ERGA

European Regulators Group for Audiovisual
Media Services

ERGA Subgroup 1
2024
Consistent implementation and enforcement of the European framework for audiovisual media services

Deliverable

Report on measures and online safety features for the Protection of Minors

A follow-up to last year's assessment of the measures adopted by audiovisual media service providers and video-sharing platforms, including age verification tools, criteria for content flagging and parental control.

ERGA

European Regulators Group for Audiovisual Media Services

Table of Contents

I. Introduction

National Regulatory Authorities (NRAs), members of the European Regulators Group for Audiovisual Media Services (ERGA), are essential in ensuring the effective implementation of the Audiovisual Media Services Directive (AVMSD).

ERGA's work program for 2024 continues to focus on monitoring and supporting the implementation of the AVMSD in the Member States. SG1s' remit this year was to consolidate ERGA´s efforts towards an enhanced and more effective enforcement of the European legal framework for audiovisual media services and to gather relevant evidence on the implementation and potential future development of this framework.

The work of SG1 for 2024 aimed to continue the discussions initiated last year on the coordination of approaches to the protection of minors on video-sharing platforms (VSPs) and on-demand services, as required by Articles 6a and 28b(3) of the AVMSD, the development of the assessment of the measures adopted, including age verification tools, criteria for content flagging and evaluation of concrete practical applications and effectiveness, to develop recommendations for ERGA and its members.

This report focuses on a follow-up assessment of the measures adopted by audiovisual media service providers and video-sharing platforms, including age verification tools, criteria for content flagging, and parental control. Moreover it aims to ascertain NRA's competence and knowledge regarding those measures and online safety features by analysing concrete practical applications and assess the effectiveness of the measures in place to develop practical recommendations for ERGA and its members.

This report is mainly based on desk research and the answers of ERGA members and observers to a questionnaire, sent out on April 19th, 2024. 29 NRAs responded to the survey (Albania, Austria, Belgium - CSA, Belgium - VRM, Bulgaria, Croatia, Cyprus, Czechia, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italia, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Norway, Poland, Portugal, Slovak Republic, Slovenia, Spain and Sweden), representing 25 EU Members, 2 EFTA countries and 1 ERGA observer country.

To improve the readability of this report, the conclusions and recommendations are presented at the end of each chapter.

Please note that the recommendations set in this report are intended as inputs for further reflection, consideration, and debate, aimed at enhancing the understanding of the protection of minors subject, in line with Subgroup 1's goal of supporting the effective implementation of AVMSD.

In this context, the recommendations outlined do not represent a common agreement or a policy course, nor do they address ERGA as a whole or any specific Member of this forum.

Instead, they are suggestions for each Member to evaluate independently and determine their applicability based on their own self-assessment.

II. Historical background

The amendments to the Audiovisual Media Services Directive, introduced by Directive (EU) 2018/1808, have significantly strengthened the protection of minors in the audiovisual environment and adapted the rules to the digital context and technological developments.

The scope of the AVMSD now includes not only traditional television services and on-demand audiovisual services, but also video-sharing platforms and social networks, which offer a significant amount of audiovisual content. This reflects the need to regulate the new means by which minors consume content.

This need led to the provision of reinforced measures for the protection of minors, under which video-sharing platforms now have explicit obligations to protect minors from content that is harmful to their physical, mental, or moral development. The measures include parental control tools and age verification mechanisms; content rating systems; and reporting and removal mechanisms.

Tighter restrictions on advertising to minors have also been introduced, including a ban on advertising that directly encourages children to buy products or services by exploiting their inexperience or credulity.

Cybersecurity and protection from harmful content are other concerns in the current text of the Directive, which has been amended to reinforce the need to protect minors from content that promotes violence, hatred, or dangerous conduct.

On the other hand, the need for platforms to share responsibility for the protection of minors has been made clear, with the Directive encouraging greater cooperation between NRAs and platforms to ensure the effective application of the rules and compliance with child protection standards.

III. Analysis of the Legal Framework

The assessment of the consistent implementation and enforcement of the European framework for audiovisual media services by the Member States – with a view to understanding and promoting best practices for the protection of minors, not only in the online environment but also in on-demand services, as well as the assessment of the adequacy of the measures adopted, from age verification mechanisms to content flagging measures and parental control – must take into account the legal analysis of the AVMSD.

As stated above, the amendments to the AVMS Directive significantly adapt audiovisual regulation to digital reality and provide a more robust framework for the protection of minors. The inclusion of video-sharing platforms and social networks, the imposition of specific measures for the protection of minors, and the emphasis on shared responsibility between platforms and regulators are some of the main advances of this legislative revision.

In this context, it is essential to differentiate the legal protection contained in the AVMSD from other national legal sources.

Articles 6a and 28b reflect an effort to adapt European legislation to new realities of digital media consumption. They extend protection against harmful content to the online environment, recognising the impact of video-sharing platforms on society, particularly on minors.

Both articles reaffirm the European Union's commitment to protecting fundamental rights, including human dignity and the safety of minors.

These provisions are part of a broader EU approach to ensuring that media legislation keeps pace with technological changes, protecting users, especially minors, and preventing the dissemination of harmful and hateful content in the digital environment.

Other European legal instruments should also be highlighted, such as the Digital Services Act (DSA), adopted in 2022, which introduces a comprehensive legal framework, with the aim of horizontal harmonization for the regulation of online platforms with a strong focus on the protection of minors. These include the prohibition of targeted advertising based on minors' profiles, which protects them from invasive commercial practices; the duty of care of online platforms, which requires them to mitigate specific risks, including the exposure of minors to harmful content and risks that could affect their well-being. Platforms must implement robust content moderation policies and systems to flag and remove dangerous content. Transparency, parental control and age verification mechanisms are also among the demands of the DSA, requiring platforms to provide users, especially parents, with clear and accessible information about parental control mechanisms and the handling of minors' data. The DSA requires to make the terms and conditions of platforms clear and easy to understand, also for minors. Finally, the obligation to carry out regular impact assessments of the risks posed by their services to minors and to adapt their policies where necessary to protect this vulnerable group.

Under the Digital Services Act (DSA), the European Commission established the Age-verification Task Force to address the challenge of protecting minors from harmful online content through effective age-verification mechanisms. The Task Force works to harmonize approaches to age assurance across the EU, ensuring compliance with legal frameworks such as the AVMSD, the GDPR, and the Better Internet for Kids (BIK+) strategy. It focuses on researching, developing, and guiding the implementation of age verification technologies across the EU.

The European Media Freedom Act (EMFA) is also one of the instruments that indirectly plays a relevant role in the protection of minors. Although this legislation aims to enhance the functioning of the internal market, thereby safeguarding media pluralism and independence and does not have explicit provisions specifically tailored to the protection of minors, it indirectly supports a safer media environment by upholding independent journalism and transparent media practices, which are crucial to ensuring trustworthy content for all audiences, including minors. In particular, EMFA sets out new cooperation and enforcement mechanisms between media regulators with the aim to foster the consistent and effective implementation of the AVMSD, including with respect to the VSPs obligations related to the protection of minors.

Also worth mentioning is the European Safer Internet Strategy[1], which aims to create a safe and stimulating digital environment for children and young people by promoting digital literacy among children, parents, and educators, enabling them to use the internet safely and responsibly, developing and making available tools for parents to supervise their children's online activities, such as content filters and parental controls, and helplines and reporting mechanisms that allow anyone to report illegal or harmful content found online.

Last but not least, other initiatives and partnerships developed by the EU are worth mentioning, such as the cooperation with the Council of Europe and UNICEF to harmonise child protection standards and exchange best practices, and the development of voluntary codes of conduct in which companies commit to additional child protection measures.

