as part of the

Landecker Digital Justice Movement

Statement on the Proposal for the

Regulation of the European Parliament and of the Council

on a Single Market for Digital Services (Digital Services Act) and

amending Directive 2000/31/EC (COM(2020) 825 final)

with focus on hate speech and digital violence

on online platforms and in social networks

Supported by ALFRED LANDECKER FOUNDATION

Our recommendations for the DSA

I. Content moderation: effective protection of those affected from digital violence (art. 14, 15)

1. **Codifying an obligation to delete illegal content in an article.** This way it is enforceable by authorities and can be subject of a complaint to the Digital Services Coordinator, even if contravention is practiced exhaustively.
2. **Codifying recital 33 in an article** to ensure its priority in the application of the law by the national courts. This creates legal certainty with regards to the contradicting regulation of the e-commerce directive. Revise art. 40 sec. 1 and create exemptions according to art. 3 sec 4 e-commerce-directive.
3. **Lowering the requirements to reverse burden of proof according to art. 14 sec. 3** and exclude the obligation to provide the user's name, e-mail address and Url of the content.
4. Creating a distinct **deadline for assessment of notifications** of 7 days.
5. Creating an obligation to provide a **statement of reasons for all content decisions**. This enables user to assess the prospects of legal action.
6. Obliging platforms to create a **notification procedure** that is **clearly visible, low- threshold and located close to the content** in question.

II. Internal complaint handling & out of court settlement (art. 17, 18)

1. **Make internal complaint handling and out of court settlement accessible to all users**, not only for uploaders. Otherwise, the situation of users that report potentially illegal content and whose notification is rejected does not improve under the DSA.
2. **Recreate the out of court settlement mechanism and replace it by mandatory summary proceedings for content decisions in all Member States.** Binding content decisions need to be reserved for courts and summary proceedings can significantly improve access to justice for all users towards online platforms.

III. Measures against misuse (art. 20)

1. **Do not overestimate the effect of suspension of users.** According to our experience most of the hate and agitation is spread by users that hold multiple profiles and will most likely proceed with another one. This is fairly easy as such profiles can be created easily with random data without any verification.
2. **Do not suspend users from notification procedure.** It cannot be expected that users can assess the lawfulness of content and it is not their task either. Negative consequences of misuse of notification results from wrongful and often automated complaint handling by the platform and waiver of human oversight.

supported by ALFRED LANDECKER FOUNDATION

IV. Point of contact (art. 10)

Oblige online platforms to maintain a point of contact in every member state that is accessible for users and not only authorities. The point of contact should guarantee for legally secure delivery of requests and be accessible in one of the countries official languages. This can be as simple as mandating a law firm.

V. Rethink oversight under consideration of the country-of-origin principle

Keep in mind that all Social Media Platforms are registered in Ireland and that alone Ireland will be in charge to manage oversight. This enables platforms to cherry-pick their regulator.

VI. Find a way to restrain harmful services registered abroad

Smaller platforms registered outside of the EU currently create a safe haven for all sorts of illegal content. In order to ensure public safety from these services we need reliable options to restrain these services or even block them in the EU.

supported by ALFRED LANDECKER FOUNDATION

Table of content

I. Who we are

HateAid gGmbH (HateAid) is a non-profit organization and was established in 2018. We are a counselling center and litigation financer that strengthens and supports people affected by digital violence. This is necessary as victims of online hate have been so far left largely unprotected in Germany. We support those affected by digital violence through emotional support as well as security and communication counselling. Furthermore, HateAid offers support in law enforcement. As a financier of legal costs, HateAid enables those affected in taking (civil) legal action against perpetrators and platforms and thus provides them with access to justice while covering the costs. As part of its advocacy work, HateAid advises the judiciary and law enforcement agencies in Germany and other European countries on how to set the appropriate framework to curb digital violence. The aim of HateAid is to strengthen the diversity of opinions on the net and thus to maintain free and open democracy.

In this position we supported more than 800 victims through counseling and enabled almost 80 victims to file more than 400 criminal complaints and more than 80 litigation cases. This is also why we have profound experience regarding the struggles and needs of people defending themselves against digital violence. The German Network Enforcement Act (NetzDG) is one of the first attempts of platform regulation in Europe and explicitly tackles hate crime on social media. HateAid has repeatedly contributed its expertise to the legislative process and its further evaluation. Similarly, we offer unique insight resulting from our active (legal) support of those affected and can hereby provide a valid testimony of the effects of the NetzDG with all its success and shortcomings.

II. The hate speech crisis: a threat to democracy

On the internet and especially on Social Media Platforms targeted hate attacks are being used to drive people with certain traits (women, Jews, immigrants etc.) out of public online debates[1]. Consequently, perpetrators are accustomed to the fact that the internet is not only a legal vacuum where their alleged freedom of expression knows no boundaries, but they can furthermore fully rely on their anonymity. The result can be observed abundantly on all-major Social Media Platforms. **73 % of the young adults** in Germany between age 18 – 24 have already witnessed hate speech, **17 % of them have already been affected themselves**. 76 % of all internet users think that hate speech increased over the last years. As a result, the victims of such hate attacks withdraw from the public debate entirely.

1. A threat to freedom of expression and democracy

But the situation is even worse. Studies show that digital violence and hate speech online do not only affect the victims of the attack themselves but also the bystanders who witness the hate attacks. As a result, **54 % of the internet users in Germany do no longer dare to express their political opinion online and 47 % rarely participate in online discussions at all out of fear of becoming victims themselves.**[2] This development does not only **restrict the freedom of expression** of the many people who have become a victim themselves but also pose **a threat to our free and open public debate** as a whole as the internet is **no longer a safe space** for everybody but for a chosen few.

