Human Rights Outsourcing and Reliance on User Activism in the DSA

Martin Senftleben

21.02.2024 doi: 10.59704/01365e5a9a872007

Introduction

Article 14(4) of the Digital Services Act (DSA) places an obligation on providers of intermediary services, including online platforms hosting user-generated content (see Article 3(g) DSA), to apply content moderation systems in "a diligent, objective and proportionate manner." The provision emphasizes that online platforms are bound to carry out content filtering with due regard to the fundamental rights of users, such as freedom of expression. Considering the central role of online platforms in the current media landscape, this regulatory attempt to safeguard the right of users to share and receive information does not come as a surprise. However, fundamental rights, including freedom of expression (Article 11(1) EU Charter of Fundamental Rights), have been designed as rights to be invoked against, and nurtured by, the state. Against this background, the approach taken in Article 14(4) DSA raises complex questions. Does the possibility of imposing fundamental rights obligations on intermediaries, such as online platforms, exempt the state power from the noble task of preventing inroads into fundamental rights itself? Can the legislator legitimately outsource the obligation to safeguard fundamental rights to private parties (see the contribution by Geiger/Frosio for a discussion of digital constitutionalism)?

In the case of user uploads to online content-sharing platforms, Article 17(7) of the Directive on Copyright in the Digital Single Market (CDSMD) adds an important guideline to the general obligation laid down in Article 14(4) DSA (see Article 2(4)(b) DSA as to the complementary application of these rules): the cooperation between online platforms and the creative industry in the area of content moderation (Article 17(4) CDSMD) must not result in the blocking of non-infringing content uploads, including situations where user-generated content falls within the scope of a copyright limitation that supports freedom of expression, such as the exemption of quotations, parodies and pastiches (explicitly mentioned in Article 17(7) CDSMD, see also the more detailed discussion here).

Joint Effort of Creative Industry and Platform Providers

Evidently, this outsourcing scheme for human rights obligations relies on a joint effort of the creative industry and the online platform industry. To set the content filtering machinery in motion, copyright holders in the creative industry must notify "relevant and necessary information" with regard to those works which they want to ban from user uploads (Article 17(4)(b) CDSMD). Once relevant and necessary information on protected works is received, the online platform is obliged to include that information in the content moderation process and ensure the unavailability of content uploads that contain traces of the protected works.

Unlike public authorities, however, the central players in this cooperation scheme are private entities that are not intrinsically motivated to safeguard the public interest in the exercise and furtherance of fundamental rights and freedoms. Despite all invocations of diligence and proportionality in Article 14(4) DSA, the decision-making in the context of content filtering is most probably much more down to earth: the moment the balancing of competing human rights positions is confidently left to industry cooperation, economic cost and efficiency considerations are likely to occupy centre stage (see already the contribution by Goldman/Schwemer).

A closer look at the different stages of industry cooperation resulting from the described regulatory model confirms that concerns about human rights deficits are not unfounded. As explained, the first step in the content moderation process is the notification of relevant and necessary information relating to "specific works and other subject matter" by copyright holders (Article 17(4)(b) CDSMD). In the light of case law precedents, in particular *Sabam/Netlog* (para. 51), use of the word "specific" can be understood to reflect the legislator's hope that copyright holders will only notify individually selected works. Otherwise, content moderation may reach proportions that violates freedom of expression and information, and other fundamental rights (see Angelopoulos & Senftleben 2021). In *Sabam/Netlog*, the Court declared content filtering based on a whole repertoire of collecting society repertoire excessive and impermissible (paras 48-51).

Seeking to avoid the evolution of an overbroad, general filtering obligation, a copyright holder could limit use of the notification system to those works that constitute cornerstones of the current exploitation strategy. As a result, other elements of the work catalogue would remain available for creative remix activities of users. This, in turn, would reduce the risk of overbroad inroads into freedom of expression and information.