The Louvain-la-Neuve Declaration[2] and the "Council conclusions on supporting influencers as online content creators"[3] are two other instruments worth highlighting.

The first - the Louvain-la-Neuve Declaration - adopted in April 2024, is a non-binding declaration that focuses on promoting a safer, more responsible and trustworthy online environment across Europe, with a strong emphasis on the protection of minors. In it, Member States committed to "develop and build upon existing harmonised technical solutions and standards across EU Member States, taking into account national initiatives taken by Member States, to provide interoperable, universal and user-friendly parental control mechanisms as well as identity and age verification through privacy-preserving technologies, including by building on the functionalities available in the forthcoming European digital identity wallets, without excluding the use of other appropriate age verification systems". In this context, it is

[1] https://digital-strategy.ec.europa.eu/en/policies/strategy-better-internet-kids

[2] https://bosa.belgium.be/sites/default/files/content/documents/LLN%20Declaration%20-%20Informal%20Telecom%20Council%20-%20v.12.04.2024.pdf

[3] https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:C_202403807

worth mentioning the *eIDAS* regulation (Electronic Identification, Authentication and Trust Services, Regulation (EU) No 910/2014[4]), which aims to create a legal framework for identifying, authenticating and ensuring secure electronic transactions between citizens, companies and governments in Member States. The *eIDAS* regulation serves as a foundational framework for digital identity solutions across the EU, providing legally recognized and secure identification standards that support interoperable and privacy-preserving age verification mechanisms.

The second - Council Conclusions on supporting influencers as online content creators, adopted in May 2014 - makes some recommendations on how to better protect minors online, in particular by addressing the increasing role that influencers and kidfluencers (underage influencers) play in shaping public opinion and online content consumption across the EU. This includes both positive and potentially harmful effects of influencers in shaping public opinion, especially among minors, and highlights the need for better media literacy, ethical guidelines and protection of minors, especially as influencers increasingly influence online content consumption.

IV. Supervision, Monitoring and Enforcement

1. State of Play of the European Audiovisual Scene – Video-on-Demand Services (VODs) and Video-Sharing Platforms (VSPs)

One of the first aspects this year's survey sought to identify was the changing landscape of VOD and VSPs in each Member State.

Looking at the **Database on audiovisual services and their jurisdiction in Europe –** MAVISE[5], it is safe to say that the VOD market has grown significantly in recent years, driven by technological advances, increased internet penetration, and changing consumer preferences. The latest data (28 August 2024) shows that there are currently 3,238 video-on-demand services available in the EEA, of which 2,684 VODs are registered under the jurisdiction of the AVMSD.

In 2022, NRAs were questioned[6] about whether they had identified VSPs under their jurisdiction, providing us with a snapshot from which to map their evolution. Since it is a new and dynamically changing landscape - with recent confluent/converging regulations regarding

[4] https://digital-strategy.ec.europa.eu/en/policies/eidas-regulation

[5] https://mavise.obs.coe.int/

[6] The implementation(s) of Article 28b AVMSD: national transposition approaches and measures by video-sharing platforms.

VSPs, such as the DSA - it seems necessary to revisit this aspect in particular before delving further.

Considering the results of the survey submitted to ERGA's members during April, we retain the following key takeaways:

- 15 NRAs have already identified VSPs under their jurisdiction, an increase from 12 to 15 countries since October 2022[7];
- 7 NRAs state that they have identified new video-sharing platforms under their jurisdiction during the last year;
- The total number of VSPs identified under ERGA members remit has risen from 49[8] to 88[9].

Regarding the latter, the increase in VSPs was indeed significant:

a) In Ireland, Coimisiún na Meán (CNAM) has identified 10 new services as VSPs by December 2023.
b) The Bavarian Regulatory Authority for New Media (BLM) in Germany has claimed jurisdiction over Twitch.
c) In Hungary, the National Media and Infocommunications Authority (NMHH) has added 5 VSPs to the list of VSPs under its jurisdiction, bringing its total to 10.
d) Similarly, in Luxembourg, the Autorité Luxembourgeoise indépendante de l'audiovisuel (ALIA), has indicated that the number of VSPs has risen from 11 to 16.
e) Now covered by the Dutch Media Act, the Commissariat voor de Media (CvdM), in the Netherlands, supervises Snapchat.
f) In Spain, the Spanish National Markets and Competition Commission (CNMC) has identified 4 owners of pornographic VSPs, although the Ministry of Digital Transformations and Civil Service is still working on the details of the registered VSPs.
g) Finally, the Swedish Agency for the Media (Memy) has registered two VSPs: Solidsport and VOYDTV.

2. Initiatives on Self/Co-Regulation

One of the critical issues addressed by the AVMSD is co- and self-regulation, which refers to the shared responsibility between Member States and content platforms on the one hand and the ability of the media sector to set and apply its standards and guidelines without the need for intervention by external bodies or government on the other hand. This approach aims to

[7] According to aforementioned report, pp.11.

[8] According to the annex list in aforementioned report, pp.23.

[9] Under the jurisdiction of the AVMSD, according to MAVISE.

promote freedom of the press, diversity of content, professional ethics, and social responsibility.

The effectiveness of self- and co-regulation can vary, depending on companies' commitment and the acceptance by the public and stakeholders, but it is considered to promote a constructive dialogue on standards and expectations concerning audiovisual content. This approach is seen as a way of balancing freedom of expression with the need to maintain standards and protect the rights of individuals in society.

2.1. Systems describing the potentially harmful nature of the content

Article 6a(3) of the AVMSD states that “Member States shall ensure that media service providers provide sufficient information to viewers about content which may impair minors' physical, mental or moral development. For this purpose, media service providers shall use a system describing the potentially harmful nature of the content of an audiovisual media service”, “[f]or the implementation of this paragraph, Member States shall encourage the use of co-regulation as provided for in Article 4a(1)”, ensuring that those codes *are broadly accepted by the main stakeholders, clearly and unambiguously set out their objectives*, ensure *regular, transparent and independent monitoring and evaluation of the achievement of the objectives*, and *provide for effective enforcement including effective and proportionate sanctions*".

This provision aims to create a safer media environment for minors by ensuring that content likely to impair their development is properly identified and communicated to users through a standardised system. It is, therefore, crucial to enable users, especially parents and guardians, to make informed decisions about what is appropriate for minors to watch.

In light of this, it was sought to ascertain if the NRAs were aware of any systems used by media service providers to describe the potentially harmful nature of the content.

Of the 29 NRAs that responded to the survey, 17 replied that, to their knowledge, media service providers currently have such a system in place.

This system could take various forms, such as content ratings, warning labels, age-appropriate symbols, detailed content descriptors, or time-based restrictions.

One possible way of further developing national systems for describing the harmful nature of the content may be to adopt a common system to be used by all media service providers. In fact, 27 of the surveyed NRAs considered it useful to adopt such a common system, recognizing the need to create a level playing field for all media service providers to the extent possible (i.e., depending on the nature of the provider), while ensuring that all potentially harmful content is flagged correctly, thus protecting minors.

In addition, it may improve compliance and enforcement by making the legislation to understand for users and providers and reducing administrative burdens/costs, given that NRAs have to set up mechanisms to monitor and enforce this provision.

Only 1 respondent (ALIA) felt that adopting a standard system would not be helpful, perhaps anticipating potential challenges to its implementation, such as the diversity of content.

2.2. Involvement of VSPs in co-/self-regulatory initiatives

Article 28b also encourages the use of co-regulation and self-regulation in its implementation. To assess the involvement of VSPs in co-/self-regulation initiatives in matters related to Article 28b of the AVMSD, NRAs were asked whether there had been any changes compared to the previous year's data. Only 3 NRAs (excluding the NRAs that did not identify VSPs under their jurisdiction) answered "Yes" for the VSPs under their jurisdiction.

