A NEW MODEL FOR MEDIA REGULATION

Online platforms are placing great pressure on safeguards to democracy, and legal remedies are on the stocks. As **KRISZTINA ROZGONYI** discusses, there is a pressing need for a new generation of media regulators to implement rules and build trust

Censorship or accountable conduct? Privatisation of law enforcement or efficient co-regulation of new media? We are witnessing differing but interrelated discussions on fundamental proposals in the European Union (EU) touching on these issues. The future regulation of online media content concerning hate speech and protection of minors, viral spreading of fake news on social media and the fight against copyright infringement on video-sharing platforms has led to heated debates among stakeholders on regulatory propositions by the European Commission. There are essential though competing concepts on the responsibility of online media platforms for illegal, harmful or misleading content hosted by them and on the role of the judiciary and of state authorities within newly emerging regulatory value chains.

In this article I give an overview of the key notions and offer critical questions for debate on how to represent the public interest in regulating online media platforms and on the role of national regulators within this framework.

RISE OF REGULATORY PRESSURE ON ONLINE PLATFORMS

The year 2018 will see an end to the regulatory impunity of internet intermediaries such as search engines (Google) and video-sharing platforms (YouTube), as well as social media networks (Facebook, Twitter) – at least in Europe. The relatively safe harbour they have enjoyed since 2010 under the liability exemptions of the EU E-Commerce Directive for third-party content, generally freeing them from the obligation to monitor such content, will come to an end soon. The regulatory proposals published by the Commission on amending the rules on audiovisual media services (draft Audiovisual Media Services Directive, AVMSD) and on reforming copyright protection and enforcement (draft Directive on Copyright in the Digital Single Market, DSMD) have been debated since 2016 and new legislation is expected to be adopted this year. Furthermore, new propositions on the policy agenda for tackling fake news and online disinformation are just emerging and suggesting new angles to the debate.[1]

The first signal to change the current regulatory regime was sent in 2015 by the European Court of Human Rights, which held platform providers liable for certain hosting activities of user-generated content.[2] Key political and social events with global and local significance – the US presidential campaign in 2016, the French elections in 2017, the migration crisis – have put pressure on online platform providers on at least two major issues: hate speech online and the spread of false information (fake news).

> **Platform providers have two major issues: hate speech and the spread of false information.**

Politicians in Europe are eager to see rapid results to stop these phenomena but have triggered hasty and misguided legislative and policy interventions. At the member state level, Germany acted first in 2017,[3] adopting a new law on immediate (24 hour) removal and blocking of "clearly violating" content and a more measured 7 days for taking down "violating content" incentivised through severe administrative penalties of up to €5 million. Next, France proposed legislation to put a halt to the spread of fake news on social media during elections.[4]

In parallel has come the reform of copyright rules for the digital age. The Commission put forward at the end of 2016 a reform package consisting of a regulation and the draft directive on copyright with the aim of "ensuring that consumers and creators can make the most of the digital world". The draft legislation has critical proposals on changing the conditions for video-sharing platforms and potentially copyright-infringing user uploads.

Though the regulatory targets of the draft and of already adopted legislation vary from tackling

dangerous expressions of hatred to pirating the latest music, the underlying concepts show similarities: assuming greater responsibility for illegal and harmful user-generated content and to make intermediaries act expeditiously and effectively in "getting rid" of such content. Moreover, the Commission contextualised the interpretation of duties of online platforms, recommending proactive detection, identification and removal of presumed illegal content online.[5] The approach is to promote co-regulative and self-regulative solutions with the aim of levelling the "playing field for comparable digital services" while expecting the "responsible behaviour of online platforms to protect core values".

HATE SPEECH AND ONLINE AUDIOVISUAL MEDIA CONTENT

The moral, philosophical and legal debates and considerations on what actually constitutes hate speech and whether it should be subject to laws criminalising such expressions (whereby most European countries restrict hate speech but the US does not) are beyond the scope of this article. Within the EU context, the Framework Decision on Combating Racism and Xenophobia defines hate speech as "public incitement to violence or hatred directed against a group of persons or a member of such a group defined on the basis of race, colour, descent, religion or belief, or national or ethnic origin" and at the same time criminalises such speech and acts taking place both online and offline.

