April 2025

Myth Busting & Facts on the Proposed Regulation on Child Sexual Abuse: Addressing Privacy Concerns with Data and Facts

This document provides an overview of the main myths and concerns that have been brought to bear in the context of the proposed Regulation to combat and prevent child sexual abuse (CSA). The paper presents the facts to address these concerns, aiming to enable a research- and data-based assessment of the proposed Regulation.

The Myths

Myth 1: The Regulation would unleash mass surveillance and 'read' all messages.

This claim builds on a misinterpretation of the process established by the proposed Regulation and a misunderstanding of the technology.

Under the proposed Regulation, detection would happen after a thorough process of risk assessment, reviews, multiple checks, and a court order, making it **virtually impossible for the detection technology to be misused** (see further Fact 2: The Regulation will establish strong oversight and ensure privacy).

The proposed Regulation would mandate all online service providers to assess the risk that their service is being used for the distribution of child sexual abuse material (CSAM) or grooming of children and to adopt preventive measures (such as safe design or user reporting) to mitigate this risk. If, despite these measures, there is still evidence of a significant risk, a national court or independent authority will determine **on a case-by-case basis** the **necessity and proportionality** of the use of specific detection tools before mandating their use through a detection order, taking into consideration the impact on users' privacy (see also Myth 2: The slippery slope of technology).

Detection would only happen in a specific part of the service (e.g., specific types of channels or specific users) which present significant risk of being used to abuse children and for a limited time. Under the proposed Regulation, the technology deployed must be **reliable** with the smallest margin of error possible and must be as **unintrusive** in terms of impact on the users' rights as possible. It cannot extract any information other than strictly necessary to detect CSAM.

In addition, this claim builds on **unfounded fears and a misunderstanding of the technology** at hand. Detection technology is built for the sole purpose of detecting CSAM and only recognises grooming patterns indicating this. It cannot and does not "read" or understand messages. It looks for matches. It either compares digital fingerprints of images via hash-matching to a database of known and verified CSAM or – in the case of unknown CSAM – it would use an AI-based machine learning (i.e. classifier) to flag content that is suspected to be CSAM. These AI classifiers are trained to be able to tell the difference between CSAM and innocent imagery. For unknown imagery that has been detected as potential CSAM, the content would undergo a multi-step process to be verified as CSAM, including human review.

In short, the technology operates like a metal detector, which can only detect metal and does not recognise or flag anything else underground.

Myth 2: The slippery slope of technology: governments could use this technology to watch political opponents or human rights defenders.

Detection technology is **built for the sole purpose of detecting CSAM** or to recognise grooming patterns. It is extremely difficult and costly to repurpose and abuse CSA detection technology.

Detection technology has been **deployed for over a decade** and is built to only detect CSA to a high level of accuracy. Over 200 companies have already deployed advanced technologies to safely detect, report and eliminate child sexual abuse.

The Regulation would put in place safeguards that would prevent misuse of detection technologies (see further below Fact 2: The Regulation will establish strong oversight and ensure privacy). Only detection technologies that meet the requirements of the Regulation (in terms notably of efficiency, reliability and scope) and are assessed as safe and privacy-preserving by a new independent EU Centre would be allowed.

Detection technologies would only be used:

1. In a specific part of the service presenting a high risk of being used to abuse children
2. After mitigation measures fail
3. Upon request of a judicial court
4. With technologies assessed as safe and privacy-preserving by an EU Centre
5. For a limited period of time (see further below Fact 2: The Regulation will establish strong oversight and ensure privacy).

In addition, the databases of indicators which providers will use to detect CSA (known CSAM, new CSAM or grooming) will be created and maintained by the EU centre itself – not the providers, nor the national law enforcement authorities.

This framework sets a high bar and ensures checks and balances to avoid misuse of detection technology.

Surveillance technology or spyware such as Pegasus already exists and has, unfortunately, been used by governments through. The deployment or non-deployment of CSA detection tools does not change the use of this surveillance technology.

Myth 3: The Regulation would compromise end-to-end encryption.

In end-to-end encrypted environments (E2EE), only the sender and receiver of communications have a 'key' to access what is being sent. With standard encryption, service providers also hold the key to the encrypted message, but can only use it in certain circumstances.

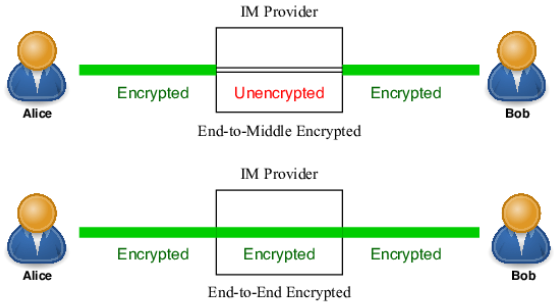


Image source: ResearchGate.