Among them, the Netherlands, CvdM noted that Snapchat is linked to the Stichting Reclame Code, which oversees the country's advertising self-regulation system. In Portugal, ERC reported preliminary meetings with Google, Meta, and TikTok are taking place. Meanwhile, in Spain, CNMC said it is currently promoting the development of a co-regulatory agreement to strengthen the protection of minors through the age classification of audiovisual content. This code will apply to all providers, including VSPs and relevant users such as influencers (Articles 98, 94, 89 of the LGCA - Audiovisual Law).

2.3. Promotion of co-regulation and self-regulation through codes of conduct

The Directive also states that Member States should encourage the use of co-regulation and the promotion of self-regulation through codes of conduct adopted at national level. Hence, the survey tried to clarify whether such encouragement had taken place. Around half of the respondents - 14 - stated that no measures had been taken, while 13 responded positively. Regarding the latter, ARCOM (France) stressed that at the end of 2022, seven online platforms ratified a standard *joint Charter to promote the information and protection of the users concerning the distribution of the image of minors on online platforms* under the auspices of ARCOM[10].

When questioned about which instruments of co and self-regulation have been adopted regarding the protection of minors, the mechanism highlighted on the survey results is the organization of roundtables and workshops (11), followed by "Convene working groups" (8) and "Enforcement Mechanisms" (7). The participation of existing regulatory bodies in co-regulation initiatives accounted for 5 of the chosen adopted instruments.

Other instruments, such as implementing guidelines and establishing reporting mechanisms (both with 4), funding and resources, and publishing best practices, were more scattered and

[10] https://www.arcom.fr/sites/default/files/2023-06/english%20version-presentation%20kit%20arcom.pdf

less relevant in proportion (1). The “Other” category received a significant number of answers, accounting for up to 10 of the answers regarding adopted instruments. Below are some of the details made available by respondents regarding these other instruments.

In the Slovak Republic, RpMS highlights the newly established Commission for the Protection of Minors[11], which functions as a unique co-regulatory body to develop and monitor a uniform and accepted labelling system for the age-appropriateness of audiovisual content. The Commission's activities are funded and administratively organised by the regulatory authority, although its mandate is carried out independently by its members, including representatives of various governmental and professional organisations, thus ensuring a comprehensive approach to child protection in the media.

Similarly, Greece states that NCRTV has been mandated by Law 4779/2021 (transposing the AVMSD) to adopt a code of conduct for all programmes. For its part, RRTV (Czechia) mentions its cooperation in drawing up the code.

In Belgium (French Speaking Community), CSA’s advisory committee is a co-regulatory body that includes two representatives of video-sharing platforms. However, in the Walloon-Brussels Federation, there are no VSPs under the jurisdiction of CSA, so there are no representatives of these services on the advisory committee. Meanwhile, the VRM, the Flemish media regulator, has published a code of ethics for Belgian bloggers, including vloggers[12].

In the Netherlands, CvdM requires VSPs to have a code of conduct that includes, where appropriate, the measures listed in Article 28b(3) of the AVMSD. This code of conduct is supervised by co-regulation and established by joint meetings with the Commissariat".

In Italy, although no VSPs have been set up or are considered to have been set up, a regulation has been adopted by AGCOM[13] to strengthen the protection of minors online, thanks to which it will be possible to report and remove videos considered to be harmful to minors within 5 working days. The regulation has been provisionally notified to the European Commission and will enter into force in 2024[14]. The importance of this regulation lies in the fact that it is possible to activate an intervention by AGCOM not only for VSPs established in a member country but also for VSPs established in another country of the European Union. It is important

[11] https://rpms.sk/rokovania-komisie-na-ochranu-maloletych

[12] Available at http://www.belgianinfluencers.be/nl/ethische-code/

[13] https://www.agcom.it/competenze/piattaforme-online/contenuti-nocivi-video-sharing-platform

[14] Following an intervention by the Italian Communications Authority (AGCOM), the video-sharing platform TikTok, based in Ireland, has proceeded to remove several videos from its platform, all related to the so-called "French scar". The videos identified involve challenges (or so-called challenges) related to the phenomenon known as the "French scar" where very young participants intentionally bruise themselves and create red marks by squeezing the skin of their cheeks around the cheekbones. The purpose behind this practice is to pretend to have been involved in a physical altercation and to appear tough, demonstrating one's courage.

to highlight the important co-regulatory role played by the VSP platforms in this context, both in the drafting phase of the text and in the implementation phase.

Also, in 2024, following a long public consultation, AGCOM adopted a regulation regulating the activity of influencers. In particular, it is therefore expected that influencers will also have to comply with the Italian provisions (articles 37 and 38 of the TUSMA) established for the protection of minors in the implementation of Directive 2018/1808[15]. The great importance of this regulation is that, as in the case of the VSP regulation, it can be applied independently of the establishment of the online platform, as it concerns the person exercising his/her influencer activity on a VSP platform. In this context, AGCOM has also highlighted the importance of co-regulation, which makes it possible to identify influencers quickly and effectively.

Finally, AKOS (Slovenia) indicates that there are no formal self-regulatory or co-regulatory instruments for the protection of minors. However, AKOS cooperates with audiovisual media service providers to facilitate classification practices and unify perceptions of harmful content. In 2018, AKOS established a committee for encoders, mainly from major TV broadcasters and VOD services, which meets informally every two months. These meetings allow encoders to openly discuss and resolve difficult classification cases, and promote a more uniform understanding of content classification.

- ***Conclusions:***

The AVMSD promotes co-regulation and self-regulation as essential approaches to balancing freedom of expression and the protection of minors. These approaches allow Member States and content platforms to share responsibilities for setting and implementing standards without the need for direct government intervention.

The effectiveness of co-regulation and self-regulation depends on the commitment of companies and the acceptance by the public and stakeholders. Although they are seen as a way of promoting constructive dialogue on standards for audiovisual content, their successful implementation can vary considerably.

Most NRAs recognise the importance of systems to describe the potentially harmful nature of audiovisual content. These systems, which may include age ratings, age-appropriate symbols, warnings or time limits, are crucial to creating a safer media environment for minors. There is strong support among NRAs for the adoption of a common content description system, with 27 of respondents considering this measure useful. A common system would help to create a

[15] https://www.agcom.it/comunicazione/comunicati-stampa/comunicato-stampa-10-gennaio-2024

level playing field for all media service providers, make the rules easier to understand for users and reduce the administrative burden on regulators.

The most commonly used co-regulatory and self-regulatory instruments include the organisation of round tables, workshops and working groups. These methods facilitate dialogue between stakeholders and help to develop codes of conduct and enforcement mechanisms. However, there is a diversity of approaches, and the implementation of such instruments varies from country to country. Successful examples include the newly established Commission for the Protection of Minors in the Slovak Republic and efforts to develop codes of conduct in Greece and Czechia.

In Belgium, VRM, the media regulator in Flanders, has published a code of ethics for bloggers and vloggers, and in the Netherlands, the implementation of the AVMSD included a requirement for VSPs to have a specific code of conduct.

In Italy, AGCOM, the Italian communications authority, has published guidelines for influencers and a regulation for VSPs. AGCOM has highlighted the importance of activating a permanent form of dialogue between all NRAs, specifically regarding video-sharing platforms. In fact, the Italian Communications Authority has observed that, unlike linear and non-linear media services, where intervention for the protection of minors is undeniably effective, in the context of VSPs, the timeliness of intervention depends on the NRA where they are established. As most of them are established in a few countries, it is often difficult to ensure timely intervention. For this reason, AGCOM proposes, on the basis of its experience, the application of the tools used in the implementation of the Electronic Commerce Directive 2000/31. To this it adds a greater involvement of VSP providers in the perspective of a greater implementation of co-regulatory tools.