The AVMSD is the EU's central and strongest legal instrument for regulating traditional broadcasting (linear AV media services) and online media content (non-linear AV media services). Since its adoption in 2010 it has set clear standards against hate speech applicable to linear and non-linear media services. All providers have to ensure that AV services do not contain any incitement to hatred based on race, sex, religion or nationality, and it is applicable also to channels using an uplink in an EU country and to satellite broadcasts. It has led to bans of television channels endorsing violence as a means of propaganda during recent conflicts.[6]

The draft AVMSD includes video-sharing platforms (VSPs) and social media services that provide a significant amount of audiovisual content as subjects in those rules (social media platforms were not subject to the original proposals, but the European Council included them in May 2017).[7] The reasoning is the increased exposure of young audiences to harmful content and hate speech on VSPs as changing patterns of online consumption become the norm. According to the new rules, VSPs will have a duty to take appropriate action when users flag up any content inciting violence or hatred. Measures would include:

- Defining and applying in the terms and conditions of the VSPs the concepts of incitement to violence or hatred
- Establishing and mechanisms for users to report to a VSP provider the content infringing those rules
- Establishing and operating systems allowing users of VSPs to rate content accordingly
- Establishing and operating systems through which providers of VSPs explain to users what effect has been given to the reporting of infringing content – through co-regulation.

The draft rules envision a new regulatory value chain whereby the interest of the public (the value) – protection against incitement to hatred and violence – is to be represented by private parties and assessed by "designated public authorities" (national regulators) in a rather vaguely described manner. Although the proposal acknowledges that the majority of the content stored on VSPs is not under the editorial responsibility of providers, they are considered to be in the position to "determine the organisation of the content, namely programmes or user-generated videos, including by automatic means or algorithms". In other words, the controlling option of VSPs is merely to legitimise a new level of responsibility related to the organisation of content to protect all citizens from incitement to violence or hatred. This means that VSP providers will have to bear obligations "in the organisational sphere and do not entail liability for any illegal information stored on the platforms as such".

The proposals came under heavy criticism especially regarding their impact on freedom of expression[8] and on media pluralism.[9] Possible "chilling effects" to user-generated content and the rise of a new generation of private censorship were quoted by many. Under the new regulatory regime, VSP and social media providers will have to maintain a compulsory self-regulatory system to deal with harmful content while corresponding with national laws on illegal content. It is envisioned that this regulatory scheme will subject providers and users to legal uncertainty about the boundaries of harmful and of illegal content, given the various cultural and political conditions of the 28 EU member states, and inevitably lead to over-removal by VSPs to avoid potential liability.

The experience since enactment of the German law[10] that has already introduced similar rules seems to support the fears projected by critics.[11] Allowing private companies to replace the courts in decision-making on inherent and fundamental speech matters has raised concerns on whether due process and ultimately the rule of law can prevail under such regulatory conditions. Furthermore, removal of harmful or illegal content could eliminate the possibility for others to counter the speech, and so could stifle public debate.[12] Also, the reliability of removal results generated by automated filtering tools, and transparency and accountability on removals as a matter of democratic media governance, have seemingly not been given due prominence by policymakers.

PARALLEL PROCEEDINGS AT EU LEVEL

There are other interrelated policy proposals on the spread of misinformation and on copyright protection (see panel, p21) that also need attention.

Fake news – public consultation and high-level expert group. At the end of 2017 the European Commission expressed concerns about the overwhelming flow of misinformation online.[13] Under the current regulatory frameworks, online platforms bear no liability for spreading misinformation unless it includes illegal elements (e.g. hate speech, incitement to terrorism). Moreover, hurdles to identification of sources of information (including illegal) make it difficult to find those liable for that content.

> **Self-regulatory measures in this area have never been tested and remain to be proven.**

The Commission sought answers to finding effective tools to identify reliable and verified information with the help of an EU-wide public consultation[1] and by setting up a high-level expert group. It aimed to explore the perception of fake news by citizens and by other stakeholders while scoping the problem; assess measures already taken by platforms, news media companies and civil society organisations to counter the spread of fake news online; and define possible actions.