This so-called **silencing effect** poses a serious threat to democracy as it shifts the debate to a very one-sided public discourse. In Germany this is without a doubt a shift towards the far right: more than

[1] https://www.isdglobal.org/wp-content/uploads/2017/10/The-Fringe-Insurgency-221017.pdf

[2] Studie des IDZ, u.a., „#Hass im Netz, Der schleichende Angriff auf die Demokratie, Eine bundesweite repräsentative Untersuchung“; 2019: https://www.idz-jena.de/fileadmin/user_upload/_Hass_im_Netz_Executive_Summary.pdf

supported by ALFRED LANDECKER FOUNDATION

70 % of online hate content reported to the Federal Police Department was assigned to right wing extremist groups in 2019[3]. As a result, the internet does no longer represent the public opinion of everybody, but of very few that are very visible and pretend to be the majority of people. Those few systematically spread hate and agitation and consider the internet to be a safe haven while they are limiting the freedom of expression of a majority through the spread of fear.

2. EU Citizens loose trust in rule of law

It is Social Media Platforms that made this behaviour possible as they stood by and did not act. Those private corporations created a de facto public space which has become indispensable to our public debate and social participation of individuals. Their power has become even more significant in times of the pandemic. This development has led us to the absurd situation that private companies like Facebook, Google or Twitter set the standards of our public debate without any regulation. This happened while those affected were left largely defenceless against hate and agitation but also against Social Media Platforms themselves.

(a) The starting point: absurd conditions for victims of digital violence in the EU

Users who are affected by illegal content such as death and rape threats, defamation and insults are put in an impossible situation. The NetzDG in Germany as a first attempt on national regulation improved their situation partially but fails to grant efficient protection in many cases. The reason is that national legislation is not able to overcome the struggles that result from transnational data transmission and private companies that refuse to be held liable outside their country of origin.

Real life example: The struggle is real

This is a situation that those affected by digital violence on Social Media Platforms experience when trying to enforce their rights online. It is exemplary for a lot of cases HateAid is attending to daily in the consultation:

After a **criminal complaint** of a famous German journalist, the platform in question **denied information about perpetrators to prosecutors**. For this reason, our client chose **civil procedure to seek a court order** that allows (but not obliges) the platform to provide information. This produced costs of more than 2.000,00 EUR. Thereupon the **platform provided information**, which is **unusual** but has been expected from previous experience with famous clients. The information provided was mainly IP addresses, usernames and in some cases phone numbers. This data however does not enable private citizens to identify the account holder and was useless to the journalist in question. That is why **we forwarded the information to law enforcement agencies** that were indeed able to identify some of the perpetrators through their phone numbers. After inspection of the files - which requires to hire and pay a lawyer – our client got the information.

For reasons beyond understanding the platform did not delete the comments in question although there was a court order considering them to be illegal. Even after two notice& takedown letters by a lawyer it was still not deleted. **This is why our client had to file a lawsuit against the platform again with our support. This lawsuit already produced costs of more than 5.000,00 EUR which is not yet**

[3] Crime statistic of politically motivated offenses 2019: https://www.bka.de/DE/UnsereAufgaben/Deliktsbereiche/PMK/PMKrechts/PMKrechts_node.html#doc121714bodyText2

the total cost risk. The oral hearing was scheduled 11 months after the lawsuit was filed – consequently the content in question will stay visible at least one more year.

Needless to say, that this situation is hard to take in for anybody and overall unacceptable under rule of law.

Insufficient options of those affected in Germany

Users who experience digital violence have the following options:

(a) A **criminal complaint** is rarely successful and will naturally never result in a removal of the content in question. In the best case the prosecution can identify the perpetrator so that the affected user can examine the file and thus get information. Identification can only be achieved in 1/3 of our cases which are already filed at a specialised cybercrime department – partly due to lack of cooperation of Social Media Platforms.

(b) If the perpetrator can be identified, affected users can **litigate** against them and pursue their claim for cease and desist and damages. We consider litigation to be an effective and comprehensive legal measure to pursue one's rights against a single perpetrator effectively. At the same time, it imposes high financial risks on the affected user, it is lengthy and complex and because of this victim rarely chose it to be a viable option.

(c) Lastly, affected users can directly get in touch with the **platform** and demand removal of the content. Evidently this option is preferable as it avoids costs and lengthy procedures. Unfortunately, content is rarely removed, and decisions are highly arbitrary. Getting in touch with platforms is already a challenge as it requires victims to search the imprint for an address and struggle with legally secure delivery of documents abroad that is necessary for litigation.

(b) Users loose trust in the ability of the EU to protect them

Not a day goes by without clients in our consultation telling us they do not file criminal complaints anymore or do not report illegal content to the platforms because of countless frustrating experiences of not being heard or ignored. Many of them express that the internet does not feel safe anymore. They have experienced to be powerless when their rights are being violated and that Social Media Platforms usually ignore their complaints and enquiries.

