speech and Democracy," Reuters Institute for the Study of Study of Journalism, University of Oxford, January 17, 2019, https://reutersinstitute.politics.ox.ac.uk/our-research/glasnost-nine-ways-facebook-can-make-itself-better-forum-free-speech-and-democracy.

77 Even if researchers secure access to the data, there would of course still be conflicting legal views among researchers about the propriety of content-moderation decisions.

78 See Oreste Pollicino, "Fundamental Rights as Bycatch—Russia's Anti–Fake News Legislation," *Verfassungsblog* (blog), March 28, 2019, https://verfassungsblog.de/fundamental-rights-as-bycatch-russias-anti-fake-news-legislation.

79 Article 10(1) of the European Convention on Human Rights states: "Everyone has the right to freedom of expression. This right shall include freedom to hold opinions and to receive and impart information and ideas without interference by public authority and regardless of frontiers. This Article shall not prevent States from requiring the licensing of broadcasting, television or cinema enterprises."

80 In addition, freedom of expression is protected under Article 11 of the EU Charter of Fundamental Rights in the context of EU law; on the national level, state constitutions also include provisions on freedom of expression.

81 See Brittan Heller and Joris van Hoboken, "Freedom of Expression: A Comparative Summary of United States and European Law," Transatlantic High Level Working Group on Content Moderation Online and Freedom of Expression, May 3, 2019, 10*ff.*, https://www.ivir.nl/publicaties/download/TWG_Freedom_of_Expression.pdf.

82 See Pieter-Jan Ombelet, "The Chilling Effects of Content Policing by Social Media," *Centre for IT & IP Law, KU Leuven* (blog), July 5, 2016, https://www.law.kuleuven.be/citip/blog/the-chilling-effects-of-content-policing-by-social-media.

83 Whereas private platforms may also negatively impact users' freedom of expression, this does not qualify as an interference under ECHR Article 10; in this case, however, positive obligations of states may be triggered.

84 Schulz, "Regulating Intermediaries," 8.

85 See Schulz, "Regulating Intermediaries," 8*ff.*

86 Schulz, "Regulating Intermediaries," 8.

87 See Schulz, "Regulating Intermediaries," 10.

88 See Schulz, "Regulating Intermediaries," 10.

89 Transparency notice: The author has been the chairman of the committee (MSI-NET) that drafted the recommendation.

90 UN Office of the High Commissioner for Human Rights, "Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression" (n.d., accessed November 4, 2019), https://www.ohchr.org/en/issues/freedomopinion/pages/opinionindex.aspx.

91 Council of Europe, "Recommendation CM/Rec(2018)2 of the Committee of Ministers to member States on the roles and responsibilities of internet intermediaries," March 7, 2018, https://rm.coe.int/1680790e14.

92 UN Office of the High Commissioner for Human Rights, "Guiding Principles on Business and Human Rights: Implementing the United Nations 'Protect, Respect and Remedy' Framework," United Nations, 2011, https://www.ohchr.org/documents/publications/GuidingprinciplesBusinesshr_eN.pdf.

93 See Amélie Pia Heldt, "Upload-Filters: Bypassing Classical Concepts of Censorship?," *Journal of Intellectual Property, Information Technology and Electronic Commerce Law* 10, no. 1 (2019): 62, par. 21.

94 See Schulz, "Regulating Intermediaries," 12.

95 It is therefore not surprising that court cases involving alleged speech-related crimes are often overturned on appeal, given that there are seldom definitive answers; see Schulz, "Regulating Intermediaries," 9.

96 For a detailed overview of current research gaps, see Joshua A. Tucker, Andrew Guess, Pablo Barberá, Cristian Vaccari, Alexandra Siegel, Sergey Sanovich, Denis Stukal, and Brendan Nyhan, "Social Media, Political Polarization, and Political Disinformation: A Review of the Scientific Literature," Hewlett Foundation, March 2018, https://hewlett.org/wp-content/uploads/2018/03/Social-Media-Political-Polarization-and-Political-Disinformation-Literature-Review.pdf.

97 See, for example, the work of Susan Benesch, especially the Dangerous Speech Project, https://dangerousspeech.org.



25 24 23 22 21 20 19 7 6 5 4 3 2 1

The preferred citation for this publication is Wolfgang Schulz, *Roles and Responsibilities of Information Intermediaries: Fighting Misinformation as a Test Case for a Human Rights–Respecting Governance of Social Media Platforms,* Hoover Working Group on National Security, Technology, and Law, Aegis Series Paper No. 1904 (November 13, 2019), available at https://www.lawfareblog.com/roles-and-responsibilities-information-intermediaries.

## About the Author

### WOLFGANG SCHULZ

Wolfgang Schulz is director of the Leibniz Institute for Media Research, Hans Bredow Institute, and professor of media law, public law, and legal theory at the University of Hamburg. He is also director of the Alexander von Humboldt Institute for Internet and Society in Berlin and is the chair for Freedom of Communication and Information at the German UNESCO Commission.

## *Working Group on National Security, Technology, and Law*

The Working Group on National Security, Technology, and Law brings together national and international specialists with broad interdisciplinary expertise to analyze how technology affects national security and national security law and how governments can use that technology to defend themselves, consistent with constitutional values and the rule of law.

The group focuses on a broad range of interests, from surveillance to counterterrorism to the dramatic impact that rapid technological change—digitalization, computerization, miniaturization, and automaticity—are having on national security and national security law. Topics include cybersecurity, the rise of drones and autonomous weapons systems, and the need for—and dangers of—state surveillance. The group's output will also be published on the Lawfare blog, which covers the merits of the underlying legal and policy debates of actions taken or contemplated to protect the nation and the nation's laws and legal institutions.

Jack Goldsmith is the chair of the National Security, Technology, and Law Working Group.

*For more information about this Hoover Institution Working Group, visit us online at http://www.hoover.org/research-teams /national-security-technology-law-working-group.*

**Hoover Institution, Stanford University**
434 Galvez Mall
Stanford, CA 94305-6003
650-723-1754

**Hoover Institution in Washington**
The Johnson Center
1399 New York Avenue NW, Suite 500
Washington, DC 20005
202-760-3200

HOOVER INSTITUTION