At a moment when trust in the internet and social media is generally low and eroding at an accelerating speed – only 34% of EU citizens trust the internet and a mere 20% trust social networks (down from 36% and 21% respectively in 2017)[14] – it is in the interest of the public to address what indeed matters in (re-)building trust. These circumstances have led to calls for policy actions in some countries – notably in France – to propose legislation "to combat fake news" on social media during elections with strict rules on content that deliberately attempts "...to blur lines between truth and lies and undermine people's faith in liberal democracy."[15]

The new French rules will require social media platforms to remove known malicious fake content that disturbs "public peace", and to reveal the sources. Moreover, they also mandate that platforms disclose the identities of advertisers who pay to spread content, as well as the amounts they are paying. It will be the French media regulator that implements the new legislation, with authority to suspend the licence of any media organisation suspected to act under foreign influence and with the aim of distorting democratic elections.

The report delivered by the high-level expert group on fake news and online disinformation[16] takes a multidimensional approach to addressing the interconnected and mutually reinforcing responsibilities of various stakeholders. The recommended actions on transparency, media literacy, user empowerment, and on a sustainable and diverse news media ecosystem, assign tasks to platforms to implement further self-regulatory measures, but it is acknowledged "that the ability of self-regulatory measures in this area has never been tested and the willingness of all parties to adhere to such an approach remains to be proven".

It is arguably in the public interest to involve online platforms in protecting democratic governance and citizens' trust in online media and the internet. But it remains to be answered who articulates that interest and in what way – how to require platforms to account for their actions and to what effect.

The similarities of EU initiatives point to a reengineering of regulatory value chains that shift

THE RELEVANCE OF MODERNISING COPYRIGHT

The EU copyright modernisation package[17] aims to achieve "a fairer marketplace for online content especially for press publications, online platforms and remuneration of authors and performers". The proposed rules address growing concerns as to whether the value generated by online video-sharing platforms is "shared in a fair manner between distributors and the rightsholders". Therefore the reform foresees appropriate and proportionate measures taken by information society providers (online platforms), such as the use of effective content recognition technologies, to counter unfair value transfer. These measures, or filtering systems, are to prevent users uploading potentially infringing content on platforms. It is envisioned that collaboration between information society service providers and rightsholders will validate the assessment of appropriateness of technologies deployed.

Many – including top academics all over the world – expressed reservations on the proposed new rules as being "incompatible with the fundamental rights and freedoms guaranteed under Articles 8 (protection of personal data), 11 (freedom of expression) and 16 (freedom to conduct a business) of the Charter of Fundamental Rights of the EU", placing a "disproportionate burden on platform providers, in particular small and medium-sized operators" and depriving "users of the room for freedom of expression that follows from statutory copyright exceptions".[18]

They suggest harmonised and more detailed rules on notice and takedown procedures and the introduction of a "counter notice" process, while leaving safe harbour immunity of platforms for uncontrolled user-generated content generally intact. The main argument lies with the need for the "maintenance of an equilibrium between all rights and freedoms involved". The critiques seem to be well founded, but it is still to be resolved who, and in which process and under what public scrutiny, can assess whether those procedures indeed preserve such an equilibrium and whether platforms operate in an accountable manner for their users.

the burden of regulating speech online to private entities. The political will of national and EU policymakers to respond fast and effectively to disturbing incidents – such as the flow of hate online, the spread of misinformation or the growing pressure from creative industries for a new "value deal" with platforms – has prompted hasty legislative interventions. However, their impact is far reaching and the consequences, the societal and true political price to be paid, is unforeseeable at this moment. Central to the critiques so far is abandonment of due process and of the rule of law as a consequence of privatisation of censorship[19] and of law enforcement.[8]

However, debates have missed some major components of the proposed new regulatory frameworks with regard to their co-regulative nature: what should the cooperative element actually entail? More precisely: what is the role of the state and of its agencies within the new regulatory value chain? Who defines public interest and how to represent it? Which standards apply and what procedures are to be followed by "designated authorities", namely national regulators, while requiring online intermediaries to account for their actions?

A NEW ERA FOR PUBLIC INTEREST

We are witnessing tipping points in a new era of regulating speech online. Internet giants such as Google, Facebook and others have grown to a critical size and entered a new level of political and economic power and influence globally. A decade ago they were praised as the trustees of freedom of expression – currently they are seen as scapegoats bearing full responsibility for pressing societal and political emergencies. An obvious response is to end their legal immunity and force them into a new regulatory regime. Yet the task of conceptualising, designing and implementing novel regulatory value chains is not straightforward – and the public needs to see their interests being represented by institutions they trust and in a manner that corresponds with realities of 21st century communication.