No effective legal defence for users

This situation poses an **exceptional challenge** to national legislators. **Digital violence brings judicial systems to their limits and puts the rule of law to a test which it is about to fail**. **Neither law enforcement nor litigation on a national level can offer effective legal defence to affected users.** Nation states are incapable to overcome the difficulties that transnational online communication and privately owned data brings along. This situation doubles the risks of digital violence to our democratic societies: It does not only drive people out of the public debate but also **weakens the trust of citizens in the rule of law and their public safety on the internet in general**. As a result, perpetrators and affected users alike consider the internet to be a lawless space and have adjusted their behaviour accordingly.

(c) Imbalance of power: The role of Social Media Platforms

In the end, the current situation of users is an evidence of incapacity. Users and law enforcement agencies are highly **dependent on the platform's goodwill.** Especially when the perpetrator's identity is unknown users are largely unprotected and have no option to obtain legal protection against defamation, insults, or death threats. The platforms are aware that users will not institute legal proceedings against then due to the cost risk and **imbalance of (financial) power**. This situation leads to **the apathy of users** as described above.

We need to acknowledge that every solution to tackle the Hate Speech crisis needs to take in account the role of Social Media Platforms and their liability for the risks they create. They chose their business model well when they decided to open their platforms to masses of users with only little precautionary measures. Now we need to find a way to deal with the harmful consequences that this business model created not only for the citizens of the EU but also for society as whole. Platforms need to take responsibility to **protect users** and **contain the negative effects** of their business model just as it is also required from industries that sell dangerous goods or services such as drugs or food. This may also require **financial efforts** from online platforms on a large scale to e.g., improve content moderation that can easily be expected from platforms such as Facebook with an annual turnover of 85 million USD (2020)[4].

III. The Digital Services Act must protect those affected from digital violence

From our perspective the Digital Services Act is **a unique chance to create equality of arms for users of Social Media Platforms**. We are convinced that a thorough compliance scheme and oversight will not eliminate digital violence. **Therefore, the protection and support of users affected by digital violence should be a central part of every step of any platform regulation.** This approach consequently reflects the basic needs of users across the European Union who are largely left without rights facing online platforms that appear superior and inaccessible.

1. Our expectations for the Digital Services Act

To protect the rights of those affected from digital violence the Digital Services Act needs to focus on the following aspects:

Protection against digital violence is protection of human rights

Taking the Human Rights Perspective, Article 8 of the ECHR obliges states to protect the personality rights that include the personal honour and identity as well as the right to one's own image (as parts of the private life) from publications in an adequate manner. States shall not tolerate infringements of Article 8 of the ECHR by private parties. This obligation is further amplified by art. 13 of the ECHR, stating that the duty to implement effective criminal law and law enforcement to prosecute violations of art. 8 ECHR is to be accompanied by appropriate possibilities to lodge a complaint. We must acknowledge that we face considerable shortcomings to live up to these principles in a digital environment in the EU.

Furthermore, according to art. 13 of the ECHR, the use of individuals' complaints must not be obstructed by authorities refraining from fulfilling their duties. Considering the central position of the

[4] https://www.statista.com/statistics/268604/annual-revenue-of-facebook/

supported by ALFRED LANDECKER FOUNDATION

fundamental rights as guaranteed by the ECHR in art. 6 sec. 3 of the TFEU, these obligations must be kept in mind as general guiding principles when elaborating the DSA.

Empowerment of users

This necessarily means that users must be offered reliable options to act against digital violence. Consequently, they must be enabled to act against wrongful content decisions and take effective legal action if their rights are violated on or by Social Media Platforms. It also means that platforms must be obliged to assess users' requests with clearly defined and transparent procedures that leave no space for arbitrariness. It needs to be ensured that platforms are obliged to install user-friendly and comprehensible procedures for complaints, content decisions are transparent and can be subjected to judicial review by users in a reasonable manner. And it is even more important to make sure that those obligations are well-defined, predictable, and specific so they cannot be evaded by shady tricks. A prominent example of such tricks is how difficult it is to find and use reporting mechanism in Germany although the NetzDG is in effect with a wording which is similar to that of the proposal.

Moreover, platforms need to be accessible for users by establishing contact points. So far, they are allowed to establish their headquarter in one member state and still deliver their services everywhere in the European Union. The victims of digital violence that want to redress content decision need to search the imprint for valid information and find a way for legally secure delivery abroad in order to take legal action against the platform.

Demand action from platforms

It must be in our mutual interest to remove illegal content from Social Media Platforms that affect the rights of other users. Therefore, obligations need to be imposed on platforms to support affected users and not obstruct their efforts to defend themselves against digital violence. Especially large platforms claim to be overstrained with this task. Platforms need to be held responsible and have a significant contribution not only to prevent digital violence but also to support users that have been affected affected by digital violence in their field.

Strengthen regulators

Effective platform regulation cannot only rely on the empowerment of users. Effective oversight needs to be installed which is able put pressure on online platforms that benefit from extensive liability privileges. The regulator can only be effective if not only privileges and exemptions, but also enforceable obligations are defined. With regards to the country-of-origin principle effective oversight needs to take into account the fact that all major Social Media Platforms are established in Ireland and that Ireland so far made no efforts to constrain the harmful effects of Social Media companies.

2. The proposal of the Digital Services Act: general assessment

The Commission's proposal offers only little improvement for the situation outlined above and falls short of our expectations:

Safety by design & compliance

In general, the proposal pursues a safety by design approach to platform regulation and sets a strong focus on transparency, supervision, and compliance. This approach is welcomed in principle as it installs supervision of Social Media Platforms for the first time and is designed to prevent structural

infringements. We should not only rely on prevention but also implement powerful means to enable users to defend their rights on online platforms.

