Follow-up questions

Shadows's meeting

14 December 2022

	EPP	S&D	Renew	Greens	ID	ECR	The Left
COM		1. What is your intention of the relationship between the EU Centre and Europol? Will they be permitted to share data without any barriers? And what do you consider it an effective firewall if the Centre resides at the Europol premises? 2. Did you consult MS' authorities on the current human resources to monitor CSAM and the potential impact this proposal will have on their resources?	No question	1. Will you propose to extend the validity of the interim Regulation, and when do you intend to make a legislative proposal on this? 2. What types of procedures and technologies, with what success or failure rates, has the Commission tested for the detection of unknown CSAM and of possible grooming attempts online, and to what extent? Please answer in detail and with references. 3. According to the proposed legislation, the EU Centre should act independently, receive reports from the services and platforms and sort out incorrectly identified material therein and forward only "relevant" (punishable?) material to the law enforcement agencies in the Member States. At the same time, the center should be responsible for deleting identified CSAM from providers and maintain a list of indicators that providers and platforms are obliged to use to find possible CSAM and to detect possible grooming attempts. On the basis of which rights and with whom can data subjects appeal wrongly reported data and false suspicions and, if necessary, claim damages and the restoration of wrongly deleted data?			1. Technologies such as client-side scanning (CSS) would need substantial rights in the operating system to scan multimedia data being transmitted through a hosting service or an interpersonal communication service in order to match it to an external database. For the process of matching, technical measures would have to be taken in order to ensure that users are protected against potential misuse. Which technological measures are already considered by the COM to ensure end user IT security and data protection and how can these technological measures guarantee end user IT security and data protection? 2. What is the accuracy of existing technologies for the detection of a) known CSAM, b) new CSAM and c) grooming with regard to false negatives and false positives and what are the possibilities and prospects for improving the accuracy, according to objective and accessible research? 3. According to various digital rights organisations, in particular edri, the Commission repeatedly refused to meet with them on the proposal. Is that true? With what other stakeholders has the Commission or the Commissioners herself met on the file? 4. According to many experts the proposal lacks a sufficient legal basis. Could you please comment on that? 5. Many stakeholders underline that the proposal fails to meet the key human rights principles of necessity and proportionality and violates several fundamental rights. As confirmed by the European Commission's own internal Regulatory Scrutiny Board (RSB), the proposal may also violate the EU prohibition of general monitoring. Can you please comment on that?

This non-paper prepared by the Commission's services aims to provide explanations with regard to the technical elements of the proposal for a Regulation on preventing and combating child sexual abuse.
This non-paper is based on the relevant Commission proposal and does not present any new positions with regard to that proposal.

Answers to questions by the S&D

1) What is your intention of the relationship between the EU Centre and Europol?

For the relationship between the EU Centre and Europol (and Coordinating Authorities), please see the dedicated paper attached.

CA-Centre-Europol_final.pdf

Will they be permitted to share data without any barriers?

Recital 71 of the proposal states that

> "Considering Europol's mandate and its experience in identifying competent national authorities in unclear situation and its database of criminal intelligence which can contribute to identifying links to investigations in other Member States, the EU Centre should cooperate closely with it, especially in order to ensure the swift identification of competent national law enforcement authorities in cases where that is not clear or where more than one Member State may be affected."

And Article 53(2) on the cooperation of the EU Centre and Europol states that

> "Europol and the EU Centre shall provide each other with the fullest possible access to relevant information and information systems, where necessary for the performance of their respective tasks and in accordance with the acts of Union law regulating such access."

This means that Europol and the EU Centre will not share data without any barriers. Europol and the EU Centre will share data only where necessary for the performance of their respective tasks, including the tasks for the EU Centre detailed in Article 43, and only in accordance with the acts of Union law regulating such access, including the Europol Regulation, the GDPR, as well as any applicable rule on professional secrecy.

And what do you consider it an effective firewall if the Centre resides at the Europol premises?

Article 53(3) on the cooperation of the EU Centre and Europol states that

> "the terms of cooperation and working arrangements shall be laid down in a memorandum of understanding."

This could include the arrangements to ensure security and data access restrictions.
That said, given the space available at the current Europol's premises and the space required to host the EU Centre as described in the proposal, it seems unlikely at this point that the Centre could reside at such Europol premises.

2) Did you consult MS' authorities on the current human resources to monitor CSAM and the potential impact this proposal will have on their resources?

Yes, Member States' authorities were consulted during the preparation of the proposal. Annex 2 of the Impact Assessments describes the multiple consultations that took place over two years in the preparation of the Impact Assessment for the proposal, including to Member States' authorities.

