Source: © Jim Bourg, Reuters.

Media and Freedom of Expression

On Public Discourse in the Digital Sphere

Supporting Freedom of Expression through a Graduated Approach to Regulating Disinformation

Tobias Schmid

Disinformation – we have all been in contact with it at one time or another, even if we were not aware of it. A forsa survey for Safer Internet Day 2021[1] reveals that 83 per cent of young internet users aged 14 to 24 have encountered fake news on social media. But what do we mean when we talk about disinformation and fake news? How much can be tolerated by a democracy before it is described as unstable? And at what point does regulation become necessary to protect this democracy and its vital process of opinion formation?

Background

These are just some of the questions that arise when considering freedom of expression and disinformation. The issue becomes even more complex when one considers how, in today's digital age, people can disseminate information across borders and share it millions of times – including anonymously.

But first things first: individual freedom of expression has to be at the heart of all considerations about creating and protecting a functioning process of opinion formation. Even before the start of the COVID-19 pandemic and the key role of digital platforms and social media in the run-up to the last US presidential election, there was an ongoing discussion about the various phenomena encompassed by the collective term "disinformation" and potential responses to them. And yet it is precisely in the run-up to elections that democratic societies rely more than ever on functioning, fair, and equal processes of opinion formation.

For many years now, the processes of discussion and argument that elections require have been gradually shifting to the digital sphere – and hence to spaces that are supported by technology. Almost inevitably, this has been accompanied by new methods of communication and by the technical means to disseminate and manipulate information, which in turn leads to the question of responsibilities.

The Role of Platforms

The question of who is responsible for the content of online platforms is not a simple one. First of all, there is the content creator, as the party who is disseminating potentially illegal or false information or using manipulative techniques. However, the platform operators also come into the equation because they provide the infrastructure that gives everyone such a wide audience. They are also often easier to identify and address.

The operators of social media platforms are the main beneficiaries of the way public opinion formation has shifted online. As a result, they have a strong interest in the outcome of the discussions about changing responsibilities. The issue of their responsibility for the opinion formation process is not a new one, and a raft of regulations have been introduced at both EU and national levels. Partly in response to this, many operators have included basic strategies to tackle hate speech and various forms of disinformation in their house rules. However, this has not yet been sufficient to eliminate manipulation because of the lack of enforcement on the part of the platforms and the lack of a basic regulatory structure for such enforcement.

Meeting these new responsibilities requires coordination between the platforms' rules and the statutory regulations. But where do we stand on this?

Disinformation as a European Issue

Like all regulatory discussions relating to occurrences in the digital and hence cross-border space, the fight against disinformation cannot be won solely at the national level. However, we should bear in mind that the EU has only very limited competences in this respect due to the sovereignty of member states over issues relating to culture and media.

> **The platform operators will always protect their business model.**

Back in 2018, the European Commission established a tougher process for tackling the phenomenon of disinformation with its Action Plan against Disinformation.[2] This was in the wake of a commitment by the major platforms to develop a self-regulatory framework for the fight against disinformation – the Code of Practice on Disinformation.[3] Among other things, the Code covers their obligations regarding transparency about political advertising, the deletion of fake accounts, and the demonetisation of those which spread disinformation. The Code was initially signed by Facebook, Google, Twitter, and Mozilla, along with sections of the advertising industry. Microsoft and TikTok followed suit in 2019 and 2020. The self-regulation contained in the Code represents an initial step. It shows that platforms are aware of their changing responsibilities and are willing to accept them to a certain extent. However, this is only a first step towards adequately addressing the problem because voluntary commitments have two major drawbacks. First, like any commercial enterprise, the platform operators will always protect their business model. And the rules tend to be so vaguely formulated (and inevitably drawn up from the company's perspective) that implementation can vary widely from platform to platform. Secondly, the Code does not provide for sanctions. Of course, it is possible to check compliance against their voluntary commitments, but a sound and meaningful assessment of implementation requires a reliable data set. As has already been shown by the assessment of the implementation of the Code conducted by the European Regulators Group for Audiovisual Media Services (ERGA)[4] in 2020, this data is not currently available.[5] The information that is currently provided by the platforms via the Self-Assessment Reports (SAR) has been previously filtered and organised by the platforms, so it is very difficult to make valid statements about the status of the implementation. However, other than verification, it is currently not possible to shift the burden of proof or even to sanction non-compliance.