So the question here is how ready are regulators to engage? What role should they be assigned within emerging new regulatory value chains? Robin Mansell[20] recently articulated the need for dialogue about ways and possibilities for humans to regain their control and authority over digital technologies. It seems inevitable at this point to consider and rethink whether regulators are in a position to exercise control and what prerequisites meet the expectations of citizens, industry and policymakers. In other words: what does it take for regulators to rebuild trust and ensure our freedoms are safeguarded, our rights are enforced, and our interests are advocated?

And this trust is not a given. The OECD has found, in a deteriorating trend, that only 43% of citizens globally trust their government.[21] Moreover, media was the least trusted institution globally for the first time.[22] Against this backdrop, regulators as trustees of public interest will play a precarious role in considering the impact of technological change and keeping pace with rapidity and complexity. Here are suggestions about changes that would improve their relationship with online media platforms.

Redesigned accountability mechanisms. Regulators need to act within the given and proposed legislative boundaries for engagement as national forums of accountability on transparency and appropriateness of removal of content by online intermediaries. They need to set up and enforce accountability procedures for platforms' actions on self-regulating speech and making decisions

about illegal and harmful content. As a minimum, online media should account for, in a transparent manner, on whose request, on what judgement and in what process they are acting – and regulators need to scrutinise those actions. Moreover, regulators should undertake impact assessments of self-regulatory actions and focus on the wider societal and policy implications.

Under redesigned accountability mechanisms, a possible new regulatory value chain should entail, as a minimum, regulatory scrutiny of the appropriateness of measures undertaken by online platforms, for both legal aspects and policy dimensions, as follows.

Legal aspects:

- Regulators need to check under which international standards and which jurisdiction platforms are acting when they delete, or refuse illegal content
- Within the boundaries of national legislation, regulators need to set rules on certain "duties of explanation" for decisions made by automated or artificially intelligent algorithmic systems operated by platforms
- Based on the multinational cooperation of regulators – possibly within established networks such as the European Regulators Group for Audiovisual Media Services (ERGA) or the European Platform of Regulatory Authorities (EPRA) – it seems necessary to develop tools to monitor online platforms, particularly for algorithmic accountability.

The results of such regulatory interventions should be made available in regular monitoring reports, with full transparency and subject to open public debate.

Self-regulatory decisions on speech made by platforms about illegal and harmful content are ongoing. Since 2016, online platform operators are under a self-regulatory regime to enforce the Code of Conduct on Countering Illegal Hate Speech agreed by the European Commission and four major companies (Facebook, Microsoft, Twitter and YouTube). The results of yearly monitoring by the Commission show removal by the platform operators is accelerating to an average removal rate of 70% of the illegal hate speech notified to them (in 2016 it was 28%, in 2017 59%).[23]

While the Commission praised this outcome, it remains unanswered what standards are actually being applied to define hate speech and under which jurisdiction; how and to what extent those removals interfered with the right to freedom of expression; where the boundaries were drawn between illegal and harmful speech; whether removals (or refusals thereof) as self-regulatory decisions on speech matters met the requirements of proportionality; and whether due process was provided to users posting speech later removed.

Most answers to these questions are buried in non-transparent and non-available codebooks used by human moderators and in the algorithmic procedures of online platforms. Revelations on the legality and accuracy of removals are worrying. As an example, the take-down of posts on Holocaust denial, which constitutes illegal hate speech in at least 14 countries but is actually enforced in only four,[24] suggests the need for inquiry on a national level as well. Moreover, there is a high risk of overreaction when removing content because of current incentives (e.g. high fines in Germany), and the potential of automated filtering systems to censor lawful content.

Regulators need to be in a position to effectively monitor the algorithmic accountability of online platforms and take back control largely handed over to software engineers.[25] Moreover, the EU's General Data Protection Regulation (GDPR), which enters into force on 25 May 2018, will not solve problems on transparency of automated or artificially intelligent algorithmic systems and no "right to explanation" will be guaranteed without further legislative steps[26] and regulatory backup. Furthermore, the role of explanation in ensuring accountability needs regular engagement and adaption to the ever-changing technology,[27] and this is where regulators could and should play a crucial role.