The proposed Regulation does not include any provision on E2EE. The Regulation is **technology neutral**, meaning it would not require any specific technology to detect child sexual abuse, but would instead set criteria for such technology to meet, including that it ensures the respect for privacy, before a deployment order can be issued. This is important to ensure the law can adapt to, and include, developing technologies.

Public authorities have the obligation to ensure children are protected from sexual abuse **in all environments**, even in the most private forms of personal communication. Two-thirds of children who received sexually explicit material online did so through private messaging, mostly on their personal mobile. Predators use **off-platforming**, meaning moving conversations with children to E2EE services to avoid detection of abuse. Our societies cannot allow the creation of a black hole where any type of crime is able to unfold undetected.

Technology to detect child sexual abuse in E2EE while respecting privacy and encryption already exists. This technology is like that used to detect malware and viruses

(see Myth 4: Client side scanning breaks encryption). WhatsApp, an E2EE service, already deploys advanced technology to detect malware and viruses without compromising E2EE.

Myth 4: Client-side scanning breaks encryption.

Client-side scanning consists of **scanning the message before it is sent to the encrypted channel**. It does not break encryption. Client-side scanning can operate on device or with the support of an external database to ensure a verified match against known or suspected CSAM. The EU Centre would ensure that any database used for client-side scanning only contains confirmed CSAM or approved classifiers.

Client-side detection allows the detection of CSAM before a message enters an encrypted environment. **This privacy-preserving technology is already deployed effectively at scale**. It is in operation on major platforms. This is how, for example, WhatsApp prevents the spread of malicious URLs on its encrypted messaging service without affecting E2EE, and how browsers like Chrome and Edge warn users of malware on https. Recently, Apple launched their 'Sensitive Content Warning' and 'Communication Safety' tool. The tool scans messages locally on children's devices to flag sent and received content containing nudity.

Detecting CSAM within end-to-end encrypted environments can also be done in a privacy-forward way through homomorphic encryption, multi-party computation, or secure enclaves. There is still room for huge innovation in this area. A multitude of solutions will mean new approaches that can be used by companies of all shapes, sizes, and scales. In its technological neutrality, the proposed Regulation will encourage **innovation** in this area.

Myth 5: Detecting new CSAM will lead to many false positives.

The tools used to combat online child sexual abuse and exploitation have been used for over a decade across many different types of platforms.

Known CSAM, i.e. that which has already been flagged and verified as CSAM and added to a database, is detected using 'hash-matching technology, which compares two images and flags (almost) identical matches.

Detection technology to detect **new or unknown CSAM** and **grooming** use 'classifiers' that are trained on confirmed CSAM, adult pornography, and legal images to be able to tell the difference between CSAM and innocent 'baby in the bathtub' pictures to a high degree of accuracy. Companies which deploy these technologies can set the threshold for detection accuracy to extremely high to avoid false positives – this is a choice that can be made by a platform.

Once content is flagged using detection technology, **human review** – analysts trained to identify illegal content under a clearly defined legal framework – will confirm that the content is criminal, ensuring that only criminal material is acted upon by law enforcement authorities.

To avoid false positives, specific threshold and accuracy requirements could be established by the **EU Centre** to ensure that a high standard is met. Under the proposal, the EU centre will assess the reports received to ensure unfounded reports are not shared with law enforcement authorities.

Ultimately, false positive rates are a trade-off between precision rates (how much of all the flagged content is CSAM) and recall rates (how much of the CSAM on a platform is detected).

These two rates are adjusted by the technology developers when training the models, and thus depends on where they would like the efficiencies of the technology to lie. In practice, detection methods are tuned to have extremely high precision rates to ensure that all children suffering sexual abuse are effectively protected. This justifies the extremely low risk of false positives.

Many technologies already in deployment – such as speed cameras which keep our roads safe – produce false positives. Societies opt to deploy them as reducing road accidents is important enough to accept a low number of false positives. A zero-error technology does not exist: protecting children from online abuse is a legitimate objective, and the use of detection technologies is **proportional**.

Myth 6: A new regulation is not necessary, extending the interim regulation is enough.

The interim Regulation was adopted in 2021 to derogate some provisions of the e-Privacy directive. This derogation allows number-independent interpersonal communications service (NI-ICS), such as webmail or chat services, to continue detecting child sexual abuse material on their platforms on a voluntary basis. The extension of the temporary derogation alone will not be sufficient to address the scale of the situation. This extension **does not apply to online service providers who start operating after 2 August 2021** and does not cover private communications. This excludes apps that children both use daily, and where they are exposed to sexual abuse. Tackling child sexual abuse should not rely solely on the initiative of online service providers. Transparency and accountability are key in the fight against child sexual abuse online. Children have the right to be protected equally on all the online platforms they use.