The European Commission is now taking further action on a number of issues.[6] This includes guidelines[7] on how the platforms can revise and strengthen the Code and a law[8] to improve transparency in sponsored, political advertising. The widely discussed proposal for a Digital Services Act,[9] which essentially deals with enforcing the rules on illegal content on the internet, also enshrines in law certain key elements of the Code relating to the transparency of advertising. In future, platforms will be obliged to carry out risk assessments and to take appropriate countermeasures relating to any systemic risk to freedom of expression that may arise from the operation or use of their services or from the deliberate technical manipulation of their infrastructure. It remains to be seen whether this somewhat piecemeal approach will work – but it is unlikely to be enough.

Disinformation as a National Issue – Information against Disinformation

The first steps in the fight against disinformation have also been taken at national level. Germany's new Interstate Media Treaty (*Medienstaatsvertrag*),[10] which entered into force in November 2020, contains a new supervisory structure regarding compliance with journalistic principles in certain telemedia. In addition to the German Press Council and, in future, possibly other institutions of voluntary self-regulation, the state media authorities are now also

Light into the darkness: To combat disinformation, the European Commission launched an action plan in 2018.
Source: © Johanna Geron, Reuters.

charged with monitoring and enforcing compliance with these principles.

The dispatch of the first letters of advice has already led to tangible results. This less formal approach has increased awareness of diligent journalism, and some telemedia providers who were contacted in this way have already made changes to their offerings. Some of those who have not done so have introduced oversight procedures. These are merely first steps, and a lengthy road lies ahead – persistence and perseverance are required to produce widespread, visible results. But the state media authorities have demonstrated these qualities more than once.