The reporting obligations for the companies (notably the requirements in relation to the information to be included in the reports) and the check function of the EU Centre will help ensure that the reports that reach law enforcement are actionable. This will alleviate the current workload and the time that national law enforcement has to dedicate to filter out non-actionable reports, mostly because they lack sufficient information for law enforcement to open an investigation.

In addition, the role that the EU Centre will play in facilitating efforts by Member States on prevention and assistance to victims, in particular by promoting the exchange of best practices and serving as a centralised knowledge hub, will likely reduce duplication of efforts and inefficiencies across the EU. National hotlines already in place could support the work of the EU Centre on prevention and assistance to victims: they currently constitute an important stakeholder in this field at national level, hence they are a key interlocutor to identify best practices and areas of improvements.

Answers to questions by the Greens

1) **Will you propose to extend the validity of the interim Regulation, and when do you intend to make a legislative proposal on this?**

On the possibility to extend the validity of the Interim Regulation, please refer to the comments made in the dedicated document.

The Commission stands ready to support the co-legislators to achieve an agreement on the proposal for the Regulation as soon as possible, to prevent the need for such an extension from arising. Depending on the progress of the negotiations, the Commission, following consultations with the co-legislators, will decide whether to present such a proposal for an extension.

2) **What types of procedures and technologies, with what success or failure rates, has the Commission tested for the detection of unknown CSAM and of possible grooming attempts online, and to what extent? Please answer in detail and with references.**

It is not within the Commission's competences to develop for commercial purposes technologies to detect unknown CSAM or grooming and test them in actual operating conditions, which are the conditions that produce the most relevant error rates. The Impact Assessment accompanying the proposal contains the data and evidence available in relation to such error rates as provided by the companies, civil society organisations and national authorities during the extensive consultations carried out in the preparation of the proposal.

Please refer to the responses to question 2) by the Left for additional details on these technologies and to question 3) by the Left on the consultations carried out in the preparation of the proposal.

3) **According to the proposed legislation, the EU Centre should act independently, receive reports from the services and platforms and sort out incorrectly identified material therein and forward only "relevant" (punishable?) material to the law enforcement agencies in the Member States. At the same time, the center should be responsible for deleting identified CSAM from providers and maintain a list of indicators that providers and platforms are obliged to use to find possible CSAM and to detect possible grooming attempts. On the basis of which rights and with whom can data subjects appeal wrongly reported data and false suspicions and, if necessary, claim damages and the restoration of wrongly deleted data?**

Data subjects affected by detection, reporting, removal or blocking of online CSA are entitled to:

- A right to complain with the service provider responsible for the data processing (Articles 10(4)(d)), 15(1), 18(3) of the proposal).
- A right to complain with the competent Coordinating Authority (Article 34 of the proposal).
- A right to effective judicial redress before national court, in accordance with the procedure available under national law (in line with the principle of national procedural autonomy). This right includes the right to challenge a detection, removal or blocking order before national courts (Articles 9(1), 15(1) and 18(1) of the proposal). Effective judicial redress necessarily encompasses a right of compensation for any damage that might result from wrongful

processing. In this respect, it might be useful to notice that the proposed regulation does not derogate from Article 82 GDPR, which provides for the data subject's right to compensation for any infringements of that regulation. Similarly, the proposal does not derogate from the rights attributed to the data subject, among others, by Articles 15 (right of access), 16 (right to rectification) and 17 (right to erasure) GDPR. Where relevant, the corresponding Articles of the Law Enforcement Directive (Article 56 on the right to compensation, Article 14 on the right of access and Article 16 on the right to rectification or erasure) are also applicable.

- A right to be informed about their right to complain to the service provider and Coordinating Authority, as well as to seek redress before national courts (Articles 10(5)(c), 12(2), 15(3)(c), 17(1)(i) of the proposal)

Data subjects affected by detection, reporting, removal or blocking of online CSA are also entitled to the rights provided by the Digital Services Act, when the service provider is also subject to the obligations of the Digital Services Act.

Answers to questions by the Left

1) **Technologies such as client-side scanning (CSS) would need substantial rights in the operating system to scan multimedia data being transmitted through a hosting service or an interpersonal communication service in order to match it to an external database. For the process of matching, technical measures would have to be taken in order to ensure that users are protected against potential misuse. Which technological measures are already considered by the COM to ensure end user IT security and data protection and how can these technological measures guarantee end user IT security and data protection?**

The proposal is technology neutral and future proof. It does not impose the use of any specific detection technology and leaves the choice of the most appropriate detection measure to providers having received a detection order, within the limits of the specifications set out in the proposal to ensure protection of data privacy, security, as well as accuracy and reliability of detection.

