December 2022

DOT Europe position paper on the
Regulation laying down rules to prevent and combat child sexual abuse

ISSUE PAPER 3

Technologies

The proposed text adopts a tech-neutral approach when it comes to methods of detection and other mitigations. DOT Europe welcomes this. However, several areas of concern arise from the proposed provisions.

While the proposal does not specify particular technologies, scanning seems to be the favoured approach. The Regulation should be open to a broader range of technological mitigations under the Article 3 and 4 assessment and mitigation process, as long as they are effective, targeted and proportionate. This would allow providers to adopt detection measures which accommodate their respective service features, including E2EE, while also respecting user security and privacy. Indeed, no technology currently exists that allows for scanning in an E2EE environment without breaking encryption.

In addition, the proposal empowers the new EU Centre to develop or endorse particular technologies and mandate their use. This merits further reflection. It will take time for the EU Centre to build the necessary expertise to develop or assess technologies so the proposal must facilitate a more flexible and pragmatic approach to ensure continuity of detection and technological innovation. Mandating by the EU Centre should, like detection orders, be a last resort and defer to providers' risk assessments and choice of mitigation measures in the first instance.

Many smaller providers who do not have the resources to develop their own detection technology will come to depend on the technology vetted and recommended by the EU Centre. It would be therefore essential to give those interested players the opportunity to participate in the process lead by the Technology Committee to issue its recommendation to the EU Centre. Interested players should be afforded the opportunity to stress test the technology they might need to deploy going forward.

The proposal also empowers the EU Centre to develop and manage databases of indicators. There are robust processes and procedures in place to ensure that databases currently used by providers are of high quality. It is important that any new databases developed by the EU Centre meet the same high standards of classification as those currently in use.

Rue du Trône, 60, 1050
Brussels, Belgium
Tel: +32 (0) 472 26 83 02

WWW.DOTEUROPE.EU
Transparency Register: 53905947933-43
DOT Europe was formerly known as EDiMA

When it comes to technology, huge investments have been made in the last 15+ years to advance hash-matching tools for the detection of known CSAM[1]. This is the only area where mature, robust and proven effective detection technology exists. Indeed, DOT Europe notes that, as it currently stands, technology to detect new CSAM based on machine learning classifiers requires heavy investment in human reviewers to verify the content. Indeed, these classifiers do not in themselves 'detect' CSAM. Instead, they are trained on the basis of certain classifiers, to flag suspicious content that is likely to contain abusive content, for review and confirmation by human reviewers. While classifiers exist that help companies prioritise content for review, this technology is different to hash matching technology for the detection of known CSAM in that the work of automated detection tools need to be systematically complemented through human review. The technology to detect solicitation remains nascent and particularly challenging as well. This is because no reliable solicitation detection technology exists that allows for an accurate and automated process, requiring limited human review (only for technical/operational purposes), and it is unclear at this stage whether/when such technology will exist. Indeed, solicitation is insidious, context-based and by nature difficult to detect.. The Regulation expects these maturing technologies to have an unrealistically low false positive rate for the detection of solicitation, which could disincentivise investment and interrupt innovation by service providers. Also considering continuous developments of circumvention techniques, it is important that the Regulation does not inadvertently block the development of new types of detection. The risks arising from testing and refining such technologies can be addressed effectively within Article 4 and appropriate mitigations could be supervised by the relevant Coordinating Authority and Data Protection Authority.

Examples of current practices

Google develops and shares cutting-edge technology, Content Safety API and CSAI Match, free of charge for qualifying organisations to make their operations better, faster and safer, and encourages interested organizations to use these child safety tools to combat child sexual abuse. Content Safety API helps organisations classify and prioritise potential abuse content for review, so they can identify problematic content faster and with more precision so they can report it to authorities as applicable. CSAI Match is an API that helps organisations identify re-uploads of previously identified child sexual abuse material in videos so they can responsibly action it in accordance with local laws and regulations.

Meta developed and open-sourced its photo- and video-matching technologies (TMK and PDQ) so that industry partners, smaller developers and non-profits can use them to more easily identify abusive content and share hashes — or digital fingerprints — of different types of harmful content.

Microsoft's PhotoDNA tool has been in use since 2009 to identify and remove child sexual exploitation and abuse imagery from online platforms and services. PhotoDNA is used by many companies across the industry, including Discord, Google, Snap Inc, Twitter, TikTok and Meta.

As a part of the **Technology Coalition** , Amazon, Apple, Discord, Dropbox, Google, Meta, Microsoft, Snap Inc, TikTok, Twitter and Yahoo actively support an industry initiative launched in 2020 that includes a multi-million-dollar investment into research and innovation to prevent online child sexual exploitation and abuse.

[1] See for instance Lee, Hee-Eun; Ermakova, Tatiana; Ververis, Vasilis; Fabian, Benjamin (2020). Detecting child sexual abuse material: A comprehensive survey. Forensic Science International: Digital Investigation, Volume 34, 301022. doi:10.1016/j.fsidi.2020.301022

DOT Europe's recommendations

- The Regulation should respect technology neutrality and enable providers to define effective, targeted and proportionate mitigation measures, including detection technologies, to address CSAM on their respective services. This is why DOT Europe welcomes Recital 26 and calls for preservation of language on technological neutrality and tailored solutions in the final text. This approach will allow the Regulation to adapt to new service types and features both now and in the future.
- The Regulation should clearly reflect the differences in maturity of technology to detect known CSAM, new CSAM and solicitation to ensure a proportionate approach that limits the risk of false positives and protects the privacy rights of users.
- Article 19 safeguards designated companies from any related legal challenge related to the use of the new EU Centre's technologies and this provision should be maintained in the final text. Moreover, companies' responsibility should be limited to ensuring deployed technology following a detection order works smoothly.
- More clarity is required when it comes to the timing foreseen to develop the databases managed by the EU Centre and the process for mandating their use and under what circumstances.
- Where the EU Centre develops new technologies, the process should be transparent and industry players should be involved in the stress-testing process, for example by industry experts joining the proposed Technology Committee. Moreover, we recommend continuing to engage with a range of experts to ensure the Regulation balances privacy, safety, and security in a way that can be implemented in practice.
- On the other hand, the Technology Committee should not replicate existing work being done elsewhere, and should take into consideration the great work done by existing groups.
- Novel technologies can initially be less accurate, so the Regulation must be flexible enough to allow such technology to be tested and developed under the Article 3 and 4 risk assessment and mitigation framework, with mitigations supervised by the relevant authorities. This approach recognises that companies are the main source of new technologies and innovation, especially with respect to solicitation and new CSAM, in this space and that the Regulation should not stifle but enable and incentivise beneficial cycles of industry investment and collaboration.
- The Regulation should avoid an overreliance on accuracy of online detection solutions as the metric of effectiveness since testing is often conducted in controlled environments, considering that technology solutions are not at a point where industry can remove the need for human intervention and review.