Despite widespread support for co- and self-regulation, there are challenges and limitations, such as the diversity of content and the complexity of implementing a common system. The lack of formal instruments in some countries, such as Slovenia, also indicates the need for further development and standardisation of child protection practices.

➢ *Points to Consider:*

Based on the conclusions presented, the following points for reflection can be highlighted:

1. Increase the participation of service providers:

- *Promoting active dialogue: It is essential to increase the involvement of service providers in the process of co- and self-regulation. To this end, more meetings, such as roundtables and workshops, should be encouraged to promote cooperation and the exchange of ideas between regulators and service providers.*

- *Encourage co-creation of solutions: The active participation of service providers in the design of regulatory solutions is essential. This co-creation can increase the acceptance and effectiveness of the systems implemented.*

2. Implement a common content classification and description system:

- *Adopt a consistent approach: It is recommended to consider adopting a common content classification and description system be adopted for linear and non-linear services and video-sharing platforms. This system should be developed in consultation with service providers to ensure uniformity and consistency in the protection of minors.*
- *Facilitate public understanding: A common classification would help the public, particularly children, young people, and parents/caretakers, better understand the nature of the content to which they have access on different platforms, thereby promoting more informed and appropriate choices.*

3. Highlight the benefits of standardisation:

- *Emphasising transparency and control: If implementing a common description system, it is important to emphasise the benefits of transparency, giving the public access to clear information about content. This not only facilitates parental control but also promotes greater trust in the platforms.*
- *Ensure a minimum level of protection: Standardising rating criteria would help to ensure that all platforms provide a minimum level of protection for children and young people from inappropriate content.*

4. Overcoming implementation challenges:

- *Prepare for challenges: While implementing a common system may face challenges, such as adapting to different types of content and platforms, it is essential to address them proactively. This can include providing clear guidelines and technical support to ease the transition.*
- *Promote education and awareness: In addition to regulatory measures, it is recommended that investments be made in media/digital literacy to help the public understand and use the new classifications effectively.*

These recommendations aim to improve the effectiveness of co- and self-regulatory systems and to ensure a more consistent and robust protection of minors in the audiovisual environment.

3. Children Protection Mechanisms / Age Verification

The AVMSD establishes a regulatory framework for video-sharing platforms and other audiovisual communication services, focusing on the protection of minors. Updated in 2018, the Directive addresses several mechanisms to ensure the safety and protection of children in the online environment.

Among the mechanisms for the protection of minors, we highlight:

(i) Inappropriate content: The Directive restricts the display of content that may be harmful to minors. This includes banning content that promotes hatred, violence, exploitation, or other harmful behaviour.

(ii) Video-sharing platforms and platforms providing video-on-demand services must implement measures to protect minors from harmful content. These may include age rating systems and parental control over what their children can see.

(iii) Advertising: There are specific rules on advertising aimed at minors, setting out what is acceptable and what should be avoided. Advertising should not be misleading or take advantage of children's inexperience.

(iv) Responsibilities of service providers: VSPs are responsible for taking appropriate measures to protect minors using their platforms. This includes content moderation policies and reporting mechanisms.

These mechanisms aim to create a safer online environment for minors, considering the constant evolution of technologies and forms of content consumption. The practical implementation of these measures is essential to protect children from the risks associated with the digital world.

3.1. Age verification mechanisms

In 2023, ERGA members were questioned about their role in the implementation of age verification mechanisms (AVM) (Articles 6a(1) and 28b(3) of the AVMSD), giving an overview of the entities with which cooperation has been established. The report highlighted the involvement of some NRAs in the discussions/preparation of AVMs (Age Verification Mechanisms) at their national level, mainly by providing input to their government and/or parliament in the legislative process and also the strong cooperation with data protection authorities.

This year's survey aimed to go a little deeper into the question of which actors in the value chain NRAs cooperate with when implementing age verification mechanisms. Although the majority of respondents (22) did not reply, the survey revealed that public sector bodies (including data protection authorities, educational institutions, government, etc.) only accounted for 4 of the cooperation reported by the inquiries.

Both technology providers (e.g., set-up boxes and smart TVs) and intermediary service providers (hosting VSPs) gathered approximately four of the answers. However, the result that really stands out from the survey is what appears to be an ongoing dialogue with industry stakeholders (audiovisual media services, businesses) – accounting for up to six NRAs that identified such cooperation.

The survey delved further to gather the overall perception of the NRAs that pursued cooperation with the abovementioned entities about the tangible results derived from this collaboration and whether there were perhaps key takeaways the respondents would like to see shared in the report.

In this context, AGCOM and the Italian Data Protection Authority have formed a joint Committee to promote a code of conduct for digital platforms to implement age verification systems, facilitating cooperation between the two authorities. Additionally, AGCOM has launched a public consultation[16] to define procedural and technical methods for age verification. The final adoption of the new regulation is expected in September and will be notified to the European Commission before its entry into force (at the end of the 90-day standstill period). In practice, the approach adopted by AGCOM was technologically neutral and aimed at leaving the subjects required to implement age verification processes, i.e. the regulated subjects, a reasonable freedom of assessment and choice, while establishing the principles and requirements that the systems implemented must comply with[17].

AGCOM's new regulation on age verification concerns all online platform services disseminating pornographic content, through VSPs and websites. AGCOM has established that a functional system to provide the "age guarantee" must comply with a number of procedural and system requirements and specifications: 1. Proportionality; 2. Protection of personal data: the implemented age guarantee systems must comply with the data protection rules and principles established by Regulation (EU) 2016/679 (data minimisation, accuracy, storage limitation, etc.), provide adequate information to users and ensure that only and exclusively the personal necessary data for the purpose is collected. Finally, it considers that regulated entities and third parties involved in the age verification process and related processes (e.g. system maintenance, service management or billing, etc.) should not profile users and, in particular, the age verification mechanisms implemented should not allow regulated entities to collect the identity, age, date of birth or other personal information of users. 3. Use of independent third parties: As a general rule, the Authority considers that an age verification

[16] https://merlin.obs.coe.int/article/10035

[17] https://www.agcom.it/competenze/consumatori/interventi-regolamentari-tutela-degli-utenti-finali-attuazione-del-nuovo/tutela-minori-age-verification

system which provides for two logically separate steps will comply with these specifications: identification and authentication of the person identified for each session of use of the regulated service.

Similarly, CPTRA (Estonia) initiated an umbrella organisation for media service providers to establish a standard symbol system to inform viewers about age-restriction requirements, which have already been developed and are in use.

In Spain, a task force led by the Ministry for Digital Transformation and Public Service, including CNMC, is developing a government-provided age verification solution. This group, including the Spanish Agency for Data Protection (AEPD), which advises on data protection issues based on its' Decalogue of Principles[18], and discusses the suitability criteria for the solution. CNMC considers self-declaration ineffective for verifying legal age and parental control mechanisms and restricted to adult labels (RTA) insufficient.

In April 2024, CNMC published a summary of responses to its public consultation on age verification systems. Most respondents (85%) prefer non-face-to-face verification, with the main industry options being official identification documents and facial age estimation[19].

In France, after the survey, ARCOM was granted, by 21st May of 2024 Law, administrative blocking powers of pornographic services accessible to minors and had to establish a technical framework setting out the minimum technical requirements for age assurance systems. In April 2024, ARCOM published a public consultation on a draft technical framework that promotes, like AGCOM, a technologically neutral approach[20]. This framework was adopted and published on 11th October 2024[21]. Pursuant to the law, the framework establishes both reliability and privacy criteria for age assurance systems; this is why ARCOM has been closely collaborating with the French Data protection authority (CNIL) while drafting the document. In addition, the CNIL must be formally consulted for its opinion before adopting the final technical framework.

