: Global Governance Spotlight

2 | 2019

sef:

International Rules for Social Media

Safeguarding human rights, combating disinformation

Matthias C. Kettemann

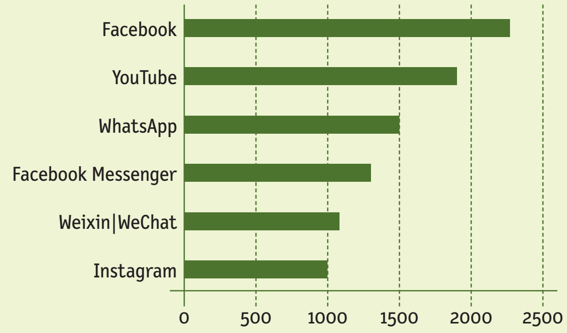
Information on the Internet is often "manipulative, full of half-truths or even used in a targeted manner as state propaganda", as German Chancellor Angela Merkel pointed out at the opening of the new Federal Intelligence Service headquarters in Berlin in February 2019. For that reason, she said, "We must learn to deal with fake news as an element of hybrid warfare." (Federal Chancellery 2019)

Facebook, Wikipedia, YouTube and Twitter are increasingly becoming vehicles for strategic content deployment by States in (dis)information operations. These platforms – collectively termed "social media" – offer their users a multitude of opportunities for information-gathering, networking, opinion-forming and communication. These processes need rules and governance – that much is obvious. What is less clear is why existing regulatory mechanisms have failed, thus far, to impose effective curbs on hate speech and disinformation.

This issue of Spotlight looks at the successes and shortcomings of current norm-building processes and efforts to regulate social media, particularly at regional and global level. On this basis, five guidelines which may be useful in overcoming the failings of previous regulatory efforts are then identified.

Communication billionaires:
Most popular social networks worldwide

Number of active users (in millions), as of January 2019



Source: https://de.statista.com/statistik/daten/studie/181086/umfrage/die-welt-weit-groessten-social-networks-nach-anzahl-der-user/, 19.02.2019

The Internet: a new theatre in the information war

The social networks' user statistics speak for themselves: real-time communication in these new online spaces worldwide is non-stop. In theory, this form of communication falls within the scope of universal basic and human rights protection and the principle of non-interference in States' domestic affairs.

In other words, there is no lack of applicable rules – from international law, regional integration law and national law to community standards and providers' general terms of service. And yet many users and some States are flouting these norms. Alongside large-scale information operations using fake news and fake accounts (social bots), hate speech – from discrimination to Holocaust denial – encourages a crude and abusive style of online communication that is harmful to the community at large.

When it comes to mastery of social media, few countries come close to rivalling Russia. Russian trolls, hackers and social bots continuously intervene in other societies' online communication. In their book LikeWar (2018, p. 204), Peter W. Singer and Emerson T. Brooking quote Vladimir Putin's former press secretary, Dmitry Peskov: "The new reality [of social media] creates a perfect opportunity for mass disturbances or for initiating mass support or mass disapproval." And as part of this reality, a febrile war of likes and dislikes is raging – and is increasingly manipulated for political purposes. This impacts on Germany too: a memorable example is the 2016 campaign that spread fake news about alleged violent crimes by refugees and even involved some of Russia's leading politicians.

Existing norms are clearly not having the desired effect. New normative processes and enforcement mechanisms seem to be required – and are to some extent already emerging – to ensure that the norms which exist in theory are applied to social media and these media are regulated in a manner consistent with human rights.

New norms for social networks: the UN and Council of Europe processes

At its latest session in December 2018, the United Nations General Assembly agreed on two parallel processes for developing rules with social media relevance. A Russian resolution (A/RES/73/27) proposed the convening of an open-ended working group (OEWG) tasked with further developing the rules, norms and principles of responsible behaviour of States in cyberspace. The resolution is infused with a state-centred conception of cybersecurity based on the protection of national interests. The resolution's recitals thus affirm "the right and duty of States" to combat "the dissemination of false or distorted news, which can be interpreted as interference in the internal affairs of other States or as being harmful to the promotion of peace, cooperation and friendly relations among States and nations".

The second of these regulatory processes centres on the responsibility of the Group of Governmental Experts (GGE), whose reconvening until 2021 was secured by a General Assembly resolution tabled by the US (A/RES/73/266). This is significant since the existing GGE process has already been successful in establishing a number of key normative waymarkers. For example, in the 2015 Report of the Group of Governmental Experts (A/70/174), States reaffirmed that international law, and in particular the United Nations Charter and the Universal Declaration of Human Rights, is applicable in its entirety to the Internet. The GGE has also developed norms, rules and principles for the responsible behaviour of States and identified various confidence-building measures.