According to Article 50(1) of the proposal, the assessment of specific detection technologies to be included in the EU Centre list will be carried out by the EU Centre (supported by expertise in the form of a dedicated Technology Committee) in cooperation with the EDPB, whose opinion will be particularly relevant to ensure that data protection and privacy is safeguarded. The deployment of detection technologies on a case-by-case basis requires a judicial or independent administrative order and must be performed, in accordance with Article 7(3) of the proposal, under the supervision of the competent data protection authorities, based on the implementation plan presented by the provider and reviewed by the other authorities involved, including the data protection authority.

The impact assessment accompanying the proposal contains examples of technologies that could guarantee end user IT security and data protection (see in particular Annex 9). These include, for example, on-device CSS partial hashing technologies with remaining hashing and matching at server. In this case, part of the hash is generated at the device and the rest at the server, where the matching also takes place. This hybrid approach could be worth considering, as (compared to full hashing at the client and matching at the server) it makes the process lighter and ensures even more end user IT security and data protection. That said, the proposal does not require the use of CSS technologies or any other specific type of technology.

2) **What is the accuracy of existing technologies for the detection of a) known CSAM, b) new CSAM and c) grooming with regard to false negatives and false positives and what are the possibilities and prospects for improving the accuracy, according to objective and accessible research?**

The accuracy rate of existing technologies is as follows:

a) **For known CSAM**, the most widely used technology is a hashing technology known as PhotoDNA, which has an extremely high accuracy rate. The rate of false positives in tests has been demonstrated to be below 1 in 50 billion. PhotoDNA has been in use for more than 10 years by over 150 organisations globally including service providers (Microsoft, Facebook, Twitter, Apple), NGOs (e.g. NCMEC, Internet Watch Foundation) and law enforcement in the EU (e.g. Europol, DE, SE and others). In these 10 years, the tool has been used daily and analysed trillions of images without any accuracy concerns being

identified. Other examples of hashing technology used for these purposes, and operating on similar principles, include YouTube CSAI Match, Facebook's PDQ and TMK+PDQF.

b) **For new CSAM**, technologies currently used include **classifiers**. A classifier is any algorithm that sorts data into labelled classes, or categories of information, through **pattern recognition**. Examples of classifiers include those that can detect nudity, shapes or colours. Classifiers need data to be trained on and their accuracy improves the more data they are fed. Hence, compulsory detection (based on a judicial or independent administrative order), coupled with incentives towards innovation in the field, is bound to further improve current accuracy rates.

   Thorn's CSAM Classifier is one example of industry's ability to detect new child sexual abuse material. The tool can be set at a 99.9% accuracy rate[1] (false positives). With that precision rate, 99.9% of the content that the classifier identifies as CSAM is CSAM, and it identifies 80% of the total CSAM in the data set. With this precision rate, only 1% of the content flagged as CSAM will end up being non-CSAM. These metrics are very likely to improve with increased utilization and feedback.

   Other tools making use of classifier technology to detect previously new CSAM include Google's Content Safety API[2], and Meta's AI technology[3].

c) **For grooming**, technologies currently used also include classifiers. Like the classifiers used to detect new CSAM, these tools can only detect patterns, which point to possible concrete elements of suspicion of online child sexual abuse without being able to deduce the substance of the content. While not identical in function, these tools use technology similar to the one used in spam filters[4].

   Text classifiers used to detect grooming are trained on Large Language Models, which involve feeding the classifiers with billions of lines of text in order to train the technology in semantic meaning, and also by inputting use-cases involving instances of grooming. Tools of this type include the tool developed under Microsoft's Project Artemis, in collaboration with The Meet Group, Roblox, Kik and Thorn. This tool analyses text-based conversations, rating them on a series of characteristics and assigning each conversation an overall probability that it constitutes grooming. These ratings can be used as a determiner, set by individual companies, to address flagged conversations for additional review. The tool was made available to companies, law enforcement, NGOs and other government entities through Thorn (Anti-grooming starter kit). Microsoft has reported that, in its own deployment of this tool in its services, its accuracy (false positives rate) is 88%.

It is important to note that in all the above technologies, the accuracy rate (false positives rate) *is a setting*, i.e. the tool can be set to detect known CSAM, new CSAM or grooming with more or less accuracy, depending on the optimal operational settings of that specific online service.

---

1 Thorn's data from bench tests.

2 Fighting child sexual abuse online

3 See here and here for more information on Facebook's tool to proactively detect child nudity and previously unknown child exploitative content using artificial intelligence and machine learning.
For more information about content spam filters see here and here and for other spam filters see here, here and here. Spam filters are usually run with the receiving end-user's consent. Some spam filters look only at the subject line of the email.