Even with a new regulation in place, there must be **a clear legal basis for voluntary detection** to ensure there are not **gaps** in child protection. This could happen, for instance, when an online service provider has to wait to “fail” the risk assessment and mitigation process to receive a detection order and thereby have a legal basis to detect. **Voluntary detection is a risk mitigation** tool and is complementary to the detection orders system proposed by the Regulation. Online service providers cannot identify the risks of child sexual abuse on their services without detecting them. They need to be able to detect to understand the scale of the issue on their platforms.

To avoid gaps in child protection and ensure the long-term feasibility of the proposed Regulation, mandatory and voluntary detection must coexist.

The Facts

Fact 1: Detection is effective and essential in preventing the spread of CSAM. Public reporting will never be sufficient.

The effectiveness of detection is evidenced by the fact that pausing detection correlates directly with falling statistics on the total amount of CSAM reported and removed. This was evident during the legislative gap in 2021 when Facebook was forced to stop detecting in the EU for 10 months resulting in a 58% reduction in CSAM being found and removed.

Public reporting will never be sufficient due to the significant barriers to reporting. Education and awareness about the value of ‘’bystanders’’ reporting can help improve reporting, but will not resolve the issue of under-reporting. Child victims are often unlikely to report their abuse. According to a prevalence study, “**83%** of young people aged 11 to 17 years old who had been sexually assaulted by a peer had **not told anyone**”. Victims may not know their abuse has been recorded, some victims are too young to speak up, and older children often do not report due to shame, stigma, fear, or threats from the offender.

Data shows that the **proactive detection** of CSAM leads to a substantially higher volume of identified and removed CSAM. In 2021, the 50 INHOPE hotlines processed 928,278 URLs reported by the public, while the IWF, the UK’s hotline and Europe’s largest hotline, handled 361,062 CSAM items alone, of which 66% resulted from proactive searching. The Canadian Project Arachnid’s automated web crawling detection tool of known CSAM or close matches processed **158 billion+** images between 2017 and March 2023.

Mandatory company reporting to the NCMEC Cybertipline amounted to 85 million files. While public reporting is crucial to discover known and previously unknown material, proactive searching can do so at a rate and scale that meets the **volume of CSAM in circulation**.

Preventative measures, such as risk assessment and mitigation measures, are crucial to build a digital environment that is safe-by-design for children. However, **prevention measures alone will not stop the proliferation of child sexual abuse online**. Prevention and detection are complementary mechanisms; both play their part in effectively protecting children from re-victimisation and ongoing abuse.

Fact 2: Detecting new CSAM and grooming saves lives.

Behind every image and video of child sexual abuse, there is a child in danger. Detecting new CSAM and grooming is crucial to stop ongoing abuse and protect children from imminent danger.

Currently, online service providers are voluntarily using detection technologies to find and report child sexual abuse to the US’ National Center for Missing and Exploited Children (NCMEC). NCMEC refers these reports to the relevant national law enforcement agencies, who can open investigations to arrest the perpetrators. New CSAM detected and reported enable the law enforcement to save children and arrest offenders every day. In the UK alone, an estimated **1,200 children are safeguarded** and 800 suspected child sex offenders arrested on average **every month.**

We must prevent re-victimisation. The redistribution of CSAM means that, for victims, the abuse not only stays in their memory, but is re-lived constantly and unendingly all over the

world. It is widely acknowledged that the **trauma** of knowing that the evidence of your abuse is recirculating is profoundly damaging, and creates immense difficulties for victims to heal. **Children exposed to grooming and sexually explicit content report similar levels of trauma symptoms (i.e. clinically diagnosable PTSD) to victims of penetrative offline sexual offences.**