Both processes are noteworthy, but given that the first (OEWG) places the emphasis on sovereignty while the second is essentially community-oriented, differing nuances can be expected in the emerging norms. On a positive note, even the OEWG resolution makes reference to the main outcomes achieved by previous expert groups (specifically, the duty of a State to abstain from any action that violates international law) and calls for input from all stakeholders, including civil society. As the large majority of UN member States voted for both processes, it is still unclear which "bloc" will ultimately be more successful as a norm entrepreneur: the States that favour sovereignty-oriented Internet governance or those which advocate for a community-based approach.

Let's turn to Europe. Here, it was the Council of Europe which established the key guidelines for the future regulation of social networks in its Recommendation CM/Rec(2018)2 of the Committee of Ministers on the roles and responsibilities of Internet intermediaries. The Recommendation starts by underlining the duty of member States to refrain from violating the right to freedom of expression and other human rights in the digital sphere when regulating social networks. Any action by public authorities addressed to Internet intermediaries that interferes with human rights and fundamental freedoms must be prescribed by law. For their part, Internet intermediaries, when developing internal rules, must act in accordance with their responsibilities to respect human rights, in particular the United Nations Guiding Principles on Business and Human Rights (Ruggie Principles). This responsibility exists independently of the States' ability or willingness to fulfil their own human rights obligations.

New approaches to norm enforcement: EU and German legislation

At European Union level and within the national framework, the focus is less on the setting of norms and more on mechanisms for their enforcement, including the assignment of responsibilities to major commercial Internet platforms.

The European Commission has long favoured a cooperative approach to the regulation of Internet intermediaries and social media service providers. In its current efforts to prevent the dissemination of terrorist content, however, it is turning its back on this approach. This is particularly evident in the draft Regulation on preventing the dissemination of terrorist content online (COM(2018) 640 final, 12.9.2018), which has come in for criticism from business and the public alike. The proposed provisions focus on hosting service providers within the EU and require them, under the threat of penalties, to remove terrorist content or disable access to it within one hour of a removal order being issued by a competent authority. Hosting service providers must also take pro-active measures to prevent re-upload of already removed content.

The problem here is the very broad definition of "terrorist" content. Whereas Directive (EU) 2017/541 on combating terrorism consistently defines terrorism as acts committed with a specific terrorist aim, the draft Regulation may also apply to communications in which "terrorist" content is reproduced for journalistic or academic purposes or in the context of historical research. What's more, the means of legal redress are not spelled out in sufficient detail in the draft. In particular, the short one-hour timeframe scarcely allows for a thorough review of human rights implications. The proposed pro-active measures would compel service providers to deploy algorithmic filtering ex ante. In practice, the new rules would create general surveillance obligations, which are prohibited under the e-Commerce Directive (2000/31/EC).

Besides combating terrorist propaganda online, the European Commission's regulatory endeavours are aimed at safeguarding the integrity of the upcoming European elections in May 2019. Following on from the report by its High Level Expert Group on Fake News and Online Disinformation, the Commission published a Communication (Tackling online disinformation: a European Approach, COM(2018) 236 final, 26.4.2018) in April 2018. The various measures that it calls for include a Code of Practice for self-regulation by major service providers such as Facebook and Google (EU Code of Practice on Disinformation). Signatories are required to make efforts to ensure that they do not accept remuneration from accounts and websites which consistently misrepresent information; to ensure transparency about political and issue-based advertising; to take pro-active measures against fake accounts and social bots; to empower users to report disinformation; and to facilitate discovery and access to reliable content.

With its lack of clear fulfilment criteria and self-regulation mechanisms, the Code of Practice undoubtedly still has shortcomings. Nevertheless, its general approach – focusing on cooperative regulatory solutions for normative fields where binding law alone cannot achieve the desired objectives – is promising in principle. The aim of empowering users to broaden their media consumption as a means of countering disinformation can only be achieved through cooperation with platform operators, who must modify their algorithms to include more content diversity in their suggestions for users.

One criticism often heard in the German debate is that cooperative agreements do not go far enough. Accordingly, the focus has recently shifted to legislative measures. Germany's Act to Improve Enforcement of the Law in Social Networks (Network Enforcement Act – NetzDG), which came into effect on 1 January 2018, applies to providers of social networks with more than two million registered users in the Germany. Among other things, providers of social networks are required to remove or block access to content that is "manifestly unlawful" within 24 hours of receiving a complaint. Companies that fail to comply or do not set up an effective complaints mechanism face fines of up to €50 million. A review of the Act is planned for next year and will look at criticisms and lessons learned during its initial period in force.