[18] Decalogue of Principles.: https://www.aepd.es/guides/decalogue-principles-age-verification-minors-protection.pdf

[19] CNMC has not yet verified the effectiveness of these solutions: https://www.cnmc.es/prensa/respuestas-cp-verificacion-edad-plataformas-20240417

[20] Available at https://www.arcom.fr/vos-services-par-media/consultations-publiques/consultation-publique-sur-le-projet-de-referentiel-determinant-les-exigences-techniques-minimales-applicables-aux-systemes-de-verification-de-lage-mis-en-place-pour-acces-contenus-pornographiques-en-ligne

[21] Available here only in FR: https://www.arcom.fr/presse/acces-des-mineurs-aux-contenus-pornographiques-larcom-publie-son-referentiel

3.2. Certification of age verification tools

The survey also questioned whether or not NRAs should certify age verification tools. In this regard, the majority of ERGA members stated that they currently did not have an official position, making it a case for possible further discussion.

Of the respondents who answered “no” (4), CPTRA and ALIA highlighted the possible excessive workload that would require significant technical expertise and resources. In their view, by maintaining current monitoring practices, NRAs ensure the adequacy of the tools used by platforms and their effectiveness in preventing minors from accessing harmful content, failing which they are subject to corrective measures. Therefore, direct certification requires work and resources and may go beyond the legal mandate.

Finally, in the French view, VSP providers are responsible for the protection of minors and therefore are expected to find the appropriate age verification mechanisms. NRAs should, therefore, be able to require providers to carry out audits and be transparent about their results to ensure compliance.

Concerning those ERGA members who argue that NRAs should certify age verification tools (also 4), DLM (Germany) emphasized that certification can provide the necessary legal certainty for providers of age verification tools and media service providers who implement these tools. In Germany, developers/providers of age verification tools can submit their tools for assessment by the Commission for the Protection of Minors (KJM). The assessment is carried out by simply giving the green light to the submitted system. However, this does not prevent companies from using different methods that meet the abstract requirements set by the German State Media Authorities.

The CvdM, on their side, considers that it may be appropriate to certify the age verification tool if it uses certain personal information – and this is often the case. It may also be appropriate to carry out such verification when the tools are used to provide or limit access to certain content.

3.3. Privacy concerns posed by AVMs

Another dimension addressed in the survey related to NRAs’ concerns about the privacy of age verification tools within their jurisdiction.

Again, excluding those for whom the question did not apply (because they had no VSPs under their jurisdiction), almost half of the respondents (9) stated they had no data available on privacy concerns. 3 NRAs – CEM (Bulgaria), DLM, and KRRiT (Poland) –, indicated that there were no concerns in their jurisdiction, while 6 (CRTA, ARCOM, ALIA, NMA, AKOS and CNMC) considered that there were concerns in theirs.

About the latter, all of the respondents (except for ALIA, in the two dimensions signalled below) considered the following practices of age verification tools to be of concern[22]:

- Data collection – collection practices and the potential for misuse of personal sensitive information;
- Data security – if information breaches occur, data can be exploited and used for malicious purposes (ALIA did not endorse this dimension);
- Data retention – the length of time this kind of data is retained, without proper justification, poses privacy concerns (ALIA did not subscribe to this dimension);
- Profiling and targeting – use of collected data for profiling and/or targeting advertisements based on age;
- Accuracy and Bias – effectiveness of the tools, leading to inaccurate age measurements

➢ ***Conclusions:***

The mechanisms for the protection of minors provided for in the AVMS Directive reflects a comprehensive commitment by the European Union to the protection of minors in the audiovisual environment, both in traditional and digital media. AVMs are a key tool mentioned in the Directive to ensure that minors do not have access to inappropriate content. Cooperation between NRAs, technology providers and other stakeholders is essential for the development and implementation of effective age verification solutions.

Last year's report stated that t*he double-blind solution and the intervention of an independent intermediary are options considered by many NRAs, showing the concerns regarding privacy. In this regard, a solution through digital ID seems preferred by most NRAs although some are not completely convinced. Self-declaration is almost unanimously discarded as an efficient AVM*.

The effective implementation of AVMs faces challenges, such as the need for certification, which some members of ERGA consider essential to ensure legal certainty for providers of age verification tools. However, other members are concerned about the work and resources required for such certification and the risk of going beyond the legal mandate of NRAs.

There are significant privacy concerns related to age verification mechanisms, particularly in areas such as data collection, security, data retention, and the use of data for profiling and

[22] CNMC added an infographic in where the Spanish Data Protection Authority identifies risks associated with age verification systems and is available in the following link: https://www.aepd.es/infographics/infographic-risks-age-verification-systems.pdf

advertising targeting. These concerns have been identified by several NRAs, reflecting the need to address these issues to ensure that age verification tools are secure and respect users' privacy.

Examples of cooperation, such as the joint efforts in Italy and Spain to develop codes of conduct and government solutions for age verification, show that cooperation between different bodies, including Data Protection Agencies and NRAs, is key to addressing regulatory challenges.

Approaches to the implementation and monitoring of age verification mechanisms vary from country to country. While some countries, such as Germany, see certification as a way to provide legal certainty, others prefer to maintain monitoring practices without formal certification, depending on national circumstances and available resources.

These findings highlight the need for a balanced and collaborative approach to implementing age verification mechanisms that protects minors while addressing privacy concerns and regulatory effectiveness. Given the shared objectives of privacy, security, and cross-border recognition of digital solutions, *eIDAS* could serve as a valuable regulatory reference for a coherent, privacy-respecting implementation of age verification mechanisms across the EU.

➢ ***Points to Consider:***

Based on the conclusions presented, the following points for reflection can be made:

- *Strengthen the effectiveness of age verification systems: Regulators and industry should work together to ensure that all age verification systems in use are effective and reliable.*
- *Continuous monitoring: Consider establishing a system of ongoing monitoring and evaluation to ensure that AVMs are working properly and meeting the objectives of protecting minors.*
- *Encourage voluntary certification and voluntary auditing of certification: Although there is ambivalence, voluntary certification of age verification systems, with a focus on data protection and privacy, can increase user confidence and transparency.*
- *Education and communication: Promote educational campaigns to inform the public about how AVMs protect privacy and the importance of age verification for online safety.*
- *Developing partnerships: Encourage stronger partnerships between regulators, public authorities, namely Data Protection Authorities, and industry to address the challenges of implementing AVMs. Sharing best practices and developing tools together can improve the effectiveness and acceptance of these systems.*

- *Stakeholder involvement: Continue to involve all stakeholders in the process of developing codes of conduct, ensuring that the voices of all interested parties, including parents and children are taken into account.*
- *Investing in research and development: Support the research and development of new technologies, such as advanced biometric methods and AI solutions, to create age verification systems that are less invasive and more accurate.*
- *Balance innovation and privacy: When adopting new technologies, ensure that they respect users' privacy rights and strike a balance between technological innovation and data protection, in particular by complying with the principle of data minimisation (Article 5 of the GDPR[23]) and the principles of data protection by design and by default (Article 25 of the GDPR).*

These recommendations can help strengthen the protection of minors in the digital environment while addressing the ethical and technological concerns associated with the use of age verification systems.

4. Flagging Content and Respective Criteria

Flagging or labelling of content in audiovisual media services is of the utmost importance because it helps protect vulnerable audiences, and minors, from exposure to inappropriate or harmful material. As we know, this is especially important given the potential negative impact such content can have on youngsters.

This not only helps to build trust and safety, encouraging more people to use media services with confidence but also ensures compliance with legal requirements and platform policies, helping to avoid penalties and maintain a respectful community standard.