Content moderation

Moreover, in terms of content moderation HateAid endorses the fundamental decision to create a notice and action mechanism (art. 14) that enables users to report illegal content on all online platforms. Nevertheless, its declaration in the proposal as an "action" mechanism cannot be approved because in fact the only "action" required by platforms after a notification is that the notification is to be processed. This cannot be qualified as a valid "action"- mechanism. Above all the proposal does not codify an obligation to block or remove illegal content. This is especially deplorable as it obstructs authorities and particularly the Digital Services Coordinator as a body of complaint for users (art. 43) to enforce the removal of illegal content. Consequently, it puts users in the position to be (again) powerless against potentially wrong content decisions.

HateAid approves that platforms should provide a **statement of reasons (art. 15)** when content is removed. By now content decisions are not comprehensible at all for users and we hope that a statement of reasons enables users to assess their options of legal protection more clearly. HateAid also approves the measures that intend to prevent the misuse of the notification mechanisms (art. 20) and the put back mechanism (art. 17,18) in principle.

HateAid also understands the intention of the installation of **an internal complaint handling system (art. 17)** and **out of court dispute settlement (art. 18)** as they obviously aim to revise content decisions and avoid overblocking. Still, it is surprising that on the one hand online platforms shall not be addressees of a binding and enforceable obligation to delete illegal content but on the other hand a largely undefined "external body" shall be enabled to make binding content decision according to art. 18. In our opinion this approach comes with the right intention to accelerate decisions and create a low threshold option for users. But it is also contrary to the proposals intention to take content decisions out of the hands of private institutions and reserve them for courts. In our opinion the uneven access to this out of court settlement only for those affected from removal but not for those affected by the denial of removal is not justified.

We consider the approach to **specify terms and conditions** (art. 12) to be helpful as this provides a possibility to subject decision based on terms and conditions, which is the vast majority, to a more profound legal review. The right to address claims by class action promises to offer new options to people affected by digital violence and is welcomed (art. 68).

Considerable shortcomings in user protection

We regret that the commission did not take the chance to offer users of online and Social Media Platforms a comprehensive solution to defend themselves legally against the violation of their (human) rights on and by Social Media Platforms. Instead, users are supposed to be burdened with the task to search for illegal content while being left with absolutely no options if the platform denies their request for removal other than litigation. Especially this task should not be imposed to civil society through trusted flagger programs alone as it was recently described by Magarethe Vestager in an interview

with Spiegel Online.[5] This approach is only to the benefit of online platforms that save money and resources and shift responsibility to civil society and thus predominantly publicly- and donor-funded organisations.

Misguided focus on overblocking and misuse: learnings from NetzDG

Further the proposal fails to address basic needs of those affected to be able to enforce their rights against Social Media Platforms. Instead, it focusses on the fear of misuse of the notification mechanism and overblocking and thereby puts users who report illegal content under general suspicion. To make things absolutely clear: It is necessary to install a viable mechanism to prevent overblocking. The risk of overblocking as an unintended consequence of the mandatory removal of content is reasonable and needs to be addressed.

In our experience, however, these are not the core problems for those affected by digital violence. Looking at the latest evaluation of the NetzDG in September 2020 we can conclude that there is no indication of overblocking because of the obligation of the platforms to delete content after a defined process[6]. This is the case although the NetzDG does not only require Social Media Platforms to install a notification mechanism but also imposes a deadline of 24 hours for deletion if content is clearly illegal. Indeed, the findings of the evaluation were the opposite as they found out that only a small fraction of notifications leads to deletion (11 -27 % depending on the platform[7]). In many cases the platforms did not react at all.

Misuse or wrongful content decision?

Another result of the evaluation of the NetzDG is that there **is no indication of significant abuse of notification procedures**. In fact, there have been only very few cases where the notification mechanisms were actually misused by prearranged targeted attacks of mass notifications. These cases had one thing in common: They all resulted in **wrongful content decisions of the platforms.** These were caused by the number of notifications that triggered a mechanism to automatically block the content or profiles that were subject to the notification without human oversight. All those decisions could have been prevented through human oversight. The few cases that we saw overturned by the platform after the initial blocking of profiles or content were possible because of the intervention and direct communication of HateAid with the platform. A privilege that the majority of users do not have. **Therefore, we must be aware that whenever misuse of reporting channels creates overblocking we are in fact dealing with wrongful content decisions of Social Media Platforms. Social Media Platforms are the main profiteers of rules that enable them to exclude users from their service as it reliefs them from extra effort to assess notifications and potentially excludes users wrongfully from their right to notify.**

Such entitlements on the other hand could be prone to error to exclude users from notification procedures, who repeatedly report abusive content which was not found to be illegal by the platform. Several aspects can motivate users to report illegal content: different understandings of “decency” or “discrimination” due to different cultural backgrounds and a lack of legal knowledge. None of these

[5] Interview with M. Vestager, Spiegel Online, 07.02.2021, https://www.spiegel.de/politik/deutschland/margrethe-vestager-soziale-netzwerke-sind-ein-systemisches-risiko-fuer-die-demokratie-a-88df56db-a167-46ea-a1b5-9058a6406d8d

[6] evaluation of the NetzDG (September 20): Eifert, Martin et al., Evaluation des NetzDG im Auftrag des BMJV, p. 53; https://www.bmjv.de/SharedDocs/Downloads/DE/News/PM/090920_Juristisches_Gutachten_Netz.pdf?__blob=publicationFile&v=3

[7] Drucksache der Bundesregierung 19/26398,https://dip21.bundestag.de/dip21/btd/19/267/1926749.pdf

supported by ALFRED LANDECKER FOUNDATION

aspects should lead to exclusion of inconvenient users from notification procedures but on the contrary should encourage online platforms to improve their content decision procedures.