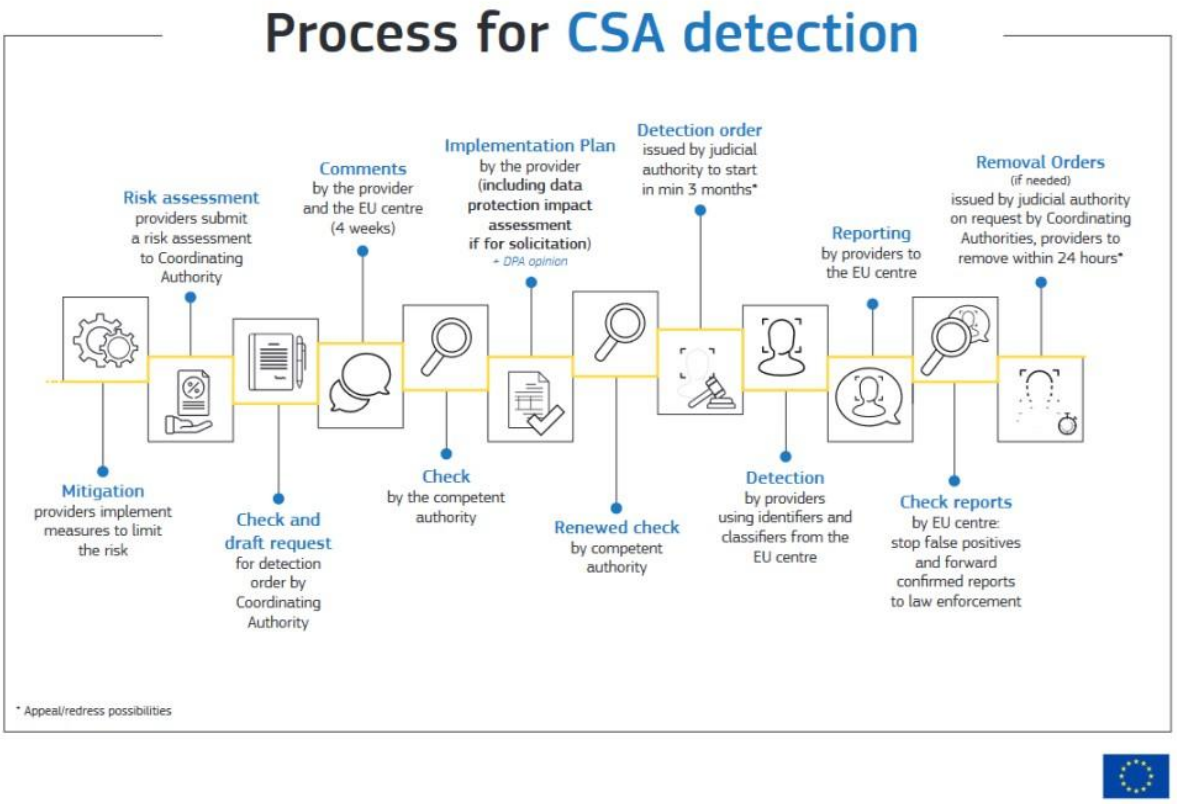
Fact 3: The Regulation will establish strong oversight mechanisms and ensure the privacy of all users.

The unsolicited contact of an adult with a child with sexual intent and the dissemination of images and videos depicting the sexual abuse of a child are breaches of the right to privacy of both the child and the victim. The Regulation will ensure **the right to privacy of children, victims and survivors** is protected.

For all other users of the internet, the Regulation does not allow indiscriminate scanning of private messages. The Regulation will establish strong **safeguards and a long review process** before any detection is authorised to ensure that **no indiscriminate detection** of illegal material is carried out. This minimises any perceived invasion of privacy. These safeguards include:

1. Detection technologies will be authorised and provided by the EU Centre established in the Regulation. Online service providers will not be able to use detection technology that infringes the minimum standards of security and privacy established by the Centre and the Regulation.
2. Detection orders will be issued by a national judicial or administrative authority, in line with national and EU law on data protection and fundamental rights.
3. National Coordination Authorities will review and give feedback on the risk management and on the implementation plan of a detection order.
4. Data Protection Authorities can provide recommendations in the detection process.

As shown in the graphic below, **multiple checks** by the EU Centre and a Data Protection Authority, including transparency reporting and oversight checks, are foreseen to ensure that detection conforms to existing EU law, including the GDPR, and respects the privacy of users.



Source: European Commission.

One must bear in mind that:

- ➔ All legislation in the EU must comply with other existing laws, including the **GDPR**, which strictly regulates the control and processing of personal data by private companies. This is no different for the CSA Regulation.
- ➔ Filtering has been accepted by the Court of Justice of the European Union in cases of high accuracy (for example in IP protection).
- ➔ The proposed legislation places the responsibility of balancing fundamental rights **with independent authorities, rather than by individual companies**.
- ➔ Multiple checks will take place to ensure that only illegal material is removed. Any content detected by the technology will be checked to ensure that they indeed constitute illegal material.
- ➔ When signing into a platform, internet users must consent to the platform's Terms of Services to use the platform.
- ➔ The proposed Regulation will mandate **transparency and accountability** of platforms, so that users are aware what a platform is doing to prevent and remove illegal material.
- ➔ The databases of indicators which will be used by providers to detect each type of CSA (known CSAM, new CSAM or solicitation of children) will be created, maintained and operated by the EU centre itself – not providers, nor national law enforcement authorities.