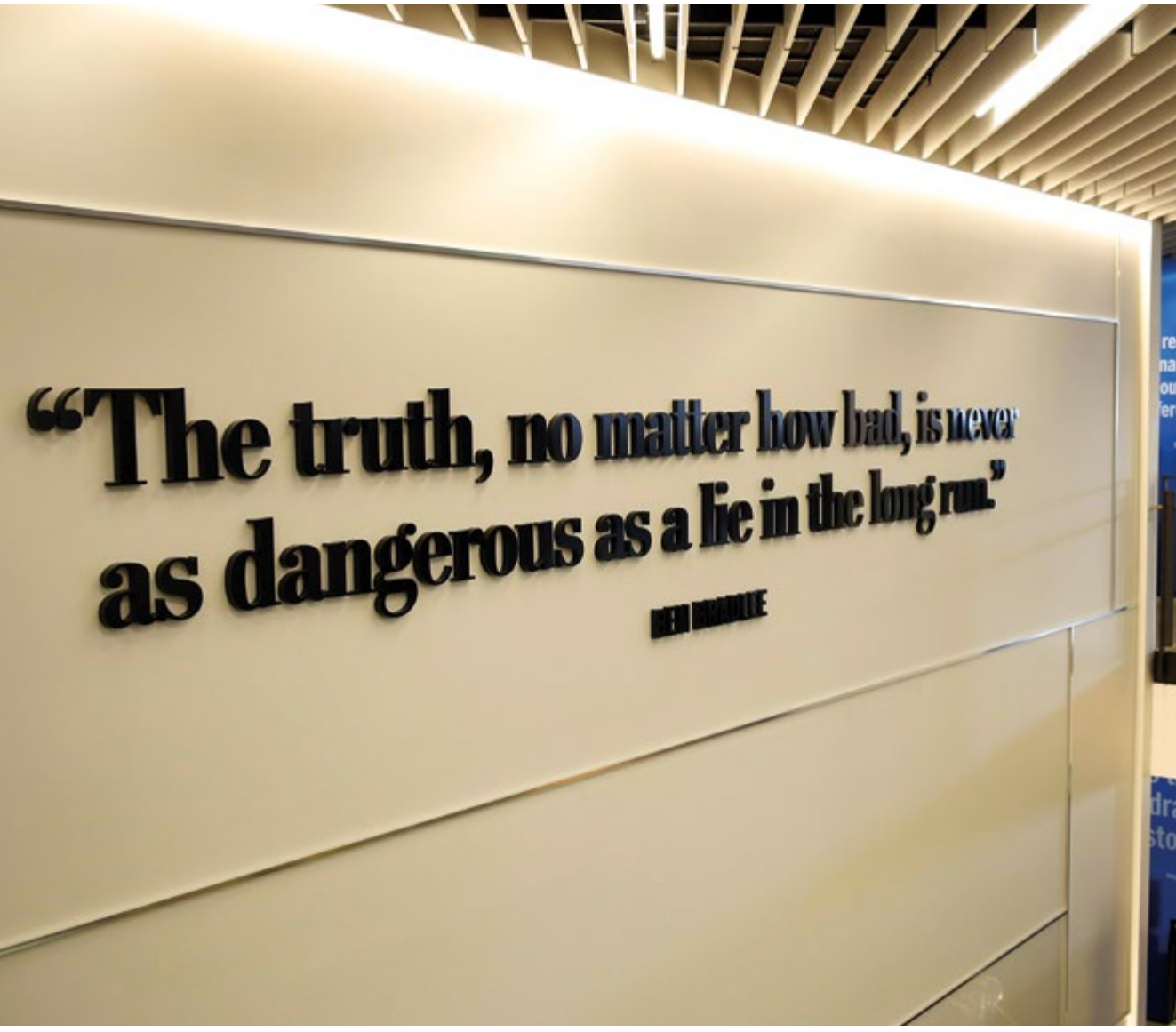
By monitoring compliance with journalistic due diligence, the media authorities are addressing an issue that lies at the heart of opinion formation. Anyone who takes advantage of people's trust in journalism and designs their offering in a way that inspires confidence also has to accept the responsibility that this entails and to work in a professional manner. Naming sources, the correct handling of citations, and meticulous research all increase the reliability of the information service and – in addition to the direct impact of sanctions for violations – provide a counterweight, because information is one of the best ways of countering disinformation.

It is also in line with the idea of easy discoverability, something that is reflected in the Interstate Media Treaty.[11] As of September 2021, all media offerings that make a significant contribution to shaping opinions in Germany must

be provided on user interfaces with easy discoverability. The criteria for this include a high proportion of news reporting, regional and local information, and the predominant use of trained, professional journalists in producing programme content. If users can find these kinds of information services quickly and easily, it makes it much more difficult for deliberate disinformation to get through. In order to maximise the effect of using information to combat disinformation, it is vital to improve and promote media literacy in society.