> **"Regulators need to monitor algorithmic accountability and take back control from engineers."**

To this end, regulators need to possess assessment tools that simulate and examine algorithmic procedures and provide insights into those systems. Such tools should be developed as a result of wider (at least European level) multinational cooperative efforts.

Policy dimensions:

- Editorial decisions made by online platforms should be monitored by national regulators and juxtaposed with ethical and professional standards
- Regulators should assess the wider societal impact of self-regulatory actions undertaken by platforms within a national context and actively pursue a dialogue with stakeholders in evaluating the results of such an assessment.

Major online platform operators such as Facebook now have news editor roles similar to a traditional media organisation, creating controversies and accusations of bias.[28] Moreover, it seems search engine operators such as Google also seek legal impunity and protection under speech regulations.[29] If that is the case, it means they are subject to different regulatory norms with very different liabilities. They have crossed the line, and national media regulators need to start weighing their operations according to the rules and norms of media under editorial responsibility.

Further, regulators need to gain insights into the national and contextual societal impact of the self-regulatory actions of platforms on speech. This inquiry necessarily has to reflect on historical and political debates on countering hate speech and limitations to freedom of expression with regard to European dimensions. Regulators should closely follow academic research when considering possible regulatory intervention.

National regulators as forums of dispute settlement. The draft AVMSD foresees online platforms as providing complaint and redress mechanisms for the settlement of disputes under self-regulatory measures. National regulators are to:

- Offer possible forums for the settlement of such disputes in agreement with platform operators and representatives of user groups, and alternative settlement procedures (e.g. mediation, conciliation)
- Provide effective oversight of disputes with regular open reporting on evaluation
- Ensure users are aware of their rights in disputes and are encouraged to seek settlement.

Democratic control over online platforms necessitates that users participate in such control, questioning and challenging self-regulatory decisions undertaken by their providers. Regulators are essential in empowering users to enjoy their rights and providing contextualisation and support for their actions. Regular and open debates and dialogue generated by regulators involving representatives of users' groups and other stakeholders could offer meaningful opportunities for such efforts.

A NEW GENERATION OF NATIONAL REGULATORS

The previous era and regime of broadcast regulation has already come to an end. Moreover, the interim period characterised by the extension of previous regulatory categories (e.g. non-linear audiovisual media services) and by balancing acts between "old" and "new" media will probably end soon as well.

It means regulators will have to transform accordingly, in both their organisational and operational setups. On a conceptual level, regulation is to become an inherently interactive exercise with much attention to dynamics between the regulated (platforms) and the beneficiaries (users).[25] It also means that regulators will have to constantly evaluate the impact of such interactions and reflect on their actions. This will need a new generation of regulatory capacity:

- **Skills and competencies:** regulating online platforms needs a different type of competence available within and next to regulators: engineers trained to understand how media distribution networks (e.g. cable, satellite) operate must be complemented by IT professionals who know about algorithms; lawyers skilled in licensing procedures must be supported by data protection specialists; and social scientists knowledgeable about human dimensions of online participation must come on board.
- **Organisational aspects:** monolithic institutional setups should be downsized and reshaped, regulatory monitoring implemented in close cooperation with user groups, and collaborative, interactive, flexible and participatory regulatory procedures need to be followed.
- **Next level of inter- and supranational cooperation:** use of shared pools of information and aligned research methods need to become primary practice; dynamic and multinational research should also engage in routine regulatory exercises.

The next generation of national regulation must be global by design: every regulatory intervention has global impacts and assessment. Regulators must change to meet expectations of their new roles in governing media and communications to build the trust of citizens and to ensure they enjoy their freedoms and rights.

CONCLUSION

The pressure to regulate online platforms is intensifying. The European Commission's latest communication on the fight against illegal content online[30] suggests the next regulatory regime is coming fast. There is not much time to reconsider and agree on a redistribution of roles within the new regulatory value chain and specifically on the representation of public interest. At the same time, established media policy principles such as plurality, diversity and localism should not be abandoned[31] but their effective articulation needs fundamentally new approaches and methods.