In practice, however, rightholders are unlikely to adopt this cautious approach. The success of the risk reduction strategy surrounding the word "specific" is doubtful. In the cooperation with online platforms, nothing seems to prevent the creative industry from sending copyright notifications that cover each and every element of impressive work catalogues. Platforms for user-generated content may thus receive long lists of all "specific" works which copyright holders have in their repertoire. Adding up all works included in these notifications, the conclusion can become inescapable that the regulatory approach underlying the described interplay of rules in the DSA and the CDSMD culminates in a filtering obligation that is very similar to the filtering measures which the CJEU prohibited in *Sabam/Netlog*. The risk of encroachments upon human rights is evident (see also Senftleben 2024).

Impact of Cost and Efficiency Considerations

Turning to the second step in the content moderation process – the act of filtering carried out by online platforms to prevent the availability of notified works – the aforementioned proportionality and diligence obligations apply: content moderation must comply with the diligence and proportionality requirements in Article 14(4) DSA. As to the practical outcome of content filtering in the light of diligence and proportionality requirements, however, it is to be recalled that online platforms will most probably align the concrete implementation of content moderation systems with cost and efficiency considerations. In reality, the subordination of concrete industry decisions to abstract diligence and proportionality imperatives – the acceptance of more costs and less profits to reduce the corrosive effect on freedom of expression and information – would come as a surprise. Online platforms can be expected to be rational in the sense that they seek to achieve content filtering at minimal costs. A test of proportionality is unlikely to occupy centre stage unless the least intrusive measure also constitutes the least costly measure. A test of professional diligence is unlikely to lead to the adoption of a more costly and less intrusive content moderation system unless additional revenues accruing from enhanced popularity among users offsets the extra investment of money.

Hence, there is no guarantee that industry cooperation in the field of user-generated content will lead to the adoption of the most sophisticated filtering systems with the highest potential to avoid unjustified removals of content mash-ups and remixes (further examined here). An assessment of liability rules also confirms that excessive filtering risks must be taken seriously. An online platform seeking to minimize the risk of liability is likely to succumb to the temptation of overblocking. Filtering more than necessary is less risky than filtering only clear-cut cases of infringement. After all, primary, direct liability for infringing user uploads follows from Article 17(1) CDSMD and dangles above the head of providers of platforms for

user-generated content like the sword of Damocles. The conclusion is thus inescapable that the outsourcing strategy underlying the EU regulation of content moderation in the DSA and the CDSMD is highly problematic. Instead of safeguarding human rights, the regulatory approach is likely to culminate in human rights violations.

Reliance on User Complaints

Against this background, it is of particular importance to analyse mechanisms that could bring human rights deficits to light and remedy shortcomings. This question requires the discussion of the role of users. Article 14(1) DSA and Article 17(9) CDSMD both make users the primary addressees of information about content moderation systems. According to Article 14(1) DSA, users shall receive information on upload and content sharing restrictions arising from the employment of content moderation tools. If they want to take measures against content restrictions, Article 17(9) CDSMD – and the complementary provisions in Article 20 DSA – ensure that complaint and redress mechanisms are available to users of OCSSP services "in the event of disputes over the disabling of access to, or the removal of, works or other subject matter uploaded by them."

Hence, users are expected to instigate complaint and redress procedures at platform level and, ultimately, go to court. The reliance placed on this mechanism, however, is surprising. Evidence from the application of the DMCA counter-notice system in the U.S. shows quite clearly that users are unlikely to file complaints in the first place. This is confirmed by data from recent transparency reports from the largest user-generated content (UGC) platforms (examined by Senftleben, Quintais & Meiring 2023). If users must wait relatively long for a final result, it is foreseeable that a complaint and redress mechanism that depends on user initiatives is incapable of safeguarding freedom of expression and information. Moreover, an overly cumbersome complaint and redress mechanism may thwart user initiatives from the outset.