Tools against Disinformation

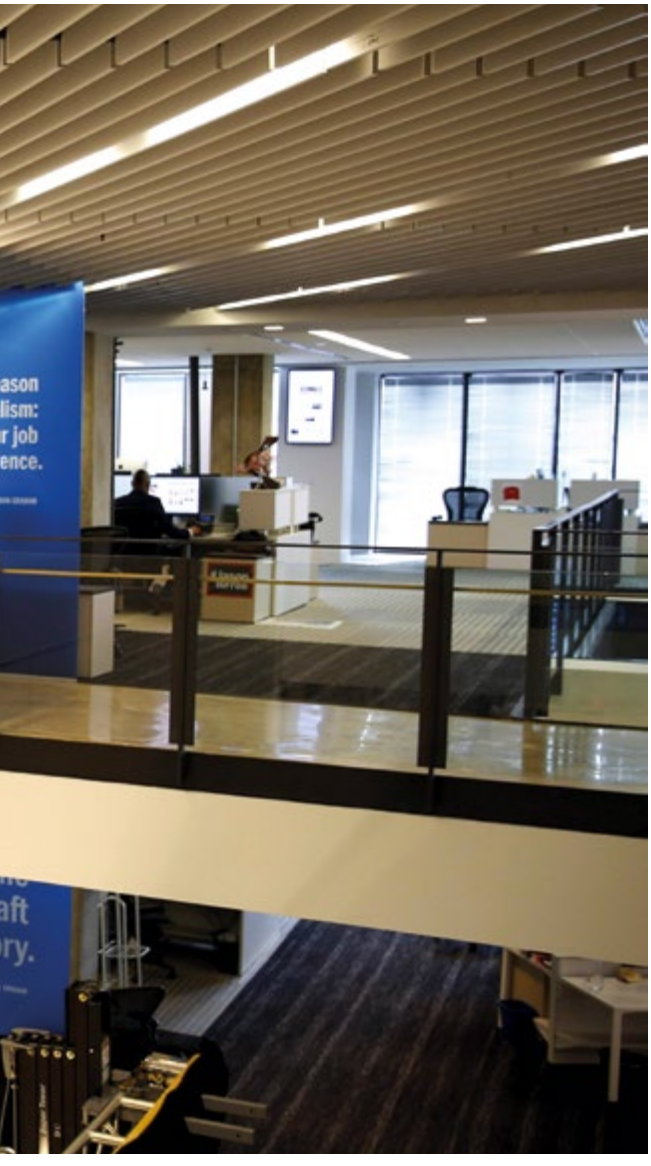
But beyond the approaches taken so far, what more can be done to mitigate the threat of disinformation while protecting freedom of expression and the opinion formation process? The change will certainly not be completed with the previously described expansion or shift of this process to the digital sphere. However, the current status of the discussions shows that the debate itself always lags behind tangible developments. This makes it all the more important



to counter disinformation, not only in its current form, but also to create a regulatory environment that contains abstract mechanisms.

Any such approach must be based on treating the expression of opinions separately from their content. It is also important for all posts to remain in the public discourse for as long as possible. Freedom of expression is a precious asset in democratic societies. Everything possible should be done to support but not interfere with the opinion formation process. As a result, many unpleasant things can and must be tolerated. Steps taken to protect freedom of expression and opinion formation always presuppose an interplay between projects promoting media literacy and legal frameworks. Regulation should only be introduced to support the opinion formation process when it is unable to deal with the factors that affect it.

One example of such support is transparency rules. They can be established and monitored independently of the content of an expressed opinion. Transparency can eliminate information deficits without changing the content itself. This can also make certain behaviours – such as covertly buying followers or likes – less attractive. A post that is displayed frequently but bears a clear indication that its reach has been artificially enhanced has much less potential for manipulation.



Interventions that prevent certain forms of expression must remain a last resort.

However, transparency also provides a basis for the discussion process. If all parties to a discussion have an equal amount of information, this enables them to classify a post correctly. For example, this might be the case with people who exert greater influence on public opinion because of their prominent position in society. Once this fact is made transparent, it is much easier to classify an expression of opinion. This is an advantage that should be available to everyone. But this does not mean an end to anonymous or pseudonymous online communication, as people who communicate in this way are unable to benefit from their social status.

Quote in the Washington Post's newsroom: "The truth, no matter how bad, is never as dangerous as a lie in the long run." Source: © Gary Cameron, Reuters.

Due diligence obligations are another instrument for combating disinformation. They already apply to broadcasting and journalistically designed telemedia in the form of the obligation to observe journalistic principles – as outlined above – and are monitored by state media authorities and self-regulatory bodies, such as the German Press Council. Due diligence obligations only indirectly target the outcome of the sourcing, aggregation, and presentation of information and take into account the underlying craft in the production of news and opinion. Akin to transparency rules, they should, therefore, be regarded as content neutral. The instrument of due diligence can be handled flexibly and proportionately in terms of both the scope of its application and the parties that it addresses in the opinion formation process.