IV. Recommendations

In the following passage we want to outline recommendations which HateAid considers to be necessary to adjust and improve the proposal regarding users of Social Media Platforms and their protection from digital violence.

1. Content moderation: Insufficient protection of those affected by digital violence

The proposal of the Digital Services Act has a very strong focus on the protection of users who are potentially subject to unjustified and disadvantageous content decisions. Consequently, they are provided with different options to appeal (art. 17,18). Unfortunately, these same rights are denied to those who are subject to and/or report illegal content and whose request is denied by the platform. Therefore, the proposal does not improve the situation of users who are subject to digital violence at all. It particularly fails to define an enforceable claim for the removal of illegal content for authorities. While it is only vaguely paraphrased in recital 22 in conjunction with art. 5, the proposal fails to provide any procedural requirements concerning the enforcement of said claim when platforms are unwilling to act on notifications. In this respect the proposal falls short of setting higher standards to protect those affected from digital violence lagging behind even the e-commerce directive from 2000. Considering the above-mentioned tendency of people affected from digital violence, hate speech and discrimination to withdraw from the digital space or keep silent, especially discriminated groups (women, jews, roma etc.) are increasingly excluded from public debate on platforms because they have no rights to appeal.

a) Notice and Action, art. 14, and statement of reasons, art. 15

The only option explicitly granted to users (affected by illegal content) is to notify the platform according to art. 14. The only action required by online platforms is to assess the notification "without undue delay" and notify the individual of their decisions and redress possibilities (art. 14 sec. 5). With regard to these sparse obligations that the proposal would impose on online platforms the following facts are to be underlined:

(1) No obligation to delete illegal content

The proposal does not codify an obligation to delete illegal content. Only recital 22 states that providers should act to remove content upon receiving actual knowledge in order to benefit from the exemption from liability. This is in accordance with art. 5 sec. 1 lit. b that is also speaking of requirements for exemption of liability. Consequently, the proposal refrains from determining an obligation to remove or block illegal content.

This leads to the conclusion that such an obligation **cannot be infringed** and therefore the removal of content **will not be enforceable by authorities** even if contravention is practiced exhaustively. What's more – since compliance cannot be monitored due to a lack of legal obligation - **it won't even be verifiable if illegal content is deleted at all upon notifications**. Only if a user is personally affected, can they seek legal protection through litigation but in all other cases such as incitement or other unconstitutional content, removal cannot be enforced.

(2) Country-of-origin principle, recital 33 and art. 40 sec. 1

In our experience the country-of-origin principle is one of the major obstacles of all measures against digital violence on the internet. Although it allows the free flow of services within the EU it simultaneously creates insurmountable barriers for member states to protect users from digital violence and enforce the law.

We believe that recital 33 of the proposal can be understood as a remarkable exemption from the country-of-origin principle[8] and enables national courts and authorities to make decisions that are not bound to the country-of-origin principle. We see that recital 33 of the proposal may also serve as a foundation for its articles 8 and 9 and welcome the possibilities for cross-border action of authorities. Nevertheless, the recital fundamentally differs and even contradicts the regime of Article 3 of the e-commerce directive (Directive 2000/31/EC) in certain aspects. **To ensure that courts and authorities can prioritize the underlying intention of recital 33 over art. 3 of the e-commerce directive it needs to be codified in an article for the purpose of clarification.**

When it comes to due diligence obligations it is to be expected that this problem will intensify because of art. 40 sec.1 of the proposal. According to our understanding art. 40 sec. 1 will override art. 3 sec.4 of the e-commerce-directive. Only Ireland will be in charge. This leads to the conclusion that contradicting outcomes are predestined. While national courts and authorities can direct specific orders against platforms anytime, Ireland is supposed to be in charge for enforcement of due diligence obligations exclusively. That is why we consider an adjustment of art. 40 sec. 1 and the creation of an exemption similar to art. 3 sec. 4 E-commerce-Directive to be necessary to avoid contradicting outcomes and enable other member states to step in when the country of origin fails.

(3) Notification procedure, art .14

Art. 14 sec. 3 reverses burden of proof under excessively **strict requirements that will hardly ever be met by users**. Especially users who have been affected by digital violence are obviously sensitive when it comes to revealing their personal data to the counter party and fear even more hate speech or oppression. That is why it is highly likely they will refuse to reveal their full name and e-mail address especially when they are not affected by the illegal content themselves. We consider this standard of actual knowledge to be unreasonably strict as it demands excessive and unnecessary information from users and gives advantages to the platforms that already profit from extensive privileges in the art. 3 – art. 7. Name and contact details of the user who reports potentially illegal content is not relevant for the assessment of potentially illegal content and actual knowledge and should therefore not be required by art. 14 sec. 3.

Further we would like to point out the following aspects:

- The proposal does not provide a deadline for the assessment of notifications. From our experience “without undue delay” is an insufficient benchmark, as it lacks a specific definition and therefore will be stretched to the limit by online platforms.
- The design of notification measures according to art. 14 sec. 1 of the proposal is defined only rudimentary as user friendly and easily accessible. From the experience with NetzDG, that chose a similar wording, we know that its definition is stretched to the disadvantage of users by all major Social Media Platforms.