Rules for an Internet that complies with international law and human rights

In view of social media reach and the changing communication practices of today's users, there is no circumventing the need to develop and enforce new rules for the governance of these media. The political sphere as we know it – the articulation and aggregation of opinions – is increasingly shifting to online platforms. The following five guidelines have the potential to provide direction for efforts aimed at addressing the shortcomings of current regulatory approaches while avoiding ethical problems:

(1) New rules for social media must be based on a commitment by all stakeholders – particularly States and businesses – to take on shared, role-specific responsibility to protect the rule of law and human rights. Simply shifting responsibility to users is not an adequate response.

(2) States should not use new processes for social media regulation as a means of asserting particular interests that conflict with the international community's goal of making the Internet human rights- and development-compliant. Above all, States should not attempt to duplicate existing processes when they start to lose control of the agenda.

(3) States must be reminded of their self-imposed duty – set out, inter alia, in the 2015 Report of the Group of Governmental Experts – to apply international law in its entirety to the Internet. Specifically, States have a duty to abstain from interventions (via and on social media) that violate international law.

(4) Businesses must participate in norm-setting processes in good faith and should not attempt to subvert these processes with references to self-regulation. While the potential benefits of self-regulation should be recognised in international negotiations, it must support and enhance States' existing rights protection infrastructure, not weaken or replace it.

(5) The regulation of algorithms governing suggestion selection and deletion practices on social media is still inadequate. Automated decision-making systems must be designed with sensitivity

to human rights in mind. Even Mark Zuckerberg acknowledged in an open letter in November 2018 that people are engaging disproportionately with "more sensationalist and provocative content". Zuckerberg plans to use artificial intelligence to reduce the dissemination of such content. Although this is, broadly speaking, the right approach, a more fundamental debate about algorithm regulation as a mechanism for governance of expression of opinion on social media is still awaited.

Using the law against the trolls

In her speech, quoted at the start of this paper, Chancellor Merkel asked: "Which information is correct? What has been manipulated? Where might propaganda by a state agency be hovering in the background?" Citizens who are social media users must be empowered to answer these questions – and democracies under the rule of law must assist them. The Internet's own dynamics, the untrustworthiness of the information warriors and the skewed logic of troll communication do not make this an easy task. As Peter W. Singer and Emerson T. Brooking rightly observe in LikeWar (2018, p. 211), Western democracies find themselves "at a distinct disadvantage": "Shaped by the Enlightenment, they seek to be logical and consistent. Built upon notions of transparency, they seek to be accountable and responsible. These are the qualities that made them so successful ... Unfortunately, they are not the values of a good troll ..."

What's the solution? It is certainly not to dispense with normative rules for online communication altogether. On the contrary, faced with the challenges of hate speech, trolls, social bots and (dis)information operations, we need political processes that produce legitimate norms. Only fair national, regional and international negotiating processes that are as inclusive as possible and sensitive to the rule of law can guarantee the emergence of new rules for social media that restore them to their rightful role as a force for social emancipation instead of being a threat to democracy.

Author

Dr Matthias C. Kettemann | Head of the Research Programme Regulatory Structures and the Emergence of Rules in Online Spaces at the Leibniz Institute for Media Research | Hans-Bredow-Institut. He is currently completing his postdoctoral thesis in Internet law, international law and legal theory at Goethe University Frankfurt's Faculty of Law. He is also a lecturer at the University of Graz, Austria.

References

Federal Chancellery 2019: Speech by Chancellor Angela Merkel at the opening of the new Federal Intelligence Service headquarters, Berlin, 8.2.2019 (in German).

European Commission 2018: Action Plan against Disinformation, JOIN(2018) 36 final, 5.12.2018.

Council of Europe 2018: Recommendation CM/Rec(2018)2 of the Committee of Ministers to member States on the roles and responsibilities of internet intermediaries, 7.3.2018.

Peter W. Singer/Emerson T. Brooking 2018: LikeWar: The Weaponization of Social Media, Boston.

Imprint

The Development and Peace Foundation (sef:) was founded in 1986 on the initiative of Willy Brandt. As a cross-party and non-profit-making organisation, the sef: provides an international high-level forum for shared thinking on urgent peace and development issues.

Global Governance Spotlight is a policy-oriented series whose purpose is to critique international negotiation processes from a global governance perspective.

Published by
Development and Peace Foundation (sef:)/
Stiftung Entwicklung und Frieden (sef:)
Dechenstr. 2 : 53115 Bonn : Germany
Phone +49 (0)228 959 25-0 : Fax -99
sef@sef-bonn.org : @sefbonn
www.sef-bonn.org

Editor
Dr Michèle Roth

Translation
Hillary Crowe

Design Basic Concept
Pitch Black Graphic Design
Berlin/Rotterdam

Layout
Gerhard Süß-Jung

Contents do not necessarily reflect the views of the publisher.

ISSN 2566-624X