Interventions that prevent certain forms of expression must remain a last resort. Prohibition can only be considered if the aforementioned obligations prove insufficient. For example, in the event that misuse of the platforms' technical infrastructure means a post is given a prominence that is not reflected in the public debate and thus only serves to distort the formation of public opinion.

Monitoring of future regulations in this area should also draw on ERGA's experience in its above-mentioned assessment of the Code. The aforementioned lack of access to information, which would make it possible to assess the Code more effectively, could also be remedied by shifting the burden of proof. When regulators identify systemic failures, they would report them to the platform operator, who would then be required to prove that no breach has occurred. This would solve a structural problem and allow operators and regulators to reduce their personnel costs. This is because it is difficult for the platforms to judge what data is necessary for the regulators to conduct a full assessment. However, the regulators' lack of knowledge about company structures means they cannot define which precise data they need to do their work. In addition to these instruments, ERGA[12] calls for the introduction of a regular review of the implementation of the Code and the possibility for the media regulators or for ERGA to issue a formal public reprimand to ensure they can point out deficits as appropriate.

A Graduated Regulatory Approach

Overall, the aforementioned instruments and the classification of various phenomena under the heading of disinformation should make it possible to take a content-neutral view and adopt an appropriate, proportionate, and graduated response. It is also important to stress that this also avoids the difficulty of evaluating whether statements are true or false, something that is highly subjective. On the basis that the right to freedom of expression protects any statement as long as it does not cross the line to become a punishable offence, these subjective standards must be disregarded in any objective regulation. The graduated regulatory approach involves measures that affect both content creators and communication platforms. In this way, it fosters an appropriate distribution of responsibilities between these two key players in the communication and opinion formation process and provides a framework for social discourse under these – no longer particularly new – conditions.

– translated from German –

Dr. Tobias Schmid is Director of the Media Authority of North Rhine-Westphalia. He is also the European Affairs Commissioner of the Conference of Directors of the Media Authorities (DLM) and Chair of the European Regulators Group for Audiovisual Media Services (ERGA), the association of national media regulators in Europe.

Disinformation – Categories, Actors, and Counterstrategies

Daphne Wolter

Disinformation often has a political background and aims to manipulate public debate or damage the reputation of a person or institution. Especially actors from authoritarian states use targeted campaigns in an attempt to exert political influence, undermine democratic debate, and increase social polarisation. Authoritarian regimes also seem to benefit from the digital revolution by using their citizens' data as a way of controlling and manipulating them. This is why the laws relating to human rights, copyright, and data privacy that apply in the analogue world also have to be constantly defended in the digital sphere.

This systemic rivalry is particularly obvious when key elections are being held and targeted disinformation campaigns are used in an attempt to influence public opinion. Germany and other EU Member States have a duty to protect their open democracies from such influence. Therefore, in addition to existing legislative initiatives and task forces,[13] it is important to educate the public on this issue and to build resilience. For instance, it is only by understanding how messaging works on such platforms that we are better equipped to identify disinformation and protect ourselves from it. Using regulation to directly combat disinformation is a difficult balancing act: a "law against disinformation" drafted in Germany to protect freedom of expression could be "repurposed" in authoritarian states to suppress and restrict freedom of expression by pushing their own narratives rather than true facts and thereby manipulating the public with fake news. In this respect, liberal democracies should ensure that potential laws are written so transparently and unambiguously that authoritarian regimes cannot interpret them in such a way that plurality of opinion and the media could be severely impaired. This also applies in the event that these democracies themselves experience unfavourable domestic power shifts.

What Are the Different Types of Manipulative Disinformation?

Fake news is false or misleading information that is circulated with the intention of harming a person, institution, or organisation. Rumours and false reports are supported by fake "evidence" and combined into one post. Corresponding posts from other users then flow into the supposed "chain of evidence". This can result in entire fake plots. Images are also often taken out of context in order to deliberately change a story.

Deepfakes are a subcategory of fake news that use the persuasive power of audiovisual media to achieve their manipulative effect. These are electronically modified moving images or photographs that alter or simulate people and events.