[8] See Holznagel, CR 2021, 123-132

(4) Recommendations for content moderation

Our essential recommendations for content moderation are as followed:

1. Codifying an obligation to delete illegal content in an article.
2. Codifying recital 33 in an article and create an exemption of art. 40 sec. 1 according to Art. 3 sec. 4 E-Commerce-Directive.
3. Lowering the requirements to reverse burden of proof according to art. 14 sec. 3 under exclusion of name, e-mail address and URL.
4. Creating a distinct deadline for assessment of notifications of at least 7 days.
5. Creating an obligation to provide a statement of reasons for all content decisions.
6. Obliging platforms to create a notification procedure that is clearly visible, low- threshold and located close to the content in question.

b) Internal complaint handling system, art. 17, and Out-of-court dispute settlement, art. 18

We fundamentally approve the decision to install an appeal and put back mechanism. The specific design nonetheless is unbalanced and raises questions and concerns about potential harm to freedom of expression.

(1) Access to redress mechanism for victims of digital violence

If the platform decides **not to act subsequently to a notification** of illegal content, the **affected users won't have access** to the internal complaint handling system, art. 17, and out-of-court dispute settlement, art. 18. Further, no statement of reasons by the platform according to art. 15 is required. This deprives the users of the opportunity to soundly assess the prospects of judicial review or seek low-threshold and inexpensive reassessment.

Put into practice, this would mean that the situation of users whose notifications are rejected by the platform does not change under the rule of the proposal. If the platforms decide not to take action upon (the) notification – no matter if manifestly unjustified or not – the users only option is to hire a lawyer and take legal action on their own costs. This is rarely going to happen because of the financial risk, so platforms don't fear civil court cases for wrongful content decisions and act accordingly as the threshold for users is too high.

This is not comprehensible and causes an **unjustified unequal treatment** of users reporting illegal content in comparison to those whose content is reported to the platform. It becomes even more clear when looking at art. 17 sec. 3 which enables online platforms to **reverse decisions according to art. 17 sec. 1 immediately** when a complaint contains **sufficient grounds to consider the addressed content as not illegal**. The article serves as a strong legal basis for service recipients who had their own content potentially wrongfully managed by the online platforms with particularly low requirements. It becomes hereby obvious that the strong fundamental ground for removal is not equipped with an equally strong mechanism in the proposal to ensure its enforcement to the benefit of users who are affected by digital violence. Certainly, this does **not offer effective protection against the spread of illegal content on online platforms**. We consider this to be a major miscalculation of the proposal.

(2) Affected users are left defenceless

We are convinced that this regulation serves as a strong incentive for online platforms to reject any notification that does not concern manifestly illegal content. The concern is even more reasonable with a look at the fact that only a fraction of the notifications has led to the removal of content under the rule of NetzDG with strict deletion periods and even fines on systematic infringements. Due to legal interpretation which is hardly ever 100 % clear this will relate to the majority of content reported. **The assumption that any user would have adequate options to take legal action against a wrongful content decision is highly theoretical due to the following reasons:**

- Legal action can only be taken when the user is personally affected and subject to the illegal content. Other cases such as incitement, racist discrimination, death threats against politicians or anti-constitutional content cannot be pursued by users.

- Even if the user is personally affected it is very unlikely that they actually take legal action due to cost risk and legal uncertainties.

- A complaint to the Digital Services Coordinator according to art. 43 cannot be successful as it is no infringement of the proposal due to the lack of an obligation to delete illegal content.

In the end it must be feared that tremendous amounts of illegal content will not be deleted, in fact will not even be noticed. The consequences for discriminated groups and for society as a whole have been explicated above.

(3) Out of court settlement vs. rule of law

We understand and appreciate the presumably underlying thought of art. 18 to create an alternative dispute resolution option which is likely to accelerate the decision-making process. Nonetheless its design is questionable, and we doubt that it will be able to live up to the high standards to protect freedom of expression which the EU rightfully intends to preserve.

We seriously doubt the “external body” can be an appropriate solution that makes legal proceedings dispensable. It is important to stress that under rule of law the final decision must be reserved for courts and that courts need to be accessible for citizens without disproportionate barriers. Our focus should not be to find workarounds that create an intermediate solution. To uphold the rule of law in order to guarantee freedom of expression it is necessary to enable users to seek judicial review at reasonable and affordable conditions. Therefore, it is necessary to improve access to justice for all users in the EU. In our opinion this can be realised **by installing mandatory summary proceedings in all member states for such content decisions which are accessible for victims and uploaders alike**.

Our concerns with regard to art. 18 to the current art. 18 of the proposal focus on the following aspects:

- **Transparency about the applicable law.** According to the country-of-origin principle it is obviously the law of the country where the platform is located. In case of Social Media Platforms that would be Ireland. It could only be otherwise if the external body is an authority according to art. 8, but that remains to be subject of interpretation. If this is not the case it needs to be questioned how the law of e.g. Ireland should be applied by an “external body” in other European countries or elsewhere since users have the choice to choose their body of complaint.

- Many questions occur concerning the legal **status of the "external body".** It remains unclear how it should be constitutionalized, how its expertise is assessed and especially how the **independence** would be guaranteed. Particularly the independence is vital as it is especially important to avoid financial or personnel involvement of online platforms with the external body. Further clarification and procedural structure are required.