In the context of user-generated content, it is often crucial to react quickly to current news and film, book and music releases. If the complaint and redress mechanism finally yields the insight that a lawful content remix or mash-up has been blocked, the decisive moment for the affected quotation or parody may already have passed. From this perspective, the elastic timeframe for complaint handling in Article 17(9) CDSMD – "shall be processed without undue delay" – gives rise to concerns. This standard differs markedly from an obligation to let blocked content reappear promptly. As Article 17(9) CDSMD also requires human review, it may take quite a while until a decision on the infringing nature of content is taken. Considering these features, the complaint and redress option may appear unattractive to users (see Senftleben 2020).

Instead of dispelling concerns about human rights deficits, the reliance on user complaints, thus, constitutes a further risk factor. Apart from being ineffective as a remedy for human rights violations, it may allow authorities to hide behind a lack of user activism and thereby conceal human rights deficits. It may also be that users refrain from complaining because they consider the mechanism too cumbersome and/or too slow. However, when taking the number of user complaints as a yardstick for assessing human rights risks, a relatively low number of user complaints may be misinterpreted as evidence that content moderation does not lead to excessive content blocking. If users refrain from taking action, human rights deficits stay under the radar. The oversimplified equation "no user complaint = no human rights problem" offers the opportunity of presenting potentially overly restrictive content moderation systems as a success. Instead of shedding light on human rights deficits, the complaint and redress mechanism can be used strategically – by platforms and regulators alike – to conceal encroachments upon freedom of expression and information.

Conclusion

In sum, closer inspection of DSA and CDSMD content moderation rules confirms a worrying tendency of reliance on industry cooperation and user activism to safeguard human rights. Instead of putting responsibility for detecting and remedying human rights deficits in the hands of the state, the EU legislature

prefers to outsource this responsibility to private entities, such as online platforms, and conceal potential violations by leaving countermeasures to users. The risk of eroding freedom of expression is further enhanced by the fact that, instead of exposing and discussing the corrosive effect of human rights outsourcing, the CJEU has already rubberstamped the described regulatory approach. In its *Poland* decision (see Quintais 2022 and Husovec 2023), the Court has even qualified problematic features of the outsourcing and concealment strategy as valid safeguards against the erosion of freedom of expression and information (see further Senftleben 2024).

To safeguard human rights, the state power itself must become much more active. Litanies of due diligence and proportionality obligations for private entities and reliance on user activism are not enough. Requirements for audit reports under Article 37 DSA should include the obligation to provide sufficiently detailed information on the implementation of human rights safeguards to allow the European Commission to exercise effective control and prevent encroachments (see Articles 42(4), 66(1), 70(1), 73(1), 74(1) DSA). The implementation of Article 17 CDSMD in national legislation should only be deemed satisfactory when the Member State has devised effective legal mechanisms to ensure that content filtering measures do not erode the freedom of users to upload quotations, parodies and pastiches (Article 17(7) CDSMD). Moreover, the research community should be encouraged to throw light on violations of freedom of expression and information when analysing platform data (Article 40(4) and (12), 34(1)(b) DSA).

ABOUT THE AUTHOR

Martin Senftleben

Martin Senftleben is Professor of Intellectual Property and the Director of the Institute for Information Law (IViR) at the University of Amsterdam. He works as an Of Counsel at Bird & Bird in The Hague.Martin Senftleben is Professor of Intellectual Property and the Director of the Institute for Information Law (IViR) at the University of Amsterdam. He works as an Of Counsel at Bird & Bird in The Hague.

SUGGESTED CITATION

Senftleben, Martin: Human Rights Outsourcing and Reliance on User Activism in the DSA, Verfassungsblog, 21.02.2024, doi: 10.59704/01365e5a9a872007.

Licensed under CC BY-SA 4.0.

Langzeitarchiviert bei IntRechtDok unter: doi.org/10.17176/20240222-004320-0.