Flagging or labelling harmful and illegal content relies on several key elements (protecting audiences, enhancing user experience and trust, ensuring compliance, preventing misinformation, etc.). Between these key elements, the survey focuses on identifying and marking/classifying certain content based on specific criteria according to its nature/quality (such as explicit material, hate speech, misinformation, or other forms of potentially harmful or illegal content).

Audiovisual media services can better protect their users, comply with legal standards, and maintain a positive and safe media environment by establishing and adhering to clear criteria for flagging or labelling of content.

[23] General Data Protection Regulation - https://eur-lex.europa.eu/eli/reg/2016/679/oj

4.1. Criteria for flagging or labelling

The criteria for flagging or labelling of content in audiovisual media services usually include the dimensions set out in the survey sent to ERGA members, including, but not limited to, violence and gore, hate speech, misinformation, unwanted commercial content, child exploitation, and mental health. These are summarised in the table below, for which the survey gives us insight into how ERGA members classify content as harmful or illegal according to different criteria.

Content	Illegal	Harmful
Violence and Gore:		
Content that displays explicit or gratuitous violence, such as murder, mutilation, torture, or animal abuse.	**45%**	**38%**
Images of real-world violence, such as wars, terrorist attacks, or severe accidents.	**14%**	**72%**
Content that promotes or glorifies violence.	**24%**	**17%**
Content that is offensive or that may cause psychological harm to users.	**31%**	**45%**
Hate Speech:		
Content that promotes hatred or discrimination against individuals or groups based on race, religion, ethnicity, nationality, sexual orientation, gender identity, disability, or any other protected characteristic.	**86%**	**3%**
Content that incites violence or hatred against individuals or groups	**83%**	**7%**
Misinformation:		
Content containing false or misleading information with the intention of manipulating or harming users[24].	**7%**	**28%**

[24] Examples: Fabricated news stories: These are entirely made-up stories disguised as legitimate news articles. They often target specific groups of people by exploiting existing biases and anxieties, aiming to sow discord or promote violence; Deepfakes: These are artificially generated videos or audio recordings that manipulate existing footage to make it appear as if someone is saying or doing something they never did. They can be used to damage someone's reputation, spread misinformation, or interfere with elections.

Content that promotes conspiracy theories or fake news[25].	**3%**	**31%**
Content that disguises itself as real news to spread misinformation[26].	**3%**	**31%**
Spam and Unwanted Commercial Content:		
Content that aims to deceive or manipulate users for profit-oriented purposes.	**55%**	**17%**
Bots or automated accounts that post spam or unwanted commercial content.	**24%**	**41%**
Sexual Abuse and Child Exploitation:		
Content depicting child sexual abuse, including child pornography and sexual exploitation of minors.	**90%**	**-**
Content that promotes or glorifies child sexual abuse.	**83%**	**-**
Images of children in risky or exploitative situations.	**21%**	**10%**
Content perceived to be harmful online, including negative body image and eating disorder content.	**14%**	**45%**
Content glamourizing unhealthy or abusive lifestyles, and the promotion of self-harm.	**31%**	**52%**
Mental health		
Content perceived to be harmful online, including negative body image and eating disorder content.	**7%**	**28%**
Content glamourizing unhealthy or abusive lifestyles, and the promotion of self-harm.	**17%**	**55%**

[25] Examples: Articles claiming that the moon landing was faked by NASA; social media posts alleging that vaccinations cause autism; websites spreading the conspiracy theory that the earth is flat; social media accounts spreading false information about election fraud without credible evidence, etc.

[26] Examples: Fake news: False or misleading articles that resemble real news; manipulated content: Images, videos, or audios edited to distort the truth; deepfakes: Videos manipulated with artificial intelligence technology to make it appear that someone is saying or doing something they did not say or do; misleading headlines: Headlines that exaggerate or distort the content of an article; opinion articles disguised as news: Articles that present opinions as facts; sensationalist news: News that exaggerates or distorts facts to attract public attention; sponsored or promoted content: Content paid for by a company or organization to promote a product, service, or agenda.

In the «Violence and Gore» dimension, the table shows that respondents, overall, are more inclined to label real-world violence as harmful rather than illegal, whereas explicit violence and gore are considered both illegal and harmful. Content that promotes violence is moderately viewed as illegal and offensive content is seen as harmful by a significant number, almost half of the respondents.

Regarding the hate speech criteria, the data indicates a clear consensus that hate speech—whether it promotes hatred and discrimination or incites violence and hatred - is overwhelmingly considered illegal rather than just harmful. The near-universal agreement on this issue reflects a strong stance against this type of content.

Misinformation, especially when disguised as real news or promoting conspiracy theories, is most often seen as harmful by the 8 of respondents who had a view on the matter. Similarly, general false or misleading information with the intention of manipulating or harming audiences is more likely to be seen more harmful than illegal. The data reflect a general attitude against the dissemination of misinformation, but not through prohibitive measures.

When it comes to spam or unsolicited commercial content, for-profit misleading or manipulative content is largely seen as illegal by more than half of respondents. In contrast, while bots and automated spam are seen as harmful by a significant proportion of respondents, fewer see them as illegal, perhaps indicating a preference for mitigation rather than outright prohibition. This highlights a nuanced approach to classifying different types of spam and unwanted commercial content.

The table above also shows a strong consensus on the illegality of content related to child sexual abuse and exploitation, with the majority of respondents classifying such content as illegal. Less serious - but still globally perceived as harmful - is content such as negative body image and the promotion of self-harm, according to a significant proportion of respondents. Overall, the data reflect a clear prioritisation of labelling as illegal the most obvious and intentional forms of content while recognising the harmful effects of other types of sensitive content.

Finally, there is a clear concern among respondents about content related to mental health issues, content promoting self-harm and unhealthy lifestyles. While there is less support for outright illegality, there is a strong recognition of its harmful effects.

4.2. Mechanisms for identification of harmful and illegal content

Beyond content labelling criteria, identifying harmful or illegal content (including for minors) in VOD and VSP is a complex challenge. However, several tools can assist in this process. The survey, therefore, asked NRAs whether they had identified any of these mechanisms in addition to the platforms under their jurisdiction.

It is notable that around a quarter of respondents (9) did not identify any of the above tools, as no answer was given. However, it is encouraging to note that 17 responses indicate that the human element remains a crucial aspect of identifying harmful content. This is evidenced by the fact that 'human moderation' tools were mentioned in 8 responses, while the ability of video-sharing platforms to allow users to report content they consider harmful was highlighted in 7 responses. Two respondents mentioned the possibility of using moderators who are well trained to identify different types of content and make fair and consistent decisions.

Other tools with little or no human intervention appear to be more fragmented. Mechanisms such as visual hash detection, convolutional neural networks or natural language processing are present in 2 responses, as one of the tools found by NRAs on platforms under their jurisdiction.

4.3. Effectiveness of the mechanisms

It would also be interesting to know which mechanisms are considered most effective by NRAs, but the survey did not collect enough hard data or case studies to suggest that regulators either believe that effectiveness comes from a mix of tools, or that there are no clear winners when it comes to detecting harmful or illegal content. In this context, we highlight Italy's response to the survey, which recognizes the effectiveness of Convolutional Neural Networks (CNN) for identifying patterns and Natural Language Processing (NLP) for keywords and phrases while also acknowledging that human moderation remains essential for understanding the context and intent behind the content. Thus, the integration of advanced technologies with human experience and judgment provides a more comprehensive and effective protection for children online.

4.4. General basis for NRAs to act upon

In addition to determining the effectiveness of the above mechanisms, the survey was also designed to understand the general basis on which NRAs act upon harmful content, either through users’ or institutions’ complaints, content monitoring, or both. At the same time, it sought to establish whether NRAs find it useful for regulators to have their tools for detecting harmful content and whether they currently have/use tools for detecting harmful content.