- **Uncertainties remain**, even if the decision could still be subject to judicial review The decisions of the "external body" are supposed to be binding although it is not defined as a court and if the affected user wants to challenge the final decision it needs to be clarified **who is to be addressed by legal action** – the platform or the legal body itself that made the binding decision? Moreover, it is **unclear who is to cover the costs if the decision of the external body is overruled in court as this is only regulated for the costs of the out of court settlement itself.**

- **Moreover, the question of assumption of costs in general is at our concern.** It is debatable why the affected user should bear any costs of this out of court settlement in the first place. The claim that they are supposed to be "reasonable" is not sufficient in our opinion as this is subject to a **wide range of interpretation** and can also be dependent on the member state concerned. From our experience users **will not use this tool if the costs are not predictable and not clearly defined**. Therefore, transparency about the costs is absolutely necessary. Otherwise, the out of court settlement will be only of little relevance. The costs also need to be significantly lower than litigation costs. Furthermore, it is objectionable that **no instruction about the costs in advance is required** and therefore they cannot be assessed properly by users.

(4) Recommendations for internal complaint handling & out of court settlement

1. Make internal complaint handling and out of court settlement accessible to all users.
2. Recreate the out of court settlement mechanisms and acknowledge that the proposed "external body" is unsuitable to meet the desired standard to ensure freedom of expression and rule of law. We recommend finding a regulation instead **that improves access to justice for all users towards online platforms** no matter if their content is removed or their notification rejected. We recommend regulating the mandatory regulation of **summary proceedings** in all member states for content decisions.

c) Trusted Flaggers, art. 19

We fundamentally welcome the approach to create a legal framework for trusted flagger institutions. Our experience with current trusted flaggers programs that some online platforms initiated voluntarily is ambivalent: On the one hand it gives Trusted Flagger organizations the opportunity to offer help to affected users and clients who unsuccessfully reported illegal content and appeal content decisions through a "direct line" to the platforms which have not been responsive to them. This way more complex matters can be enlightened easily in some cases. On the other hand, the trusted flagger status does not guarantee priority and fast assessment. In some cases organizations do not even receive an answer.

However, it needs to be pointed out that a legal framework of trusted flaggers **can only be a supplement to** a broader strategy to tackle the spread of illegal content. The legislator should not rely on it as a key strategy. Since the trusted flagger status is not compensated, only NGOs can have an incentive to apply for it. These NGOs are by nature either publicly or donor funded, oftentimes even relying on voluntary commitment of civil society and are overall, chronically underfunded. To rely on NGOs to proactively search for illegal content instead of the online platforms themselves or law enforcement implies shifting the burden onto civil society and taxpayers with extra effort and costs while sparing the online platforms who are profiting economically from the traffic on their platforms but are not obliged to also economically cover the negative side effects that are caused on these same social media spaces.

d) Measures against misuse, art. 20

It is in general useful to install a mechanism to aim to preclude certain accounts or users who have already attracted attention due to infringements. Nonetheless we explicitly warn against putting users under general suspicion if the consequence is to relieve platforms from content moderation efforts.

(1) Suspension of users, art. 20 section 1

With regard to Social Media Platforms this measure can be part of a comprehensive package to prevent the spread of illegal content much more efficiently**. Though from our experience the effect of this measurement should not be overestimated. The majority of profiles that are spreading manifestly illegal content are fake profiles**. It quickly becomes obvious that it is either one of several profiles with only little personal information that one user maintains simultaneously or that it is even just a random fake profile without any personal interest. Consequently, the effect of blocking such profiles cannot be sustainable and the users in question change their profiles or simply create a new one. This is extremely easy for them as they can employ random data since there is no obligation to verify their identity on online platforms.

(2) Suspension from notification procedure, art. 20 section 2

We consider the suggestion to block users from the notification procedure to be highly problematic. This is partly due to the broad decision-making scope of online platforms to assess if content is illegal and partly because this assessment cannot be expected from users with no legal educational background. It is highly likely that users – especially from marginalised groups – repeatedly report content that they experience to be discriminatory but which in fact is not illegal or subject to a violation of terms and conditions. Moreover, under the rule of the country-of-origin principle it can be assumed that the legal standard of the country where the headquarter is located needs to be applied, which cannot be expected to be taken into account by users at all.

We understand the underlying intention to prevent misuse of notification procedures. We consider this objective to be misguided as it again gives advantages and relief to platforms at the expense of their users.

Instead of raising the demands and requirements on the side of overall disadvantaged users the proposal should compel the platforms to assess all notification of sufficient duty of care and improve their content moderation procedures.

2. Point of Contact is insufficient, art. 10

The proposal contains mandatory establishment of several institutions aiming to facilitate contacting intermediary services. This is a step in the right direction when it comes to ensuring compliance and transparency. Nevertheless, said institutions (single point of contact, art. 10 (1), legal representative, art. 11, compliance officers, art. 32, Digital Service Coordinator, art. 38) exclusively regulate communications of authorities with online platforms. Individuals seeking to contact intermediary services do not gain any options that facilitate access to online platforms.

To introduce access to justice for individuals towards online platforms it is essential to guarantee easy, transparent and comprehensible access to platforms. In this regard the proposal falls short to acknowledge the practical difficulties individuals face when seeking to contact intermediary services in general. Not being able to find reliable and helpful information on where to address legal claims will most certainly lead to frustration and prevent individuals from enforcing their rights. Other regulations simply lack practical relevance for individuals as they are forced to search the website and hidden imprints for a postal address and then face the practical question of how to ensure a legally certain delivery abroad.