Fact 4: Technology already exists to tackle child abuse while respecting privacy.

Technologies already exist that effectively detect CSAM with high accuracy rates. This includes PhotoDNA, YouTube CSAI Match, Facebook's PDQ and TMK+PDQF for known CSAM and

Thorn's Safer Tool, Google's Content Safety API and Facebook's AI Technology, for new or unknown CSAM and grooming. **These technologies are already deployed at scale with no issue of misuse or privacy concerns.**

Client-side scanning is already deployed at scale in E2EE for various legitimate purposes, such as viruses and malwares. Some online service providers, such as Apple, already use it to flag CSAM and grooming conversations in their messaging apps. Detecting CSA in E2EE could be done in the exact same manner, using the same technology (see Myth 5: Client-side scanning breaks encryption and Myth 3: Undermining end-to-end encryption).

The Regulation provides **a framework to check** that the technology used to detect CSA and grooming will **minimise privacy intrusion** through the intervention of experts' opinions and judicial courts, while ensuring that the detection is targeted and effective (see Fact 3: The Regulation will establish strong oversight and ensure privacy). Thanks to these requirements, companies will have a powerful **incentive** to develop privacy preserving technologies that can be deployed without friction in any platform, including E2EE environments.

The privacy of all users, including children, victims and survivors, is essential. **The right to privacy of children, victims, and survivors is infringed when pictures and videos of their abuse are shared online** without their consent and when they receive unsolicited contact from adults. **The privacy concerns of child victims and survivors of CSA should be equally valued** by privacy-rights organisations and data-protection authorities.

Fact 5: Most child sexual abuse occurs in private messaging

CSAM and grooming mostly occurs through the use of private messaging. **Two-thirds of children** who received sexually explicit material online did so **through private messaging,** mostly on their personal mobile.

Tools targeting private messaging are key to detect and remove images and videos and to flag potentially grooming conversations. Detection technologies would not be able to 'read' the messages, but instead predict the probability that grooming is happening in a conversation (see Myth 1: The Regulation would unleash mass surveillance and 'read' all the messages). Detecting CSA in private messages thus plays a crucial role in keeping children safe and disclosure of their abuse on behalf of the victim of sexual abuse.

A common tactic used by perpetrators is called **'off-platforming'**, meaning that perpetrators initiate contact with children from public platforms, then entice them to applications that use end-to-end encryption or where detection tools are not in operation. This is a deliberate tactic to obtain CSAM from their victims **undetected**. If we do not include private messaging in the scope of this Regulation, we risk private communications becoming a haven for perpetrators to abuse children.

Fact 6: Citizens overwhelmingly back the EU Regulation.

In 2023, analysts at the Internet Watch Foundation reviewed 101,988 webpages hosted in the EU containing child sexual abuse material between January and August alone. The EU continues to be the largest hub hosting this material, with over 60% of CSAM reported in 2024 being traced to an EU country. Europeans are aware that child sexual abuse is a rising problem in their countries. There is overwhelming support for online service providers to proactively fight against this crime.

The recent Eurobarometer survey shows that 78% of respondents approve of the Commission's legislative proposal to prevent and combat child sexual abuse and 96% see the ability to detect child abuse as equally important or more important than the right to online privacy. Moreover, between **84% and 89% support that service providers use tools to automatically detect** images and videos of known CSAM (89%), new images and videos (85%), and grooming (84%), even if those tools could be perceived to interfere with the privacy of users.

Similarly, a recent ECPAT and NSPCC poll showed that 81% of European respondents support obliging online service providers to detect, report, and remove child sexual abuse online. According to 86% surveyed Europeans, children are increasingly at risk of child sexual abuse and exploitation online, and data reveal that the majority of polled EU citizens see online service providers as one of the most important actors in preventing and protecting children from sexual abuse and exploitation online. **95% say it is key that there are regulations to prevent online child sexual abuse.** These findings, in alignment with the Eurobarometer results, underscore a critical message: European citizens are deeply concerned about child sexual abuse online.

More than half of all Europeans surveyed declare that the issue of child sexual abuse and exploitation online will **influence how they vote at a future election.** There is a clear and urgent demand for decisive action to address this issue. With the European Parliament elections on the horizon, MEPs (Members of the European Parliament) face a duty and a moral responsibility to enact meaningful legislation for child safety online.

More resources

Fact-check: Top 9 claims made on the Regulation to fight Child Sexual Abuse