Social bots are machine-controlled and programmed profiles on social media. They pretend to be normal human users, so they usually have a photo and a made-up name. Their aim is to influence social interaction and opinion formation on social networks by spreading fake news.

Trolls are human users. They specifically try to disrupt or interrupt discussions on social media. Trolls try to polarise, provoke, and vilify other users by calling them trolls.

Who Are the Perpetrators and What Are the Motives?

Disinformation often has a **political background**. It is organised directly by **state or non-state actors**. In countries without stable democratic conditions, for example, content can also be disseminated and thus amplified by the state-controlled media.

Other motivations can be entertainment (in the negative sense) and **attention-seeking**. Deliberate provocations designed to annoy and challenge have, unfortunately, long been a hallmark of the online culture. In most cases, however, these campaigns also have a substantive political goal; the vehicle for this – often via memes or deepfakes – is **entertainment**.

Finally, **advertising** can also be a financial motive behind disinformation campaigns. These campaigns aim to generate as much traffic as possible. They manipulate content to get a higher click-through rate for their adverts. Politically emotive topics are often used as clickbait.

What Increases the Effectiveness of Disinformation?

Disinformation campaigns are run by different groups of perpetrators and have various motivations. But the breeding ground is always the same:

- The growing importance of social media as a source of news
- A polarised political landscape
- Lack of trust in traditional media

Emotive topics have strong potential to go viral. In order to appear as genuine as possible, fake sources are quoted and media logos can sometimes be misused.

What Can Civil Society Do about it?

Digital disinformation is an ongoing threat. It will also evolve in line with technological advances.

News, research, and information literacy must be expanded in every age group. Through systematic clarification, state institutions, authorities, and above all journalists in their reporting can contribute to highlighting and preventing the problem of disinformation.

Personal responsibility – every single person can take responsibility for preventing the spread of fake news. If the source of a news item is unknown or cannot be traced, there is a good chance that it is fake news. Linguistic inaccuracies are also often a hallmark of disinformation.

What gives us hope? There is a growing demand for **quality journalism** among internet users. This offers a great opportunity for newspapers and broadcasters to also provide reliable information on the web. This would require the legal possibilities relating to discoverability to be adapted accordingly.

– translated from German –

Daphne Wolter is Policy Advisor Media in the Analysis and Consulting Department at the Konrad-Adenauer-Stiftung.

1 Klicksafe 2021: forsa survey on Safer Internet Day 2021, in: https://bit.ly/3BXGR8X [30 Jul 2021]
2 European Commission 2019: Action Plan against Disinformation: Progress Report, Jun 2019, in: https://bit.ly/3il16Wx [3 Aug 2021].
3 European Commission 2021: Shaping Europe's digital future: Code of Practice on Disinformation, 13 Jul 2021, in: https://bit.ly/3BVhjte [30 Jul 2021].
4 European Regulators Group for Audiovisual Media Services (ERGA): https://erga-online.eu [3 Aug 2021].
5 ERGA 2019: ERGA Report on Disinformation: Assessment of the Implementation of the Code of Practice, in: https://bit.ly/3iejJLI [30 Jul 2021].
6 European Commission 2020: Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions: European Democracy Action Plan, 3 Dec 2020, in https://bit.ly/3zO0OgI [30 Jul 2021].
7 European Commission 2021, n. 3.
8 European Commission 2020, n. 6.
9 European Commission: The Digital Service Act: ensuring a safe and accountable online environment, 15 Dec 2020, in: https://bit.ly/378T2lj [30 Jul 2021].
10 Die Medienanstalten 2020: Interstate Media Treaty (MStV), 7 Nov 2020, in: https://bit.ly/3j2nmEO [3 Aug 2021].
11 Ibid., paragraph 84 MStV.
12 ERGA 2021: ERGA position on the next instalment of the Code of Practice on Disinformation, May 2021, in: https://bit.ly/3zP6Eyq [30 Jul 2021].
13 EUvsDISINFO: https://euvsdisinfo.eu [3 Aug 2021].