Therefore, we insistently demand the following: All providers of online platforms must be obliged to **appoint an authorised recipient for all kinds of legal proceedings concerning content moderation or related measures (art. 17, 18, 20) in every Member State in at least one official language of that state which dedicatedly serves as a contact point for legally binding delivery in order to enforce user rights.** This can be reached by providing information on local offices or by appointing national law firms or similar representatives for reception of legal documents. The latter option allows minimum effort for platforms while maximizing benefits of individuals. Consequently, it is reasonable to extend this measure to all intermediary services. Exceptions can be made in case intermediary services qualify as micro and small enterprises within the meaning of the Annex to Recommendation 2003/361/EC.

3. Oversight under the rule of the country-of-origin principle

The proposal creates a complex and high-volume system of oversight and compliance while – apart from very little exemptions – keeping up the country-of-origin principle. With regards to huge Social Media Platforms where the majority of our public debate takes place, that principle will fall under the jurisdiction of Ireland. In practice this means that Ireland will be responsible for oversight of Facebook, Youtube, Instagram, Twitter and TikTok. All other member states will be excluded from oversight over all major Social Media Platforms. Even when disregarding the fact that Irish authorities did not take any action against Social Media Platforms because of their handling of illegal content this situation is not reasonable. It does not only enable online platforms to cherry-pick their regulator, but also raises valid concerns that this task will simply overstrain Irish authorities.

4. Find a way to restrain harmful services registered abroad

Certain online platforms do not only demonstrate lack of law enforcement but even create a safe haven for all sorts of illegal content. Those platforms serve anti-constitutional movements as a place to organise themselves, to recruit new members and commit serious crimes. The most famous example in Germany is the messenger service Telegram which serves especially deniers of covid 19 and gave rise to various conspiracy theory movements during the pandemic. The channels on Telegram

do no longer only act as a messaging service but have become social networks long ago[9]. Many channels have 100.000 – 200.000 followers who actively comment and share content. One of the most famous accounts on Telegram in Germany is conducted by an antisemitic conspiracy theorist with more than 114.000 followers who is posting antisemitic, inciting, holocaust denying, demagogic and other clearly illegal content every minute[10]. Still there is nothing German law enforcement authorities can do about it. Telegram itself is registered in Dubai, has no servers or bank accounts in the EU and does not respond to law enforcement agencies or other authorities. It does offer an e-mail address for complaints but does not react on those. As a result, it is impossible to enforce any law towards this kind of platforms and even if the German regulator tried to enforce the NetzDG towards Telegram the prospects of success are extremely low.

In our experience this tactic is employed by many others and very small platforms such as extremist blogs or other websites that are completely aware of the legal loopholes that they profit from.

In order to ensure public safety from these services we need reliable options to restrain these services or even block them in the EU which is not yet granted by the legal framework proposed by the DSA.

The proposal so far does not require any credible solutions that can guarantee law enforcement towards such platforms. Neither art. 40 sec. 2, 3 nor art. 41 sec. 3 offer any solutions for their enforcement if a platform is registered abroad and does not react to orders. Especially it does not allow the involvement of third parties such as access providers, app stores or payment providers which we consider to be necessary to get hold of such services that cannot be held liable otherwise.

The Landecker Digital Justice Movement is a joint project of HateAid and the Alfred Landecker Foundation.

[9] https://techcrunch.com/2020/12/23/telegram-to-launch-an-ad-platform-as-it-approaches-500-million-users/?guccounter=1&guce_referrer=aHR0cHM6Ly93d3cuZ29vZ2xlLmNvbS8&guce_referrer_sig=AQAAAHkQbAWZz6YUGvOFKDotDHgpf6ZMQnLllQ2o6OKV7gPM_182T7M101T1LcVNf1k7fdCKsG7tbv6Opyn0xl_H6ZnvWXP0llUImh7DVprw4pLAY9SvzYSXMF0cWfUliO6R9JqcbfBNGP2TWhqtz5ZC-sLS8jOt9WJ5W7ZWIvGtKBlj

[10] https://t.me/s/ATTILAHILDMANN

HateAid gGmbH was initiated in 2018. We are the first organization in Germany to offer protection from digital violence to those affected and at the same time to support effective sanctioning of the perpetrators. Moreover, we create social awareness of the destructive effects of digital hatred on our democracy. HateAid's aim is to relieve the burden of the victims of attacks, enforce their rights, deter the perpetrators, and overall strengthen our democracy and society.

ALFRED LANDECKER
FOUNDATION

The Alfred Landecker Foundation promotes and accelerates the development of an open, democratic and discrimination-free society – in an innovative, fearless and disruptive way.
As an advocate for democracy in the digital age, the foundation puts technological progress and comprehensive expertise at the service of open societies, the fight against anti-Semitism and racism and a contemporary culture of remembrance. The Foundation sets up networks, creates spaces and knowledge by supporting, promoting, networking and professionalising interdisciplinary projects. The Alfred Landecker Foundation was founded in 2019 by the Reimann family and is based in Berlin. Previously, the business historian Professor Paul Erker, who was commissioned by them, had determined that the manager of the family-owned company Joh. A. Benckiser GmbH was a supporter of the Nazi regime during the National Socialist era. To face up to their responsibility arising from this part of their family history, the company heirs are concerned to support victims of the Holocaust who are still alive today and, as a lesson of history for the present days, to promote democracy and human rights and to contribute to the preservation and strengthening of a pluralistic society.