ECLAG
Ending Child Sexual Abuse Online

June 2023

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

There are a lot of misconceptions out there about the proposed Regulation to fight Child Sexual Abuse and the technology that is deployed to detect child sexual abuse material (CSAM). Here are facts.

Claim #1 *Detection technology won't be effective in stopping the spread of CSAM.*

> ***Fact check*** *On the contrary: Detection can help dramatically stop the spread of child sexual abuse material online, as part of a toolbox of solutions needed to tackle this complex crisis. In 2021, detection efforts slowed down as the legal basis for detection expired. As a result, the number of incidents of reporting went down, despite the fact that data shows the volume of abusive material increased.* ***Increasing detection efforts means more CSAM being found, removed and reported****, which ensures dignity for victims and survivors and increases children's online safety.*

Claim #2 *This legislation is establishing general and ungrounded mass surveillance. The EU wants to "open every letter" and read each and every private message.*

> ***Fact check*** *This claim builds on unfounded fears and on a misunderstanding of the technology at hand.* ***Detection technology doesn't "read" messages.*** *It either compares digital fingerprints of images via hash-matching to a database of known and verified CSAM - or it uses a classifier to flag content that is suspected to be CSAM which then undergoes* ***a multi-step process*** *to get verified as CSAM including human review. What's more, the legislation sets out a safeguard usage process through triangulation of private companies, public authorities and the courts. Additionally, the legislation builds in* ***safeguards*** *to ensure transparency in the use of detection tools, including the input of independent bodies or courts and of the data protection authorities.*

Claim #3 *CSAM detection tools are easy to reengineer for other purposes.*

> ***Fact check*** *The tools to detect CSAM are* ***highly targeted*** *at finding CSAM - and only that. If someone wanted to detect other content,* ***they would need to design entirely different tools and resources****. Under the legislation, the EU Centre will provide access to accredited state of the art tech which is designed to only detect CSAM. It will also*

periodically review their effectiveness. What's more, the use of these technologies will only be permitted on a case by case basis under the review of public authorities.

Claim #4 *CSAM detection tools have a high false positive rate which leads to innocent people getting prosecuted.*

Fact check *F**alse positive rates are a trade-off between precision rates** (how much of the flagged content is CSAM) **and recall rates** (how much of the CSAM on a platform is being detected). In practice, detection methods are tuned to have extremely high precision rates. For all unknown CSAM, there is a **multi-step system** in place ensuring only CSAM gets flagged as CSAM: First, service providers should conduct human review of newly flagged CSAM. Secondly, the US-American Center for Missing and Exploited Children (NCMEC) or in future the EU Centre - with analysts trained to identify illegal content - get to review to ensure the material flagged is actually CSAM. It is highly unlikely that these two instances misinterpret an image.*

Claim #5 *CSAM detection tools wrongly flag consensually shared imagery or pictures of kids in bathtubs.*

Fact check *CSAM detection tools are **specifically trained** to not find "kid in the bathtub" type innocent images. These tools are trained on known CSAM, adult pornography and benign images particularly to tell the difference between them and to keep benign imagery from being misinterpreted as CSAM.*

Claim #6 *CSAM is hosted primarily on the dark web, therefore it is not helpful to detect and fight CSAM on the open web.*

Fact check ***Tens of millions of pieces of CSAM are distributed on the open web**, which for example the NCMEC report numbers show. The dark web is similar to the internet from the early nineties before search engines - it is slow! This means that uploading and downloading content like CSAM, videos in particular, takes very long. One of the most **common use cases for the dark web is therefore to share links to CSAM that is hosted on the open web**. Some sites intentionally mix CSAM with benign imagery to hide CSAM in plain sight. Fighting grooming and sextortion also makes most sense on the open web since this is where most kids are and where they are approached by perpetrators.*

Claim #7 *There is no detection possible for E2EE environments, so we should carve out E2EE from this legislation*.

Fact check ***Detecting for CSAM within end-to-end encrypted environments can be done in a privacy-forward way*** *through homomorphic encryption, multi-party computation, secure enclaves, or client-side-scanning (or a combination of these) with client-side-scanning as the most feasible option currently. And,* ***there may be more ways we have yet to discover:*** *it will take a multitude of solutions from industry to tackle the problem so that they can be used by a variety of companies of different sizes and scales. Knowing how fast technology evolves, it would be fatal to exclude any technology from the scope of this legislation. To incentivize innovation and remain future-proof,* ***this legislation must stay tech-neutral.***

Claim #8 *The algorithms of a classifier are a black box and cannot be trusted. Also, it is biased.*

Fact check *We rely on algorithms already in numerous ways in our daily life. They are by no means a black box: The algorithms deployed to detect CSAM are carefully chosen by engineers. In this particular use case, algorithms cannot be made public to avoid that perpetrators use this knowledge to circumvent them. We support that the EU Centre provides exemplary sets of data on which these algorithms must* ***meet certain benchmarks to be deployed in the EU.*** *Any bias in such algorithms would simply be a lack of diverse enough data.* ***The more data available for training, the better.***

Claim #9 *This legislation would further overburden law enforcement agencies due to the rising number of reports. All we need is more resources for law enforcement.*

Fact check *As for every crime,* ***rising numbers are a reason to intensify efforts*** *and not to look away from the evidence. This said, law enforcement officers involved in the fight against CSAM are dependent on more information and details to quickly identify and rescue children at immediate risk. Sufficient funding for law enforcement goes hand in hand with technologies that support their efforts. Beyond law enforcement investigation, the detection of CSAM plays a crucial role in reducing the further spread of CSAM online, which ensures dignity for victims and survivors, and reduces access to CSAM.* ***It requires the whole child protection ecosystem ranging from hotlines to survivor support to research to technology and more to effectively fight child sexual abuse****.*