Regarding the first question - excluding the 5 NRAs in the total universe who did not answer this question - only 2 (KommAustria and CRTA) said that they rely on user complaints, while the overwhelming majority – 22 - confirmed that their action on harmful content is based on both content monitoring and complaints.

On the second question of whether regulators find it useful to have their own tools for identifying harmful content, 19 respondents answered affirmatively whilst 6 did not find it useful.

Among those in favour of NRAs having their own tools, the underlying motives for this choice were related, but not limited, to increased effectiveness (8 respondents), while 1 saw an increase in standardization and consistency and, finally, 11 (the largest number of responses) considered it to be a matter of independent supervision.

On the side of defendants who feel that it is not particularly useful for regulators to have their tools for dealing with harmful content, many see cost and complexity (6) and technical challenges (5) as reasons for moving away from these tools, closely followed by the potential for censorship (3). ARCOM added that the responsibility for implementing these tools lies primarily with the content providers themselves and that, in addition, NRAs have the possibility of occasionally relying on third parties, such as private companies or civil organizations specializing in the screening of online content, if necessary.

Following the questions above, NRAs were asked whether they were already using tools for detecting harmful content. 22 NRAs provided an answer, revealing an interesting twist: even though 19 respondents think it would be useful for NRAs to have tools, only 4 have such mechanisms available within their organization.

For the majority of the respondents that do not have or use harmful content detection tools, there is no clear trend in the intention to acquire them in the future. Responses are split between those who plan to do so and those who have no plans at all. Some raise the issue of resources and the technical expertise needed to use them.

For those who already have these detection tools, we highlight the one developed in Germany and used since 2021. An AI-supported tool named "KIVI", searches for potential breaches of the law online and thus works ahead of its employees. The tool focuses on the protection of human dignity and the protection of minors. Following the relevant German legal framework, the specific offense categories include, for example, depictions of violence, incitement to hatred, the use of anti-constitutional symbols, or freely accessible pornography.

Belgium's CSA has developed a cooperation to also use "KIVI", to monitor pornographic content on X and is currently being trained with datasets to detect hate speeches circulating in French on YouTube.

➢ ***Conclusions:***

Content labelling in audiovisual media services is essential to protect vulnerable audiences, in particular minors, from illegal or harmful material. This practice not only protects users, but also contributes to compliance with legal requirements and platform policies, as well as it contributes to maintain respectable community standards.

The criteria for flagging or labelling of content include categories such as explicit violence, hate speech, disinformation, child exploitation and mental health issues. There is consensus on the need for clear, standardized criteria for classifying and labelling such content, which helps to protect users and ensure a safe media environment. Most regulators consider explicit violence, hate speech and child exploitation to be illegal content. Other types of content, such as disinformation, spam and content harmful to mental health, are often classified as harmful but not always illegal, indicating a more nuanced approach to their management.

Human moderation remains an important part of identifying harmful content, with many responses indicating its importance. However, there is also recognition of the usefulness of automated tools, such as visual hash detection and neural networks, which complement human experience.

There is no clear consensus on which mechanisms are most effective in detecting harmful content. Italy's response highlights the combination of advanced technologies with human moderation as the most complete approach to protecting minors.

Although 19 of regulators consider it useful to have their own tools to detect harmful content, only 4 have such tools. The remainder rely mainly on monitoring and complaint mechanisms. Barriers to adopting their own tools include cost, technical complexity and concerns about censorship. The German tool "KIVI" is highlighted as an example of an effective AI-based solution used to monitor potential violations of the law online, especially in relation to the protection of human dignity and minors.

➢ ***Points to Consider:***

On the basis of the conclusions presented, a number of points for reflection can be made in order to improve the protection of users, in particular minors, in audiovisual media services:

1. Improve content labelling criteria:

- *Develop clear and consistent guidelines: Consider developing guidelines for labelling content, such as violence, hate speech and disinformation. This will help ensure consistent and effective enforcement and protect users from harmful content.*
- *Promote moderator training: Ensure that human moderators are adequately trained to correctly identify different types of harmful content. Combining advanced technology with human judgement can increase the accuracy and fairness of moderation decisions.*

2. Encourage the integration of advanced tools with human oversight:

- *Combining AI with human moderation: Integrating technologies such as convolutional neural networks (CNN) and natural language processing (NLP) with human moderation can provide a more comprehensive and effective approach to detecting harmful content.*
- *Adopt advanced detection tools: Regulators who do not already use harmful content detection tools should consider testing such technologies, in order to implement the most effective mechanism for their local use.*

3. Strengthen cooperation and sharing of resources:

- *Cooperation between regulators and institutions: Encourage cooperation between different regulators and platforms to share tools and best practices. While there's no one size fits all technology, using technologies developed by other countries or organisations can help reduce technical and financial barriers.*
- *Establish partnerships with third parties: Consider working with private companies or civil society organisations that specialise in monitoring online content, especially when internal resources are limited.*

4. Support independent monitoring and transparency:

- *Promote independent oversight: Regulators should have their own tools to ensure independent oversight and avoid over-reliance on external platforms or vendors that may compromise impartiality.*
- *Ensure transparency of labelling processes: Regulators and platforms should transparently communicate the criteria and processes for flagging or labelling of content to the public in order to increase user confidence in the system.*

5. Regularly evaluate and update tools and criteria:

- *Monitor the effectiveness of the tools: It is important to regularly evaluate the effectiveness of detection tools and flagging or labelling criteria to ensure that they remain appropriate in the face of evolving content and technologies.*
- *Adapt to new threats: Regulators should be prepared to adapt labelling criteria and detection technologies in response to new forms of harmful content that may emerge.*

These recommendations can help create a safer and more trusted media environment, especially for the most vulnerable users, such as children.

5. Measures to be taken by VSPs

Article 28b of the AVMSD was introduced as part of the reforms to adapt media legislation to the new realities of the digital environment and changes in content consumption. It deals with the protection of minors from exposure to harmful content. It sets out requirements for providers to ensure that content accessible to children and adolescents does not contain material that could be considered harmful to their physical, mental, or moral well-being.

The main concerns surrounding this article include the effective implementation of protection standards, the balance between freedom of expression and the protection of minors, and questions regarding the responsibility of digital platforms and streaming services with the content they provide.

In short, Article 28b is a key element in the attempt to modernize audiovisual media regulation in the EU, reflecting the need for the protection of minors in an increasingly diverse and accessible media environment.

One of the specific features of the obligations imposed on the VSPs is the one provided for in Article 28b(3) of the Directive, namely the requirements applicable to the terms and conditions of the platform, the requirements imposed on the uploading systems, the characteristics of the systems and tools that should be made available to users, all of which are subject to scrutiny by NRAs.

Today, in a more robust environment of VSPs registered with their respective NRAs, the verification of terms and conditions following Article 28b(3) of AVMSD appears to be a good indicator of a first approach and step to ascertain the action of NRAs towards VSPs about child protection. In this regard, it is also relevant to mention the DSA, which complements the AVMSD by requiring video-sharing platforms and other online platforms to implement accessible and user-friendly reporting mechanisms and transparent content moderation practices. These requirements align with Article 28b of the AVMSD, ensuring that platforms take due care in preventing minors’ exposure to harmful content while facilitating effective user reporting and platform accountability.

The provision in Article 28b(3)(a) of AVMSD establishes that, among others, VSPs must adopt measures to include and apply, in the terms and conditions of the services of video-sharing platforms, requirements for the protection of minors against programs, user-generated videos and audiovisual commercial communications that may harm their physical, mental or moral development.

5.1. Terms and conditions

The results of the survey showed that 12 of the NRAs with VSPs under their jurisdiction had not reviewed the terms and conditions of video-sharing platforms. Among the reasons given for not carrying out the review were "other priorities" or a "different philosophy of action".