The higher the threshold set, the lower the number of false positives. For example, if a 99% precision is set, only 1% of the images identified as new CSAM will be a false positive. However, a higher rate of new CSAM images will be left undetected (false negatives). It is estimated that for a 99% precision rate, the false negative rate would currently be around 23%, meaning that 77 of every 100 new CSAM images would be identified[5], while 23 would be left undetected). Idem for grooming: whereas the tool could be set to detect conversations that have a 99% chance of constituting grooming, the higher the accuracy rate set (less false positives), the higher the number of grooming conversations that will be left undetected (more false negatives).

The maximum accuracy rate at which the tool can operate in optimal conditions increases the broader the dataset on which the classifier is trained. Hence the creation of a data set of child sexual abuse images and videos and grooming conversations that have each been verified by a court in an EU Member State, which will set a new standard of quality that does not exist to date, will be a key contribution to further increasing accuracy over time.

3) ***According to various digital rights organisations, in particular edri, the Commission repeatedly refused to meet with them on the proposal. Is that true? With what other stakeholders has the Commission or the Commissioners herself met on the file?***

The Commission met and discussed with EDRI on several occasions concerning the proposal. In particular:

- EDRI participated in the workshop organised by the Commission to prepare the drafting of the proposal, together with other privacy NGOs, on 26 February 2021.

- Two members of DG HOME, Unit D.4 (in charge of the file) met with EDRI on behalf of the Commissioner on 17 February 2022.

- The Commission responded to EDRI's open letter on Protecting digital rights and freedoms in the legislation to effectively tackle child abuse (Ares(2022)4863937).

The Commission also exchanged with EDRi on a number of occasions on various public panels and roundtables. The Commission met with a number of stakeholders on the file, including privacy organisations, NGOs working on the rights of children, relevant service providers, technology experts and law enforcement from different Member States. Annex 2 of the Impact Assessments describes the multiple consultations that took place over two years in the preparation of the Impact Assessment for the proposal, from February 2020 to January 2022, which continued until the adoption of the proposal in May 2022.

4) According to many experts the proposal lacks a sufficient legal basis. Could you please comment on that?

The proposal is correctly based solely on Article 114 TFEU, allowing the EU to take measures which have as their object the establishment and functioning of the internal market. In particular, Article 114 is the appropriate legal basis to address differences between provisions of Member States' laws which obstruct the fundamental freedoms and thus have a direct effect

[5] Thorn's data from bench tests.

on the functioning of the internal market, and to prevent the emergence of future obstacles to trade resulting from differences in the way national laws would otherwise develop.

The main aim of the proposal is to ensure the proper functioning of the internal market, including through the harmonisation of rules and obligations concerning certain online service providers in relation to providing services which are at high risk of being used for child sexual abuse and exploitation online. As explained in the explanatory memorandum and impact assessment accompanying the proposal, Member States have started taking action unilaterally, adopting or considering rules to deal with the challenge posed by child sexual abuse online, which are necessarily national in scope and risk fragmenting the Digital Single Market.

The main content of the proposal consists of (i) obligations on online service providers, meant to create the best conditions for maintaining a safe online environment and (ii) the establishment of the EU Centre, to facilitate the relevant service providers' compliance with their obligations and ensure coordination and cooperation of the activities under the proposal at EU level. As such, the initiative should increase legal certainty, trust, innovation and growth in the single market for digital services.

It should be added that the choice of in internal market legal basis to ensure a level playing field and a high level of security on the digital single market is in line with the Commission's practice. Relevant examples are the TCO regulation (2021/784) and the DSA (Regulation 2022/2065).

It should be added that Articles 82 and 83 TFEU, which constitute the legal basis for the Child Sexual Abuse Directive (Directive 2011/93/EU), provide a basis for criminal law rules concerning, inter alia, the rights of victims of crime and the definition of criminal offences and sanctions in the areas of particularly serious crime with a cross-border dimension such as sexual exploitation of children. As the proposal does not harmonise criminal law, Articles 82 and 83 TFEU could not be considered as appropriate legal basis.

5) **Many stakeholders underline that the proposal fails to meet the key human rights principles of necessity and proportionality and violates several fundamental rights. As confirmed by the European Commission's own internal Regulatory Scrutiny Board (RSB), the proposal may also violate the EU prohibition of general monitoring. Can you please comment on that?**

The European Commission's Regulatory Scrutiny Board does not provide opinions on legislative proposals but on draft Impact Assessments which are then adopted by the Commission to accompany and explain the legislative proposal and considerations that have gone into its preparation once any concerns of the Board have been comprehensively addressed. The RSB's Opinion is published on the Register of Commission Documents with reference SEC(2022)209 (https://ec.europa.eu/transparency/documents-register/detail?ref=SEC(2022)209&lang=en).