There are becoming fewer meaningful opportunities for regaining control over digital technologies, as Mansell comments.[20] As the Council of Europe says, "States have clear responsibilities vis-à-vis internet intermediaries and bear the ultimate obligation to protect human rights and fundamental freedoms in the digital environment."[32] Regulators acting on behalf of citizens enjoy a unique and precious privilege: they are tasked and mandated by the state to exercise such control. It is high time that they find their place for the next era of regulation.

KRISZTINA ROZGONY *is an assistant professor at the Media Governance and Media Industries Research Lab, Department of Communications, University of Vienna. She is also a senior regulatory and legal/policy adviser.*

REFERENCES **1** European Commission (2017). Next steps against fake news: Commission sets up high-level expert group and launches public consultation. **2** Delfi vs Estonia (2015). ECHR 64669/09. **3** Germany: The act to improve enforcement of the law in social networks of 2017. **4** La ministre de la culture précise les contours de la loi contre les 'fake news'. Le Monde, 13 February 2018. **5** European Commission (2018). Commission recommendation on measures to effectively tackle illegal content online. **6** See e.g. European Commission decision on upholding the suspension of retransmission of a Russian language channel RTR Planeta by the Lithuanian regulator. **7** See European Parliamentary Research Service. The Audiovisual Media Services Directive [EU Legislation in Progress]. 28 June 2017. **8** European Digital Rights (2014). Human rights and privatised law enforcement. **9** Bard P, Bayer J (2016). A comparative analysis of media freedom and pluralism in the EU member states. **10** See e.g. Two members of Germany's far-right party investigated by state prosecutor. Guardian, 2 January 2018. **11** Article 19 (2017). Germany: The act to improve enforcement of the law in social networks. Article19.org **12** Center for Democracy & Technology (2017). Tackling 'illegal' content online: The EC continues push for privatised law enforcement. **13** See questions posed by the European Commission in: Public consultation on fake news and online disinformation. (Closed 28 February 2018). **14** European Broadcasting Union (2018). Trust in media 2018. **15** Emmanuel Macron: French president announces 'fake news' law. BBC News, 3 January 2018. **16** High level expert group (2018). Final report of the high level expert group on fake news and online disinformation. **17** European Commission. Modernisation of the EU copyright rules. bit.ly/2cUY9Nf **18** Senftleben M et al. (2017). The recommendation on measures to safeguard fundamental rights and the open internet in the framework of the EU copyright reform. SSRN. bit.ly/2qqS51M **19** Tambini D et al. (2007). The privatisation of censorship? Self-regulation and freedom of expression. In: Codifying Cyberspace: Communications self-regulation in the age of internet convergence. Routledge/UCL Press. **20** Mansell R (2017). Are we losing control? Intermedia, 45 (3): 4–7. **21** OECD (2017). Trust in Government. bit.ly/1Wxr8lE **22** Edelman Trust Barometer 2018. bit.ly/2qlR9Na **23** European Commission (2018). Countering illegal hate speech online – Commission initiative shows continued improvement, further platforms join. Press release. bit.ly/2Du4RoU **24** How Facebook flouts Holocaust denial laws except where it fears being sued. Guardian, 26 May 2017. **25** Helberger N et al. (2015). Regulating the new information intermediaries as gatekeepers of information diversity. Info 17 (6): 50–71. **26** Wachter et al. (2017). Why a right to explanation of automated decision-making does not exist in the General Data Protection Regulation. International Data Privacy Law 7 (2): 76–99. **27** Doshi-Velez F, Kortz M (2017). Accountability of AI under the law: the role of explanation. Berkman Klein Center Working Group on Explanation and the Law. **28** Votavová K, Janda J (2017). Making online platforms responsible for news content. Prague: Kremlin Watch. **29** Google and Facebook don't qualify for first amendment protections. Guardian, 9 March 2018. **30** European Commission (2018). Measures to further improve the effectiveness of the fight against illegal content online. Ares(2018)1183598. **31** Napoli P (2014). Digital intermediaries and the public interest standard in algorithm governance. Media Policy Project Blog, LSE. **32** Council of Ministers (2018). Recommendation CM/Rec(2018)2 of the Committee of Ministers to member states on the roles and responsibilities of internet intermediaries.