In fact, RpMS stated other priorities at the moment, since it is currently focusing on market analysis and communication with potential VSPs to be designated. CNMC is giving priority to monitoring the obligation to implement age verification systems, mainly due to the characteristics of the VSPs under Spanish jurisdiction. Meanwhile, Memy and NMA (Norway) are in the process of setting up monitoring mechanisms, which should start soon, in 2024-2025.

As for those with a different approach to the verification of terms and conditions, there is the case of DLM, which addresses VSPs on the basis of complaints, stating an example of a VSP that changed certain paragraphs of its terms and conditions that had been the subject of a complaint.

For its part, CvdM is currently working on the Code of Conduct with Snap (co-regulation of the only VSP under its jurisdiction). The Code should include, where appropriate, the measures listed in Article 28b(3) of the Directive.

On the other hand, 4 of the respondents (again excluding those for whom the question did not apply) stated they had verified the terms and conditions of VSPs and found them generally adequate in meeting the provisions of the AVMSD, particularly complying with Article 28b (ERC and KRRiT).

In Luxembourg, ALIA, despite not having conducted a detailed analysis of the terms and conditions of the concerned VSPs, reported engaging in information exchanges and providing guidance to VSP providers to help them fulfil their obligations, including information regarding terms and conditions.

In Hungary, NMHH noted that some VSPs have terms and conditions that are difficult to understand.

5.2. Consequences of breach of the terms and conditions

Remaining within the scope of supervision of terms and conditions, the third question of the questionnaire was handed out to collect information about the consequences of breaching them and whether they were clearly and easily explained.

When users are aware of the specific consequences of breaching terms and conditions, they are, in principle, more likely to adhere to the rules. This can help prevent misconduct and promote a healthier and more respectful user community.

The answers received showed that most respondents had no data available (10)[27]. 4 considered that the consequences of an eventual breach are easily understandable (CEM, NMHH, CvdM and KRRiT) and only ERC observed these to be poorly explained and presented to users.

5.3. Mechanisms to report or flag content

The survey produced a very similar set of results on the abovementioned issue upon the question handed out to learn if the mechanisms established by the VSPs to report or flag content are transparent and user-friendly.

Article 28b(1) of the AVMSD sets out mandatory measures for which a lack of user-friendliness may hinder the desired effects; hence, the responses are interesting and it is important to know if they are monitored by all NRAs.

Transparent and user-friendly mainly refers to the mechanisms that provide accessible reporting options (quick with minimal steps, with guidance, easy-to-find locations, use of consistent imagery such as flags or exclamation marks) while also providing users with confirmation of the steps they have taken (submitting complaints/reports).

The majority of NRAs (11)[28] replied that they had no data on the issue and 4 (CEM, NMHH, CvdM and KRRiT) considered that the mechanisms mentioned were transparent and user-friendly.

5.4. Handling of user complaints

As for the systems implemented by VSPs to explain to users the follow-up given to reported or flagged content and whether the procedures for dealing with user complaints to VSPs were transparent, easy to use, and effective, the results for both questions were very similar. 13 respondents said they had no data on the subject, while 2 did not answer, which represents 88% of respondents who did not provide any information/input on the subject. As for the remaining 12%, CEM considered them "[q]uick", while KKRiT said that "[t]here was room for improvement".

Finally, concerning the transparency, user-friendliness, and effectiveness of the procedures for dealing with user complaints to VSPs, CEM considered them to be positive, while ERC answered "No" to the same question. The lack of substantial responses may be an indication of the need for further discussions.

[27] Does not include NRAs that do not have VSPs under their jurisdiction.

[28] Does not include NRAs that do not have VSPs under their supervision.

➢ *Conclusions:*

Article 28b of the AVMSD is crucial for the modernisation of audiovisual media regulation in the European Union, in particular with regard to the protection of minors from harmful content. It sets out clear requirements for video-sharing platforms (VSPs) to protect minors, balancing freedom of expression with the need for safety.

One of the main challenges is the effective implementation of the protection standards. Many NRAs have not yet reviewed the terms and conditions of VSPs under their jurisdiction, with 12 not having done so. Reasons for this include other priorities or different approaches, such as a focus on market analysis or communication with VSPs.

Some countries, such as Germany, take a complaint-based approach, while others, such as Portugal and Poland, consider the terms and conditions of VSPs sufficient to comply with the Directive's provisions. The Netherlands is working on a code of conduct in cooperation with platforms such as Snap.

The transparency and user-friendliness of content flagging tools and complaints procedures are critical areas, but many regulators (11) do not have sufficient data on their effectiveness. Where evaluation has taken place, some regulators have found the tools to be transparent and easy to use, but there is room for improvement, particularly in the follow-up of user complaints.

The lack of substantive responses on the effectiveness of signalling systems and complaints suggests a significant gap in the monitoring and enforcement of these measures. Most regulators do not have data on the functioning of the systems, suggesting the need for further discussion and possibly more supervision or regulation.

While some authorities found the consequences of breaching terms and conditions to be clear and understandable, others, such as Portugal, found them to be poorly explained. This highlights the need to improve communication to ensure that users fully understand the rules and consequences.

➢ *Points to Consider:*

Based on the concerns and findings presented, here are some points for reflection to improve the implementation of Article 28b of the AVMSD and the protection of minors on video-sharing platforms (VSPs):

1. Prioritize review of terms and conditions by NRAs:

- *Encourage regular review: NRAs should consider giving priority to reviewing the terms and conditions of VSPs under their jurisdiction. This review should ensure that the requirements of Article 28b(3) are clearly implemented and adapted to digital realities.*
- *Set deadlines and guidelines: Setting clear deadlines for reviewing the terms and conditions and detailed guidelines that may help NRAs fulfil their responsibilities more effectively.*

2. Improve transparency and clarity on the consequences of non-compliance:

- *Clear and effective communication: VSPs should be encouraged to explain the consequences of breaches of terms and conditions in a clear and accessible manner. This may include automated notifications or explanatory guides for users to improve compliance and promote a safer online environment.*
- *Ongoing monitoring: NRAs should consider actively monitoring whether the consequences of non-compliance are effectively communicated to users and take corrective action where necessary. This could be done by having platforms notifying/reporting from time to time also to NRAs.*

3. Strengthening content signaling mechanisms:

- *Improved usability and accessibility: It is essential that the mechanisms for flagging and reporting content on VSPs are transparent, easy to use and accessible. Platforms should simplify the reporting process and provide clear feedback to users on the status of their reports.*
- *Common standards for reporting mechanisms: Consider creating common standards for the transparency and usability of these mechanisms to ensure that all VSPs provide effective tools to protect minors.*

4. Encourage further discussion and study of reporting practices:

- *Undertaking studies and sharing best practices: NRAs should consider carrying out further studies and exchange best practices on reporting procedures and the follow-up of user complaints. This could include regular consultations with VSPs and reviews of current practices to identify areas for improvement.*
- *Regular discussion forums: Organizing regular discussion fora between NRAs and VSPs can help align expectations and improve the effectiveness of safeguards.*

5. Training of NRAs in the use of monitoring tools:

- *Tool development and implementation: NRAs should consider developing or adopting technological tools to proactively monitor VSPs for compliance with Article 28b. These* *tools could include automated systems that alert on potential breaches, allowing for a faster and more effective response.*

- *Training and resources: Consider investing in training and capacity building for NRAs to use these monitoring tools and ensure that they have the necessary resources to carry out their supervisory tasks effectively.*

These recommendations aim to strengthen the protection of minors, improve the transparency and effectiveness of reporting systems in VSPs and ensure that NRAs play an active role in the supervision and enforcement of the measures set out in Article 28b of the AVMSD.